Activator-blocker model of transcriptional regulation by pioneer-like factors

Aileen Julia Riesle*[,1,8], Meijiang Gao*[,1], Marcus Rosenblatt*[,2,4], Jacques Hermes*[,2,4] Helge Hass[2,4], Anna Gebhard[1], Marina Veil[1], Björn Grüning[5,6], Jens Timmer**[,2,3,4], Daria Onichtchouk**[1,3,7]

***contributed equally**

[1]Department of Developmental Biology, Albert-Ludwigs-University of Freiburg, 79104 Freiburg, Germany

[2]Institute of Physics, Albert-Ludwigs-University of Freiburg, 79104 Freiburg, Germany

[3]Signalling Research centers BIOSS and CIBSS, 79104 Freiburg, Germany

[4]Freiburg Center for Data Analysis and Modelling (FDM), 79104 Freiburg, Germany

[5]Department of Computer Science, University of Freiburg, 79110 Freiburg, Germany

[6]Center for Biological Systems Analysis (ZBSA), University of Freiburg, 79104 Freiburg, Germany

[7]Institute of Developmental Biology RAS, 119991 Moscow, Russia

[8]Current address: Epigenetics and Neurobiology Unit, European Molecular Biology Laboratory, EMBL Rome, Adriano Buzzati-Traverso Campus, Via Ramarini 32, 00015 Monterotondo (RM), ltaly

**Correspondence to: daria.onichtchouk@biologie.uni-freiburg.de, jeti@fdm.uni-freiburg.de

Summary.

Zygotic genome activation (ZGA) in the development of flies, fish, frogs and mammals depends on pioneer-like transcription factors (TFs). Those TFs create open chromatin regions, promote histone acetylation on enhancers, and activate transcription. Here, we use the panel of single, double and triple mutants for zebrafish genome activators Pou5f3, Sox19b and Nanog, multi-omics and mathematical modeling to investigate the combinatorial mechanisms of genome activation. We show that Pou5f3 and Nanog act differently on synergistic and antagonistic enhancer types. Pou5f3 and Nanog both bind as pioneers on synergistic enhancers, promote histone acetylation and activate transcription. Antagonistic enhancers are activated by pioneer binding of one of these factors. The other TF binds as non-pioneer, competes with the activator and blocks all its effects, partially or completely. This activator-blocker mechanism mutually restricts widespread transcriptional activation by Pou5f3 and Nanog and prevents premature expression of late developmental regulators in the early embryo.

- Pou5f3 and Nanog compete as pioneer-like and non-pioneer factors on common genomic sites.
- Pioneer-like binding activates transcription, non-pioneer binding blocks transcription.
- Competition between Pou5f3 and Nanog establishes the order of gene expression.

Introduction.

Awakening of zygotic transcription during maternal-to-zygotic transition is a universal feature of multicellular organisms. The major wave of Zygotic Genome Activation (ZGA) is driven by sequence-specific transcription factors (TFs), different across species (Schulz and Harrison, 2018; Vastenhouw et al., 2019). Nucleosomes protect enhancers from binding of most TFs and are the main obstacles for transcriptional activation. The small group of “elite” TFs, called pioneer factors, directly bind to nucleosomes and create open, nucleosome-free chromatin regions on enhancers (Iwafuchi-Doi and Zaret, 2016). Zygotic genome activators act as pioneer-like factors, by opening chromatin and initiating the sequence of events which lead to enhancer activation and widespread zygotic gene expression (Gaskill et al., 2021; Miao et al., 2022). Direct nucleosome binding was shown for some of the genome activators (Fernandez Garcia et al., 2019; Gassler et al., 2022).

In animals, gene products synthetized at ZGA enable gastrulation and subdivide the embryo into the three germ layers, ectoderm, mesoderm and endoderm. The regulators of later developmental programs, organogenesis and cell lineage specification, are kept silent at ZGA. They will be synthesized shortly before the beginning of appropriate developmental stages (Domazet-Loso and Tautz, 2010; Heyn et al., 2014; Neyfakh, 1964). The expression waves of transcripts encoding cohorts of transcriptional regulators follow each other in precise order in the developmental time course (White et al., 2017). While the gene activation of the major ZGA is plausibly explained by widespread binding of zygotic genome activators to the genomic regulatory elements, molecular mechanisms which keep lineage specifying genes silent at ZGA are unknown.

Maternal transcription factors Pou5f3, Sox19b and Nanog (PSN) together activate zygotic gene expression in zebrafish (Lee et al., 2013; Leichsenring et al., 2013). Zebrafish genome activators are homologs of mammalian pluripotency factors, which play a central role in reprogramming and in the transition from pluripotency to differentiation (Onichtchouk and Driever, 2016). Sox19b is the only maternally expressed member of the SoxB1 family of transcription factors. Shortly after ZGA, early zygotic SoxB1 factors Sox3, Sox19a and Sox2 are expressed and act redundantly with Sox19b (Okuda et al., 2010). Pou5f3 and Nanog are homeodomain transcription factors, with Pou5f3 also possessing the second POU DNA-binding domain. Pou5f3 and Nanog recognize overlapping yet not identical sets of sequence-specific motifs (Veil et al., 2019). SoxB1 bind *sox* motifs, and co-bind composite *sox:pou* motifs together with Pou5f3 (Gao et al., 2022; Leichsenring et al., 2013).

PSN, as well as their mammalian counterparts, frequently colocalize on the same genomic sites; therefore, they initially were thought to cooperate (Boyer et al., 2005; Leichsenring et al., 2013). However, recent studies in mouse and fish revealed that the factors may act alone, additively or interchangeably (Friman et al., 2019; Gao et al., 2022; Miao et al., 2022). The mechanisms of how PSN act together are still under debate. PSN open chromatin on more than half of active enhancers at ZGA, but are

necessary for expression of only a fraction of early zygotic genes (Miao et al., 2022). It is unclear how widespread pioneer-like activity of PSN on enhancers relates to early zygotic transcription. We had previously observed that late developmental regulators are prematurely activated in the zebrafish mutants by Pou5f3 or in the double mutants by Pou5f3 and Sox19b (Gao et al., 2022; Onichtchouk et al., 2010); molecular mechanisms underlying this phenomenon remained unclear.

The purpose of this study was to understand the mechanisms of combined Pou5f3, Sox19b and Nanog action, and the connection between PSN pioneer activity, enhancer activation and selection of the early zygotic gene expression repertoire. We systematically investigated the chromatin state in the single, double and triple maternal-zygotic null mutants by Pou5f3, Sox19b and Nanog using ATAC-seq. Further, we characterized time-resolved transcription in all these mutants and used mathematical modeling to predict differentially regulated target genes. We found that Pou5f3 and Nanog in combination not only activate part of zygotic transcripts, but reciprocally restrict or block transcriptional activation by each other on hundreds of developmental enhancers.

Results.

Study design and overview of the ATAC-seq data set

For multi-omic assays, we used the panel of eight zebrafish genotypes: zebrafish wild-type embryos (WT), maternal-zygotic mutants for Pou5f3 (P, MZ*spg*), Sox19b (S, MZ*sox19b*), and Nanog (N, MZ*nanog*), double mutants MZ*ps*, MZ*pn,* MZ*sn,* and the triple mutant MZ*triple*. Sox19b was dispensable for normal development, the absence of Pou5f3, Pou5f3/Sox19b or Nanog resulted in abnormal gastrulation as previously reported (Gao et al., 2022; Lunde et al., 2004; Veil et al., 2018). At the absence of Pou5f3/Nanog, Sox19b/Nanog or all three factors zygotic development was arrested: epiboly and gastrulation movements did not start (Fig. 1A).

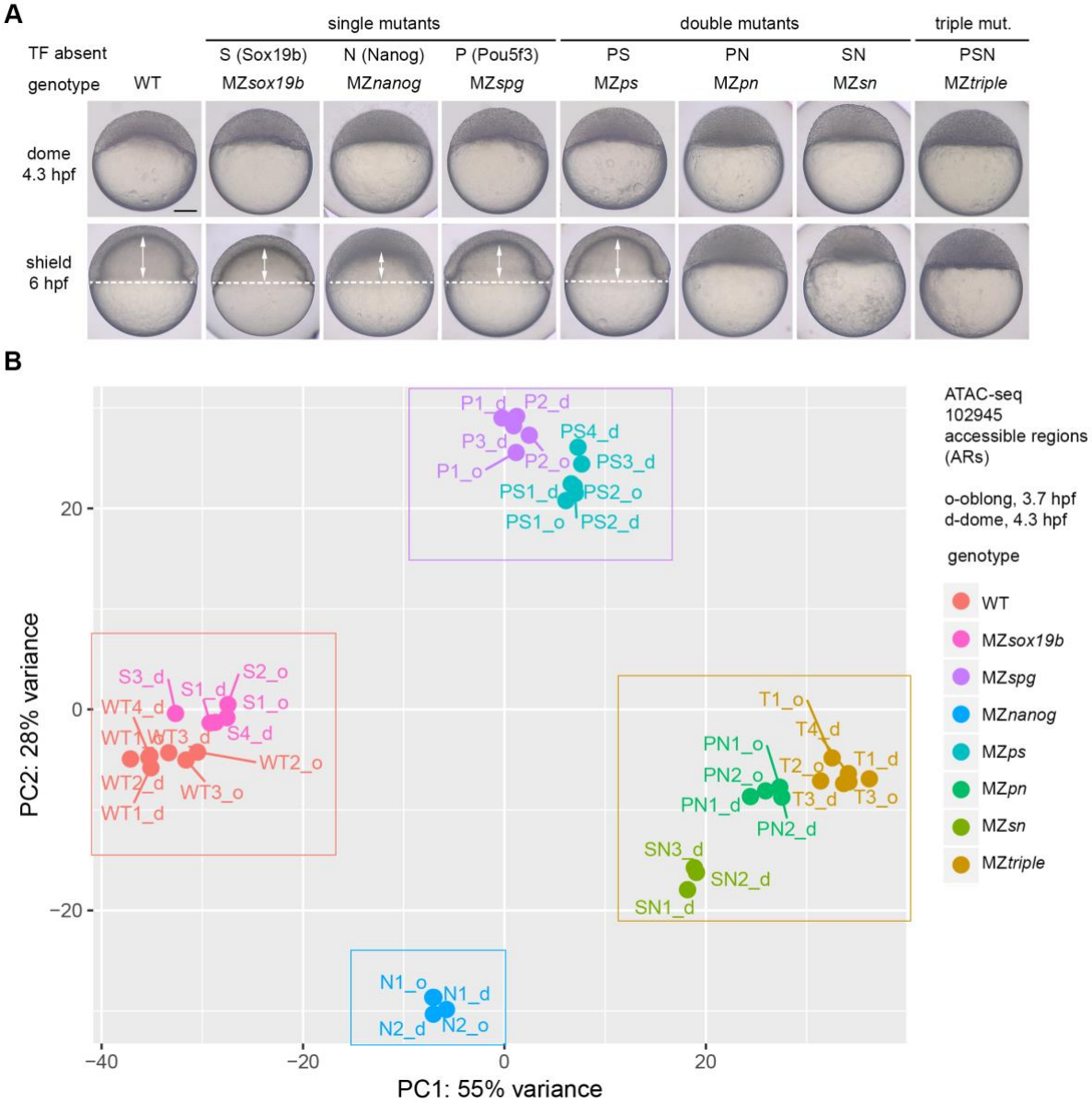

Figure 1. Single and combined mutants by zebrafish zygotic genome activators used in this study.

A. Mutant phenotypes at blastula (dome) and midgastrula (shield) stages. White dotted line shows epiboly border; double arrows show the internalized yolk. Scale bar=100µm. **B.** Principal Component Analysis of ATAC-seq data in all mutants, on the

genomic regions accessible in the wild-type (ARs). The data in eight genotypes clustered into four groups (boxed).

We used this comprehensive panel of mutants and the wild-type embryos to assess individual and combined effects of Pou5f3, Sox19b, and Nanog on enhancer regulation and transcription. We performed Omni ATAC-seq in all genotypes in two late blastula stages: 3.7 hpf (hours post-fertilization), and 4.3 hpf. We also performed RNA-seq profiling in all genotypes at eight time points, every 30 min starting from pre-ZGA (2.5 hpf) till mid-gastrulation (6 hpf), in at least two biological replicates per genotype. The outline of data analysis is shown in Fig. S1. In brief, we aimed to derive direct enhancer groups differentially regulated by PSN using genomic data (Table S1), and differentially regulated target gene groups using transcriptome data (Table S2). Finally, for cross-validation, we put together the independently obtained results of the genome and transcriptome parts (Table S3).

We started the genomic analysis by estimating the differences in the chromatin accessibility between the genotypes, on 102 945 Accessible Regions (ARs) in the wild-type. We used Principal Component Analysis (PCA, Fig. 1B), to estimate how different the chromatin profiles are, and made three initial observations. First, PCA revealed a little change in chromatin accessibility in MZ*sox19b* mutants to wild-type, as compared to all other mutants (Fig.1B, red box). Sox19b is the major SoxB1 factor expressed at 3.7 and 4.3 hpf, we therefore assumed that the SoxB1 family proteins act not alone, but as cofactors or redundantly with Pou5f3 or Nanog. Second, chromatin profiles of MZ*sn,* MZ*pn,* and MZ*triple* mutants clustered together in PCA space (Fig. 1B, brown box), corresponding to the most severe phenotypes of these mutants. Third, chromatin accessibility profiles in MZ*nanog* were quite different from MZ*spg* and MZ*ps* (Fig. 1B, blue and violet boxes). This suggested that Pou5f3 and Nanog change chromatin accessibility on different regions, or on the same regions but in the opposite directions.

Analysis of MZ*triple* mutant suggests the presence of additional genome activators in early zebrafish embryo

To tease apart the apparently complex regulation of chromatin accessibility by Pou5f3, SoxB1 and Nanog, we first characterized the effects of their combined activity on chromatin using MZ*triple* mutant. We divided all ARs to three groups: “down”, where the chromatin accessibility was reduced in MZ*triple* compared to the wild-type (31% of all ARs), “up”, where the chromatin accessibility was increased (15% of all ARs), and “same” –unchanged (Fig. 2A). Sequence-specific binding motifs for all three factors were enriched in “down” regions, as expected from the previous observation that direct binding of PSN opens chromatin (Palfy et al., 2020; Veil et al., 2019). “Up” regions had the highest content of G and C nucleotides and of GC-rich TF-binding motifs (Fig. 2B, C). The genes closest to “down” and “up” ARs were both enriched in transcriptional regulatory functions among the others (Fig. 2D).

Next, we analyzed the time-resolved transcriptome in MZ*triple* mutant compared to the wild-type, using our previously developed R package RNA-sense (www.bioconductor.org/packages/release/bioc/html/RNAsense.html, (Gao et al., 2022)). We selected 4777 transcripts zygotically expressed in the wild-type or in MZ*triple,* and split them into three groups (Fig. 2E): “DOWN” (downregulated compared to the wild-type or not expressed in MZ*triple,* 37% of all zygotic transcripts), “SAME” (unchanged, 46%), and “UP” (upregulated in MZ*triple* compared to the wild-type, or expressed only in MZ*triple,* 17%).

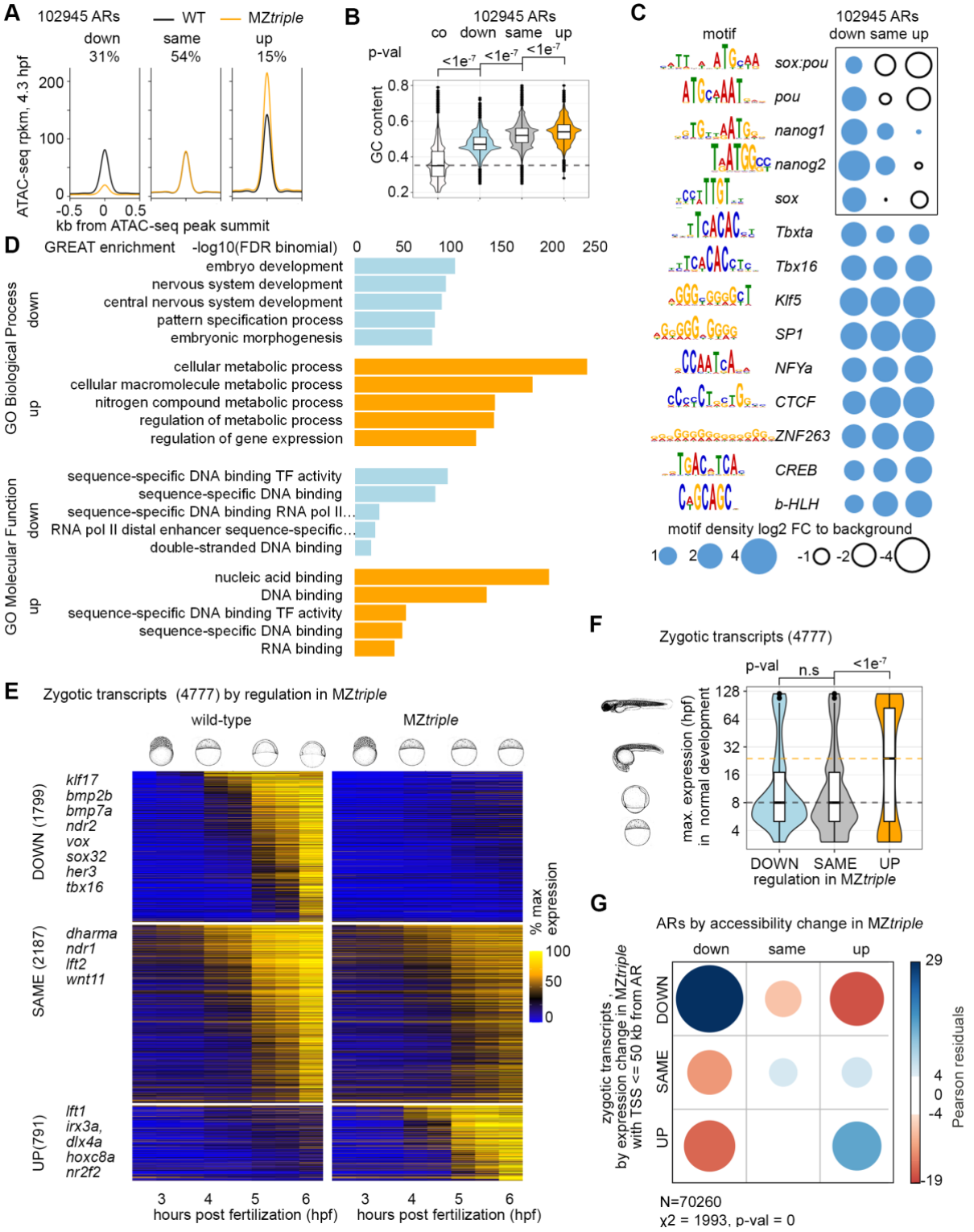

Figure 2. Changes in chromatin accessibility parallel the changes in gene expression in Pou5f3, Sox19b and Nanog triple mutant. A. Three groups of accessible regions (ARs) by chromatin accessibility change in the MZ*triple:* "down", "same", "up". **B.** ARs of "up" group had the highest GC content. P-values for Tukey-Kramer test are shown; p-value for 1-way ANOVA was $<2e^{-16}$. Co – control non-accessible genomic regions (dotted line). **C.** Pou5f3, Sox and Nanog-binding motifs are enriched in "down" ARs (black rectangle). **D.** GREAT analysis. Note that "down" and "up" ARs are both enriched around DNA-binding proteins. **E.** Three groups of zygotic transcripts by expression change in MZ*triple*, from total 4777 zygotic transcripts in WT or MZ*triple*. The heatmap shows normalized expression at eight time points,

from pre-ZGA (2.5 hpf) till 6 hpf; example developmental genes at the left. **F.** Group of zygotic genes is prematurely expressed in MZ*triple*. Y-axis: time of maximal expression in the normal development (3 hpf - 120 hpf), schematic embryo drawings illustrate the stages. Median max. expression time of transcripts upregulated in MZ*triple* was 24 hpf (yellow dotted line), versus 8 hpf for down- or unchanged transcripts (gray dotted line). P-values in Tukey-Kramer test are shown; p-value in 1-way ANOVA was $<2e^{-16}$. **G.** Down- or upregulation of chromatin accessibility in MZ*triple* correlates with respective transcriptional changes of linked genes. χ^2 test; positive correlations are shown in blue and negative in red.

Out of 791 "UP" group transcripts, 67% were zygotically expressed only in MZ*triple* but not in the wild-type (i.e. *hoxc8a*, *nr2f2*). We wondered if regulatory genes, promoting differentiation, were prematurely expressed MZ*triple,* as we observed previously in MZ*spg* and MZ*ps* mutants (Gao et al., 2022; Onichtchouk et al., 2010). To test that, we used published data (White et al., 2017) to calculate the maximal expression time during normal zebrafish development for each zygotic gene and compared the groups. We found that zygotic genes were indeed prematurely expressed in MZ*triple:* the median expression time in the "UP" group was 24 hpf, versus 8 hpf in the groups "DOWN" and "SAME" (Fig. 2F). Similar analysis for other mutants revealed that except for MZ*sox19b,* all of them prematurely expressed diverse sets of late regulatory genes (Fig. S2).

Finally, we checked if the changes in chromatin accessibility on the putative regulatory regions of zygotic genes correlated with the changes in their expression in MZ*triple*. This was indeed the case: "down" ARs were enriched around the promoters of "DOWN" genes, and "up" ARs around "UP" genes in MZ*triple* (Fig. 2G).

Taken together, our analysis of MZ*triple* mutants confirmed previous findings that Pou5f3, Sox19b and Nanog pioneer chromatin accessibility on the enhancers of early zygotic genes (Miao et al., 2022; Palfy et al., 2020; Veil et al., 2019). The increase in transcription in MZ*triple* mutants was unexpected. We concluded that additional activators were present at ZGA. These unknown activators could bind to GC-rich regulatory elements.

Four types of regulation of chromatin accessibility by Pou5f3, SoxB1 and Nanog. To investigate how Pou5f3, SoxB1 and Nanog create the regions of open chromatin, we restricted our analysis to the ARs, where any combination of the three factors was required for chromatin accessibility (group "down" in MZ*triple,* Fig. 2A). To study only direct TF effects, we further selected 20131 TdARs (TF-bound Accessible Regions, downregulated in the MZ*triple* mutant), overlapping with ChIP-seq peaks for Pou5f3, SoxB1 (Leichsenring et al., 2013), or Nanog (Xu et al., 2012).
We then classified TdARs by non-redundant requirements for Pou5f3 and Nanog for chromatin accessibility into four groups 1.PN, 2.P, 3.N, and 4.- (Fig. 3A). Accessibility of the 1.PN group TdARs was reduced in MZ*spg* and MZ*nanog*, so we concluded that Pou5f3 and Nanog were both required for accessibility in these regions. Pou5f3 was required in the 2.P group, Nanog in the 3.N group, and none of the single factors was required in the 4.- group (meaning that several factors were required redundantly).

To further characterize these groups, we cross-compared motif density, TF occupancy, dynamics of chromatin opening and GC content (Fig. 3B-E). Out of the four groups, 2.P and 3.N were the most enriched for the Pou5f3- and Nanog – motifs, and the most

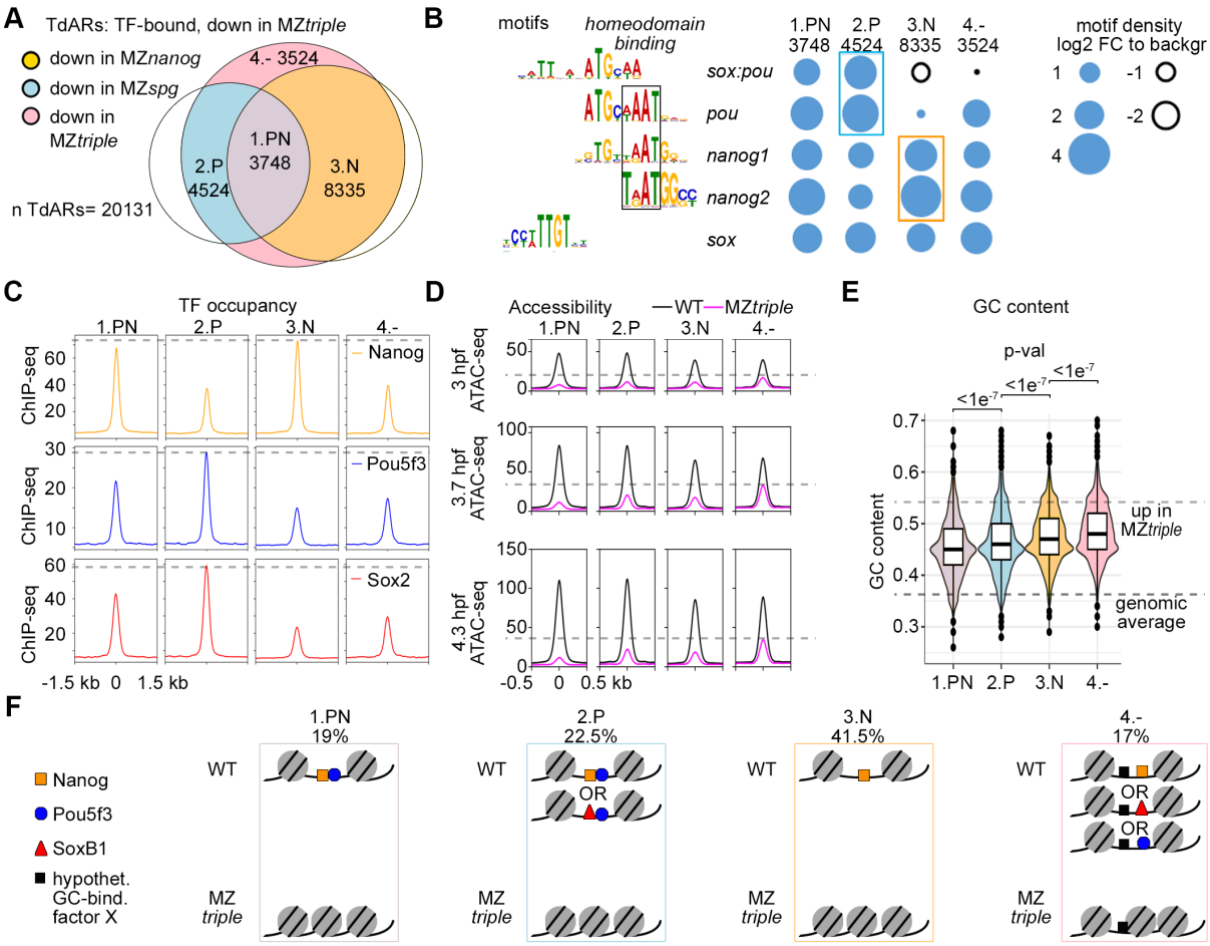

Figure 3. Four types of TdARs by pioneer-like activity of Pou5f3, SoxB1, and Nanog. A. Venn diagram. TdARs (TF-bound ARs, downregulated in MZ*triple),* divided into 4 groups by non-redundant requirements for Pou5f3 and/or Nanog for chromatin accessibility. **B.** Frequencies of PSN motifs in the 4 groups. Black box - common TAAT sequence, recognized by homeodomains. Note that 2.P and 3.N groups are the most enriched in the Pou5f3- and Nanog- specific motifs, respectively (colored boxes). **C.** 2.P and 3.N groups are the mostly occupied by Pou5f3/SoxB1 and Nanog, respectively (gray dashed lines). ChIP-seq (rkpm) summary profiles for indicated TFs. **D.** PSN establish accessible chromatin of all 4 TdAR groups starting from major ZGA (3 hpf). ATAC-seq summary profiles in WT and MZ*triple* at 3, 3.7 and 4.3 hpf. Note that chromatin in group 4.- regions in MZ*triple* is more accessible than in the other groups (gray dashed line). ATAC-seq signal in rpkm. **E.** 4.- group has the highest GC content out of all groups. P-values for Tukey-Kramer test are shown; p-value for 1-way ANOVA was $<2e^{-16}$. Lower dashed line shows median genomic control GC content, upper line – median GC content of ARs upregulated in MZ*triple*. **F.** Schematic drawing of four types of regulation of chromatin accessibility ny PSN binding. percentage of each group from all TdARs is shown on top. "OR" – logical operator.

highly occupied by Pou5f3 and Nanog, respectively (Fig. 3B, C). To address if chromatin becomes accessible in all groups at the same time, we performed Omni-ATAC-seq at major ZGA (3 hpf), and compared ATAC-seq signals at 3, 3.7 and 4.3 hpf, in the wild-type and MZ*triple*. While all groups depended on PSN already at ZGA, chromatin in 4.- group was more accessible in MZ*triple* in all stages than on the other groups (Fig. 3D, dashed lines). Therefore, it was conceivable that factors other than PSN also contribute to chromatin accessibility in the group 4.-. We compared GC

content in the TdAR groups and found that GC content increased in order 1.PN < 2.P < 3.N < 4.-, (Fig. 3E). As our analysis of MZ*triple* mutants suggested that GC-binding activators are present at ZGA, we hypothesized that GC-binding factor X may be involved in opening chromatin in group 4.- regions (Fig. 3F).

We next wondered whether the factors required for chromatin accessibility in a given group of TdARs were also sufficient to open chromatin. We used rescue experiments of (Miao et al., 2022) to answer this question. To address which factors can restore chromatin accessibility in the triple Pou5f3, Sox19b, and Nanog mutants MZ*nps*, the authors microinjected Pou5f3, Sox19b, and Nanog mRNAs, individually and in combinations, and compared ATAC-seq signals in the injected and non-injected MZ*nps* embryos. Using their data, we scored the chromatin accessibility rescue in each of four groups. Nanog alone rescued more than 90% of TdARs in the 3.N group. Pou5f3 alone rescued only 26% of TdARs in the 2.P group; most of the rest could be rescued by the Pou5f3/Sox19b or Pou5f3/Nanog combinations (Fig. S3A-F). We concluded that Nanog alone was required and sufficient for establishing chromatin accessibility in the majority of its target regions. In contrast, Pou5f3 needed assistance of either Nanog or SoxB1 to establish accessible chromatin in the majority of its target regions.

In agreement with this conclusion, in 59% of TdARs in the 2.P group chromatin accessibility was downregulated in the double Sox19b/Nanog mutant MZ*sn* (Fig. S3I, J). This explained the molecular similarity of MZ*sn* and MZ*triple* (Fig.1 A,B): although Pou5f3 was present in MZ*sn*, it was apparently not sufficient to open chromatin in most of its target regions.

It was suggested that Pou5f3, Sox19b and Nanog pioneer activity is facilitated by nucleosomes, as three factors can replace for each other on the nucleosome-rich sites (Miao et al., 2022). In agreement, we found that the percentages of non-specific rescue (with more than one single TF) and predicted nucleosome occupancy were the highest in 4.- group TdARs (Fig. S3G, H). It was also in line with the highest GC content of the 4.- group (Fig. 3E), as nucleosomes prefer GC-rich regions (Kaplan et al., 2009).

Fig. 3F summarizes four types of establishment of chromatin accessibility by pioneer-like factors. Pou5f3 and Nanog are both required on the 1.PN group TdARs. Pou5f3 and redundant contribution of Nanog or SoxB1 are required on most of the 2.P group TdARs. Nanog is required on most of the 3.N group TdARs. Pou5f3, SoxB1 or Nanog are redundantly required on 4.- group TdARs, together with a hypothetical GC-binding factor X.

Pou5f3 and Nanog act as Activator and Blocker and vice versa on a fraction of common enhancers

Pou5f3, SoxB1, and Nanog together promote histone acetylation on their binding sites, which is critical for the enhancer activation (Miao et al., 2022). However, individual roles of the TFs remain unclear. We used H3K27ac ChIP-seq data in the single mutants ((Gao et al., 2022) and this work) and MZ*nps* triple mutant (Miao et al., 2022) to analyze how H3K27 acetylation relates to the pioneer-like activity of the factors.

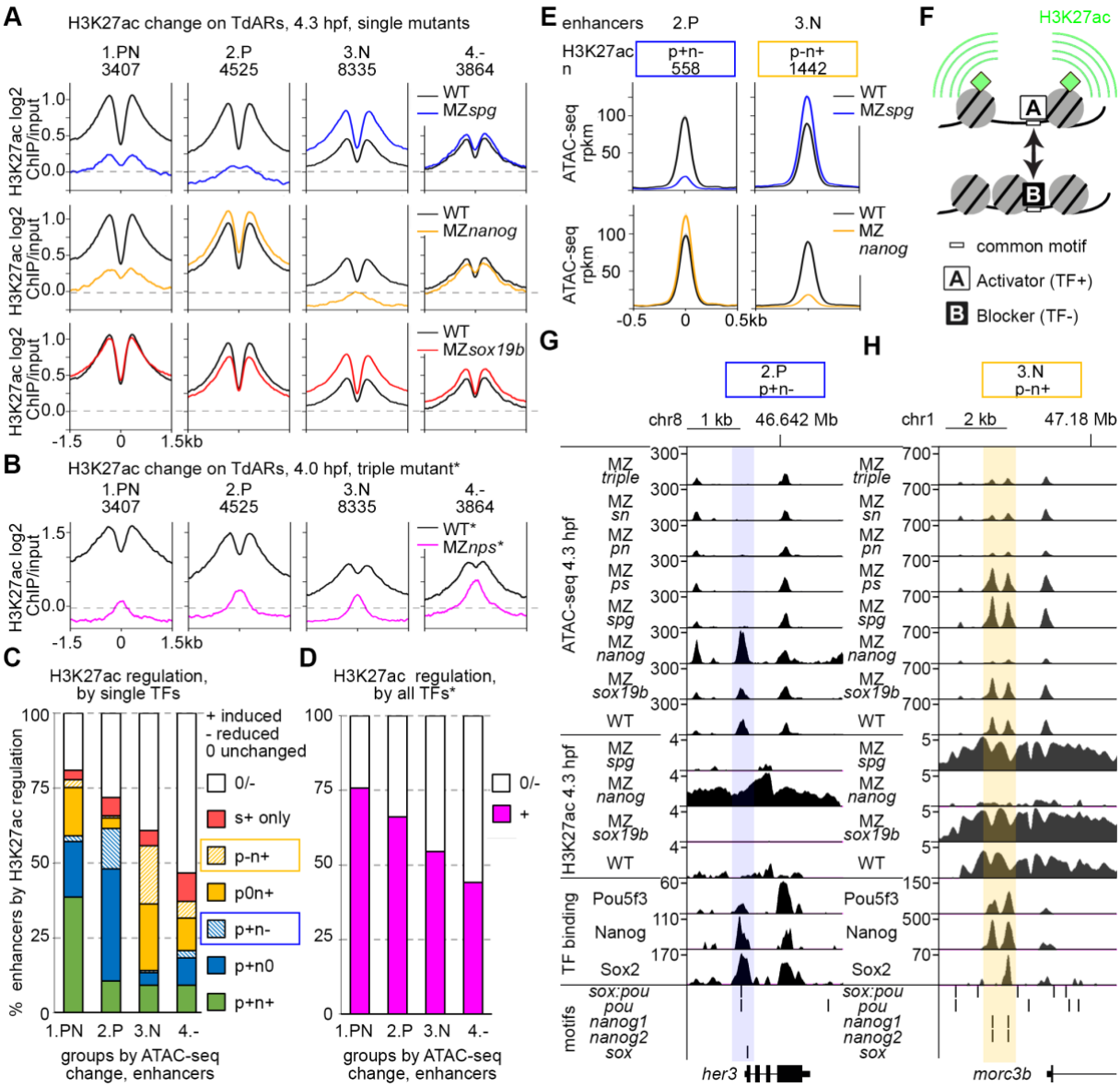

Figure 4. Activator-blocker model: Pou5f3 and Nanog oppose each other effects on H3K27 acetylation and chromatin accessibility on a fraction of enhancers.

A. Summary profiles of H3K27ac histone mark in the wild-type and single mutants, in four types of TdARs: 1.PN, 2.P, 3.N, 4.-. Note the opposite effects of Pou5f3 and Nanog on H3K27 acetylation on the 2.P and the 3.N groups. **B.** Summary profiles of H3K27ac mark in the WT and triple mutant MZ*nps*. The PSN combined activity is required for H3K27ac in all four TdAR groups. **C.** Effects of the single TFs on H3K27 acetylation of 1.PN, 2.P, 3.N, and 4.-. groups of enhancers. Percentages of (+) enhancers where at least one TF is required for H3K27ac acetylation are shown in colors indicated in the legend, antagonistic enhancers are boxed. **D.** Effects of PSN combined activity on H3K27ac. **E.** Pou5f3 and Nanog have the opposite effects on chromatin accessibility on antagonistic enhancers. ATAC-seq summary profiles of 2.P p+n- and 3.N p-n+ groups. **F.** Schematic illustration of activator-blocker model (explanations in the text). **(G,H)** UCSC Genomic browser views show (from bottom to top) motif occurrence, TF binding, H3K27ac and ATAC-seq in the indicated genotypes. (G) All three TFs bind to *sox:pou* motif on *her3* enhancer (blue shading). Pou5f3/Sox19b act as Activators (pioneer-like) and Nanog as Blocker (non-pioneer). (H) Pou5f3 and Nanog, or all three factors bind to two *nanog* motifs on *morc3b*

enhancer (yellow shading). Nanog acts as an Activator (pioneer-like) and Pou5f3 as Blocker (non-pioneer). In B, D * -data from (Miao et al., 2022).

We found that the same TF which was strictly required for chromatin accessibility was required for H3K27 acetylation: both Pou5f3 and Nanog in the 1.PN group, Pou5f3 in the 2.P group, Nanog in the 3.N group, and none of the single factors in the 4.- group. Unexpectedly, single mutant analysis revealed reciprocal antagonistic relationships between Pou5f3/Sox19b and Nanog: Nanog reduced H3K27ac in the 2.P group, and Pou5f3 reduced H3K27ac in the 3.N group (Fig. 4A). The PSN combined activity was required for H3K27 acetylation in all groups (Fig. 4B).

To further investigate the individual roles of the TFs, we selected H3K27-acetylated regions from the four TdAR groups and defined them as enhancers (STAR methods, Table S1). We then scored the enhancers on which H3K27ac was differentially regulated by each TF (+: induced, 0: unchanged, or -: reduced,Fig. 4C), or by the PSN combined activity (Fig. 4D). We noted that the percentage of enhancers activated by PSN was the highest in the 1.PN and the lowest in the 4.- group and thus inversely correlated with GC content (Fig. 4C, D, Fig.3E). Importantly, our analysis uncovered a large number of previously uncharacterized “antagonistic enhancers”, where Pou5f3 and Nanog regulated H3K27ac in the opposite directions (2504 p+n- and p-n+ enhancers in sum, boxed in Fig. 4C).

To understand how per se transcriptional activators and pioneer-like factors could negatively regulate enhancer activity, we took a closer look on chromatin accessibility changes on the 2.P p+n- and the 3.N p-n+ antagonistic enhancers (Fig. 4E). We found that the factor which reduced H3K27ac also reduced chromatin accessibility with high statistical significance (Fig. S4 A-D, F). We confirmed this result using MNase-seq nucleosome positioning data set (Veil et al., 2019, Fig. S4E, F). Together, we have shown by two independent methods that only one of the TFs, either Pou5f3 or Nanog, acts as pioneer-like factor on the antagonistic enhancers. The second TF reduces chromatin accessibility, perhaps by competing for common genomic site with the first TF and binding as a non-pioneer.

Our next question was whether the positive and negative effects of Pou5f3 and Nanog on chromatin accessibility correlate with sequence features of the bound sites. As shown in Fig. S4G, Pou5f3 induced chromatin accessibility on its own motifs and reduced it on Nanog motifs *nanog1* and *nanog2*. Vice versa, Nanog reduced chromatin accessibility on Pou5f3/SoxB1 and Pou5f3 motifs *sox:pou* and *pou.* This observation suggested that Pou5f3 and Nanog alternate their binding mode, acting as pioneer-like factors on some genomic sites and as non-pioneers on the others. Indeed, while common homeodomain-binding sequence in Pou5f3 and Nanog cognate motifs (Fig. 3B) is recognized by both TF, exact match to the motif is necessary for the pioneer-like activity.

Summarizing our findings, we suggest the general mechanism of antagonistic interactions between Pou5f3 and Nanog, as shown on Fig. 4F. Two TFs, A (activator) and B (blocker) compete for binding on the common motif within an enhancer. The activator acts as a pioneer-like factor on this motif: it displaces nucleosomes and promotes H3K27 acetylation. Non-pioneer binding of the blocker protects the motif from the the activator, thereby reducing nucleosome displacement and H3K27 acetylation. Two opposite examples illustrate the model. On *her3* regulatory region,

Pou5f3 together with SoxB1 act as an activator and Nanog as a blocker on *sox:pou* motif (Fig. 4G). On *morc3b* regulatory region, Nanog acts as an activator and Pou5f3 as a blocker on two nearby *nanog* motifs (Fig. 4H). Hence, Pou5f3 and Nanog compete as pioneer-like and non-pioneer factors on common genomic sites, eliciting opposite effects on enhancer activity.

Mathematical modelling of transcriptional regulation by genome activators

We wondered next, whether the antagonistic enhancers are relevant for zygotic gene expression. To answer this question, we chose the strategy to on the one hand sort zygotic transcripts into groups according to their regulation, without involving genomic information. On the other hand, we linked enhancers to transcripts and determined correlations between the enhancer and transcript groups.

To sort the transcriptome, we used a mathematical modeling approach in two subsequent steps. In the first step, we developed a “core model” to describe the dynamic behavior of the known core components of the system: Pou5f3, Nanog, and SoxB1 group genes (including Sox19b, Sox19a, Sox3 and Sox2). In the second step, dynamics of these core components were then used to inform ODE-based mini models to formalize all possible regulatory inputs from Pou5f3, Nanog or SoxB1 on the individual target genes.

The core model was based on ordinary differential equations (ODEs), where initial assumptions on the regulatory interactions between the core components were taken from the literature (STAR methods, Fig. S5A for model cartoon and Table S4 for model equations). To infer the model’s parameters, we used the measured time-resolved transcriptional profiles for *pou5f3*, *nanog*, *sox19b*, *sox19a*, *sox3* and *sox2*. After model calibration, the model trajectories resulting from the best fit parameters were in good agreement with the experimental data (Fig. S5B, C, Fig. S6A, and Table S5).

For the second step of ODE-based mini models, we assumed that at least one factor of Pou5f3, Nanog or SoxB1 should activate the target, while the other two can activate (+), repress (-) or do nothing (0). This gives 19 possible regulatory combinations. Each of the 19 mini models was separately fitted to experimental time-resolved RNA-seq data for each of the 1799 zygotic transcripts that had been found to be directly or indirectly activated by Pou5f3, Sox19b or Nanog (“DOWN” group in Fig. 2E, Table S6). The dynamics of Pou5f3, Nanog, and SoxB1 group genes (SOX) obtained in step one of the analysis were used as input to the mini models, while the remaining parameters were again calibrated based on the RNA-seq data. To simplify the analysis with respect to our goal of connecting chromatin state and transcription for Pou5f3 and Nanog, we merged the 19 mini-models down to six mini-model groups (Fig. 5A) capturing all possible variants of target regulations by Pou5f3 and Nanog (targets 1. P+N+, 2a. P+N0, 2b. P+N-, 3a. P0N+, 3b. P-N+) and the regulation by SOXB1 alone (4. S+ only).

For each transcript, we selected the best candidate mini-model group according to Bayesian Information Criterion (BIC, see Table S7 and Table S2 for selected mini-model groups). Fits of selected mini-model groups are shown for exemplary targets in Fig. 5B and Fig. S6B. 1799 transcripts were then classified into six groups according to the best fitting direct regulatory scenario (Fig.5C).

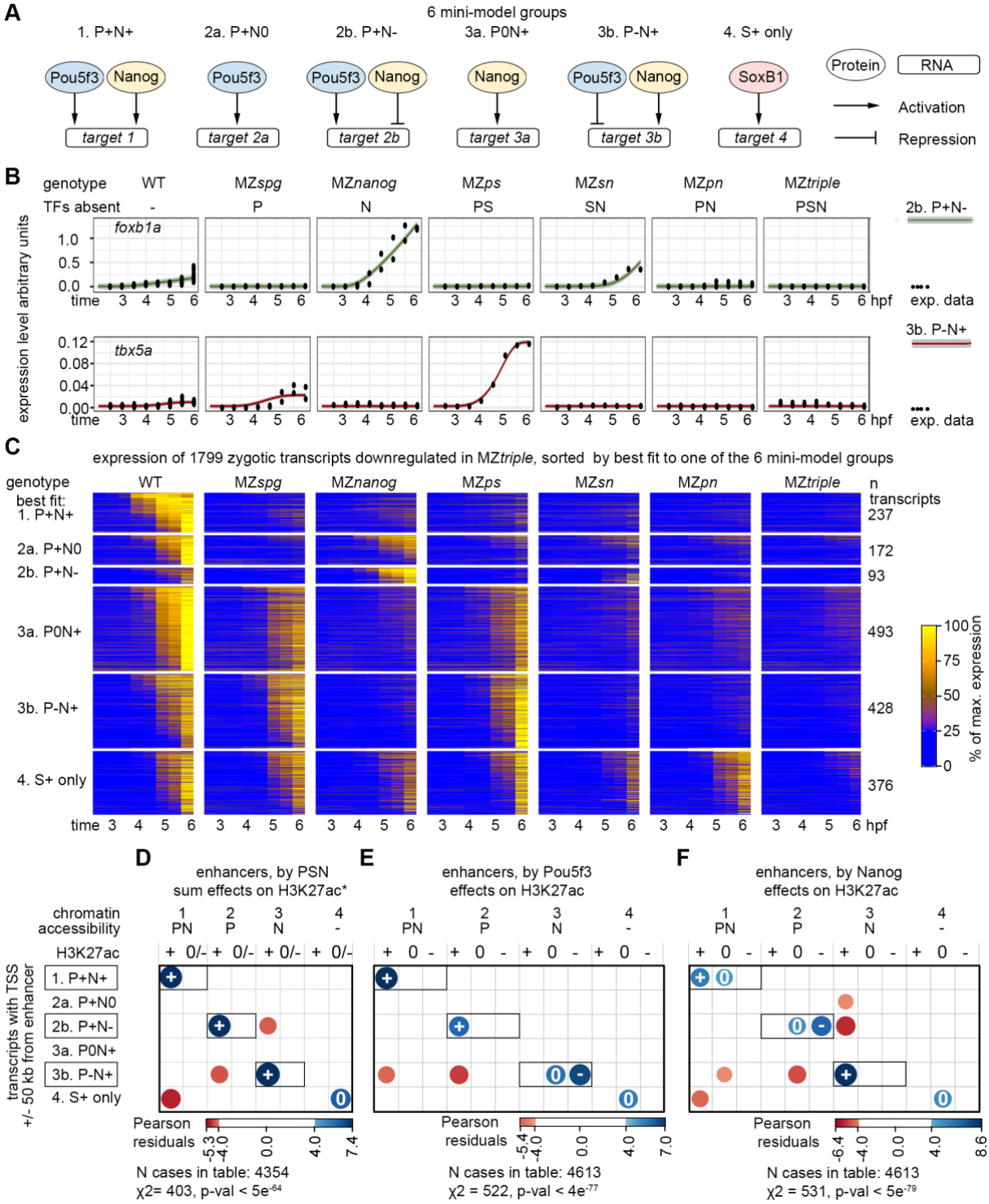

Figure 5. Zygotic gene expression is balanced by synergy and competition of Pou5f3 and Nanog on common enhancers. A. Schemes of six alternative mini-model groups for target regulation by Pou5f3, Nanog and SoxB1. **B**. *foxb1a* and *tbx5a* transcripts fitted best to the antagonistic groups 2b. P+N- and 3b. P-N+ respectively. **C.** Heatmap of 1799 zygotic genes**,** downregulated in MZ*triple***,** sorted by best fit to one of the groups shown in A. **D-F**. χ^2 tests, positive correlations between the transcriptomics and genomics groups are shown in blue and negative in red. Vertical axis: transcriptomics groups. Direct enhancers were linked to the promoters of zygotic genes within +/-50 kb and sorted according to best fit transcriptional model. The synergistically regulated targets 1. P+N+ and antagonistically regulated targets 2b. P+N- and 3b. P-N+ are boxed. Horizontal axis: direct enhancers were sorted to four

groups (1.PN, 2.P, 3.N and 4.-) by chromatin accessibility, and by H3K27ac regulation by sum of PSN activities (D, *- H3K27ac data from (Miao et al., 2022)), by Pou5f3 (E) or by Nanog (F). “+”,”0” and “–“ indicate that TF induces H3K27ac, does nothing, or blocks it, respectively; “0/-“ - merged group. Note that the enhancers where H3K27ac is induced by one factor and blocked by the other (2.P p+n- and 3.N p-n+) are enriched around TSS of targets activated by one factor and repressed by the other, respectively.

Interestingly, three out of six groups of targets significantly associated with gene ontology terms. Groups 1. P+N+ and 3b. P-N+ were enriched in the regulators of transcription and development, while the largest group 3a. P0N+ was enriched in the housekeeping genes (Fig. S7A). In sum, we sorted PSN-dependent zygotic transcripts to six groups with biologically relevant functions, using formal and unbiased mathematical modelling approach.

Zygotic gene expression is balanced by synergistic and antagonistic effects of Pou5f3 and Nanog on common enhancers.

To inquire if synergistic and antagonistic chromatin regulation by Pou5f3 and Nanog on common enhancers is directly linked to synergistic and antagonistic regulation of transcription, we put together the results of genomic and transcriptomic analyses. As shown in Fig. 5D-F, the predicted direction of transcriptional regulation of targets by Pou5f3 and Nanog (activation or repression) strongly correlated with Pou5f3 and Nanog-dependent H3K27ac changes on their enhancers (activation or blocking), respectively.

Notably, although redundant enhancers of the 4.- group were specifically enriched in the putative regulatory regions of the transcriptional targets fitting to the 4. (Sox+ only) regulatory scenario, the deposition H3K27ac mark on these enhancers did not depend on PSN (Fig. 5D-F, Fig. S7B, C). Considering the high GC content of group 4.- enhancers (Fig. 3E), we assumed that PSN modulate the pioneer-like activity of hypothetical GC-binding factor X, which activates transcription of the 4. “Sox+ only” group (Fig. 3F). Taken together, using cross-validation of independent genomic and transcriptomic data sets we demonstrated that the transcription of PSN target genes is directly regulated by synergistic and antagonistic enhancers.

Pou5f3 and Nanog competition on antagonistic enhancers prevents expression of late developmental genes in the early embryo.

We have noticed that all mutants except MZ*sox19b* prematurely transcribe poorly overlapping sets of genes, enriched for DNA-binding and transcription regulatory functions (Fig. S2). We also demonstrated that Pou5f3 and Nanog compete on antagonistic enhancers as activator and blocker of chromatin accessibility, H3K27 acetylation, and transcription. Putting these two findings together, we wondered again whether Pou5f3 could directly block premature transcriptional activation by Nanog and vice versa. To answer this question, we examined putative regulatory regions of the transcripts, upregulated in the Pou5f3 and Nanog mutants, for the enrichment of antagonistic enhancers of the opposite types.

In Pou5f3 mutant MZ*spg*, 17% of transcripts were upregulated compared to the wild-type (Fig. 6A, Table S2); the group of upregulated transcripts contained prematurely expressed genes (Fig. 6B).

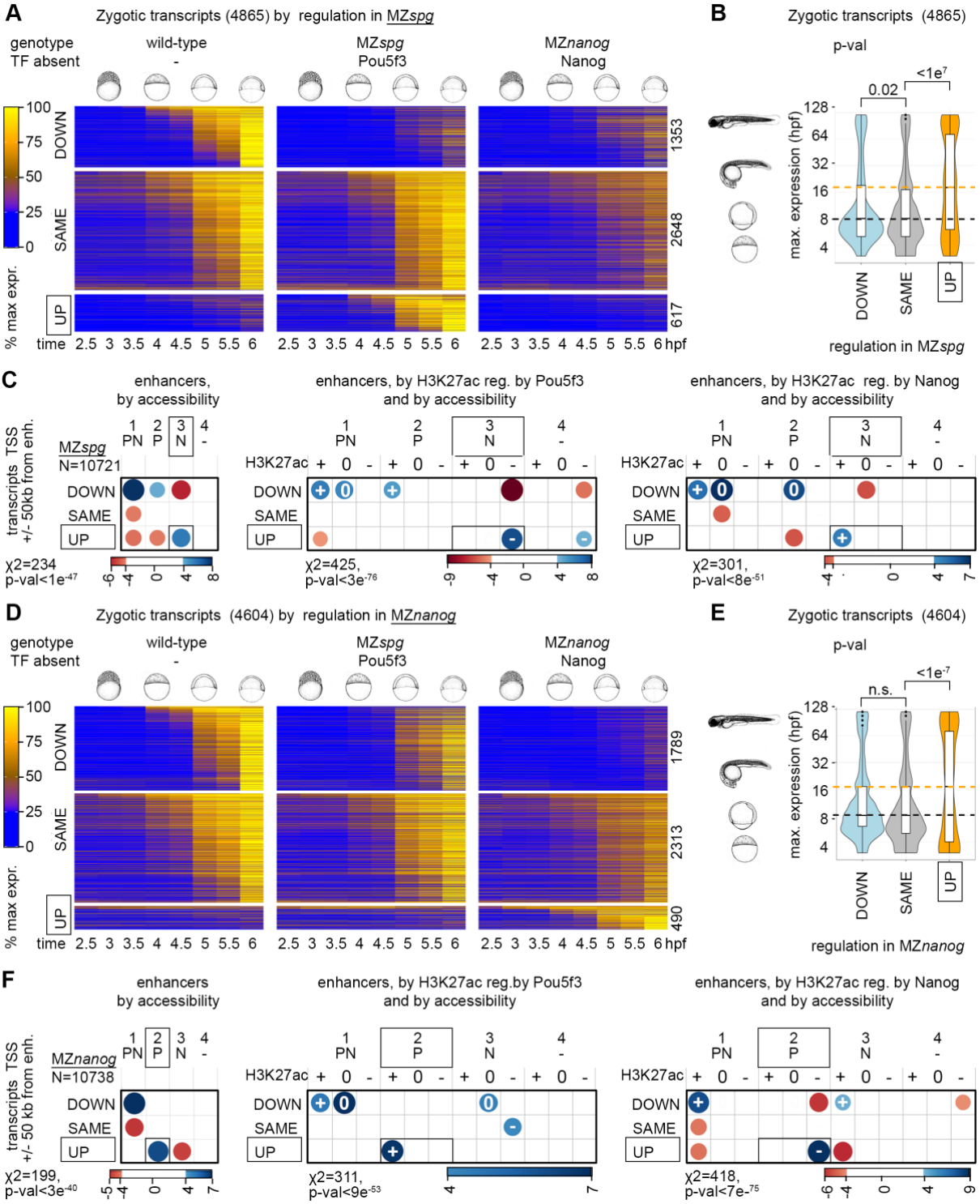

Figure 6. Pou5f3 blocks premature transcriptional activation by Nanog and vice versa. (**A-C**). Pou5f3 blocks transcription from Nanog-dependent enhancers. (**D-F**). Nanog blocks transcription from Pou5f3-dependent enhancers. **A, D.** Zygotic transcripts were sorted to "DOWN", "SAME" and "UP" groups, by expression change in MZ*spg* (A) or MZ*nanog* (D) compared to the wild-type. Heatmaps show normalized expression at eight time points, from pre-ZGA (2.5 hpf) till 6 hpf; numbers of transcripts in each group are shown at the right. **B, E**. Premature expression of zygotic genes in MZ*spg* (B) or in MZ*nanog* (E). Median expression time in the normal development was compared in "UP" "DOWN" and "SAME" groups. P-values in Tukey-Kramer test; p-value in 1-way ANOVA was $<2e^{-16}$. **C, F.** χ^2 tests, positive correlations between the transcriptomics and genomics groups are shown in blue and negative in red, scales

show Pearson residuals. Vertical axis: transcriptomics groups. Direct enhancers were linked to the promoters of zygotic genes within +/-50 kb and sorted according to DOWN", "SAME" and "UP" transcriptional groups. Horizontal axis: genomic groups. Direct enhancers were sorted by changes in chromatin accessibility, and by H3K27ac regulation by Pou5f3 or Nanog, as indicated.

Strikingly, the putative regulatory regions of these genes were highly enriched for antagonistic 3.N p-n+ enhancers (Fig. 6C, boxed). 3.N p-n+ enhancers were also enriched in the putative regulatory regions of genes upregulated in the double Pou5f3/Sox19b mutant MZ*ps* (Fig. S7D-F). Thus, Pou5f3 blocked inappropriate transcriptional activation of late developmental genes by Nanog.

To test if the reverse is also true, we examined the transcripts upregulated in MZ*nanog.* 10% of transcripts were upregulated compared to the wild-type in MZ*nanog*; the group of upregulated transcripts contained prematurely expressed genes (Fig. 6D, E). The putative regulatory regions of these genes were highly enriched for antagonistic 2.P p+n- type enhancers (Fig. 6F, boxed). Thus, Nanog blocked inappropriate transcriptional activation of late developmental genes by Pou5f3.

Discussion

Our study provides the mechanistic links between the pioneer-like activity of genome activators Pou5f3, Sox19b (together with zygotic SoxB1 factors) and Nanog, and their potential to activate transcription at ZGA and during the first half of gastrulation. We demonstrate that direct binding of PSN to genome is not only responsible for transcriptional activation at ZGA, but also for adjusting the levels of early zygotic transcripts and for the timing of gene expression in development. We distinguish three types of PSN-dependent direct enhancers, as summarized at Fig.7.

For the synergistic 1.PN p+n+ enhancers, Pou5f3 and Nanog are both required to establish chromatin accessibility and promote H3K27 acetylation. Synergistic enhancers are enriched around the early zygotic genes, activated by Pou5f3 and Nanog (Fig. 7A).

For the antagonistic 2.P p+n- enhancers (Fig. 7B), Pou5f3 pioneer-like binding, often with the assistance of SoxB1, creates accessible chromatin and promotes H3K27 acetylation. Nanog competes with Pou5f3 on the common binding sites, binds as a non-pioneer and blocks enhancer activation. 2.P p+n- enhancers are enriched around the early zygotic genes activated by Pou5f3 and attenuated by Nanog, and also around the lineage-specific regulators blocked by Nanog (Fig. 7B).

For the antagonistic 3.N p-n+ enhancers, Nanog pioneer binding creates accessible chromatin and promotes H3K27 acetylation. Pou5f3 competes with Nanog on the common binding sites, binds as non-pioneer and blocks enhancer activation. 3.N p-n+ enhancers are enriched around the early zygotic genes activated by Nanog and attenuated by Pou5f3, and also around the lineage-specific regulators blocked by Pou5f3 or Pou5f3/SoxB1 (Fig. 7C).

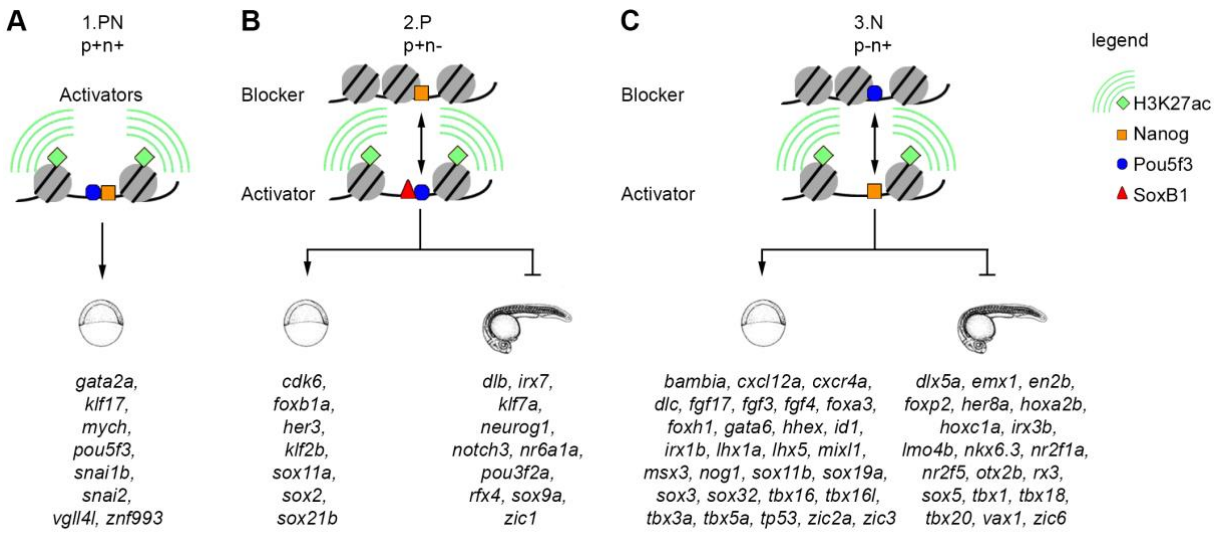

Figure 7. Synergistic and antagonistic Pou5f3/Nanog direct enhancers: the mechanism of action and direct target genes.
A. Synergistic enhancers 1.PN p+n+: Pou5f3 and Nanog act as activators. **B.** Antagonistic enhancers 2.P p+n-: Pou5f3 acts as an activator, Nanog as a blocker. **C.** Antagonistic enhancers 3.N p-n+: Nanog acts as an activator, Pou5f3 as a blocker. Representative early targets (expressed in the wild-type during 2.5-6 hpf time course), and late targets (prematurely expressed in the mutants by blocker TF during 2.5-6 hpf time course) are indicated by schematic early and late embryo drawings.

Synergistic interactions were long ago suggested for zebrafish genome activators and their mammalian homologs (Boyer et al., 2005; Leichsenring et al., 2013). Antagonistic interactions of Pou5f3 and Nanog on chromatin are a new finding. On synergistic enhancers, both proteins bind as pioneer-like factors. On antagonistic enhancers, “activator” and “blocker” compete for binding to the same sequence-specific motif, but only “activator” binds as a pioneer-like factor (Fig. 4F).

A reasonable question is which features define pioneer- and non-pioneer binding in each case. We found a correlation with two sequence features. One feature is GC content around TF binding motifs: on synergistic enhancers, GC content is lower than on the other enhancer types (Fig. 3E, Fig.4C, 1.PN group). Another feature is the different frequency of sequence-specific motifs for ”activator” and “blocker” - on antagonistic enhancers. Pou5f3- and Nanog- sequence-specific motifs all contain homeodomain-binding site TAAT, which can be recognized by both TFs (Fig. 3B). We found that a given TF acts more frequently as activator on its own motifs, and as blocker as it binds to the motifs of the other factor (Fig. 4 F-H, Fig. S4G). In a sense, “activator” and “blocker” TFs can be compared with a precise key and a key blank to the locked door (which is an enhancer): both keys fit to the lock, but only a precise key can open the door (activate an enhancer).

The correlating sequence features alone are definitely not enough to predict how and whether TFs will regulate H3K27 acetylation on each given enhancer. In whole embryo assays, we could not distinguish whether the ATAC-seq signals come from the whole embryo or from the subpopulation of cells. Therefore, we cannot exclude that cell-specific differences, such as local concentrations of Pou5f3, SoxB1 and Nanog, presence of transcriptional cofactors or cross-talk with signaling pathways affect the binding mode cell type-specifically. In line with that, it was shown that nucleosome displacement activity of known mammalian pioneer factors, including Oct4/Pou5f1, is

conditional: all of them bind to different genomic locations in different cell types (Donaghey et al., 2018). Further, it is also possible that antagonistic enhancers act as bifunctional cis-regulatory elements, i.e. as activators or as silencers, depending on the embryonic cell type (Segert et al., 2021). Single cell multi-omics and reporter assays are needed to clarify this issue.

The activator-blocker competitive model easily fits to the picture of discontinuous transcriptional regulation emerging from recent molecular imaging studies. Transcription occurs in bursts, and bursting frequencies are regulated by developmental enhancers (Fukaya et al., 2016; Suter et al., 2011). Transcription factors occupy their binding sites only for short residence times (Chen et al., 2014; Dufourt et al., 2018; Huang et al., 2021; Kuznetsova et al., 2022). An increase of the residence time of transcriptional activators on enhancers was shown to increase transcription bursting frequency in yeast and mammalian cell culture (Donovan et al., 2019; Popp et al., 2021). The residence time of the activator on the antagonistic enhancer would be likely reduced upon the presence of the competing blocker, and may become too short to trigger transcriptional bursting. For Pou5f3 and Nanog, it will be interesting to measure whether the residence time of one factor depends on the presence of the other one, as well as to compare transcriptional outputs of synergistic and antagonistic enhancers in the presence of only Pou5f3, only Nanog, or both TFs.

The dual role of Pou5f3 and Nanog in activating early zygotic genes and blocking premature transcription of lineage specifiers provides an immediate parallel to the biological roles of their mammalian homologues. Oct4/Pou5f1 and Nanog are expressed in tightly controlled levels in mammalian embryos and ES cells, where they regulate the seemingly opposite processes of pluripotency maintenance and cell lineage decisions (Mulas et al., 2018; Niwa et al., 2000; Radzisheuskaya and Silva, 2014; Tiana et al., 2022; Torres-Padilla and Chambers, 2014). Although zebrafish Pou5f3 is not equivalent to mammalian pioneer factor Oct4/Pou5f1 in terms of nucleosome-binding and reprogramming capacity (Esch et al., 2013), Pou5f3 and Oct4 recognize the same motifs and colocalize with Nanog in the respective systems. It is tempting to speculate that mammalian Pou5f1 and Nanog may reciprocally block some of each other activities in a direct way, similarly to their zebrafish homologues. Recent study supports this possibility: acute depletion of Oct4 in ES cells results in increased genomic binding of Nanog, suggesting that Oct4 outcompetes Nanog on chromatin and antagonizes its function (Bates et al., 2021).

Our results suggest that the list of zebrafish genome activators is not complete: hundreds of transcripts, enriched in transcriptional regulatory functions, are prematurely expressed in MZ*triple* mutant and linked to regulatory elements with high GC content (Fig. 2). Identification of putative GC-binding factors regulating these genes and dissecting their interplay with Pou5f3, SoxB1 and Nanog will further clarify the mechanisms of gene regulation at ZGA. Finally, the concepts presented in this work will assist the studies of combinatorial genome activation by TFs in other organisms and deepen our understanding of two global and currently unresolved questions: conditional activity of pioneer and pioneer-like factors in different biological settings, and of how the gene expression is regulated in developmental time.

Acknowledgments

We are grateful to Dr. Rainer Duden for commenting on the manuscript, to Sabine Götter for excellent fish care, and to Andrea Buderer and Cornelia Wagner for

administrative support. D.O is grateful to Wolfgang Driever for hosting her lab. This work was supported by DFG-ON86/4-2 and DFG-ON86/6-1 for DO, and DFG-EXC2189 – Project ID: 390939984 for D.O. and J.T. BG is funded by DFG grant SFB 992/1 2012 and BMBF grant 031 A538A RBC.

Author Contributions

AJR, MG, MV, and AG - experiments; AJR, MG, BG and DO - data analysis, MR and JH -mathematical modeling, DO - design of the study, supervision of the biological part, JT – supervision of the modeling part. DO wrote the manuscript, all authors edited the manuscript.

Declarations of Interests

The authors declare no competing interests

Main figure titles and legends

Figure 1. Single and combined mutants by zebrafish zygotic genome activators used in this study.
A. Mutant phenotypes at blastula (dome) and midgastrula (shield) stages. White dotted line shows epiboly border; double arrows show the internalized yolk. Scale bar=100µm. **B.** Principal Component Analysis of ATAC-seq data in all mutants, on the genomic regions accessible in the wild-type (ARs). The data in eight genotypes clustered into four groups (boxed).

Figure 2. Changes in chromatin accessibility parallel the changes in gene expression in Pou5f3, Sox19b and Nanog triple mutant.
A. Three groups of accessible regions (ARs) by chromatin accessibility change in the MZ*triple:* "down", "same", "up". **B.** ARs of "up" group had the highest GC content. P-values for Tukey-Kramer test are shown; p-value for 1-way ANOVA was $<2e^{-16}$. Co – control non-accessible genomic regions (dotted line). **C.** Pou5f3, Sox and Nanog-binding motifs (black rectangle) are enriched in "down" ARs. **D.** GREAT analysis. Note that "down" and "up" ARs are both enriched around DNA-binding proteins. **E.** Three groups of zygotic transcripts by expression change in MZ*triple,* from total 4777 zygotic transcripts in WT or MZ*triple*. The heatmap shows normalized expression at eight time points, from pre-ZGA (2.5 hpf) till 6 hpf; example developmental genes at the left. **F.** Group of zygotic genes is prematurely expressed in MZ*triple*. Y-axis: time of maximal expression in the normal development (3 hpf - 120 hpf), schematic embryo drawings illustrate the stages. Median max. expression time of transcripts upregulated in MZ*triple* was 24 hpf (yellow dotted line), versus 8 hpf for down- or unchanged transcripts (gray dotted line). P-values in Tukey-Kramer test are shown; p-value in 1-way ANOVA was $<2e^{-16}$. **G.** Down- or upregulation of chromatin accessibility in MZ*triple* correlates with respective transcriptional changes of linked genes. χ^2 test; positive correlations are shown in blue and negative in red.

Figure 3. Four types of TdARs by pioneer-like activity of Pou5f3, SoxB1, and Nanog.
A. Venn diagram. TdARs (TF-bound ARs, downregulated in MZ*triple),* divided into 4 groups by non-redundant requirements for Pou5f3 and/or Nanog for chromatin accessibility. **B.** Frequencies of PSN motifs in the 4 groups. Black box - common TAAT sequence, recognized by homeodomains. Note that 2.P and 3.N groups are the most enriched in the Pou5f3- and Nanog- specific motifs, respectively (colored boxes). **C.** 2.P and 3.N groups are the mostly occupied by Pou5f3/SoxB1 and Nanog, respectively (gray dashed lines). ChIP-seq (rkpm) summary profiles for indicated TFs. **D.** PSN establish accessible chromatin of all 4 TdAR groups starting from major ZGA (3 hpf). ATAC-seq summary profiles in WT and MZ*triple* at 3, 3.7 and 4.3 hpf. Note that chromatin in group 4.- regions in MZ*triple* is more accessible than in the other groups (gray dashed line). ATAC-seq signal in rpkm. **E.** 4.- group has the highest GC content out of all groups. P-values for Tukey-Kramer test are shown; p-value for 1-way ANOVA was $<2e^{-16}$. Lower dashed line shows median genomic control GC content, upper line – median GC content of ARs upregulated in MZ*triple*. **F.** Schematic drawing of four types of regulation of chromatin accessibility ny PSN binding. percentage of each group from all TdARs is shown on top. "OR" – logical operator.

Figure 4. Activator-blocker model: Pou5f3 and Nanog oppose each other effects on H3K27 acetylation and chromatin accessibility on a fraction of enhancers.
A. Summary profiles of H3K27ac histone mark in the wild-type and single mutants, in four types of TdARs: 1.PN, 2.P, 3.N, 4.-. Note the opposite effects of Pou5f3 and Nanog on H3K27 acetylation on the 2.P and the 3.N groups. **B.** Summary profiles of H3K27ac mark in the WT and triple mutant MZ*nps*.The PSN combined activity is required for H3K27ac in all four TdAR groups. **C.** Effects of the single TFs on H3K27 acetylation of 1.PN, 2.P, 3.N, and 4.-. groups of enhancers. Percentages of (+) enhancers where at least one TF is required for H3K27ac acetylation are shown in colors indicated in the legend, antagonistic enhancers are boxed. **D.** Effects of PSN combined activity on H3K27ac. **E.** Pou5f3 and Nanog have the opposite effects on chromatin accessibility on antagonistic enhancers. ATAC-seq summary profiles of 2.P p+n- and 3.N p-n+ groups. **F.** Schematic illustration of activator-blocker model (explanations in the text). (**G,H)** UCSC Genomic browser views show (from bottom to top) motif occurrence, TF binding, H3K27ac and ATAC-seq in the indicated genotypes. (G) All three TFs bind to *sox*:*pou* motif on *her3* enhancer (blue shading). Pou5f3/Sox19b act as Activators (pioneer-like) and Nanog as Blocker (non-pioneer). (H) Pou5f3 and Nanog, or all three factors bind to two *nanog* motifs on *morc3b* enhancer (yellow shading). Nanog acts as an Activator (pioneer-like) and Pou5f3 as Blocker (non-pioneer). In B, D * -data from (Miao et al., 2022).

Figure 5. Zygotic gene expression is balanced by synergy and competition of Pou5f3 and Nanog on common enhancers.
A. Schemes of six alternative mini-model groups for target regulation by Pou5f3, Nanog and SoxB1. **B**. *foxb1a* and *tbx5a* transcripts fitted best to the antagonistic groups 2b. P+N- and 3b. P-N+ respectively. **C.** Heatmap of 1799 zygotic genes**,** downregulated in MZ*triple***,** sorted by best fit to one of the groups shown in A. **D-F**. χ^2 tests, positive correlations between the transcriptomics and genomics groups are shown in blue and negative in red. Vertical axis: transcriptomics groups. Direct enhancers were linked to the promoters of zygotic genes within +/-50 kb and sorted according to best fit transcriptional model. The synergistically regulated targets 1. P+N+ and antagonistically regulated targets 2b. P+N- and 3b. P-N+ are boxed. Horizontal axis: direct enhancers were sorted to four groups (1.PN, 2.P, 3.N and 4.-) by chromatin accessibility, and by H3K27ac regulation by sum of PSN activities (D, *-H3K27ac data from (Miao et al., 2022)), by Pou5f3 (E) or by Nanog (F). “+”,”0” and “–“ indicate that TF induces H3K27ac, does nothing, or blocks it, respectively; “0/-“ - merged group. Note that the enhancers where H3K27ac is induced by one factor and blocked by the other (2.P p+n-, and 3.N p-n+) are enriched around TSS of targets activated by one factor and repressed by the other, respectively.

Figure 6. Pou5f3 blocks premature transcriptional activation by Nanog and vice versa.
(**A-C**). Pou5f3 blocks transcription from Nanog-dependent enhancers. (**D-F**). Nanog blocks transcription from Pou5f3-dependent enhancers. **A, D.** Zygotic transcripts were sorted to “DOWN”, “SAME” and “UP” groups, by expression change in MZ*spg* (A) or MZ*nanog* (D) compared to the wild-type. Heatmaps show normalized expression at eight time points, from pre-ZGA (2.5 hpf) till 6 hpf; numbers of transcripts in each group are shown at the right. **B, E**. Premature expression of zygotic genes in MZ*spg* (B) or in MZ*nanog* (E). Median expression time in the normal development was compared in “UP” “DOWN” and “SAME” groups. P-values in Tukey-Kramer test; p-value in 1-way ANOVA was $<2e^{-16}$. **C, F.** χ^2 tests, positive correlations between the transcriptomics

and genomics groups are shown in blue and negative in red, scales show Pearson residuals. Vertical axis: transcriptomics groups. Direct enhancers were linked to the promoters of zygotic genes within +/-50 kb and sorted according to DOWN", "SAME" and "UP" transcriptional groups. Horizontal axis: genomic groups. Direct enhancers were sorted by changes in chromatin accessibility, and by H3K27ac regulation by Pou5f3 or Nanog, as indicated.

Figure 7. Synergistic and antagonistic Pou5f3/Nanog direct enhancers: the mechanism of action and direct target genes.
A. Synergistic enhancers 1.PN p+n+: Pou5f3 and Nanog act as activators. **B.** Antagonistic enhancers 2.P p+n-: Pou5f3 acts as an activator, Nanog as a blocker. **C.** Antagonistic enhancers 3.N p-n+: Nanog acts as an activator, Pou5f3 as a blocker. Representative early targets (expressed in the wild-type during 2.5-6 hpf time course), and late targets (prematurely expressed in the mutants by blocker TF during 2.5-6 hpf time course) are indicated by schematic early and late embryo drawings.

Supplemental figure titles (including “related to” info) and legends

Figure S1. Study design and outline of the data analysis (related to all figures)
A. Genomic analysis. We performed Omni ATAC-seq in eight genotypes (WT, MZ*spg*, MZ*sox19b*, MZ*nanog*, MZ*ps*, MZ*sn*, MZ*pn* and MZ*triple*) and took 102 945 regions accessible in the wild-type (ARs) for further analysis. To select and group direct enhancers, we combined chromatin accessibility changes on ARs in each mutant with additional genomic information, which was the direct binding of TFs (ChIP-seq for Pou5f3, SoxB1, and Nanog) and changes in enhancer activation (ChIP-seq for H3K27ac active enhancer mark in the WT and selected mutants). This genomic analysis resulted in a list of direct enhancers bound by Pou5f3, SoxB1, or Nanog, classified by differential regulation. Data related to genome analysis are in the Table S1. **B.** Transcriptomic analysis. We performed time-resolved RNA-seq profiling at 8 developmental points, every 30 min starting from pre-ZGA (2.5 hpf) till midgastrulation (6 hpf), in 8 genotypes (WT, MZ*spg*, MZ*sox19b*, MZ*nanog*, MZ*ps*, MZ*sn*, MZ*pn,* and MZ*triple*), in at least two biological replicates per genotype. First, we derived differentially expressed gene groups (“DOWN”, “SAME”, “UP”) by pairwise comparisons of transcriptional profiles in each mutant to the wild-type. Second, to decipher combinatorial regulatory inputs of Pou5f3, SoxB1, and Nanog factors to transcription, we applied dynamic mathematical modeling to classify 1799 transcripts, downregulated in MZ*triple*, by type of regulation. This transcriptomic analysis resulted in a list of differentially regulated transcriptional targets of Pou5f3, SoxB1, or Nanog. Data related to transcriptome analysis are in the Table S2. **C.** Cross-validation. We put together the results of independent genome and transcriptome analysis parts to cross-validate them. We linked all zygotic genes to ARs within +/- 50 kb from the gene promoter, and computed correlations between enhancer and transcript groups (Table S3).

Figure S2. DNA-binding proteins are prematurely expressed in most PSN mutants (related to the main Fig. 2).
A. Heatmap shows expression profiles of zygotic transcripts, upregulated in at least one mutant relatively to the WT, in all investigated genotypes. Transcripts were sorted to 7 non-overlapping groups by maximal expression (e.g. if the transcript was upregulated in more than one mutant, it was assigned to the mutant where the expression was higher, Table S2). Groups were named by absent TF(s) in the mutant (vertical labels). **B.** Premature transcription of zygotic genes occurs in all mutants except MZ*sox19b*. The max. expression in the time window from 3 hpf to 5 days (120 hpf) of normal development was calculated for each gene. Median expression time for each group was compared to the control group (Co): 896 zygotic transcripts expressed in the WT and not significantly changed in any mutants. 1-way ANOVA, p-val<$2e^{-16}$, p-values in Tukey-Kramer test are shown. **C.** DNA-binding proteins are prematurely expressed in the mutants for P, S, and N. DAVID, “Molecular function” GO analysis. Note the enrichment for DNA-binding proteins and regulatory function in transcription in upregulated groups, and enrichment for housekeeping genes in the unchanged control group.

Figure S3. Pou5f3, Sox19b and Nanog have different pioneer-like factor properties (related to the main Fig. 3).
A. TdARs (TF-bound ARs, downregulated in MZ*triple*) which were also downregulated in the triple mutant MZ*nps* (Miao et al., 2022) were divided to 4 groups by non-redundant requirements for Pou5f3 and/or Nanog for chromatin accessibility (as in Fig. 3A, main text). **B**-**D.** Percentages of TdARs, for four groups, on which chromatin accessibility could be rescued by microinjection of different concentrations of single TFs into 1-cell stage MZ*nps* embryos in the experiments of Miao et al., 2022. RNA concentrations: low, normal, high for Pou5f3 and Nanog (B, C); normal and high for Sox19b (D). **E.** Percentages of TdARs, for four groups, which could be rescued by microinjection of double or triple combinations of P,S and N TFs into 1-cell stage MZ*nps* embryos in the experiments of Miao et al., 2022. TFs in normal concentration. **F.** Summary of specific rescue by one, two, or three factors, for four TdAR groups (normal concentration). **G**. Summary of promiscuous rescue (by more than one single factor in high concentration, Miao et al., 2022). Note that chromatin accessibility of more than 45% of group 4.0 TdARs can be rescued promiscuously. **H**. Predicted nucleosome occupancy (Kaplan et al., 2009) is the highest in group 4.0 from all four groups of TdARs. Note that predicted nucleosome occupancy of all four groups of TdARs is significantly higher than genomic average (lower dotted line): thus, all TFs bind to nucleosome-rich sites as shown previously (Veil et al., 2019). Still, it is yet significantly lower than of TdARs where accessibility is upregulated in MZ*triple* (upper dotted line). p-value<$2e^{-16}$ for 1-way ANOVA, p-values in Tukey-Kramer test. **I, J.** top: Venn diagrams show the overlaps between the 2.P

and 3.N groups of TdARs, downregulated in the respective single mutants, with TdARs, downregulated in the double mutants by the other two factors. Bottom: The same as top, but 2.P and 3.N groups were split to two groups by rescue by Pou5f3 or Nanog, respectively, in MZ*nps* (Miao et al., 2022) **I**. Pioneer-like activity of Pou5f3 requires assistance of SoxB1 or Nanog in the majority of cases. 59% of 2.P group were downregulated in Sox19b/Nanog double mutant MZ*sn* (top); out of them 31% of 1066 regions which could be rescued by Pou5f3 in MZ*nps* and 82% of 3016 regions which could not be rescued by Pou5f3. **J**. Pioneer-like activity of Nanog does not require assistance of SoxB1 or Pou5f3 in the majority of cases. Only 23% of 3.N group were also downregulated in Pou5f3/Sox19b double mutant MZ*ps* (top); out of them 19% of 6841 regions which could be rescued by Nanog in MZ*nps* and 69% of 685 regions which could not be rescued by Nanog.

Figure S4. Pou5f3 and Nanog compete as activator (+) and blocker (-) on antagonistic enhancers (related to the main Fig.4).
A. Summary profiles and heatmaps for H3K27ac in the indicated genotypes, on the 558 2.P p+n- and 1442 3.N p-n+ antagonistic enhancers. **B.** Left: Summary profiles for H3K27ac in the WT and MZ*sox19b* on the 2.P p+n- and 3.N p-n+ enhancers (heatmaps in A). Right: Sox19b cooperates with Pou5f3 in more cases than with Nanog: on 2.P p+n- enhancers Sox19b induces H3K27ac in 50% and represses in 15% cases, on 3.P-N+ enhancers Sox19b induces H3K27ac in 25% and represses in 50% of cases. **C.** 2.P p+n- and 3.N p-n+ antagonistic enhancers lose chromatin accessibility and H3K27ac in the triple mutants for Pou5f3, Sox19b, and Nanog. * - H3K27ac data from (Miao et al). **D, E**. Summary profiles and heatmaps for ATAC-seq (D) and MNase-seq (E) in the indicated genotypes, for the 2.P p+n- and the 3.P-N+ antagonistic enhancers. Note that chromatin accessibility (ATAC-seq) is increased and nucleosome occupancy (MNase-seq) is reduced on 2.P p+n- enhancers in MZ*nanog* mutant: this demonstrates that Nanog blocks Pou5f3 pioneer-like activity. Conversely, chromatin accessibility is increased and nucleosome occupancy is reduced on 3.N p-n+ enhancers in MZ*spg* mutant: this demonstrates that Pou5f3 blocks Nanog pioneer-like activity. For the heatmaps in A, D, E, enhancers were sorted by descending ATAC-seq signal in the WT. **F**. The opposite effects of Pou5f3 and Nanog on chromatin accessibility and nucleosome occupancy on the 2.P p+n- and 3.N p-n+ antagonistic enhancers are statistically significant. Violin plots show the distribution of ATAC-seq signal (top) and MNase-seq signal (bottom) in the WT, MZ*spg* and MZ*nanog*. P-values in Tukey-Kramer test are shown above the graphs. 1-way ANOVA, ATAC-seq p-val $<2e^{-16}$; 1-way ANOVA, MNase-seq p-val $=7.2e^{-9}$ for 2.P p+n- enhancers, p-val$<2e^{-16}$ for 3.N p-n+ enhancers. **G.** TdARs downregulated (down) or upregulated (up) in each mutant compared to wild-type were selected to score enrichment of sequence -specific motifs (scheme at the left). Motif frequencies in down- and upregulated regions for the indicated genotypes; all TdARs - TF-bound accessible regions downregulated in MZ*triple*. Number of regions at the bottom. Note the enrichment for Pou5f3 cognate motifs in the regions upregulated in MZ*nanog* (yellow box) and enrichment for Nanog cognate motifs in the regions upregulated in MZ*spg* and MZ*ps* (blue boxes).

Figure S5. Core model for dynamics of zygotic genome activators (related to the main Fig. 5).
A. Model cartoon of core interactions between SoxB1 family class (orange), Nanog (violet) and Pou5f3 (green). Corresponding mechanistic model equations are summarized in Table S4. **B.** Fit results of the core model are shown as a waterfall plot sorted by their likelihood values. Out of 100 fits, 38 fits ended up at the global optimum. Parameter values of the best fit are given in Table S5. **C-D.** Model trajectories (lines) of RNA targets (C) and protein targets (D) of the best fit are shown together with RNA-seq data (points) used for model calibration. The seven experimental conditions are displayed color coded and in different columns. Different rows correspond to different transcription factors and proteins, respectively. Shaded areas in C correspond to uncertainties predicted by the model for each target (compare estimated error parameters in Table S5).

Figure S6. Figure S6: Identifiability of core model parameters and exemplary mini-model fits. (related to the main Fig. 5).
A. Parameter values (dots) of the best fit are shown with their profile likelihoods (lines) for the core model. Likelihood-ratio test based confidence levels for one degree of freedom are shown as dashed grey lines. Parameter values (x-axis) are displayed on log10 scale. Out of 37 estimated parameters, 30 showed parabolic profile likelihoods corresponding to finite confidence intervals, and are thus identifiable (Raue et al., 2009). The remaining seven were practically non-identifiable given the available data. The resulting parameter confidence intervals are summarized in Table S5. **B**. For each mini-model group (colors), two examples of the best fit mini models (lines) are shown together with the corresponding RNA-seq data (points). Columns refer to experimental conditions, and rows indicate different target genes. Model equations representing the different mini-model groups are summarized in Table S6, with bestfit parameter values for the individual target genes shown in Table S7**.**

Figure S7. Pou5f3/Sox19b block transcription activated by Nanog. **(related to the main Fig. 5 and Fig. 6)**

A. DAVID, GO analysis in the categories "Biological process" and "Molecular function" for the groups of transcripts 1. P+N+, 3a. P0N+ and 3b P-N+ (see main text). **B, C**. χ^2 tests, positive correlations between the transcriptomics and genomics groups are shown in blue and negative in red. Vertical axis: transcriptomics groups. Direct enhancers were linked to the TSS of zygotic genes within +/-50 kb and sorted according to best fit transcriptional model (see main text and Fig. 6). The synergistically regulated targets 1. P+N+ and antagonistically regulated targets 2b. P+N- and 3b. P-N+ are boxed. Horizontal axis: genomic groups. Direct enhancers were sorted to four groups (1.PN, 2.P, 3.N and 4.-) by changes in chromatin accessibility (B), and by H3K27ac regulation by Sox19b (C). +, 0 and – indicate that Sox19b induces H3K27ac, does nothing, or blocks it, respectively. Note that the s+ enhancers are enriched only in the putative regulatory regions of 2b. P+N- transcript group. **D**. Zygotic gene expression changes in MZ*ps* double mutant. Heatmap shows normalized expression of three groups of zygotic transcripts ("DOWN", "SAME" and "UP" compared to the wild-type) at eight time points, from pre-ZGA (2.5 hpf) till 6 hpf; numbers of transcripts are shown at the right. **E.** Premature zygotic expression in MZ*ps***.** Expression times in the normal development were compared in "DOWN", "SAME" and "UP", groups; and were significantly higher in "UP" group of transcripts (dotted lines show the median values). P-values in Tukey-Kramer test. **F.** χ^2 tests, positive correlations between the transcriptomics and genomics groups are shown in blue and negative in red. Vertical axis: transcriptomics groups. Direct enhancers were linked to the TSS of zygotic genes within +/-50 kb and sorted according to DOWN", "SAME" and "UP" transcriptional groups. Horizontal axis: genomic groups. Direct enhancers were sorted by changes in chromatin accessibility, and by H3K27ac regulation by Pou5f3 or Nanog, as indicated. Note that the Nanog-activated enhancers where H3K27 acetylation is blocked by Pou5f3 (3.N p-n+) are enriched around TSS of genes upregulated in MZ*ps* (boxed).

STAR Methods

RESOURCE AVAILABILITY

Lead contact

Requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Daria Onichtchouk (daria.onichtchouk@biologie.uni-freiburg.de).

Materials availability

Zebrafish lines generated in this study are available from lead contact, Daria Onichtchouk (daria.onichtchouk@biologie.uni-freiburg.de) under the restrictions of German Animal Protection law (TierSchutzGesetz).

Data and code availability

- Raw and processed RNA-seq and ATAC-seq data generated in this study have been deposited in GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- The original code is deposited in https://github.com/vandensich/zebrafish-minimodels and will be made available to the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact Daria Onichtchouk (daria.onichtchouk@biologie.uni-freiburg.de) upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Zebrafish maintenance and generation of the double and triple mutants

Wild-type fish of AB/TL and mutant strains were raised, maintained and crossed under standard conditions as described by Westerfield (Westerfield, 2000). Embryos were obtained by natural crossing (4 males and 4 females in 1,7 l breeding tanks, Techniplast). Wild-type and mutant embryos from natural crosses were collected in parallel in 10-15 minutes intervals and raised in egg water at 28.5°C until the desired stage. Staging was performed following the Kimmel staging series (Kimmel et al., 1995). Stages of the mutant embryos were indirectly determined by observation of wild-type embryos born at the same time and incubated under identical conditions. All experiments were performed in accordance with German Animal Protection Law (TierSchG) and European Convention on the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes (Strasburg, 1986). The generation of double and triple mutants for Pou5f3, Sox19b and Nanog was approved by the Ethics Committee for Animal Research of the Koltzov Institute of Developmental Biology RAS, protocol 26 from 14.02.2019. Maternal-Zygotic (MZ) homozygous mutant embryos MZ*sn* were obtained in three subsequent crossings. First, MZ*nanog*^{m1435} (Veil et al., 2018) homozygous males were crossed with MZ*sox19b*^{m1434}(Gao et al., 2022)

homozygous females. The double heterozygous fish were raised to sexual maturity and incrossed. The progeny developed into phenotypically normal and fertile adults. Genomic DNA from tail fin biopsies was isolated and used for genotyping. We first selected *sox19b* homozygous mutants, by PCR-with Sox19b-f1/Sox19b-r1 primers followed by restriction digest with BbsI (Gao et al., 2022). To select the double homozygous fish, we used the genomic DNA from *sox19b* homozygous mutants for by PCR with Nanog1-f1/Nanog1-r1 primers followed by restriction digest with NdeI. MZ*sn* embryos for experiments were obtained from incrosses of double homozygous *sox19b-/-*; *nanog-/-* fish. The line was maintained by crossing *sox19b-/-*; *nanog -/-*; males with *sox19b-/-*; *nanog* +/- females.

MZ*pn* homozygous mutant embryos were obtained in three subsequent crossings. First, MZ*nanog* ^{m1435} homozygous males (Veil et al., 2018) were crossed with MZ*spg* ^{m793} (Lunde et al., 2004) homozygous females. The double heterozygous fish were raised to sexual maturity and incrossed. To bypass the early requirement for Pou5f3 in the *spg793* homozygous mutants, one-cell stage embryos were microinjected with 50-100 pg synthetic Pou5f3 mRNA as previously described (Lunde et al., 2004). The fish were raised to sexual maturity, genomic DNA from tail fin biopsies was isolated and used for genotyping. We first selected *nanog* homozygous mutants, by PCR with Nanog1-f1/Nanog1-r1 primers followed by restriction digest with NdeI. To select the double homozygous fish, we used the genomic DNA from *nanog* homozygous mutants to PCR-amplify the region, flanking the *spg*^{m793} allele. *Spg*^{m793} allele carries an A->G point mutation in the splice acceptor site of the first intron of Pou5f3 gene, which results in the frameshift starting at the beginning of the second exon, prior to the DNA-binding domain. *Spg*^{m793} is considered to be null allele. We used the following PCR primers: spg-f1 5'-` GTCGTCTGACTGAACATTTTGC -3' and spg-r1 5'-` GCAGTGATTCTGAGGAAGAGGT -3'. Sanger sequencing of the PCR products was performed using commercial service (Sigma). The sequencing traces were examined and the fish carrying A to G mutation were selected. MZ*pn* embryos for experiments were obtained from incrosses of double homozygous *spg-/-*; *nanog-/-* fish. The line was maintained by crossing *spg-/-*; *nanog* -/- males with *spg-/-*; *nanog* +/- females and microinjecting Pou5f3 mRNA in each generation.

To obtain the triple Maternal-Zygotic (MZ) homozygous mutant embryos MZ*triple*, M*Zsn* double homozygous males were crossed with MZ*ps* double homozygous females (Gao et al., 2022). The *sox19b-/-*; *spg+/-*; *nanog+/-* progeny was incrossed, microinjected with 50-100 pg synthetic Pou5f3 mRNA and genotyped for *spg* and *nanog* as described above. MZ*triple* embryos for experiments were obtained from incrosses of triple homozygous fish. The line was maintained by crossing *sox19b-/-*; *spg-/-*; *nanog-/-* males with *sox19b-/-*; *spg-/-*; *nanog+/-* females and microinjecting Pou5f3 mRNA in each generation.

Genomic DNA isolation and PCR for genotyping

Genomic DNA was isolated from individual tail fin biopsies of 3 months old fish. Tail fin biopsies or embryos were lysed in 50µl lysis buffer (10 mM Tris pH 8, 50 mM KCl, 0.3%

Tween20, 0.3% NP-40, 1mM EDTA) and incubated at 98°C for 10 min. After cooling down Proteinase K solution (20 mg/ml, A3830, AppliChem) was added and incubated overnight at 55°C. The Proteinase K was destroyed by heating up to 98°C for 10 min. The tail fin biopsies material was diluted 20x with sterile water. 2µl of was used as a template for PCR. PCR was performed in 25-50 µl volume, using MyTag polymerase (Bioline GmbH, Germany) according to the manufacturer instructions, with 30-35 amplification cycles. The primer sequences are listed in the key resources table.

METHOD DETAILS

ATAC-seq and library preparation

Omni-ATAC-seq was performed as described in (Gao et al., 2022). Briefly, embryos were enzymatically dechorionated with Pronase E (30 mg/ml). 30 embryos were deyolked manually with an eyebrow needle in Danieaus medium, centrifuged (500 x g, 5 min, 4°C), washed with 100 µl ice-cold PBS, and then centrifuged again. The cells were resuspended by pipetting in 50 µl ice-cold lysis buffer (ATAC-RSB (10 mM TrisHCl pH 7.4, 10 mM NaCl, 3 mM MgCl2), with 0.1% NP40, 0.1% Tween20, and 0.01% digitonin) and incubated on ice for 5 minutes. Then, the lysate was diluted with ice-cold dilution buffer (1 ml, ATAC-RSB containing 0.1% Tween-20). Nuclei were pelleted (500 x g, 10 min, 4°C), the supernatant carefully removed, and the nuclei resuspended in 16.5 µl PBST. Tagmentation reaction was assembled according to the manufacturer instructions for Illumina small Tn5 buffer and enzyme kit in 50 µl volume and incubated in a thermomixer for 30 min (37°C, 800 rpm). The sample was purified using Qiagen PCR MinElute Purification kit. The purified transposed mix was pre-amplified for 5 cycles (72°C for 5 min, 98°C for 30 sec, 5 cycles of (98°C for 10 sec, 63°C for 30 sec, 72°C for 1 min) using NEBNext 2x Master Mix (NEB) and Nextera adapters and put on ice. The number of additional cycles was determined using 5 µl of the pre-amplified mixture as in (Buenrostro et al., 2013). The total number of PCR cycles was 10-12. The amplified libraries were purified using SPRI beads (Beckmann) according to the manufacturer instructions. ATAC-seq was performed at three developmental stages, 3 hpf (1k), 3.7 hpf (oblong) and 4.3 hpf (dome). At 3 hpf, three biological replicates for WT and MZ*triple* were used. At 3.7 hpf, two biological replicates for MZ*nanog* and MZ*pn* and three biological replicates for MZ*triple* were used. At 4.3 hpf, three biological replicates for MZ*sn* and MZ*triple* and two for MZ*pn* and MZ*nanog* were used. Paired-end 150 bp sequencing of resulting 23 ATAC-seq libraries was performed on NovaSeq6000 (Illumina) by Novogene company, with the sequencing depth 80 million reads per library. The raw and processed ATAC-seq data are deposited in GEO (NCBI), accession number GSE215956.

ATAC-seq data processing

ATAC-seq data processing was performed as in (Gao et al., 2022) on the european Galaxy server (Afgan et al., 2016). Adapters were removed, the reads were cropped to 30 bp from start, the reads with average quality less that 30 and length less than 25 bp were removed using Trimmomatics. The trimmed reads were aligned to the danrer11/ GRCz11 genome assembly without alternative contigs using Bowtie2

(Langmead and Salzberg, 2012), with the parameters: '—dovetail, --very-sensitive, --no-unal'. The aligned reads were filtered using BamTools Filter (Barnett et al., 2011) with the parameters -isProperPair true, -mapQuality >=30, -isDuplicate false, -reference !chrM. The aligned reads were filtered further using the alignmentsieve tool from the deepTools2 suite (Ramirez et al., 2016), with the parameters —ATACshift, --minMappingQuality 1, --maxFragmentLength 110. Thus, only fragments with a size of 110 bp or less were kept. The filtering steps were performed for the individual replicates as well as and for the merged biological replicates BAM files. Bigwig files for each stage and genotype were obtained from merged BAM files using Bamcoverage and used for data visualization in deepTools2. Bigwig and peak files for each stage and genotype are included in GEO GSE215956 submission as processed files.

Definition of regulated ARs and TdARs

The ATAC-seq data from this work at 3.7 hpf and 4.3 hpf were processed together with our previously published ATAC-seq data set: three biological replicates for the wild-type and two biological replicates for MZ*spg,* MZ*sox19b* and MZ*ps* at 3.7 hpf, three biological replicates for the wild-type, MZ*spg* and MZ*sox19b* and four for MZ*ps* at 4.3 hpf (Gao et al., 2022), GEO accession number GSE188364. For seven wild-type biological replicates, filtered reads were converted from BAM to BED using Bedtools (Quinlan and Hall, 2010) and were used as inputs for peak calling. Peaks were called from individual replicates and merged replicates using MACS2 callpeak (Feng et al., 2012) with the additional parameters '--format BED, --nomodel, --extsize 200, --shift -100' and a 5% FDR cutoff. Regions mapping to unassembled contigs were excluded. Only peaks that were overlapping in merged and each of the single replicates were kept. ATAC-seq peaks overlapping between 3.7 and 4.3 hpf stages in the wild-type were centered on ATAC-seq peak summits (4.3 hpf) and cut to 110 bp length. This set of 102 965 regions accessible in the wild-type (ARs) was used in all subsequent analyses. The coverage of ATAC-seq reads on ARs was scored using MultiCovBed (Bedtools) in all ATAC-seq experiments (3.7 and 4.3 hpf). Deseq2 (Love et al., 2014) was used to normalize the reads and compare chromatin accessibility in each mutant to the wild-type. For each mutant genotype, we made four Deseq2 cross-normalizations: across all replicates, across the replicates of 3.7 hpf or 4.3 hpf separately, and across three samples (i.e. wild-type, MZ*nanog* and MZ*triple*). The region was scored as downregulated in the mutant, if the fold change to the wild-type was negative with FDR 5% in four Deseq2 normalizations. The region was scored as upregulated in the mutant, if the fold change to the wild-type was positive with FDR 5% in four Deseq2 normalizations. The remaining ARs were considered unchanged. To define TdARs, (TF-bound Accessible Regions, downregulated in MZ*triple*), we selected ARs, downregulated in MZ*triple* and overlapping with any of Sox2, Pou5f3 (Leichsenring et al., 2013) or Nanog (Xu et al., 2012) ChIP-seq peaks for at least 1 bp. The genomic coordinates of TF-binding peaks remapped to danrer11/GRCz11 assembly were taken from (Gao et al., 2022). Genomic coordinates of ARs and their properties are provided in the Table S1.

Scoring the motif enrichment

Sequence-specific *nanog1* and *nanog2* motif matrices were taken from (Veil et al., 2019), the rest of the motif matrices from (Gao et al., 2022). Genomic coordinates of all occurrences for each motif in the 3 kb regions around ATAC-seq peaks were obtained using FIMO (Grant et al., 2011) with p-value threshold 10^{-4} and converted to BED files. To calculate the background frequency of the motifs, we used the background control peak file, where the genomic coordinates of all ATAC-seq summits were shifted 1 kb downstream. The motif densities were calculated as the average number of motifs overlapping with 60 bp around the ATAC-seq or control peak summits. The log2 ratio of ATAC-seq to background motif densities was taken as a motif enrichment value.

GC content and *in-vitro* nucleosome predictions

GC content for 110 bp around AR summits was calculated with geecee program, *in-vitro* nucleosome prediction with the algorithm of Kaplan et al. (Kaplan et al., 2009) in the european Galaxy instance usegalaxy.org. The genomic coordinates of ARs and control peaks were extended to +/-5 kb to account for the edge effects, and *in-vitro* nucleosome prediction value was derived for every base pair. Nucleosome predictions were converted to bedgraph files for scoring. GC content and nucleosome prediction values per 110 bp around ARs are provided in the Supplementary Table 1.

Scoring the rescue of chromatin accessibility

Track hub with Omni-ATAC rescue processed data in UCSC bigwig format was downloaded from the home page of Antonio Giraldez lab (https://www.giraldezlab.org/data/miao_et_al_2022_molecular_cell/) and used for scoring. We calculated ATAC-seq signals in all conditions for 110 bp around ARs. For ARs, downregulated in MZ*nps* compared to the WT B1 replicate in (Miao et al., 2022), we estimated the rescue. We considered AR as rescued by TF(s) microinjection, if the log2 of ATAC-seq signals (injected MZ*nps*/control MZ*nps*) was exceeding log2 of ATAC-seq signals (control WT B1/injected MZ*nps*). The rescue status by all combinations of TFs for ARs downregulated in MZ*nps* is provided in the Table S1; the statistics for rescue on TdARs is provided on the Fig. S3.

ChIP-seq for H3K27ac in MZ*nanog*

The ChIP-seq was performed as in (Gao et al., 2022). Embryos were obtained from natural crossings in mass-crossing cages (4 males + 4 females). The freshly laid eggs of MZ*nanog* mutants and wild-type were collected in 10-15 min intervals. The embryos were incubated at 28.5°C and enzymatically dechorionated with 30 mg/ml Pronase shortly before 4.3 hpf. Embryos were homogenized in 10 ml 0.5 % Danieau's (for 1L of 30X stock: 1740 mM NaCl, 21 mM KCl, 12 mM MgSO4, 18 mM Ca(NO3)2, 150 mM HEPES buffer, pH 7.6) containing 1x protease inhibitor cocktail (PIC, Roche) using a Dounce tissue grinder, and immediately treated with 1% (v/v) Methanol-free Formaldehyde (Pierce) for exactly 10 min shaking at room temperature. The fixation was stopped with 0.125 M Glycine by shaking for 5 min on a rotating platform. Lysate was centrifuged for 5 min, 4700 rpm at 4 °C, the pellet was resolved in cell lysis buffer (10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 0.5 % NP-40, 1-4 ml/1000 embryos), followed

by 5 min incubation on ice, and centrifuged 1 min, 2700 g at 4 °C. The pellet was washed 2 times with 1 ml ice cold 1x PBST (for 1 L: 40 ml PO4 buffer (0.5 M), 8 g NaCl, 0.2 g KCl, 0.1% Tween20, pH 7.5) and resuspended in 1 ml PBST. 5 µl aliquote of nuclei suspension was stained with Sytox green, nuclei were scored under fluorescence microscope using the Neubauer counting chamber. The residual nuclei were pelleted by 1 min centrifugation at 2700 g at 4 °C, snap frozen in liquid nitrogen and stored at -80 °C. The nuclei collected in different days were pooled together to reach the total number of 2.5-3 million, which was used to start one ChIP experiment. The chromatin was thawn, pooled in 2 ml of nuclei lysis buffer (50 mM Tris-HCl (pH 7.5), 10 mM EDTA, 1 % SDS) and incubated 1 h on ice. In order to shear the chromatin to 200 bp fragments (on average), the chromatin was sonicated using the Covaris S2 sonicator (DC 20 %, Intensity 5, Cycles of Burst 200, Time = 3*40 cycles with 30 sec each (3*20 min)). To control the sonication, 30 µl of the sheared chromatin was de-crosslinked with 250 mM NaCl overnight at 65 °C and then analyzed using the Agilent Expert 2100 Bioanalyzer® and Agilent high sensitivity DNA Chip kit.

The lysed and sheared samples were centrifuged for 10 min, 4°C at max. speed in table top centrifuge. 60 µl of each sample were kept as input control. The chromatin was then concentrated to 100 µl using the Microcon centrifugal filters (Merck Millipore MRCF0R030) and diluted 1:3 by adding ChIP dilution buffer (16.7 mM Tris-HCl pH 7.5, 167.0 mM NaCl, 1.2 mM EDTA) containing protease inhibitors. The H3K27ac antibody was added and incubated overnight at 4 °C on a rotating wheel. 150 µl of magnetic Dynabeads coupled to protein G (Stock 30 mg/ml; invitrogen DynaI 10003D) were transferred into a 2 ml eppendorf tube and placed on a magnetic rack. The beads were washed 3x with 5 mg/ml specially purified BSA in PBST and 1x with 500 µl ChIP dilution buffer. After removing the ChIP dilution buffer, the chromatin-antibody mix was added and incubated with the beads at 4 °C overnight on a rotating wheel. Beads were pulled down by placing the eppendorf tubes on the magnetic rack. The beads were resuspended in 333 µl RIPA buffer (10 mM Tris-HCl (pH 7.6), 1 mM EDTA, 1 % NP-40, 0.7 % sodium deoxycholate, 0.5M LiCl) containing PIC. The Protein G-antibody-chromatin complex was washed 4x5 min on a rotating platform with 1 ml of RIPA buffer, followed by 1x1 ml TBST buffer (25 mM Tris-HCl,150 mM NaCl, 0.05% Tween 20, pH 7.6). The beads were pulled down again. To elute the chromatin, 260 µl elution buffer (0.1 M NaHCO3, 1% SDS) was added and incubated for 1 h at 65° C in a water bath. The samples were vortexed every 10 - 15 min. The beads were pulled down, supernatant was transferred to a fresh eppendorf tube. 12.5 µl 5M NaCl was added to de-crosslink the chromatin and incubated overnight at 65 °C in a water bath. The input samples were treated in parallel (230 µl elution buffer per 30 µl input). Purification of the de-crosslinked chromatin was performed using the QIAquick PCR Purification Kit from Qiagen. The concentration was determined using the Qubit fluorometer and Quanti-iT™ PicroGreen® dsDNA Kit according to manufacturer instructions.

To estimate the signal to background ratio in each ChIP experiment, we used the positive and negative reference genomic regions, near *tiparp* and *igsf2* genes, respectively enriched in or devoid of H3K27ac (Gao et al., 2022). We performed

quantitative PCR in ChIP and Input control material, using the primers for these regions. PCR primers used were: tiparp_f_1 5' CGCTCCCAACTCCATGTATC-3', tiparp_r_1 5'-AACGCAAGCCAAACGATCTC-3',igsf2_f_2 5'-GAACTGCATTAGAGACCCAC-3', igsf2_r_2 5'-CAATCAACTGGGAAAGCATGA-3'. QPCR was carried out using the SsoAdvanced™ Universal SYBR® Green Supermix from BIO-RAD. ChIP and input were normalized by ddCT method, using negative reference region *(igsf2*). The ChIP experiment was considered successful, if the enrichment in ChIP over input control on the positive reference region (*tiparp*) was more than 5-fold.

The library preparation was carried out with NEBNext® Ultra™ DNA Library Prep Kit according to manufacturer instructions, with the modifications indicated below. As the DNA outcome of individual MZ*nanog* K27ac ChIP experiments did not reach the input DNA limit for the kit (5 ng), we pooled together the material from two successful ChIP experiments, as well as corresponding inputs. Since the DNA input was <100 ng, in the adaptor ligation step, the NEBNext Adaptor for Illumina® (15 µM) was diluted 10-fold in 10 mM Tris-HCl (pH 7.8) to a final concentration of 1.5 µM and used immediately. At the final clean-up step, the reaction was purified by mixing the samples with AMPure XP Beads (45 µl) before incubating at room temperature for 5 min. After a quick spin using the tabletop centrifuge, samples were placed on a magnetic rack and supernatant was discarded. 200 µl of 80 % Ethanol were added and removed after 30 sec, two times. The beads were air dried for 3 min and the DNA target was eluted by adding 33 µl of 0.1x TE (pH 8) and incubating at room temperature for 2 min. 28 µl of the library were transferred to a fresh PCR tube and stored at -20 °C. 2 µl of the sample were diluted 5-fold with 0.1x TE and used to check the size distribution of the library using Agilent Expert 2100 Bioanalyzer® and Agilent high sensitivity DNA Chip kit. The ChIP-seq library was sequenced at 70 mln paired end 150bp reads the input library were sequenced to 30 mln reads. Sequencing was performed by the Novogene company (China). The raw and processed data have been deposited in GEO under the number GSE143439.

Definition of regulated enhancers

Sequenced reads were mapped to the zebrafish danrer11 assembly using Bowtie2 in the european Galaxy server. Alternate loci and unassembled scaffolds were excluded. H3K27ac data for MZ*nanog* from this work was further processed together with H3K27ac for the wild-type, MZ*spg* and MZ*sox19b* (Gao et al., 2022). To create bigwig files for scoring, the log2 ratio between each ChIP and merged inputs (in rpkm) was obtained using BAMcompare program in deepTools, bin size =10. The mean H3K27ac signal (log2 ChIP/input per bin) around each AR was calculated for 1 kb around AR summit. Values smaller than 0.1 were rounded to 0.1. We considered AR as H3K27-acetylated, if the mean H3K27ac signal was higher than arbitrary threshold 0.2. All H3K27-acetylated TdARs were considered as enhancers. Enhancer was considered as activated by TF, if log2 (mutant by TF/WT) ratio of H3K27ac signals was less than arbitrary threshold -0.7, was considered as blocked by TF, if log2 (mutant by TF/WT) ratio of H3K27ac signals was more than 0.7, and was considered as unchanged by TF

all other cases. Enhancers activated, blocked or unchanged by the sum of TFs were assigned by applying the same procedure and thresholds to H3K27ac ChIP-seq data on the WT and MZ*nps* (Miao et al., 2022). H3K27ac status of all ARs is provided in the Supplementary Table 1.

RNA-seq time curves

RNA-seq time curves were obtained as described previously (Gao et al., 2022). Freshly laid eggs were obtained from natural crossings in mass-crossing cages (4 males + 4 females). 5-10 cages were set up per genotype, the eggs from different cages were pooled. At least 600 freshly laid eggs per each genotype collected within 10-15 minutes were taken for single experiment. To match the developmental curves as precisely as possible, the material was simultaneously collected for two genotypes in parallel. 45-60 minutes after the egg collection, we ensured that the embryos of both genotypes are at 2-4 cell stage, removed non-fertilized eggs and distributed the embryos to 8 dishes per genotype, 40-45 embryos per dish, to obtain 8 time points per genotype: 2.5, 3, 3.5, 4, 4.5, 5, 5.5 and 6 hpf. The temperature was kept at 28.5°C throughout the experiment. The embryos from one dish per genotype were snap-frozen in liquid nitrogen every 30 minutes, starting from 2.5 hpf (pre-ZGA, 256 cell stage, 8th cell cycle) till midgastrula (6 hpf). Two biological replicates for each of MZ*nanog* and MZ*sn* were previously collected in parallel to the wild-type replicates B1 and B2 from (Gao et al., 2022). In addition, we collected six biological replicates of the time curve for the wild-type, and three biological replicates for MZ*pn* and MZ*triple* mutants from six independent experiments. Total RNA was isolated with QIAGEN RNeasy kit following the user's manual. We checked RNA quantity and quality by using Agilent RNA 6000 nano Kit on Agilent Bioanalyzer, according to manufacturer instructions. Poly-A enriched library preparation and sequencing on Illumina platform was performed by Novogene company (China). The MZ*nanog* and MZ*sn* sampes were single-end sequenced at 35 million reads per sample (50 bp), similarly to the previously published samples WT, MZ*sox19b*, MZ*spg* and MZ*ps* (Gao et al., 2022). The MZ*pn* and MZ*triple* samples and corresponding six replicates of the wild-type were paired-end sequenced at 58 million reads per sample (150 bp).

RNA-seq time curves: data processing

FASTQ files were processed on european Galaxy server usegalaxy.eu. All sequenced clean reads data were trimmed on 5' end for 5 bp by Trim Galore! program according to the sequencing quality. Trimmed reads were mapped to danrer11 genome assembly using RNA STAR (Dobin et al., 2013) and ENSEMBL gene models. Number of reads for each gene was counted using Feature count (Liao et al., 2014). Single-end and paired-end sequencing data sets were further processed separately as the part 1 and the part 2. For the part 1, feature counts for the MZ*nanog* and MZ*sn* samples were cross-normalized with MZ*sox19b*, MZ*spg* and MZ*ps* and four biological replicates of the wild-type (Gao et al., 2022) using Deseq2. For the part 2, feature counts for the MZ*pn* and MZ*triple* samples and corresponding six replicates of the wild-type were cross-normalized using Deseq2. The raw data and two normalized data tables were deposited in GEO with the accession number GSE162415.

Finding differentially expressed zygotic genes

We used 3-step RNA-sense program to process normalized RNA-seq time curves as described in (Gao et al., 2022). Shortly, RNA-sense takes RNA-seq time curves in two conditions (WT and mutant) as an input. The transcripts expressed below a user-defined threshold are excluded from the analysis. In the first step, RNA-sense creates the lists of transcripts switching UP or DOWN in time, in one of the genotypes; the genotype and switch p-value are defined by user. We found that 4247 genes are switching UP in the wild-type, in both part1 and part 2 of the analysis (thresholds >100 for expression and <0.15 for switch p-value). Switching UP transcripts were considered as zygotically expressed (Gao et al., 2022). 1668 transcripts were switching up in at least one mutant but not in the wild-type; we considered them as zygotically expressed in the respective mutants. The regulation status was assigned to each zygotic transcript on the second step of RNA-sense. Zygotic transcript was considered to be downregulated in the mutant, if it was expressed at least two-fold lower compared to the wild-type with p-value <0.05 in Student's t-test in at least one time point at or after the switch UP. Zygotic transcript was considered to be upregulated in the mutant, if it was expressed at least two-fold higher compared to the wild-type with p-value <0.05 in Student's t-test in at least one time point. Some transcripts were up – and down regulated in the same mutant in different time points and were considered as "undefined" ("U"). The list of zygotic genes, their switch and regulation status in all the mutants is provided in the Table S2.

Enrichment statistics

Zygotic transcripts (Table S2) were linked to Chromatin Accessible Regions (ARs, Table S1) within +/-50 kb from Transcription Start Site (TSS). Resulting 70419 AR-transcript pairs were uniquely numbered (Table S3). To estimate the over- or underrepresentation of different types of enhancers in the putative regulatory regions of zygotic genes, Chi-squared (χ^2) test was performed using chisq.test function and visualized using corrplot function in R. We considered only strong positive correlations with Pearson residuals more than 4. To reproduce our results, the input files for all χ^2 tests can be easily derived from the supplementary table 3 as explained below. The input file is a tab-separated text file with three columns: "AR-transcript_pair" (which is always column A in the Table S3), "AR", and "transcript". To create the input file for χ^2 test, delete the first two rows of the Supplementary Table 3, keep column A and two columns specified below (delete all other columns), delete the rows specified below, change the headings to AR" and "transcript" and export the input file as tab-separated text. Fig. 2G: "AR" – column N, "transcript" – column AF, delete all rows with "U". Fig. 5D-F: "AR" – columns O, Q, R, respectively, delete all rows with "none"; "transcript" – column AG, delete all rows with "N/A". Fig. 6C: "AR" – columns G, Q, R, delete all rows with "none"; "transcript" – column Z, delete all rows with "U". Fig. 6C: "AR" – columns G, Q, R, delete all rows with "none"; "transcript" – column AB, delete all rows with "U". Fig. S7B, C: "AR" – columns G, P, respectively, delete all rows with "none"; "transcript" – column AG, delete all rows with "N/A". Fig. S7F: "AR" – columns G, Q, R, delete all rows with "none"; "transcript" – column AC, delete all rows with "U".

Gene Ontology analysis.

For the genomic regions we used GREAT analysis (Hiller et al., 2013) at http://great.stanford.edu/great/public-3.0.0/html/. Genomic regions were converted to zv9/danrer9 genomic assembly, which is the only assembly available for this server, using Liftover Utility from UCSC and associated with genes using 20 kb single nearest gene association rule. The categories were ranked by ascending FDR value. For the transcript groups we used DAVID (Huang da et al., 2009).

Data visualization

Omni-ATAC-seq: to create bigwig files for data visualization, Omni-ATAC-seq signal (in rpkm) was obtained using BAMcoverage program in deepTools2, bin size =10. H3K27ac: to create bigwig files for data visualization, the log2 ratio between each ChIP and merged inputs (in rpkm) was obtained using BAMcompare program in DeepTools, bin size =10. The heatmaps or mean profiles were plotted using plotheatmap or plotprofile programs in DeepTools. H3K27ac, TF ChIP-seq and omni-ATAC-seq profiles in single genes were visualized in UCSC browser. RNA-seq heatmaps were plotted using R with custom script. For the heatmaps, the transcripts were sorted by ascending RNA-sense switch p-value then by ascending RNA-sense switch time then by group. Biological replicates from all experiments were cross-normalized using Deseq2. Replicates for each genotype were averaged and the maximal expression value across the time curve in all genotypes was calculated. Expression/max ratio for each timepoint was plotted.

Mathematical Modeling

The mathematical modeling process was performed in two main steps, (i) the core model and (ii) the mini models. The idea behind developing a core model was to use the resulting dynamics of relevant transcription factors or their combinations as input functions to the mini models that later describe the dynamical behavior of a huge amount of 1799 preselected target genes. Generation of ODE models, parameter estimation and identifiability analysis was performed by means of the R package dMod (Kaschek et al., 2019). For each model, the best parameter set was obtained by the method of maximum likelihood performing a deterministic multi-start optimization (Raue et al., 2013) from multiple randomly chosen starting points in parameter space. The trust region optimizer *trust* (Geyer CJ (2004) trust: Trust Region Optimization. ***R package*** Version 0.1-8) was used constituting the standard optimizer in the R package dMod. Model parameters were log-transformed to ensure positivity and enable optimization over a broad range of magnitudes (Raue et al., 2013). Parameter values were not restricted by fixed borders. Instead, in order to prevent the optimizer from finding solutions with very low or high parameter values, we constrained the model parameters with a weak L_2 prior that contributed with one to the likelihood, if the parameter differed by three orders of magnitude from a value of zero on the logarithmic scale. For the computation of confidence intervals and to test identifiability of the model's parameters by means of the profile likelihood (Raue et al., 2009) the contribution of these L_2 priors was subtracted to ensure a purely data-based result.

Construction of the core model

In total, the core model consists of 11 model states and 17 reactions (Fig. S5A, Table S4). Kinetic expressions of the model reactions were derived based on state-of-the-art transcription factor network modeling (Alon, 2007) and prior knowledge from published literature (Gao et al., 2022; Okuda et al., 2010; Onichtchouk et al., 2010; Veil et al., 2018). For simplicity, the degradation of Sox19b was incorporated via a switch-like input function dependent on the time t as a parameter. Model states were assumed to be directly observed, and uncertainty of the observations were estimated jointly with the kinetic parameters assuming an absolute error model for the measurements of each observed state. Model equations, as summarized in Table S4, include three boolean variables MZ*spg*, MZ*nanog* and MZ*sox* that were used to distinguish between the seven experimental conditions including wild type, single, double and triple knockdowns. In total, a number of 888 RNA-seq data points for six targets was used for model calibration. The first time point of the data set was at 2.5 hpf. As an assumption, the initial values of the ODE model were defined at 2 hpf. The initial values of downstream targets Sox2, Sox3 and Sox19a were assumed to be zero, while for the values of the zygotic transcription factors Pou5f3, Sox19b and Nanog a free parameter was estimated by the model, with the constraint of being the same over all experimental conditions.

During the model development, a couple of parameters was fixed to zero or one, following likelihood-ratio test based model reduction (Maiwald et al., 2016) since they were not necessary in order to describe the available data. Hill-kinetic parameters were fixed to a value of five corresponding to a strong threshold behavior. Further parameter transformations that have been applied are summarized in https://github.com/vandensich/zebrafish-minimodels. Parameter estimation was performed using deterministic multi-start optimization (Raue et al., 2013) starting from 192 randomly chosen sampled positions in parameter space. Parameter values of the best obtained fit are summarized in Table S5. To show reliability of the optimization, the 100 best optimization runs were displayed as a waterfall plot (Fig. S5B) sorted by their negative log-likelihood values (Raue et al., 2013). The global optimum was found in 38 of the 101 converged optimizations runs. Considering a confidence level of 95%, profile likelihoods showed finite confidence intervals for 30 out of 37 parameters (see Fig. S6A).

Construction of the mini models

Assuming at least one of Pou5f3, Nanog and SOX to be activating, a set of 19 mini models was constructed. Model equations for each mini model (Table S6) were derived from the theory of transcription factor network modeling (Alon, 2007). When more than one transcription factor was either activating or repressing, the combination was modeled using appropriate logical AND-gate functions (Maiwald et al., 2016). Dynamics of the proteins Pou5f3, Nanog and the sum of SoxB1 class genes SOX were taken from the core model for the different experimental conditions and used as input functions for the mini models. To avoid numerical issues with infinite derivatives of hill-kinetic terms, a small number of 1e-4 was added to the input functions where necessary.

Analogously to the core model, RNA-seq data of seven experimental conditions was used for model fitting of each mini model and for each target gene. Per target gene, a total number of 192 was available. In contrast to the core model, the different experimental conditions were not incorporated via boolean variables, instead the differential equations stayed the same over all conditions. Thus, differences in

dynamics were only possible via the differences in the given input dynamics of the core model.

For each mini model and for each target gene, a total number of 64 optimization runs was performed out of which 99.97% converged and at least 98.67% reached the global optimum. Finally, using the Bayesian Information criterion (BIC), out of the 19 analyzed mini models, the best fitting one was selected for every target. Parameter values of the selected models are summarized in Table S7.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical comparisons of three or more samples were performed using one-way ANOVA, followed by Tukey-Kramer Test to evaluate the statistical significance of pairwise differences. Two samples were compared using Student 2-tailed t-test.

Supplemental item titles.

Table S1. Accessible_regions.xlsx
Related to main Fig. 1, Fig. 2, Fig. 3, Fig. 4, Fig. S3, Fig. S4, the legend in the file

Table S2. Zygotic_transcripts.xlsx
Related to main Fig. 2, Fig. 5, Fig. 6, Fig. S2, Fig. S6, Fig. S7, the legend in the file.

Table S3. ARs_linked_to_transcripts.xlsx
Related to main Fig. 2, Fig. 5, Fig. 6, Fig. 7, Fig. S7, the legend in the file.

Table S4. Coremodel_equations.csv
Related to Fig.S5 and Fig.S6
.

Table S5. Coremodel_Bestfitparameters.csv.
Related to Fig.S5 and Fig.S6.

Table S6. Minimodel_equations.csv
Related to main Fig. 5, Fig. S6 and Fig. S7.

Table S7. Minimodel_parameters.csv.
Related to main Fig. 5, Fig. S6 and Fig. S7
.

References

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., *et al.* (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res *44*, W3-W10.

Alon, U. (2007). Network motifs: theory and experimental approaches. Nature Reviews Genetics *8*, 450-461.

Barnett, D.W., Garrison, E.K., Quinlan, A.R., Stromberg, M.P., and Marth, G.T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics *27*, 1691-1692.

Bates, L.E., Alves, M.R.P., and Silva, J.C.R. (2021). Auxin-degron system identifies immediate mechanisms of OCT4. Stem Cell Reports *16*, 1818-1831.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., *et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature methods *10*, 1213-1218.

Chen, J., Zhang, Z., Li, L., Chen, B.C., Revyakin, A., Hajj, B., Legant, W., Dahan, M., Lionnet, T., Betzig, E., *et al.* (2014). Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. Cell *156*, 1274-1285.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Domazet-Loso, T., and Tautz, D. (2010). A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. Nature *468*, 815-818.

Donaghey, J., Thakurela, S., Charlton, J., Chen, J.S., Smith, Z.D., Gu, H., Pop, R., Clement, K., Stamenova, E.K., Karnik, R., *et al.* (2018). Genetic determinants and epigenetic effects of pioneer-factor occupancy. Nature Genetics *50*, 250-258.

Donovan, B.T., Huynh, A., Ball, D.A., Patel, H.P., Poirier, M.G., Larson, D.R., Ferguson, M.L., and Lenstra, T.L. (2019). Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. EMBO J *38*.

Dufourt, J., Trullo, A., Hunter, J., Fernandez, C., Lazaro, J., Dejean, M., Morales, L., Nait-Amer, S., Schulz, K.N., Harrison, M.M., *et al.* (2018). Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. Nat Commun *9*, 5194.

Esch, D., Vahokoski, J., Groves, M.R., Pogenberg, V., Cojocaru, V., Vom Bruch, H., Han, D., Drexler, H.C., Arauzo-Bravo, M.J., Ng, C.K., *et al.* (2013). A unique Oct4 interface is crucial for reprogramming to pluripotency. Nature cell biology *15*, 295-301.

Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nat Protoc *7*, 1728-1740.

Fernandez Garcia, M., Moore, C.D., Schulz, K.N., Alberto, O., Donague, G., Harrison, M.M., Zhu, H., and Zaret, K.S. (2019). Structural Features of Transcription Factors Associating with Nucleosome Binding. Mol Cell *75*, 921-932 e926.

Friman, E.T., Deluz, C., Meireles-Filho, A.C., Govindan, S., Gardeux, V., Deplancke, B., and Suter, D.M. (2019). Dynamic regulation of chromatin accessibility by pluripotency transcription factors across the cell cycle. Elife *8*.

Fukaya, T., Lim, B., and Levine, M. (2016). Enhancer Control of Transcriptional Bursting. Cell *166*, 358-368.

Gao, M., Veil, M., Rosenblatt, M., Riesle, A.J., Gebhard, A., Hass, H., Buryanova, L., Yampolsky, L.Y., Gruning, B., Ulianov, S.V., *et al.* (2022). Pluripotency factors determine gene expression repertoire at zygotic genome activation. Nat Commun *13*, 788.
Gaskill, M.M., Gibson, T.J., Larson, E.D., and Harrison, M.M. (2021). GAF is essential for zygotic genome activation and chromatin accessibility in the early Drosophila embryo. Elife *10*.
Gassler, J., Kobayashi, W., Gaspar, I., Ruangroengkulrith, S., Mohanan, A., Gomez Hernandez, L., Kravchenko, P., Kummecke, M., Lalic, A., Rifel, N., *et al.* (2022). Zygotic genome activation by the totipotency pioneer factor Nr5a2. Science, eabn7478.
Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017-1018.
Heyn, P., Kircher, M., Dahl, A., Kelso, J., Tomancak, P., Kalinka, A.T., and Neugebauer, K.M. (2014). The earliest transcribed zygotic genes are short, newly evolved, and different across species. Cell reports *6*, 285-292.
Hiller, M., Agarwal, S., Notwell, J.H., Parikh, R., Guturu, H., Wenger, A.M., and Bejerano, G. (2013). Computational methods to detect conserved non-genic elements in phylogenetically isolated genomes: application to zebrafish. Nucleic Acids Res *41*, e151.
Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.
Huang, S.K., Whitney, P.H., Dutta, S., Shvartsman, S.Y., and Rushlow, C.A. (2021). Spatial organization of transcribing loci during early genome activation in Drosophila. Curr Biol *31*, 5102-5110 e5105.
Iwafuchi-Doi, M., and Zaret, K.S. (2016). Cell fate control by pioneer transcription factors. Development *143*, 1833-1837.
Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., *et al.* (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *458*, 362-366.
Kaschek, D., Mader, W., Fehling-Kaschek, M., Rosenblatt, M., and Timmer, J. (2019). Dynamic Modeling, Parameter Estimation, and Uncertainty Analysis in R. Journal of Statistical Software *88*, 1 - 32.
Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., and Schilling, T.F. (1995). Stages of embryonic development of the zebrafish. Dev Dyn *203*, 253-310.
Kuznetsova, K., Chabot, N.M., Ugolini, M., Wu, E., Lalit, M., Oda, H., Sato, Y., Kimura, H., Jug, F., and Vastenhouw, N.L. (2022). Nanog organizes transcription bodies. Curr Biol.
Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.
Lee, M.T., Bonneau, A.R., Takacs, C.M., Bazzini, A.A., Divito, K.R., Fleming, E.S., and Giraldez, A.J. (2013). Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. Nature, 360-364.
Leichsenring, M., Maes, J., Mossner, R., Driever, W., and Onichtchouk, D. (2013). Pou5f1 transcription factor controls zygotic gene activation in vertebrates. Science *341*, 1005-1009.
Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
Lunde, K., Belting, H.G., and Driever, W. (2004). Zebrafish pou5f1/pou2, homolog of mammalian Oct4, functions in the endoderm specification cascade. Current biology : CB *14*, 48-55.
Maiwald, T., Hass, H., Steiert, B., Vanlier, J., Engesser, R., Raue, A., Kipkeew, F., Bock, H.H., Kaschek, D., Kreutz, C., *et al.* (2016). Driving the Model to Its Limit: Profile Likelihood Based Model Reduction. PLOS ONE *11*, e0162366.

Miao, L., Tang, Y., Bonneau, A.R., Chan, S.H., Kojima, M.L., Pownall, M.E., Vejnar, C.E., Gao, F., Krishnaswamy, S., Hendry, C.E., *et al.* (2022). The landscape of pioneer factor activity reveals the mechanisms of chromatin reprogramming and genome activation. Mol Cell *82*, 986-1002 e1009.
Mulas, C., Chia, G., Jones, K.A., Hodgson, A.C., Stirparo, G.G., and Nichols, J. (2018). Oct4 regulates the embryonic axis and coordinates exit from pluripotency and germ layer specification in the mouse embryo. Development *145*.
Neyfakh, A.A. (1964). Radiation Investigation of Nucleo-Cytoplasmic Interrelations in Morphogenesis and Biochemical Differentiation. Nature *201*, 880-884.
Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet *24*, 372-376.
Okuda, Y., Ogura, E., Kondoh, H., and Kamachi, Y. (2010). B1 SOX coordinate cell specification with patterning and morphogenesis in the early zebrafish embryo. PLoS Genetics *6*, e1000936.
Onichtchouk, D., and Driever, W. (2016). Zygotic Genome Activators, Developmental Timing, and Pluripotency. Current topics in developmental biology *116*, 273-297.
Onichtchouk, D., Geier, F., Polok, B., Messerschmidt, D.M., Mossner, R., Wendik, B., Song, S., Taylor, V., Timmer, J., and Driever, W. (2010). Zebrafish Pou5f1-dependent transcriptional networks in temporal control of early development. Mol Syst Biol *6*, 354.
Palfy, M., Schulze, G., Valen, E., and Vastenhouw, N.L. (2020). Chromatin accessibility established by Pou5f3, Sox19b and Nanog primes genes for activity during zebrafish genome activation. PLoS Genet *16*, e1008546.
Popp, A.P., Hettich, J., and Gebhardt, J.C.M. (2021). Altering transcription factor binding reveals comprehensive transcriptional kinetics of a basic gene. Nucleic Acids Res *49*, 6249-6266.
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.
Radzisheuskaya, A., and Silva, J.C. (2014). Do all roads lead to Oct4? the emerging concepts of induced pluripotency. Trends in cell biology *24*, 275-284.
Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-165.
Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., and Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics *25*, 1923-1929.
Raue, A., Schilling, M., Bachmann, J., Matteson, A., Schelker, M., Kaschek, D., Hug, S., Kreutz, C., Harms, B.D., Theis, F.J., *et al.* (2013). Correction: Lessons Learned from Quantitative Dynamical Modeling in Systems Biology. PLOS ONE *8*, 10.1371/annotation/ea0193d1378-1371f1377f-1492a-b1370b1377-d877629fdcee.
Schulz, K.N., and Harrison, M.M. (2018). Mechanisms regulating zygotic genome activation. Nature Reviews Genetics.
Segert, J.A., Gisselbrecht, S.S., and Bulyk, M.L. (2021). Transcriptional Silencers: Driving Gene Expression with the Brakes On. Trends Genet *37*, 514-527.
Suter, D.M., Molina, N., Gatfield, D., Schneider, K., Schibler, U., and Naef, F. (2011). Mammalian genes are transcribed with widely different bursting kinetics. Science *332*, 472-474.
Tiana, M., Lopez-Jimenez, E., de Aja, J.S., Barral, A., Victorino, J., Badia-Careaga, C., Rollan, I., Rouco, R., Santos, E., Sanchez-Iranzo, H., *et al.* (2022). Pluripotency factors regulate the onset of Hox cluster activation in the early embryo. Sci Adv *8*, eabo3583.
Torres-Padilla, M.E., and Chambers, I. (2014). Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. Development *141*, 2173-2181.

Vastenhouw, N.L., Cao, W.X., and Lipshitz, H.D. (2019). The maternal-to-zygotic transition revisited. Development *146*.
Veil, M., Schaechtle, M.A., Gao, M., Kirner, V., Buryanova, L., Grethen, R., and Onichtchouk, D. (2018). Maternal Nanog is required for zebrafish embryo architecture and for cell viability during gastrulation. Development *145*.
Veil, M., Yampolsky, L., Gruening, B., and Onichtchouk, D. (2019). Pou5f3, SoxB1, and Nanog remodel chromatin on High Nucleosome Affinity Regions at Zygotic Genome Activation. Genome Res.
Westerfield, M. (2000). The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio). 4th ed edn (Univ. of Oregon Press, Eugene.).
White, R.J., Collins, J.E., Sealy, I.M., Wali, N., Dooley, C.M., Digby, Z., Stemple, D.L., Murphy, D.N., Billis, K., Hourlier, T., *et al.* (2017). A high-resolution mRNA expression time course of embryonic development in zebrafish. Elife *6*.
Xu, C., Fan, Z.P., Muller, P., Fogley, R., Dibiase, A., Trompouki, E., Unternaehrer, J., Xiong, F., Torregroza, I., Evans, T., *et al.* (2012). Nanog-like Regulates Endoderm Formation through the Mxtx2-Nodal Pathway. Developmental cell *22*, 625-638.