

RESEARCH ARTICLE

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

Máté Pálffy^{1‡}, Gunnar Schulze^{2‡}, Eivind Valen^{2,3}, Nadine L. Vastenhouw^{1*}

1 Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany, **2** Computational Biology Unit, Department of Informatics, University of Bergen, Bergen, Norway, **3** Sars International Centre for Marine Molecular Biology, University of Bergen, Bergen, Norway

‡ These authors share first authorship on this work.

* 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 remain poorly understood. Here, we generated chromatin accessibility maps with ATAC-seq 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 hypertranscribed 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.

OPEN ACCESS

Citation: Pálffy M, Schulze G, Valen E, Vastenhouw NL (2020) Chromatin accessibility established by Pou5f3, Sox19b and Nanog primes genes for activity during zebrafish genome activation. *PLoS Genet* 16(1): e1008546. <https://doi.org/10.1371/journal.pgen.1008546>

Editor: A. Aziz Aboobaker, University of Oxford, UNITED KINGDOM

Received: June 20, 2019

Accepted: December 2, 2019

Published: January 15, 2020

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pgen.1008546>

Copyright: © 2020 Pálffy et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression

Author summary

In eukaryotes, DNA is packed inside the cell nucleus in the form of chromatin, which consists of DNA, proteins such as histones, and RNA. Chromatin accessibility influences when and where DNA-binding proteins such as transcription factors and RNA polymerase II find their targets in order to activate transcription. It is unclear, however, whether the accessibility of regulatory regions precedes and predicts future transcription. To address these questions, we took advantage of ATAC-seq, which is a powerful technique to probe chromatin accessibility states. We analyzed zebrafish embryos as they gear up to start transcription for the first time and related the accessibility of regulatory regions to the activity of associated genes. This revealed that chromatin at regulatory regions is often accessible prior to gene activity and that this has predictive value for future transcription.

Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE130944.

Funding: This work was supported by the Max Planck Society (<https://www.mpg.de/en>), a Human Frontier Science Program Career Development Award to NLV, CDA00060/2012 (<https://www.hfsp.org/>), the Bergen Research Foundation to EV (<https://www.mohnfoundation.no/?lang=en>) and the Norwegian Research Council to EV, #250049 (<https://www.forskningsradet.no/en/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Analyzing transcription factor mutants, we further determined that Pou5f3, Sox19b and Nanog play an important role in opening up chromatin. Thus, our study shows that transcription factors open regulatory elements to prime gene activity during development.

Introduction

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 [15–18]. It starts as early as the 64-cell stage, when the miR-430 gene cluster is activated [19–22], 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 [15,23–25]. Furthermore, zebrafish Pou5f3 and Nanog have been shown to play a role in promoting an open chromatin state at regulatory elements [26,27]. 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 [28]. Data from bulk ATAC-seq reflects the average accessibility of cells in a population [3]. 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) (Fig 1A). Biological replicates showed a large degree of correlation and samples from adjacent stages clustered together (S1A and S1B Fig), demonstrating the high quality of our data. We called ATAC-seq peaks by MACS2 [29], using genomic DNA as an input control (S2 Fig and Methods) and used k-means to cluster these peaks. We identified five clusters with different chromatin accessibility dynamics (Fig 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 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 (S3 Fig), 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 (Fig 1C and 1D). This dynamic opening of regulatory elements is also apparent at individual gene loci (Fig 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 that we observe 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 (S4 Fig), as previously described [15,30,31]. 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 (Fig 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 (Fig 2A). Interestingly, the expression level of miR-430 is orders of magnitude higher than that of any other gene during early zebrafish development [19]. 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* [32–34], 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 (Fig 2B). This is exemplified in genome browser snapshots of chromatin accessibility at the miR-430 locus (hypertranscribed), *id1* (strongly expressed), and *fbxw4* (weakly expressed) loci in transcription-inhibited and untreated embryos (Fig 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.

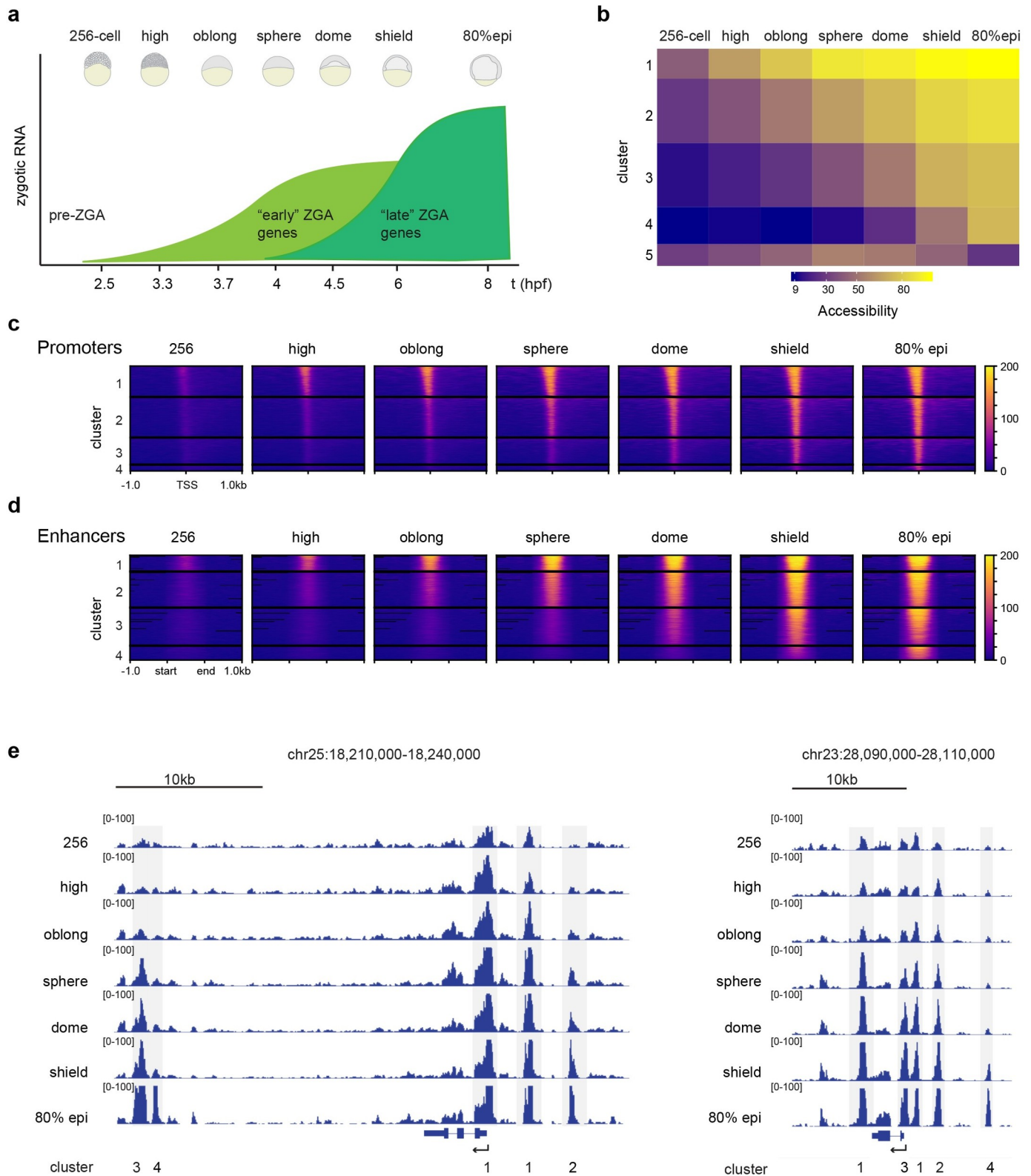


Fig 1. Chromatin accessibility at regulatory elements increases during zygotic genome activation. a) Schematic of timepoints and embryonic stages chosen for ATAC-seq. "Early" ZGA genes correspond to genes activated until sphere stage and "late" ZGA genes correspond to genes activated after sphere stage. 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.

<https://doi.org/10.1371/journal.pgen.1008546.g001>

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 Fig 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 (Fig 3A). Moreover, chromatin accessibility increases during the stages preceding transcription activation (Fig 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 (Fig 3B). As for promoters, accessibility of

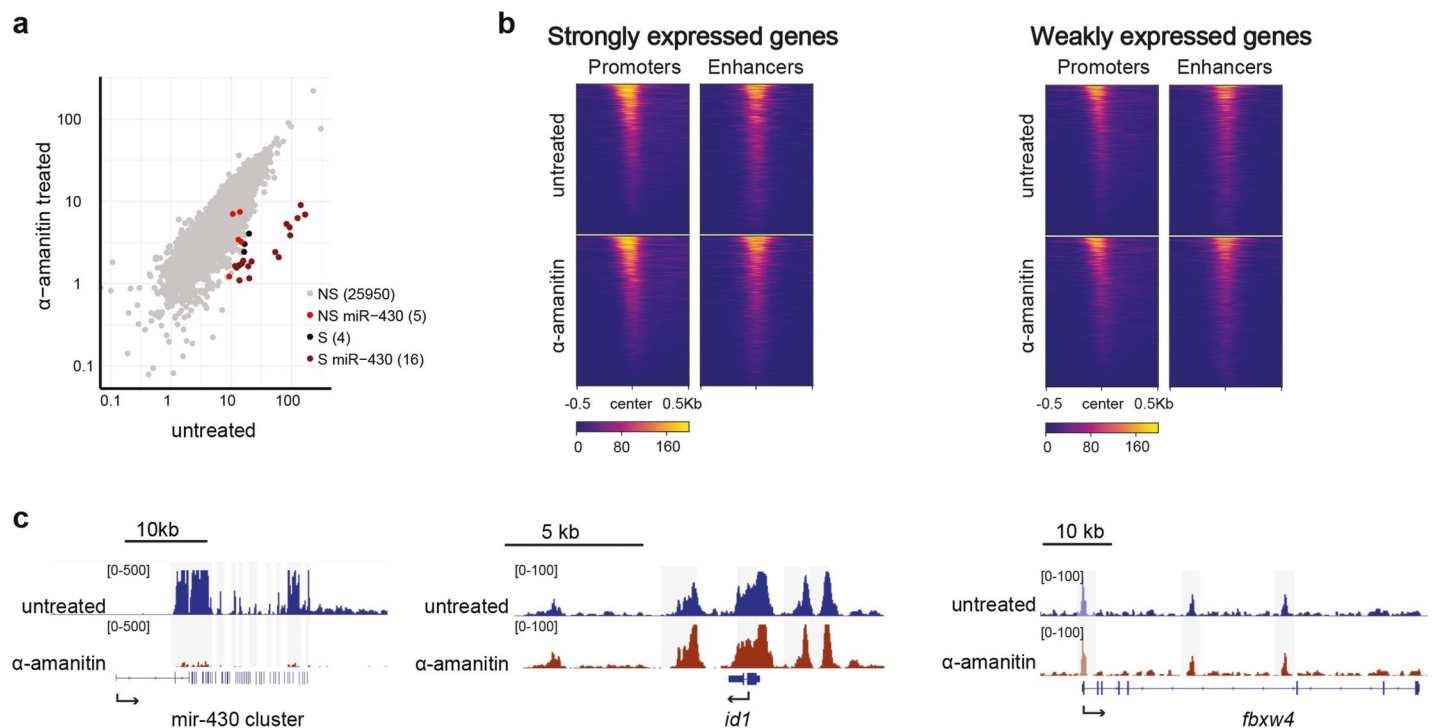


Fig 2. Transcription inhibition affects chromatin accessibility only at the miR-430 locus. a) Correlation dot plot comparing accessibility in α -amanitin-treated embryos with untreated embryos. Significantly (S) and non-significantly (NS) affected peaks (DESeq2 log₂ fold change > 1.5 , 5% FDR), as well 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 hypertranscribed miR-430 gene cluster; *id1*, a gene with strong zygotic expression; and *fbxw4*, a gene with weak zygotic expression.

<https://doi.org/10.1371/journal.pgen.1008546.g002>

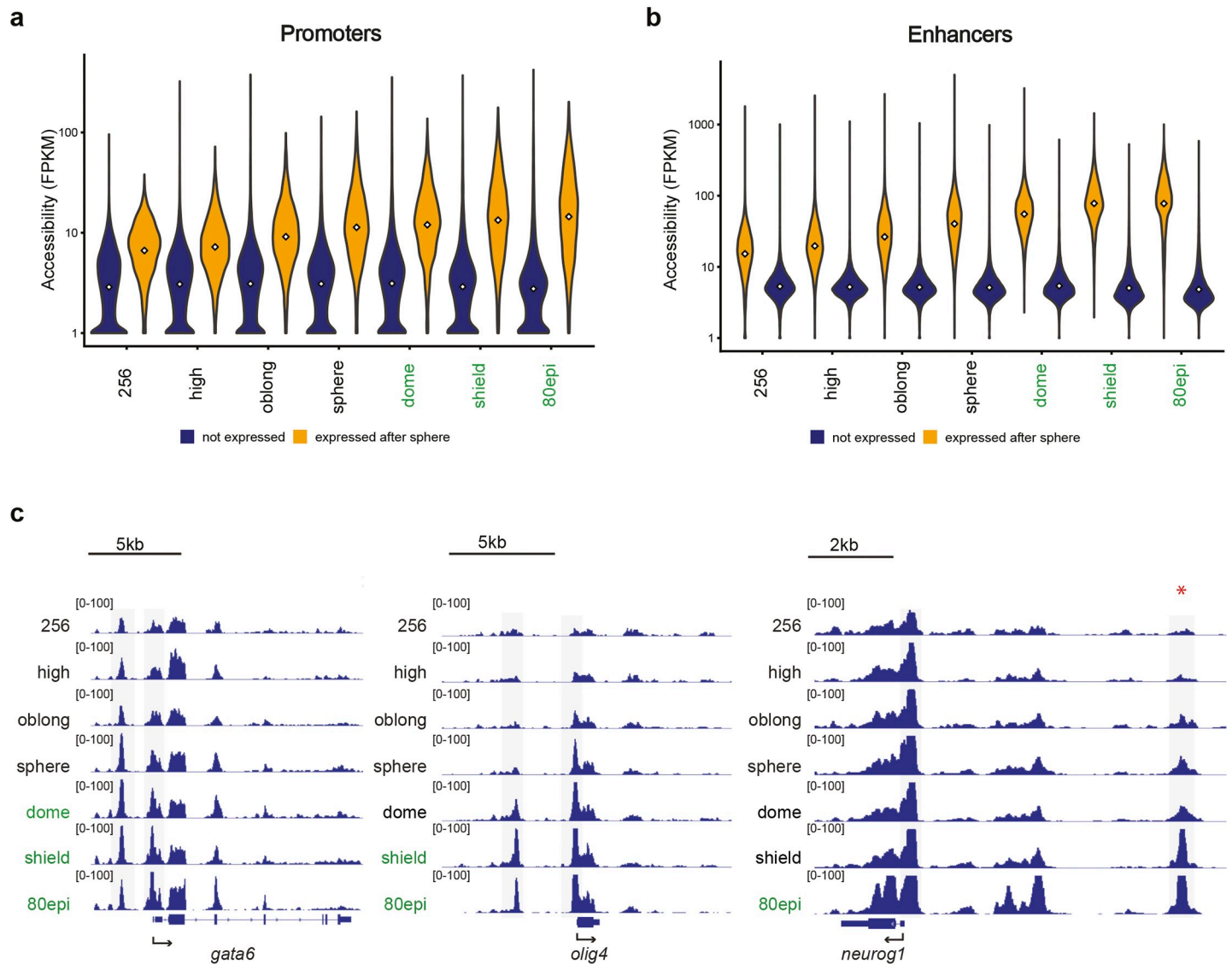


Fig 3. Chromatin accessibility at regulatory elements increases prior to gene activity. a) Distribution of accessibility values (ATAC-seq FPKM) of promoters associated with genes that become activated after sphere stage (yellow), and genes that are not expressed (blue). White squares represent the median 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 snapshots 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 an experimentally validated *neurog1* enhancer [37].

<https://doi.org/10.1371/journal.pgen.1008546.g003>

putative enhancers increases during the stages prior to gene activation. Here, we used zebrafish enhancers from adult tissues as a control [35]. 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 [16,36], are accessible at 256-cell stage and their accessibility increases prior to transcription (Fig 3C). A similar relationship between accessibility and transcription activation can be observed for the regulatory elements of *olig4* (activated at shield stage), and experimentally validated enhancers of *neurog* (activated at 80%

epiboly stage), *ta* (activated at sphere stage), and *otx2* (activated at 80% epiboly stage) [37–39] (Fig 3C, S5 Fig). 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 [15,23,25]. Moreover, their motifs are enriched in the distal regions that increase in accessibility between 256-cell and oblong stage (S6 Fig). We created *pou5f3* and *sox19b* mutants using CRISPR-Cas9 [40] (S7 Fig). 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 [41,42], or the generation of unexpected transcripts that escape nonsense-mediated decay [43]. Maternal-zygotic (MZ) *pou5f3* embryos arrest in their development during late gastrulation, as previously described [44–46]. 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 [47]. For our studies this redundancy does not cause a problem, because at the early stages of development only Sox19b is present at high levels [15,47]. For *nanog*, we obtained a recently published null mutant that harbors a 10 base pair deletion in the first exon, causing a frameshift [48]. Importantly, the *nanog* mutant phenotype has previously been shown to be identical to embryos in which Nanog is knocked down by translation-blocking morpholinos [25,48], 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. Oblong stage is the last embryonic stage during which development relies entirely on maternal products, which makes it unlikely that effects observed in the mutants would be indirect. Upon the loss of Pou5f3, we observed a widespread decrease in chromatin accessibility (Fig 4A, S8 Fig). Similarly, the loss of Sox19b and Nanog resulted in a strong decrease in accessibility at hundreds of genomic regions (Fig 4A, S8 Fig). When analyzing changes in accessibility at regulatory elements, we found that promoters and enhancers are strongly affected (Fig 4B), as is apparent at genomic loci at Pou5f3-, Sox19b- and Nanog-regulated genes (Fig 4C). As expected, these changes in accessibility are significantly larger at promoters and enhancers of genes that are regulated by the respective factors [15] (S9 Fig).

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 [27]. As a consequence, the effects on promoter accessibility were not addressed. To determine whether the effect of Pou5f3 on promoter accessibility that we observed (Fig 4C) could also be observed in the dataset that was previously generated [27], 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 (S10 Fig). We conclude that Pou5f3, Sox19b, and Nanog are important for the opening of chromatin at regulatory elements during zygotic genome activation.

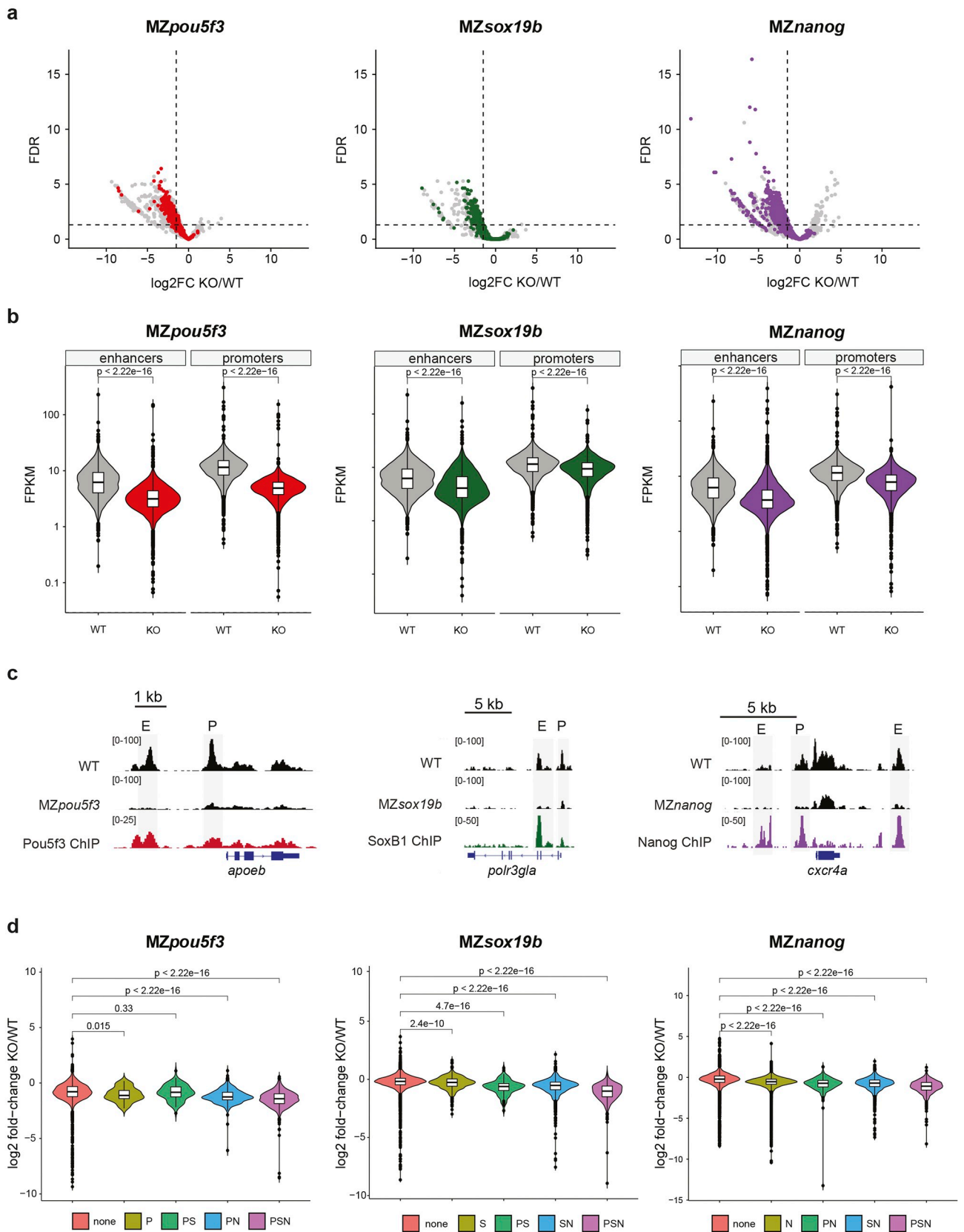


Fig 4. Pou5f3, Sox19b and Nanog regulate chromatin accessibility at gene regulatory elements. a) Volcano plots showing the log₂ fold change in accessibility in *MZpou5f3*, *MZsox19b* and *MZnanog* mutants compared to wild-type embryos at oblong stage. Sites bound by Pou5f1, SoxB1, and Nanog are indicated in red, green and purple respectively. Non-bound sites are indicated in grey. b) Violin plots show the accessibility value (FPKM) at promoters and putative enhancers in wild-type embryos compared to the respective TF mutants. Significance of differences was tested using paired t-tests. c) Genome browser snapshots show accessibility tracks in *MZpou5f3*, *MZsox19b* and *MZnanog* mutants and wild-type embryos at oblong stage, for target genes of the respective TFs [15,25]. ChIP-seq tracks for the respective TFs at blastula stage are from data in [23,24]. d) Violin plots show the log₂ fold change in mutants over wild-type at sites bound by different combinations of Pou5f3, SoxB1 and Nanog. p-values are shown for differences between binding sites as assessed by one-sided Wilcoxon tests.

<https://doi.org/10.1371/journal.pgen.1008546.g004>

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 [27]. 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 (Fig 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 *MZpou5f3*, *MZsox19b*, and *MZnanog* 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.

Finally, we analyzed the motifs at sites where Pou5f3, Sox19b or Nanog increase accessibility. Pioneer factors often target partial motifs [49]. During human stem cell reprogramming for example, Oct4 and Sox2 target partial DNA motifs on nucleosomes [50]. Similarly, a fraction of nucleosomal FoxA binding occurs at non-optimal motifs which are degenerate in at least two base positions [51]. In contrast, our de novo motif analysis identified only minimal deviations from the canonical binding sites with no significantly enriched partial motifs (S11 Fig). In addition, we observe the largest change in accessibility at sites with canonical binding motifs (S12 Fig). These results are in agreement with another study in zebrafish [27]. Together, this suggests that in zebrafish, Pou5f3, Sox19b and Nanog regulate chromatin accessibility predominantly at sites with high motif strength.

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 shows that larger increases in chromatin accessibility between 256-cell and oblong stage are correlated with higher expression levels at sphere stage (Fig 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 (Fig 5B). Similar results were obtained at shield stage (S13A and S13B Fig). 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 (Fig 5A and 5B, S13A and S13B Fig) 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

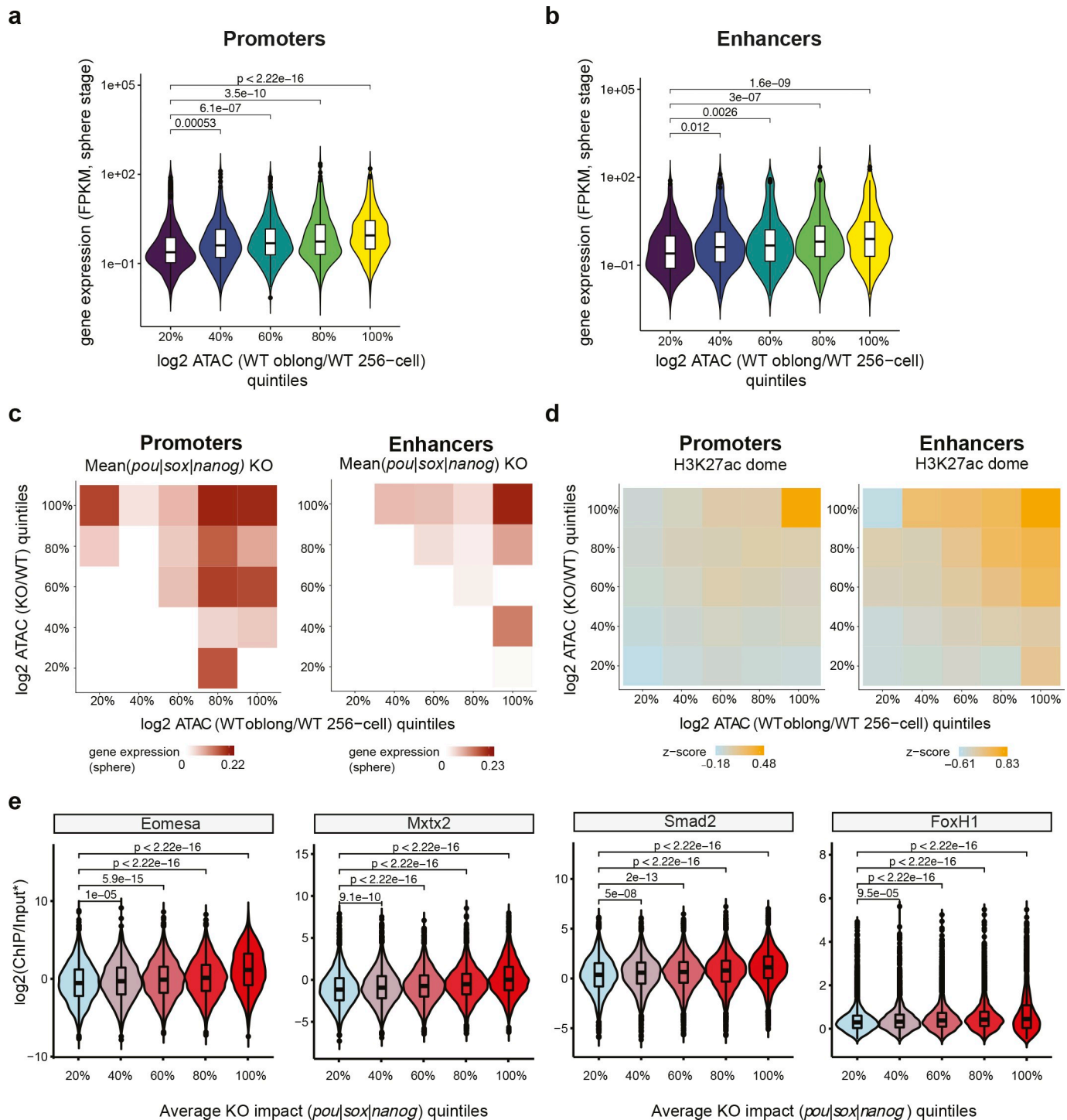


Fig 5. An increase in chromatin accessibility predicts future gene expression, especially at *Pou5f3/Sox19b/Nanog*-regulated regions. **a**) Promoter regions were sorted into 20% quintiles based on accessibility increase between 256-cell and oblong stage, and violin plots show the expression value of associated genes at sphere stage. p-values are shown for differences in expression between the quintiles as assessed 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. p-values are shown for differences in expression between quintiles as assessed 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 of accessibility increase between 256-cell and oblong stage (x-axis), and 20% bins of accessibility change in MZ*pou5f3*, MZ*sox19b* and MZ*nanog* 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 [24,81,82]. *FoxH1 signal was normalized to the genome-wide mean as no input sample was available. p-values are shown for differences in TF binding strength between the different bins as assessed by one-sided Wilcoxon tests.

<https://doi.org/10.1371/journal.pgen.1008546.g005>

by the average fold change in accessibility in Pou5f3, Sox19b, and Nanog mutants compared to wild-type embryos (Fig 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 loss of either Pou5f3, Sox19b or Nanog, have the highest expression level at sphere stage (Fig 5C). Here too, analysis of shield stage embryos led to the same conclusion (S13C Fig). Because of the observed correlation between zygotic transcription and H3K27ac levels on regulatory elements during zebrafish ZGA [22,52], 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 (Fig 5D). Together, this shows 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 *MZpou5f3*, *MZsox19b* or *MZnanog* mutants are the most highly enriched for the binding of the transcription factors Mxtx2, Eomesa, Smad2 and Foxh1 at blastula stages (Fig 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 [4,10–13,26,53]. Because bulk ATAC-seq measurements reflect an ensemble average of distinct molecular states [3,54,55], an increase in accessibility over time may be due to an increase in the number of cell types in which specific regions are accessible, or a change to higher accessibility states of specific regions in most, if not all, cells of the embryo [3]. Individual zebrafish cells are transcriptionally largely uniform during early developmental stages [56,57], which suggests that the chromatin accessibility landscape may also be largely uniform, similar to what has been described for the early stages of *Drosophila*

development [58]. Moreover, the maternally loaded transcription factors Pou5f3, Sox19b and Nanog, which play an important role in opening chromatin, are uniformly expressed in the embryo at the time of ZGA [24,47,59]. Together, this favors a scenario in which the increase in chromatin accessibility during zebrafish genome activation is caused by an embryo-wide change in chromatin accessibility states.

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

Our observation that chromatin accessibility is independent of transcriptional activity is consistent with studies that failed to detect a causal link between the transcription process and promoter chromatin accessibility at a genome-wide level [60,61]. 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 [32–34]. 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 [20–22,62], which might facilitate the formation of RNA-pol II ‘trains’ capable of evicting nucleosomes, as shown *in vitro* [34].

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, histone modifications H3K4me3 [63,64] and H3K27ac [62,65], a hypomethylated state [26,66–68], and well-positioned nucleosomes [69] 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 [58]. Furthermore, histones can be acetylated prior to the major wave of ZGA at sites of early Zelda binding [70,71]. 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 [50,72], and the finding that their homologues are involved in zebrafish zygotic genome activation [15,23], 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 [26,27]. In contrast to the previously reported preference of Pou5f3 and Nanog to open up distal regions that are DNA methylated [26], 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 [27] 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 [27]. 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 [73,74], promotes chromatin accessibility and enhances the binding of Oct4, Sox2 and Nanog in embryonic stem cells [75], 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 [10,76,77], 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 show a correlation with transcription at later stages of zebrafish development [65,69].

Priming is strongest for genes 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 [53]. 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 [62,65]. Pou5f3, Sox19b and Nanog, as well as H3K27ac are essential for zygotic genome activation [15,22,65], and the highest H3K27ac enrichment occurs at Pou5f3-, Sox19b-, and Nanog-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 [26], similar to what has been reported in ES cells [78]. 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 [79] and its capability for reprogramming [79,80].

Materials and methods

Ethics statement

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; license number DD24.1-5131/354/92).

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 [83].

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 [84,85]. The following gRNA-s were used to delete the *sox19b* and *pou5f3* genes (the **NGG** PAM sequence indicated in bold):

N-terminal *sox19b*: GGGTTGCGTTGAGCGCTCATT**GG**

C-terminal *sox19b*: GGTGTCCAACAGCACTATCTT**GG**

N-terminal *pou5f3*: GGCCATGTATCCTCAAGCC**CGG**

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

gRNAs were cloned into the T7cas9gRNA2 vector [40] 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 [40] 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 [31], and MZ*pou5f3* embryos were obtained by incrossing homozygous *pou5f3* mutant fish.

To obtain MZ*nanog* embryos, homozygous zygotic *nanog* mutants from [48] were incrossed.

ATAC-seq

The ATAC-seq protocol was adapted from [86] and performed on approximately 70,000 cells. Zebrafish embryos were first deyolked using buffers as described in [87]; 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 [86]. 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 (S1 Fig), 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. 1 microgram of RNA was reverse transcribed to produce cDNA using the iScript cDNA Synthesis kit (Bio-Rad). For qPCR, a readily made SYBR green Master mix with ROX (100nM) was used. Primers for *fam212aa*, *sox19a*, *mxtx2* and the housekeeping gene *ef1g2a* were from [31].

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 ($q \geq 30$) 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 transcription 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` [23], `Nanog` [24] or `Pou5f1` [23] and one additional developmental transcription factor (`Smad2` [81], `FoxH1` [82], `Mxtx2` [24], `Eomesodermin` [81], `Ta`, `Tbx16` or `Mixl1` [88]).

Clustering of accessible regions

We calculated the average open chromatin fragments (fragment length ≤ 130 bp) 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 [89]. 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 [24] 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 [90]. This custom collection was then used in motif enrichment and comparative analyses.

Proximal (\pm 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—value-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.

De novo motif analysis for differentially accessible TFBS

We centered a 200bp window on summits of transcription factor binding sites for Pou5f1 (post ZGA) and Sox2 (post ZGA), counted ATAC-seq open chromatin fragments within these regions and called regions significantly affected in our MZ*pou5f3* and MZ*sox19b* mutants compared to wt condition (DESeq2, log₂ fold-change ≤ -1.5 , FDR $< 5\%$). We subsequently performed *de novo* motif analysis on sequences extracted from both significantly affected and unaffected (log₂ fold-change ≥ 0) regions using DREME (MEME-suite) with non-default parameters: *-m 5 -mink 3 -maxk 8*. For all significant motifs, we additionally calculated centralized enrichment relative to the peak summit using CENTRIMO (MEME-suite) and selected motifs with an e-value of 0.01 or less as centrally enriched.

Binned motif score analysis

We extracted canonical motifs for Pou5FB1 (MA0792.1), Sox2 (MA0143.3) and the Pou5f1--Sox2 double motif (MA0142.1) from the JASPAR 2018 database. We then performed low-stringency motif scanning using FIMO (MEME-suite) across 200bp windows, centered on TFBS summits (see previous section) using these motifs. We reported all motif hits at a p-value threshold of 0.05. Positions of significant motif matches were parsed into R and the highest scoring match for each sequence and motif was selected. For each motif, we divided the range of observed motif scores into 10 bins of approximately equal frequency (quantile binning) and plotted the distribution of associated accessibility fold-changes in mutant compared to wild-type conditions (violin plots).

Gene expression sets and transcriptional priming

We obtained gene expression data from [15] and ([16,24] 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] were 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 [35] as a control.

We additionally compared the accessibility fold-changes in our maternal-zygotic mutants per gene, stratified by whether the same gene's gene expression was negatively affected in mutants as reported before [15]. We tested the statistical significance of this difference by two-sample Welch t-tests and visualized the fold-change distributions using boxplots.

Supporting information

S1 Fig. 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.

(TIF)

S2 Fig. Pipeline for analyzing chromatin accessibility. Chromatin and naked genomic DNA from zebrafish embryos was tagged 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.

(TIF)

S3 Fig. 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), at promoters (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.

(TIF)

S4 Fig. 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. α -amanitin treatment results in transcription inhibition. 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.

(TIF)

S5 Fig. 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 [38,39] are marked by a red asterisk.

(TIF)

S6 Fig. Motif analysis for regions with increasing accessibility. Enrichment of TF motifs from the JASPAR vertebrates database (with the custom Nanog motif from [24]) at proximal (within +/- 1kb of TSS) and distal (enhancer/intergenic) 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.

(TIF)

S7 Fig. 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.

(TIF)

S8 Fig. Global change in chromatin accessibility in MZ*pou5f3*, MZ*sox19b* and MZ*nanog* mutants. Violin plots show the aggregated fold change in accessibility in MZ*pou5f3*, MZ*sox19b* and MZ*nanog* mutants compared to wild-type embryos at oblong stage. Significance of differences was tested using paired, one-sided, t-tests.

(TIF)

S9 Fig. Chromatin accessibility in mutants is most affected at regulatory regions of Pou5f3-, Sox19b- and Nanog-regulated genes. Boxplots show the log₂ fold change in accessibility in mutants compared to wild-type embryos (oblong stage) at putative enhancers and promoters associated with genes that are downregulated upon Pou5f3, Sox19b and Nanog knock-down (red) and genes that are not downregulated (grey). Expression data is from [15]. Statistical significance of the differences was tested by two-sample Welch t-tests.

(TIF)

S10 Fig. 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 MZ*pou5f3* 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 *pou5f3* (MZ*spg*) mutants compared to wild-type at dome stage, using data from [27]

(TIF)

S11 Fig. De novo motif analysis at loci with reduced chromatin accessibility in mutants identifies canonical motifs as most enriched. De novo motif analysis was performed for peaks that are a) bound by Pou5f3 and strongly reduced (blue) or unaffected (red) in MZ*pou5f3* mutants; b) bound by Sox1 and strongly reduced (blue) or unaffected (red) in MZ*sox19b* mutants; or c) bound by Nanog and strongly reduced (blue) or unaffected (red) in MZ*nanog* mutants. Sequence logos of significant (e-value < 0.01) motifs identified are shown along with their associated E-value.

(TIF)

S12 Fig. Regulation of accessibility by Pou5f3 and Sox19b is associated with high motif strength for these factors. To determine the correlation between loss of accessibility in mutants and motif strength, we used motifs from the JASPAR 2018 database to perform low-stringency motif scanning. Binned motif scores are visualized from low (left) to high (right). Violin plots show the distribution of associated fold-changes in accessibility between a) MZ*pou5f3* and wild-type for the Pou5f1-Sox2 double motif; b) MZ*sox19b* and wild-type for the Pou5f1-Sox2 double motif; c) MZ*pou5f3* and wild-type for the Pou5FB1 motif; and d) MZ*sox19b* and wild-type for the Sox2 motif.

(TIF)

S13 Fig. An increase in chromatin accessibility predicts future gene expression, especially at *Pou5f3/Sox19b/Nanog*-regulated regions. **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. p-values are shown for differences in expression between quintiles as assessed 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. p-values are shown for differences in expression between quintiles as assessed by one-sided Wilcoxon tests **c)** Heatmaps show the median expression value for genes associated with regulatory regions at shield stage. Genomic regions are resolved by 20% bins of accessibility increase between 256-cell and oblong stage (x-axis), and 20% bins of accessibility change in *MZpou5f3*, *MZsox19b* and *MZnanog* mutants compared to wild-type embryos at oblong stage (y-axis). (TIF)

Acknowledgments

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, three anonymous reviewers for constructive feedback, and the following facilities and services: Fish Facility (MPI-CBG) and the DRESDEN-concept Genome Center.

Author Contributions

Conceptualization: Máté Pálffy, Nadine L. Vastenhouw.

Data curation: Máté Pálffy, Gunnar Schulze.

Formal analysis: Gunnar Schulze.

Funding acquisition: Eivind Valen, Nadine L. Vastenhouw.

Investigation: Máté Pálffy, Gunnar Schulze.

Methodology: Máté Pálffy, Gunnar Schulze.

Project administration: Eivind Valen, Nadine L. Vastenhouw.

Resources: Máté Pálffy, Gunnar Schulze, Eivind Valen, Nadine L. Vastenhouw.

Software: Gunnar Schulze.

Supervision: Eivind Valen, Nadine L. Vastenhouw.

Validation: Máté Pálffy, Gunnar Schulze.

Visualization: Máté Pálffy, Gunnar Schulze.

Writing – original draft: Máté Pálffy, Nadine L. Vastenhouw.

Writing – review & editing: Máté Pálffy, Gunnar Schulze, Eivind Valen, Nadine L. Vastenhouw.

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. <https://doi.org/10.1186/gb-2011-12-4-r34> PMID: 21473766

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. <https://doi.org/10.1038/nature11232> PMID: 22955617
3. Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet*. 2019; 20: 207–220. <https://doi.org/10.1038/s41576-018-0089-8> PMID: 30675018
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: 652–657. <https://doi.org/10.1038/nature18606> PMID: 27309802
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. *Nat Commun*. 2018; 9:781 <https://doi.org/10.1038/s41467-018-03149-4> PMID: 29472610
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. <https://doi.org/10.1038/s41467-018-08244-0> PMID: 30664750
7. McKay DJ, Lieb JD. A common set of DNA regulatory elements shapes *Drosophila* appendages. *Dev Cell*. 2013; 27: 306–318. <https://doi.org/10.1016/j.devcel.2013.10.009> PMID: 24229644
8. Haines JE, Eisen MB. Patterns of chromatin accessibility along the anterior-posterior axis in the early *Drosophila* embryo. *PLoS Genet*. 2018; 14: e1007367. <https://doi.org/10.1371/journal.pgen.1007367> PMID: 29727464
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: 943–949. <https://doi.org/10.1126/science.1256271> PMID: 25103404
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: 1375–1388. <https://doi.org/10.1016/j.cell.2016.05.050> PMID: 27259149
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: 256–260. <https://doi.org/10.1038/s41586-018-0080-8> PMID: 29720659
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: 847–858. <https://doi.org/10.1038/s41556-018-0123-2> PMID: 29915357
14. Kane DA, Kimmel CB. The zebrafish midblastula transition. *Development*. 1993; 119: 447–456. PMID: 8287796
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: 360–364. <https://doi.org/10.1038/nature12632> PMID: 24056933
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. Pálffy M, Joseph SR, Vastenhouw NL. The timing of zygotic genome activation. *Curr Opin Genet Dev*. 2017; 43: 53–60. <https://doi.org/10.1016/j.gde.2016.12.001> PMID: 28088031
18. Vastenhouw NL, Cao WX, Lipshitz HD. The maternal-to-zygotic transition revisited. *Development*. 2019; 146.
19. 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: 285–292. <https://doi.org/10.1016/j.celrep.2013.12.030> PMID: 24440719
20. Hilbert L, Sato Y, Kimura H, Jülicher F, Honigsmann 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. <https://doi.org/10.1038/s41467-019-08487-5> PMID: 30741925
22. 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. <https://doi.org/10.1016/j.devcel.2019.05.037> PMID: 31211993
23. Leichsenring M, Maes J, Mössner R, Driever W, Onichtchouk D. Pou5f1 transcription factor controls zygotic gene activation in vertebrates. *Science*. 2013; 341: 1005–1009. <https://doi.org/10.1126/science.1242527> PMID: 23950494

24. 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. pp. 625–638. <https://doi.org/10.1016/j.devcel.2012.01.003> PMID: 22421047
25. Gagnon JA, Obbad K, Schier AF. The primary role of zebrafish nanog is in extra-embryonic tissue. *Development*. 2018; 145.
26. 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. <https://doi.org/10.1101/gr.228833.117> PMID: 29844026
27. 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. <https://doi.org/10.1101/gr.240572.118> PMID: 30674556
28. 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: 1213–1218. <https://doi.org/10.1038/nmeth.2688> PMID: 24097267
29. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008; 9: R137. <https://doi.org/10.1186/gb-2008-9-9-r137> PMID: 18798982
30. Kane DA, Hammerschmidt M, Mullins MC, Maischein HM, Brand M, van Eeden FJ, et al. The zebrafish epiboly mutants. *Development*. 1996; 123: 47–55. PMID: 9007228
31. Joseph SR, Pálffy 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.
32. 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: 541–552. [https://doi.org/10.1016/s1097-2765\(02\)00472-0](https://doi.org/10.1016/s1097-2765(02)00472-0) PMID: 11931762
33. 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: 11325–11330. <https://doi.org/10.1073/pnas.1001148107> PMID: 20534568
34. van den Berg AA, Depken M. Crowding-induced transcriptional bursts dictate polymerase and nucleosome density profiles along genes. *Nucleic Acids Res*. 2017; 45: 7623–7632. <https://doi.org/10.1093/nar/gkx513> PMID: 28586463
35. 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: 259–268. <https://doi.org/10.1101/gr.203679.115> PMID: 27965291
36. 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: 577–591. <https://doi.org/10.1101/gr.133009.111> PMID: 22110045
37. 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: 5627–5637. <https://doi.org/10.1242/dev.01455> PMID: 15496438
38. Harvey SA, Tümpel S, Dubrulle J, Schier AF, Smith JC. no tail integrates two modes of mesoderm induction. *Development*. 2010; 137: 1127–1135. <https://doi.org/10.1242/dev.046318> PMID: 20215349
39. 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. pp. 19350–19355.
40. 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: 13904–13909. <https://doi.org/10.1073/pnas.1308335110> PMID: 23918387
41. Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, et al. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature*. 2015; 524: 230–233. <https://doi.org/10.1038/nature14580> PMID: 26168398
42. 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: 193–197. <https://doi.org/10.1038/s41586-019-1064-z> PMID: 30944477
43. 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. <https://doi.org/10.1371/journal.pgen.1007105> PMID: 29161261

44. Reim G, Brand M. Maternal control of vertebrate dorsoventral axis formation and epiboly by the POU domain protein Spg/Pou2/Oct4. *Development*. 2006; 133: 2757–2770. <https://doi.org/10.1242/dev.02391> PMID: 16775002
45. 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. [https://doi.org/10.1016/s1534-5807\(03\)00396-4](https://doi.org/10.1016/s1534-5807(03)00396-4) PMID: 14723850
46. 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. <https://doi.org/10.1016/j.cub.2003.11.022> PMID: 14711414
47. 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. <https://doi.org/10.1371/journal.pgen.1000936> PMID: 20463883
48. 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.
49. Mayran A, Drouin J. Pioneer transcription factors shape the epigenetic landscape. *J Biol Chem*. 2018; 293: 13795–13804. <https://doi.org/10.1074/jbc.R117.001232> PMID: 29507097
50. 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: 555–568. <https://doi.org/10.1016/j.cell.2015.03.017> PMID: 25892221
51. Meers MP, Janssens DH, Henikoff S. Pioneer Factor-Nucleosome Binding Events during Differentiation Are Motif Encoded. *Mol Cell*. 2019; 75: 562–575.e5. <https://doi.org/10.1016/j.molcel.2019.05.025> PMID: 31253573
52. 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.
53. Gentsch GE, Spruce T, Owens NDL, Smith JC. Maternal pluripotency factors initiate extensive chromatin remodelling to predefine first response to inductive signals. *Nat Commun*. 2019; 10: 4269. <https://doi.org/10.1038/s41467-019-12263-w> PMID: 31537794
54. Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science*. 2015; 348: 910–914. <https://doi.org/10.1126/science.aab1601> PMID: 25953818
55. Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*. 2015; 523: 486–490. <https://doi.org/10.1038/nature14590> PMID: 26083756
56. Farrell JA, Wang Y, Riesenfeld SJ, Shekhar K, Regev A, Schier AF. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science*. 2018;360.
57. 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: 981–987. <https://doi.org/10.1126/science.aar4362> PMID: 29700229
58. Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferrerres R, et al. The cis-regulatory dynamics of embryonic development at single-cell resolution. *Nature*. 2018; 555: 538–542. <https://doi.org/10.1038/nature25981> PMID: 29539636
59. Lippok B, Song S, Driever W. Pou5f1 protein expression and posttranslational modification during early zebrafish development. *Dev Dyn*. 2014; 243: 468–477. <https://doi.org/10.1002/dvdy.24079> PMID: 24130110
60. 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: 2399–2408. <https://doi.org/10.1101/gr.138776.112> PMID: 23100115
61. 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.
62. Sato Y, Hilbert L, Oda H, Wan Y, Heddleston JM, Chew T-L, et al. Histone H3K27 acetylation precedes active transcription during zebrafish zygotic genome activation as revealed by live-cell analysis. *Development*. 2019; 146.
63. 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: 922–926. <https://doi.org/10.1038/nature08866> PMID: 20336069
64. 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. <https://doi.org/10.1016/j.devcel.2011.10.008> PMID: 22137762

65. 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: 673–686.e6. <https://doi.org/10.1016/j.molcel.2018.10.017> PMID: 30444999
66. 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. <https://doi.org/10.1016/j.cell.2018.01.022> PMID: 29456083
67. 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. <https://doi.org/10.1186/s13059-016-1013-1> PMID: 27381023
68. 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. <https://doi.org/10.1186/gb-2012-13-7-r65> PMID: 22830626
69. 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: 260–266. <https://doi.org/10.1101/gr.157750.113> PMID: 24285721
70. 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.
71. 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. <https://doi.org/10.1101/gr.192682.115> PMID: 26335634
72. Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell*. 2012; 151: 994–1004. <https://doi.org/10.1016/j.cell.2012.09.045> PMID: 23159369
73. 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.
74. Stanney W, Ladam F, Donaldson IJ, Parsons TJ. TALE and NF-Y co-occupancy marks enhancers of developmental control genes during zygotic genome activation in zebrafish. *bioRxiv*. 2019.
75. 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: 708–722. <https://doi.org/10.1016/j.molcel.2014.07.005> PMID: 25132174
76. 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.
77. 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: 858–871.e8. <https://doi.org/10.1016/j.molcel.2018.06.044> PMID: 30078726
78. 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.:
79. Esch D, Vahokoski J, Groves MR, Pogenberg V, Cojocar V, Vom Bruch H, et al. A unique Oct4 interface is crucial for reprogramming to pluripotency. *Nat Cell Biol*. 2013; 15: 295–301. <https://doi.org/10.1038/ncb2680> PMID: 23376973
80. 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. <https://doi.org/10.1038/ncomms2229> PMID: 23232409
81. Nelson AC, Cutty SJ, Niini M, Stemple DL, Flicke 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. <https://doi.org/10.1186/s12915-014-0081-5> PMID: 25277163
82. 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.
83. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. *Dev Dyn*. 1995; 203: 253–310. <https://doi.org/10.1002/aja.1002030302> PMID: 8589427
84. 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. <https://doi.org/10.1093/nar/gku410> PMID: 24861617
85. 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. <https://doi.org/10.1093/nar/gkw398> PMID: 27185894
86. 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.

87. Link V, Shevchenko A, Heisenberg C-P. Proteomics of early zebrafish embryos. *BMC Dev Biol.* 2006; 6: 1. <https://doi.org/10.1186/1471-213X-6-1> PMID: [16412219](https://pubmed.ncbi.nlm.nih.gov/16412219/)
88. 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. pp. 2782–2795. <https://doi.org/10.1016/j.celrep.2017.06.011> PMID: [28658625](https://pubmed.ncbi.nlm.nih.gov/28658625/)
89. 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. <https://doi.org/10.1093/nar/gkp335> PMID: [19458158](https://pubmed.ncbi.nlm.nih.gov/19458158/)
90. Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Research.* 2018. pp. D260–D266. <https://doi.org/10.1093/nar/gkx1126> PMID: [29140473](https://pubmed.ncbi.nlm.nih.gov/29140473/)