The Zelda Interactome Reveals Diverse Co-factors Essential for Pioneer Factor-Mediated Zygotic Genome Activation

Xiao-Yong Li[1] and Michael B. Eisen[1,2,3,+]

1: Howard Hughes Medical Institute, University of California, Berkeley, CA
2: Department of Molecular and Cell Biology, University of California, Berkeley, CA
3: Department of Integrative Biology, University of California, Berkeley, CA
+: Correspondence: mbeisen@gmail.com

ABSTRACT

Zelda (Zld) is the master pioneer transcription factor essential for zygotic genome activation (ZGA) during the *Drosophila* maternal-to-zygotic transition (MZT). While Zld is known to promote chromatin accessibility at early enhancers and possess multiple conserved functional domains, the specific protein complexes it recruits to execute its functions remain poorly defined. Using an optimized immunoprecipitation and mass spectrometry (IP-MS) pipeline in Stage 4-5 embryos, we identified a diverse Zld interactome. This repertoire includes the coactivators dCBP and Fsh (the *Drosophila* Brd4 ortholog), subunits of major nucleosome remodeling complexes(e.g. PBAP/BAP, NURF, FACT), RNA polymerase II (RNAPII), the corepressor Smrter (Smr), and the Tousled-like kinase (Tlk). We confirmed the enrichment of ten key factors at Zld-bound regions using CUT&RUN. Notably, our results demonstrate that RNAPII associates not only with Zld-bound promoters but also with distal Zld-binding sites. This finding, which contrasts with previous ChIP-based studies, suggests that Zld facilitates transcription by actively scaffolding or pre-recruiting the transcriptional machinery at enhancers. Functional analyses revealed that dCBP and Fsh are required for Zld-mediated activation, while RNAi knockdown of *smr* led to the broad derepression of Zld-target genes, suggesting that Smr modulates Zld activity to prevent premature expression. Finally, we show that Tlk, a kinase typically associated with DNA replication, directly interacts with a specific Zld domain *in vitro* and co-localizes with Zld *in vivo*. This suggests a novel mechanism by which Zld may coordinate transcriptional activation with the rapid mitotic cycles of early embryogenesis. Collectively, our findings provide a comprehensive map of the Zld interactome and reveal how a pioneer factor integrates diverse chromatin and transcriptional regulators to orchestrate zygotic genome activation.

Keywords: Zelda, ZGA, Interactome, Tlk, dCBP/p300, Fsh/BRD4, Chromatin remodeling, Transcriptional hubs, Pioneer factor

INTRODUCTION

In multicellular organisms, precise spatial and temporal gene expression is governed by enhancers - regulatory DNA sequences harboring binding sites for sequence-specific transcription factors(TFs) [1–3]. Most enhancers utilize unique combinations of these sites to integrate diverse developmental cues into specific transcriptional outputs. Within the nuclear chromatin context, distinct TFs may serve specialized roles in enhancer activity.

Eukaryotic genomes are organized into chromatin through the assembly of nucleosomes [4], and inactive enhancers, in particular, are characterized by high nucleosome occupancy *in vivo* [5–9]. This wrapping of DNA around histone octamers in nucleosomes can occlude DNA binding motifs and restrict TF binding through steric hindrance. Consequently, many TFs are unable to access their motifs embedded in nucleosomes *in vitro* [10,11], and bind predominantly to accessible sites *in vivo* [12–15]. However, a specialized class of TFs termed "pioneer factors" possess the intrinsic ability to bind motifs embedded in nucleosomes *in vitro* and within closed chromatin *in vivo* [16,17]. Upon binding, such factors are able to promote local chromatin accessibility, facilitating the binding by other factors to achieve transcription activation. Pioneer factors are thus essential for establishing the initial regulatory landscape during zygotic genome activation (ZGA) and for reshaping gene networks during developmental processes such as cell fate determination and lineage specification, as well as cell reprogramming [16–22].

Zld is a master pioneer factor essential for zygotic genome activation in *D. melanogaster* [23]. It binds broadly across the genome, targeting the majority of zygotic enhancers [24,25]. Notably, Zld binding occurs very early - often before the onset of robust transcription and prior to the recruitment of non-pioneer patterning factors such as Dorsal (Dl) and Caudal (Cad) [24,26]. Zld is required for histone depletion, increased chromatin accessibility, and histone acetylation at zygotic enhancers [7,8,26–30]; and it exerts a significantly more profound effect on chromatin accessibility at enhancers than typical embryo patterning factors [8,30]. By establishing accessible chromatin, Zld facilitates the subsequent binding of other factors and the activation of zygotic genes [7,8,28,30–33]. Additionally, recent studies suggest Zld may also function by forming nuclear hubs that increase the local concentration of other transcriptional regulators [33–36].

Beyond sequence-specific TFs, enhancer function requires a diverse repertoire of factors, including coactivators, chromatin modifiers, and nucleosome remodelers [37–47]. To date, however, the specific protein complexes that collaborate with pioneer factors, such as Zld, to execute their functions remain poorly defined. Given Zld's role in actively promoting chromatin accessibility, we anticipated that it interacts with specific histone modifiers and nucleosome remodelers, as has been observed for other pioneer factors [20,22]. Structurally, Zld contains multiple conserved domains, including a putative activation domain [48] and a domain with repressive activity [49]. These domains likely recruit the basal transcription machinery, coactivators, and potentially unknown regulators to coordinate transcription during the rapid, synchronous nuclear divisions of early development.

In this study, we utilize an optimized immunoprecipitation and high-resolution mass spectrometry (IP-MS) approach to provide a comprehensive proteomic and genomic map of the Zld interactome during the D. melanogaster maternal-to-zygotic transition (MZT). Our findings reveal that Zld recruits a diverse repertoire of co-factors, including coactivators, nucleosome remodelers, RNA polymerase II and its associated factors, the corepressor Smarter (Smr), and the Tousled-like kinase (Tlk). We performed CUT&RUN analysis for ten of these factors, demonstrating their association with Zld binding sites in embryos. Furthermore, we investigated the functional relationships between Zld and several enriched factors. We found that the coactivators dCBP (encoded by *nej*) [50] and Fsh (encoded by *fs(1)h*) [51]—the latter being the ortholog of the mammalian BET (bromodomain and extra-terminal domain) factor Brd4—are essential for the transcription of Zld target genes. We also found that Smr [52,53] exerts a broad repressive effect on Zld target gene transcription. Finally, based on *in vitro* binding assays, we found that Tlk, a cell-cycle regulated kinase with activity peaking during S-phase [54], interacts directly with a specific Zld domain. This suggests a novel mechanism by which Zld may coordinate transcriptional activation with the rapid mitotic cycles during early embryogenesis. Collectively, these results elucidate how Zld integrates chromatin remodeling with the active recruitment of the transcriptional apparatus to orchestrate the initial wave of zygotic genome activation.

RESULTS

Identification of the Zld interactome via optimized IP-MS

To identify protein factors associated with Zld in embryos, we developed an optimized IP-MS approach based on established protocols [37–40,55,56](Fig.1A-B). To ensure efficient fixation, crosslinking was performed by homogenizing embryos in a formaldehyde-containing buffer followed by an extended incubation. For lysate preparation from isolated nuclei, we employed only mild sonication to avoid the disruption of protein complexes. Full solubilization was achieved via Benzonase treatment; by completely digesting DNA and RNA, this step prioritized the identification of factors associating with Zld through protein-protein interactions rather than tethering via nucleic acids. Following immunoprecipitation, the isolated proteins were digested with trypsin, and the resulting peptides were analyzed by high-resolution LC-MS/MS. Data analysis was performed using MaxQuant with label-free quantification (LFQ) [57], followed by statistical analysis in Perseus [58].

To define the protein complexes associated with Zld, we utilized a Zld antibody previously validated for high efficiency in ChIP-seq analysis [24]. This antibody exhibits high specificity; immunostaining analysis produced no signal in *zld1-/-* germline clone embryos, in contrast to the robust signal observed in wild-type (WT) embryos. As a positive control for our IP-MS pipeline, we performed IPs using an antibody against Fibrillarin (Fib), a nucleolar protein involved in rRNA processing with a well-characterized set of associated factors [59]. Additionally, we

conducted parallel analyses using anti-GFP antibodies in embryos endogenously expressing GFP-Zld.

General characteristics of the enriched protein factors

Immunoprecipitation using the anti-Zld antibody was performed following both a standard protocol and a modified version incorporating stringent 2M urea wash steps to identify stable

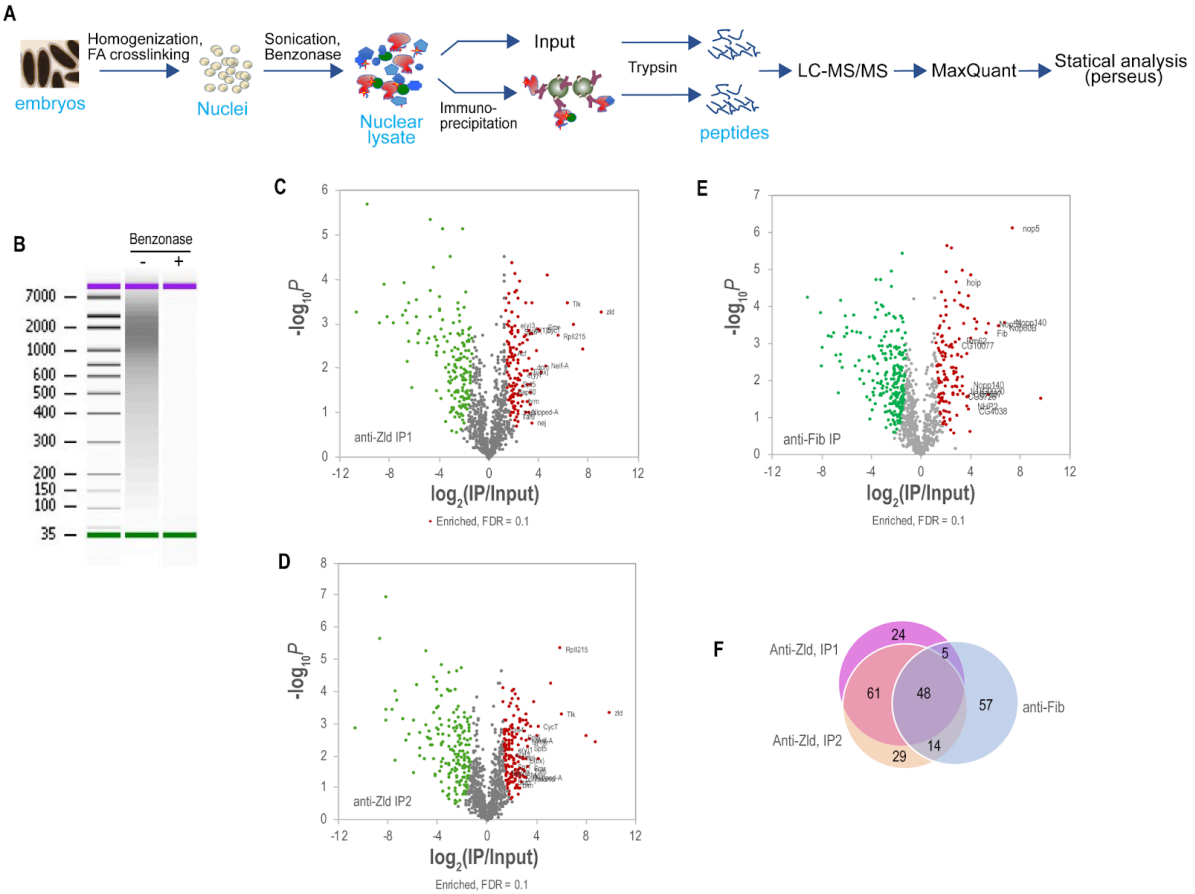

Fig. 1. Identification of the Zelda (Zld) interactome in early *D. melanogaster* embryos.

A) **Schematic of the IP-MS workflow.** Embryos were crosslinked, nuclei isolated, and chromatin solubilized with Benzonase prior to immunoprecipitation and LC-MS/MS analysis.

B) **Efficient DNA digestion by Benzonase.** Bioanalyzer traces show chromatin DNA fragments from the nuclear lysate before (–) and after (+) treatment with Benzonase.

C-E) **Volcano plots showing protein enrichment in Zld and Fib IPs.** Data are presented as –log10 p-value vs. log2(IP/Input). IPs were performed using an anti-Zld antibody under standard conditions (**C**), with stringent 2M urea washes (**D**), or using an anti-Fib antibody as a control (**E**). Enriched proteins (FDR < 0.1) are highlighted in red. The FDR was determined using a combination of p-values and enrichment ratios in Perseus(Tyanova, Temu, Sinitcyn, et al. 2016). Key factors associated with chromatin modulation and transcription (**C, D**) or rRNA processing (**E**) are labeled.

F) **Overlap of enriched factors.** Venn diagram showing the intersection of protein factors identified across the three IP experiments.

interactors. Under these conditions, we identified 138 and 152 enriched proteins, respectively (FDR < 0.1; Fig. 1, Supplementary Table 1). The two datasets showed a high degree of overlap,

Table 1. Top 10 enriched GO terms for the protein factors enriched in the anti-Zld IP-MS and anti-Fib IP-MS

anti-Zld IP1

	No. factors	Enrichment	raw P-value
GO biological process complete			
transcription initiation at RNA polymerase II promoter (GO:0006367)	7	4.4	2.87E-04
negative regulation of transcription by RNA polymerase II (GO:0000122)	10	2.9	1.03E-03
positive regulation of transcription by RNA polymerase II (GO:0045944)	13	2.9	1.85E-04
regulation of mRNA splicing, via spliceosome (GO:0048024)	17	2.8	2.07E-05
positive regulation of gene expression (GO:0010628)	17	2.7	4.04E-05
chromatin remodeling (GO:0006338)	22	2.3	8.05E-05
mRNA splicing, via spliceosome (GO:0000398)	25	2.2	4.62E-05
GO molecular function complete			
RNA polymerase II general transcription initiation factor activity (GO:0016251)	4	6.0	1.36E-03
transcription corepressor activity (GO:0003714)	8	4.3	1.22E-04
RNA polymerase II transcription regulatory region sequence-specific DNA binding (GO:0000977)	9	4.2	5.19E-05
transcription coactivator activity (GO:0003713)	8	3.8	4.13E-04
DNA-binding transcription factor activity (GO:0003700)	10	3.8	7.32E-05
protein heterodimerization activity (GO:0046982)	8	3.3	1.11E-03
transcription factor binding (GO:0008134)	11	3.1	3.12E-04
chromatin binding (GO:0003682)	28	2.9	1.20E-08
histone binding (GO:0042393)	10	2.9	1.03E-03
mRNA binding (GO:0003729)	31	2.1	1.17E-05
GO cellular component complete			
transcription repressor complex (GO:0017053)	6	7.5	5.09E-06
SAGA-type complex (GO:0070461)	4	6.0	1.36E-03
polytene chromosome interband (GO:0005705)	8	3.5	6.93E-04
heterochromatin (GO:0000792)	7	3.5	1.64E-03
RNA polymerase II, holoenzyme (GO:0016591)	7	3.3	2.59E-03
precatalytic spliceosome (GO:0071011)	19	2.0	1.46E-03

anti-Fib

	No. factors	Enrichment	raw P-value
GO biological process complete			
rRNA pseudouridine synthesis (GO:0031118)	4	7.6	2.92E-04
regulation of transcriptional start site selection at RNA polymerase II promoter (GO:0001178)	5	5.4	6.27E-04
regulation of mRNA 3'-end processing (GO:0031440)	6	5.0	2.83E-04
mRNA destabilization (GO:0061157)	7	3.8	9.14E-04
nuclear-transcribed mRNA catabolic process (GO:0000956)	8	3.6	6.37E-04
regulation of alternative mRNA splicing, via spliceosome (GO:0000381)	15	3.4	3.54E-06
mRNA splicing, via spliceosome (GO:0000398)	27	3.0	7.41E-09
regulation of translation (GO:0006417)	15	2.7	1.18E-04
GO molecular function complete			
snoRNA binding (GO:0030515)	6	7.6	4.74E-06
mRNA 3'-UTR binding (GO:0003730)	8	3.6	6.37E-04
GO cellular component complete			
box C/D RNP complex (GO:0031428)	4	7.6	2.92E-04
box H/ACA RNP complex (GO:0072588)	4	7.6	2.92E-04
nuclear speck (GO:0016607)	12	6.1	5.32E-09
small-subunit processome (GO:0032040)	9	5.2	4.13E-06
precatalytic spliceosome (GO:0071011)	21	2.9	1.09E-06
catalytic step 2 spliceosome (GO:0071013)	19	2.9	4.23E-06

anti-Zld IP2

	No. factors	Enrichment	raw P-value
GO biological process complete			
positive regulation of transcription elongation by RNA polymerase II (GO:0032968)	6	5.3	1.65E-04
larval somatic muscle development (GO:0007526)	7	4.5	2.05E-04
transcription elongation by RNA polymerase II (GO:0006368)	7	4.5	2.05E-04
transcription initiation at RNA polymerase II promoter (GO:0006367)	7	4.1	4.35E-04
regulation of chromatin organization (GO:1902275)	8	3.8	3.70E-04
negative regulation of transcription by RNA polymerase II (GO:0000122)	11	3.0	3.76E-04
regulation of alternative mRNA splicing, via spliceosome (GO:0000381)	15	2.9	3.66E-05
mRNA splicing, via spliceosome (GO:0000398)	27	2.8	5.38E-08
positive regulation of gene expression (GO:0010628)	16	2.7	7.41E-05
epigenetic regulation of gene expression (GO:0040029)	12	2.6	7.59E-04
negative regulation of gene expression (GO:0010629)	23	2.2	7.07E-05
GO molecular function complete			
RNA polymerase II-specific DNA-binding transcription factor binding (GO:0061629)	6	5.3	1.65E-04
DNA-binding transcription factor activity, RNA polymerase II-specific (GO:0000981)	9	4.9	8.09E-06
transcription coactivator activity (GO:0003713)	10	4.7	3.99E-06
RNA polymerase II transcription regulatory region sequence-specific DNA binding (GO:0000977)	9	4.2	4.38E-05
basal RNA polymerase II transcription machinery binding (GO:0001099)	6	4.2	9.63E-04
transcription corepressor activity (GO:0003714)	8	3.8	3.70E-04
chromatin DNA binding (GO:0031490)	9	3.5	2.90E-04
protein heterodimerization activity (GO:0046982)	8	3.1	1.72E-03
mRNA binding (GO:0003729)	31	2.2	5.84E-06
GO cellular component complete			
transcription repressor complex (GO:0017053)	5	5.9	2.91E-04
polytene chromosome band (GO:0005704)	4	5.6	1.76E-03
SAGA-type complex (GO:0070461)	4	5.6	1.76E-03
transcription elongation factor complex (GO:0008023)	9	4.9	8.09E-06
polytene chromosome interband (GO:0005705)	10	4.4	9.36E-06
heterochromatin (GO:0000792)	9	4.2	4.38E-05
RNA polymerase II, holoenzyme (GO:0016591)	8	4.0	1.96E-04
polytene chromosome puff (GO:0005703)	12	3.4	4.32E-05
nuclear speck (GO:0016607)	7	3.1	3.80E-03
SWI/SNF superfamily-type complex (GO:0070603)	11	3.0	3.76E-04
catalytic step 2 spliceosome (GO:0071013)	21	2.9	1.02E-06
euchromatin (GO:0000791)	9	2.9	1.78E-03
precatalytic spliceosome (GO:0071011)	22	2.8	1.44E-06

		Anti-Fib	Anti-Zld IP1	Anti-Zld IP2	
box C/D snoRNP complex	Fib	5.3	0.5	0.3	Fib associated (STRING)
	nop5	7.3	0.1	0.6	
	Nop56	6.4	-1.6	-1.9	
	Nopp140	6.8	-0.8	1.3	
	hoip	2.8	-0.8	-0.6	
box H/ACA snoRNP complex	Nop60B	6.2	-0.5	-0.7	
	CG4038	4.3	-1.5	-0.5	
	NHP2	3.6	NA	NA	

Fig. 1_Supplement 1. Enrichment of protein factors predicted to interact with Fib in anti-Fib IP-MS experiment. The enrichment values ($\log_2$(IP/Input)) for the seven factors that are among the top 10 predicted Fib associated factors in the STRING database [60] in the anti-Fib and anti-Zld IP-MS experiments are shown.

with 109 factors in common. In contrast, the control Fib-IP enriched 125 factors; of these, only 53 and 62 overlapped with the two anti-Zld IP-MS experiments, respectively.

Gene Ontology (GO) analysis revealed that Zld-associated factors were primarily related to RNA polymerase II (RNAPII) transcription, chromatin remodeling, and the regulation of chromatin organization (Table 1). Conversely, Fib-enriched factors were associated with rRNA

modification and snoRNP complexes. We further validated our pipeline by cross-referencing Fib-enriched factors with the STRING database[60]; seven of the top ten predicted Fib interactors were significantly enriched (Fig. 1_Supplement 1), while the remaining three were not detected in our MS dataset. Collectively, our GO and STRING analyses demonstrate that the Zld and Fib interactomes are distinct and align with their respective biological roles.

Functional categories of the Zld interactome

To generate a consolidated Zld interactome, we integrated results from both anti-Zld IP experiments, excluding factors with less than a two-fold enrichment ($\log_2$(IP/Input) < 1) in either dataset. This yielded 155 factors, 93 of which were specific to the Zld IPs (Fig. 2A).

As shown in the interaction map (Fig. 2B), the most prominent group of interactors consists of factors involved in histone acetylation and nucleosome remodeling. These include the coactivators dCBP and Fsh, as well as subunits of major nucleosome remodeling complexes [61,62], such as the NURF complex (E(bx)), the PBAP/BAP complexes (Brm, SAYP, Polybromo, and Bap60), and the ACF complex. We also observed strong enrichment of Domino (DOM), a subunit of the TIP60 and SWR1 complexes [63], and the histone chaperone FACT complex (Dre4 and Ssrp) [64]. The recruitment of such a diverse array of remodelers suggests that Zld may directly recruit these machines to establish accessibility at enhancers, reminiscent of the pioneer factor GAF [65,66].

While general transcription factors (GTFs) were sparsely represented, we observed striking enrichment of the RNA polymerase II large subunit (Rpb1; hereafter referred to as RNAPII), which ranked as the second-highest Zld-specific interactor. Several of its associated factors were also enriched, including the negative elongation factor (NELF-A), the elongation factor P-TEFb (CycT), Spt5, Spt6(Fig. 2B-C). NELF-A and CycT were also enriched in the anti-Fib IP, but at much lower levels than observed in the Zld IPs. Beyond these factors required for chromatin accessibility and transactivation - the coactivators (e.g., dCBP and Fsh), nucleosome remodelers, and the RNAPII complex - several factors involved in transcriptional repression were also enriched. Among the most highly enriched proteins was the corepressor Smrter (Smr), identified along with its established partners Ebi and Sin3A [52,53]. This finding is particularly noteworthy as Zld has been shown to contain a transcription repression domain [49], suggesting that Zld may recruit Smr to modulate or "buffer" the output of its target genes.

Most intriguingly, the single most enriched factor in our anti-Zld IPs, surpassing even RNAPII, was the kinase Tlk. Tlk is an evolutionarily conserved protein whose activity is tightly coupled to the cell cycle, peaking during S-phase to regulate chromatin assembly [54,67]. Given the rapid, synchronous nuclear divisions characteristic of early embryo development, Tlk is a prime candidate for coordinating Zld-mediated ZGA with the embryonic mitotic cycle.

We also detected several factors associated with chromatin insulators and the maintenance of Topologically Associating Domain (TAD) boundaries, including BEAF-32, CP190, Pzg, and Row. While the enrichment for these factors was moderate, their presence is supported by several

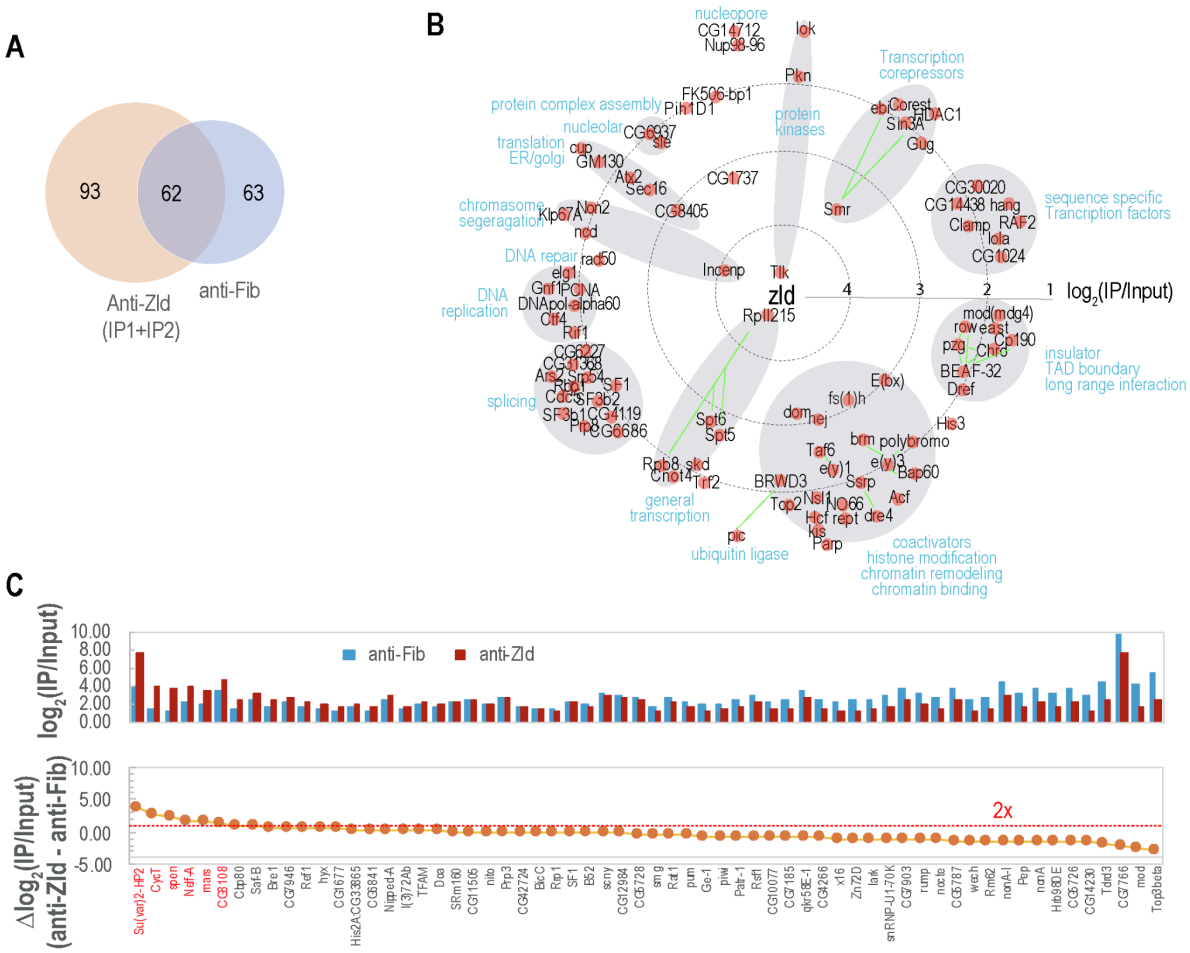

Fig. 2. Quantitative and functional analysis of the Zelda (Zld) interactome.

A) **Overlap of protein factors enriched in anti-Zld and anti-Fib IPs.** A total of 155 factors met the dual criteria of an FDR < 0.1 in at least one anti-Zld experiment and a minimum two-fold enrichment ($\log_2$(IP/Input)) ≥ 1) in the other. in the other. Of these, 62 factors were also enriched in the anti-Fib control, resulting in 93 Zld-specific interactors.

B) **Interaction map of Zld-specific factors.** Factors specific to the anti-Zld experiments are grouped by functional relatedness. The distance of each factor from the center is inversely proportional to its average enrichment in the two anti-Zld IP-MS datasets. Green lines indicate, in selected cases, subunits of the same complex or previously reported physical interactions.

C) **Relative enrichment of shared factors.** Differential enrichment for proteins identified in both anti-Zld and anti-Fib IPs. Data are plotted as the difference in $\log_2$(IP/Input) values between the two IPs. The red horizontal line indicates a two-fold enrichment difference ($\Delta\log_2$(IP/Input) = 1).

genomic observations. Zld is required for the *de novo* formation of TAD boundaries in the early embryo, and its binding motifs are significantly enriched at these sites [68,69]. Furthermore, the BEAF-32 motif was identified as a top-ranking enriched sequence within Zld ChIP-seq peaks in both larval neuroblasts and early embryos [70]. Collectively, our data provide a biochemical link between Zld and the spatial organization of higher-order chromatin structure during the MZT.

Other enriched factors further underscore Zld's role as a cooperative regulator. For example, the transcription factor Clamp, which cooperatively binds and activates enhancers with Zld

[70–72], was also identified. Additionally, the weak enrichment of DNA replication factors may be explained by the widespread stalling of replication forks at transcriptionally engaged, Zld-activated genes [73]. We also noted the enrichment of CG1737, a factor of unknown function, which was also previously identified in pull-downs for Chro [74] and HP1 [75]. However, whether this factor is functionally related to Zld requires further investigation.

In summary, our Zld interactome comprises a diverse repertoire of factors, suggesting that Zld does not merely "open" chromatin but acts as a central scaffold that bridges the maternal-to-zygotic transition by integrating transcriptional activation with the unique temporal demands of early embryogenesis.

Validation of the Zld interactome using endogenously tagged GFP-Zld

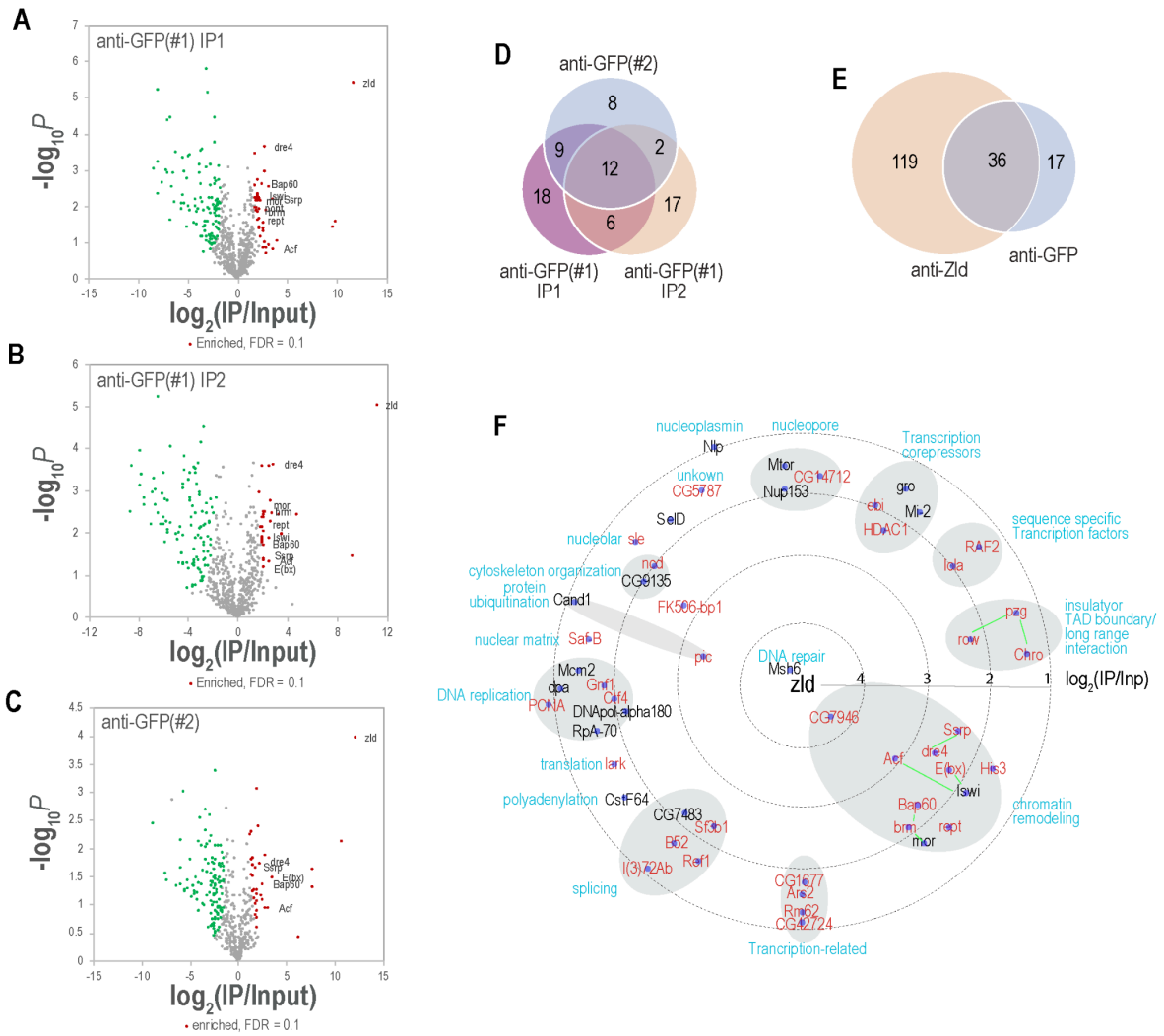

Fig. 3. IP–MS analysis using anti-GFP antibodies in embryos expressing endogenously tagged GFP-Zld Immunoprecipitation was performed using two distinct anti-GFP antibodies: anti-GFP(#1) and anti-GFP(#2). For anti-GFP(#1), experiments were conducted under two different washing conditions to assess interaction stability. Protein enrichment was analyzed as described in Figure 1.

A) **Enrichment of protein factors using anti-GFP (#1) under standard conditions.** Volcano plot showing $-\log_{10}P$ vs. $\log_2$(IP/Input). Factors enriched at an FDR < 0.1 are highlighted in red; select interactors associated with chromatin structure are labeled.

B) **Enrichment using anti-GFP (#1) under high-stringency conditions.** Similar to (A), but including a 2M urea wash step during the immunoprecipitation.

C) **Enrichment using anti-GFP (#2).** Similar to (A), but utilizing a different anti-GFP antibody.

D) **Overlap of GFP-Zld interactors across experiments.** Venn diagram showing the distribution of enriched factors across the three anti-GFP IP-MS experiments.

E) **Comparison of anti-Zld and anti-GFP interactomes.** The factors from the three anti-GFP experiments were consolidated and compared to the total list of factors enriched in the anti-Zld IP-MS experiments.

F) **Functional interaction map of GFP-Zld-associated factors.** Factors enriched in the anti-GFP IP-MS experiments are arranged by functional relatedness, with distance from the center inversely proportional to average enrichment. Factors that were also identified in the anti-Zld experiments are highlighted in red.

Table 2. Top 10 enriched GO terms for the proteins enriched in anti-GFP IP-MS

anti-GFP(#1) IP1

	No. factors	Enrichment	raw P-value
GO biological process complete			
ATP-dependent chromatin remodeling (GO:0043044)	7	14.7	1.38E-06
regulation of chromatin assembly or disassembly (GO:0001672)	4	13.5	4.11E-04
regulation of neuroblast proliferation (GO:1902692)	5	11.6	1.34E-04
nucleosome organization (GO:0034728)	8	11.0	1.41E-06
chromatin assembly or disassembly (GO:0006333)	5	8.8	4.07E-04
	7	6.0	2.13E-04
DNA repair (GO:0006281)	7	5.3	4.38E-04
negative regulation of transcription, DNA-templated (GO:0045892)	8	5.0	2.51E-04
GO molecular function complete			
nucleosome binding (GO:0031491)	5	18.3	2.19E-05
DNA helicase activity (GO:0003678)	4	14.7	3.19E-04
transcription corepressor activity (GO:0003714)	5	14.7	5.23E-05
DNA-binding transcription factor activity, RNA polymerase II-specific (GO:0000981)	4	11.7	6.47E-04
transcription factor binding (GO:0008134)	6	6.4	4.64E-04
double-stranded DNA binding (GO:0003690)	8	6.3	5.46E-05
ATP binding (GO:0005524)	13	2.7	6.89E-04
GO cellular component complete			
SWI/SNF complex (GO:0016514)	3	18.9	1.10E-03
histone deacetylase complex (GO:0000118)	5	16.9	2.99E-05
brahma complex (GO:0035060)	4	16.0	2.43E-04

anti-GFP(#1) IP2

	No. factors	Enrichment	raw P-value
GO biological process complete			
nuclear cell cycle DNA replication initiation (GO:1902315)	3	40.1	1.89E-04
regulation of chromatin assembly or disassembly (GO:0001672)	4	16.5	1.93E-04
protein-DNA complex subunit organization (GO:0071824)	7	8.1	3.24E-05
DNA conformation change (GO:0071103)	7	6.2	1.52E-04
regulation of transcription by RNA polymerase II (GO:0006357)	11	4.5	2.34E-05
negative regulation of cellular metabolic process (GO:0031324)	11	3.6	1.77E-04
GO molecular function complete			
sequence-specific double-stranded DNA binding (GO:1990837)	6	7.0	2.76E-04
GO cellular component complete			
protein-DNA complex (GO:0032993)	5	9.9	2.21E-04
nuclear chromatin (GO:0000790)	9	6.7	8.65E-06
polytene chromosome (GO:0005700)	7	4.7	7.81E-04

anti-GFP(#2)

	No. factors	Enrichment	raw P-value
GO biological process complete			
ecdysone receptor-mediated signaling pathway (GO:0035076)	3	47.9	1.11E-04
nucleosome mobilization (GO:0042766)	4	28.4	3.00E-05
regulation of chromatin assembly or disassembly (GO:0001672)	4	19.7	9.58E-05
regulation of neuroblast proliferation (GO:1902692)	5	16.8	2.15E-05
ATP-dependent chromatin remodeling (GO:0043044)	5	15.2	3.26E-05
positive regulation of transcription by RNA polymerase II (GO:0045944)	10	12.5	8.38E-09
negative regulation of neurogenesis (GO:0050768)	4	12.2	4.67E-04
chromatin assembly or disassembly (GO:0006333)	4	10.2	8.39E-04
regulation of defense response (GO:0031347)	5	8.9	3.16E-04
regulation of immune response (GO:0050776)	5	8.0	4.93E-04
regulation of response to biotic stimulus (GO:0002831)	5	7.3	7.36E-04
regulation of transcription, DNA-templated (GO:0006355)	19	5.4	7.00E-11
GO molecular function complete			
nucleosome binding (GO:0031491)	5	26.6	3.38E-06
chromatin DNA binding (GO:0031490)	4	13.5	3.34E-04
histone binding (GO:0042393)	5	10.0	1.92E-04
double-stranded DNA binding (GO:0003690)	6	6.8	2.71E-04
transcription regulator activity (GO:0140110)	7	6.3	1.25E-04
sequence-specific DNA binding (GO:0043565)	6	6.2	4.52E-04
GO cellular component complete			
FACT complex (GO:0035101)	2	63.9	1.35E-03
ACF complex (GO:0016590)	2	63.9	1.35E-03
nuclear euchromatin (GO:0005719)	3	27.4	3.70E-04
SWI/SNF complex (GO:0016514)	3	27.4	3.70E-04
RSC-type complex (GO:0016586)	3	21.3	6.65E-04
histone deacetylase complex (GO:0000118)	4	19.7	9.58E-05
transcription repressor complex (GO:0017053)	3	17.4	1.08E-03
brahma complex (GO:0035060)	3	17.4	1.08E-03
polytene chromosome (GO:0005700)	9	7.2	3.93E-06

To corroborate our findings, we performed IP-MS using a fly line expressing endogenously eGFP-tagged Zld [36]. Across three biological replicates, we identified 52 high-confidence interactors (Fig. 3), 32 of which overlapped with our primary anti-Zld IP-MS results. Consistent with our initial findings, the GFP-Zld IPs were significantly enriched for subunits of the NURF, FACT, and BAP/PBAP complexes, as well as TAD boundary factors including Row, Pzg, and Chro.

However, the total number of enriched factors was notably smaller than in the anti-Zld IPs. We determined that this reduced coverage was primarily due to the proteolytic truncation of the GFP-Zld protein in the embryo, as revealed by immunostaining analysis (Fig. 3_Supplement 1).

Quantitative image analysis indicated that these truncated forms localize to the nucleus alongside the full-length protein. This protein fragmentation was confirmed by western blot and MS peptide mapping, which showed a sharp decline in spectral counts C-terminal to amino acid 1023 (Fig. 3_Supplement 2). This intrinsic instability suggests that while endogenous tagging is a powerful tool, it can impact the efficiency of interactome discovery and must be accounted for in both biochemical and live-imaging colocalization analyses.

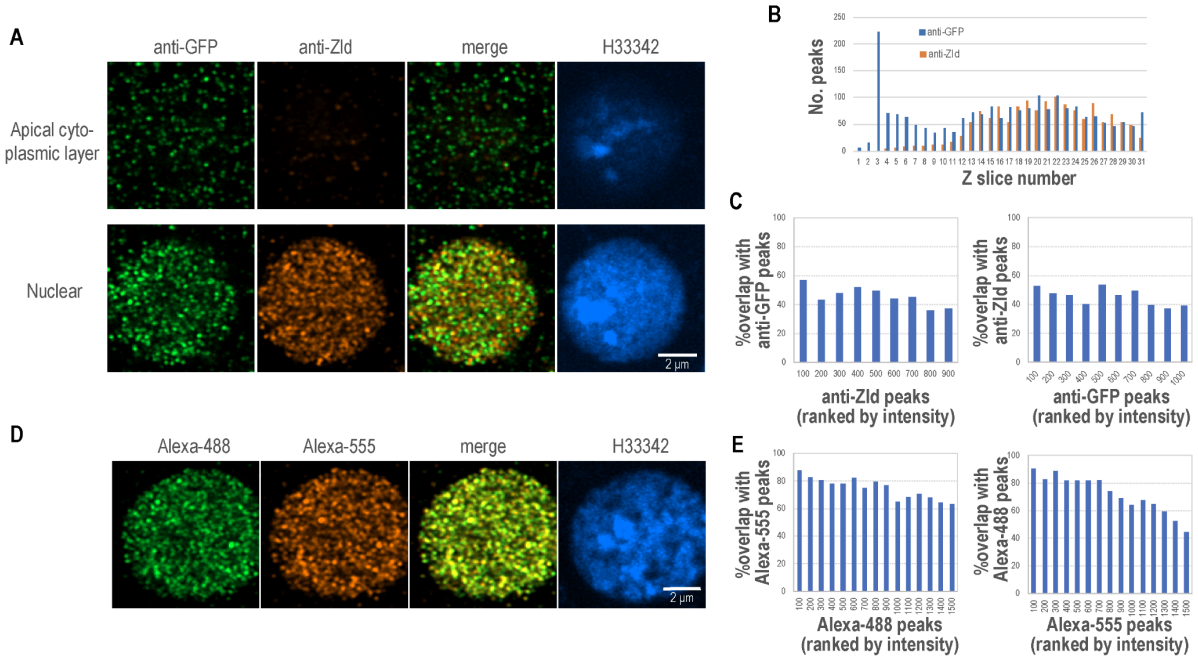

Fig. 3_Supplement 1. In vivo evidence of GFP-Zld fragmentation and spatial segregation.
Embryos were stained with the indicated antibodies and high-resolution images were acquired using the Zeiss Airyscan detection system. Images were processed and analyzed using Fiji software as detailed in the Methods.

A) **Spatial segregation of Zld fragments in the cytoplasm and nucleus.** Maximum intensity projections (MIP) of Z-slices (n = 2) from either the apical cytoplasmic layer or the nuclear center. In the cytoplasm, immunofluorescent signals were detected exclusively by the anti-GFP antibody. The absence of anti-Zld signal (targeting the C-terminal DNA binding domain) indicates these are isolated N-terminal fragments. In the nuclei, signals were detected by both antibodies.

B) **Z-profile of Zld fragments.** Distribution of fluorescent signals detected by anti-GFP and anti-Zld antibodies across the full Z-stack, illustrating the relative localization of fragments throughout the embryo depth.

C) **Quantitative colocalization analysis of nuclear Zld fragments.** Frequency of colocalization between foci detected by anti-GFP and anti-Zld antibodies. The distance from each spot in one channel to the nearest spot in the other was calculated; the chart shows the percentage of spots colocalizing within a 0.14 µm cutoff, binned by cohorts of 100 peaks ranked by intensity. The relatively low overlap demonstrates that the Zld fragments, not just full length Zld, are present in the embryo nuclei.

D) **Technical control for colocalization efficiency.** Wild-type embryos were incubated with a single primary anti-Zld antibody followed by two secondary antibodies conjugated to different fluorophores. MIP images for both secondary antibody channels are shown.

E) **Validation of colocalization methodology.** Quantification of the overlap between the two secondary antibody channels from (D), using the same proximity-based analysis described in (C). The high overlap serves as a baseline for maximum achievable colocalization.

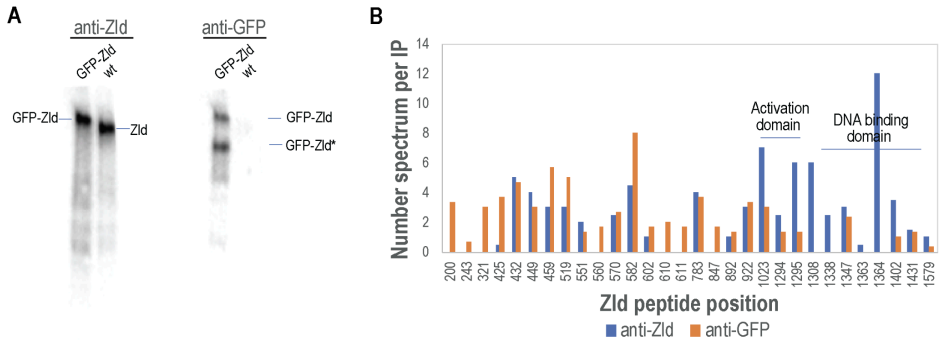

Fig. 3_Supplement 2. Evidence of GFP-Zld fragmentation in embryonic lysates.

A) **Detection of full-length and truncated GFP-Zld by Western blot.** Full-length GFP-Zld is detected by the anti-Zld antibody (targeting the C-terminus), whereas both the full-length protein and a truncated N-terminal fragment (GFP-Zld*) are detected by the anti-GFP antibody.

B) **Peptide distribution across the Zld protein sequence.** Distribution of peptide spectral counts across the Zld primary sequence as identified in anti-Zld and anti-GFP IP-MS experiments. The sharp decrease in spectral counts in the anti-GFP samples corresponds to the C-terminal region. The putative activation domain and the DNA binding domain (zinc-finger cluster) [48] are indicated for reference.

Genomic profiling confirms Zld co-localization with key factors

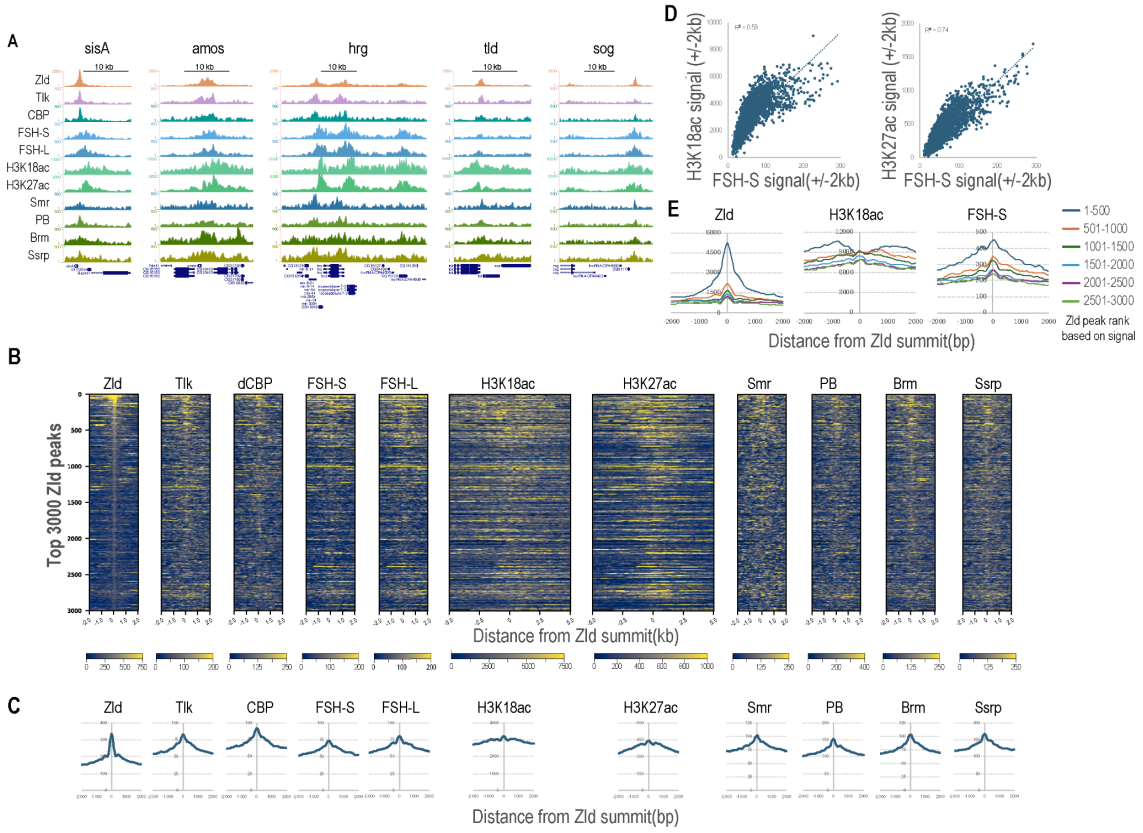

Figure 4. Genomic validation of Zld-associated factors via CUT&RUN

A) **Representative genomic tracks of Zld-associated factors.** CUT&RUN signal profiles for the indicated antibodies around the *sisA*, *amos*, *hrg*, *tld*, and *sog* loci. The experiments were performed in formaldehyde-fixed cycle 12 embryos.

B) **Global occupancy of Zld interactors at Zld-bound sites.** Heatmaps displaying the CUT&RUN signal for the indicated factors centered on the top 3,000 Zld peaks. Peaks are sorted by fold enrichment as determined by MACS..

C) **Average occupancy profiles at Zld peaks.** Composite signal profiles for each antibody centered on all identified Zld peaks.

D) **Correlation of Fsh occupancy with histone acetylation.** Correlation of Fsh-S signal with H3K18ac and H3K27ac enrichment around Zld-bound regions. The result for Fsh-L is very similar (not shown). The experiments were performed in formaldehyde-fixed cycle 13 embryos.

E) **Fsh recruitment correlates with Zld peak intensity despite local histone depletion.** **Fsh recruitment correlates with Zld peak intensity despite local histone depletion.** Zld peaks were ranked by fold enrichment and divided into cohorts of 500. Average signal profiles for each cohort demonstrate that at high-intensity Zld peaks, Fsh-S signal remains robust directly over the peak center, even where H3K18ac levels are significantly lower than in flanking sequences. Results for Fsh-L were highly similar (data not shown). The experiments were performed in formaldehyde-fixed cycle 13 embryos.

To confirm these biochemical associations *in vivo*, we performed CUT&RUN for a subset of the most highly enriched factors. We found that dCBP, Fsh, Bap, Brm, Ssrp, Smr, and Tlk all co-localized extensively at Zld binding sites (Fig. 4). Consistent with the molecular signature of active enhancers, we observed robust H3K18ac and H3K27ac signals flanking these Zld-occupied regions. Notably, Fsh occupancy generally correlated with broad domains of histone acetylation across the genome. However, a specific discrepancy emerged at the centers of strong Zld peaks: despite the local depletion of acetylated histones directly over these sites - likely due to nucleosome displacement - focal Fsh signals remained highly concentrated, often exceeding the intensity of flanking regions (Fig. 4D, E). This pattern became significantly more pronounced in cycle 13 than in cycle 12. These observations suggest that while Fsh may associate with the genome via histone acetylation globally, its recruitment to these specific centers is likely mediated by a distinct mechanism - direct physical interaction with Zld or certain other factors associated with the enhancers..

Recruitment of RNAPII to distal Zld binding sites

To investigate the striking RNAPII enrichment observed in our IP-MS, we performed RNAPII CUT&RUN across early embryonic stages. Surprisingly, we found that the RNAPII signal was not restricted to promoters; instead, it was broadly observed across nearly all Zld-bound regions, including distal intergenic peaks (Fig. 5A–C). Notably, when comparing RNAPII enrichment at these intergenic Zld peaks to their nearest potential target promoters, the signal at the distal Zld sites was often significantly more robust (Fig. 5D). This pattern was consistent at active enhancers for genes such as *Kr* and *sog*, as well as at distal enhancers for genes like *soxN*, which are Zld-bound but not yet actively transcribed. This widespread distal localization contrasts with previous ChIP-chip [76] and ChIP-seq [77] studies, which primarily localized RNAPII to transcription start sites(TSS).

RNAPII is not alone in its association with these distal regions. As shown in Fig. 5, the RNAPII-associated factor NELF-A, which was also identified in our IP-MS, as well as the

Mediator subunit Med14 and the general transcription factor (GTF) TBP, all co-localized with Zld at distal enhancers. Collectively, these CUT&RUN results suggest that Zld may serve as a scaffold to recruit RNAPII, the NELF complex, the Mediator complex, and, to a lesser extent, TBP, directly to its binding sites. This distal assembly of the transcriptional machinery likely facilitates the rapid transactivation of Zld target genes by pre-configuring the transcriptional hub prior to or concurrent with promoter communication.

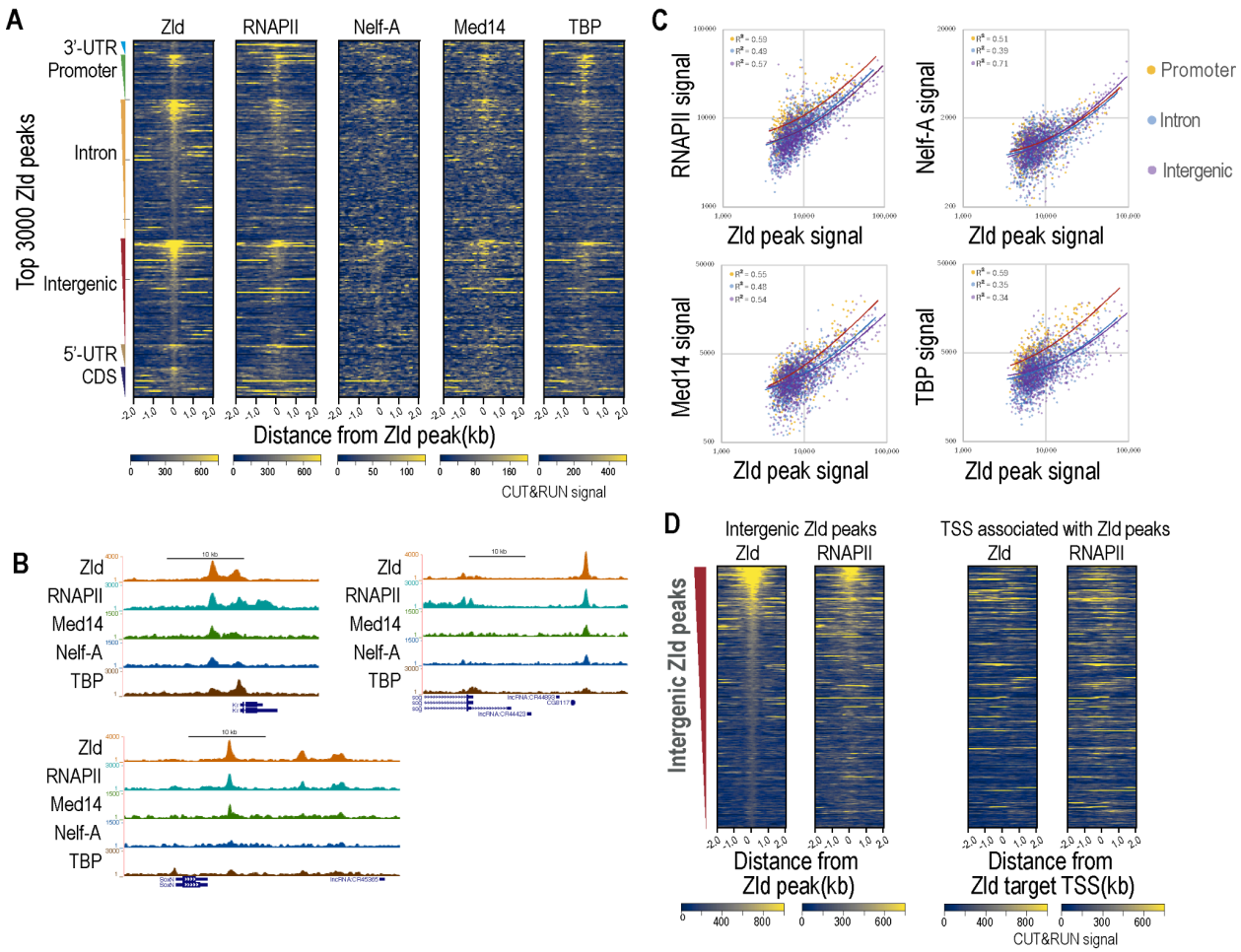

Figure 5. Zld recruits RNAPII, NELF-A, Med14, and TBP to distal and proximal binding sites.

A) **Global occupancy of transcription machinery across genomic features.** Heatmaps displaying CUT&RUN signals for the indicated factors. Zld peaks are categorized by genomic location (3' UTR, intron, intergenic, CDS, and 5' UTR) and sorted within each category by Zld fold enrichment.

B) **Representative genomic tracks.** Signal profiles for Zld, RNAPII, NELF-A, Med14, and TBP around selected gene loci, demonstrating co-localization at both promoters and distal enhancers.

C) **Correlation of factor occupancy with Zld binding.** Scatter plots and correlation coefficients showing the relationship between Zld signal and each identified factor within a ±250 bp window around Zld peaks. While all factors correlate with Zld, TBP displays a relatively higher signal bias toward promoters.

D) **Comparison of RNAPII occupancy at intergenic enhancers versus target promoters.** Heatmaps centered on intergenic Zld peaks (left) and their corresponding nearest transcription start sites (TSS; right). RNAPII signals are generally more robust at the intergenic Zld-bound enhancers than at the potential target promoters.

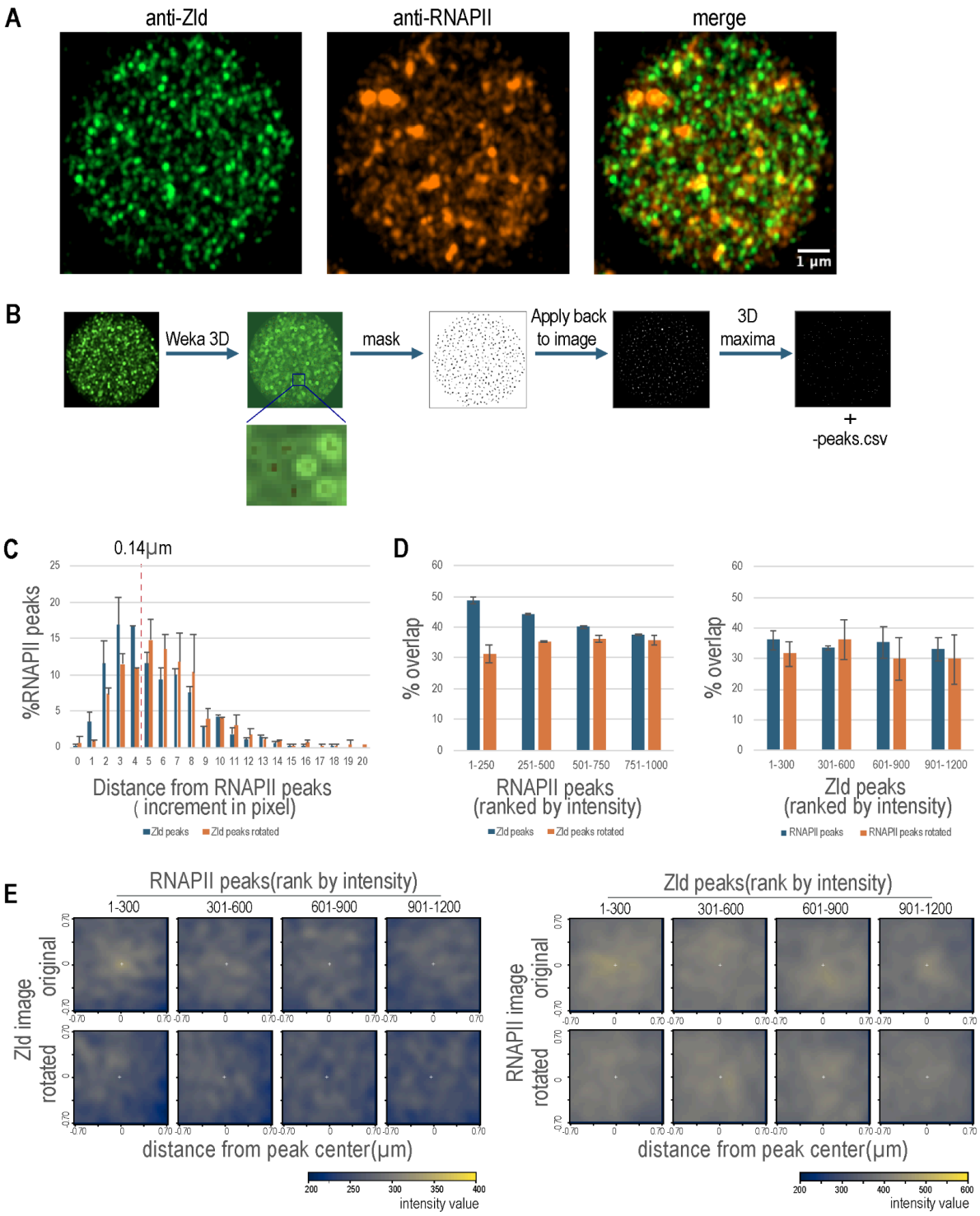

Fig. 6. Colocalization of Zld and RNAPII in embryo.

The embryos were stained with anti-Zld antibody and an anti-RNAPII (4H8) antibody, and high resolution images were acquired using the Zeiss airyscan detection system.

A) **Visualizing sub-nuclear clusters.** Maximum intensity projection (MIP) images of 5 Z-slices from a mid-cycle 12 embryo are shown.

B) **Image analysis workflow**. Schematic illustrating the peak identification and cluster segmentation process performed in Fiji.
C) **Object-based 3D distance analysis.** The 3D distance from each RNAPII peak to the nearest Zld peak was calculated. To establish a noise floor, distances were recalculated after rotating the Zld coordinates by 90° (rotated control). The distribution shows a significantly higher percentage of RNAPII peaks within 4 pixels (0.14 µm) of a Zld peak compared to the rotated control.
D) **Enrichment of brightest RNAPII hubs.** Charts showing the percentage of peak overlap (using a 0.14 µm cutoff) for peak cohorts ranked by intensity. The most significant colocalization is observed within the top 250 brightest RNAPII spots.
E) **Signal intensity heatmaps.** Average RNAPII signal centered around Zld peak cohorts and Zld signal centered around RNAPII peak cohorts. For comparison, heatmaps generated after rotating target images by 90° are shown. There was a significantly higher Zld signal centered around the top 250 RNAPII peaks.

High-resolution imaging for Zld-RNAPII colocalization

To visualize the spatial relationship between Zld and the transcriptional machinery, we performed immunostaining in early embryos using the Zeiss Airyscan detector system for super-resolution imaging. We observed that Zld and RNAPII form distinct sub-nuclear clusters (Fig. 6A), consistent with the specialized transcription "hubs" previously described [34–36,78,79].

To rigorously quantify this association and account for potential chance overlaps in the crowded nuclear environment, we employed object-based 3D distance analysis compared against a rotated control. In this control, one channel was rotated 90° relative to the other to randomize the spatial relationship while maintaining the original cluster density and morphology. We found that a significantly higher proportion of the top 250 brightest RNAPII spots were located in close proximity (within 0.14 µm) to a Zld cluster compared to the randomized controls (Fig. 6C, D).

As a secondary validation, we used a conventional heatmap analysis to examine the distribution of fluorescent signals around peak centers (Fig. 6E). Consistent with our object-based findings, we observed a significant enrichment of the Zld signal centered around the top 250 RNAPII peaks, compared to rotated controls. Conversely, RNAPII signal was modestly enriched around the brightest Zld peaks.

Taken together - and notwithstanding the high level of stochastic overlap observed in the randomized controls - this dual-methodological approach demonstrates that the core transcriptional machinery and Zld are spatially coordinated within the nucleus.

Zld clusters preferentially associate with target gene transcription loci

We next investigated whether these Zld hubs are specifically localized to Zld-target genes. Using smiFISH to detect nascent RNA transcripts, we measured the 3D distance between Zld protein clusters and active transcription foci, comparing Zld-target genes directly to non-target controls.

We found that transcription foci for the endogenous Zld-target gene *zen* were located significantly closer to Zld clusters than those of the non-target gene *dfd*. This Zld-dependent spatial coordination was further confirmed using a *lacZ* reporter driven by the *hb* stripe enhancer

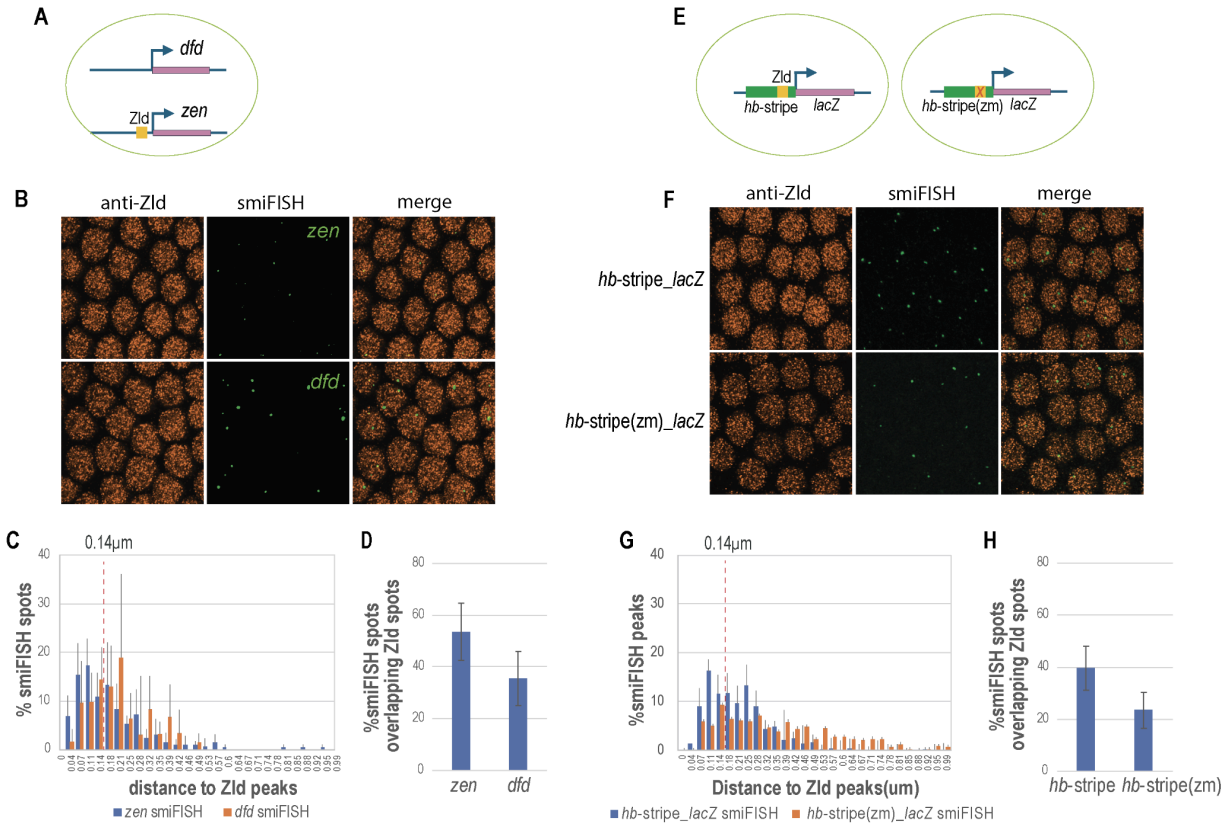

Fig. 7. Co-localization of Zld clusters with active transcription foci of Zld target genes.

High-resolution smiFISH and immunostaining using an anti-Zld antibody were performed in fixed embryos. The smiFISH method detects both individual mRNA molecules and high-intensity clusters (transcription foci) representing sites of nascent transcription. Zld peaks and transcription foci were identified using the 3D Maxima plugin in Fiji. Nearest-neighbor distances between transcription foci and their closest Zld peak were calculated as described in Figure 6.

A) **Experimental design for endogenous target validation.** Schematic of *zen* (a Zld target) and *dfd* (a non-target). Probes for both genes were used simultaneously in the same embryos, allowing for intra-nuclear comparison.

B) **Visualization of transcription foci and Zld hubs.** Representative regions (25 µm × 25 µm) are shown from original 100 µm × 100 µm maximum intensity projections (MIP) of an early c14 embryo. Images were thresholded to specifically highlight high-intensity smiFISH signals at active transcription sites.

C) **Spatial distribution of transcription foci relative to Zld.** Frequency distribution of the distances between *zen* or *dfd* smiFISH spots and the nearest Zld cluster.

D) **Quantitative co-localization of endogenous foci.** Percentage of *zen* and *dfd* transcription foci co-localizing with Zld clusters within a 0.14 µm center-to-center cutoff.

E) **Reporter constructs for Zld-dependency testing.** Schematics of the *lacZ* reporter transgenes driven by either the wild-type *hb* stripe enhancer (*hb*-stripe) or a version with mutated Zld binding sites (*hb-stripe(zm)*).

F) **Visualization of reporter transcription.** MIP images of 25 µm × 25µm showing *lacZ* smiFISH signals in cycle 14 embryos for both the wild-type and Zld-binding site mutant transgenes.

G) **Distance distribution of reporter foci.** Distribution of distances between *lacZ* transcription foci and the closest Zld clusters for both transgenes.

H) **Distance distribution of reporter foci.** Distribution of distances between *lacZ* transcription foci and the closest Zld clusters for both transgenes.

[8] - while the wild-type reporter showed close proximity to Zld clusters, the mutated *lacZ* reporter lacking Zld-binding sites was located significantly further away (Fig. 7). These results demonstrate that Zld clusters are functional transcriptional hubs recruited to specific genomic loci through direct Zld-DNA interactions, and that the loss of these binding sites results in the physical dissociation of the gene from the hub.

dCBP mediates Zld-dependent histone H3K18 acetylation and Fsh cluster formation

In our IP-MS analysis, Fsh and dCBP were among the most strongly enriched Zld-associated factors. Since Zld binding is known to correlate with H3K18ac [26], a mark deposited by dCBP[50], and because Fsh/BRD4 orthologs recognize acetylated histones to facilitate enhancer function [51,80–83], we hypothesized that dCBP and Fsh represent early-acting factors in Zld-dependent activation.

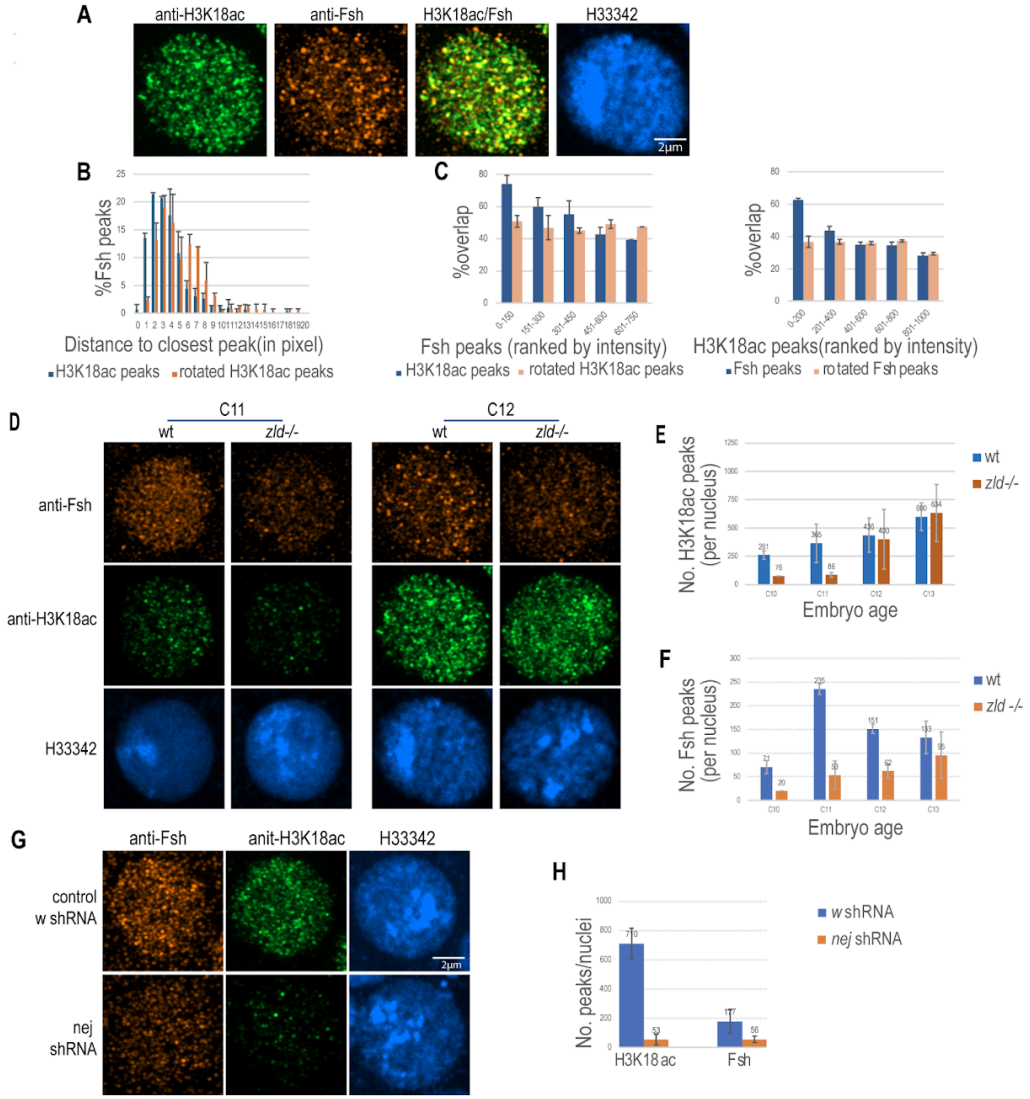

Fig. 8. dCBP and Zld are required for H3K18ac and the formation of Fsh clusters.
Immunostaining was performed in fixed embryos using anti-H3K18ac and anti-Fsh antibodies. Images were acquired and analyzed as described in Figure 6.

A) **Representative sub-nuclear distribution.** Maximum intensity projection (MIP) of a single nucleus from a cycle 12 embryo showing H3K18ac and Fsh clusters.
B) **Spatial correlation between Fsh and H3K18ac.** Distribution of nearest-neighbor distances from the top 150 Fsh peaks to the closest H3K18ac peaks. The distribution obtained after rotating H3K18ac peak coordinates by 90° is shown as a control.
C) **Intensity-dependent colocalization.** Percentage of Fsh peaks in each intensity-ranked cohort colocalizing with H3K18ac peaks (left), and vice versa (right), using a 0.14 μm center-to-center cutoff.
D) **Zld-dependency of Fsh and H3K18ac.** MIP images from wild-type and *zld*–/– germline clone mutant embryos. The wild-type cycle 12 image is the same as shown in (A).
E) **Temporal dynamics of H3K18ac loss in *zld* mutants.** Quantitative change in the number of H3K18ac clusters per nucleus in *zld* mutants compared to wild-type from cycles 10 to 13.
F) **Temporal dynamics of Fsh cluster loss in *zld* mutants.** Quantitative change in the number of Fsh clusters per nucleus in *zld* mutants compared to wild-type from cycles 10 to 13.
G) **Loss of clusters upon dCBP knockdown.** MIP images of embryos expressing shRNA targeting *nej* (encoding dCBP) or *w* (negative control).
H) **Quantification of dCBP-dependency.** Comparison of the number of Fsh and H3K18ac clusters per nucleus in *nej* versus *w* RNAi embryos at cycle 13.

BRD4 has been shown to form nuclear puncta with liquid-like condensate properties at super-enhancers in various mammalian cell types, including in mouse embryonic stem cells [84–86]. To characterize the relationship between dCBP, H3K18ac, Fsh, and Zld, we performed immunostaining analysis in wild-type, *nej* (*dCBP*) RNAi-knockdown, and *zld* mutant embryos.

High-resolution Airyscan imaging in wild-type embryos revealed a vast number of distinct H3K18ac, and Fsh clusters reminiscent of the BRD4 condensates observed in mammalian cells (Fig. 8A). Given the specificity of the anti-H3K18ac signal for chromatin, these clusters likely represent individual or small groups of co-localized enhancers. Analysis across early mitotic cycles showed that weak H3K18ac clusters are detectable as early as cycle 8 and increase in both number and intensity through subsequent cycles (Fig. 8_Supplement 1). Because H3K18ac is stable during mitosis [78], this cumulative deposition suggests that chromatin opening is a progressive process, likely reflecting the transition of zygotic enhancers from a "primed" to a "fully active" state.

Fsh exhibited similar dynamics, with signal near background at cycle 8 before forming increasingly bright clusters. This increase likely results from both rising nuclear Fsh concentrations and increased H3K18ac density around enhancers. Given that the Fsh ortholog Brd4 is a limiting factor for ZGA in zebrafish [87], the low initial levels of Fsh observed here suggest it may similarly limit the pace of ZGA in *Drosophila*.

Consistent with Fsh recruitment by acetylated histones, object-based colocalization analysis revealed that Fsh and H3K18ac clusters show significant spatial overlap, with high-intensity peaks exhibiting the strongest colocalization (Fig. 8B–C). This indicates a more stable association of Fsh with extensively acetylated, high-occupancy enhancers.

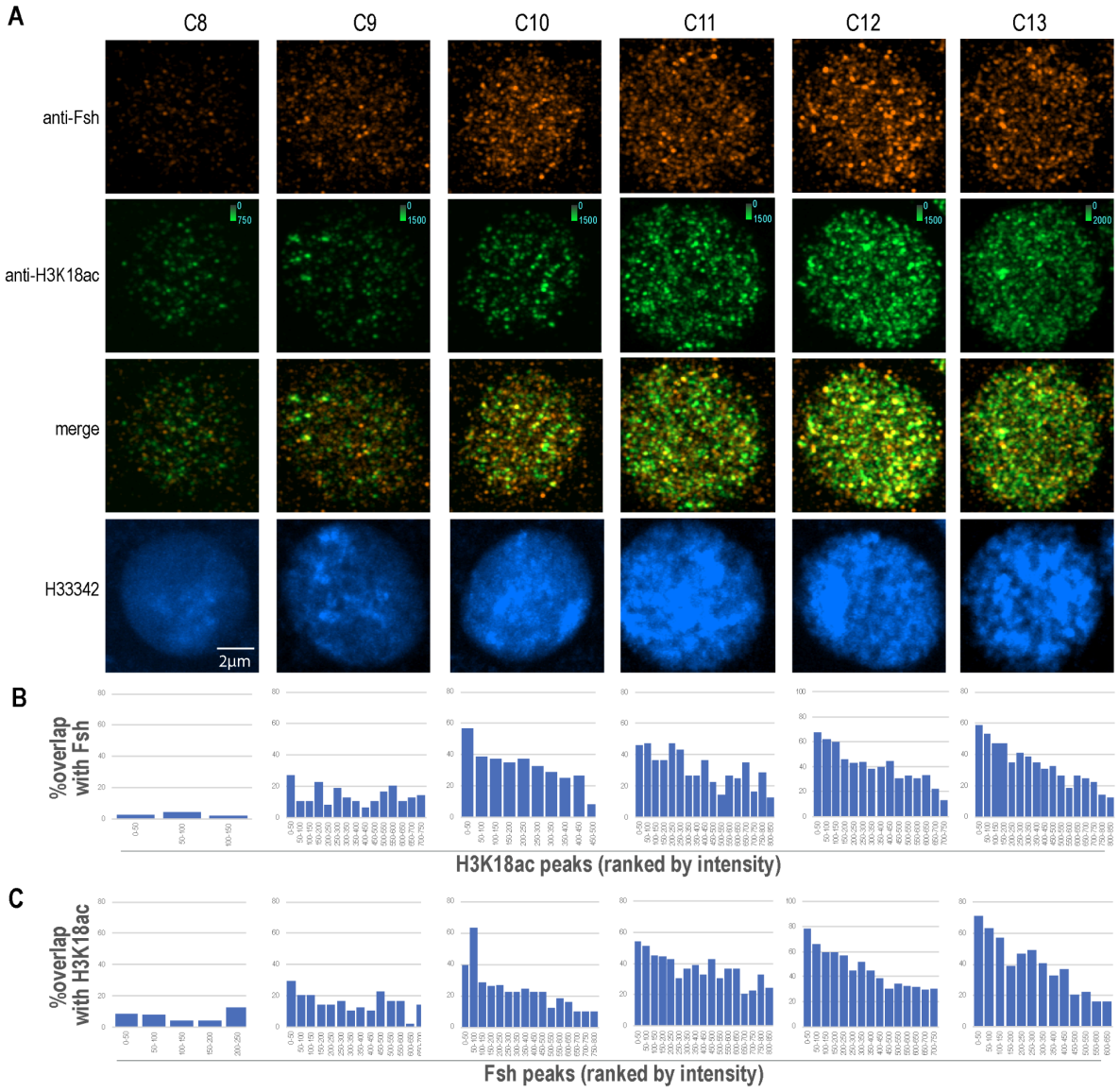

Fig. 8_Figure Supplement 1. Dynamics of H3K18ac and Fsh clusters in early mitotic cycles. Similar to the main figure Fig.8.

A) **Developmental trajectory of H3K18ac and Fsh clusters.** Representative maximum intensity projection (MIP) images of individual nuclei from embryos at the indicated mitotic cycles (C8–C14). Embryos were co-immunostained with anti-H3K18ac and anti-Fsh antibodies, illustrating the progressive increase in cluster number and intensity over developmental time.

B) **Intensity-dependent colocalization of Fsh with H3K18ac.** Quantitative analysis showing the percentage of Fsh peaks that colocalize with H3K18ac peaks (within a 0.14 µm center-to-center cutoff). Fsh peaks were ranked by intensity and divided into cohorts of 50. The data demonstrate that stronger Fsh clusters are more likely to be associated with H3K18ac-enriched regions.

C) **Intensity-dependent colocalization of H3K18ac with Fsh.** Quantitative analysis showing the percentage of H3K18ac peaks that colocalize with Fsh clusters (within a 0.14 µm cutoff). H3K18ac peaks were ranked by intensity and divided into cohorts of 50. Similar to (B), higher levels of histone acetylation correlate with a higher frequency of Fsh recruitment.

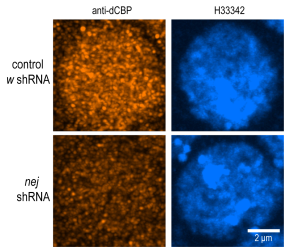

Fig. 8_Supplement 2. RNAi knockdown of dCBP. Immunostaining was performed with an anti-CBP antibody in embryos that expressed shRNA for the *nej* gene or the *w* gene (negative control). Maximum intensity projection images of individual nuclei are shown.

In *zld* mutant embryos (germline clones), both H3K18ac and Fsh signals were significantly reduced (Fig. 8D-F). The loss of H3K18ac was most pronounced during cycles 10–11, with a partial recovery by cycle 12. However, the brightest Fsh clusters remained absent, suggesting that while other factors may eventually contribute to acetylation, the assembly of large, Fsh-rich clusters is highly Zld-dependent. Furthermore, RNAi knockdown of *nej* (dCBP) dramatically decreased both H3K18ac and Fsh clusters (Fig. 8G-H, Fig. 8_supplement 2), confirming the dependency of Fsh recruitment on dCBP-mediated acetylation.

Fsh and dCBP are critical for the transactivation of Zld-dependent genes

To directly test the requirement of Fsh and dCBP in Zld-mediated transactivation, we analyzed the effects of RNAi knockdown on the transcription of Zld-dependent genes (*sala, tld,* and *hb*) and a Zld-independent control (*btd*) [23]. We performed smiFISH in embryos expressing shRNA targeting either *fs(1)h* or *nej*, using *white (w)* shRNA as a negative control. In our smiFISH analysis, individual mRNA molecules were detected as distinct puncta, while active transcription sites appeared as one or two high-intensity foci per nucleus. For *hb*, which has significant maternal RNA deposition, we quantified only the active transcription foci. For all other genes, we quantified the total number of both spot types across the Z-stack using the Fiji 3D Maxima plugin.

In addition to smiFISH, we performed immunostaining with an anti-RNAPII antibody to monitor transcriptional activity at the histone gene cluster, visualized as the Histone Locus Body (HLB). As shown in Fig. 9B-C, *fs(1)h* knockdown (Fig. 9_Supplement 1) significantly reduced the transcription of *sala, tld,* and *hb*, but not that of the Zld-independent gene *btd*. Furthermore, the RNAPII signal at the HLB was only modestly reduced in *fs(1)h* RNAi embryos compared to controls. These results indicate that Fsh is specifically required for the transcription of Zld-dependent genes.

The effects of dCBP (*nej*) knockdown mirrored those of Fsh but were markedly more severe (Fig. 9D-E). Depletion of dCBP nearly abolished the transcription of *sala, hb,* and *tld*. While we observed a slight reduction in *btd* transcription and RNAPII signal at the HLB, these effects were minor compared to the near-total loss of Zld-target expression. Collectively, these data demonstrate that while dCBP has a broader impact on the embryonic transcriptome, both dCBP and Fsh are essential for the robust transactivation of Zelda's primary targets.

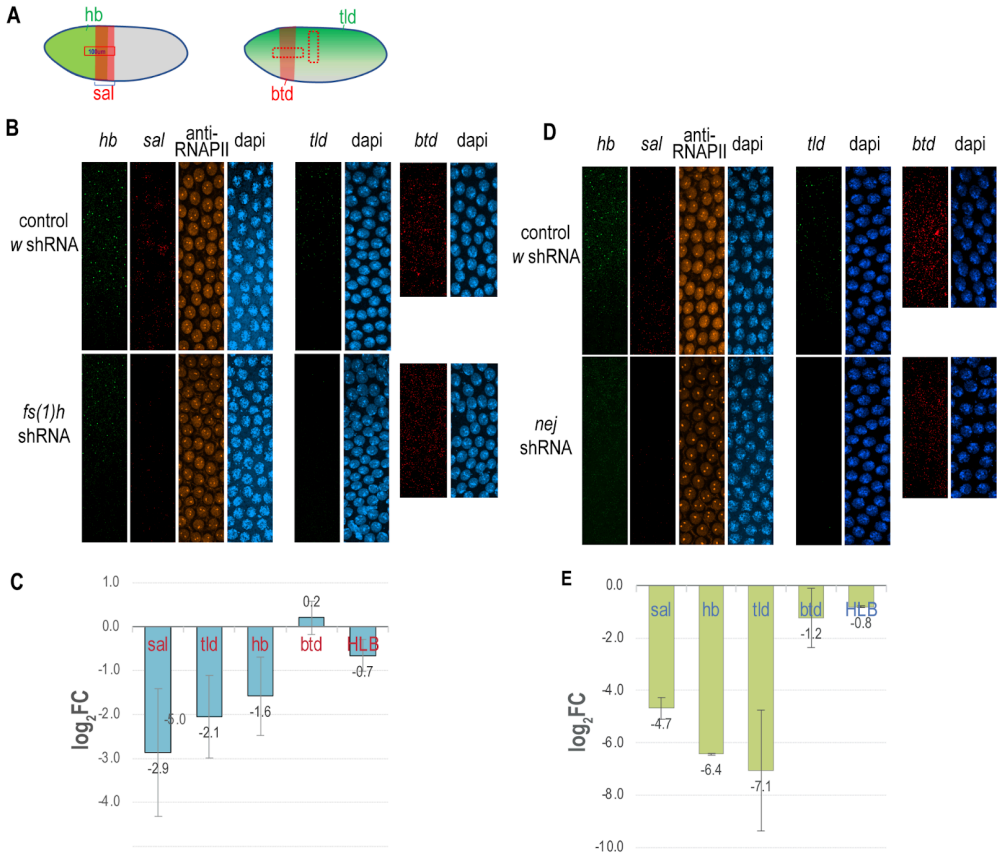

Fig. 9. Requirement of Fsh and dCBP for the transactivation of Zld-dependent genes.
smiFISH for *sala*, *tld*, *hb*, and *btd* was performed in embryos expressing shRNA targeting *fs(1)h* or *nej* (*dCBP*), with *white* (*w*) shRNA serving as a negative control. Following smiFISH hybridization, embryos were immunostained with an anti-RNAPII antibody. The number of individual mRNA molecules for *sala*, *tld*, and *btd* was quantified within the indicated embryonic regions using the Fiji 3D Maxima plugin. For *hb*, the number and intensity of bright nuclear foci (representing active transcription) were determined, and the integrated intensity (sum of focus intensities) was calculated for each scanned area. Additionally, the average intensity of RNAPII signal at the Histone Locus Body (HLB) was measured per embryo.

A) **Spatial domains of analyzed genes.** Schematics illustrating the expression patterns of the target genes and the specific embryonic regions scanned for quantification.

B and D) **Visualizing transcriptional output.** Representative maximum intensity projections (MIP) generated from Z-stacks spanning the full depth of the cortical nuclear layer for the indicated genotypes.

C and E) **Quantitative impact of Fsh and dCBP depletion.** Charts displaying the average fold change (FC) in smiFISH or HLB-RNAPII signals in *fs(1)h* or *nej* RNAi embryos relative to *w* RNAi controls. Fold changes were first normalized between embryos at matched developmental stages (cycles 12 and 13, determined by nuclear diameter) prior to calculating the mean FC for each gene.

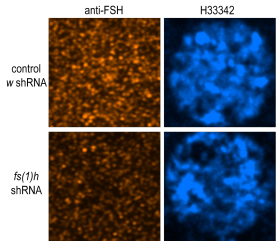

Fig. 9_Supplement 1. RNAi knockdown of Fsh. Maximum intensity projection images from immunostaining with anti-Fsh antibody in an embryo that expressed shRNA for the *fs(1)h* gene or the *w* gene (control).

Smr modulates the transcriptional output of Zld-target genes

We next investigated the corepressor Smr. RNA-seq in *smr* RNAi embryos revealed broad up-regulation of genes associated with the strongest Zld peaks (Fig. 10A, B). To specifically investigate the functional relationship between Smr and Zld, we associated Zld peaks with their

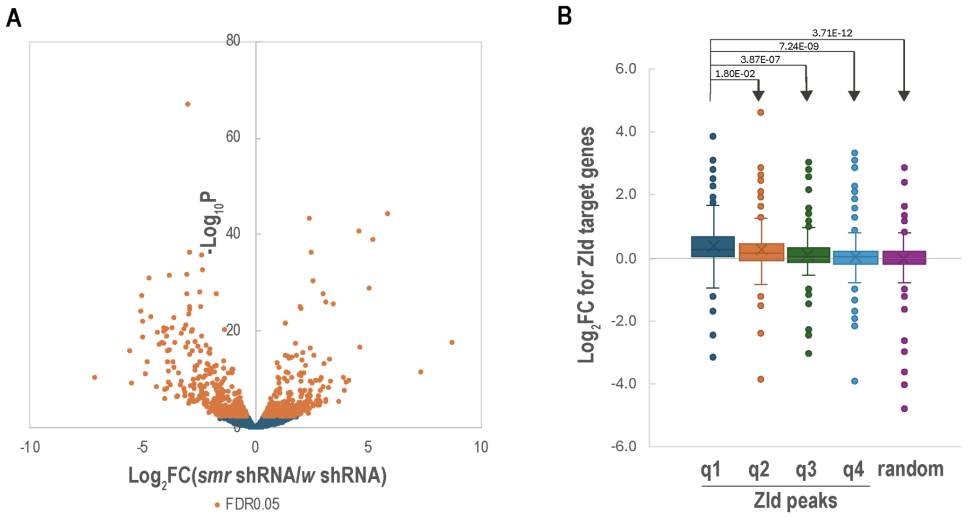

Fig. 10. Depletion of Smr leads to the derepression of genes associated with strong Zld binding.

A) **Global transcriptomic changes upon *smr* knockdown.** Volcano plot displaying differential gene expression in *smr* RNAi embryos (cycle 12-13) compared to *w* RNAi controls. Data were generated from RNA-seq and analyzed using the edgeR exactTest. Significantly up- and down-regulated genes are highlighted.

B) **Correlation between Zld occupancy and Smr-mediated repression.** The top 3,000 Zld peaks were associated with their nearest neighbor genes. These peak-gene pairs were consolidated by assigning only the strongest Zld peak to each unique gene, resulting in a high-confidence dataset of 1,160 pairs. Genes were ranked by the intensity of their associated Zld peak and divided into four quantiles ($q1$–$q4$, with $q1$ representing the strongest Zld binding). Box plots show the distribution of fold changes in transcript levels following *smr* knockdown for each quantile, compared to a randomly selected set of control genes. *P*-values from a Student's t-test are indicated above each quantile, demonstrating that genes with the strongest Zld occupancy exhibit the highest degree of derepression upon Smr depletion.

nearest genes within a 15 kb window. We observed a positive correlation between Zld occupancy and Smr-mediated repression: genes associated with the strongest Zld peaks were significantly more likely to be up-regulated upon *smr* knockdown (Fig. 10B). This suggests that Smr plays a critical role in modulating Zld-dependent transactivation, potentially preventing premature or ectopic expression. This matches the presence of a conserved repression domain within the Zld protein [49].

Tlk interacts directly with Zld *in vitro*

Our IP-MS and CUT&RUN analyses indicated that the kinase Tlk is physically associated with Zld at genomic target sites. To determine if this interaction is direct, we performed *in vitro* GST

pull-down assays. We partitioned Zld into four segments and expressed them as Glutathione S-Transferase (GST) fusion proteins. Immobilized fusion proteins were incubated with *in vitro* translated Tlk, and binding was assessed via Western blot.

As shown in Figure 11, untreated Tlk failed to bind any Zld segments. However, upon treatment with Calf Intestinal Phosphatase (CIP) to remove autophosphorylation or inhibitory phosphates, we observed significant and selective binding of Tlk to the Zld segment encompassing residues 401–750. A kinase-dead version of Tlk (Tlk^{dead}) similarly exhibited robust binding to the same segment. These findings mirror the biochemical behavior of the nucleosome assembly factor Asf1, a known Tlk substrate [88,89], suggesting that Tlk's kinase activity or its own phosphorylation state can autoinhibit its association with substrates. Collectively, these results suggest that Tlk directly interacts with Zld in a phosphorylation-dependent manner, providing a potential biochemical link between the cell-cycle machinery and the regulation of Zld activity during the rapid mitotic cycles of early ZGA.

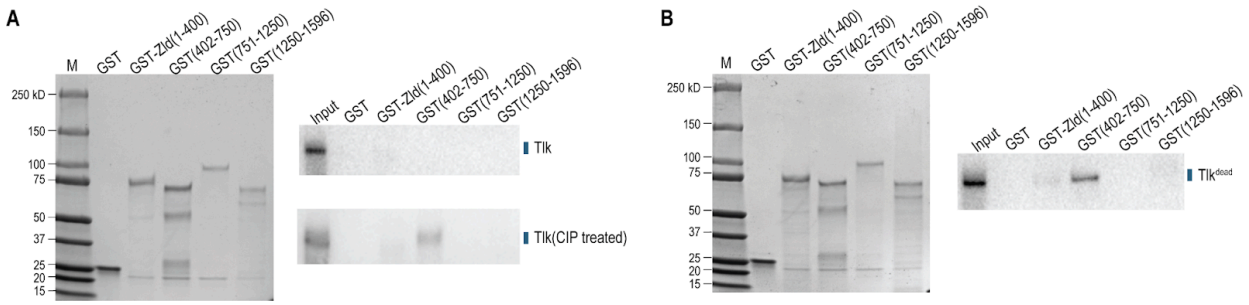

Fig. 11. Direct interaction of Tlk with a Zld domain in vitro.
GST-Zld fusion proteins containing indicated segments of Zld were immobilized on glutathione magnetic beads and incubated with *in vitro* translated Tlk protein, either wild-type (WT) or harboring a kinase-domain mutation (Tlk^{dead}). Bound Tlk was analyzed by Western blot.

A) Tlk specifically interacts with Zld(401-750). (Left) Coomassie-blue stained SDS-PAGE gel showing 20% of the immobilized GST or GST-Zld fusion proteins used as bait in the binding reactions. (Upper right) Western blot analysis of binding reactions with untreated WT Tlk. (Lower right) Western blot analysis of binding reactions with WT Tlk pre-treated with Calf Intestinal Phosphatase (CIP). The Input lane represents 20% of the total TNT product used per reaction.

B) Mutation in the Tlk kinase domain enhances binding to Zld(401-750). Binding reactions performed as in (A), using the Tlk kinase-dead mutant as the prey protein. The Input lane represents 20% of the total TNT product used per reaction.

DISCUSSION

In this study, we utilized an optimized IP-MS pipeline to define the diverse protein interactome of Zld. We identified multiple groups of functionally distinct factors, primarily those involved in chromatin remodeling and transcriptional control. Our functional validation of dCBP and Fsh extends recent reports of their roles in the MZT [78,90,91]. Furthermore, our identification of Smr provides a likely mechanistic explanation for Zld's conserved repressive domain [49].

Finally, the discovery of a direct, phosphorylation-sensitive interaction between Zld and Tlk suggests a novel regulatory link between DNA replication and ZGA.

Zld recruits specialized machinery to drive chromatin accessibility

Zld is uniquely required for nucleosome depletion and chromatin accessibility at zygotic enhancers [7,26,27,30,92,93], exerting a more profound effect on accessibility than other embryo patterning factors [30,92]. While the classical definition of a pioneer factor centers on the ability to bind motifs embedded in "closed" chromatin, recent evidence suggests this distinction is not absolute. Non-pioneer TFs can occasionally access closed chromatin depending on motif positioning [10,11,94,95], nucleosomal structures [96,97], and cellular concentrations [93,98,99]. Therefore, the hallmark of a pioneer factor should be defined not just by its occupancy of closed chromatin, but by its functional capacity to actively remodel the chromatin landscape.

Our IP-MS and CUT&RUN data provide a biochemical basis for this distinction. We demonstrate that Zld recruits a specific suite of SWI/SNF-family nucleosome remodelers - including PBAP/BAP and NURF, as well as the FACT complex to induce accessibility at its target sites. Such partnership with chromatin remodelers is a conserved feature among high-potency pioneer factors across species [20,100]. For instance, the *Drosophila* pioneer factor GAF interacts directly with the PBAP and NURF complexes to drive accessibility [65,66]. Similarly, mammalian pioneer factors like Oct4, PU.1, and GATA3 have been shown to physically recruit SWI/SNF (BAF) complexes to initiate chromatin opening, which in turn has been shown to stabilize the binding of the pioneer factors in some cases to sustain a permissive state [101–104].

Crucially, TFs appear to differ significantly in their inherent ability to recruit these machines [100,105]. While many early patterning factors may occupy the same enhancers, our results suggest that Zld acts as the primary "recruitment hub" for the remodeling apparatus. By integrating chromatin remodeling with the establishment of accessibility, Zld creates a permissive environment for other TFs that lack the specialized capacity to interface with the remodeling machinery. Our data reinforce a model where the "pioneer" status of Zld is defined by its role as a scaffold that bridges the gap between the closed chromatin state and the active transcriptional machinery.

Recruitment of RNAPII by Zld to enhancers

A striking finding of our study was the robust RNAPII signal at distal Zld-bound enhancers. This contrasts with previous ChIP-chip [76] and ChIP-seq [77] studies where RNAPII was localized almost exclusively to promoters and gene bodies. We propose that RNAPII at enhancers may not be traditionally engaged with DNA in a pre-initiation complex, leading to the low cross-linking efficiency that rendered it "invisible" in prior ChIP assays.

Our results suggest that Zld co-occupies regulatory sites with the transcription machinery. While the spatial overlap between Zld and RNAPII/transcription foci was modest—reflecting the high stochastic background and the likely dynamic nature of the partnership—this finding nonetheless stands in contrast to the "hit-and-run" model proposed based on previous live-imaging studies [34,36,91]. The failure of these studies to detect significant co-localization may be explained by several technical constraints.

First, earlier works focused on either transcription foci from large genes (e.g., hb) or a few exceptionally bright RNAPII clusters, which likely correspond to elongating RNAPII on the gene bodies of highly transcribed loci. For such genes, the physical distance between the distal, enhancer-associated Zld hub and the elongating transcription focus can be significant (> 200 nm); consequently, low spatial correlations are expected [86,106]. Such measurements may not accurately reflect the initial activation events, such as activator-assisted transcription complex assembly at the enhancer/promoter.

Another challenge is the high density of both Zld hubs and RNAPII clusters, which creates substantial background noise in colocalization analysis. Zld clusters are not only more numerous than other factors, but they are also relatively uniform in appearance compared to clusters of Dl [79] or Bicoid(Bcd) [34,35]. This uniformity likely stems from the fact that Zld-target enhancers typically contain a small number of its binding sites, whereas Dl- and Bcd-target enhancers often possess large clusters of their sites. Beyond these common technical challenges, we have crucially shown that the GFP-Zld lines used in previous studies contain fragmented protein. This fragmentation increases diffuse background noise and likely masks the subtle colocalization signals that characterize these dynamic interactions.

Together, our data support a model where enhancers serve as a "docking site" or reservoir for RNAPII. This recruitment may be mediated by NELF-A, which was highly enriched in our IP-MS and associated with Zld-bound enhancers, similar to reports for the ecdysone receptor [107]. Through this mechanism, Zld may recruit RNAPII to enhancers without directly activating transcription-an ability Zld largely lacks. This echoes the behavior of pioneer factors Foxa3 and PHA-4/FoxA, which associate with paused RNAPII to facilitate target gene expression [108,109]. Recruitment by Zld may thus facilitate the rapid assembly of transcription complexes at promoters, promoting the high-level activation required within the short mitotic windows of early development.

dCBP and Fsh are critical for Zld-mediated transcriptional activation

Our functional analysis confirms that Zld initiates a regulatory cascade: Zld → dCBP → H3K18ac → Fsh, consistent with recent findings [78,90,91]. Interestingly, our CUT&RUN data reveals Fsh peaks directly overlapping Zld binding sites even where histone acetylation is depleted due to nucleosome displacement. This suggests that Fsh recruitment may also involve direct protein-protein interactions with Zld, potentially explaining why bromodomains or acetylation marks appear partially dispensable for activation in certain contexts [90,110].

High-resolution imaging revealed hundreds of discrete, weak H3K18ac clusters as early as cycle 8, consistent with the low-level ChIP-seq signals detected at this developmental stage[26]. Both the number and intensity of these clusters increase consecutively through subsequent mitotic cycles. Because H3K18ac is maintained during mitosis [78], this accumulation likely results from continuous *de novo* deposition. Consequently, our results suggest that Zld-driven ZGA is an incremental process characterized by the gradual accumulation of histone acetylation and a progressive increase in enhancer accessibility [29], which collectively culminate in the full activation of target genes.

Interaction of Zld with chromatin boundary factors

Our identification of multiple chromatin boundary factors in the Zld interactome provides a biochemical basis for recent genomic studies. Zld motifs are highly enriched at Topologically Associating Domain (TAD) boundaries, and Zld is essential for the *de novo* formation of these boundaries during the MZT [68,69]. Furthermore, the BEAF-32 motif was previously identified as a top-ranking enriched sequence within Zld ChIP-seq peaks in both early embryos and larval neuroblasts [70].

Among these factors, Row is particularly noteworthy. While not a classical boundary protein, Row was enriched in both our anti-Zld and anti-GFP IP-MS experiments. Recent work has shown that Row interacts with BEAF-32 and other boundary factors to facilitate long-range interactions between regulated genes and TAD boundaries [111]. Interestingly, Row is also critical for neuronal development [112], a role that mirrors Zld's function in maintaining neuroblast identity [70]. Together, our data strengthen the model that Zld organizes higher-order chromatin structure, a function that likely extends beyond early ZGA to influence cell-type-specific development in the nervous system.

Coordination of Zld Activity with the cell cycle

Most intriguingly, the kinase Tlk - the most significantly enriched factor in our IP-MS - interacts directly with Zld via residues 401 - 750. Tlk is a highly conserved protein kinase whose activity is tightly coupled to the cell cycle, peaking during S-phase [67]. Notably, DNA replication has been shown to be required for Zld-mediated activation in early embryos [78]. Taken together, our observation that Tlk directly interacts with Zld suggests that Tlk provides a vital biochemical bridge, ensuring that Zld-driven ZGA is precisely coordinated with the rapid nuclear divisions in early embryos.

Interaction of multiple coordinated factors with Zld

While our study identifies several critical factors involved in Zld activity, a recent study identified an additional cofactor, Ldb1, which acts as a bridge between Zld and dCBP [113]. We

did not detect Ldb1 in our MS analysis, which may be attributed to its low abundance in our embryonic nuclear extracts or the inherent challenges in detecting certain proteins via mass spectrometry. Interestingly, Galouzis et al. found that Zld-induced chromatin accessibility generally does not require Ldb1, suggesting that Zld utilizes distinct sets of cofactors to manage different aspects of its pioneer function. Taken together, these findings reveal a sophisticated network of coordinated factors that facilitate Zld function

In summary, we have provided a comprehensive proteomic and genomic map of the Zelda interactome during the onset of ZGA. Our data reveal that Zld serves as a central hub, integrating the recruitment of nucleosome remodelers to open chromatin, the docking of RNAPII at distal enhancers to prime the transcriptional machinery, and the association of Smr to temper transcriptional output. Furthermore, the direct, cell-cycle-sensitive interaction between Zld and the Tlk kinase provides a potential biochemical link to coordinate these events with the rapid divisions of the early embryo. Collectively, these results elucidate the multi-layered mechanisms by which a master pioneer factor orchestrates the first major transition in embryonic development

Materials and Methods

Antibodies

The antibodies for Zld [114], Tlk [115], dCBP [116], Fsh [117], Smr [118], Brm and PB [119], Ssrp1 [120], NELF-A [121], and Med14 [122] have been described in previous studies, while the other antibodies including those against GFP (#1: Millipore, #AB3080P; #2: Thermo Fisher, #A11122), Fib (Abcam, #ab5821), H3K18ac (PTM Biolabs, #PTM-158), H3K27ac (Thermo, Catalog #MA5-23516), and RNAPII(clone: 4H8; Abcam, #ab5408) were purchased from commercial sources.

Fly stocks

Genotype	source	catalog no.
w[*]; P{w[+mC]=matalpha4-GAL-VP16}V37	BDSC	7063
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00017}attP2	BDSC	33623
y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01570}attP2/TM3, Sb[1]	BDSC	36682
y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMS02340}attP40	BDSC	41943
y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00842}attP2/TM3, Sb[1]	BDSC	34087
zld294, FRT19A/FM7	Liang et al., 2008[23]	
ovoD1,hsFLP112,FRT19A/Y	Liang et al., 2008[23]	
sfGFP-Zld	Hamm et al. 2017[49]	

.

IP - MS analysis for identification of Zld associated factors

Embryo collection

Wild-type (*Oregon-R*) or a line harboring endogenously GFP-tagged Zld [36] were used for IP-MS experiments. Flies were maintained in large population cages at 25°C. Embryos were collected for 1 h and allowed to develop for an additional 1 h 50 min at 25°C(to reach approximately cycle 12–14). Following harvest, embryos were dechorionated in 50% bleach for 3 min and rinsed thoroughly with water.

Embryos were resuspended in a formaldehyde-containing crosslinking buffer (15 mM HEPES, pH 7.6, 60 mM KCl, 15 mM NaCl, 4 mM $MgCl_2$, 3% formaldehyde (Sigma-Aldrich, #F8775), 0.5mM DTT, and cOmplete protease inhibitor cocktail (Sigma-Aldrich, #4693159001)) and homogenized in a Dounce homogenizer (Pestle A, 7 strokes). The homogenate was incubated for 8 min at room temperature. Nuclei were pelleted by centrifugation at 2,500 rpm for 5 min (Eppendorf 5810R, S-4-104 rotor), resuspended in 10 mL PBST (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4, 0.5% Triton X-100), supplemented with 0.125 M glycine, and 0.2 mM PMSF, and Dounce-homogenized again. After 5 min at room temperature, nuclei were pelleted at 3,000 rpm for 5 min, washed three times in PBST + 0.2 mM PMSF, and stored at -80°C.

Preparation of lysate from crosslinked nuclei

Nuclear lysates were prepared as previously described [40] with modifications. Crosslinked nuclei from 5 g of embryos were resuspended in 40 mL LB1 (50mM HEPES-KOH, pH 7.5, 140mM NaCl, 1mM EDTA, 20% glycerol, 0.5% NP-40, 0.25% Triton X-100, 1mM DTT, and 0.2mM PMSF), and homogenized in a dounce homogenizer (Pestle B, 20 strokes). After centrifugation at 2,500 rpm for 5 min, the pellet was washed in 20 mL LB2 (10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and 0.2mM PMSF) and finally resuspended in 3.5 mL LB3(10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-deoxycholate, 0.5% N-lauroylsarcosine, and 0.2mM PMSF).

Samples were sonicated in 400 µL aliquots for 10 cycles (15 s ON/45 s OFF) using a Bioruptor (Diagenode) at medium power. Following sonication, Triton X-100 (1% final), $MgCl_2$ (3 mM final), and Benzonase (100 U/mL) were added, followed by incubation at 30°C for 20 min. After centrifugation (10,000 × g, 10 min), the supernatant was collected. The pellet was re-extracted with 0.8 mL LB3, sonicated for 5 additional cycles, and the supernatants were combined. 5× RIPA buffer was added to a final 1× concentration, and the lysate was cleared by centrifugation at 20,000 × *g* for 10 min.

Immunoprecipitation (IP)reactions

For each IP, 0.35 ml of the nuclear lysate was diluted to 1 ml with RIPA buffer (50 mM TrisCl, pH 8.0, 0.15M NaCl, 0.1%SDS, 1% Triton X-100, 0.1% sodium deoxycholate, 1mM EDTA, 0.2mM PMSF), and incubated with 50 µL Protein A Dynabeads (ThermoFisher, #10002D) pre-bound with 8 µg of antibody: anti-Zld, anti-Fib, anti-GFP #1(Millipore, #AB3080P), or anti-GFP #2(ThermoFisher, #A6455). Reactions were incubated overnight at 4°C. Beads were washed consecutively: twice with RIPA, three times with RIPA + 1.0 M NaCl

(or 2.0 M urea for high-stringency anti-Zld and anti-GFP #1 samples), three times with LiCl buffer (10 mM Tris-Cl, pH 8.0, 0.5 M LiCl, 1% NP-40, 1% Na-deoxycholate, 0.2 mM PMSF), and once with TE buffer.

Preparation of samples for mass spectrometry

Beads were eluted in 60 µL of the elution buffer (10 mM Tris-Cl, 1 mM EDTA, 200 mM NaCl, 1% SDS) at 99°C for 20 min, followed by a second elution in 30 µL. Proteins were reduced with 5 mM DTT (45°C, 30 min) and alkylated with 10 mM iodoacetamide (room temperature, 30 min, dark). Proteins were purified using the SP3 procedure [123] on paramagnetic beads and digested with 0.5 µg trypsin (Thermo Fisher, #90057)in 50 mM ammonium bicarbonate for 16 h at 37°C. Peptides were recovered, and concentrations were determined using Thermo Scientific Pierce Quantitative Colorimetric Peptide Assay Kit (#23275).

Mass Spectrometry analysis and data processing

Peptides were analyzed on a ThermoFisher Orbitrap Fusion Lumos Tribrid system interfaced with an Easy nLC 1200. Separation was performed on a C18 reverse-phase column (25 cm) using an organic gradient (5–30% ACN) over 180 min at 300 nL/min. Spectra were continuously acquired in a data-dependent manner throughout the gradient, acquiring a full scan in the Orbitrap (at 120,000 resolution with an AGC target of 400,000 and a maximum injection time of 50 ms) followed by as many MS/MS scans as could be acquired on the most abundant ions in 3 s in the dual linear ion trap (rapid scan type with an intensity threshold of 5000, HCD collision energy of 32%, AGC target of 10,000, maximum injection time of 30 ms, and isolation width of 0.7 *m/z*). Singly and unassigned charge states were rejected. Dynamic exclusion was enabled with a repeat count of 2, an exclusion duration of 20 s, and an exclusion mass width of ±10 ppm.

Data were analyzed using MaxQuant for label-free quantification (LFQ)[57] against the *D. melanogaster* proteome (FlyBase r6.29). Modifications Oxidation (M), Acetyl (Protein N-term), and Carbamidomethyl (C) were included in protein quantification. Matching between runs was enabled (0.7 min match window, 20 min time window). LFQ min ratio count was set at 1. All other parameters were set as default.

Perseus [58] was used for statistical enrichment. In the analysis, the filters for proteins include to have a minimum 2 valid values in each group. Imputation was used to replace missing values with parameters set at (1.8, 0.3) for the anti-GFP IP-MS data, but not for the anti-Zld and anti-Fib IP-MS data. For T-tests, $S0 = 2$ and FDR = 0.05 were applied, with a final enrichment cutoff of FDR 0.1. GO term enrichment was performed via the Gene Ontology online tool (http://geneontology.org/) using the Fisher's Exact method against the input sample reference list.

CUT&RUN

CUT&RUN experimental procedures

Embryos were fixed and sorted as described in [90]. For each sample we used 75 sorted early (cycle 10-12) or late(cycle 12-13) stage 4 embryos. CUT&RUN experiments were conducted using a kit from Cell Signaling Technology (#86652S) with the buffer recipes following the original CUT&RUN protocol [124].

Briefly, fixed embryos were homogenized in a wash buffer (supplemented with spermidine and protease inhibitor cocktail). Nuclei were pelleted at 1500 × *g* for 5 min at 4°C, washed once, and then incubated with activated Concanavalin A magnetic beads for 20 min at room temperature on a rotator. Bead-bound nuclei were resuspended in antibody binding buffer (supplemented with spermidine, protease inhibitors, and digitonin) and incubated with primary antibodies overnight at 4°C.

Following incubation, beads were washed twice with digitonin buffer and resuspended in 50 µL pAG-MNase pre-mix. After a 1 h incubation at 4°C, beads were washed twice. The pAG-MNase cleavage reaction was performed in 50 µL digitonin buffer supplemented with 1 µL of 100 mM $CaCl_2$ for 1 h at 4°C. Reactions were stopped by adding 50 µL of 2×Stop Buffer (containing digitonin, RNase A, and yeast spike-in DNA) and incubated at 37°C for 10 min. To release DNA fragments, 1.5 µL of 20% SDS and 2 µL of 20 mg/mL Proteinase K were added, followed by overnight incubation at 65°C.

Samples were centrifuged at maximum speed for 5 min, and the supernatant was recovered on a magnetic rack. To enrich for short fragments, Ampure beads were added at a ratio of 65 µL per 100 µL of sample (0.65×). After 3 min, the beads (containing large DNA fragments > 600 bp) were discarded, and the DNA in the supernatant was recovered and purified using the Qiagen PCR purification kit. Sequencing libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit (NEB, #E7645S) and barcoded with NEBNext Multiplex Oligos Dual Index Primers. Libraries were sequenced on an Illumina NextSeq P1 platform.

CUT&RUN data analysis

The 150 bp paired-end (PE) reads were mapped to the *D. melanogaster* genome (BDGP6) and the *S. cerevisiae* genome (R64-1-1; for spike-in normalization) using Bowtie2 with the following parameters: "--end-to-end --no-mixed --no-discordant -I 10 -X 700 -p 8 -S sample-name.sam". Wig files were generated and normalized by scaling to 10 million aligned reads, followed by secondary normalization using the ratio of spike-in reads relative to the Zld sample. For visualization, signals were smoothed using a 0.3 kb window.

Peaks were called using MACS2[125] with the following options: -c Input-DNA-reads -f BEDPE -g dm --min-length 500 --max-gap 500 --keep-dup all --call-summits. Inclusion of input DNA reads as control suppressed the calling of peaks over repeated sequences, but not entirely. To further eliminate artifacts from repetitive sequences, we used an IgG CUT&RUN sample as a control; any peaks showing a five-fold or greater enrichment in the IgG sample were discarded.

Fly crosses for maternal RNAi knockdown

To achieve maternal knockdown, virgin females carrying *UASp-shRNA* transgenes were crossed to *Mat-tub-Gal4* males to drive shRNA expression during the later stages of oogenesis. Progeny siblings were intercrossed, and F1 progeny were grown at 27°C starting from the larval stage, to increase the Gal4/UASp induction [126]. Embryos were collected for 2.5 h, aged for 40 min, and fixed according to the protocol described by [127].

Zld mutant germline clone generation

Maternal *zld* mRNA depletion was achieved using the FLP-DFS (dominant female sterile) technique [128,129] as previously described [26]. Briefly, zld^{294}, *FRT19A*/*FM7* [23] virgin females were crossed with *ovoD1,hsFLP112,FRT19A*/Y [23] males. The larvae developed from embryos laid by females from these crosses were heat-shocked twice at 37°C for 2 h each (at 24–48 h and 48–72 h post-hatching) to induce mitotic recombination. Embryos from the resulting progeny were collected and fixed as described for the RNAi experiments.

Immunofluorescence, smiFISH, and 3D co-localization analysis

Immunostaining

Embryos were collected from wild-type Oregon-R, *zld−/−* germline clones, pLEP_*hb*-stripe, pLEP_*hb*-stripe(zm), or RNAi knockdown flies (see below) and fixed as previously described [127]. For each sample, 20 µL were rehydrated, washed in PBST buffer (1×PBS buffer + 0.3% Triton X-100), and blocked for 45 min at room temperature in Western Blocking Reagent (Sigma Aldrich, #11921673001) diluted in 1×PBS. Primary antibody incubation was performed overnight on a rocker at 4°C in 200 µL blocking solution using the following antibodies at recommended dilutions: anti-Zld, anti-Fsh[117], anti-H3K18ac (PTM Biolabs, #PTM-158), anti-RNAPII (4H8) (Abcam, #ab5408), and anti-GFP (Aves Labs, Inc, #1010).

The embryos were briefly washed three times with PBST, followed by six washes of 20 min each (1 mL PBST per wash). After a second blocking step, embryos were incubated with Alexa Fluor-conjugated secondary antibodies (647, 555, or 488; Thermo Fisher), diluted as recommended in PBST for 2 hrs at room temperature. After the incubation, the embryos were washed five times for 20 min each with PBST. Nuclei were stained with Hoechst 33342 (Thermo Fisher, #H3570) at the concentration of 1 µg/ml for 30 min. Embryos were mounted in ProLong™ Diamond Antifade Mountant (Thermo Fisher, #P36961) (60 µL suspension in mountant per slide) and cured for 24 hrs in the dark before imaging.

smFISH

To analyze colocalization of Zld clusters with transcription foci of *zen, dfd, and lacZ*, immunostaining with anti-Zld antibody was performed following smiFISH as described in Calvo

et al. [130]. For smiFISH, 20 nt gene-specific primary probes used for the assay (42 for *zen*-RA, 36 for *dfd*-RA, and 49 for *lacZ*) were designed using the Stellaris® RNA FISH Probe Designer (Biosearch Technologies, Inc.), and synthesized by IDT(Integrated DNA Technologies); all probe sequences are provided in Supplementary Table 3. Each probe included a 5′ common sequence (CCTCCTAAGTTTCGAGCTGGACTCAGTG) for annealing to a X-FLAP oligo (CACTGAGTCCAGCTCGAAACTTAGGAGG) labeled at both 5′ and 3′ ends with Alexa Fluor® 488 to allow detection.

For detection of the RNA of *sala*, *hb*, and *tld* in RNAi knockdown embryos, the smiFISH method was employed, with the use of a set of 20 probes for *sala-RA*, 43 for *hb-RA*, and 40 for *tld-RA*, respectively. X-FLAP oligo was labeled at both 5′ and 3′ ends with Alexa Fluor® 488(*hb*) or 647(*sal* and *tld*). For *btd*, smFISH using Custom Stellaris® FISH probes (Biosearch Technologies, Inc.) labeled with Quasar 670 was performed following the standard smFISH method (Biosearch Technologies, Inc.). For co-detection of RNAPII, smFISH was followed by standard immunostaining using the RNAPII 4H8 antibody (1:400) and Alexa Fluor 555-conjugated secondary antibody (1:1000). Samples were mounted in ProLong™ Diamond Antifade as described above.

Image acquisition

High-resolution imaging was performed on a Zeiss LSM800 Airyscan microscope using a Plan-Apochromat 63×1.4 NA oil DIC M27 objective. Images were acquired in Airyscan SR mode. The standard acquisition settings were as follows: pixel size of 0.035 µm, z-slice thickness of 0.13 µm, frame scan mode, 16-bit depth, line averaging of 2, digital gain of 2, master gain of approximately 850V. Laser power, digital gain, and master gain were adjusted such that the peak signal intensity was approximately 25% of the saturation level. Experiment-specific variations to these settings are detailed below.

For Zld and RNAPII colocalization analysis (Fig. 6), a 6× digital zoom and a scan speed of 5 were used. Laser power was set to 0.1% for both Alexa Fluor 555 and 488 channels. The image size was 16 µm × 16µm(454 × 454 pixels). and fields of view were selected from the central region of the embryos. To analyze the effects of *zld* mutation and *nej* RNAi knockdown on H3K18ac and Fsh cluster formation(Fig. 8), settings were identical except for a scan speed of 6, a digital gain of 1, and laser powers of 0.2% (Alexa 555) and 0.1% (Alexa 488).

For colocalization analysis of Zld with target gene transcription foci(Fig. 7), the zoom was 2× with a scan speed of 5. Laser powers were 1% for smFISH channels and 0.3% for the anti-Zld immunofluorescence channel. The image size was 50 µm × 50 µm(1412 × 1412 pixels). Images for *zen* and *dfd* were acquired from the dorsal-anterior region where their expression domains overlap. The *lacZ* imaging was performed over the posterior stripe controlled by the *hb* stripe enhancer.

To analyze the effects of *fs(1)h* and *nej* RNAi knockdown on transcription of Zld target and non-target genes, a 1.4× zoom and a scan speed of 6 were used for *btd* smFISH in *fs(1)h* RNAi and its corresponding w RNAi control embryos. Other images were acquired with a 1.0× zoom

and scan speed of 5. Laser powers were 1% for all smFISH channels and 0.1% for RNAPII. The image size was 50 μm × 72 μm (1400 × 2028 pixels) for *btd*, and 25 μm × 100 μm(688 × 2850 pixels) for all others. Imaging for *hb*, *sala*, and RNAPII was centered around 30%–50% egg length(EL). For *btd*, imaging was centered at approximately 30% EL (a 75 μm × 25 μm section centered on the *btd* stripe was subsequently cropped for analysis). For *tld*, the imaging area was centered at 25 μm toward the dorsal side at the embryo midline.

3D Object-based co-localization analysis

Image processing was performed in Fiji. To eliminate background, signals outside the nuclear boundaries in each Z-slice were removed using the "Clear Outside" function. Transcription foci (typically 1–2 per nucleus) (for *zen*, *dfd*, and *lacZ*) were identified as distinct 3D maxima using Fiji following background subtraction; intensity cutoffs were established for each gene to distinguish signal from noise and remained consistent across all images in an experiment. Nuclear spots from immunostaining images (anti-Zld, RNAPII 4H8, H3K18ac, Fsh, anti-GFP for GFP-Zld) were segmented using the Trainable Weka Segmentation 3D plugin. The resulting classified image stacks were converted into masks and applied to the original images using the 3D Object Counter. To identify local signal maxima within segmented objects, a 3D Maxima Filter plugin was applied to the masked images. For the Fsh-H3K18ac analysis, we further refined the segmentation by applying dilate and watershed functions to the classified masks to ensure the separation of fused clusters; this yielded fewer but more spatially distinct spots. For the Zld-RNAPII analysis, a one-slice Z-shift between the Alexa 555 and Alexa 488 channels was corrected prior to image processing.

Co-localization was quantified by calculating the Euclidean distance from each peak in the source channel to the nearest peak in the target channel. The search radius included the immediate z-slice and adjacent slices (assuming a z-dimension distance of 0.067 μm, half the slice thickness). To ensure we analyzed meaningful protein clusters rather than single molecules or background, we applied an intensity cutoff; only peaks with intensities at least twice the average intensity of single-molecule clusters (identified as small, uniform-intensity spots outside the nuclei) were included.

To statistically validate colocalization, we utilized a 90-degree rotation control. The target channel image and its associated peak coordinates were rotated by 90 degrees, and Euclidean distances to the reference channel were recalculated. To ensure a fair comparison, only peaks that remained within the boundaries of the rotated image frame were included in the final analysis. This approach provides a more stringent control for stochastic overlap than random-coordinate assignment, as it preserves the original spatial architecture and cluster density while specifically disrupting biological coordination.

Image analysis for the RNAi experiments

Individual RNA molecules and transcription foci (typically 1–2 per nucleus) were identified as distinct 3D maxima using Fiji following background subtraction, Intensity cutoffs

were established for each gene to distinguish signal from noise and remained consistent across all images in an experiment. RNAPII clusters corresponding to the Histone Locus Body (HLB) were analyzed by applying a Gaussian filter followed by segmentation with the Trainable Weka Segmentation 3D plugin. Masks generated from classified images were processed using the "Analyze Particles" and "Measure" functions to determine cluster intensities.

RNA-seq in Smr RNAi knockdown embryos

Embryos from RNAi knockdown flies for the *smr* gene or the *white* (*w*) gene (control) were generated and collected as described above. Three biological replicates were prepared for both the *smr* RNAi and *w* RNAi groups, using 50 sorted Stage 4 embryos per sample. Total RNA was extracted using TRIzol (Thermo Fisher Scientific) and treated with TURBO DNase I.

Library preparation and sequencing were performed by the QB3-Berkeley Genomics core facility. Briefly, poly(A) mRNA was enriched using poly-dT beads, and libraries were constructed using the KAPA mRNA HyperPrep kit (Roche, #KK8581). Truncated universal stub adapters were ligated to cDNA fragments, followed by PCR extension using unique dual-indexing primers to generate full-length Illumina adapters. Libraries were pooled by molarity and sequenced on an Illumina NovaSeq X (10B flow cell) for 2 × 150 cycles.

Sequencing reads were mapped to the *D. melanogaster* genome (r6.59) using the STAR (v2.7.11b) aligner [131]. RNA-seq aligner to map the resulting reads to the *D. melanogaster* genome (r6.59). Gene-level coverage was generated from the resulting BAM files using featureCounts[132]. Differential expression analysis was performed using the exactTest method within the edgeR pipeline [133] to identify genes significantly affected by *smr* knockdown.

To investigate the specific regulatory link between Zld and Smr, Zld peaks from CUT&RUN experiments were assigned to the nearest gene within a 10 kb window. For genes associated with multiple Zld peaks, only the peak with the highest signal intensity was retained. Zld peaks were then partitioned into quantiles based on intensity. We analyzed the distribution of $\log_2$FC values for genes within each quantile to assess the relationship between Zld binding strength and Smr-dependent regulation. As a control, $\log_2$FC distributions were compared against randomly selected gene sets of equal size drawn from the complete edgeR output.

In vitro protein binding assay for Tlk and Zld interaction.

To generate protein expression plasmids, DNA sequences encoding specific Zld segments (aa 1-400, aa 401-750, aa 751-1250, and aa 1251-1596) were cloned into pGEX-6P-1-H-RBS vector (GenScript). The GST-fusion proteins were expressed in *E. coli* (NEB, #C3037) and purified using Pierce™ Glutathione Magnetic Agarose Beads (Thermo Scientific, catalog#: 78601) according to the manufacturer's instructions. Briefly, bacterial lysates were prepared in lysis buffer (20 mM Tris-HC1 at pH 8.0, 0.2 M NaC1, 20 μM $ZnC1_2$, 1 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride (PMSF), and cOmplete protease inhibitor cocktail

(Sigma-Aldrich, #4693159001)) and incubated with the magnetic beads. Following extensive washing, the GST-fusion proteins remained coupled to the beads for use in the binding assays. Protein concentrations on the beads were estimated by SDS-PAGE analysis of 5 µL bead aliquots alongside a bovine serum albumin (BSA) standard curve.

Tlk in vitro transcription/translation plasmid was constructed by inserting the Tlk encoding sequence into a pTnT Vector (Promega, cat no. L5610). Tlk protein was generated using the TnT® Coupled Reticulocyte Lysate System (Promega, #L4610) following the manufacturer's protocol.

For the binding assay, approximately 3 µg of the purified GST or GST-fusion proteins immobilized on glutathione-agarose beads were incubated with 5 µL of the TnT product. The reaction was performed in 0.6 mL of protein-binding buffer (20 mM HEPES at pH 7.6, 100 mM KC1, 20 µM $ZnC1_2$, 2.5 mM $MgC1_2$, 10% glycerol, 0.025% NP-40, 0.5 mM DTT, 0.5 mM PMSF) supplemented with 600 µg of BSA to minimize non-specific interactions. After incubation at 4°C for 2 h with continuous mixing, the beads were washed four times with 1 mL of binding buffer. Bound Tlk protein was eluted in SDS sample buffer, resolved by SDS-PAGE, and detected by immunoblotting with an anti-Tlk antibody.

Acknowledgements

We would like to thank members of the Eisen lab for helpful discussion and inputs throughout the project, and Drs. Susumu Hirose, Alexander Mazo, Jenn-Yah Yu, Ong Chin Tong, and Nicola Iovino for generous gifts of antibodies; We thank Lori Kohlstaedt of the Proteomics/Mass Spectrometry Laboratory at the University of California at Berkeley for mass spectrometry, and the QB3 Genomics, UC Berkeley, Berkeley, CA, RRID:SCR_022170 for RNA-seq library preparation and the next generation sequencing services. This was funded by an HHMI Investigator award to MBE.

Author contributions

X.Y.Li and M.B.Eisen conceived the project. X.Y.Li performed all the experiments and data analysis. M.B. Eisen supervised the project. X.Y.Li and M.B.Eisen prepared the manuscript.

Funding

This was funded by an HHMI Investigator award to MBE

References

1. Lelli KM, Slattery M, Mann RS. Disentangling the many layers of eukaryotic transcriptional regulation. Annu Rev Genet. 2012;46: 43–68.

2. Spitz F, Furlong EEM. Transcription factors: from enhancer binding to developmental control. Nat Rev Genet. 2012;13: 613–626.

3. Levine M, Cattoglio C, Tjian R. Looping back to leap forward: transcription enters a new era. Cell. 2014;157: 13–25.

4. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997;389: 251–260.

5. Tillo D, Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Field Y, et al. High nucleosome occupancy is encoded at human regulatory sequences. PLoS One. 2010;5: e9129.

6. Barozzi I, Simonatto M, Bonifacio S, Yang L, Rohs R, Ghisletti S, et al. Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. Mol Cell. 2014;54: 844–857.

7. 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;25: 1703–1714.

8. Li X-Y, Eisen MB. Effects of the maternal factor Zelda on zygotic enhancer activity in the*Drosophila*embryo. bioRxiv. bioRxiv; 2018. doi:10.1101/385070

9. Veil M, Yampolsky LY, Grüning B, Onichtchouk D. Pou5f3, SoxB1, and Nanog remodel chromatin on high nucleosome affinity regions at zygotic genome activation. Genome Res. 2019;29: 383–395.

10. Zhu F, Farnung L, Kaasinen E, Sahu B, Yin Y, Wei B, et al. The interaction landscape between transcription factors and the nucleosome. Nature. 2018;562: 76–81.

11. Fernandez Garcia M, Moore CD, Schulz KN, Alberto O, Donague G, Harrison MM, et al. Structural Features of Transcription Factors Associating with Nucleosome Binding. Mol Cell. 2019;75: 921–932.e6.

12. 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;12: R34.

13. Guertin MJ, Lis JT. Chromatin landscape dictates HSF binding to target DNA elements. PLoS Genet. 2010;6: e1001114.

14. Guertin MJ, Lis JT. Mechanisms by which transcription factors gain access to target

sequence elements in chromatin. Curr Opin Genet Dev. 2013;23: 116–123.

15. Biddie SC, John S, Sabo PJ, Thurman RE, Johnson TA, Schiltz RL, et al. Transcription factor AP1 potentiates chromatin accessibility and glucocorticoid receptor binding. Mol Cell. 2011;43: 145–155.

16. Bulyk ML, Drouin J, Harrison MM, Taipale J, Zaret KS. Pioneer factors - key regulators of chromatin and gene expression. Nat Rev Genet. 2023;24: 809–815.

17. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. Genes Dev. 2011;25: 2227–2241.

18. Iwafuchi-Doi M, Zaret KS. Pioneer transcription factors in cell reprogramming. Genes Dev. 2014;28: 2679–2692.

19. Larson ED, Marsh AJ, Harrison MM. Pioneering the developmental frontier. Mol Cell. 2021;81: 1640–1650.

20. Barral A, Zaret KS. Pioneer factors: roles and their regulation in development. Trends Genet. 2024;40: 134–148.

21. Zaret KS. Pioneer Transcription Factors Initiating Gene Network Changes. Annu Rev Genet. 2020;54: 367–385.

22. Mayran A, Drouin J. Pioneer transcription factors shape the epigenetic landscape. J Biol Chem. 2018;293: 13795–13804.

23. 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;456: 400–403.

24. 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;7: e1002266.

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;7: e1002339.

26. 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. 2014;3. doi:10.7554/eLife.03737

27. 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;25: 1715–1726.

28. 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;24: 1341–1346.

29. Blythe SA, Wieschaus EF. Establishment and maintenance of heritable chromatin structure during early Drosophila embryogenesis. Elife. 2016;5. doi:10.7554/eLife.20148

30. Brennan KJ, Weilert M, Krueger S, Pampari A, Liu H-Y, Yang AWH, et al. Chromatin accessibility in the Drosophila embryo is determined by transcription factor pioneering and enhancer activation. Dev Cell. 2023;58: 1898–1916.e9.

31. 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;28: 608–621.

32. Yáñez-Cuna JO, Dinh HQ, Kvon EZ, Shlyueva D, Stark A. Uncovering cis-regulatory sequence requirements for context-specific transcription factor binding. Genome Res. 2012;22: 2018–2030.

33. Yamada S, Whitney PH, Huang S-K, Eck EC, Garcia HG, Rushlow CA. The Drosophila pioneer factor Zelda modulates the nuclear microenvironment of a dorsal target enhancer to potentiate transcriptional output. Curr Biol. 2019;29: 1387–1393.e5.

34. Mir M, Stadler MR, Ortiz SA, Hannon CE, Harrison MM, Darzacq X, et al. Dynamic multifactor hubs interact transiently with sites of active transcription in Drosophila embryos. Elife. 2018;7. doi:10.7554/eLife.40497

35. 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;31: 1784–1794.

36. Dufourt J, Trullo A, Hunter J, Fernandez C, Lazaro J, Dejean M, et al. Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. Nat Commun. 2018;9: 5194.

37. Engelen E, Brandsma JH, Moen MJ, Signorile L, Dekkers DHW, Demmers J, et al. Proteins that bind regulatory regions identified by histone modification chromatin immunoprecipitations and mass spectrometry. Nature Communications. 2015. doi:10.1038/ncomms8155

38. Wierer M, Mann M. Proteomics to study DNA-bound and chromatin-associated gene regulatory complexes. Hum Mol Genet. 2016;25: R106–R114.

39. Mohammed H, D'Santos C, Serandour AA, Ali HR, Brown GD, Atkins A, et al. Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. Cell Rep. 2013;3: 342–349.

40. Mohammed H, Taylor C, Brown GD, Papachristou EK, Carroll JS, D'Santos CS. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. Nature Protocols. 2016. pp. 316–326. doi:10.1038/nprot.2016.020

41. Rivera-Reyes R, Kleppa M-J, Kispert A. Proteomic analysis identifies transcriptional cofactors and homeobox transcription factors as TBX18 binding proteins. PLoS One.

2018;13: e0200964.

42. Lomaev D, Mikhailova A, Erokhin M, Shaposhnikov AV, Moresco JJ, Blokhina T, et al. The GAGA factor regulatory network: Identification of GAGA factor associated proteins. PLoS One. 2017;12: e0173602.

43. Göös H, Kinnunen M, Salokas K, Tan Z, Liu X, Yadav L, et al. Human transcription factor protein interaction networks. Nat Commun. 2022;13: 766.

44. Li X, Fu X-D. Chromatin-associated RNAs as facilitators of functional genomic interactions. Nat Rev Genet. 2019;20: 503–519.

45. Hou TY, Kraus WL. Spirits in the Material World: Enhancer RNAs in Transcriptional Regulation. Trends Biochem Sci. 2021;46: 138–153.

46. Xiao R, Chen J-Y, Liang Z, Luo D, Chen G, Lu ZJ, et al. Pervasive Chromatin-RNA Binding Protein Interactions Enable RNA-Based Regulation of Transcription. Cell. 2019;178: 107–121.e18.

47. Kalkat M, Resetca D, Lourenco C, Chan P-K, Wei Y, Shiah Y-J, et al. MYC Protein Interactome Profiling Reveals Functionally Distinct Regions that Cooperate to Drive Tumorigenesis. Mol Cell. 2018;72: 836–848.e7.

48. Hamm DC, Bondra ER, Harrison MM. Transcriptional activation is a conserved feature of the early embryonic factor Zelda that requires a cluster of four zinc fingers for DNA binding and a low-complexity activation domain. J Biol Chem. 2015;290: 3508–3518.

49. Hamm DC, Larson ED, Nevil M, Marshall KE, Bondra ER, Harrison MM. A conserved maternal-specific repressive domain in Zelda revealed by Cas9-mediated mutagenesis in Drosophila melanogaster. PLoS Genet. 2017;13: e1007120.

50. Goodman RH, Smolik S. CBP/p300 in cell growth, transformation, and development. Genes Dev. 2000;14: 1553–1577.

51. Huang DH, Dawid IB. The maternal-effect gene fsh is essential for the specification of the central region of the Drosophila embryo. New Biol. 1990;2: 163–170.

52. Tsai CC, Kao HY, Yao TP, McKeown M, Evans RM. SMRTER, a Drosophila nuclear receptor coregulator, reveals that EcR-mediated repression is critical for development. Mol Cell. 1999;4: 175–186.

53. Tsuda L, Nagaraj R, Zipursky SL, Banerjee U. An EGFR/Ebi/Sno pathway promotes delta expression by inactivating Su(H)/SMRTER repression during inductive notch signaling. Cell. 2002;110: 625–637.

54. Segura-Bayona S, Stracker TH. The Tousled-like kinases regulate genome and epigenome stability: implications in development and disease. Cell Mol Life Sci. 2019;76: 3827–3841.

55. Papachristou EK, Kishore K, Holding AN, Harvey K, Roumeliotis TI, Chilamakuri CSR, et al. A quantitative mass spectrometry-based approach to monitor the dynamics of endogenous chromatin-associated protein complexes. Nat Commun. 2018;9: 2311.

56. Rafiee M-R, Girardot C, Sigismondo G, Krijgsveld J. Expanding the Circuitry of Pluripotency by Selective Isolation of Chromatin-Associated Proteins. Mol Cell. 2016;64: 624–635.

57. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc. 2016;11: 2301–2319.

58. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote) omics data. Nat Methods. 2016;13: 731–740.

59. Galardi S, Fatica A, Bachi A, Scaloni A, Presutti C, Bozzoni I. Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA in vitro. Mol Cell Biol. 2002;22: 6663–6668.

60. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51: D638–D646.

61. Bouazoune K, Brehm A. ATP-dependent chromatin remodeling complexes in Drosophila. Chromosome Res. 2006;14: 433–449.

62. Clapier CR, Cairns BR. The biology of chromatin remodeling complexes. Annu Rev Biochem. 2009;78: 273–304.

63. Scacchetti A, Schauer T, Reim A, Apostolou Z, Campos Sparr A, Krause S, et al. Drosophila SWR1 and NuA4 complexes are defined by DOMINO isoforms. Elife. 2020;9. doi:10.7554/eLife.56325

64. Jeronimo C, Robert F. The histone chaperone FACT: a guardian of chromatin structure integrity. Transcription. 2022;13: 16–38.

65. Judd J, Duarte FM, Lis JT. Pioneer-like factor GAF cooperates with PBAP (SWI/SNF) and NURF (ISWI) to regulate transcription. Genes Dev. 2021;35: 147–156.

66. Chetverina D, Erokhin M, Schedl P. GAGA factor: a multifunctional pioneering chromatin protein. Cell Mol Life Sci. 2021;78: 4125–4141.

67. Silljé HH, Takahashi K, Tanaka K, Van Houwe G, Nigg EA. Mammalian homologues of the plant Tousled gene code for cell-cycle-regulated kinases with maximal activities linked to ongoing DNA replication. EMBO J. 1999;18: 5691–5702.

68. Espinola SM, Götz M, Bellec M, Messina O, Fiche J-B, Houbron C, et al. Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate

during early Drosophila development. Nat Genet. 2021;53: 477–486.

69. Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. Cell. 2017;169: 216–228.e19.

70. Larson ED, Komori H, Gibson TJ, Ostgaard CM, Hamm DC, Schnell JM, et al. Cell-type-specific chromatin occupancy by the pioneer factor Zelda drives key developmental transitions in Drosophila. Nat Commun. 2021;12: 7153.

71. Duan J, Rieder L, Colonnetta MM, Huang A, Mckenney M, Watters S, et al. CLAMP and Zelda function together to promote Drosophila zygotic genome activation. Elife. 2021;10. doi:10.7554/eLife.69937

72. Colonnetta MM, Abrahante JE, Schedl P, Gohl DM, Deshpande G. CLAMP regulates zygotic genome activation in Drosophila embryos. Genetics. 2021;219. doi:10.1093/genetics/iyab107

73. Blythe SA, Wieschaus EF. Zygotic genome activation triggers the DNA replication checkpoint at the midblastula transition. Cell. 2015;160: 1169–1181.

74. Kaushal A, Dorier J, Wang B, Mohana G, Taschner M, Cousin P, et al. Essential role of Cp190 in physical and regulatory boundary formation. Sci Adv. 2022;8: eabl8834.

75. Alekseyenko AA, Gorchakov AA, Zee BM, Fuchs SM, Kharchenko PV, Kuroda MI. Heterochromatin-associated interactions of Drosophila HP1a with dADD1, HIPP1, and repetitive RNAs. Genes Dev. 2014;28: 1445–1460.

76. 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;6: e27.

77. 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;2: e00861.

78. Cho C-Y, Kemp JP Jr, Duronio RJ, O'Farrell PH. Coordinating transcription and replication to mitigate their conflicts in early Drosophila embryos. Cell Rep. 2022;41: 111507.

79. Yamada S, Whitney PH, Huang S-K, Eck EC, Garcia HG, Rushlow CA. The Drosophila pioneer factor Zelda modulates the nuclear microenvironment of a Dorsal target enhancer to potentiate transcriptional output. bioRxiv. 2018. p. 471029. doi:10.1101/471029

80. Haynes SR, Mozer BA, Bhatia-Dey ~. Naina, Dawid IB. The Drosophila fsh Locus, a Maternal Effect Homeotic Gene, Encodes Apparent Membrane Proteins'. Dev Biol.

81. Chang Y-L, King B, Lin S-C, Kennison JA, Huang D-H. A double-bromodomain protein, FSH-S, activates the homeotic gene ultrabithorax through a critical promoter-proximal region. Mol Cell Biol. 2007;27: 5486–5498.

82. Wu S-Y, Chiang C-M. The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. J Biol Chem. 2007;282: 13141–13145.

83. Dey A, Nishiyama A, Karpova T, McNally J, Ozato K. Brd4 marks select genes on mitotic chromatin and directs postmitotic transcription. Mol Biol Cell. 2009;20: 4899–4909.

84. Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, et al. Coactivator condensation at super-enhancers links phase separation and gene control. Science. 2018;361. doi:10.1126/science.aar3958

85. Han X, Yu D, Gu R, Jia Y, Wang Q, Jaganathan A, et al. Roles of the BRD4 short isoform in phase separation and active gene transcription. Nat Struct Mol Biol. 2020;27: 333–341.

86. Li J, Hsu A, Hua Y, Wang G, Cheng L, Ochiai H, et al. Single-gene imaging links genome topology, promoter-enhancer communication and transcription control. Nat Struct Mol Biol. 2020;27: 1032–1040.

87. Chan SH, Tang Y, Miao L, Darwich-Codore H, Vejnar CE, Beaudoin J-D, et al. Brd4 and P300 Confer Transcriptional Competency during Zygotic Genome Activation. Dev Cell. 2019;49: 867–881.e8.

88. Silljé HH, Nigg EA. Identification of human Asf1 chromatin assembly factors as substrates of Tousled-like kinases. Curr Biol. 2001;11: 1068–1073.

89. Carrera P, Moshkin YM, Gronke S, Sillje HHW, Nigg EA, Jackle H, et al. Tousled-like kinase functions with the chromatin assembly pathway regulating nuclear divisions. Genes Dev. 2003;17: 2578–2590.

90. Ciabrelli F, Rabbani L, Cardamone F, Zenk F, Löser E, Schächtle MA, et al. CBP and Gcn5 drive zygotic genome activation independently of their catalytic activity. Sci Adv. 2023;9: eadf2687.

91. Cho C-Y, O'Farrell PH. Stepwise modifications of transcriptional hubs link pioneer factor activity to a burst of transcription. Nat Commun. 2023;14: 4848.

92. Li X-Y, Eisen MB. Zelda potentiates transcription factor binding to zygotic enhancers by increasing local chromatin accessibility during early Drosophila melanogaster embryogenesis. bioRxiv. 2018. p. 380857. doi:10.1101/380857

93. Gibson TJ, Larson ED, Harrison MM. Protein-intrinsic properties and context-dependent effects regulate pioneer factor binding and function. Nat Struct Mol Biol. 2024;31: 548–558.

94. Michael AK, Grand RS, Isbel L, Cavadini S, Kozicka Z, Kempf G, et al. Mechanisms of OCT4-SOX2 motif readout on nucleosomes. Science. 2020;368: 1460–1465.

95. Dodonova SO, Zhu F, Dienemann C, Taipale J, Cramer P. Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function. Nature. 2020;580: 669–672.

96. Mariani L, Liu X, Lee K, Gisselbrecht SS, Cole PA, Bulyk ML. DNA bendability regulates transcription factor binding to nucleosomes. Nat Struct Mol Biol. 2025;32: 2185–2195.

97. Tsompana M, Wilson PD, Murugaiyan V, Handelmann CR, Buck MJ. Defining transcription factor nucleosome binding with Pioneer-seq. PLoS Genet. 2025;21: e1011813.

98. Hansen JL, Cohen BA. A quantitative metric of pioneer activity reveals that HNF4A has stronger in vivo pioneer activity than FOXA1. Genome Biol. 2022;23: 221.

99. Hansen JL, Loell KJ, Cohen BA. A test of the pioneer factor hypothesis using ectopic liver gene activation. Elife. 2022;11. doi:10.7554/elife.73358

100. Gourisankar S, Krokhotin A, Wenderski W, Crabtree GR. Context-specific functions of chromatin remodellers in development and disease. Nat Rev Genet. 2024;25: 340–361.

101. Frederick MA, Williamson KE, Fernandez Garcia M, Ferretti MB, McCarthy RL, Donahue G, et al. A pioneer factor locally opens compacted chromatin to enable targeted ATP-dependent nucleosome remodeling. Nat Struct Mol Biol. 2023;30: 31–37.

102. Takaku M, Grimm SA, Shimbo T, Perera L, Menafra R, Stunnenberg HG, et al. GATA3-dependent cellular reprogramming requires activation-domain dependent recruitment of a chromatin remodeler. Genome Biol. 2016;17: 36.

103. King HW, Klose RJ. The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. Elife. 2017;6. doi:10.7554/eLife.22631

104. Friman ET, Deluz C, Meireles-Filho AC, Govindan S, Gardeux V, Deplancke B, et al. Dynamic regulation of chromatin accessibility by pluripotency transcription factors across the cell cycle. Elife. 2019;8. doi:10.7554/eLife.50087

105. Jain SU, Williamson KE, Ying AW, Turner AM, Jiang RJ, Raval S, et al. A SWI/SNF-specific Ig-like domain, SWIFT, is a transcription factor binding platform. Science. 2026; eaeb3627.

106. Yang JH, Hansen AS. Enhancer selectivity in space and time: from enhancer-promoter interactions to promoter activation. Nat Rev Mol Cell Biol. 2024;25: 574–591.

107. Mazina MY, Kovalenko EV, Vorobyeva NE. The negative elongation factor NELF promotes induced transcriptional response of Drosophila ecdysone-dependent genes. Sci Rep. 2021;11: 172.

108. Hsu H-T, Chen H-M, Yang Z, Wang J, Lee NK, Burger A, et al. TRANSCRIPTION. Recruitment of RNA polymerase II by the pioneer transcription factor PHA-4. Science. 2015;348: 1372–1376.

109. Horisawa K, Udono M, Ueno K, Ohkawa Y, Nagasaki M, Sekiya S, et al. The dynamics of transcriptional activation by hepatic reprogramming factors. Mol Cell. 2020;79:

660–676.e8.

110. Zheng B, Gold S, Iwanaszko M, Howard BC, Wang L, Shilatifard A. Distinct layers of BRD4-PTEFb reveal bromodomain-independent function in transcriptional regulation. Mol Cell. 2023;83: 2896–2910.e4.

111. Herman N, Kadener S, Shifman S. The chromatin factor ROW cooperates with BEAF-32 in regulating long-range inducible genes. EMBO Rep. 2022;23: e54720.

112. Stessman HAF, Willemsen MH, Fenckova M, Penn O, Hoischen A, Xiong B, et al. Disruption of POGZ Is Associated with Intellectual Disability and Autism Spectrum Disorders. Am J Hum Genet. 2016;98: 541–552.

113. Galouzis CC, Kherdjemil Y, Forneris M, Viales RR, Marco-Ferreres R, Furlong EEM. Chip (Ldb1) is a putative cofactor of Zelda forming a functional bridge to CBP during zygotic genome activation. Mol Cell. 2025;85: 2425–2441.e9.

114. Harrison MM, Botchan MR, Cline TW. Grainyhead and Zelda compete for binding to the promoters of the earliest-expressed Drosophila genes. Dev Biol. 2010;345: 248–255.

115. Yeh T-H, Huang S-Y, Lan W-Y, Liaw G-J, Yu J-Y. Modulation of cell morphogenesis by tousled-like kinase in the Drosophila follicle cell. Dev Dyn. 2015;244: 852–865.

116. Takakura M, Nakagawa R, Ota T, Kimura Y, Ng MY, Alia AG, et al. Rpd3/CoRest-mediated activity-dependent transcription regulates the flexibility in memory updating in Drosophila. Nat Commun. 2021;12: 628.

117. Kockmann T, Gerstung M, Schlumpf T, Xhinzhou Z, Hess D, Beerenwinkel N, et al. The BET protein FSH functionally interacts with ASH1 to orchestrate global gene activity in Drosophila. Genome Biol. 2013;14: R18.

118. Johnston DM, Sedkov Y, Petruk S, Riley KM, Fujioka M, Jaynes JB, et al. Ecdysone- and NO-mediated gene regulation by competing EcR/Usp and E75A nuclear receptors during Drosophila development. Mol Cell. 2011;44: 51–61.

119. Nakayama T, Shimojima T, Hirose S. The PBAP remodeling complex is required for histone H3.3 replacement at chromatin boundaries and for boundary functions. Development. 2012;139: 4582–4590.

120. Shimojima T, Okada M, Nakayama T, Ueda H, Okawa K, Iwamatsu A, et al. Drosophila FACT contributes to Hox gene expression through physical and functional interactions with GAGA factor. Genes Dev. 2003;17: 1605–1616.

121. Ngian Z-K, Lin W-Q, Ong C-T. NELF-A controls Drosophila healthspan by regulating heat-shock protein-mediated cellular protection and heterochromatin maintenance. Aging Cell. 2021;20: e13348.

122. Marr SK, Lis JT, Treisman JE, Marr MT 2nd. The metazoan-specific mediator subunit 26

(Med26) is essential for viability and is found at both active genes and pericentric heterochromatin in Drosophila melanogaster. Mol Cell Biol. 2014;34: 2710–2720.

123. Hughes CS, Foehr S, Garfield DA, Furlong EE, Steinmetz LM, Krijgsveld J. Ultrasensitive proteome analysis using paramagnetic bead technology. Mol Syst Biol. 2014;10: 757.

124. Skene PJ, Henikoff JG, Henikoff S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. Nat Protoc. 2018;13: 1006–1019.

125. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9: R137.

126. Ni J-Q, Zhou R, Czech B, Liu L-P, Holderbaum L, Yang-Zhou D, et al. A genome-scale shRNA resource for transgenic RNAi in Drosophila. Nat Methods. 2011;8: 405–407.

127. Kosman D, Mizutani CM, Lemons D, Cox WG, McGinnis W, Bier E. Multiplex detection of RNA expression in Drosophila embryos. Science. 2004;305: 846.

128. Chou TB, Perrimon N. Use of a yeast site-specific recombinase to produce female germline chimeras in Drosophila. Genetics. 1992;131: 643–653.

129. Chou TB, Perrimon N. The autosomal FLP-DFS technique for generating germline mosaics in Drosophila melanogaster. Genetics. 1996;144: 1673–1679.

130. Calvo L, Ronshaugen M, Pettini T. smiFISH and embryo segmentation for single-cell multi-gene RNA quantification in arthropods. Commun Biol. 2021;4: 352.

131. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29: 15–21.

132. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30: 923–930.

133. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26: 139–140.