Chromatin accessibility established by Pou5f3, Sox19b and Nanog primes genes for activity during zebrafish genome activation.

Máté Pálfy[1,4], Gunnar Schulze[2,4], Eivind Valen[2,3], Nadine L. Vastenhouw[1#]

1 Max Planck Institute of Molecular Cell Biology and Genetics, Dresden

2 Computational Biology Unit, Department of Informatics, University of Bergen

3 Sars International Centre for Marine Molecular Biology, University of Bergen

4 equal contribution, listed in alphabetical order

corresponding author and lead contact: vastenhouw@mpi-cbg.de

ABSTRACT

In many organisms, early embryonic development is driven by maternally provided factors until the controlled onset of transcription during zygotic genome activation. The regulation of chromatin accessibility and its relationship to gene activity during this transition remains poorly understood. Here, we generated chromatin accessibility maps from genome activation until the onset of lineage specification. During this period, chromatin accessibility increases at regulatory elements. This increase is independent of RNA polymerase II-mediated transcription, with the exception of the hyper-transcribed miR-430 locus. Instead, accessibility often precedes the transcription of associated genes. Loss of the maternal transcription factors Pou5f3, Sox19b, and Nanog, which are known to be required for zebrafish genome activation, results in decreased accessibility at regulatory elements. Importantly, the accessibility of regulatory regions, especially when established by Pou5f3, Sox19b and Nanog, is predictive for future transcription. Our results show that the maternally provided transcription factors Pou5f3, Sox19b, and Nanog open up chromatin and prime genes for activity during zygotic genome activation in zebrafish.

Keywords: ATAC-seq, zebrafish, zygotic genome activation, priming, promoters, enhancers, transcription factors

BACKGROUND

Chromatin structure is an important determinant of eukaryotic transcription regulation. At gene regulatory regions such as promoters and enhancers, chromatin needs to be in an accessible state to allow for binding of the transcriptional machinery. Consequently, the vast majority of transcription factor (TF) binding co-localizes with accessible chromatin regions [1,2]. Chromatin accessibility is regulated through the interplay of modifications to nucleosomes, such as posttranslational histone modifications and the incorporation of histone variants, as well as the binding of trans-acting factors [3].

While a number of studies suggest that chromatin accessibility generally correlates with gene activity [4–6], others have suggested that this correlation is, at least globally, rather weak. For example, chromatin accessibility landscapes are broadly similar in cell types and tissues with different expression signatures, and in embryonic cells exposed to different transcription factor concentrations [2,7,8]. This argues against the use of chromatin accessibility as a proxy for gene activity. In fact, open chromatin can also be detected at regulatory elements that become active only later during development. During hematopoiesis, for example, accessibility is more highly correlated with the 'poised' enhancer mark H3K4me1 than with the active H3K27ac mark [9]. Furthermore, *Drosophila*, mouse, and human embryos harbor accessible chromatin at regulatory elements prior to gene activation [4,10–13]. This raises the question whether chromatin accessibility can prime genes for activation.

The onset of transcription during zygotic genome activation (ZGA) provides an excellent system to address this question. In zebrafish, the bulk of zygotic transcription is initiated around three hours post fertilization (hpf) at the 1000-cell stage [14]. However, zygotic genome activation is a gradual process with different genes starting to be transcribed at different times [15–17]. It starts as early as the 64-cell stage, when the miR-430 gene cluster is activated [18–21], and ends when the cells in the embryo start to adopt different fates during gastrulation. Recently, Pou5f3, SoxB1 and Nanog have been identified as key transcription factors involved in activation of the zygotic genome; they bind thousands of putative regulatory elements, and loss of these TFs results in the reduced expression of many genes [15,22–24]. Furthermore, zebrafish Pou5f3 and Nanog have

been shown to play a role in promoting an open chromatin state at regulatory elements [25,26]. It remains unclear, however, whether Pou5f3, Sox19b and Nanog prime genes for activity.

Here we show that chromatin accessibility at promoters and enhancers often precedes transcriptional activity during zebrafish ZGA. The establishment of accessible regions requires the transcription factors Pou5f3, Sox19b and Nanog, and primes genes for future activation.

RESULTS

Chromatin accessibility at gene regulatory regions increases during genome activation

To study the dynamics of the chromatin landscape during zebrafish zygotic genome activation (ZGA), we generated genome-wide chromatin accessibility maps by ATAC-seq [27]. We analyzed seven stages from before the onset of transcription (256-cell), during genome activation (high, oblong, sphere, dome), until the cells in the embryo start to adopt different fates during gastrulation (shield, 80% epiboly) (Figure 1a). Biological replicates showed a large degree of correlation and samples from adjacent stages clustered together (Figure S1ab), demonstrating the high quality of our data. We called ATAC-seq peaks by MACS2 [28], using genomic DNA as an input control (Figure S2 and Methods) and used k-means to cluster these peaks. We identified five clusters with different chromatin accessibility dynamics (Figure 1b). The genomic regions in cluster 5 are only moderately accessible throughout the time series, while the accessibility of genomic regions in clusters 1 through 4 increases with different temporal dynamics. The regions in cluster 3 and 4 become accessible at different times during gastrulation. Cluster 2 regions gradually become accessible during ZGA, while regions in cluster 1 are already moderately accessible before ZGA (at 256-cell stage) and increase rapidly in accessibility in subsequent stages. Focusing on regulatory elements (Figure S3), we found that promoters (defined as +/- 1kb of the TSS) and putative enhancers (defined based on accessibility and binding of developmental TFs, see Methods) open with different temporal dynamics in clusters 1 through 4 (Figure 1cd). This dynamic opening of regulatory elements is also apparent over individual gene loci (Figure 1e). We conclude that during genome activation, chromatin accessibility dynamically increases at gene regulatory regions.

RNA polymerase II activity is generally not required for chromatin accessibility

The increase in chromatin accessibility during genome activation suggests that accessibility could be a consequence of transcriptional activity. To test this possibility, we inhibited transcription by injecting embryos with α-amanitin at the one-cell stage. α-amanitin treatment resulted in transcription inhibition and developmental arrest at sphere stage (Figure S4), as previously described [15,29,30]. We collected embryos at oblong stage, performed ATAC-seq, and compared chromatin accessibility in α-amanitin-treated and untreated embryos. This revealed that in general, chromatin accessibility at oblong stage is not a consequence of transcriptional activity (Figure 2a). Of all regions that are accessible in untreated embryos at oblong stage, only the miR-430 cluster shows a dramatic loss in accessibility upon transcription inhibition (Figure 2a). Interestingly, the expression level of miR-430 is orders of magnitude higher than that of any other gene during early zebrafish development [18]. This suggests that the transcription-dependent chromatin accessibility at the miR-430 cluster could be the result of RNA polymerase II complexes that evict nucleosomes, as has been demonstrated *in vitro* [31–33], and is reminiscent of the transcription-dependent accessibility at highly transcribed MERVL retrotransposons at the onset of ZGA in mouse embryos [4]. To determine whether transcribing RNA polymerase II contributes to chromatin accessibility at other highly transcribed genes as well, we grouped genes based on their expression level at oblong stage and associated putative enhancers with these genes using the nearest neighbor approach (see Methods). This revealed that chromatin accessibility at promoters and enhancers is not affected by transcription inhibition, with the exception of the miR-430 locus (Figure 2b). This is also visible in genome browser snapshots of chromatin accessibility at the miR-430, *id1*, and *fbxw4* loci in transcription-inhibited and untreated embryos (Figure 2c). We conclude that, apart from the miR-430 cluster, the chromatin accessibility landscape that is present during zebrafish genome activation does not depend on transcribing RNA polymerase II.

Chromatin accessibility at regulatory elements precedes transcriptional activity

The observation that the increase in chromatin accessibility at promoters and enhancers is independent of transcriptional activity, prompted us to investigate the temporal relationship between the opening of chromatin at regulatory elements and transcription onset. We selected genes that become transcriptionally active between sphere and shield stage [15] (“late” ZGA

genes, see Figure 1a) and related chromatin accessibility at promoters with the time of gene activation. We found that promoters are accessible several stages prior to the activation of transcription (Figure 3a). Moreover, chromatin accessibility increases during the stages preceding transcription activation (Figure 3a). As a control, we selected promoters of genes that are not expressed until two days post fertilization [16]. At these promoters, chromatin accessibility is extremely low and does not change from 256-cell to 80% epiboly. Thus, the changes in accessibility at promoters are not due to a general increase in accessibility. Next, we used the nearest neighbour approach to identify the enhancers associated with the late ZGA genes and compared their accessibility with the time of gene activation. We found that enhancers, like promoters, are accessible prior to gene activation (Figure 3b). As for promoters, accessibility of putative enhancers increases during the stages prior to gene activation. Here, we used zebrafish enhancers from adult tissues as a control [34]. At these enhancers, chromatin accessibility did not increase from 256-cell to 80% epiboly, confirming that the changes in accessibility at enhancers are not due to a general increase in accessibility. The accessibility of promoters and enhancers prior to gene activation, as well as the increase in accessibility prior to transcription activation, are easily visible at specific gene loci. For example, the promoter and putative enhancer of *gata6*, a gene that is activated at dome stage [16,35], are accessible at 256-cell stage and their accessibility increases prior to transcription (Figure 3c). A similar relationship between accessibility and transcription activation can be observed for the regulatory elements of *olig4* (activated at shield stage), and the experimentally validated enhancers of *neurog* (activated at 80% epiboly stage), *ta* (activated at sphere stage), and *otx2* (activated at 80% epiboly stage) [36–38] (Figure 3c, Figure S5). We conclude that promoters and enhancers are accessible prior to gene activation, and that their accessibility increases during the stages preceding transcription activation.

Our results so far raise two important questions. First, which factors are involved in opening up chromatin at promoters and enhancers? And second, does accessibility predict future transcription?

Pou5f3, Sox19b, and Nanog regulate chromatin accessibility

To address the first question, we set out to investigate the role of Pou5f3, Sox19b and Nanog in opening up chromatin. These transcription factors are maternally loaded, and have been shown to

be required for the activation of many zygotic genes [15,22,24]. Moreover, their motifs are enriched in the distal regions that increase in accessibility between 256-cell and oblong stage (Figure S6). We created *pou5f3* and *sox19b* mutants using CRISPR-Cas9 [39] (Figure S7). We excised the entire coding region of the *pou5f3* and *sox19b* genes, thereby preventing any potential genetic compensation that could be triggered by mutant mRNA [40,41], or the generation of unexpected transcripts that escape nonsense-mediated decay [42]. Maternal-zygotic (MZ) *pou5f3* embryos arrest in their development during late gastrulation, as previously described [43–45]. In contrast, MZ*sox19b* embryos develop normally until adulthood. The lack of a strong MZ*sox19b* phenotype can be explained by the zygotic expression of *sox19a*, *sox3* and *sox2*, which all belong to the SoxB1 family and can compensate for each other's function when expressed [46]. For our studies this redundancy does not cause a problem, because at the early stages of development only Sox19b is present at high levels [15,46]. For *nanog*, we obtained a recently published null mutant that harbors a 10 base pair deletion in the first exon, causing a frameshift [47]. Importantly, the *nanog* mutant phenotype has previously been shown to be identical to embryos in which Nanog is knocked down by translation-blocking morpholinos [24,47], suggesting that no mechanisms are involved in compensating for this mutation. Having obtained bonafide mutants for *pou5f3*, *sox19b* and *nanog*, we performed ATAC-seq on MZ*pou5f3,* MZ*sox19b* and MZ*nanog* embryos at oblong stage and compared chromatin accessibility in mutants and wild-type embryos. Upon the loss of Pou5f3, we observed a widespread decrease in chromatin accessibility (Figure 4a). Similarly, the loss of Sox19b and Nanog resulted in a strong decrease in accessibility at hundreds of genomic regions (Figure 4a). When analyzing changes in accessibility at regulatory elements, we found that promoters and enhancers are strongly affected (Figure 4b), as is apparent at genomic loci at Pou5f3-, Sox19b- and Nanog-regulated genes (Figure 4c). In a previous study, analysis of the effect of transcription factors on chromatin accessibility was limited to sites that are bound by Pou5f3, Sox19b and Nanog [26]. As a consequence, the effects on promoter accessibility were not addressed. To determine whether the effect of Pou5f3 on promoter accessibility that we observed (Figure 4c) could also be observed in the dataset that was previously generated [26], we reanalyzed the MNase-seq data generated in this study. This revealed that distal as well as proximal regions that decrease in accessibility in our MZ*pou5f3* mutant, show increased nucleosome occupancy in the *pou5f3* mutant (MZ*spg*) at dome stage (Figure S8). We conclude that Pou5f3, Sox19b, and

Nanog are important for the opening of chromatin at regulatory elements during zygotic genome activation.

It was recently suggested that Pou5f3 and Nanog affect chromatin accessibility only at sites that are bound by a combination of Pou5f3, SoxB1, and Nanog [26]. To determine whether the effect of Pou5f3, Sox19b, and Nanog is indeed limited to these so called PSN sites, we compared chromatin accessibility in the transcription factor mutants at sites bound by one, a combination of two, or all three transcription factors. Consistent with the previous study, accessibility is affected significantly at PSN sites (Figure 4D). In addition, however, sites bound by a single TF, or a combination of two of the TFs, also show a significant decrease in accessibility in the MZ*pou5f3*, MZ*sox19b,* and MZ*nanog* mutants. Together, these data suggest that Pou5f3, Sox19b, and Nanog regulate chromatin accessibility at promoters and enhancers that are bound by one or more of these factors.

Chromatin accessibility established by Pou5f3, Sox19b and Nanog predicts future transcription

Next, we addressed the question whether an increase in accessibility has predictive value for future transcription. To that end, we selected promoter regions that show an increase in ATAC-seq signal between 256-cell and oblong stage. We binned these regions into 20% quintiles, and analyzed the median expression intensity for associated genes at sphere and shield stages. This analysis showed that larger increases in chromatin accessibility between 256-cell and oblong stage are associated with higher expression levels at sphere stage (Figure 5a). To assess whether this correlation is also valid for distal elements, we linked ZGA genes with putative enhancers and observed that, as for promoters, the mean expression intensity is higher for genes that are linked to enhancers that show larger changes in accessibility (Figure 5b). Similar results were obtained at shield stage (Figure S9ab). We conclude that an increase in chromatin accessibility at regulatory regions is a predictor of gene activity.

The observed role of Pou5f3, Sox19b and Nanog in regulating chromatin accessibility in early embryos suggests that these transcription factors might be involved in priming. To test this possibility, we extended our analysis of the relationship between accessibility and future gene

expression (Figure 5ab, Figure S9ab) with the effect of transcription factor loss. To this end, we resolved gene expression levels at sphere stage by the fold change in chromatin accessibility at promoters and enhancers between 256-cell stage and oblong, as well as by the average fold change in accessibility in Pou5f3, Sox19b, and Nanog mutants compared to wild-type embryos (Figure 5c). This revealed that regulatory elements that show both the largest increase in accessibility between 256-cell and oblong stage, and the largest reduction in accessibility upon the individual loss of either Pou5f3, Sox19b or Nanog, have the highest expression level at sphere stage (Figure 5c). Because of the observed correlation between zygotic transcription and H3K27ac levels on regulatory elements during zebrafish ZGA [19,48], we repeated our analysis using H3K27Ac as a proxy for transcriptional activity. This independent analysis confirmed the important role of Pou5f3, Sox19b and Nanog in priming gene expression (Figure 5d). Here too, analysis of shield stage embryos resulted in the same conclusion (Figure S9c). Together, these data show that Pou5f3, Sox19b and Nanog open up gene regulatory regions to facilitate gene activation.

Finally, if Pou5f3, Sox19b, and Nanog prime genes for activation it would be predicted that they enable the binding of other transcription factors. To test this hypothesis, we assessed whether regions where Pou5f3, Sox19b or Nanog open up chromatin are enriched for the binding of transcription factors involved in mesendoderm specification. This analysis showed that regulatory elements which experience the greatest loss in accessibility in MZ*pou5f3*, MZ*sox19b* or MZ*nanog* mutants are the most highly enriched for the binding of the transcription factors Mxtx2, Eomesa, Smad2 and Foxh1 at blastula stages (Figure 5e). Together, our results show that Pou5f3, Sox19b and Nanog prime genes for activation by opening up chromatin, which facilitates the binding of sequence-specific transcription factors.

DISCUSSION

In this study, we analyzed the dynamics and regulation of the chromatin accessibility landscape that underlies transcription activation during embryo development, using zebrafish ZGA as a model system. We find that promoters and enhancers increase in accessibility during genome activation. This is not due to transcriptional elongation by RNA polymerase II. In fact, accessibility at regulatory elements increases prior to gene activity. The increase in accessibility

primes genes for activation, and Pou5f3, Sox19b and Nanog play a major role therein. Our results provide important insights into the relationship between chromatin accessibility and gene activity, and the molecular factors that shape the chromatin landscape during zebrafish genome activation.

Zygotic genome activation is accompanied by a dynamic increase in chromatin accessibility at regulatory elements

Our finding that chromatin accessibility increases during genome activation, especially at promoter and enhancer elements, is consistent with recent studies in *Drosophila*, zebrafish, *Xenopus*, mouse and human embryos [4,10–13,25,49]. An increase in the number and strength of ATAC-seq peaks over time may either be due to an increase in the number of cells in which specific regions are accessible, or an actual increase in the accessibility of specific regions in all cells of the embryo. Single cell ATAC-seq in *Drosophila* embryos has shown that at the onset of genome activation, cells are largely homogeneous in terms of chromatin accessibility [50]. Although such analysis has not yet been done in zebrafish embryos, individual zebrafish cells are transcriptionally largely uniform during early developmental stages [51,52], suggesting that the accessibility landscape may be largely uniform as well. Furthermore, the maternally loaded transcription factors Pou5f3, Sox19b and Nanog that play an important role in opening chromatin are uniformly expressed in the embryo at the time of ZGA [23,46,53]. Together, this favors a scenario in which the increase in chromatin accessibility during zebrafish genome activation is caused by an embryo-wide increase in accessibility at individual loci.

Transcribing RNA polymerase II evicts nucleosomes only at the hyper-transcribed miR-430 locus

Our observation that chromatin accessibility is independent of transcriptional activity is consistent with studies that failed to detect a causative link between the transcription process and promoter chromatin accessibility at a genome-wide level [54,55]. However, *in vitro* and modeling studies have proposed that transcription by RNA polymerase II can form complexes which are able to evict nucleosomes and increase chromatin accessibility [31–33]. Our data suggest that such a mechanism may only take place at genomic regions that are very highly expressed. We find that the mir-430 cluster, which is transcribed two orders of magnitude higher

than any other region, undergoes a dramatic loss in accessibility when elongation of RNA polymerase II is blocked. The miR-430 cluster is one of the first to be transcribed and initially, all transcribing RNA polymerase II localizes to the two alleles of the miR-430 cluster [19–21,56], which might facilitate the formation of RNA-pol II 'trains' capable of evicting nucleosomes, as shown *in vitro* [33].

Chromatin accessibility precedes transcriptional activity

Our finding that regulatory elements can be accessible prior to transcription of associated genes is in agreement with the establishment of other "active" chromatin features prior to gene expression. At zebrafish promoters, for example, the histone modifications H3K4me3 [57,58] and H3K27ac [56,59], a hypomethylated state ([25,60–62], and well-positioned nucleosomes [63] can be established prior to the time their associated genes become expressed. Similar observations have been made in other species. In *Drosophila* embryos, for example, 47% of enhancers are accessible prior to large-scale genome activation [10], and lineage-specific chromatin regulatory landscapes are established prior to gastrulation [50]. Furthermore, histones can be acetylated prior to the major wave of ZGA at sites of early Zelda binding [64,65]. Similarly, a subset of promoters in human and mouse embryos exhibit open chromatin prior to activation of the genome [4,11,12]. We conclude that the establishment of accessible chromatin prior to gene activity is a common theme during early development.

Pou5f3, Sox19b and Nanog establish chromatin accessibility

We found that zebrafish Pou5f3, Sox19b and Nanog are important for the increase in chromatin accessibility during ZGA. Ever since the discovery that human Oct4 and Sox2 act as pioneer factors during cell reprogramming [66,67], and the finding that their homologues are involved in zebrafish zygotic genome activation [15,22], it has been hypothesized that Pou5f3, Sox19b and Nanog could play a role in remodeling chromatin during genome activation. We now provide direct evidence for this hypothesis. Our results are in agreement with two recent studies reporting that the genetic loss of *pou5f3* and *nanog,* or their knockdown by translation-blocking morpholinos, influences the accessibility of chromatin at sites bound by these factors [25,26]. In contrast to the previously reported preference of Pou5f3 and Nanog to open up distal regions that are DNA methylated [25], we find that Pou5f3, Sox19b and Nanog are also important for

regulating chromatin accessibility at promoters of zygotic genes, which are generally hypomethylated. This might explain why a third study [26] found no preference for Pou5f3 and Nanog affecting accessibility at hypermethylated sites. Our finding that the loss of Pou5f3, Sox19b and Nanog has the largest effect on accessibility at sites bound by all three factors, is consistent with a previous study [26]. However, while this study reported that Pou5f3 and Nanog exclusively remodel chromatin accessibility at PSN sites, we find that they also regulate accessibility at regions that are bound by only one or two factors.

We note that our analysis of motifs that are enriched in regions that show an increase in chromatin accessibility during early ZGA suggests that in addition to Pou5f3, Sox19b, and Nanog, there are other transcription factors that could play a role in priming genes for activity. One such candidate factor is NFYA. It has previously been shown that NFYA binds to DNA in combination with TALE factors during zebrafish blastula stages [68], promotes chromatin accessibility and enhances the binding of Oct4, Sox2 and Nanog in embryonic stem cells [69], and establishes open chromatin during mouse ZGA [11]. Future studies will reveal whether NFYA cooperates with Pou5f3, Sox19b and Nanog in transcriptional priming during zebrafish ZGA.

Pou5f3, Sox19b and Nanog prime genes for activation

We found that an increase in chromatin accessibility at regulatory elements primes genes for future activity. Although changes in chromatin accessibility have been correlated with transcription levels before [10,70,71], it was not clear whether chromatin accessibility at regulatory regions primes genes for activation. Here, we show directly that the increase in accessibility during the earliest stages of zygotic genome activation has a strong predictive value for future transcription. This is in contrast with the absolute level of chromatin accessibility prior to transcription, which does not correlate well with transcription at later stages of zebrafish development [59,63].

Priming is strongest at sites where accessibility depends on Pou5f3, Sox19b and Nanog. Our data suggests that by opening up chromatin at regulatory elements, Pou5f3, Sox19b and Nanog facilitate the binding of other transcription factors involved in lineage specification and

patterning. This is in agreement with recent work showing that Pou5f3 and Sox3 remodel chromatin during *Xenopus* ZGA to predefine competence for germ layer formation [49]. Future studies will be required to explore in more detail how Pou5f3, Sox19b and Nanog prime genes for activity. One potential mechanism is that these transcription factors mediate the establishment of H3K27ac at regulatory elements [56,59]. PSN as well as H3K27ac are essential for zygotic genome activation [15,19,59], and the highest H3K27ac enrichment occurs at PSN-primed loci (this study). The interdependence of these factors in the context of priming, however, remains unclear. Another potential mechanism involves the recruitment of chromatin remodelers to establish accessibility at regulatory elements. Recently, it was suggested that BRG1 could be involved in Pou5f3-mediated opening of chromatin in zebrafish [25], similar to what has been reported in ES cells [72]. However, it is likely that Pou5f3-mediated chromatin remodeling in zebrafish mechanistically differs from other species, given that zebrafish Pou5f3 lacks the linker region between the POU domains, which appears to be important for the interaction with Brg1 [73] and its capability for reprogramming [73,74].

CONCLUSIONS

During zebrafish zygotic genome activation, promoters and enhancers are often accessible before their associated genes are transcribed. The increase in accessibility at regulatory elements primes genes for transcription. This priming is strongest where the increase in accessibility is regulated by the transcription factors Pou5f3, Sox19b and Nanog.

Acknowledgements

We thank Daria Onichtchouk for kindly providing zebrafish *nanog* -/- mutant fish, Stefan Hans for advice on CRISPR experiments and reagents, members of the Vastenhouw and Valen labs for discussions, Carine Stapel for manuscript comments, and the following facilities and services: Fish Facility (MPI-CBG) and the DRESDEN-concept Genome Center.

Funding

This work was supported by the Max Planck Society, a Human Frontier Science Program Career Development Award (to NLV, CDA00060/2012), the Bergen Research Foundation (to EV) and the Norwegian Research Council (to EV, #250049).

Author's contributions

MP conceived the project, performed experiments, interpreted the data, and wrote the manuscript. GS performed bioinformatics analysis and interpreted the data. EV obtained funding and supervised bioinformatics analysis. NLV conceived the project, obtained funding, supervised the project, and wrote the manuscript. All authors read the manuscript, edited the manuscript and approved the final manuscript.

Ethics approval and consent to participate

All animal experiments were conducted according to the German Animal Welfare act and the European Communities Council Directive of 22 September 2010, and approved and licensed by the local ethics committee (Landesdirektion Sachsen, Germany).

METHODS

Zebrafish husbandry and manipulation

Zebrafish (TLAB) were maintained at 28.5°C according to standard protocols. Embryos were dechorionated with pronase and were grown in Danieau's medium on agarose dishes until the desired stage. For transcription inhibition, α-amanitin (A2263; Sigma) was injected at the 1-cell stage at a concentration of 0.2 ng per embryo. Staging of embryos was according to [77].

Generation of MZ*pou5f3*, MZ*sox19b* and MZ*nanog* embryos

MZ*pou5f3* and MZ*sox19b* embryos were generated by CRISPR-Cas9. Two gRNAs were designed to target sequences close to the N-and C-terminal ends of both the *pou5f3* and *sox19b* genes using the CHOP-CHOP tool [78,79]. The following gRNA-s were used to delete the *sox19b* and *pou5f3* genes (the **NGG** PAM sequence indicated in bold):

N-terminal *sox19b*: GGGTTGCGTTGAGCGCTCAT**TGG**

C-terminal *sox19b*: GGTGTCCAACAGCACTATCT**TGG**

N-terminal *pou5f3*: GGCCATGTATCCTCAAGCCG**CGG**

C-terminal *pou5f3*: GGCTTGCACCCTGGTTTGGT**GGG**

gRNAs were cloned into the T7cas9gRNA2 vector [39] and in vitro transcribed using the MEGAshortscript T7 kit (Ambion/Invitrogen), followed by purification with the mirVana miRNA Isolation kit (Ambion/Invitrogen). Embryos at the early 1-cell stage were injected with 150 pg of Cas9 mRNA and 30 pg of each gRNA. Cas9 mRNA for the injections was produced by in vitro transcription of the p3T3S-nls-zCas9-nls plasmid [39] using the T3 mMessage machine transcription kit (Ambion/Invitrogen).

Embryos injected with Cas9 mRNA and gRNAs targeting *sox19b* and *pou5f3* were screened for mutant founders by outcrossing to wild-type fish. Maternal-zygotic mutants for *sox19b* were produced by incrossing homozygous mutant fish. Homozygous *pou5f3* mutant fish were rescued with injection of 40pg of Pou5f3-2xHA mRNA [30], and MZ*pou5f3* embryos were obtained by incrossing homozygous *pou5f3* mutant fish.

To obtain MZ*nanog* embryos, homozygous zygotic *nanog* mutants from [47] were incrossed.

ATAC-seq

The ATAC-seq protocol was adapted from [80] and performed on approximately 70,000 cells. Zebrafish embryos were first deyolked using buffers as described in [81]; briefly, embryos were dissociated in deyolking buffer by brief vortexing and shaking at 1100 rpm at 4°C. Cells were collected by centrifugation at 400 rcf and the supernatant was removed. The cells were washed two more times with deyolking wash buffer. For the 256-cell stage, 300 embryos were manually deyolked as we were not able to obtain sufficient number of cells by buffer deyolking.

Following dissociation, cells were lysed in cell lysis buffer. Nuclei were spun down at 500 rcf for 10 minutes at 4°C, resuspended in tagmentation mix and incubated at 37°C for 30 minutes with gentle shaking at 800 rpm. Tagmented DNA was cleaned up using the Qiagen Minelute Reaction Cleanup kit and DNA was eluted in 10 ul of EB. DNA libraries were made by a limited cycle PCR Using the NEBNext High-fidelity 2X PCR Master Mix (NEB). To maximize the complexity of DNA libraries, the number of PCR cycles were determined by qPCR as described in [80]. Following 11-13 cycles of PCR, the libraries were purified using the Minelute Reaction cleanup kit (Qiagen) and measured with a fragment analyzer. Samples were sequenced following adaptor cleanup on an Illumina HiSeq25000 instrument with 75bp reads.

We performed two or three biological replicates for wild-type samples from different stages (: Figure S1), two biological replicates for amanitin-treated embryos (Pearson correlation=0.95), two biological replicates for MZ*pou5f3* embryos (Pearson correlation=0.96), two biological replicates for MZ*sox19b* embryos (Pearson correlation= 0.88), and two biological replicates for MZ*nanog* embryos (Pearson correlation=0.96).

RT-qPCR

25 embryos per stage were flash-frozen in liquid nitrogen. RNA was extracted using the RNeasy Mini kit from Qiagen. 1microgram of RNA was reverse transcribed to produce cDNA using the iScript cDNA Synthesis kit (Bio-Rad). For qPCR, a readily made SYBR green Mastermix with ROX (100nM) was used. Primers for *fam212aa*, *sox19a*, *mxtx2* and the housekeeping gene *eif4g2a* were from [30].

ChIP-seq processing

We obtained ChIP-sequencing datasets for transcription factors from GEO studies GSE84619 (Ta, Tbx16 and Mixl1), GSE34683 (Nanog, Mxtx2), GSE51894 (Smad2, Eomesodermin), GSE67648 (Foxh1), GSE39780 (Pou5f1, Sox2). In addition we downloaded histone ChIP-seq for H3K27ac and H3K4me3 from GSE32483 for dome stage. Fastq-files were adapter and quality trimmed using cutadapt and aligned to the zebrafish genome build Zv10 using bowtie2 with parameters --no-mixed --no-discordant (paired-end datasets only). We deduplicated and filtered alignments by quality (q >= 30) using samtools. We called peaks using MACS2 at 5% FDR using ChIP-Input libraries as background (-c flag), if available. For histone ChIP-seq, we called peaks using the --broad flag.

ATAC-seq processing

ATAC-seq paired-end reads were adapter- and quality trimmed using cutadapt and aligned to zebrafish genome build Zv10 using bowtie2 with a maximum insert size of 2000 (-X 2000) and parameters '--no-mixed --no-discordant'. Alignments were further deduplicated using 'samtools rmdup'. Only properly paired alignments with alignment quality >= 30 were used in downstream analyses. Additionally, alignments to the mitochondrial genome were filtered out.

Alignment files were converted to BED format using the 'bedtools bamtobed' utility in combination with a custom awk script. Start positions of reads mapping to the plus strand were adjusted by +4bp and end positions of reads mapping to the minus strand were adjusted -5bp to account for the transposase cutting event. Adjusted alignments were further subdivided into nucleosomal and open chromatin fragments corresponding to nucleosome-free regions (NFR) by taking a fragment size cut-off of 130bp (determined by visual inspection of the fragment length distributions).

ATAC-seq peak calling

ATAC peaks were called for both all fragments and the NFR fraction only with the program MACS2 at an FDR of 5% with additional parameters "--no-model --no-lambda" and using the control digestion of genomic DNA as background (-c parameter). For each stage/condition we called peaks based on pooled replicates and subsequently overlapped the resulting peaks with those called in individual replicates, retaining only peaks that were confirmed in all replicates. We then merged all NFR peaks called in wild-type conditions into a consensus set (min 1bp overlap between peaks). We called a peak from the consensus set "present" in a particular stage if it overlapped by a replicated peak identified in that stage.

Differential accessibility analysis

For differential analysis, we counted ATAC-seq reads over the consensus peak set using the featureCounts function from the Rsubread package. To enrich for fragments associated with nucleosome free regions, only fragments up to a length of 130bp length were counted. Alignments from multimapping reads were additionally discarded for this analysis. Differential analysis was performed using DEseq2, setting the size factor for each replicate to the total fragment count in that replicate. We considered peaks with a log2 fold change of -1.5 or 1.5 at an FDR of 5% as significantly decreasing or increasing, respectively.

Definition of cis-regulatory elements and putative enhancers

We extracted transcriptions start sites (TSS) for all genes based on Danio rerio Zv10 genome assembly and ensembl v85 gene build. Promoters were defined as 2kb windows centered on the

TSS. We determined putative enhancer elements as distal (min 1kb distance from any TSS) accessible regions from the consensus peak set that overlap at least one ChIP-seq peak of either Sox2[22], Nanog[23] or Pou5f1[22] and one additional developmental transcription factor (Smad2 [75], FoxH1[76], Mxtx2[23], Eomesodermin[75], Ta, Tbx16 or Mixl1[82]).

Clustering of accessible regions

We calculated the average open chromatin fragments (fragment length <= 130bp) per kilobase per million reads mapped (FPKM) for each consensus peak in each stage and clustered the matrix of accessibility values by k-means clustering ('kMeans' function in R). We set the number of clusters k to 5 and initialized 10 random starts with a maximum of 1000 iterations. This resulted in 60% of the total variance explained by differences between clusters. All other parameters were kept at their default. The resulting clustering was then visualized in a heatmap. For this purpose, the matrix values were further binned into 100 quantiles ranging from low (0) to high (100) accessibility.

Accessibility heatmaps

For each ATAC-seq library, we generated coverage tracks in bigwig format using the 'deeptools bamCoverage' utility at a bin size of 20bp. We extended alignments to cover the full fragment width (-e) and normalized to the individual library sizes using RPKM normalization. To separately analyze the open chromatin fraction, we additionally generated tracks using only fragments smaller or equal to 130bp in length. We plotted the open chromatin coverage signal over promoters and enhancers and at transcription factor binding sites, using the 'deeptools plotHeatmap' utility. For promoters and transcription factors, we centered the heatmap on the TSSs and ChIP-seq peak summits, respectively. Enhancer regions were scaled to each represent a 1000bp window. We used the 'deeptools bigwigCompare' utility to generate log2 ratio tracks of MZpou5f3 ATAC-seq and MZspg MNase-seq (external) samples over their wild-type counterparts at oblong and dome stage, respectively.

Motif analysis

Motif analysis was performed using the MEME suite v5.03 [83]. To include the zebrafish Nanog-like motif in our analyses, we used a 60bp window centered on the top 5000 peaks called

from 3.5hpf Nanog ChIP-seq data [23] as input to MEME (-minw 6 -maxw 15 -nmotifs 10). We identified the consensus motif corresponding to the previously reported Nanog-like motif by visual inspection and added the corresponding position weight matrix to a local copy of the JASPAR 2018 core vertebrate collection of motifs. This custom collection was then used in motif enrichment and comparative analyses.

Proximal (+/-1kb around TSS) and distal (enhancer and intergenic) ATAC-seq peaks were sorted by their relative increase in accessibility between 256-cell and oblong stage and the top 20% of peaks within each category were selected for motif analysis. Enrichment analysis against the modified JASPAR database (containing Nanog, see above) was performed using AME (MEME-suite) using non-standard parameters *--kmer 2 --evalue-report-threshold 10*. A custom R script was used in connection with R/TFBStools to summarize enriched factors at the motif family level. For each TF-family we reported the minimal adjusted p-value as well as the maximal % of peaks in the input set that have at least 1 motif from this family. *De novo* motif analysis was performed using the MEME-CHIP utility with parameters *-order 1 -meme-p 4 -ccut 0 -meme-mod zoops -meme-minw 6 -meme-maxw 15 -meme-nmotifs 10.*

We selected the top *de novo* motif for each TF family detected by matching motifs to their closest known counterparts in the database (TOMTOM). A custom R-script was used in connection with the R/ggseqlogo package to extract and plot position weight matrices for these motifs.

Gene expression sets and transcriptional priming

We obtained gene expression data from [15] and ([16,23] and sorted genes based on their zygotic expression strength in sphere and shield stage, respectively. We then plotted the distribution of normalized ATAC-seq signal over promoters and enhancers of sphere and shield activated genes. Genes with a zygotic expression RPKM>=1.0 [15] where considered expressed in sphere or shield stage, respectively. Shield-expressed genes were additionally filtered to only contain genes not previously expressed in sphere. Enhancers were assigned to promoters by proximity within a maximal distance of 20kb. We summed the expression level of genes reported in [16]

until 8hpf and defined a cut-off of RPKM < 0.1 to select genes not expressed until shield stage as a control group. For enhancers we used late-activated enhancers [34] as a control.

REFERENCES

1. 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.

2. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. Nature. 2012;489:75–82.

3. Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. Nat Rev Genet. 2019;20:207–20.

4. Wu J, Huang B, Chen H, Yin Q, Liu Y, Xiang Y, et al. The landscape of accessible chromatin in mammalian preimplantation embryos. Nature. 2016;534:652–7.

5. Clark SJ, Argelaguet R, Kapourani C-A, Stubbs TM, Lee HJ, Alda-Catalinas C, et al. scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. Nature Commun. 2018; 9:781.

6. Liu L, Leng L, Liu C, Lu C, Yuan Y, Wu L, et al. An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. Nat Commun. 2019;10:364.

7. McKay DJ, Lieb JD. A common set of DNA regulatory elements shapes Drosophila appendages. Dev Cell. 2013;27:306–18.

8. Haines JE, Eisen MB. Patterns of chromatin accessibility along the anterior-posterior axis in the early Drosophila embryo. PLoS Genet. 2018;14:e1007367.

9. Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, Zaretsky I, Jaitin DA, David E, et al. Immunogenetics. Chromatin state dynamics during blood formation. Science. 2014;345:943–9.

10. Blythe SA, Wieschaus EF. Establishment and maintenance of heritable chromatin structure

during early Drosophila embryogenesis. Elife. 2016;5.

11. Lu F, Liu Y, Inoue A, Suzuki T, Zhao K, Zhang Y. Establishing Chromatin Regulatory Landscape during Mouse Preimplantation Development. Cell. 2016;165:1375–88.

12. Wu J, Xu J, Liu B, Yao G, Wang P, Lin Z, et al. Chromatin analysis in human early development reveals epigenetic transition during ZGA. Nature. 2018;557:256–60.

13. Li L, Guo F, Gao Y, Ren Y, Yuan P, Yan L, et al. Single-cell multi-omics sequencing of human early embryos. Nat Cell Biol. 2018;20:847–58.

14. Kane DA, Kimmel CB. The zebrafish midblastula transition. Development. 1993;119:447–56.

15. Lee MT, Bonneau AR, Takacs CM, Bazzini AA, DiVito KR, Fleming ES, et al. Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. Nature. 2013;503:360–4.

16. White RJ, Collins JE, Sealy IM, Wali N, Dooley CM, Digby Z, et al. A high-resolution mRNA expression time course of embryonic development in zebrafish. Elife 2017;6.

17. Pálfy M, Joseph SR, Vastenhouw NL. The timing of zygotic genome activation. Curr Opin Genet Dev. 2017;43:53–60.

18. Heyn P, Kircher M, Dahl A, Kelso J, Tomancak P, Kalinka AT, et al. The earliest transcribed zygotic genes are short, newly evolved, and different across species. Cell Rep. 2014;6:285–92.

19. Chan SH, Tang Y, Miao L, Darwich-Codore H, Vejnar CE, Beaudoin J-D, et al. Brd4 and P300 regulate zygotic genome activation through histone acetylation. bioRxiv. 2018

20. Hilbert L, Sato Y, Kimura H, Jülicher F, Honigmann A, Zaburdaev V, et al. Transcription organizes euchromatin similar to an active microemulsion. bioRxiv. 2018

21. Hadzhiev Y, Qureshi HK, Wheatley L, Cooper L, Jasiulewicz A, Van Nguyen H, et al. A cell cycle-coordinated Polymerase II transcription compartment encompasses gene expression before global genome activation. Nat Commun. 2019;10:691.

22. Leichsenring M, Maes J, Mössner R, Driever W, Onichtchouk D. Pou5f1 transcription factor controls zygotic gene activation in vertebrates. Science. 2013;341:1005–9.

23. Xu C, Fan ZP, Müller P, Fogley R, DiBiase A, Trompouki E, et al. Nanog-like Regulates Endoderm Formation through the Mxtx2-Nodal Pathway. Developmental Cell. 2012. p. 625–38.

24. Gagnon JA, Obbad K, Schier AF. The primary role of zebrafish nanog is in extra-embryonic tissue. Development. 2018;145 (1).

25. Liu G, Wang W, Hu S, Wang X, Zhang Y. Inherited DNA methylation primes the establishment of accessible chromatin during genome activation. Genome Res. 2018;28:998–1007.

26. 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;29:383–95.

27. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10:1213–8.

28. 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.

29. Kane DA, Hammerschmidt M, Mullins MC, Maischein HM, Brand M, van Eeden FJ, et al. The zebrafish epiboly mutants. Development. 1996;123:47–55.

30. Joseph SR, Pálfy M, Hilbert L, Kumar M, Karschau J, Zaburdaev V, et al. Competition between histone and transcription factor binding regulates the onset of transcription in zebrafish embryos. Elife. 2017;6.

31. Kireeva ML, Walter W, Tchernajenko V, Bondarenko V, Kashlev M, Studitsky VM. Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. Mol Cell. 2002;9:541–52.

32. Kulaeva OI, Hsieh F-K, Studitsky VM. RNA polymerase complexes cooperate to relieve the nucleosomal barrier and evict histones. Proc Natl Acad Sci U S A. 2010;107:11325–30.

33. van den Berg AA, Depken M. Crowding-induced transcriptional bursts dictate polymerase and nucleosome density profiles along genes. Nucleic Acids Res. 2017;45:7623–32.

34. Pérez-Rico YA, Boeva V, Mallory AC, Bitetti A, Majello S, Barillot E, et al. Comparative analyses of super-enhancers reveal conserved elements in vertebrate genomes. Genome Res. 2017;27:259–68.

35. Pauli A, Valen E, Lin MF, Garber M, Vastenhouw NL, Levin JZ, et al. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. Genome Res. 2012;22:577–91.

36. Blader P, Lam CS, Rastegar S, Scardigli R, Nicod J-C, Simplicio N, et al. Conserved and acquired features of neurogenin1 regulation. Development. 2004;131:5627–37.

37. Harvey SA, Tümpel S, Dubrulle J, Schier AF, Smith JC. no tail integrates two modes of mesoderm induction. Development. 2010;137:1127–35.

38. Kurokawa D, Sakurai Y, Inoue A, Nakayama R, Takasaki N, Suda Y, et al. Evolutionary constraint on Otx2 neuroectoderm enhancers-deep conservation from skate to mouse and unique divergence in teleost. Proceedings of the National Academy of Sciences. 2006. p. 19350–5.

39. Jao L-E, Wente SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc Natl Acad Sci U S A. 2013;110:13904–9.

40. Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, et al. Genetic compensation induced by deleterious mutations but not gene knockdowns. 2015;524:230–3.

41. El-Brolosy MA, Kontarakis Z, Rossi A, Kuenne C, Günther S, Fukuda N, et al. Genetic compensation triggered by mutant mRNA degradation. Nature. 2019;568:193–7.

42. Anderson JL, Mulligan TS, Shen M-C, Wang H, Scahill CM, Tan FJ, et al. mRNA processing in mutant zebrafish lines generated by chemical and CRISPR-mediated mutagenesis

produces unexpected transcripts that escape nonsense-mediated decay. PLoS Genet. 2017;13:e1007105.

43. Reim G, Brand M. Maternal control of vertebrate dorsoventral axis formation and epiboly by the POU domain protein Spg/Pou2/Oct4. Development. 2006;133:2757–70.

44. Reim G, Mizoguchi T, Stainier DY, Kikuchi Y, Brand M. The POU domain protein spg (pou2/Oct4) is essential for endoderm formation in cooperation with the HMG domain protein casanova. Dev Cell. 2004;6:91–101.

45. Lunde K, Belting H-G, Driever W. Zebrafish pou5f1/pou2, homolog of mammalian Oct4, functions in the endoderm specification cascade. Curr Biol. 2004;14:48–55.

46. Okuda Y, Ogura E, Kondoh H, Kamachi Y. B1 SOX coordinate cell specification with patterning and morphogenesis in the early zebrafish embryo. PLoS Genet. 2010;6:e1000936.

47. Veil M, Schaechtle MA, Gao M, Kirner V, Buryanova L, Grethen R, et al. Maternal Nanog is required for zebrafish embryo architecture and for cell viability during gastrulation. Development. 2018;145 (1)

48. Bogdanović O, Fernandez-Miñán A, Tena JJ. Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. Genome Res. 2012 10:2043-53

49. Gentsch GE, Spruce T, Owens NDL, Smith JC. The role of maternal pioneer factors in predefining first zygotic responses to inductive signals. bioRxiv. 2018

50. Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferreres R, et al. The cis-regulatory dynamics of embryonic development at single-cell resolution. Nature. 2018;555:538–42.

51. Farrell JA, Wang Y, Riesenfeld SJ, Shekhar K, Regev A, Schier AF. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. Science. 2018;360.

52. Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, Klein AM. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. Science. 2018;360:981–7.

53. Lippok B, Song S, Driever W. Pou5f1 protein expression and posttranslational modification during early zebrafish development. Dev Dyn. 2014;243:468–77.

54. Fenouil R, Cauchy P, Koch F, Descostes N, Cabeza JZ, Innocenti C, et al. CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters. Genome Res. 2012;22:2399–408.

55. King HW, Fursova NA, Blackledge NP, Klose RJ. Polycomb repressive complex 1 shapes the nucleosome landscape but not accessibility at target genes. Genome Res. 2018;10:1494-1507

56. Sato Y, Hilbert L, Oda H, Wan Y, Heddleston JM, Chew T-L, et al. Quantitative measurements of chromatin modification dynamics during zygotic genome activation. bioRxiv. 2019.

57. Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, Liu XS, et al. Chromatin signature of embryonic pluripotency is established during genome activation. Nature. 2010;464:922–6.

58. Lindeman LC, Andersen IS, Reiner AH, Li N, Aanes H, Østrup O, et al. Prepatterning of developmental gene expression by modified histones before zygotic genome activation. Dev Cell. 2011;21:993–1004.

59. Zhang B, Wu X, Zhang W, Shen W, Sun Q, Liu K, et al. Widespread Enhancer Dememorization and Promoter Priming during Parental-to-Zygotic Transition. Mol Cell. 2018;72:673–86.e6.

60. Murphy PJ, Wu SF, James CR, Wike CL, Cairns BR. Placeholder Nucleosomes Underlie Germline-to-Embryo DNA Methylation Reprogramming. Cell. 2018;172:993–1006.e13.

61. Kaaij LJT, Mokry M, Zhou M, Musheev M, Geeven G, Melquiond ASJ, et al. Enhancers reside in a unique epigenetic environment during early zebrafish development. Genome Biol. 2016;17:146.

62. Andersen IS, Reiner AH, Aanes H, Aleström P, Collas P. Developmental features of DNA methylation during activation of the embryonic zebrafish genome. Genome Biol. 2012;13:R65.

63. Zhang Y, Vastenhouw NL, Feng J, Fu K, Wang C, Ge Y, et al. Canonical nucleosome organization at promoters forms during genome activation. Genome Res. 2014;24:260–6.

64. 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.

65. 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;25:1715–26.

66. Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell. 2012;151:994–1004.

67. Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. Cell. 2015;161:555–68.

68. Ladam F, Stanney W, Donaldson IJ, Yildiz O, Bobola N, Sagerström CG. TALE factors use two distinct functional modes to control an essential zebrafish gene expression program. Elife. 2018;7.

69. Oldfield AJ, Yang P, Conway AE, Cinghu S, Freudenberg JM, Yellaboina S, et al. Histone-fold domain protein NF-Y promotes chromatin accessibility for cell type-specific master transcription factors. Mol Cell. 2014;55:708–22.

70. Jänes J, Dong Y, Schoof M, Serizay J, Appert A, Cerrato C, et al. Chromatin accessibility dynamics across C. elegans development and ageing. Elife. 2018;7.

71. Pliner HA, Packer JS, McFaline-Figueroa JL, Cusanovich DA, Daza RM, Aghamirzaie D, et al. Cicero Predicts cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data. Mol Cell. 2018;71:858–71.e8.

72. 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.

73. Esch D, Vahokoski J, Groves MR, Pogenberg V, Cojocaru V, Vom Bruch H, et al. A unique Oct4 interface is crucial for reprogramming to pluripotency. Nat Cell Biol. 2013;15:295–301.

74. Tapia N, Reinhardt P, Duemmler A, Wu G, Araúzo-Bravo MJ, Esch D, et al. Reprogramming to pluripotency is an ancient trait of vertebrate Oct4 and Pou2 proteins. Nat Commun. 2012;3:1279.

75. Nelson AC, Cutty SJ, Niini M, Stemple DL, Flicek P, Houart C, et al. Global identification of Smad2 and Eomesodermin targets in zebrafish identifies a conserved transcriptional network in mesendoderm and a novel role for Eomesodermin in repression of ectodermal gene expression. BMC Biol. 2014;12:81.

76. Dubrulle J, Jordan BM, Akhmetova L, Farrell JA, Kim S-H, Solnica-Krezel L, et al. Response to Nodal morphogen gradient is determined by the kinetics of target gene induction. Elife. 2015;4.

77. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. Dev Dyn. 1995;203:253–310.

78. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res. 2014;42:W401–7.

79. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Res. 2016;44:W272–6.

80. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol. 2015;109:21.29.1–9.

81. Link V, Shevchenko A, Heisenberg C-P. Proteomics of early zebrafish embryos. BMC Dev Biol. 2006;6:1.

82. Nelson AC, Cutty SJ, Gasiunas SN, Deplae I, Stemple DL, Wardle FC. In Vivo Regulation

of the Zebrafish Endoderm Progenitor Niche by T-Box Transcription Factors. Cell Reports. 2017. p. 2782–95.

83. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37:W202–8.

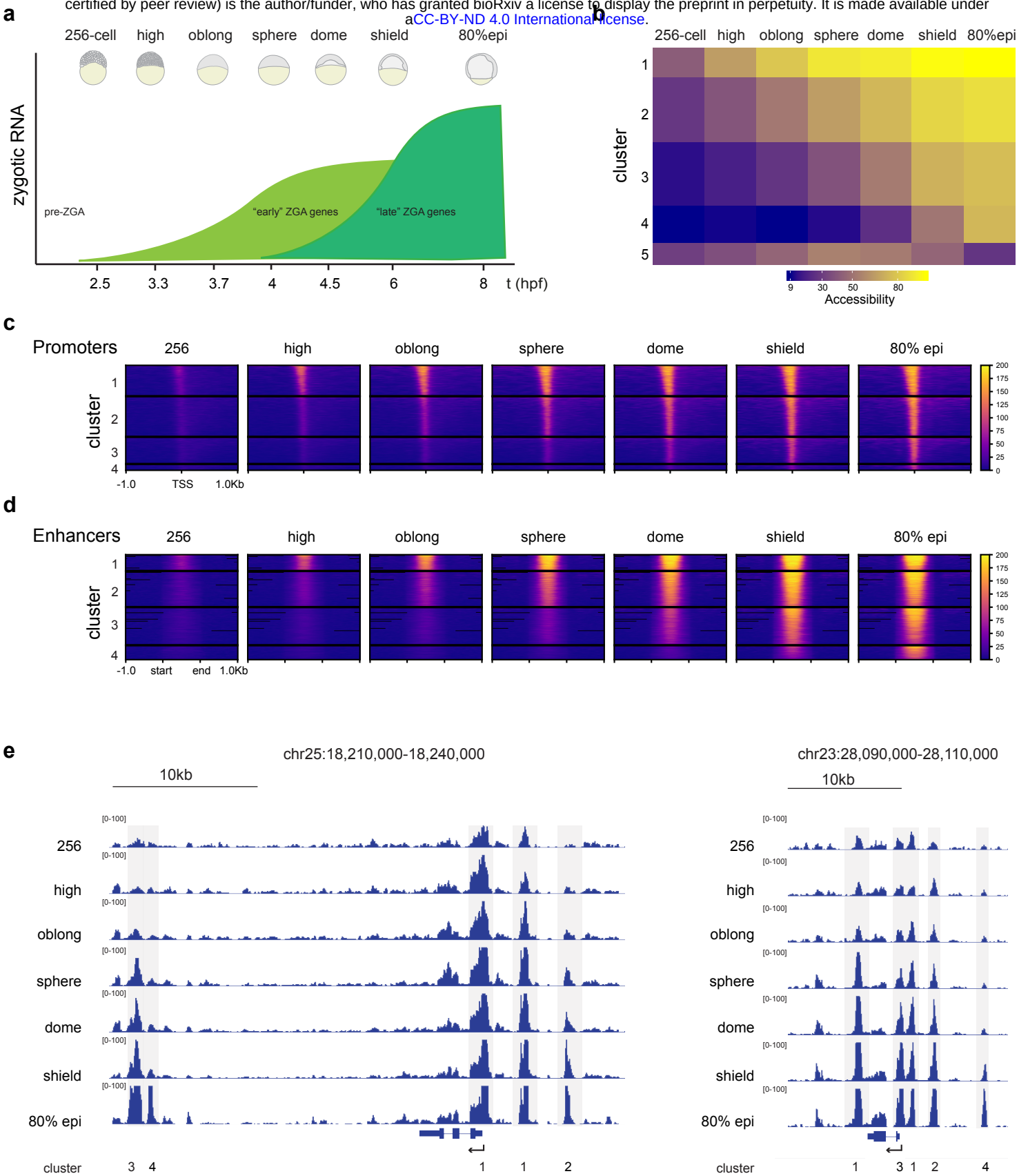


Figure 1. Chromatin accessibility at regulatory elements increases during zygotic genome activation.
a) Schematic of timepoints and embryonic stages chosen for ATAC-seq. "Early" ZGA genes correspond to genes activated until sphere stage and "late" ZGA genes correspond to genes activated after sphere stage. **b)** Five clusters of chromatin accessibility over the ATAC-seq time-series. **c)** TSS-centered heatmaps of ATAC-seq peaks in clusters 1 through 4. **d)** Heatmaps of ATAC-seq peaks centered on putative enhancer regions in clusters 1 through 4. **e)** Genome browser snapshots of two genomic regions over the ATAC-seq time-series. Peaks belonging to clusters 1-4 are indicated.

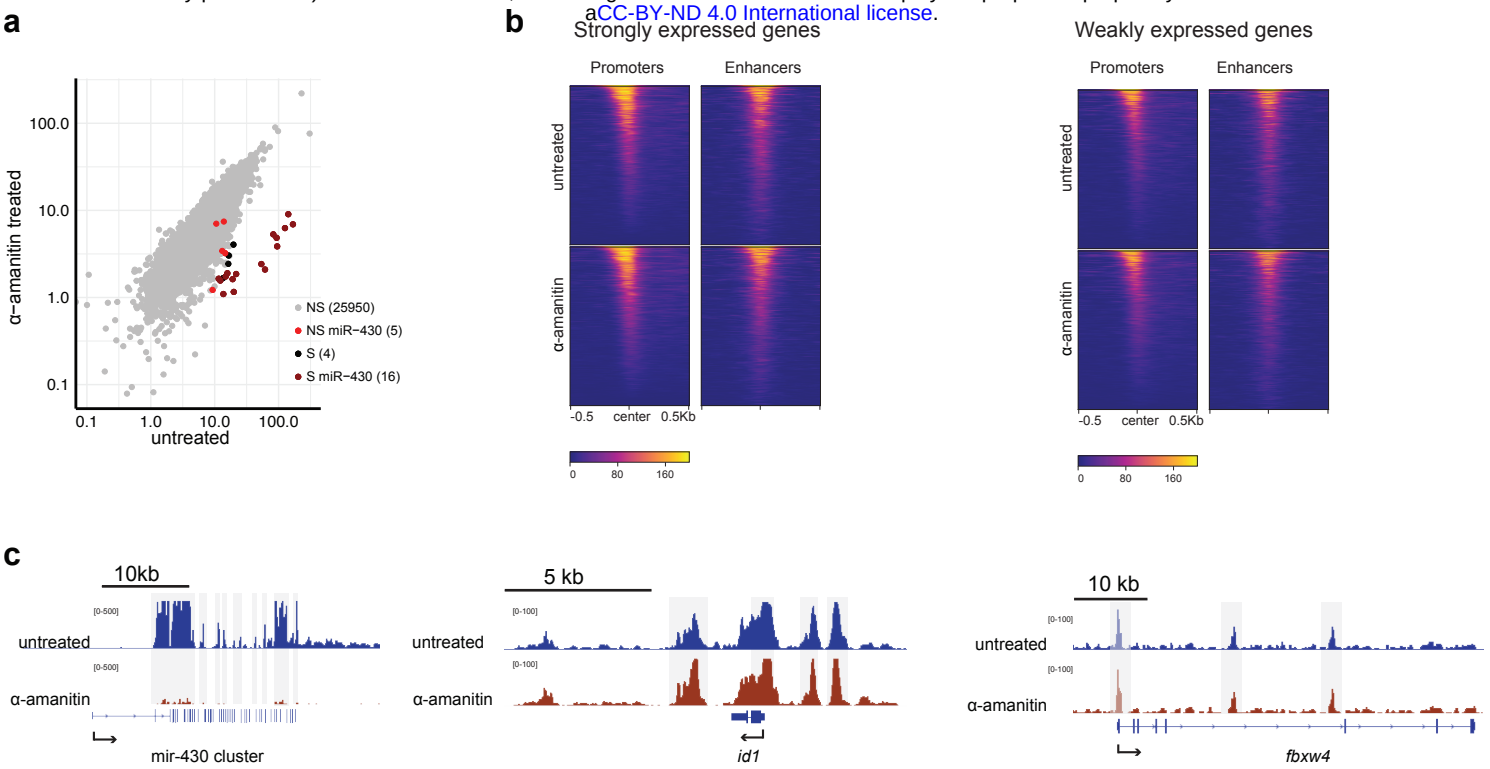


Figure 2. Transcription inhibition affects chromatin accessibility only at the miR-430 locus.
a) Correlation dot plot comparing accessibility over ATAC-seq peaks in α-amanitin-treated with untreated embryos. Significantly (S) and non-significantly (NS) affected peaks (DESeq2 log2 fold change -1.5, 5% FDR), as well as peaks that fall into the miR-430 cluster are indicated. **b)** Heatmaps of chromatin accessibility in untreated and α-amanitin-treated embryos, grouped based on the expression level of associated genes at sphere stage (expression data taken from [15]; strongly expressed genes are defined as the top 20% genes expressed at sphere stage (RPKM >= 0.419), whereas weakly expressed genes are all other genes (with a minimal expression of RPKM >= 0.004 at sphere stage. **c)** Genome browser snapshots of untreated and α-amanitin-treated embryos at oblong stage for the hyper-transcribed miR-430 gene cluster; *id1*, a gene with strong zygotic expression; and *fbxw4*, a gene with weak zygotic expression.

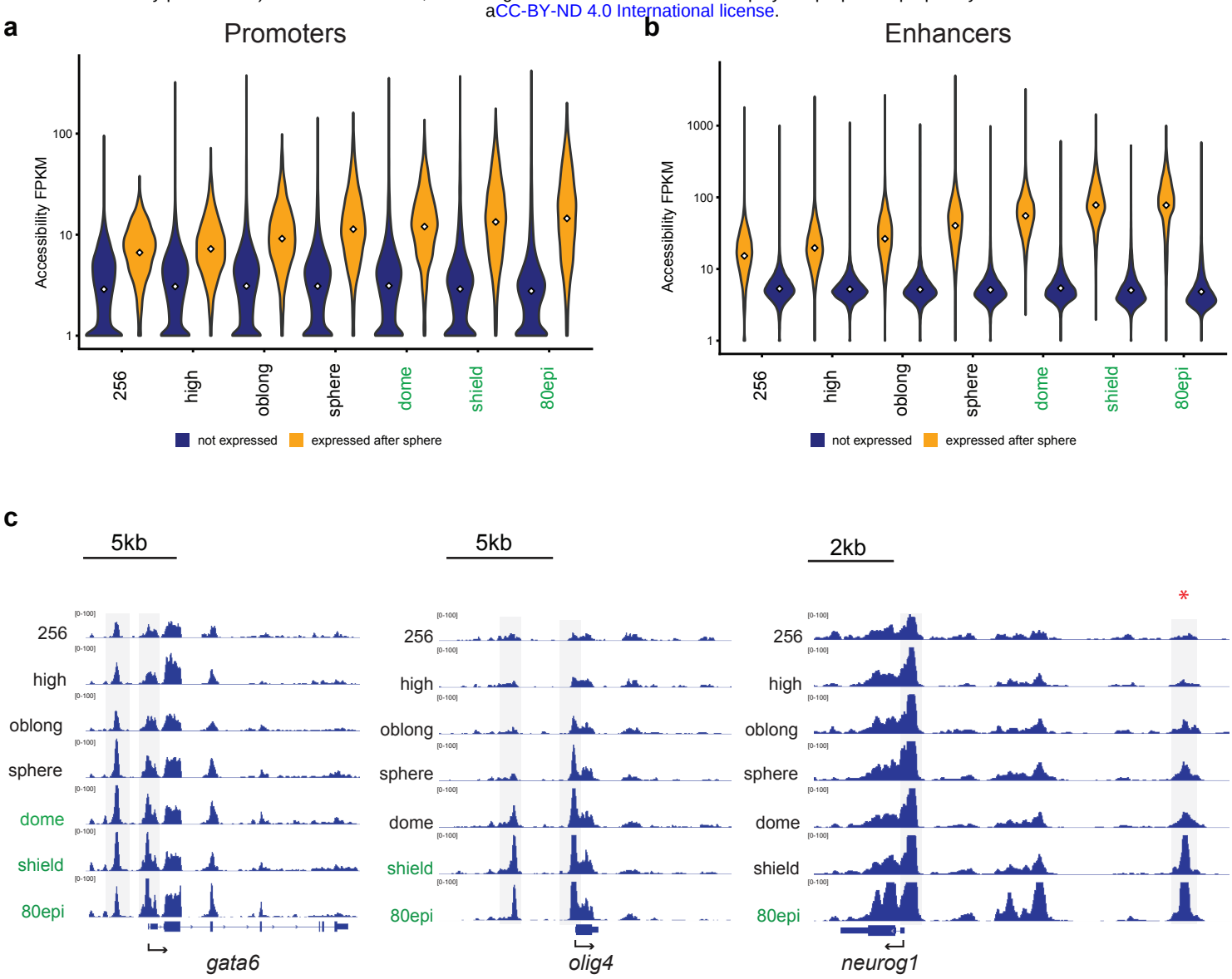


Figure 3. Chromatin accessibility at regulatory elements increases prior to gene activity.
a) Distribution of accessibility values (ATAC-seq FPKM) of promoters associated with genes that become activated after sphere stage (yellow), and genes that are not expressed (blue). White squares represent the median value. **b)** Distribution of accessibility values of enhancers associated with genes that become activated after sphere stage (yellow), and genes that are not expressed during early development (blue). White squares represent the median value. **c)** Genome browser snapshot of chromatin accessibility at the *gata6*, *olig4* and *neurog1* gene loci; stages at which genes are transcribed are indicated with green fonts, the red asterisk marks the experimentally validated *neurog1* enhancer [36]

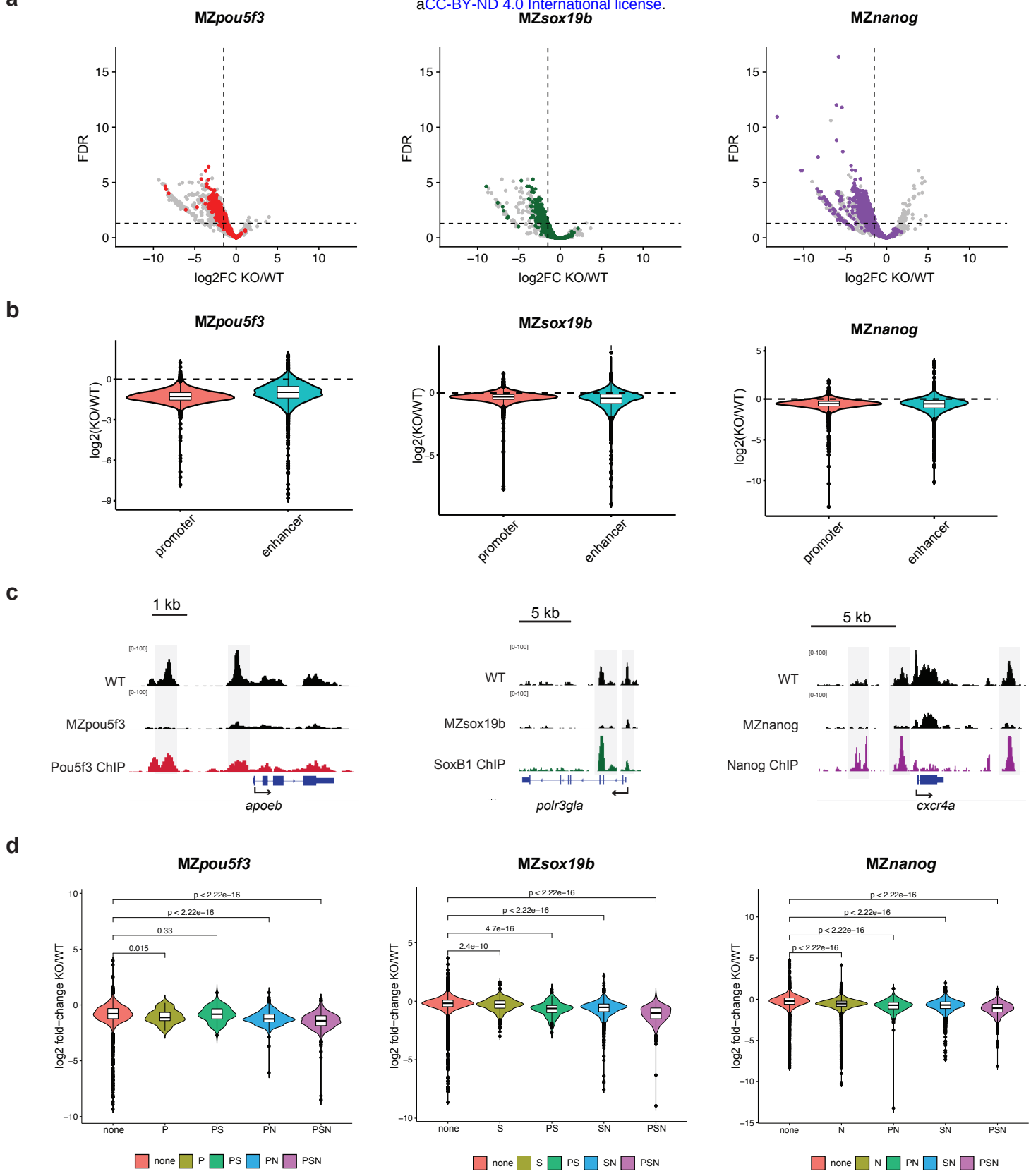


Figure 4. Pou5f3, Sox19b and Nanog regulate chromatin accessibility at gene regulatory elements.
a) Volcano plots showing the log2 fold change in accessibility in MZpou5f3, MZsox19b and MZnanog mutants compared to wild-type at oblong stage. Sites bound by Pou5f1, SoxB1, and Nanog are indicated in red, purple, and green, respectively. Non-bound sites are indicated in grey. **b)** Violin plots show the log2 fold change in accessibility at promoters and putative enhancers in the respective TF mutants compared to wild-type. **c)** Genome browser snapshots show ATAC accessibility tracks in MZpou5f3, MZsox19b and MZnanog mutants and wild-type embryos at oblong stage, for target genes of the respective TFs [15,24]. ChIP-seq tracks for the respective TFs at blastula stage are from data in [22,23]. **d)** Violin plots show the log2 fold change in mutants over wild-type at sites bound by different combinations of Pou5f3, SoxB1 and Nanog. Bars with p-values indicate comparisons of different sets of binding sites by one-sided Wilcoxon tests.

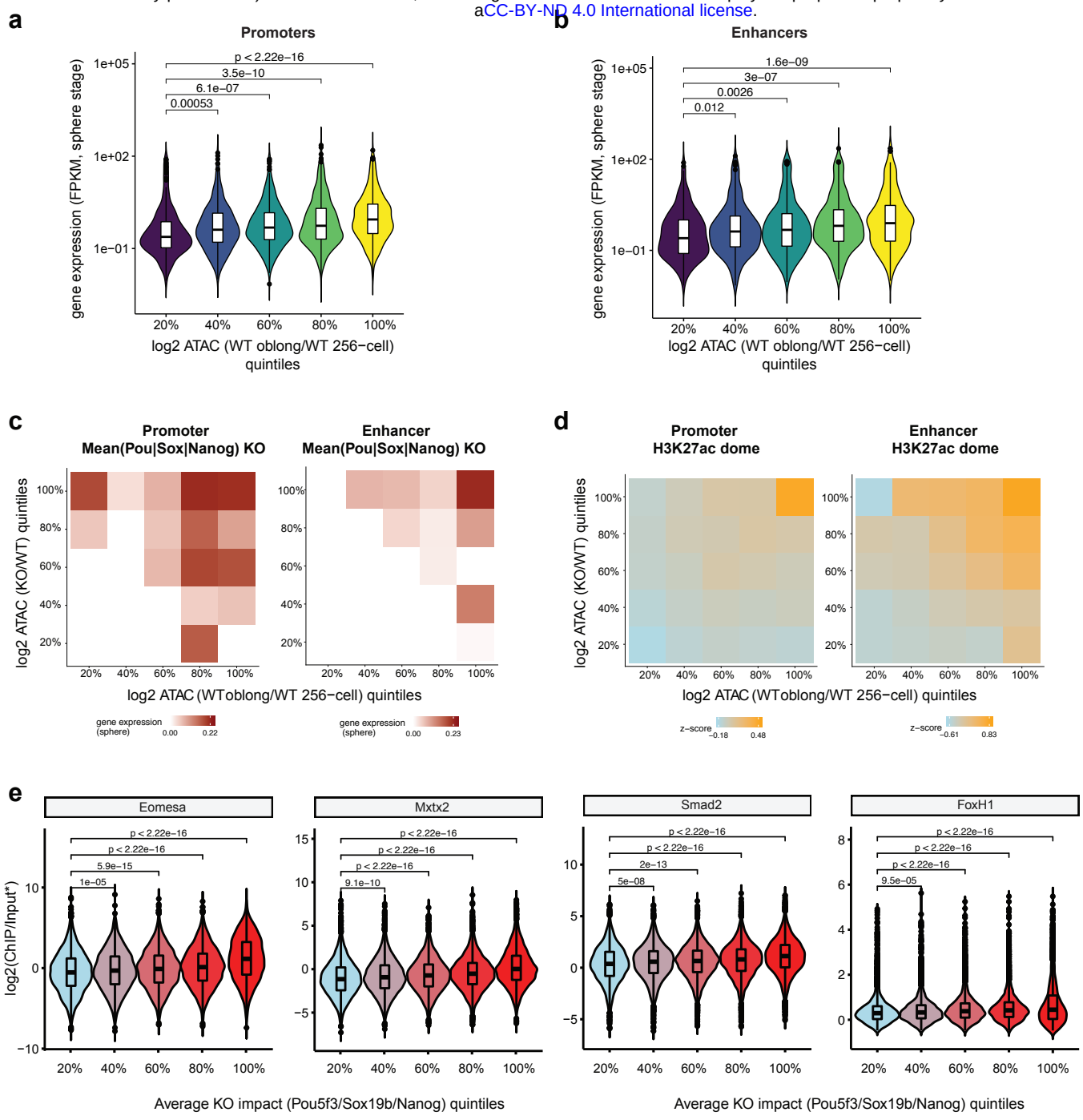
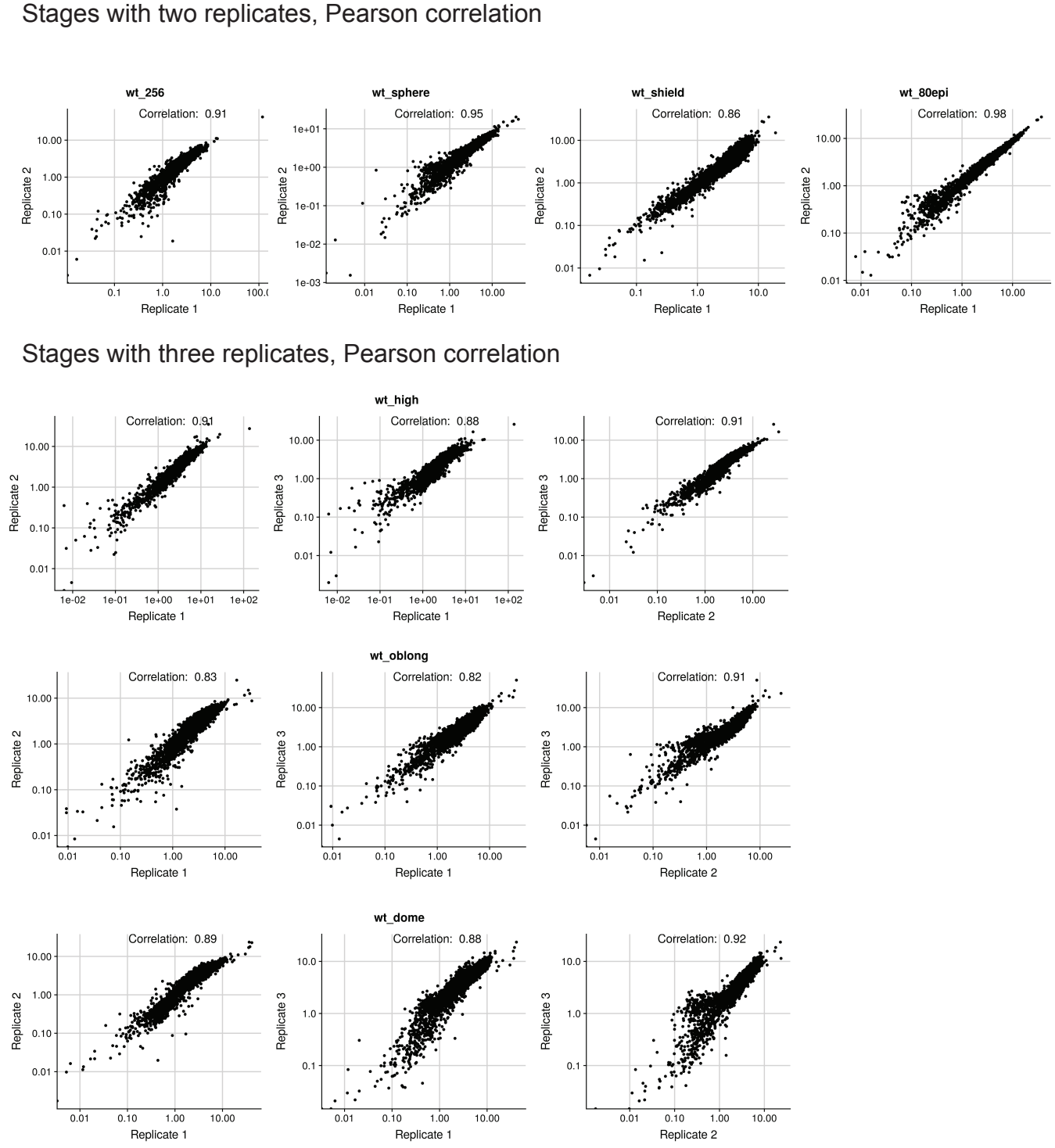


Figure 5. An increase in chromatin accessibility that is regulated by Pou5f3, Sox19b and Nanog most strongly predicts future gene expression.

a) Promoter regions were sorted into 20% quintiles based on accessibility increase between 256-cell and oblong stage, and violin plots show the expression value of associated genes at sphere stage. Bars with p-values indicate comparisons of the different sets of promoter regions by one-sided Wilcoxon tests. **b)** Putative enhancer regions were sorted into 20% quintiles based on accessibility increase between 256-cell and oblong stage and violin plots show the expression value of associated genes at sphere stage. Bars with p-values indicate comparisons of different sets of putative enhancer regions by one-sided Wilcoxon tests. **c)** Heatmaps show the median expression value for genes associated with regulatory regions at sphere stage, genomic regions are resolved by 20% bins for accessibility increase between 256-cell and oblong stage (x-axis), and accessibility change in MZpou5f3, MZsox19b and MZnanog mutants compared to wild-type embryos at oblong stage (y-axis). **d)** Heatmaps show the mean z-score normalized H3K27ac signal at the same genomic regions as in c. **e)** Regulatory elements were binned into 20% quintiles based on accessibility change in MZpou5f3, MZsox19b and MZnanog mutants compared to wild-type embryos, and violin plots show the enrichment (ChIP/Input) of Eomesa, Mxtx2, Smad2 and FoxH1 binding at blastula stages [23,75,76]. *FoxH1 signal was normalized to the genome-wide mean as no input sample was available. Bars with p-values indicate significant differences in TF binding strength between the different bins as assessed by one-sided Wilcoxon tests.

Supplemental Figures

a



b

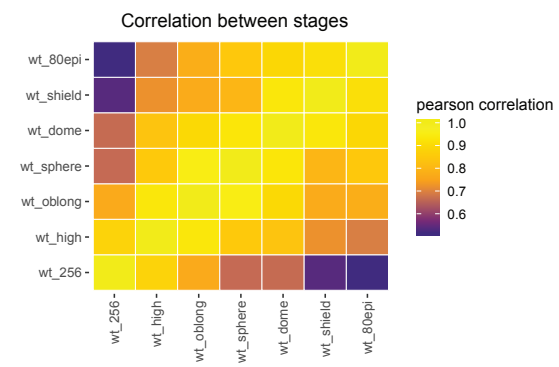


Figure S1. Correlation of chromatin accessibility between ATAC-seq samples in early zebrafish embryos.
a) Pearson correlation between individual ATAC-seq replicates. Two biological replicates were generated for 256-cell, sphere, shield and 80% epiboly stage, and three biological replicates for high, oblong and dome stage. b) Pearson correlation between pooled replicates of different stages.

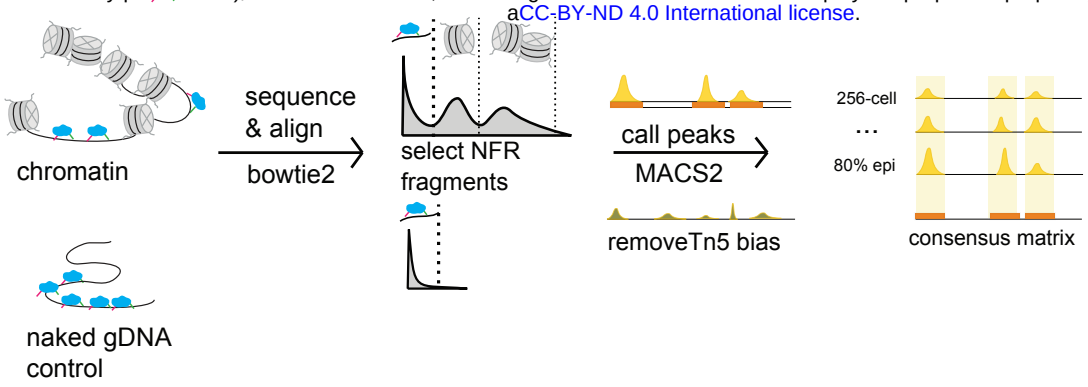


Figure S2. Pipeline for analyzing chromatin accessibility.
Chromatin and naked genomic DNA from zebrafish embryos was tagmented using Tn5 transposase, followed by sequencing and alignment to the genome (bowtie2). Fragments of sub-nucleosomal size (nucleosome free regions, NFR) were selected by applying a 130bp cut-off after visual inspection of the fragment length profiles. Peaks were called using MACS2 using the signal obtained from the naked DNA sample as background. Peaks called in each wild-type stage were merged across the time-course to generate a consensus peak set for further analysis.

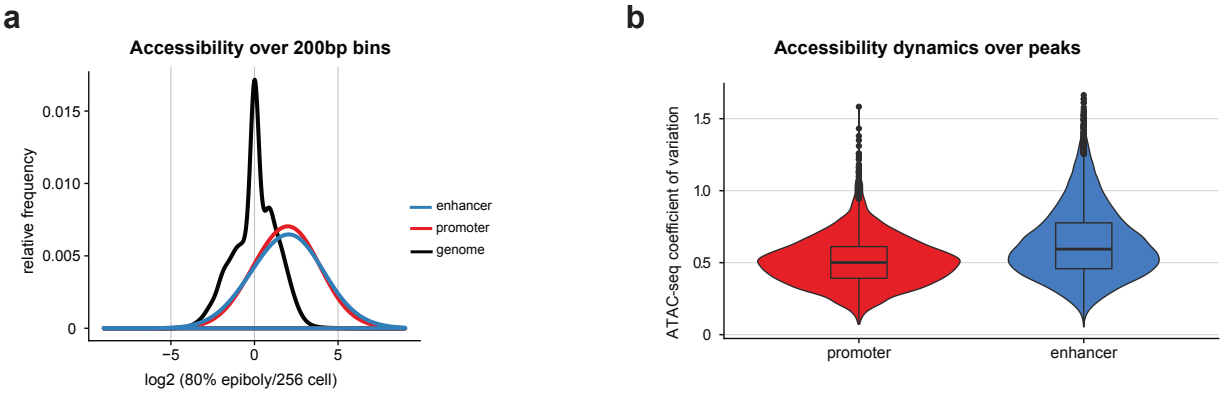


Figure S3. Dynamics of accessibility increase at regulatory elements.
a) Average accessibility increase across the time-series (from 256-cell to 80% epiboly) over 200bp bins genome-wide (black), and at promoter (red) and putative enhancer elements (blue). b) The coefficient of variation of normalized accessibility across the time course for promoter and putative enhancer associated peaks. A higher value indicates more variation in accessibility across stages.

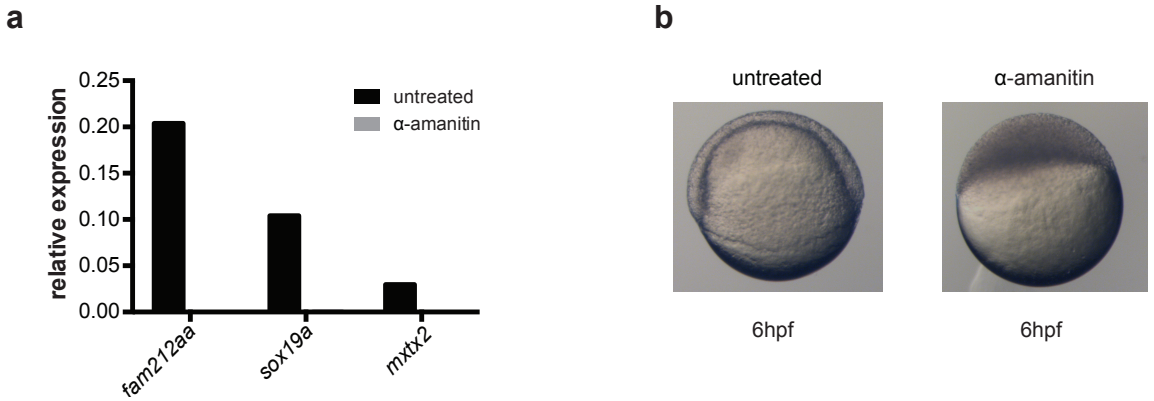


Figure S4. Transcription inhibition in zebrafish embryos.
a) Gene expression levels of the zygotically activated genes fam212aa, sox19a and mxtx2 in untreated and α-amanitin-treated embryos. b) Brightfield images of zebrafish embryos at 6hpf in untreated and α-amanitin-treated embryos. α-amanitin treatment results in embryonic arrest at sphere stage.

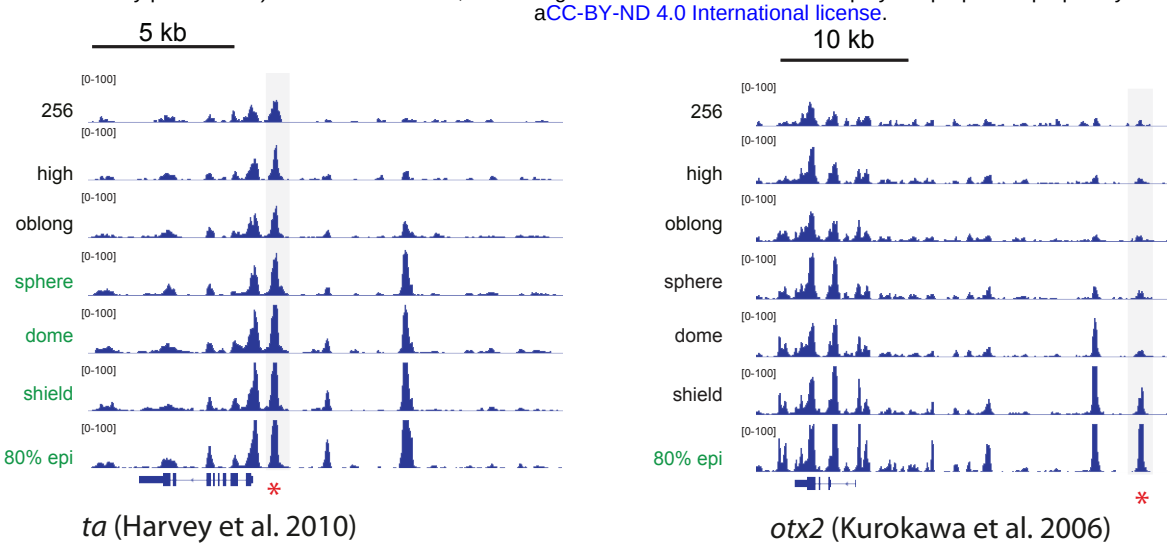


Figure S5. Additional examples for the observed increase in accessibility at regulatory elements prior to gene activity.
Genome browser snapshots showing the ta and otx2 gene loci. Green fonts indicate the stages when the gene is active. Experimentally verified enhancers for these genes [37,38] are marked by a red asterisk.

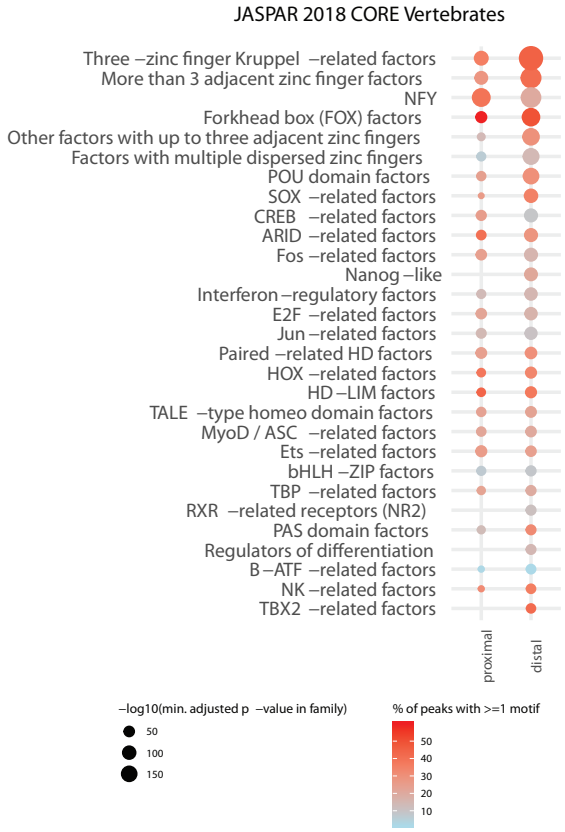


Figure S6. Motif analysis for regions with increasing accessibility.
Enrichment of TF motifs from the JASPAR vertebrates database (with the custom Nanog motif from [23]) at distal (enhancer/intergenic) and proximal (within +/- 1kb of TSS) regions that show an increase in accessibility between 256-cell and oblong stage. Individual motifs were summarized at the motif-family level and the minimum adjusted p-value for each family is indicated by the circle size in the plot. The color scale ranging from blue (low) to red (high) indicates the % of peaks with at least one motif from the respective motif-family.

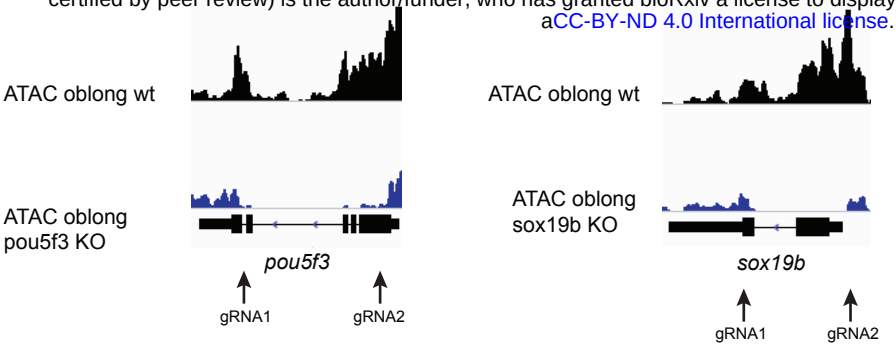


Figure S7. CRISPR deletion of pou5f3 and sox19b genes.
Genome browser snapshots showing ATAC accessibility tracks at pou5f3 and sox19b in wild-type and knock-out embryos. The absence of reads over the gene body of pou5f3 and sox19b in knock-outs confirms the deletion of the loci. Location of the gRNAs used for deleting the genes are indicated with arrows.

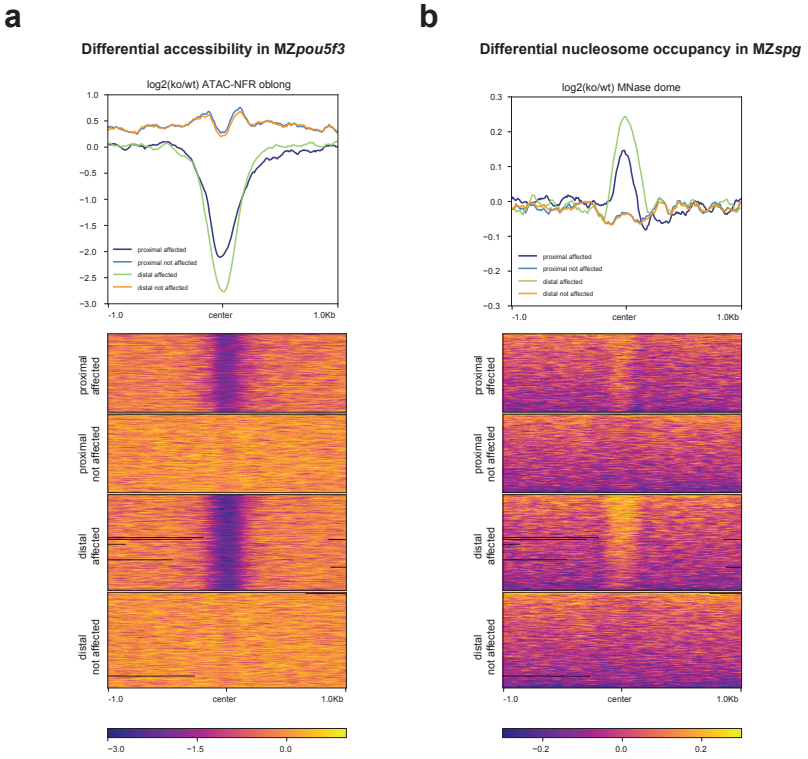


Figure S8. Comparison of chromatin accessibility changes in maternal-zygotic pou5f3 mutants assayed by ATAC-seq and MNase-seq
a) Metagene profiles and heatmaps showing the log2 fold change in accessibility in MZpou5f3 mutants compared to wild-type at oblong stage. b) Metagene profile and heatmaps for the same regions as in a), showing the log2 fold change in nucleosome occupancy in MZspg mutants compared to wild-type at dome stage, using data from [26]

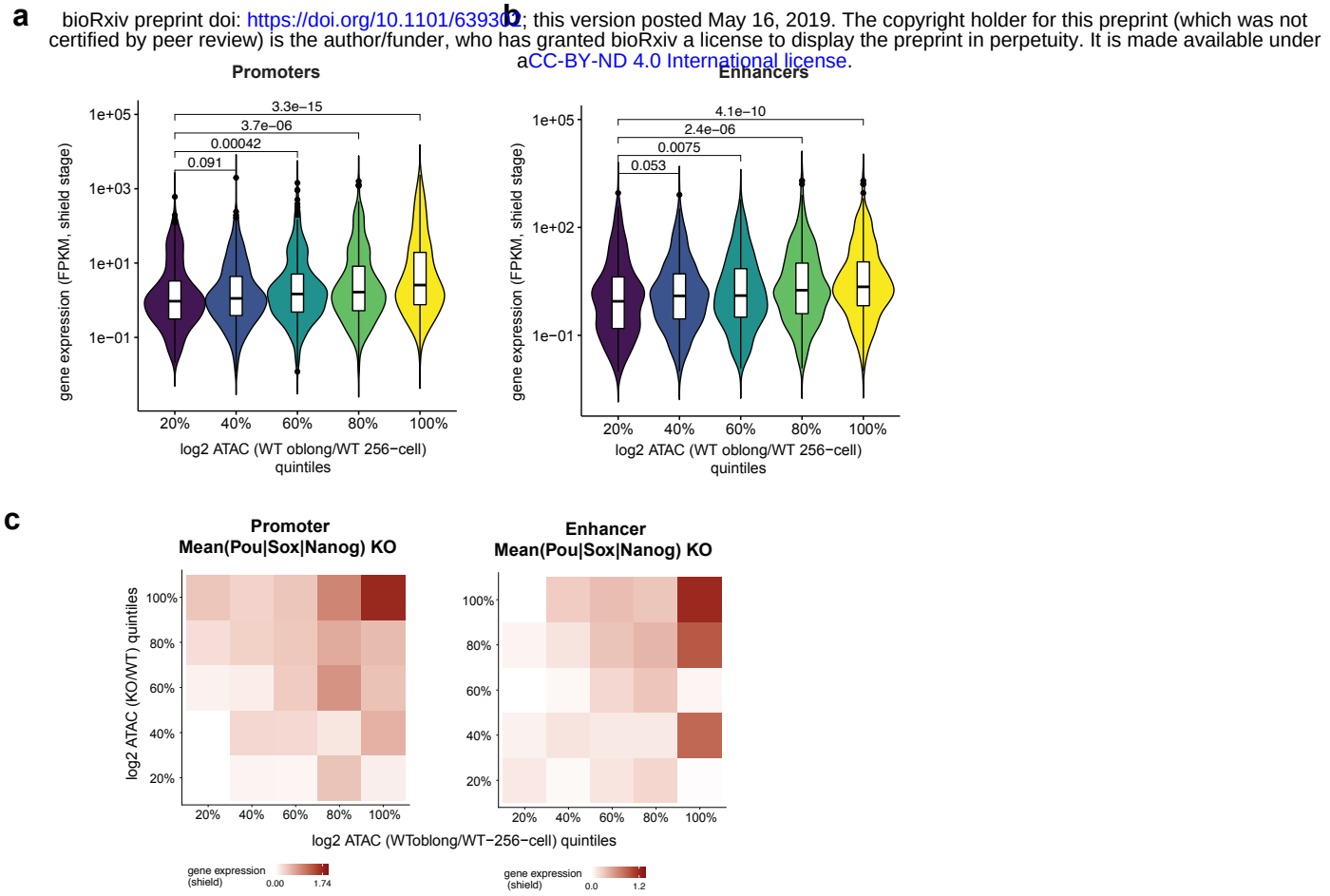


Figure S9. An increase in chromatin accessibility that is regulated by Pou5f3, Sox19b and Nanog most strongly predicts future gene expression.

a) Promoter regions were sorted into 20% quintiles based on accessibility increase between 256-cell and oblong stage, and violin plots show the expression value of associated genes at shield stage. Bars with p-values indicate comparisons of the different sets of promoter regions by one-sided Wilcoxon tests. b) Putative enhancer regions were sorted into 20% quintiles based on accessibility increase between 256-cell and oblong stage, and violin plots show the expression value of associated genes at shield stage. Bars with p-values indicate comparisons of different sets of putative enhancer regions by one-sided Wilcoxon tests c) Heatmaps show the median expression value for genes associated with regulatory regions at shield stage, genomic regions resolved by 20% bins for accessibility increase between 256-cell and oblong stage (x-axis), and accessibility change in MZpou5f3, MZsox19b and MZnanog mutants compared to wild-type embryos at oblong stage (y-axis).