

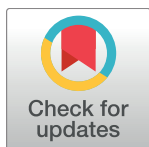
RESEARCH ARTICLE

# Patterns of chromatin accessibility along the anterior-posterior axis in the early *Drosophila* embryo

Jenna E. Haines<sup>1</sup>, Michael B. Eisen<sup>1,2,3\*</sup>

**1** Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, United States of America, **2** Department of Integrative Biology, University of California, Berkeley, Berkeley, United States of America, **3** Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, United States of America

\* [mbeisen@berkeley.edu](mailto:mbeisen@berkeley.edu)



**OPEN ACCESS**

**Citation:** Haines JE, Eisen MB (2018) Patterns of chromatin accessibility along the anterior-posterior axis in the early *Drosophila* embryo. PLoS Genet 14(5): e1007367. <https://doi.org/10.1371/journal.pgen.1007367>

**Editor:** Juan M. Vaquerizas, Max Planck Institute for Molecular Biomedicine, GERMANY

**Received:** November 2, 2017

**Accepted:** April 17, 2018

**Published:** May 4, 2018

**Copyright:** © 2018 Haines, Eisen. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Raw sequence files and normalized wig files can be accessed from GEO, accession number GSE104957 at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE104957>. Protocols are available at [protocols.io](https://www.protocols.io/view/atac-seq-on-nuclei-from-frozen-sliced-drosophila-m-j9zcr76) at <https://www.protocols.io/view/atac-seq-on-nuclei-from-frozen-sliced-drosophila-m-j9zcr76>. All code and commands are available in the supplementary information.

**Funding:** This research was supported by the MBE's Investigator Award from the Howard Hughes Medical Institute and JEH's Graduate

## Abstract

As the *Drosophila* embryo transitions from the use of maternal RNAs to zygotic transcription, domains of open chromatin, with relatively low nucleosome density and specific histone marks, are established at promoters and enhancers involved in patterned embryonic transcription. However it remains unclear how regions of activity are established during early embryogenesis, and if they are the product of spatially restricted or ubiquitous processes. To shed light on this question, we probed chromatin accessibility across the anterior-posterior axis (A-P) of early *Drosophila melanogaster* embryos by applying a transposon based assay for chromatin accessibility (ATAC-seq) to anterior and posterior halves of hand-dissected, cellular blastoderm embryos. We find that genome-wide chromatin accessibility is highly similar between the two halves, with regions that manifest significant accessibility in one half of the embryo almost always accessible in the other half, even for promoters that are active in exclusively one half of the embryo. These data support previous studies that show that chromatin accessibility is not a direct result of activity, and point to a role for ubiquitous factors or processes in establishing chromatin accessibility at promoters in the early embryo. However, in concordance with similar works, we find that at enhancers active exclusively in one half of the embryo, we observe a significant skew towards greater accessibility in the region of their activity, highlighting the role of patterning factors such as Bicoid in this process.

## Author summary

DNA in the nuclei of animals and other eukaryotes is not floating around freely. Rather it is wrapped around proteins called histones that are in turn compacted into higher order structures called chromatin. Highly compact chromatin can restrict which genes and regulatory regions are accessible to the machinery that turns DNA into RNA, effectively shutting genes off. In contrast, chromatin at active sequences tends to be more open and accessible. However, it is still unknown whether open chromatin enables genes to be

Research Fellowship from the National Science Foundation and NHGRI T32<sup>®</sup> training grant in Genomics and Computational Biology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** I have read the journal's policy and the authors of this manuscript have the following competing interests: Michael Eisen is a founder and member of the board of directors of PLOS.

active, or if open chromatin is a byproduct of activity. Here we use a genomic technique to compare chromatin in the anterior and posterior halves of young fly (*Drosophila melanogaster*) embryos and show that two important classes of regulatory sequences—promoters and enhancers—are accessible in parts of the embryo where they are transcriptionally active and, often, in parts where they are not. This suggests that chromatin accessibility is not a direct consequence of activity per se, and supports a model in which chromatin around active genes and enhancers is systematically opened by specific regulator proteins and is later refined by activity.

## Introduction

During early embryogenesis all animal genomes undergo a transition from a largely quiescent to a highly active state with widespread zygotic transcription [1]. This process, known as the maternal-to-zygotic transition (MZT), involves a major reorganization of chromatin, during which active and inactive regions are established which are distinguished by nucleosome composition, density and post-translational modifications [2–6]. It is generally thought that active—or “open”—chromatin facilitates the binding of polymerases, transcription factors and other proteins to target sequences, while inactive—or “closed”—chromatin limits the scope of their activity, although the degree to which chromatin state is instructive remains controversial [7,8]. Two important open questions are how genomic locations of active and inactive chromatin are encoded in the genome and how their active state is established, especially during the MZT, which follows early embryonic mitotic divisions where little or no differentiation into open and closed chromatin has been observed [2].

In *Drosophila melanogaster*, zygotic transcription largely begins at the seventh syncytial mitotic cycle (although there is evidence for low levels of transcription from the beginning of embryogenesis [9]) and gradually increases until the end of mitotic cycle 13, when the embryo has several thousand nuclei and widespread zygotic transcription is observed [10,11]. Many of the genes activated during the MZT produce mRNAs that have spatially restricted distributions. These patterns are established through the activity of transcriptional enhancers, *cis*-regulatory sequences that integrate activating and repressive inputs from well-characterized, patterning transcription factors to produce novel, increasingly precise transcriptional outputs [12–15].

It is widely assumed that the interactions among patterning factors and the DNA to which they bind plays a central role in determining which sequences will function as enhancers, and that their competition with nucleosomes and recruitment of chromatin remodeling factors establish chromatin accessibility at selected sites [16–19]. The anterior morphogen Bicoid, for example, has been shown to create open chromatin at a subset of its targets [19] in the early embryo.

However, we and others have shown that a parallel system involving the ubiquitously expressed, maternally-deposited pioneer factor Zelda also plays a role in this process [2,20–24]. Zelda binds prior to the MZT to a large fraction of the enhancers and promoters that become active once widespread zygotic transcription begins [20,25]. Most MZT enhancers and promoters contain conserved Zelda binding sites that are highly predictive of both transcription factor activity and chromatin accessibility [20]. Furthermore, Zelda binding is associated with changes to chromatin, including nucleosome depletion and specific post-translational modifications of histones [2,20–24].

Although abundant genetic, genomic and biochemical data support the importance of Zelda in establishing enhancer and promoter activity, many aspects of Zelda activity remain

unresolved. While embryos lacking Zelda show severe defects prior to gastrulation, patterned, enhancer-driven transcription is not eliminated in Zelda<sup>-</sup> embryos [22,25–28]. This could reflect the activity of additional pioneer factors [21] such as the ubiquitously expressed trithorax-like/GAGA Factor (or GAF) which plays an important role in establishing accessibility at promoters [21,29–31] and is likely associated with changes in the nucleo-cytoplasmic ratio [32,33].

Because Zelda and GAF are ubiquitously expressed, while patterning factors have spatially restricted expression, we reasoned that we could probe their relative contributions to the establishment of chromatin accessibility by measuring spatial patterns of chromatin accessibility in the early embryo. As a first step towards this end, here we compare chromatin accessibility in anterior and posterior regions of the *D. melanogaster* embryo.

## Results

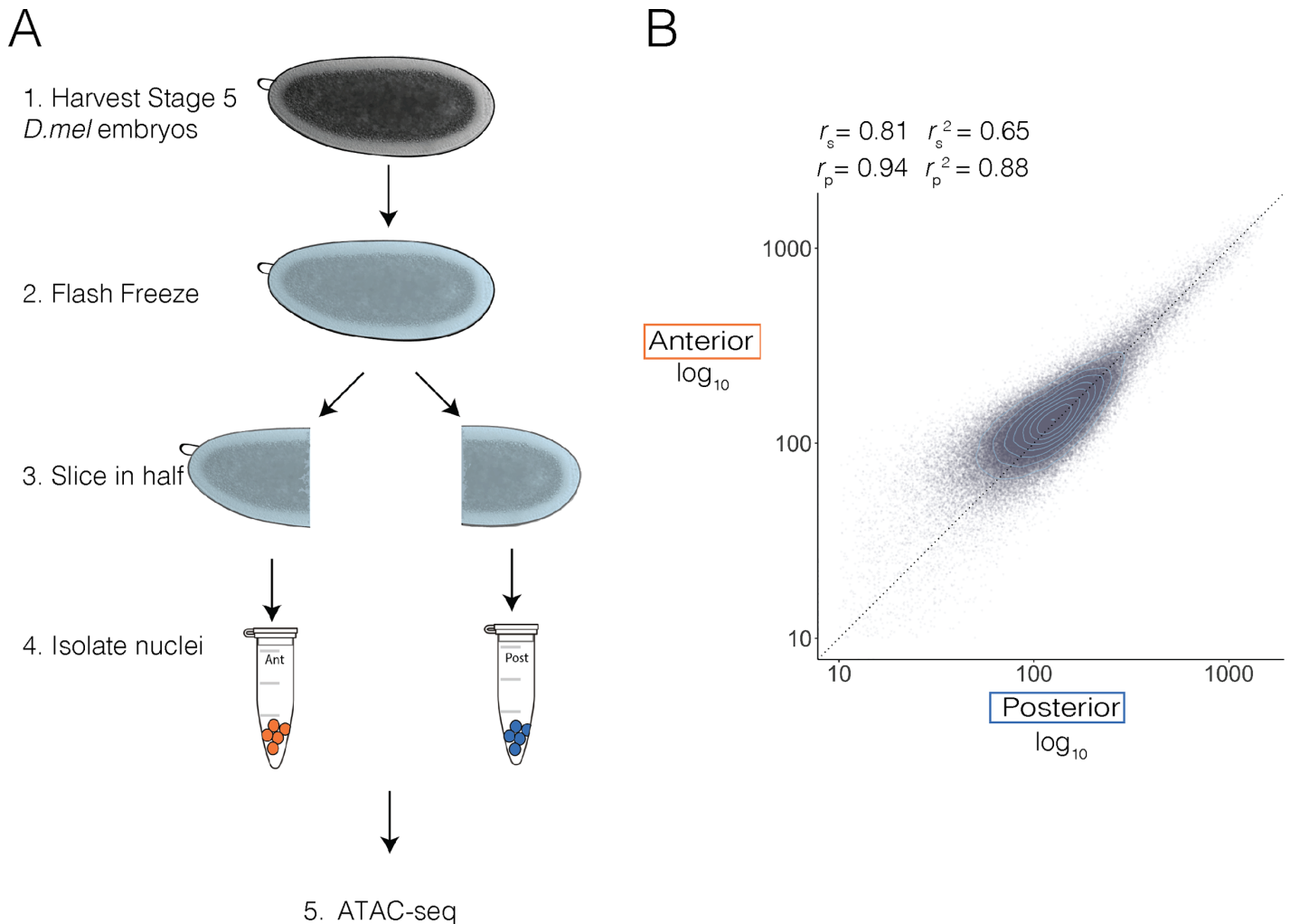
### Spatially resolved ATAC-seq is robust and consistent with whole embryo measurements of chromatin accessibility

To determine the extent to which chromatin accessibility is spatially patterned along the A-P axis in the early embryo, we manually separated anterior and posterior embryo halves and performed a modified ATAC-seq [34] protocol on each half separately. Briefly, we collected cellular blastoderm embryos (mitotic cycle 14, embryonic stage 5), flash froze them in liquid nitrogen, and then sliced each embryo at approximately 50% egg length (ascertained by eye) with a chilled scalpel, separating anterior and posterior halves into separate pools (Fig 1A). We isolated nuclei from 20 anterior halves (in duplicate), 20 posterior halves (in duplicate), 10 frozen unsliced embryos, and a mixed sample containing a subset of nuclei from anterior and posterior samples and applied the ATAC-seq “tagmentation” process to each sample. We sequenced the resulting libraries, mapped reads to the *D. melanogaster* genome and normalized the data using standard methods (S1 Fig).

ATAC-seq accessibility profiles generated from sliced and unsliced whole embryo samples correlated highly, demonstrating that the slicing process does not introduce any biases ( $r_p = 0.95$ , S2C Fig). Both halves correlate with published DNaseI hypersensitivity data from similar embryo stages [35], demonstrating that our embryo preparation protocol coupled with ATAC-seq can accurately map accessibility in the equivalent of 10 whole frozen embryos ( $r_p > 0.78$ , S2A Fig). Biological replicates of anterior and posterior halves that were collected, sliced, and tagmented independently moderately correlate with each other (anterior replicates  $r_p = 0.88$ , posterior replicates  $r_p = 0.80$ , S2B Fig). To call peaks using MACS2, we first merged replicates to increase the total read number and decrease spurious peaks that arise from low coverage regions [32,36]. We then filtered our peaks for those that were found in both replicates (methods).

### Globally similar chromatin accessibility patterns in anterior and posterior embryo halves

Genome-wide, chromatin accessibility in the anterior and posterior halves is remarkably similar (Fig 1B;  $r_p = 0.94$  on data binned into 1kb windows and  $r_p = 0.90$  for all whole embryo peaks S3 Fig). Dramatic changes in chromatin accessibility have been observed between early (stage 5) and later stage (stage 14) *Drosophila* embryos [17,35]. Expectedly, A-P halves are more similar to each other than embryos from stage 5 and 14 ( $r_p = 0.66$ , S4 Fig). The conservation of chromatin accessibility patterns between halves is detailed in Fig 2, which shows the results of our ATAC-seq experiments near loci of three A-P patterning genes (*even-skipped*, *giant*, and *hunchback*) and one dorsoventral patterning gene (*dpp*).



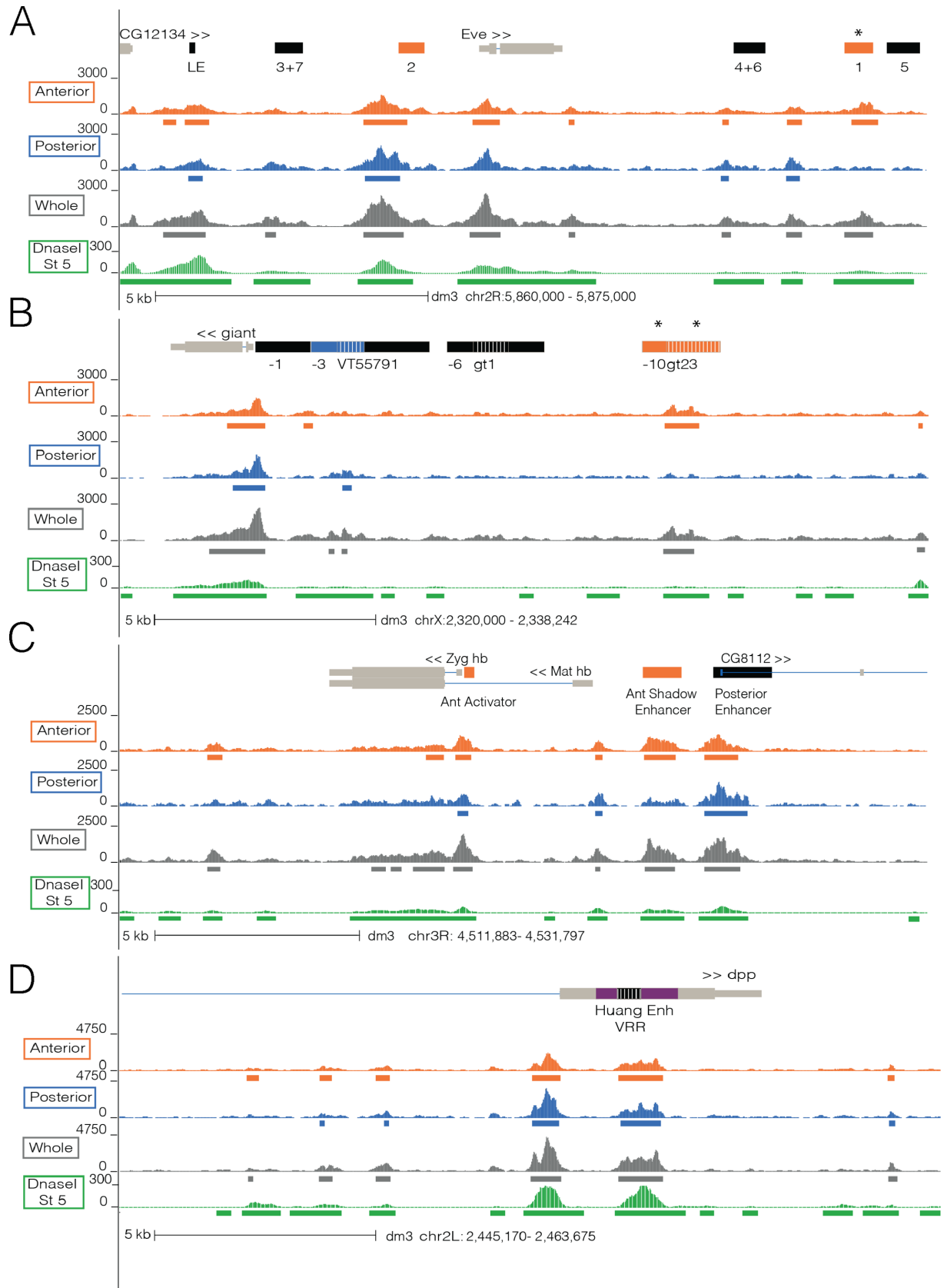
**Fig 1. ATAC-seq on dissected, frozen, embryo halves.** (A) Stage 5, hand sorted *Drosophila* embryos were flash frozen over dry ice in a buffer containing 5% glycerol and manually sliced in half with a scalpel. Twenty anterior and posterior halves were collected, homogenized, and the nuclei were isolated. ATAC-seq was then performed as described in [34] with three times Tn5 transposase. (B) Scatter plot of normalized ATAC-seq signal over 1kb adjacent windows that tile the *Drosophila* genome in posterior (x) and anterior (y) samples shows high degree of correlation between the anterior and posterior halves. The Spearman correlation coefficient (denoted by  $r_s$ ) is 0.81. The Pearson correlation coefficient (denoted by  $r_p$ ) is 0.94. X and Y are log transformed. Light blue circles denote point density.

<https://doi.org/10.1371/journal.pgen.1007367.g001>

Each of these A-P loci contains enhancers that are active exclusively in the anterior or posterior half of the embryo (denoted by colored in Fig 2). For some, the peaks are of similar heights in both halves, such as at *eve* stripe 2 (anterior ATAC-seq signal / posterior ATAC-seq Signal—755/686). However there are some examples where accessibility is clearly reduced in the inactive half, such as at *eve* stripe 1 (1083/336), the *gt* anterior enhancers 23 (648/211) and -10 (513/175) (Fig 2, marked by asterisks).

### Most A-P enhancers are open in both halves of embryo but tend to be more accessible where they are active

To get a more systematic view of the relationship between transcriptional activity and spatial patterns of chromatin accessibility, we used available genome annotation and published *in situ* hybridization experiments to systematically identify A-P and dorsal-ventral (D-V) (as a control) patterning enhancers whose transcriptional outputs are restricted to one half of the



**Fig 2. Chromatin accessibility differences and similarities at A-P and D-V patterning loci.** Normalized ATAC-seq signal of anterior (orange), posterior (blue), whole embryo (gray) is depicted at *even-skipped* (A), *giant* (B), and *hunchback* (C) loci which contain A-P patterning enhancers and promoters and at *decapentaplegic* (D), a D-V patterning gene. Chromatin accessibility signal derived from DNaseI data for stage 5 *Drosophila* embryos is depicted in green [35]. Colored bars represent peaks called in anterior (orange), posterior (blue), whole (gray), and in DNaseI data (green). Asterisks denote annotated features that show significant changes in accessibility between the anterior and posterior halves. Light gray bars denote the gene annotation while the black bars denote annotated enhancers. Colored annotation bars represent enhancers analyzed in Fig 3. Dashed lines in the enhancer bars signify overlapping enhancers.

<https://doi.org/10.1371/journal.pgen.1007367.g002>

embryo [37–48] (S1 File). We excluded enhancers and promoters of genes expressed only around 50% egg length because the precision of manual slicing is most likely variable. We also excluded enhancers that did not overlap peaks called in any of the anterior, posterior, or whole samples, leaving 85 A-P and D-V patterning enhancers.

Patterning enhancers clearly trend towards greater accessibility in the embryo half where they are active (Fig 3). Normalized ATAC-seq signal at anterior enhancers (anterior  $r_p = 0.81$ ) is less correlated than all 1kb regions genome-wide (gray;  $r_p = 0.94$ ) or D-V patterning enhancers ( $r_p = 0.97$ ) while posterior patterning enhancers ( $r_p = 0.99$ ) correlate similar to the genome-wide measurements (anterior  $n = 30$ , orange; posterior  $n = 9$ , blue; dorsal  $n = 16$ , purple; ventral  $n = 27$ , green; Fig 3A and 3B). From this, it is clear that at anterior patterning enhancers, chromatin accessibility is greater in the anterior half.

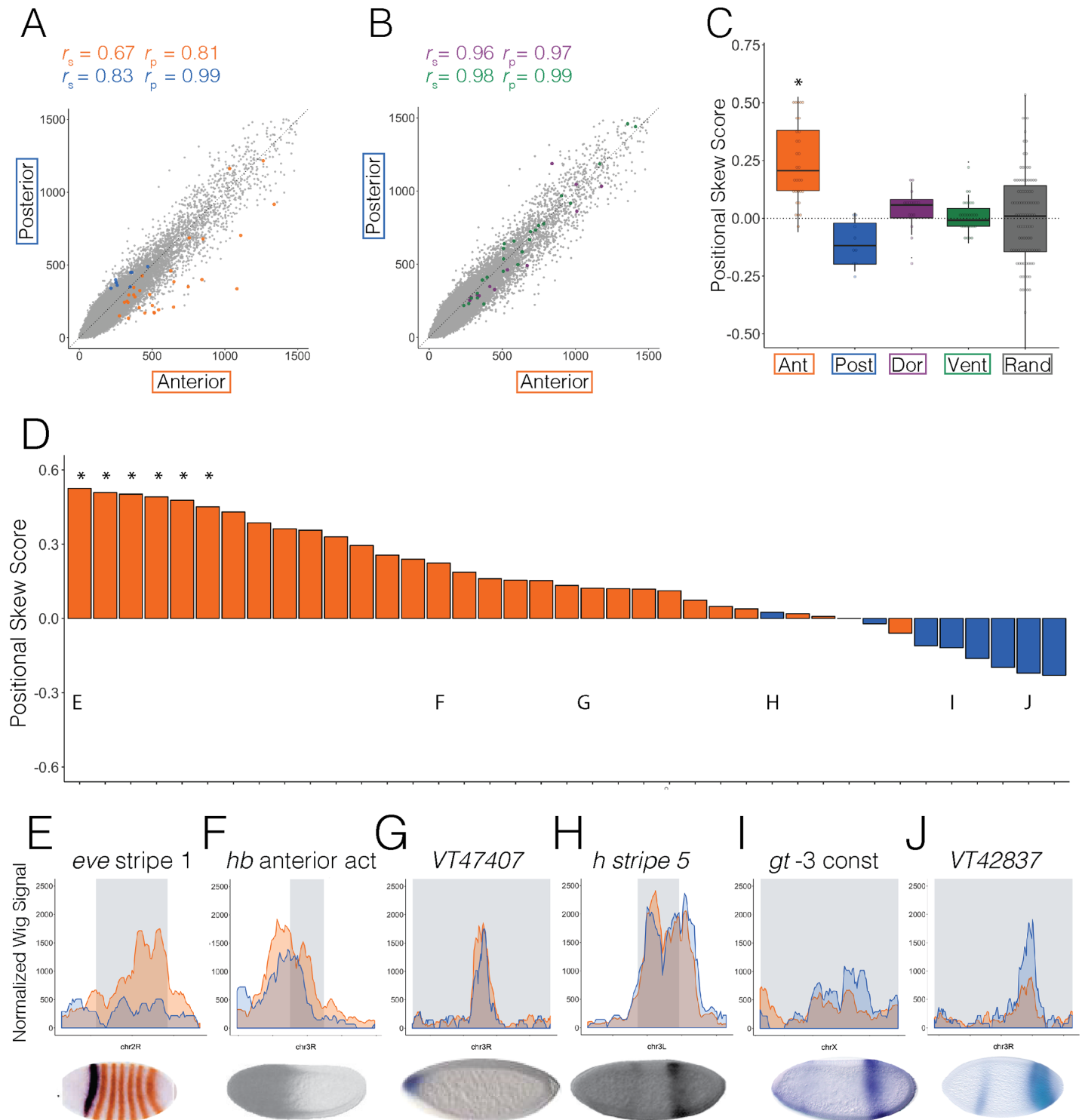
We computed a measure of differential accessibility (positional skew score) for each A-P enhancer by dividing the difference in accessibility in the anterior and posterior half by total accessibility, such that positive scores denote loci that are more accessible in the anterior half, negative scores signify loci that are more accessible in the posterior half, and loci with a score of zero have no difference in accessibility. We found that only anterior enhancers have a significantly greater mean positional skew score than D-V enhancers ( $p_{\text{ant}} < 0.006$  versus dorsal;  $5.57e-05$  versus ventral) or random genomic regions with similar total accessibility ( $p_{\text{ant}} < 6.68e-08$ , Fig 3C and S1 Table).

Accessibility at almost all anterior enhancers is skewed towards the anterior while that of posterior enhancers is skewed towards the posterior (Fig 3D). This pattern is in contrast to D-V patterning enhancers and promoters (S5 Fig) and A-P patterning promoters (Fig 4D). Although individually only six anterior enhancers had significant skews relative to random regions, it is remarkable that almost all of these enhancers are skewed towards the active half regardless of the degree of skew. Similar trends were seen when positional skew scores calculated from replicates were examined (S6 Fig) as well as in accessibility profiles derived from single embryo halves (S7 Fig).

In order to understand if this phenomena extends beyond annotated A-P patterning enhancers, we evaluated whether whole peaks that have chromatin accessibility skewed towards the anterior or posterior show specific patterning activity in the embryo. We overlapped significantly skewed peaks with regions identified in a genome-wide screen for enhancer activity [45]. We then utilized available *in situ* hybridization experiments to evaluate whether these fragments have spatially patterned activity (<http://enhancers.starklab.org/>). Indeed nine out of twelve significantly skewed enhancers showed spatially patterned activity (8 anterior and 1 posterior, S8D Fig).

### Promoters of A-P patterning genes are similarly accessible both when active and inactive

We next examined the promoters of A-P patterning genes using expression data from sections of embryos cryosliced along the A-P axis to curate lists of A-P patterning gene promoters [28]. Similar to our enhancer set, we only included promoters that overlapped accessibility peaks



**Fig 3. A-P patterning enhancers tend to be more accessible where they are active.** Scatter plots showing normalized ATAC-seq signal in anterior (x axis) and posterior (y axis) halves at (A) anterior (orange) and posterior (blue) and (B) dorsal (purple) and ventral (green) patterning enhancers active in Stage 5 embryos and at 1kb adjacent windows tiling the genome (A and B, gray). (C) Box plots showing the difference in mean and variation between average positional skew scores (methods) for anterior (orange), posterior (blue), dorsal (purple), and ventral (green) enhancers. Positional skew scores at random genomic regions (excluding genes and enhancers) that are selected so that their total ATAC-seq signal is of the same magnitude and distribution as the patterning enhancer and promoter set depicted here and in Fig 4 (methods). Pairwise t-tests confirm that the means of anterior patterning enhancers are significantly different than dorsal and ventral patterning enhancers and selected random regions. The mean positional skew score of posterior enhancers is not significantly different than dorsal, ventral, or random regions. (D) Bar graph

shows the positional skew scores (methods) calculated for all anterior (orange) and posterior (blue) patterning enhancers in the dataset (S1 File). Asterisks denote enhancers whose accessibility skew scores show statistical significance over random regions ( $p < 0.05$ ). (E–J) Normalized ATAC-seq signal across 1kb windows centered around *eve* stripe 1 (E), the *hunchback* anterior activator (F), Kvon region VT47407 (G), *hairy* stripe 5 (H), *giant* -3 construct (I), and Kvon region VT42837 (J) with anterior signal in orange and posterior signal in blue. Gray rectangles denote the location and size of the enhancer. Published *in situ* hybridization images depicting gene expression patterns driven by each enhancer are below each graph [39,86–88].

<https://doi.org/10.1371/journal.pgen.1007367.g003>

called in whole, anterior, or posterior samples and that are also associated with patterned expression confirmed by *in situ* hybridization assays ( $n = 19$  anterior promoters,  $n = 25$  posterior promoters, S1 File).

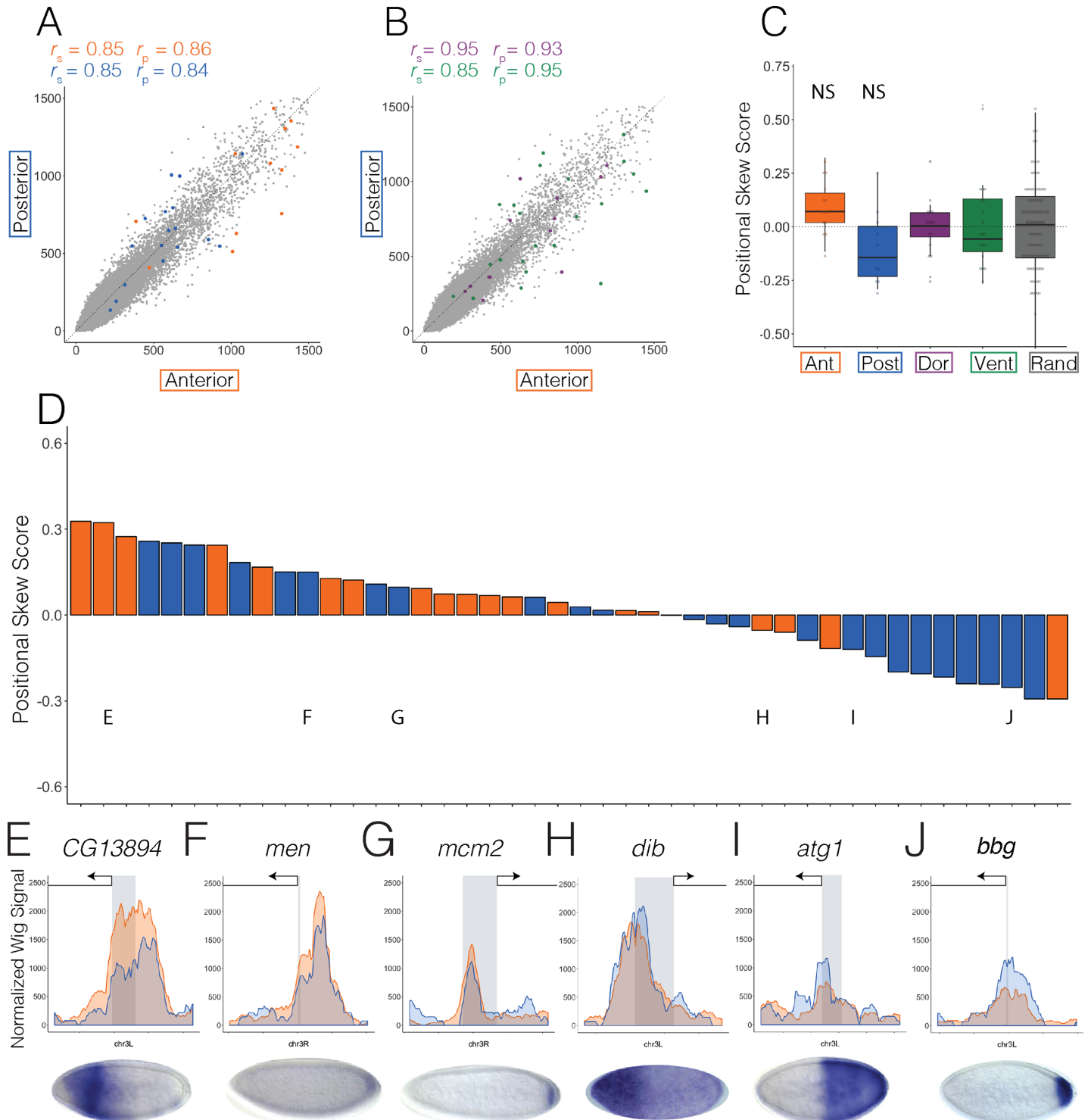
Though accessibility in the active and inactive halves is only slightly more similar at anterior promoters ( $r_p = 0.86$ ) than at anterior enhancers ( $r_p = 0.81$ ), it is comparable to posterior promoters ( $r_p = 0.84$ ) and D-V enhancers and promoters (Fig 4A and 4B, dorsal promoters,  $r_p = 0.93$ ; ventral promoters,  $r_p = 0.95$ ). We confirmed these trends by showing that the mean positional skew score of anterior and posterior patterning promoters is not significantly different than D-V patterning promoters or random regions (Fig 4C and S1 Table). Though accessibility at promoters of A-P expressed, zygotic genes seems to show a very slight trend in the direction of activity, their means are not significantly different than that of random regions (Fig 4C). What is more telling is that there is no distinct skew of accessibility in the expected direction of activity, in contrast to what we observed for A-P enhancers (Fig 4D). From this we conclude that accessibility at promoters is not as correlated with transcriptional activity as enhancers.

### Anterior accessibility is associated with Bicoid binding while similarly accessible regions are enriched for Zelda binding

We then used published ChIP-seq data of A-P patterning factors from stage 5 *Drosophila* embryos [49] to examine binding patterns at similarly and differentially accessible A-P enhancers. We analyzed Bicoid, Caudal, Knirps, Giant, Hunchback, Kruppel, and Zelda binding data, normalized by the mean signal for each factor. A-P patterning enhancers that are more accessible in the anterior (shades of orange) generally are dominated by Bicoid binding, with strikingly little binding from other transcription factors, although there are some exceptions (Fig 5, S9 Fig). Enhancers more accessible in the posterior (shades of blue) generally have high Caudal, Knirps, Giant, and Kruppel binding, with more diversity in factors bound than at anteriorly accessible enhancers. Interestingly, enhancers with similar accessibility in both halves (shades of white) have a high diversity of factors binding—including Zelda (Fig 5, S9B Fig). These patterns reveal that while transcription factor binding clearly does not completely explain differential chromatin accessibility, there are clear differences in factor composition and density between differentially and similarly accessible enhancers.

We next examined transcription factor binding at peaks called in whole embryo ATAC-seq samples. Using a stringent statistical cutoff, we found 107 anteriorly skewed peaks, 9 posteriorly skewed peaks, and 6640 peaks that were not significantly skewed. Anteriorly skewed peaks were enriched for Bicoid, GAF, CF2 and to a lesser extent Zelda binding sites while the unskewed peaks were enriched for Zelda, pnr (GATA factor homolog), Dref (associated with insulators), and CF2. Due to so few peaks being skewed towards the posterior, none of the posterior peaks had significant motifs called, although GAF and Hunchback were enriched (S8 Fig). We then overlapped transcription factor peaks with anteriorly skewed, posteriorly skewed, and unskewed peaks and found that there is a significant depletion in Hunchback, Kruppel, Caudal, Knirps, and Zelda peaks in the anteriorly skewed peaks ( $p = 0.03, 0.005, 5.14E-10, 0.01, \text{ and } 0.00001735$  respectively, S8 Fig). These data further demonstrate the observation that, while transcription factor binding does not completely explain the differences





**Fig 4. A-P patterning gene promoter accessibility does not correlate with activity.** Scatter plots showing normalized ATAC-seq signal in anterior (x axis) and posterior (y axis) halves at (A) anterior (orange) and posterior (blue) and (B) dorsal (purple) and ventral (green) patterning promoters active in Stage 5 embryos and at 1kb adjacent windows tiling the genome (A and B, gray). (C) Box plots showing the difference in mean and variation between overall positional skew scores (methods) for anterior (orange), posterior (blue), dorsal (purple), and ventral (green) promoters of zygotically expressed genes only. Zygotic genes were classified as such from previously published transcriptome data [11]. Positional skew scores at random genomic regions (excluding genes and enhancers) with the same number and distribution of total ATAC-seq signal as the patterning enhancer and promoter set are in gray (methods). NS (not significant) refers to pairwise t-tests which confirm that the mean positional skew score of anterior and posterior patterning promoters is not significantly different than dorsal and ventral patterning promoters or random

regions. (D) Bar graph shows positional skew scores calculated for all anterior (orange) and posterior (blue) patterning promoters in the dataset (S1 File). Asterisks denote promoters whose accessibility skew scores show statistical significance over random regions ( $p < 0.05$ ). (E-J) Normalized ATAC-seq signal across 1kb windows centered around *CG13894* (E), *men* (F), *mcm2* (G), *dib* (H), *atg1* (I), and *bbg* (J) with anterior signal in orange and posterior signal in blue. Gray denotes the location and size of the promoter. Arrow denotes the direction of the gene. Published *in situ* hybridization images depicting gene expression patterns driven by each promoter are below each graph [46–48].

<https://doi.org/10.1371/journal.pgen.1007367.g004>

between skewed and unskewed peaks, Bicoid, Zelda, and GAF are likely playing a role in shaping chromatin accessibility, as has already been reported by several recent studies [19,21,32].

## Discussion

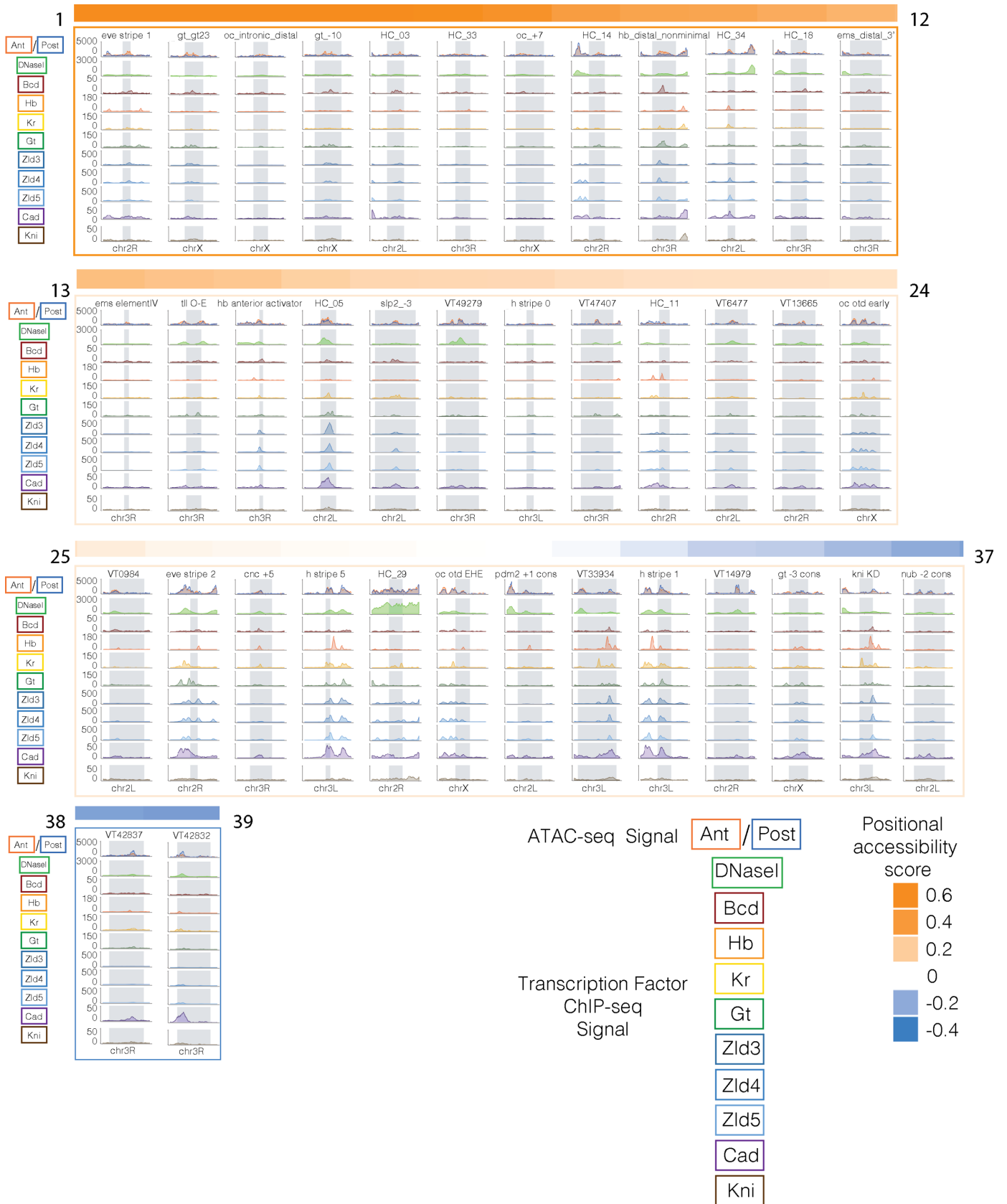
We designed this experiment to ask if the open chromatin observed at active enhancers and promoters is found in every nucleus, suggesting a dominant role for ubiquitous factors in establishing regions of genomic activity in the early embryo, or if open chromatin is spatially restricted, suggesting a dominant role for patterning factors.

The data presented here offer several clarifying observations about the early *Drosophila* embryo. First, we find that the vast majority of regions observed to be accessible in whole embryos are equally accessible in anterior and posterior halves, including the promoters of many genes active in only one half. Second, for a curated set of enhancers driving patterns along the anterior-posterior axis, we find that chromatin is more accessible in nuclei where the enhancer is active. This is especially true of anterior patterning enhancers regulated by the primary anterior morphogen Bicoid, but is also observed for several posterior patterning enhancers. Third, in most cases where we see accessibility biased towards the embryo half where an element is active we also see clear chromatin in the “inactive” half.

Each of these observations comes with the caveat that the signal measured in each pool of halves is an average from approximately 60,000 nuclei that clearly limits the conclusions that we can draw. For example, for a given, region equal accessibility in both halves could indicate uniform accessibility in nuclei across the embryo, but could also arise from similar fractions of nuclei active in both halves. Similarly, it is impossible to resolve whether quantitative differences between the halves are the result of different numbers of nuclei with open chromatin, differences in the levels of accessibility, or both. And, finally, we likely cannot detect cases where only a small fraction of nuclei are active, although with our current methods it is difficult to accurately estimate what our resolving power is.

Nonetheless, a study published while this paper was under review that applied a single-nucleus chromatin accessibility assay, which is largely immune to these caveats, to a variety of embryonic stages largely confirms our finding [50]. They observe that early (2–4 hour) chromatin is more homogeneous than at later embryonic stages, and that the majority of regions of open chromatin in blastoderm embryos show no clear spatial pattern. Furthermore, when they applied an unbiased clustering method to single nuclei ATAC-seq data, nuclei from blastoderm embryos separate into two broad populations, each with increased accessibility in enhancers with anterior and posterior biases respectively.

We mapped their data to the regions we analyze above and find general agreement with ours (Fig 6). Skew scores computed with their data are correlated with our skew scores, although the skews from single-cell data are more extreme than those from the hand-dissected embryos (Fig 6B). Notably, when these skew scores are plotted around A-P patterning regions, we detect a similar overall skew in the direction of activity at enhancers but not promoters (Fig 6C and 6D). It is interesting that single cell data show greater magnitude of skew score at these regions. This likely reflects the increased spatial precision of single cell methods which are able to subdivide the embryo into anterior and posterior domains by accessibility profile instead of



**Fig 5. A-P patterning transcription factor binding at similarly and differentially accessible A-P patterning enhancers.** A-P patterning enhancers from Fig 3 are ordered by positional skew score. Positional skew score is indicated by the colored bar above each panel—orange indicates more accessible in the anterior, blue indicates more accessible in the posterior, and white is similarly accessible in both halves. Each panel consists of normalized wig signal in a 3kb window around each enhancer (the actual enhancer region is denoted by a gray rectangle). The first panel shows normalized, merged, ATAC-seq signal in the anterior (orange) and posterior (blue) halves. The second panel shows DNaseI signal from stage 5 embryos [35] in green. The third through eleventh panels are normalized wig signal from ChIP-seq experiments of the following proteins: Bicoid (red), Hunchback (orange), Kruppel (yellow), Giant (green), Zelda from stage 3,4, and 5 embryos (dark, medium, and light blue), Caudal (purple), and Knirps (brown). The name of the enhancer is above each panel.

<https://doi.org/10.1371/journal.pgen.1007367.g005>

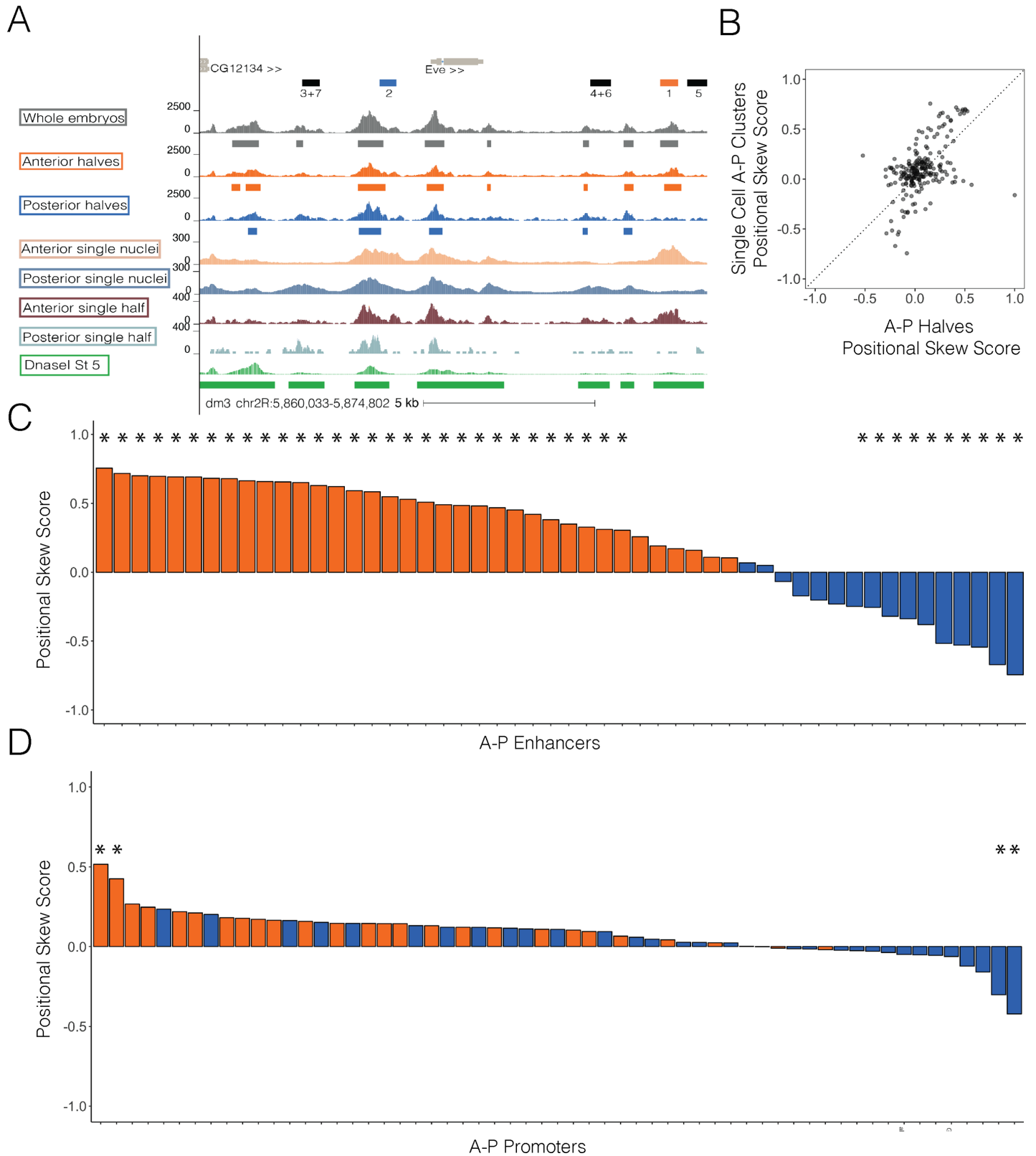
by approximately 50% egg length as we have done. Though there are numerous advantages to single-nucleus experiments, one of the benefits of the experiment reported here is that spatial information about pools of nuclei is determined independently from their accessibility profiles. As such, both experiments taken together reveal that though most genomic regions show similar patterns of accessibility, most A-P patterning enhancers are more accessible in their active half.

The strong anterior skew we and others observe for Bicoid targets is consistent with a recent study showing that chromatin accessibility at a set of around 100 early embryonic enhancers is primarily dependent on Bicoid [19]. Given the strong anterior bias in Bicoid protein levels, it would have been surprising not to find an anterior bias in chromatin accessibility for these regions, although we note that there is significant chromatin accessibility in the posterior for many of these regions in both our data and that of [50], perhaps reflecting activity by the low levels of Bicoid in the posterior [51].

A more comprehensive understanding of how chromatin accessibility is established requires better resolved data on how closely chromatin accessibility tracks with activity in these regions, as neither our data nor that of [50] currently provides greater spatial precision than anterior vs. posterior. There is a large body of literature showing, for example, that tissue specific enhancers often have open chromatin in tissues in which they are not active [52,53], and that this is often a result of enhancer priming by pioneer factors. Furthermore, the transcriptional output of enhancer in the early *Drosophila* embryo [52], and in many other systems, is determined by a balance between the binding of activators and repressors (reviewed in [54]). We expect activators to be bound in parts of the embryo where the enhancer is active, but repressors will bind, by definition, in parts of the embryo where they are repressing enhancer activity. Assuming that open chromatin is associated both with activator and repressor binding implies that the nuclei where chromatin is open for a given enhancer will always be a superset of the nuclei where it is transcriptionally active—the question for the future is how wide these regions are and what their specific patterns tell us about the mechanisms of how they were established.

It will also be interesting to look at the emergence of spatial patterns and biases over time. We have previously shown that nucleosome depletion at enhancers and other aspects of their chromatin state is established prior to the expression of most patterning factors [2]. It has also recently been shown that most enhancers and promoters for patterning genes are accessible by nuclear cycle 11, that this state is maintained through DNA replication and mitosis, and that this open state is associated with ubiquitous factors Zelda and GAF [32]. This leaves open the possibility that the binding of ubiquitous pioneer factors plays a particularly important role in determining chromatin accessibility at early cycles.

In conclusion, our data, as well as that of [19] and [50], establish that though most genomic regions do not show any difference in accessibility, there is significant spatial patterning of chromatin along the anterior-posterior axis in blastoderm embryos, with a clear coupling of activity and accessibility at spatially patterned enhancers. Whether these patterns of chromatin accessibility are instructive for patterning transcription, or merely reflect patterns of activity, remains to be determined.



**Fig 6. Single nuclei ATAC-seq from cellular blastoderm embryos largely agrees with data from embryo halves.** Recently published single nuclei ATAC-seq data [50] corresponding to the cellular blastoderm were separated into anterior and posterior groups following designations determined by [50]. (A) Genome browser trace at the eve locus. Whole embryos, anterior halves, and posterior halves are in grey, orange, and blue respectively. Merged anterior single nuclei and merged posterior single

nuclei from [50] are shown in peach and grey blue or cesious. Finally, single anterior and single posterior halves from our study are shown in rusty red and light blue. DnaseI hypersensitivity data is in green. Peaks are depicted in bars below the pooled halves and whole data. Light gray bars denote the gene annotation while the black bars denote annotated enhancers. Colored annotation bars represent enhancers analyzed in Fig 3. (B) Positional skew scores calculated for single nuclei and halves ATAC-seq data at all A-P and D-V patterning regions (enhancers and promoters) are plotted with halves on X and single nuclei on Y. Dotted line indicates  $X = Y$ . (C) Bar graph shows the positional skew scores calculated from single nuclei ATAC-seq data at all anterior (orange) and posterior (blue) patterning enhancers in the dataset. Asterisks denote enhancers whose accessibility skew scores show statistical significance over random regions ( $p < 0.05$ ). (D) Bar graph shows the positional skew scores calculated from single nuclei ATAC-seq data at all anterior (orange) and posterior (blue) patterning promoters in the dataset. Asterisks denote promoters whose accessibility skew scores show statistical significance over random regions ( $p < 0.05$ ).

<https://doi.org/10.1371/journal.pgen.1007367.g006>

## Methods and materials

### Fly lines

*Drosophila melanogaster* OregonR embryos were collected for 2 hours and aged for 90 minutes on molasses agar plates. Embryos were then dechorionated with 30%-50% bleach solution for three minutes. Embryos were hand staged at 20x magnification at 14°C to be mitotic cycle 14 (NC14) using previously established methods [2].

### Slicing frozen embryos

NC14 embryos were placed in a custom freezing buffer consisting of ATAC-seq lysis buffer [34] without detergent, 5% glycerol, and 1ul of bromoblue dye. Embryos were then taken out of the freezing buffer and placed onto a glass slide which was then put on dry ice for 2–5 minutes. Once embryos were completely frozen, the glass slide was removed and embryos were sliced with a chilled razorblade. Sliced embryo halves were moved to tubes containing ATAC-seq lysis buffer with 0.15mM spermine added to help stabilize chromatin.

### ATAC-seq on frozen embryo halves

Embryo halves were then homogenized using Kimble Kontes Pellet Pestle (cat no. K749521-1590). IGEPal CA-630 was added to a final concentration of 0.1%. After a 10 minute incubation, nuclei were spun down and resuspended in water. Twenty halves were added to the transposition reaction containing 25ul of 2x TD buffer (Illumina), and 7.5ul of Tn5 enzyme (Illumina). The reaction was incubated at 37°C for 30 minutes. Transposed DNA was purified using Qiagen Minelute kit. Libraries were then amplified using Phusion (NEB cat no. F531S) and Illumina Nextera index kit (cat no. FC-121-1011). Libraries were then purified with Ampure Beads at a 1.2: 1 beads to sample ratio and sequenced on the Hiseq4000 using 100bp paired end reads. Fragments over 500bp were removed from libraries using a Pippin prep to reduce sequencing bias with the Hiseq4000.

### ATAC-seq data preprocessing and normalization

Fastq files were aligned to *Drosophila* dm3 genome with Bowtie2 using the following parameters -5 5 -3 5 -N 1 -X 2000—local—very-sensitive-local. Mapping metrics are provided in supplementary S2 Table. Sam files were then sorted and converted to Bam files using Samtools, only keeping uniquely mapped reads with a MAPQ score of 30 or higher using -q 30, proper pairs with -f 2, and removing unmapped, not primary alignment, reads that fail platform vendor quality checks, and optical duplicates with sam flag -F 1804. Duplicates were removed with Picard (<http://broadinstitute.github.io/picard/>). Bams were then converted to bed files with bedtools [55] and shifted using a custom shell script to reflect a 4bp increase on the plus strand and a 5bp decrease on the minus strand as recommended by [34]. Replicate bed files were merged. Finally shifted bed files were converted into wig files using custom scripts (S3 File) and wig files which were uploaded to the UCSC genome browser. Merged wig files were

normalized to reflect 10 million mapped *Drosophila melanogaster* reads. Anterior and posterior samples were normalized by linear regression to the whole embryo sample not including the y intercept (S1 Fig).

### A-P and D-V patterning enhancer and promoter annotations

A-P and D-V patterning enhancer and gene annotations were compiled from many sources (S1 File) [28,37–48,56]. In order to provide the most accurate promoter annotations possible for our analysis we used RACE, CAGE, and EST data performed in *Drosophila melanogaster* embryos [57] to identify promoters preferentially used by the fly embryo. When there were multiple promoters per gene (as was frequently seen), we chose the promoter that was verified by all three methods, denoted by a “V” in Hoskins et. al. (2011) supplementary file 3. There were three genes that did not have annotated promoters in the Hoskins et al. (2011) dataset that were used in our analysis. Instead, these promoter annotations came from the Eukaryotic Promoter Database converted to dm3 annotations [58,59].

In order to further validate our A-P and D-V patterning enhancer and promoter annotations we manually curated *in situ* hybridization images corresponding to 678 genomic regions from multiple sources [38,42,46–48,60–83]. Each region was manually inspected such that only regions with both an *in situ* hybridization image that shows spatially restricted expression as well as moderate accessibility signal (wig signal > 200) were kept for further analysis leaving 253 enhancers and promoters with spatially restricted expression. Anterior and posterior patterning enhancers and promoters were categorized as either completely spatially restricted or mostly spatially restricted. A report PDF for each region, including *in situ* hybridization images, accessibility browser traces, and Z-score and p-value, are found in S2 File.

Promoters used in our analysis were categorized as maternal, maternal-zygotic, or zygotic using previously published RNA-seq data from single embryos [11]. Only zygotically expressed gene promoters were used in Fig 4C. However, all classes of promoters were used in the rest of Fig 4.

### Differential ATAC-seq analysis

All graphs were made with R scripts (S3 File) and Deeptools [84]. Accessibility skew score was calculated by the following equation:

$$\frac{X_{active} - X_{inactive}}{X_{active} + X_{inactive}} \quad (1)$$

where  $X_{active}$  is the wig signal in the half where the region is activating gene expression and  $X_{inactive}$  is wig signal in the half where the region is not supposed to activate gene expression. Accessibility skew score measures whether a region is differentially accessible in the expected direction. This score is useful when comparing differential accessibility regardless of which half is favored (for example when comparing accessibility skew at anterior to posteriorly patterned regions).

Positional skew score provides information about the direction of the skew such that regions that are more accessible in the anterior have a positive positional skew score while those that are more accessible in the posterior have a negative positional skew score. Positional skew score is calculated by the following equation:

$$\frac{X_{anterior} - X_{posterior}}{X_{anterior} + X_{posterior}} \quad (2)$$

Where  $X_{anterior}$  is the wig signal in the anterior sample and  $X_{posterior}$  is the wig signal in the

posterior sample. Significance for each region was determined by computationally matching each region to a random region that has the same total normalized wig score (S10 Fig). Positional skew score was calculated for each random region (termed RandSkewScore). These scores were distributed normally and were used to determine a Z-score for each region of interest ( $Z_{ROI}$ ) by the following equation:

$$Z_{ROI} = \frac{\text{AccessibilitySkewScore} - \mu}{\sigma}$$
 where  $\mu$  is the mean of the random region distribution and  $\sigma$  is the standard deviation of the random region distribution. Two-tailed p-values were then calculated from the Z score.

## Peaks

Replicates were merged and peaks were called on the merged bed file using MACS2 with the following parameters:—nomodel—call-summits—bdg -p 1e-3. Reproducible peaks were selected using bedtools to intersect peaks called in both replicates and the merged samples. Anterior and posterior accessibility signal was averaged using custom scripts around the set of reproducible whole peaks and positional skew scores were calculated for each peak region as described above. Significantly skewed peaks were determined using random regions as with A-P patterning regions. Significantly skewed anterior and posterior peaks as well as unskewed peaks were intersected with experimentally derived promoters, Kvon predicted enhancers [45], and transcription factor ChIP-seq binding data [15,49,85]. One-tailed fisher exact tests were performed to determine whether there was a significant depletion or enrichment of transcription factor peaks in significantly skewed accessibility peaks. Regulatory sequence analysis tools (RSAT) peak motif tool was used to find motifs in each peak set.

## ChIP-seq data analysis

Wig files from previously published ChIP-seq data were obtained for Kruppel, Hunchback, Giant, Knirps, Caudal, Bicoid [15,49,85], and Zelda data from stage 3, 4, and 5 embryos [20]. Wig files were normalized by the mean signal for each sample, assuming that the mean signal over the entire genome is similar to that of background. This normalization essentially transforms the data into deviations from the mean such that signal from different experiments can be compared to each other. Wig signal around regions of interest was determined and graphed in R (S3 File). For the heat maps, normalized wig signal was averaged over 3kb windows around regions of interest for each factor before being scaled such that the region with the highest value is equal to 1 and the lowest to 0 for each factor.

## DNaseI data analysis

For all DnaseI comparisons to ATAC-seq (S2 Fig and Fig 5), previously published DnaseI data from stage 5 embryos (replicate 1) normalized to 10 million reads was used. For S4 Fig, DNaseI data was downloaded from the following SRA datasets: SRA:SRP002474.1, SRA:SRX020691.4, SRA:SRX020692.1, SRA:SRX020693.1, SRA:SRX020694.1, SRA:SRX020695.1, SRA:SRX020696.1, SRA:SRX020697.1, SRA:SRX020698.1, SRA:SRX020699.1, SRA:SRX020700.1, SRA:SRX041410. Reads were processed similarly to ATAC-seq data. Reads were aligned to the dm3 genome with Bowtie2 with the following parameters -p 10-5 5-3 5 -N 1 -X 2000—local—very-sensitive-local. Sam files were then sorted and converted to Bam files using Samtools using the same filters as ATAC-seq samples. Duplicates were removed with Picard. Bams were then converted to bed files using Bedtools and converted into wig files using custom scripts (S1 Code). All replicates were merged and were normalized to 10 million mapped reads.



## Single nuclei ATAC-seq analysis

Bam files labeled to correspond with published clusters were shared by [50] with the authors. Bam files corresponding to anterior and posterior clusters were merged respectively. Merged bam files were then converted into wig files and then normalized to 1 million reads. The anterior wig file was then normalized by linear regression to the posterior wig file. Positional skew score was measured for all A-P and D-V patterning regions in the same manner as was done for halves ATAC-seq data (S4 File).

## Supporting information

**S1 Fig. Linear regression normalization.** Scatter plots showing 1kb genomic bins (gray) with 2D density plots (light blue) indicating areas of increased point density. The  $X = Y$  line is indicated by a dashed line. Linear regression model is indicated by a solid dark blue line. Anterior and posterior halves were normalized to whole samples. (A,C) Scatter plots show data before normalization. (B,D) Scatter plots show data after normalization. Spearman correlation coefficients ( $r_s$ ), Pearson correlation coefficients ( $r_p$ ), and  $r$  squared values are shown above each plot.

(TIF)

**S2 Fig. ATAC-seq on embryo halves correlates highly with DNaseI and whole embryos.** (A) Scatter plots showing merged normalized wig signal values from ATAC-seq experiments performed on anterior halves, posterior halves, and whole embryos compared with DNase-I hypersensitivity from stage 5 *Drosophila melanogaster* embryos [35] binned into 1kb regions. Spearman correlation coefficients ( $r_s$ ), Pearson correlation coefficients ( $r_p$ ), and  $r$  squared values are shown above each plot. The line  $X = Y$  is shown as a dotted line. (B) Scatter plots showing wig signal values for anterior, posterior, and whole replicates normalized to 1 million reads. (C) Normalized wig signal from ATAC-seq data for the combined anterior and posterior halves sample compared to the merged whole embryo sample.

(TIF)

**S3 Fig. Chromatin accessibility in anterior and posterior halves is highly correlated at whole embryo peaks.** (A) Scatter plots showing normalized ATAC-seq signal in anterior (x axis) and posterior (y axis) pooled halves at whole embryo accessibility peaks (gray). Spearman correlation coefficients ( $r_s$ ), Pearson correlation coefficients ( $r_p$ ), and  $r$  squared values are shown above each plot. The line  $X = Y$  is shown as a dotted line. Light blue circles denote point density.

(TIF)

**S4 Fig. Chromatin accessibility changes during development.** Scatter plots showing merged normalized wig signal values from DNaseI hypersensitivity experiments on stage 5,9,11, and 14 *Drosophila melanogaster* embryos [35] binned into 1kb regions. Spearman correlation coefficients ( $r_s$ ), Pearson correlation coefficients ( $r_p$ ), and  $r$  squared values are shown above each plot. The line  $X = Y$  is shown as a dotted line. 2D density plot (light blue) indicates areas of increased point density.

(TIF)

**S5 Fig. D-V patterning enhancers and promoters are similarly accessible in both halves.** Bar graphs showing positional skew score calculated for dorsal (purple) and ventral (green) patterning enhancers and promoters. The enhancer or promoter names are below the graph. (A) D-V enhancers (B) D-V promoters.

(TIF)

**S6 Fig. Replicate and merged positional skew scores at A-P enhancers.** The bar graph represents positional skew scores calculated from the merged replicate samples for all anterior (orange) and posterior (blue) patterning enhancers in the dataset (S1 File). Dots show the positional skew score calculated for both biological replicates. (TIF)

**S7 Fig. Single anterior and posterior halves show similar trends as pooled embryo halves.** (A) Scatter plots showing normalized ATAC-seq signal in anterior (x axis) and posterior (y axis) single halves at 1kb adjacent windows tiling the genome (gray). Spearman correlation coefficients ( $r_s$ ), Pearson correlation coefficients ( $r_p$ ), and  $r$  squared values are shown above each plot. The line  $X = Y$  is shown as a dotted line. (B) Bar graph shows the positional skew scores calculated for all anterior (orange) and posterior (blue) patterning enhancers in the dataset (File S1) and for dorsal (purple) and ventral (green) patterning enhancers. (C) Bar graph shows the positional skew scores calculated for all anterior (orange) and posterior (blue) patterning promoters in the dataset (File S1) and for dorsal (purple) and ventral (green) patterning promoters. The region names are below each graph. (TIF)

**S8 Fig. Significantly skewed peaks show similar differences in transcription factor binding and are predictive of enhancer activity.** (A) ChIP-seq data for Bicoid, Caudal, Knirps, Giant, Hunchback, Kruppel, and Zelda from three stages (stage 3,4, and 5) from [20,49] normalized to the mean of each factor and scaled between 0 and 1 summed over a 3kb window around each reproducible whole peak. White represents the minimum signal and black represents the maximum ChIP signal for that transcription factor. Above each heat map is a colored bar that represents the positional skew score for each reproducible whole peak with orange representing peaks that are more accessible in the anterior, blue representing those that are more accessible in the posterior, and white representing those that are not differentially accessible between the two halves. Peaks are ordered by positional skew score. (B) Positional skew scores were calculated for reproducible whole peaks and z-scores and p-values were determined using random region distributions like for patterning regions. Peaks were then divided into three categories based on skew score and p-value. AntSig peaks are those that are significantly more accessible in the anterior (positive positional skew score and  $p$ -value  $< 0.05$ ). PostSig peaks are those that are significantly more accessible in the posterior (negative positional skew score and  $p$ -value  $< 0.05$ ). NotSig peaks are those that are not significantly skewed in any direction ( $p$ -value  $> 0.05$ ). These three groups of peaks were then intersected using Bedtools with peaks called on ChIP-seq data from (A). One-tailed Fisher's exact tests were performed on these overlaps and the resulting odds ratio are depicted on the logarithmically transformed Y axis such that negative values represent depletions of transcription factor binding in whole peaks and positive values represent enrichments. Asterisks indicate odds ratios that are significant. (C) Transcription factor motifs were searched for in AntSig, PostSig, and NotSig peaksets (methods). The top five predicted motifs are shown with associated number of sites, e-values, and predicted factor match. (D) Replicated whole peaks were intersected with known Kvon enhancers [45]. Matches were then filtered for ones that had annotated expression in stage 4–6 embryos. 12 matches remained. For each of these 12, published slides from *in situ* hybridization experiments were analyzed for stage 5 embryos which are depicted below the bar graph. Using these *in situ* images, the location of activity was determined (represented by bar color—orange represents those active in the anterior, blue represents those active in the posterior, purple represents activity in both halves, gray represents no visible expression). Y-axis depicts positional skew scores calculated for each matching peak region. (TIF)

**S9 Fig. Transcription factor binding and differential accessibility at A-P patterning enhancers.** ChIP-seq data for Bicoid, Caudal, Knirps, Giant, Hunchback, Kruppel, and Zelda from three stages (stage 3,4, and 5) from [20,49] normalized to the mean of each factor and scaled between 0 and 1 summed over a 3kb window around each A-P enhancer. White represents the minimum signal and black represents the maximum ChIP signal for that transcription factor. Above each heat map is a colored bar that represents the positional skew score for each A-P enhancer with orange representing enhancers that are more accessible in the anterior, blue representing those that are more accessible in the posterior, and white representing those that are not differentially accessible between the two halves. (A) Enhancers are ordered by positional skew score. (B) Enhancers are hierarchically clustered using complex heat map package in R [89]. (TIF)

**S10 Fig. Positional skew scores of enhancers and promoters compared to random regions.** Histograms showing the distribution of positional skew scores of random regions compared to (A) A-P patterning enhancers (B) D-V patterning enhancers (C) A-P patterning promoters (D) D-V patterning promoters. Anterior is orange, posterior is blue, dorsal is purple, ventral is green. (E) Histogram showing the distribution of random regions with the fitted normal curve in black. Mu and std from the normal curve are shown above the graph. (TIF)

**S1 Table. Comparisons between A-P and D-V enhancer and promoter mean positional skew scores using Tukey HSD (multiple testing correction).** Pairwise comparisons of mean positional skew score across the different predicted locations (anterior, posterior, dorsal, ventral) as well as sets of random genomic regions chosen to mirror the same total signal distribution of the tested regions. The table is derived from the output of the functions below that can also be found in the supplemental code file (S4 File).

```
Anova_Enhancers_lm = lm(as.numeric(Anova_Enhancers_dataframe
$ATACSkewScore) ~
Anova_Enhancers_dataframe$Location)
Anova_Enhancer = aov(formula = as.numeric(Anova_Enhancers_dataframe$ATACSkewScore) ~
Anova_Enhancers_dataframe$Location)
Tukey_Enhancers = TukeyHSD(Anova_Enhancer)
write.table(Tukey_Enhancers$`Anova_Enhancers_dataframe$Location`,
"020118_Anova_ALLLocations_Enhancers.txt")
```

The output of the Tukey HSD function supplies the identity of the pair-wise comparison, “diff” which is the difference in the observed means, “lwr” which is the lower end point for the interval, “upper” is the upper end point for the interval, and “p adj” is the p-value after adjustments are made for multiple comparisons.

(XLSX)

**S2 Table. ATAC-seq halves mapping metrics.** Cells A3 to D11 show the number of reads that were aligned by Bowtie2 (methods). B3 to B9 contain the total number of reads for each sample. C3 to C9 show the number of concordant read pairs that aligned. D3 to D11 shows the overall alignment rate (including multiple and discordant reads) to represent what percent of the sample came from our expected genome (*Drosophila melanogaster*). Rows 13 to 34 shows the read number changes after filtering out unmapped, not primary alignment, duplicates, failed platform checks, and improper pairs. Rows 23 to 31 show read numbers of each sample that are less than 130 bp. C14 to C34 shows the normalization factor used to normalize each

sample to the number in column D (1 million reads for all replicate samples and 10 million reads for samples where the replicates are merged).

(XLSX)

**S1 File. Average signal and ATAC-seq analysis for A-P and D-V patterning enhancers and promoters—all replicates.** This file contains all of the patterning regions used to make Figs 2,3,4 and 5. These are regions that were filtered to overlap a peak and have spatial localization of activity confirmed by *in situ* hybridization experiments. The values in each column is described in the table included in [S1 File](#).

(XLSX)

**S2 File. Random region file.** Random genomic regions chosen to correspond to A-P and D-V combined patterning enhancers and promoters. File columns are as included in the file.

(XLSX)

**S3 File. Report file for all patterning regions used in Figs 3 and 4.** Reports consist of *in situ* hybridization images, ATAC-seq traces, and calculated p-value and Z Score for each region used in the final analysis.

(ZIP)

**S4 File. Scripts and code.** HTML file of all scripts used in the final analysis.

(HTML)

**S1 Text. Supplementary references.**

(DOCX)

## Acknowledgments

The authors would like to thank members of MBE's research group for thoughtful discussions and comments on an earlier draft of this manuscript. Special thanks to Mawupedzro K Tekpa for help curating *in situ* hybridization data and to Xiao-Yong Li, Michael Stadler, Colleen Hannon, Steven Kuntz, Peter Combs, and Ashley Albright for help optimizing ATAC-seq on frozen embryo slices, data analysis, use of scripts, and figure design. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. Special thanks to the facility director Shana McDevitt for help running ATAC-seq libraries on the Hiseq 4000.

## Author Contributions

**Conceptualization:** Michael B. Eisen.

**Data curation:** Jenna E. Haines.

**Formal analysis:** Jenna E. Haines, Michael B. Eisen.

**Funding acquisition:** Michael B. Eisen.

**Investigation:** Jenna E. Haines, Michael B. Eisen.

**Methodology:** Jenna E. Haines.

**Project administration:** Michael B. Eisen.

**Resources:** Jenna E. Haines, Michael B. Eisen.

**Software:** Jenna E. Haines, Michael B. Eisen.

**Supervision:** Michael B. Eisen.

**Validation:** Jenna E. Haines.

**Visualization:** Jenna E. Haines, Michael B. Eisen.

**Writing – original draft:** Jenna E. Haines, Michael B. Eisen.

**Writing – review & editing:** Jenna E. Haines, Michael B. Eisen.

## References

1. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. *Development*. 2009 Sep; 136(18):3033–42. <https://doi.org/10.1242/dev.033183> PMID: 19700615
2. Li X-Y, Harrison MM, Villalta JE, Kaplan T, Eisen MB. Establishment of regions of genomic activity during the *Drosophila* maternal to zygotic transition. *Elife* [Internet]. 2014 Oct 14; 3. Available from: <http://dx.doi.org/10.7554/eLife.03737>
3. Pérez-Montero S, Carbonell A, Morán T, Vaquero A, Azorín F. The embryonic linker histone H1 variant of *Drosophila*, dBigH1, regulates zygotic genome activation. *Dev Cell*. 2013 Sep 30; 26(6):578–90. <https://doi.org/10.1016/j.devcel.2013.08.011> PMID: 24055651
4. Chen K, Johnston J, Shao W, Meier S, Staber C, Zeitlinger J. A global change in RNA polymerase II pausing during the *Drosophila* midblastula transition. *Elife*. 2013 Aug 13; 2:e00861. <https://doi.org/10.7554/eLife.00861> PMID: 23951546
5. Harrison MM, Eisen MB. Transcriptional Activation of the Zygotic Genome in *Drosophila*. In: *The Maternal-to-Zygotic Transition*. Elsevier; 2015. p. 85–112. (Current Topics in Developmental Biology; vol. 113).
6. Andersen IS, Lindeman LC, Reiner AH, Østrup O, Aanes H, Aleström P, et al. Epigenetic Marking of the Zebrafish Developmental Program. In: *Epigenetics and Development*. Elsevier; 2013. p. 85–112. (Current Topics in Developmental Biology; vol. 104).
7. Bartman CR, Blobel GA. Perturbing Chromatin Structure to Understand Mechanisms of Gene Expression. *Cold Spring Harb Symp Quant Biol* [Internet]. 2015 Sep 14; Available from: <http://dx.doi.org/10.1101/sqb.2015.80.027359>
8. Li B, Carey M, Workman JL. The role of chromatin during transcription. *Cell*. 2007 Feb 23; 128(4):707–19. <https://doi.org/10.1016/j.cell.2007.01.015> PMID: 17320508
9. Ali-Murthy Z, Lott SE, Eisen MB, Kornberg TB. An essential role for zygotic expression in the pre-cellular *Drosophila* embryo. *PLoS Genet*. 2013 Apr; 9(4):e1003428. <https://doi.org/10.1371/journal.pgen.1003428> PMID: 23593026
10. Pritchard DK, Schubiger G. Activation of transcription in *Drosophila* embryos is a gradual process mediated by the nucleocytoplasmic ratio. *Genes Dev*. 1996 May 1; 10(9):1131–42. PMID: 8654928
11. Lott SE, Villalta JE, Schroth GP, Luo S, Tonkin LA, Eisen MB. Noncanonical compensation of zygotic X transcription in early *Drosophila melanogaster* development revealed through single-embryo RNA-seq. *PLoS Biol*. 2011 Feb 8; 9(2):e1000590. <https://doi.org/10.1371/journal.pbio.1000590> PMID: 21346796
12. Nüsslein-Volhard C, Frohnhofer HG, Lehmann R. Determination of anteroposterior polarity in *Drosophila*. *Science*. 1987 Dec 18; 238(4834):1675–81. PMID: 3686007
13. Small S, Levine M. The initiation of pair-rule stripes in the *Drosophila* blastoderm. *Curr Opin Genet Dev*. 1991 Aug; 1(2):255–60. PMID: 1822273
14. Small S, Kraut R, Hoey T, Warrior R, Levine M. Transcriptional regulation of a pair-rule stripe in *Drosophila*. *Genes Dev*. 1991 May; 5(5):827–39. PMID: 2026328
15. Li X-Y, MacArthur S, Bourgon R, Nix D, Pollard DA, Iyer VN, et al. Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol*. 2008 Feb; 6(2):e27. <https://doi.org/10.1371/journal.pbio.0060027> PMID: 18271625
16. Fu D, Ma J. Interplay between positive and negative activities that influence the role of Bicoid in transcription. *Nucleic Acids Res*. 2005 Jul 19; 33(13):3985–93. <https://doi.org/10.1093/nar/gki691> PMID: 16030350
17. Li X-Y, Thomas S, Sabo PJ, Eisen MB, Stamatoyannopoulos JA, Biggin MD. The role of chromatin accessibility in directing the widespread, overlapping patterns of *Drosophila* transcription factor binding. *Genome Biol*. 2011 Apr 7; 12(4):R34. <https://doi.org/10.1186/gb-2011-12-4-r34> PMID: 21473766
18. Liu X, Lee C-K, Granek JA, Clarke ND, Lieb JD. Whole-genome comparison of Leu3 binding in vitro and in vivo reveals the importance of nucleosome occupancy in target site selection. *Genome Res*. 2006 Dec; 16(12):1517–28. <https://doi.org/10.1101/gr.5655606> PMID: 17053089

19. Hannon CE, Blythe SA, Wieschaus EF. Concentration Dependent Chromatin States Induced by the Bicoid Morphogen Gradient. bioRxiv [Internet]. 2017; Available from: <http://biorxiv.org/content/early/2017/05/02/133348.abstract>
20. Harrison MM, Li X-Y, Kaplan T, Botchan MR, Eisen MB. Zelda binding in the early *Drosophila melanogaster* embryo marks regions subsequently activated at the maternal-to-zygotic transition. PLoS Genet. 2011 Oct; 7(10):e1002266. <https://doi.org/10.1371/journal.pgen.1002266> PMID: 22028662
21. Schulz KN, Bondra ER, Moshe A, Villalta JE, Lieb JD, Kaplan T, et al. Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early *Drosophila* embryo. Genome Res. 2015 Nov; 25(11):1715–26. <https://doi.org/10.1101/gr.192682.115> PMID: 26335634
22. Foo SM, Sun Y, Lim B, Ziukaite R, O'Brien K, Nien C-Y, et al. Zelda potentiates morphogen activity by increasing chromatin accessibility. Curr Biol. 2014 Jun 16; 24(12):1341–6. <https://doi.org/10.1016/j.cub.2014.04.032> PMID: 24909324
23. Sun Y, Nien C-Y, Chen K, Liu H-Y, Johnston J, Zeitlinger J, et al. Zelda overcomes the high intrinsic nucleosome barrier at enhancers during *Drosophila* zygotic genome activation. Genome Res. 2015 Nov; 25(11):1703–14. <https://doi.org/10.1101/gr.192542.115> PMID: 26335633
24. Xu Z, Chen H, Ling J, Yu D, Struffi P, Small S. Impacts of the ubiquitous factor Zelda on Bicoid-dependent DNA binding and transcription in *Drosophila*. Genes Dev. 2014 Mar 15; 28(6):608–21. <https://doi.org/10.1101/gad.234534.113> PMID: 24637116
25. Nien C-Y, Liang H-L, Butcher S, Sun Y, Fu S, Gocha T, et al. Temporal coordination of gene networks by Zelda in the early *Drosophila* embryo. PLoS Genet. 2011 Oct; 7(10):e1002339. <https://doi.org/10.1371/journal.pgen.1002339> PMID: 22028675
26. Struffi P, Corado M, Kaplan L, Yu D, Rushlow C, Small S. Combinatorial activation and concentration-dependent repression of the *Drosophila* even-skipped stripe 3+7 enhancer. Development. 2011 Oct; 138(19):4291–9. <https://doi.org/10.1242/dev.065987> PMID: 21865322
27. Liang H-L, Nien C-Y, Liu H-Y, Metzstein MM, Kirov N, Rushlow C. The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*. Nature. 2008 Nov 20; 456(7220):400–3. <https://doi.org/10.1038/nature07388> PMID: 18931655
28. Combs PA, Eisen MB. Genome-wide measurement of spatial expression in patterning mutants of *Drosophila melanogaster*. F1000Res. 2017 Jan 12; 6:41. <https://doi.org/10.12688/f1000research.9720.1> PMID: 28299188
29. Shopland LS, Hirayoshi K, Fernandes M, Lis JT. HSF access to heat shock elements in vivo depends critically on promoter architecture defined by GAGA factor, TFIID, and RNA polymerase II binding sites. Genes Dev. 1995 Nov 15; 9(22):2756–69. PMID: 7590251
30. Bhat KM, Farkas G, Karch F, Gyurkovics H, Gausz J, Schedl P. The GAGA factor is required in the early *Drosophila* embryo not only for transcriptional regulation but also for nuclear division. Development. 1996 Apr; 122(4):1113–24. PMID: 8620838
31. Tsai S-Y, Chang Y-L, Swamy KBS, Chiang R-L, Huang D-H. GAGA factor, a positive regulator of global gene expression, modulates transcriptional pausing and organization of upstream nucleosomes. Epigenetics Chromatin. 2016 Jul 27; 9:32. <https://doi.org/10.1186/s13072-016-0082-4> PMID: 27468311
32. Blythe SA, Wieschaus EF. Establishment and maintenance of heritable chromatin structure during early *Drosophila* embryogenesis. Elife [Internet]. 2016 Nov 23; 5. Available from: <http://dx.doi.org/10.7554/eLife.20148>
33. Blythe SA, Wieschaus EF. Zygotic genome activation triggers the DNA replication checkpoint at the midblastula transition. Cell. 2015 Mar 12; 160(6):1169–81. <https://doi.org/10.1016/j.cell.2015.01.050> PMID: 25748651
34. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013 Dec; 10(12):1213–8. <https://doi.org/10.1038/nmeth.2688> PMID: 24097267
35. Thomas S, Li X-Y, Sabo PJ, Sandstrom R, Thurman RE, Canfield TK, et al. Dynamic reprogramming of chromatin accessibility during *Drosophila* embryo development. Genome Biol. 2011 May 11; 12(5):R43. <https://doi.org/10.1186/gb-2011-12-5-r43> PMID: 21569360
36. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008 Sep 17; 9(9):R137. <https://doi.org/10.1186/gb-2008-9-9-r137> PMID: 18798982
37. Chen H, Xu Z, Mei C, Yu D, Small S. A system of repressor gradients spatially organizes the boundaries of Bicoid-dependent target genes. Cell. 2012 Apr 27; 149(3):618–29. <https://doi.org/10.1016/j.cell.2012.03.018> PMID: 22541432

38. Ochoa-Espinosa A, Yucel G, Kaplan L, Pare A, Pura N, Oberstein A, et al. The role of binding site cluster strength in Bicoid-dependent patterning in *Drosophila*. *Proc Natl Acad Sci U S A*. 2005 Apr 5; 102(14):4960–5. <https://doi.org/10.1073/pnas.0500373102> PMID: 15793007
39. Schroeder MD, Pearce M, Fak J, Fan H, Unnerstall U, Emberly E, et al. Transcriptional control in the segmentation gene network of *Drosophila*. *PLoS Biol*. 2004 Sep; 2(9):E271. <https://doi.org/10.1371/journal.pbio.0020271> PMID: 15340490
40. Riddihough G, Ish-Horowitz D. Individual stripe regulatory elements in the *Drosophila* hairy promoter respond to maternal, gap, and pair-rule genes. *Genes Dev*. 1991 May; 5(5):840–54. PMID: 1902805
41. Hartmann B, Reichert H, Walldorf U. Interaction of gap genes in the *Drosophila* head: tailless regulates expression of empty spiracles in early embryonic patterning and brain development. *Mech Dev*. 2001 Dec; 109(2):161–72. PMID: 11731230
42. Biemar F, Zinzen R, Ronshaugen M, Sementchenko V, Manak JR, Levine MS. Spatial regulation of microRNA gene expression in the *Drosophila* embryo. *Proc Natl Acad Sci U S A*. 2005 Nov 1; 102(44):15907–11. <https://doi.org/10.1073/pnas.0507817102> PMID: 16249329
43. Berman BP, Pfeiffer BD, Laverty TR, Salzberg SL, Rubin GM, Eisen MB, et al. Computational identification of developmental enhancers: conservation and function of transcription factor binding-site clusters in *Drosophila melanogaster* and *Drosophila pseudoobscura*. *Genome Biol*. 2004 Aug 20; 5(9):R61. <https://doi.org/10.1186/gb-2004-5-9-r61> PMID: 15345045
44. Gallo SM, Gerrard DT, Miner D, Simich M, Des Soye B, Bergman CM, et al. REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in *Drosophila*. *Nucleic Acids Res*. 2011 Jan; 39(Database issue):D118–23. <https://doi.org/10.1093/nar/gkq999> PMID: 20965965
45. Kvon EZ, Kazmar T, Stampfel G, Yáñez-Cuna JO, Pagani M, Schernhuber K, et al. Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo. *Nature*. 2014 Aug 7; 512(7512):91–5. <https://doi.org/10.1038/nature13395> PMID: 24896182
46. Hammonds AS, Bristow CA, Fisher WW, Weiszmann R, Wu S, Hartenstein V, et al. Spatial expression of transcription factors in *Drosophila* embryonic organ development. *Genome Biol*. 2013 Dec 20; 14(12):R140. <https://doi.org/10.1186/gb-2013-14-12-r140> PMID: 24359758
47. Tomancak P, Beaton A, Weiszmann R, Kwan E, Shu S, Lewis SE, et al. Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biol*. 2002 Dec 23; 3(12):RESEARCH0088. <https://doi.org/10.1186/gb-2002-3-12-research0088> PMID: 12537577
48. Tomancak P, Berman BP, Beaton A, Weiszmann R, Kwan E, Hartenstein V, et al. Global analysis of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biol*. 2007; 8(7):R145. <https://doi.org/10.1186/gb-2007-8-7-r145> PMID: 17645804
49. Bradley RK, Li X-Y, Trapnell C, Davidson S, Pachter L, Chu HC, et al. Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related *Drosophila* species. *PLoS Biol*. 2010 Mar 23; 8(3):e1000343. <https://doi.org/10.1371/journal.pbio.1000343> PMID: 20351773
50. Cusanovich DA, Reddington JP, Garfield DA, Daza R, Marco-Ferreres R, Christiansen L, et al. The cis-regulatory dynamics of embryonic development at single cell resolution [Internet]. *bioRxiv*. 2017 [cited 2017 Sep 1], p. 166066. Available from: <http://www.biorxiv.org/content/early/2017/07/20/166066.abstract>
51. Mir M, Reimer A, Haines JE, Li X-Y, Stadler M, Garcia H, et al. Dense Bicoid hubs accentuate binding along the morphogen gradient. *Genes Dev*. 2017 Sep 1; 31(17):1784–94. <https://doi.org/10.1101/gad.305078.117> PMID: 28982761
52. Koenecke N, Johnston J, He Q, Meier S, Zeitlinger J. *Drosophila* poised enhancers are generated during tissue patterning with the help of repression. *Genome Res*. 2017 Jan; 27(1):64–74. <https://doi.org/10.1101/gr.209486.116> PMID: 27979994
53. John S, Sabo PJ, Thurman RE, Sung M-H, Biddie SC, Johnson TA, et al. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nat Genet*. 2011 Mar; 43(3):264–8. <https://doi.org/10.1038/ng.759> PMID: 21258342
54. Reynolds N, O’Shaughnessy A, Hendrich B. Transcriptional repressors: multifaceted regulators of gene expression. *Development*. 2013 Feb 1; 140(3):505–12. <https://doi.org/10.1242/dev.083105> PMID: 23293282
55. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010 Mar 15; 26(6):841–2. <https://doi.org/10.1093/bioinformatics/btq033> PMID: 20110278
56. Koenecke N, Johnston J, Gaertner B, Natarajan M, Zeitlinger J. Genome-wide identification of *Drosophila* dorso-ventral enhancers by differential histone acetylation analysis. *Genome Biol*. 2016 Sep 27; 17(1):196. <https://doi.org/10.1186/s13059-016-1057-2> PMID: 27678375

57. Hoskins RA, Landolin JM, Brown JB, Sandler JE, Takahashi H, Lassmann T, et al. Genome-wide analysis of promoter architecture in *Drosophila melanogaster*. *Genome Res.* 2011 Feb; 21(2):182–92. <https://doi.org/10.1101/gr.112466.110> PMID: 21177961
58. Dreos R, Ambrosini G, Groux R, Cavin P rier R, Bucher P. The eukaryotic promoter database in its 30th year: focus on non-vertebrate organisms. *Nucleic Acids Res.* 2017 Jan 4; 45(D1):D51–5. <https://doi.org/10.1093/nar/gkw1069> PMID: 27899657
59. Dreos R, Ambrosini G, P rier RC, Bucher P. The Eukaryotic Promoter Database: expansion of EPDnew and new promoter analysis tools. *Nucleic Acids Res.* 2015 Jan; 43(Database issue):D92–6. <https://doi.org/10.1093/nar/gku1111> PMID: 25378343
60. Doctor JS, Jackson PD, Rashka KE, Visalli M, Hoffman FM. Sequence, biochemical characterization, and developmental expression of a new member of the TGF-  superfamily in *Drosophila melanogaster*. *Dev Biol.* 1992 Jun 1; 151(2):491–505. PMID: 1601181
61. Rothberg JM, Hartley DA, Walther Z, Artavanis-Tsakonas S. slit: an EGF-homologous locus of *D. melanogaster* involved in the development of the embryonic central nervous system. *Cell.* 1988 Dec 23; 55(6):1047–59. PMID: 3144436
62. Wimmer EA, Simpson-Brose M, Cohen SM, Desplan C, J ckle H. Trans- and cis-acting requirements for blastodermal expression of the head gap gene buttonhead. *Mech Dev.* 1995 Oct; 53(2):235–45. PMID: 8562425
63. Johnson K, Knust E, Skaer H. bloated tubules (blot) encodes a *Drosophila* member of the neurotransmitter transporter family required for organisation of the apical cytotocortex. *Dev Biol.* 1999 Aug 15; 212(2):440–54. <https://doi.org/10.1006/dbio.1999.9351> PMID: 10433833
64. Kearney JB, Wheeler SR, Estes P, Parente B, Crews ST. Gene expression profiling of the developing *Drosophila* CNS midline cells. *Dev Biol.* 2004 Nov 15; 275(2):473–92. <https://doi.org/10.1016/j.ydbio.2004.08.047> PMID: 15501232
65. Klingler M, Soong J, Butler B, Gergen JP. Disperse versus compact elements for the regulation of runt stripes in *Drosophila*. *Dev Biol.* 1996 Jul 10; 177(1):73–84. <https://doi.org/10.1006/dbio.1996.0146> PMID: 8660878
66. Szab  K, J kely G, R rth P. Cloning and expression of sprint, a *Drosophila* homologue of RIN1. *Mech Dev.* 2001 Mar; 101(1–2):259–62. PMID: 11231087
67. Kellerman KA, Miller KG. An unconventional myosin heavy chain gene from *Drosophila melanogaster*. *J Cell Biol.* 1992 Nov; 119(4):823–34. PMID: 1429838
68. Starz-Gaiano M, Cho NK, Forbes A, Lehmann R. Spatially restricted activity of a *Drosophila* lipid phosphatase guides migrating germ cells. *Development.* 2001 Mar; 128(6):983–91. PMID: 11222152
69. Howard KR, Struhl G. Decoding positional information: regulation of the pair-rule gene hairy. *Development.* 1990 Dec; 110(4):1223–31. PMID: 2100260
70. Sahut-Barnola I, Pauli D. The *Drosophila* gene stand still encodes a germline chromatin-associated protein that controls the transcription of the ovarian tumor gene. *Development.* 1999 May; 126(9):1917–26. PMID: 10101125
71. Renault AD, Zhang X-H, Alphey LS, Frenz LM, Glover DM, Saunders RDC, et al. giant nuclei is essential in the cell cycle transition from meiosis to mitosis. *Development.* 2003 Jul; 130(13):2997–3005. PMID: 12756181
72. Ch vez VM, Marqu s G, Delbecque JP, Kobayashi K, Hollingsworth M, Burr J, et al. The *Drosophila* disembodied gene controls late embryonic morphogenesis and codes for a cytochrome P450 enzyme that regulates embryonic ecdysone levels. *Development.* 2000 Oct; 127(19):4115–26. PMID: 10976044
73. Chang T, Mazotta J, Dumstrei K, Dumitrescu A, Hartenstein V. Dpp and Hh signaling in the *Drosophila* embryonic eye field. *Development.* 2001 Dec; 128(23):4691–704. PMID: 11731450
74. Gould AP, Brookman JJ, Strutt DI, White RA. Targets of homeotic gene control in *Drosophila*. *Nature.* 1990 Nov 22; 348(6299):308–12. <https://doi.org/10.1038/348308a0> PMID: 1979146
75. Doyle HJ, Kraut R, Levine M. Spatial regulation of zerkn llt: a dorsal-ventral patterning gene in *Drosophila*. *Genes Dev.* 1989 Oct; 3(10):1518–33. PMID: 2612903
76. Schnepf B, Grumblin G, Donaldson T, Simcox A. Vein is a novel component in the *Drosophila* epidermal growth factor receptor pathway with similarity to the neuregulins. *Genes Dev.* 1996 Sep 15; 10(18):2302–13. PMID: 8824589
77. Mart n-Blanco E, Gampel A, Ring J, Virdee K, Kirov N, Tolkovsky AM, et al. puckered encodes a phosphatase that mediates a feedback loop regulating JNK activity during dorsal closure in *Drosophila*. *Genes Dev.* 1998 Feb 15; 12(4):557–70. PMID: 9472024
78. Lu Y, Settleman J. The *Drosophila* Pkn protein kinase is a Rho/Rac effector target required for dorsal closure during embryogenesis. *Genes Dev.* 1999 May 1; 13(9):1168–80. PMID: 10323867



79. Ozon S, Guichet A, Gavet O, Roth S, Sobel A. *Drosophila* stathmin: a microtubule-destabilizing factor involved in nervous system formation. *Mol Biol Cell*. 2002 Feb; 13(2):698–710. <https://doi.org/10.1091/mbc.01-07-0362> PMID: 11854423
80. Perry MW, Boettiger AN, Levine M. Multiple enhancers ensure precision of gap gene-expression patterns in the *Drosophila* embryo. *Proc Natl Acad Sci U S A*. 2011 Aug 16; 108(33):13570–5. <https://doi.org/10.1073/pnas.1109873108> PMID: 21825127
81. Saebøe-Larssen S, Mohebi BU, Lambertsson A. The *Drosophila* ribosomal protein L14-encoding gene, identified by a novel Minute mutation in a dense cluster of previously undescribed genes in cytogenetic region 66D. *Mol Gen Genet*. 1997; 255(2):141–51. PMID: 9236770
82. Lécuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, et al. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell*. 2007 Oct 5; 131(1):174–87. <https://doi.org/10.1016/j.cell.2007.08.003> PMID: 17923096
83. Wilk R, Hu J, Blotsky D, Krause HM. Diverse and pervasive subcellular distributions for both coding and long noncoding RNAs. *Genes Dev*. 2016 Mar 1; 30(5):594–609. <https://doi.org/10.1101/gad.276931.115> PMID: 26944682
84. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res*. 2016 Jul 8; 44(W1):W160–5. <https://doi.org/10.1093/nar/gkw257> PMID: 27079975
85. MacArthur S, Li X-Y, Li J, Brown JB, Chu HC, Zeng L, et al. Developmental roles of 21 *Drosophila* transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions. *Genome Biol*. 2009 Jul 23; 10(7):R80. <https://doi.org/10.1186/gb-2009-10-7-r80> PMID: 19627575
86. Fujioka M, Emi-Sarker Y, Yusibova GL, Goto T, Jaynes JB. 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*. 1999 Jun; 126(11):2527–38. PMID: 10226011
87. Driever W, Thoma G, Nüsslein-Volhard C. Determination of spatial domains of zygotic gene expression in the *Drosophila* embryo by the affinity of binding sites for the bicoid morphogen. *Nature*. 1989 Aug 3; 340(6232):363–7. <https://doi.org/10.1038/340363a0> PMID: 2502714
88. Häder T, La Rosée A, Ziebold U, Busch M, Taubert H, Jäckle H, et al. Activation of posterior pair-rule stripe expression in response to maternal caudal and zygotic knirps activities. *Mech Dev*. 1998 Feb; 71(1–2):177–86. PMID: 9507113
89. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016 Sep 15; 32(18):2847–9. <https://doi.org/10.1093/bioinformatics/btw313> PMID: 27207943