bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Pluripotency factors select gene expression repertoire at Zygotic Genome Activation

Meijiang Gao*[,1], Marina Veil*[,1], Marcus Rosenblatt[2], Anna Gebhard[1], Helge Hass[2], Lenka Buryanova[1], Lev Y. Yampolsky[3,4], Björn Grüning[5,6], Jens Timmer[2,7], Daria Onichtchouk**[1,7,8]

***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]Department of Biological Sciences, East Tennessee State University, Johnson City, Tennessee 37614-1710, USA

[4]Zoological Institute, Basel University, Basel CH-4051, Switzerland

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

[6]Center for Biological Systems Analysis (ZBSA), Albert-Ludwigs-University of Freiburg, 79110 Freiburg, Germany

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

[8] Koltzov Institute of Developmental Biology RAS, 119991 Moscow, Russia

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Awakening of zygotic transcription in animal embryos relies on maternal pioneer transcription factors. The interplay of global and specific functions of these proteins remains poorly understood. Here, we analyzed nucleosome positioning, H3K27 acetylation, transcription, and gastrulation rates in zebrafish embryos lacking pluripotency factors Pou5f3 and Sox19b. We show that the bulk transcriptional onset does not require Sox19b and Pou5f3, but is sensitive to their balance. Pou5f3 docks H3K27ac on the enhancers of genes involved in gastrulation and ventral fate specification. Sox19b facilitates Pou5f3 access to one-third of these enhancers. The genes regulating mesendodermal and dorsal fates are primed for activation independently on Pou5f3 and Sox19b. Strikingly, the loss of either factor results in activation of silent enhancers; simultaneous loss of both leads to premature expression of differentiation genes. Our results uncover how independent activities of maternal Pou5f3 and Sox19b add up or antagonize to determine the early gene expression repertoire.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Following fertilization, the differentiated cells, egg and sperm, are reprogrammed into the totipotent state of the resulting zygote. The genome remains transcriptionally silent, while reprogramming takes place. The zygotic genome awakens through a process known as maternal-to-zygotic transition, during which the degradation of maternal transcripts is coordinated with Zygotic Genome Activation (ZGA). The current models of ZGA take into account the gradual increase in the ratio of transcriptional activators to transcriptional repressors, accompanied with local changes of chromatin accessibility, which together create a permissive environment for transcription[1,2]. In zebrafish, *Xenopus* and *Drosophila*, where development starts with rapid cell cycles, excessive maternal core histones serve as general transcriptional repressors before ZGA[3-5]. Activators that are translated after fertilization and reach critical levels at ZGA include three types of proteins: basal transcription factors[6,7], the regulators of H3K27ac enhancer mark[8,9] and maternal enhancer-binding transcription factors (TFs). TFs that broadly activate zygotically expressed genes have been identified in *Drosophila*[10], zebrafish[11,12], *Xenopus*[13], human[14] and mouse[15].

Nucleosome positioning plays a dominant role in regulating DNA access by TFs[16]. The widespread action of sequence-specific genome activators is thought to reside on their ability to act as pioneer factors, first displacing nucleosomes, so that other TFs can bind[17-19]. In case of several genome activators, like Pou5f3, Sox19b and Nanog in zebrafish, it remains an open question, how their broad pioneer activity at ZGA relates to their different functions later in development. In zebrafish embryos, a block of maternal and zygotic Pou5f3 function leads to epiboly defects and arrest during gastrulation[20], while the block of redundant SoxB1 family members (Sox19a, Sox19b, Sox3 and Sox2) leads to developmental arrest at the beginning of organogenesis[21].

Mammalian homologs of zebrafish genome activators, Pou5f1 and Sox2, are used for reprogramming of somatic cells to pluripotency and are proven sufficient for reprogramming in several cases[20,22,23]. The mechanisms underlying their partnership *in-vivo* are still not resolved. Until recently, Pou5f1 and Sox2 were thought to act cooperatively, binding as heterodimers to bipartite *pou:sox* cognate motifs[24,25]. This view was challenged by four studies which suggested four different scenarios of how Pou5f1 and Sox2 interact with each other and with chromatin in ES cells[26-29]. The embryonic phenotypes of mouse Pou5f1 and Sox2 mutants are different, suggesting unique biological roles of each TF[30,31].

Before the terms MZT and ZGA were coined[32-34], Alexander Neyfakh described two distinct periods of zygotic gene activity in teleost fish[35,36]. The earliest period spanned from mid to late blastula and provided the instructions for whole gastrulation. The second period started at midgastrula and provided the instructions for organogenesis. From the standpoint of Neyfakh's concept, Pou5f3 function in zebrafish is critical for the earliest zygotic period, and SoxB1 for the second. The earlier function for SoxB1 proteins was suggested by dominant-negative approach[37], and by combined morpholino knockdowns of SoxB1 with Nanog and/or Pou5f3[11]. The mechanisms of SoxB1 activity at ZGA and its molecular connection to Pou5f3 were not characterized. In this study, we investigate epistatic relationships between Pou5f3 and Sox19b, the only maternally expressed member of the SoxB1 family

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Results

ZGA onset and gastrulation timing depends on the maternal Sox19b/Pou5f3 balance.

To abolish the expression of Sox19b in zebrafish, we introduced a mutation in *sox19b* by using the TALEN technique[38](Fig. 1a). The MZ*sox19b* embryos lacking both maternal and zygotic Sox19b developed into fertile adults. However, they developed slower than controls (Movie S1-1).

Zebrafish embryos are first going through 9 synchronous cell cycles and a longer 10th cell cycle, followed by ZGA at 3 hours post-fertilization (hpf)[39,40]. Duration of cycles 6-9 was the same in MZ*sox19b* and wild-type (Fig. S1a), but the 10th cell cycle was detectably longer in MZ*sox19b* (Fig.1b). The time gap between MZ*sox19b* and wild-type increased up to two hours at the end of gastrulation, with no further delay during organogenesis (Fig. 1c, d).

Normal development of MZ*sox19b* embryos is plausibly explained by the presence of zygotic SoxB1 factors (Fig.2a, Fig.S1). We removed zygotic SoxB1 activity in MZ*sox19b* mutants to address if SoxB1 function is required for gastrulation. SoxB1-deficient embryos showed only post-gastrulation defects (Fig. 2b, Fig.S1d). We concluded that maternal and zygotic functions of SoxB1 proteins are uncoupled: zygotic SoxB1 activity becomes critical starting from the end of gastrulation, maternal Sox19b protein sets the timing of gastrulation.

To investigate the interplay of maternal Sox19b with Pou5f3, we obtained a double mutant MZ*sox19bspg* by crossing MZ*sox19b* to Pou5f3 mutant MZ*spg* [20]. MZ*spg*^{m793} mutants develop severe epiboly defects[41,42] and are weakly dorsalized[43,44]. Epiboly defects in double mutant were similar to MZ*spg* (Fig. 2c, Fig. S1e), while dorsalization was stronger (Fig.S1f). Unexpectedly, the double mutants were less delayed in gastrulation than the single mutants (Movie S1-2, 3). We produced maternal-only M*sox19bspg* mutants by fertilizing the mutant eggs with wild-type sperm, to find out if a combined maternal contribution of Pou5f3 and Sox19b is critical for developmental timing, epiboly, or dorso-ventral (D/V) patterning. Single maternal mutants were delayed in gastrulation (in order M*sox19b*>M*spg*>WT) but developed normally. The double mutant embryos were not delayed (M*sox19bspg*=WT), but were dorsalized (Fig. 2d,e, Movie S1-4, 5). The genetic compensation of developmental delays in the double mutant suggested that maternal Sox19b and Pou5f3 balance the developmental rate (Fig. 2f). The dorsalization of the double mutant suggested that either Pou5f3 or Sox19b should be maternally present to safeguard normal dorso-ventral patterning of the embryo (Fig. 2g). Finally, Pou5f3 was required for epiboly, while SoxB1 was dispensable (Fig. 2h). To understand the molecular nature of Pou5f3 and Sox19b interactions, we examined the changes in chromatin accessibility, histone modifications and transcription in the mutants.

Sox19b and Pou5f3 act as independent pioneer factors

Zygotic genome activators Pou5f3 and SoxB1 and Nanog recognize their consensus motifs in the context of High Nucleosome Affinity Regions (HNARs)[47]. Pou5f3, Sox19b, and Nanog act as pioneer factors, displacing nucleosomes from the regions they bind[45-47].

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We performed MNase-seq experiment on 4.3 hpf MZ*sox19b* mutants to compare nucleosome positioning in MZ*sox19b*, MZ*spg* and the wild-type[47]. We ranked genomic regions bound by Pou5f3, SoxB1 and Nanog (PSN)[47] by descending chromatin accessibility in the wild-type (Fig.3a) and took HNAR centers as summits (Fig.3b). To estimate nucleosome displacement by each TF, we subtracted the wild-type MNase signals from those in the respective mutant (Fig. 3cd, Fig.S1). Pou5f3 displaced nucleosomes on the most accessible, "open" regions, while Sox19b affected both "open" and "closed" regions (Fig. 3cd, red arrowheads, Fig. S1). This suggested that Sox19b is able to engage into inaccessible chromatin more efficiently than Pou5f3, perhaps by using different recognition cues. SoxB1 proteins can recognize two types of cognate sequences: bipartite *pou:sox* motifs, also bound by Pou5f3, and *sox* motifs[47,67](Fig.3e). Pou5f3 and Sox19b can also interact with non-consensus DNA recognition cues within HNARs (Fig.3f, Fig.S1).

We assumed that, if the effects of TF were sequence-specific, the strength of the nucleosome displacement would increase with the number of cognate motifs per genomic region occupied by TF. We then compared the mean nucleosome displacement values by each TF for the PSN regions with no motifs, one motif, two motifs, three or more motifs of each type (*sox* or *pou:sox*, Fig.3e). We found that nucleosome displacement by Sox19b strongly increases with the number of *sox* motifs and weaker with *pou:sox* motifs per region. Pou5f3 displaced nucleosomes only on *pou:sox* motifs (Fig. 3g, Table S1 for statistics). These results fit two scenarios: either Sox19b strongly binds to *sox* and weakly to *pou:sox* motifs, or Sox19b facilitates Pou5f3 binding to some *pou:sox* motifs.

To infer non-consensus binding effects of Sox19b and Pou5f3 on chromatin, we ranked the genomic regions by ascending *in-vitro* predicted nucleosome occupancy and compared the mean nucleosome displacement values. Nucleosome displacement by Sox19b was directly proportional to *in-vitro* predicted nucleosome occupancy, while nucleosome displacement by Pou5f3 showed the opposite tendency, on PSN regions and genome-wide (Fig. 3h, Fig. S2, Table S1 for statistics).

In sum, we found that Sox19b and Pou5f3 act as independent pioneer factors during maternal-to-zygotic transition. They displace nucleosomes by binding to different consensus motifs. Non-consensus binding of two factors results in the opposing effects of local chromatin accessibility genome-wide.

Genome-wide H3K27ac changes in MZ*sox19b* and MZ*spg* mutants define four types of differentially regulated embryonic enhancers.

In parallel with the increase of chromatin accessibility during ZGA, zebrafish genome acquires H3K27ac histone tail mark on active enhancers[8,9,48]. To test if histone mark deposition depends on Pou5f3 and Sox19b, we immunoprecipitated embryonic chromatin originating from of the WT, MZ*spg* and MZ*sox19b* late blastula stage embryos with H3K27ac antibodies, sequenced and compared the mapped data.

To infer the genome-wide enhancer activity changes in the mutants, we identified H3K27ac peaks in each genotype, removed promoters and clustered unique peaks by H3K27ac signal in the three genotypes (Table S2). This resulted in five clusters of peaks: 1) Pou5f3-activated, 2) Pou5f3 and Sox19b – activated, 3) Pou5f3-repressed, 4) Sox19b-repressed and 5) independent on Pou5f3 and Sox19b (Fig. 4a, Fig. S3a).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

H3K27ac peaks of clusters 1, 2, 4 and 5 associated with genes expressed during gastrulation, while cluster 3 associated with neural patterning genes expressed during organogenesis in GREAT[49](Fig.4a, Table S2). Genomic regions around H3K27ac peaks were enriched in enhancer features: they were accessible by ATAC-seq, bound by maternal TFs, and had elevated *in-vitro* nucleosome prediction values (Fig.S3).

In sum, we delineate five enhancer types by genome-wide redistribution of H3K27ac mark in MZ*spg* and MZ*sox19b* mutants. We further focused on differential enhancers, where H3K27 acetylation depended on Pou5f3 or Sox19b (types 1- 4).

Pou5f3 pioneer activity, independent on Sox19b on type 1 enhancers, or assisted by Sox19b on type 2 enhancers, induces local H3K27 acetylation.
To distinguish between direct and indirect effects of Pou5f3 and Sox19b pioneer activity on H3K27ac, we inquired if the nucleosome displacement by Pou5f3 and Sox19b induces the local H3K27ac change. We selected PSN binding regions within +/- 500 bp from H3K27ac peaks and took HNAR centers as summits. We compared the H3K27ac occupancy in three genotypes with the nucleosome displacement by each TF, and with the density of *sox* and *pou:sox* motifs (Fig. 4b, Fig. S3e,f). Nucleosome displacement by Pou5f3 and enrichment in *pou:sox* motifs was detectable on Pou5f3-activated (type 1) and Pou5f3 and Sox19b-activated (type 2) enhancers (Fig. 4b). Hence, Pou5f3 binding to type 1 and 2 enhancers displaced nucleosomes and directly induced local H3K27 acetylation. Nucleosome displacement by Sox19b was dispensable for H3K27ac deposition on type1 enhancers, but was required on type 2 enhancers.

Sox19b and Pou5f3 can act on type 2 enhancers by either co-binding *pou:sox* motif, or by binding different motifs. To distinguish between the scenarios, we asked if H3K27ac deposition by each TF depended on the presence of *sox* or *pou:sox* motifs in the genomic regions. We found that local H3K27ac deposition by Sox19b depended only on *sox* motifs, and by Pou5f3 only on *pou:sox* motifs (Fig.4c, Table S1). The most parsimonious explanation for type 2 enhancers is that Sox19b binding to *sox* motif facilitates Pou5f3 loading on *pou:sox* motifs nearby. Pou5f3 acts downstream of Sox19b, binds to *pou:sox* motifs and induces local H3K27 acetylation.

Pou5f3 and Sox19b indirectly prevent H3K27ac deposition on enhancer types 3 and 4.
Repression of H3K27ac by Pou5f3 (type 3 enhancers), or by Sox19b (type 4 enhancers) was most likely indirect, as it was not associated with pioneer activity of repressing TFs or enrichment for their motifs (Fig.4b, Fig.S3). However, chromatin on type 3 and 4 enhancers was accessible (Fig. S3). Sox19b displaced nucleosomes on *sox* motifs on type 3 enhancers (Fig.4b, Fig.S3), suggesting that SoxB1 factors may activate some of them. Since type 4 enhancers were bound by Nanog[50] and two maternal TFs implicated in mesendoderm specification, Eomesodermin[51-53] and FoxH1[54,55](Fig. S3), we assumed that these factors may displace nucleosomes on type 4 enhancers at ZGA.

Four types of differential enhancers define four types of transcriptional response to Pou5f3 and Sox19b.
As a rule, all enhancers types were intermingled in the proximity of developmental genes (Table S2), suggesting complex combinatorial regulation. Only in a handful of

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cases, H3K27ac was changed over 5-50 kb around a gene according to one enhancer type pattern. To get clear regulatory examples, we picked several "top genes" for each differential enhancer type (Fig.4d-g and Fig. S4). Notably, 7/13 top cluster 3 genes were previously characterized SoxB1 transcriptional targets[21] and 4/7 top cluster genes were Eomesodermin/FoxH1/Nodal targets (Fig.S4). Using selected "top genes", we addressed the question, if H3K27ac changes at four differential enhancer types are predictive for the transcriptional outcome in the mutants. To compare the single and double mutant effects on gene expression, we performed RNA-seq time-series analysis using the wild-type (w), MZ*sox19b* (s), MZ*spg* (p) and double (d) mutants (Fig.5a). The change of H3K27ac on enhancer types 1-3 was predictive for transcriptional regulation of selected top genes; top type 4 genes were upregulated in the double mutant at most (Fig.5b, compare to Fig.4d-g; Fig.S5, compare to Fig.S4). Our next step was to extend time-resolved transcriptional analysis to the whole genome.

RNA-seq time resolved analysis reveals that bulk transcriptional timing at ZGA is sensitive to the Sox19b/Pou5f3 balance.

Two processes shape the transcriptional landscape of embryos at ZGA: the burst of zygotic transcription and regulated decay of maternal mRNAs. About 70% of zebrafish mRNAs are maternally loaded[56] so that the mRNA present in the embryo at each given early time point is a mix of maternal and zygotic transcripts for most genes. To account for the maternal and zygotic differences between the wild-type and mutants, we developed a novel tool for dynamic RNA-seq data analysis, which we called RNA-sense.

The RNA-sense tool compares RNA-seq time curves in two experimental conditions, *i.e*. the wild-type and mutant, and works in three steps. At Step 1, it builds a time profile for each transcript in one condition (*i.e.* wild-type) and tests if the transcript abundance grows or decays significantly. Dynamic transcripts are then sorted to non-overlapping groups by the time point of switch up (mostly zygotic transcripts), or switch down (mostly maternal transcripts). At Step 2, RNA-sense outputs the groups of differentially expressed transcripts, which are up- or downregulated in the mutant compared to the wild-type, for each time point. Correlations between the outputs of Step 1 and Step 2 are calculated at Step 3. The results of the correlation analysis are printed as two-dimensional color plot, with time profiles and differential expression groups at Y- and X-axis, which facilitates the interpretation of the data (Fig. S6).

"Switch up" and "switch down" time profile groups generated by RNA-sense matched to zygotic and maternal transcripts identified in three studies[11,56,57] (Fig. S6, Table S3), and will be referred as "zygotic" and "maternal" genes below.

RNA-sense analysis using the wild-type dynamic time profiles revealed general delays in zygotic transcription and maternal mRNA degradation in all mutants (Fig.5c, Fig.S6, Table S3). Unexpectedly, the depletion of single TFs resulted in a stronger transcriptional delay than simultaneous depletion of both TFs, indicating the genetic compensation of Pou5f3 and Sox19b in regard to the bulk ZGA timing (Fig.5d, Fig.S6). The general gene expression delays provided molecular explanation for gastrulation delays we observed, as the order of transcriptional delays MZ*sox19b* >MZ*spg*>MZ*sox19bspg* was mirrored by morphological delays (Movie S1). Taken together, our data uncovered unanticipated genome-wide antagonism between two

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pluripotency factors. Pou5f3/Sox19b balance appeared to be more important for timely onset of bulk transcription, than the presence of both factors in the embryo.

Pou5f3 and Sox19b are dispensable for the zygotic activation of maternal ß-catenin and Eomesodermin/FoxH1 target genes.

To distinguish between specific downregulation and general ZGA delay, we grouped the zygotic genes downregulated only in the single mutants, or in the double mutant (Fig. 5d, Fig. S7). Strikingly, the early dorsal targets of maternal ß-catenin and components of endomesoderm specification network were not repressed in the double mutant, while ventral specifying factors and BMP pathway components were repressed (Fig. 5e,f, Fig.S7). The repression in the double mutant was non-additive, as expected when genes are activated by one of the two TFs (Fig. S7). The proportion of type 1 and 2 enhancers in the regulatory regions exceeded genomic average only for the group of genes repressed in the double mutant (Fig.S7), indicating that only these genes are specifically activated by Pou5f3 and Sox19b.

Maternal Pou5f3 and Sox19b prevent the premature expression of late differentiation genes and restrict transcription of dorsal genes.

To take into account the transcripts present only in the mutants, we obtained the mutant time profiles with RNA-sense and scored the zygotic transcripts upregulated in the mutants at 5.5 hpf (Table S4). Zygotic transcripts, upregulated in the single or double mutants, had elevated proportions of type 3 and type 4 enhancers within the putative regulatory regions (Fig. S7). The group of transcripts upregulated in the double mutant was enriched in late developmental regulators, normally expressed during organogenesis, and in dorsal genes (Fig. 6a-c, Fig. S7). The regulatory regions of the genes, prematurely activated in the double mutant, gained H3K27ac in both single mutants at dome stage (Fig.S7). These results suggested that Pou5f3 and Sox19b synergistically prevent H3K27 acetylation and activation of two functionally distinct gene groups: the late developmental regulators and dorsal genes.

In zebrafish, similar to other vertebrates, Chordin blocks the flow of BMPs to the dorsal side of the embryo during gastrulation[58]. Noggin1 redundantly supports the action of Chordin[59,60]. The synergistic dorsalization in double maternal mutants (Fig. 2c,d) could either result from early overexpression of Noggin1 (Fig.5b) or, alternatively, from the disbalance between the dorsal and ventral gene regulatory networks, which converge on defining the size of the Chordin domain. The normal development of M*sox19bspg* could only be rescued by reduction of Chordin, but not of Noggin (Fig 6g). We concluded that synergistic activity of Pou5f3 and Sox19b at ZGA controls D/V axis formation by balancing the expression onset of multiple zygotic dorsal versus ventral developmental regulators.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

Gene expression starts as a result of a multi-step regulatory programme, in which changes in chromatin accessibility, histone modifications and transcription quickly follow each other[16]. In this study, we dissected the contributions of Pou5f3 and Sox19b to each of these three steps, resulting in four novel insights.

First, we found that Sox19b and Pou5f3 bind independently to DNA (Fig.3). Pou5f3 activates early zygotic genes, by binding to two types of enhancers – Pou5f3-dependent type 1 and Sox19b/Pou5f3 dependent type 2 (26% and 13% of active enhancers at blastula stage, respectively, Fig.4a). Sox19b assists Pou5f3 loading on type 2 enhancers (Fig.7a).

Second, we reveal unanticipated antagonism of two pioneer factors at ZGA. ZGA timing is defined by competition between histones and genome activators[3-5], and we expected that ablation of two major activators would additively delay ZGA. We observed exactly the opposite: transcriptional and developmental delays were compensated in the double mutant (Movie S1, Fig.5d, summarized at Fig.7b), revealing a hidden antagonism between Pou5f3 and Sox19b. As non-consensus binding of two factors results in the opposing effects on chromatin (Fig. 3h, Fig. S2), we speculate that Pou5f3 and Sox19 may mutually restrict each other's non-consensus binding to HNARs genome-wide (Supplementary Notes).

Third, we show that Pou5f3 and Sox19b do not prime dorsal and mesendodermal genes for activation at ZGA or even repress them (Fig.5, Fig.S6). Taken together, our observations suggest that the establishment of transcriptional competency at major ZGA is regional and mediated by different enhancer types (Fig.7b, see Supplementary Notes). This view is compatible with recent findings in mouse[61,62] and Xenopus[63,64], which suggested independent early epigenetic priming for different embryonic tissues.

Fourth, we show that genome activators do select rather than activate zygotic genes. The absence of maternal Pou5f3 and Sox19b results in activation of alternative set of enhancers (types 3 and 4) at ZGA (Fig.4a), leading to premature expression of genes which normally regulate tissue differentiation during organogenesis stages (Fig.6, Fig. 7d,e). From the point of view of Neyfach's concept[36], Pou5f3 and Sox19b ensure the time gap between the first and second periods of zygotic transcription by shutting down the expression of genes which elicit later developmental programs. Balancing of early and late developmental programs by Pou5f3 and Sox19b provides a direct parallel to reprogramming of fibroblasts, where Pou5f1 and Sox2 shut down the somatic gene expression and activate early embryonic genes[65]

Acknowledgments

We are grateful to Sebastian Arnold and Rainer Duden for commenting on the manuscript, to Sabine Götter for excellent fish care, and to Andrea Buderer and Cornelia Wagner for administrative support. DO is grateful to Wolfgang Driever for hosting her lab. This work was supported by DFG-ON86/4-2 for DO, and DFG-EXC2189 – Project ID: 390939984 for DO and JT. The Freiburg Galaxy Team is funded by DFG grant SFB 992/1 2012 and BMBF grant 031 A538A RBC.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Author Contributions

GM, MV, AG and LB performed the experiments; GM, BG, LY and DO analysed the data, MR and HH wrote RNA-sense program, DO - design of the study, JT and DO – supervision, funding acquisition, MV and DO wrote the manuscript, GM, MR, LY and JT edited the manuscript.

Declarations of Interests

The authors declare no competing interests

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure titles and legends

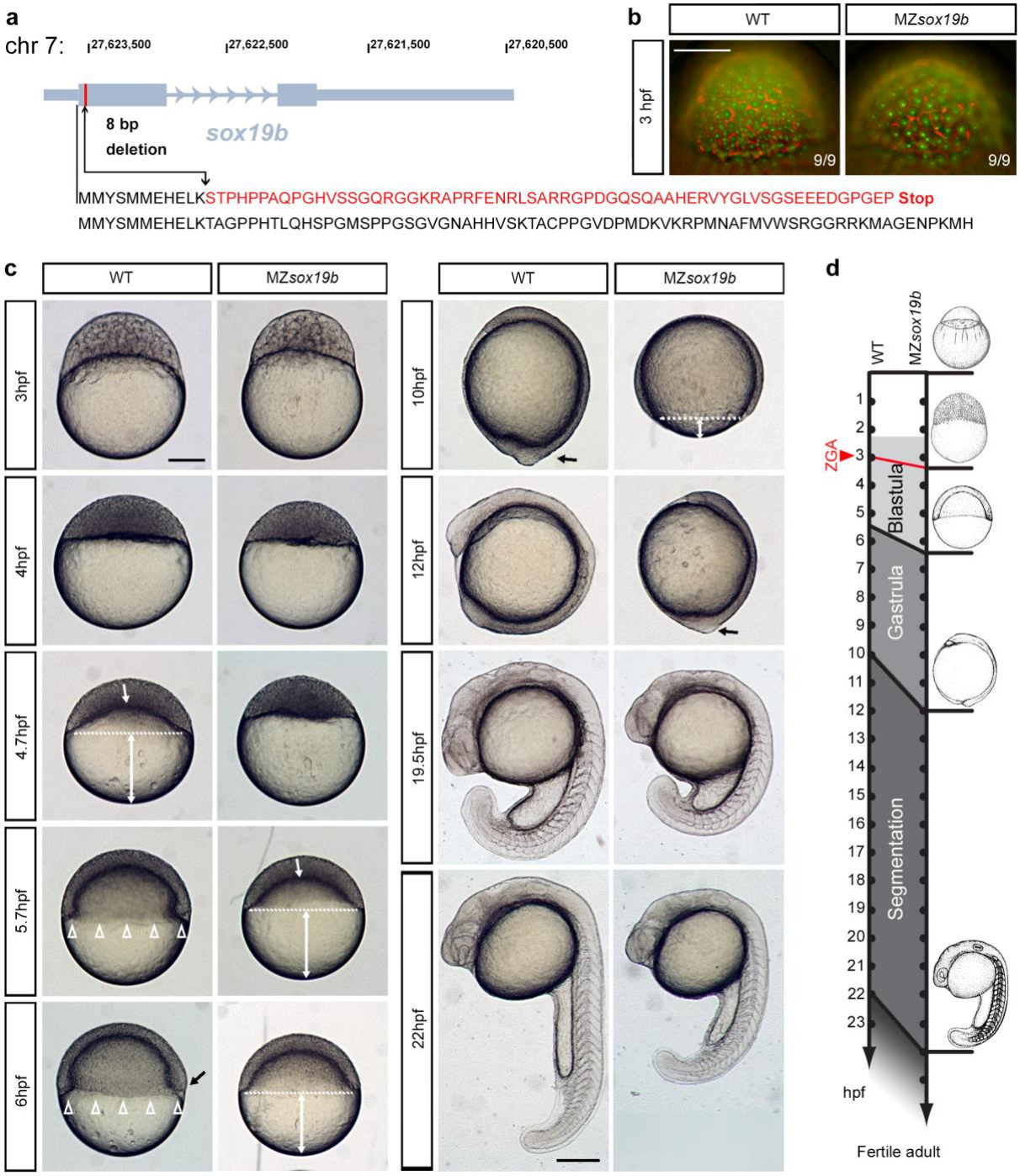

Figure 1: Mutation in maternal *sox19b* gene causes developmental delay starting at ZGA. a. Disruption of the *sox19b* gene on chromosome 7 by introducing an 8 bp deletion with TALEN technique. We chose an 8 bp deletion after the codon for amino acid 11 which resulted in introduction of a premature stop codon after another 62 amino acids. **b**. Difference in cell size and nuclei number between WT and MZ*sox19b* detectable at 3.0 hpf. SYTOX green- nuclei, rhodamine-phalloidin- submembrane cytoskeleton. To compare the duration of 10th cell cycle, we fixed WT and MZ*sox19b* embryos every 15 min starting from 2.5 hpf till 4 hpf, and compared the size of the cells and number of nuclei between the genotypes. We could detect that the MZ*sox19b* were still at the 512 cell stage (10th cell cycle), while the WT

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

embryos proceeded to the 1024 cell stage. MZ*sox19b* completed 10th cell cycle within the next 15 minutes. **c**. Phenotypes of the WT and MZ*sox19b* embryos at the indicated time points. Morphogenetic movements in zebrafish blastula start from doming of the yolk (white arrow) followed by epiboly. Double white arrows show the distance from epiboly border (white dotted line) to the vegetal pole. Involution of mesendodermal layer marks gastrulation onset (hollow arrowheads in the wild-type, 5.7 hpf and 6 hpf), is followed by shield formation (black arrow in the wild-type, 6 hpf). Gastrulation ends with tail bud formation (black arrow at 10 hpf, WT, and at 12 hpf in MZ*sox19b*). Scale bar 200 µm. **d**. Altered developmental timing in MZ*sox19b*: 10th cell cycle is delayed at <=15 min, gastrulation starts 45 min-1 hour* later and completes 2 hours later than in the control.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

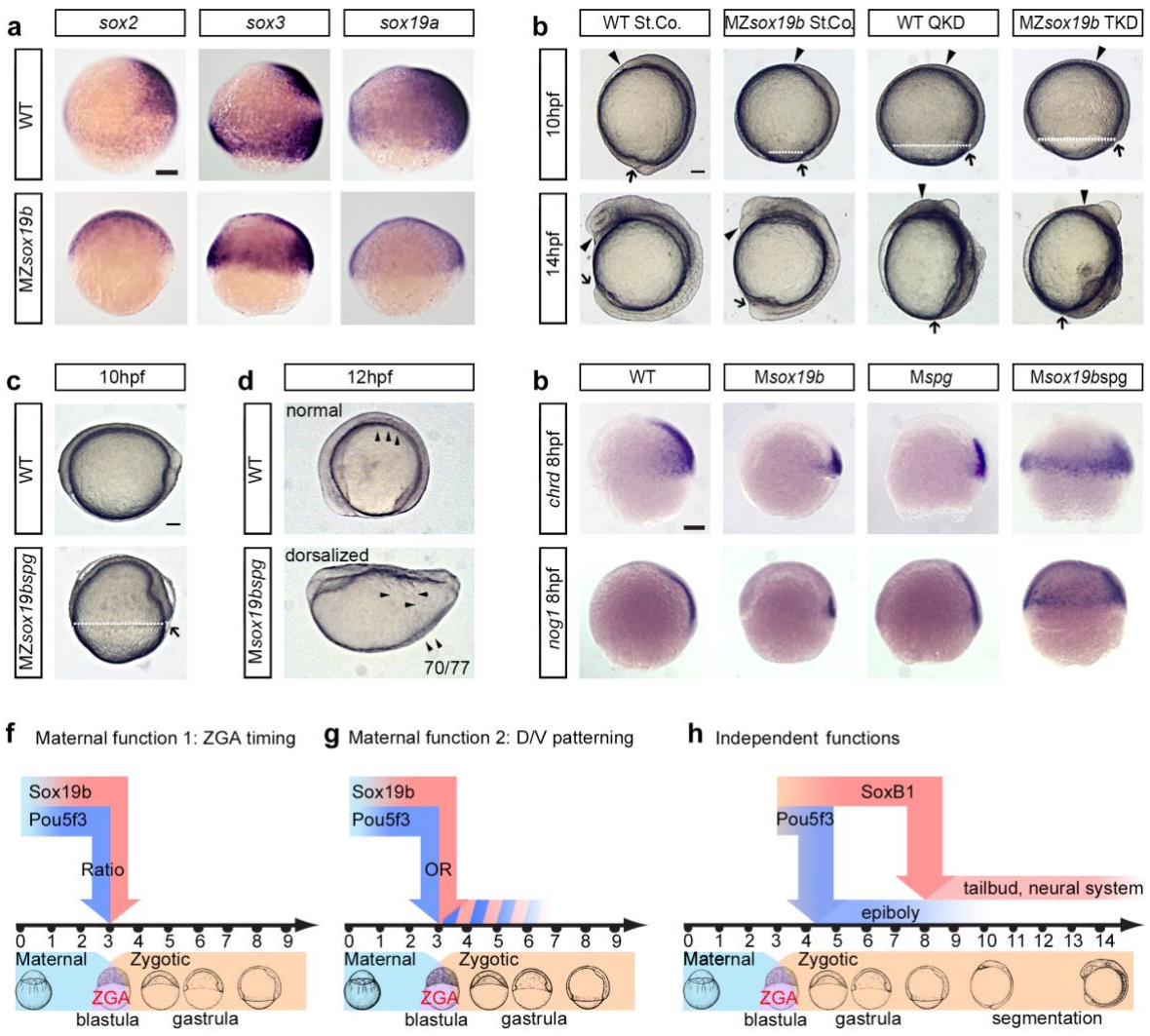

Figure 2: Common and separate biological functions of Pou5f3 and SoxB1 factors. a. *In situ* hybridization for *sox2, sox3 and sox19a*, 8 hpf, lateral views, dorsal to the right. **b.** SoxB1 knockdown embryos complete epiboly, but show defects in later stages, tailbud formation and axis elongation. Arrow - shield or tailbud position; arrowhead -head position; white dotted line - epiboly border. One cell stage wild-type or MZ*sox19b* embryos were injected with control Morpholino (StCo), or QKD (Quadriple Knockdown) mix (Sox2, Sox3, Sox19a and Sox19b Morpholinos[21]), or TKD (Triple Knockdown, Sox2, Sox3, Sox19a Morpholinos), as indicated. The early functions of SoxB1 genes could be masked by maternal Sox19b protein, which remains In WT- QKD[12], but not MZ*sox19b*-TKD. Axis elongation defects in MZ*sox19b*-TKD can be rescued by injection of 20pg *sox19b* mRNA (see Fig.S1d) **c**. MZ*sox19bspg* mutants are arrested in epiboly and dorsalized. White dotted line shows epiboly border, arrow shows enlarged shield. **d.** M*sox19bspg* +/- mutants are dorsalized. Somites (arrowheads) form on the dorsal side in WT, but spread over the M*sox19bspg* embryo. Dorsal up, anterior to the left. **e.** *In-situ* hybridization for the dorsal transcripts *noggin1* and *chordin.* Note the expansion of dorsal domain in the double mutants. **f-h.** Common and separate functions of Pou5f3 and Sox19b. **f.** Balance of maternal Pou5f3 and Sox19b is required for correct developmental timing. **g.** Maternal Sox19b and Pou5f3 safeguard correct D/V patterning. **h.** Independent

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

functions: Pou5f3 is critical for gastrulation, SoxB1 for organogenesis. Scale bars in A-E: 100 µm.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

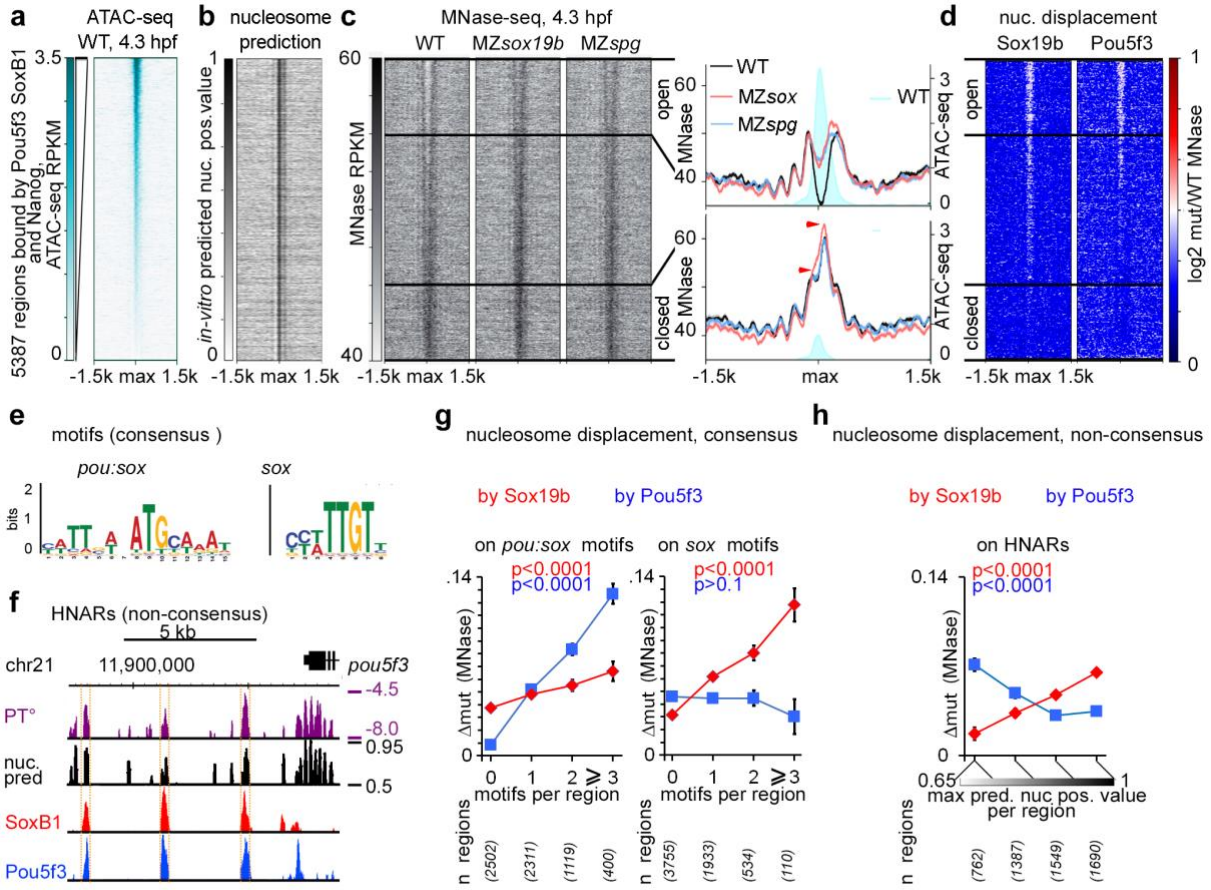

Figure 3: Sox19b and Pou5f3 bind independently across the genome. a-d. Heat maps (PSN) 3 kb genomic regions. **a.** ATAC-seq at 4.3 hpf[45] **b**. *in-vitro* predicted nucleosome occupancy[66]. **c.** Left: MNase-seq signals at 4.3 hpf. Right: summary MNase profiles of "open" and "closed" regions in the indicated genotypes. **d**. Nucleosome displacement by Pou5f3 and Sox19b. **e.** Consensus binding clues: motifs. SoxB1 proteins can recognize two types of cognate sequences: bipartite *pou:sox* motifs, also bound by Pou5f3, and *sox* motifs[47,67]. **f.** Non-consensus binding clues: High Nucleosome Affinity Regions (HNARs). HNARs are featuring high values of predicted *in-vitro* nucleosome occupancy (nuc.pred)[66] and DNA shape parameter Propeller Twist (PT°)[47]. **g**. Nucleosome displacement by Sox19b (red) and Pou5f3 (blue) on PSN regions, depending on the number of *pou:sox* (left) and *sox* (right) motifs per region. Numbers of PSN genomic regions with 0,1,2 and 3 and more motifs are indicated below the graphs. **h.** Non-consensus nucleosome displacement by Sox19b increases with the *in-vitro* predicted nucleosome positioning value of HNAR, Pou5f3 shows the opposite tendency. P-values for the 1-way ANOVA (Table S1). Nucleosome displacement by each TF in (g,h) was calculated as a normalized difference in MNase signals between respective mutant and the wild-type, per 320 bp region.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

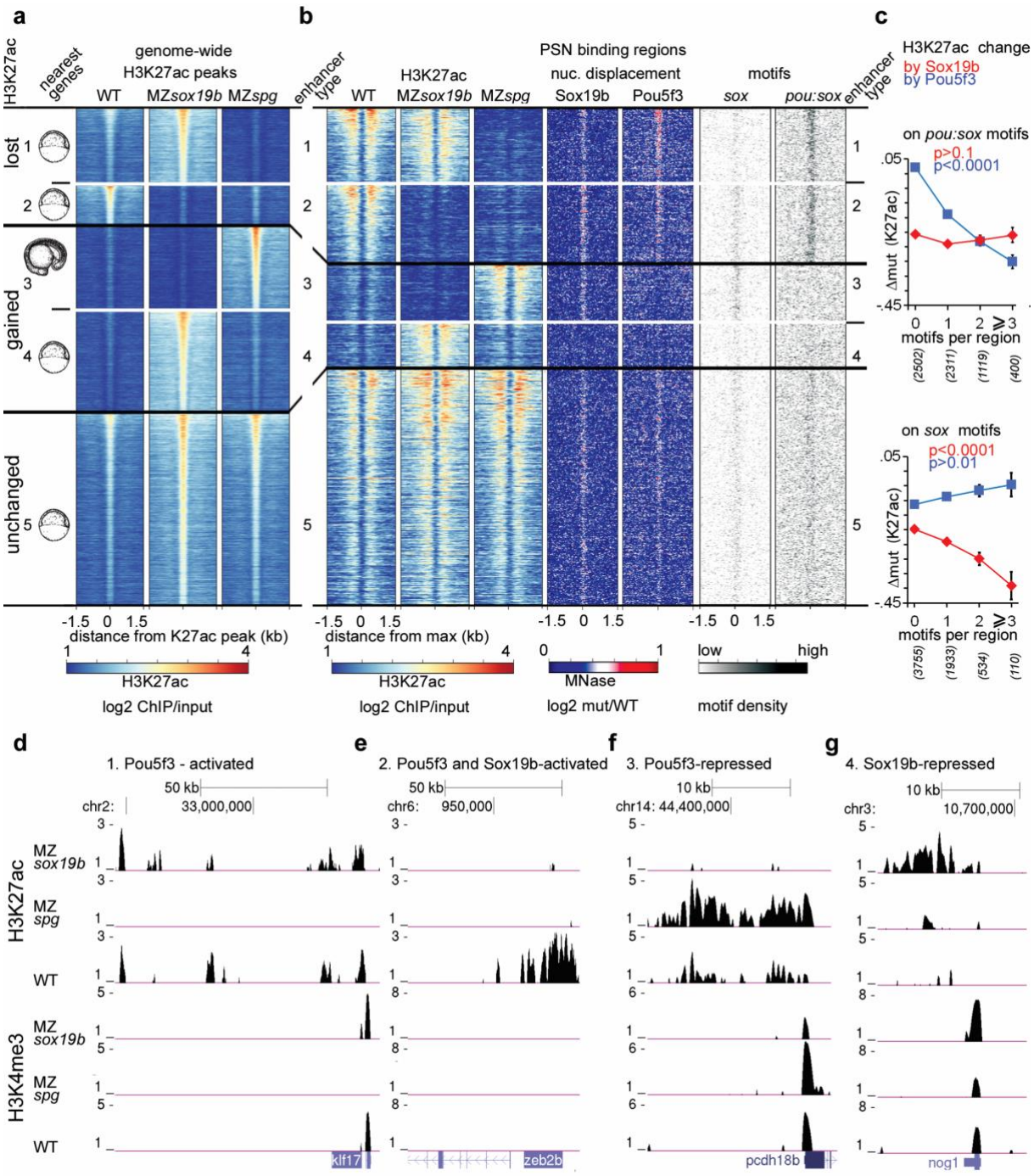

Figure 4: H3K27 acetylation is differentially regulated by Pou5f3 and Sox19b at four types of enhancers. a. 5 clusters of H3K27ac peaks. The icon at the left shows the developmental stage of expression of genes associated with each cluster in GREAT analysis[49] (Table S2). **b.** Heat maps of 5 enhancer types overlapping with PSN genomic regions, aligned on max. predicted nucleosome position of HNAR. **c**. H3K27ac change by Sox19b (red) and Pou5f3 (blue) on PSN regions, depending on the number of *pou:sox* (top) and *sox* (bottom) motifs per region. H3K27ac change by TF (ΔH3K27ac) was calculated as a normalized difference in H3K27ac signals between respective mutant and the wild-type, per each PSN region (+/- 500bp from max). The numbers of genomic regions with 0,1,2 and 3 and more motifs are indicated below the graphs, P-values for the 1-way ANOVA (Table S1). **d-g.** Genes

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

regulated by enhancer types 1-4 (UCSC genome browser, signal units are log2 ChIP/Input).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

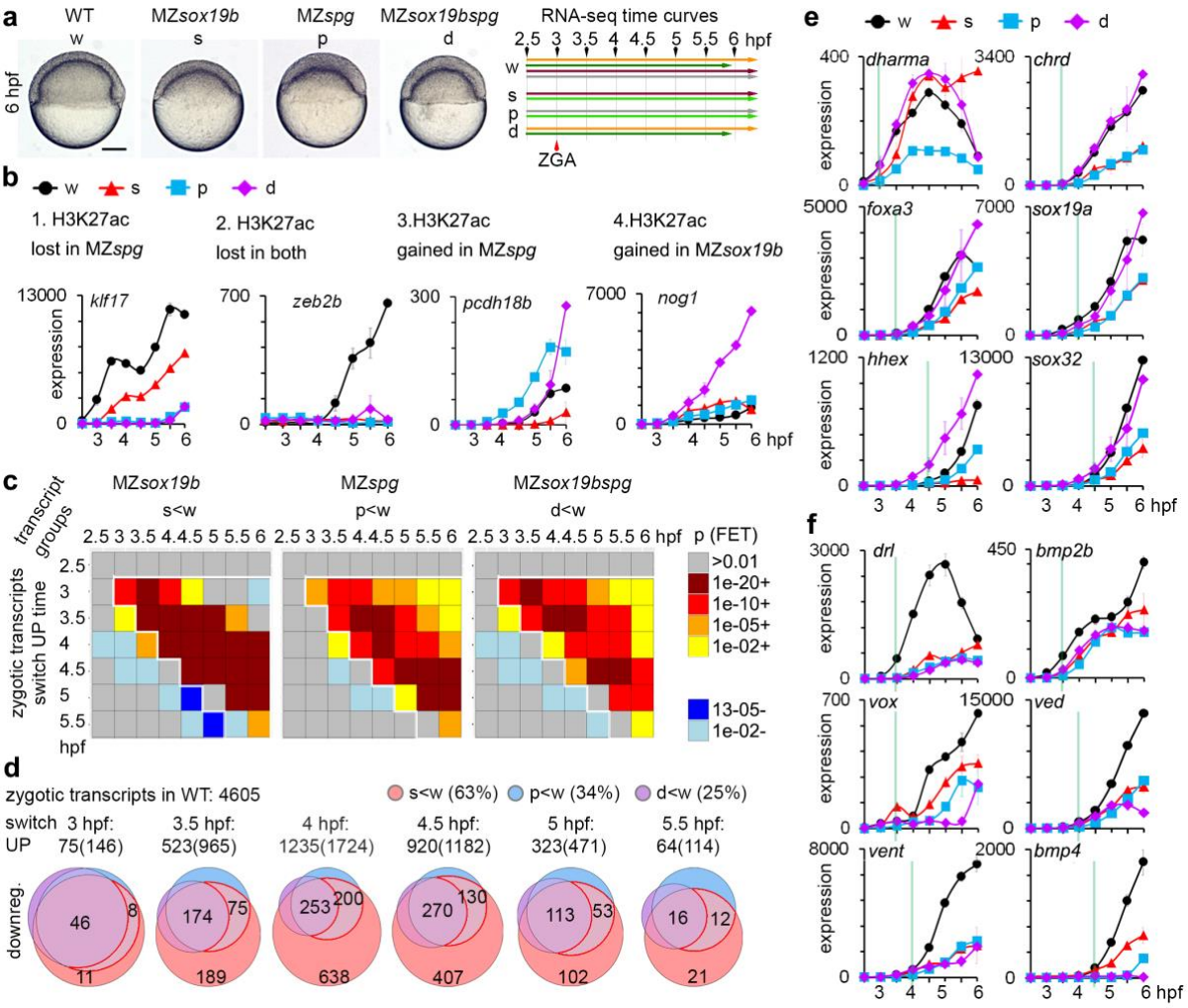

Figure 5: Direct activation of zygotic transcription by Pou5f3 and Sox19b is non-additive. a. left: live phenotypes of MZ Pou5f3 (p), Sox19b (s), double mutant (d), and the wild-type (w). Right - RNA-seq experiments; samples collected in one experiment have the same color. Scale bar 200 µm. **b.** Expression of the top genes regulated by enhancers types 1- 4 in the WT, single and double mutants (compare to Fig.4d-g, which shows H3K27ac in the regulatory regions of the same genes). Note, that type 4 gene *noggin1* is strongly upregulated in the double mutant. **c.** RNA-sense plots for zygotic genes downregulated in the indicated mutants. Y-axis - time profile groups by WT switch up time, X-axis – groups of downregulated transcripts per time point. FET- Fisher's Exact Test. **d.** The scope of ZGA delays in single and double mutants, per switch UP time group. Zygotic transcripts were scored as repressed, if they were downregulated in at least one point at or after the switch up (to the right from the white line in Fig.5c). Numbers (n) of repressed transcripts (n total transcripts) per group are shown above the diagrams. Group of transcripts repressed in both single mutants, but not in double mutant (s,p<w) is highlighted by red frame. **e.** Maternal ß-catenin and Eomesodermin targets are not repressed in the double mutant (s,p<w group) **f.** Zygotic ß-catenin targets, BMPs and ventral marker *draculin* are repressed in single and double mutants (s,p,d<w group). RNA-sense switch up time point is indicated by green vertical line in e,f.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

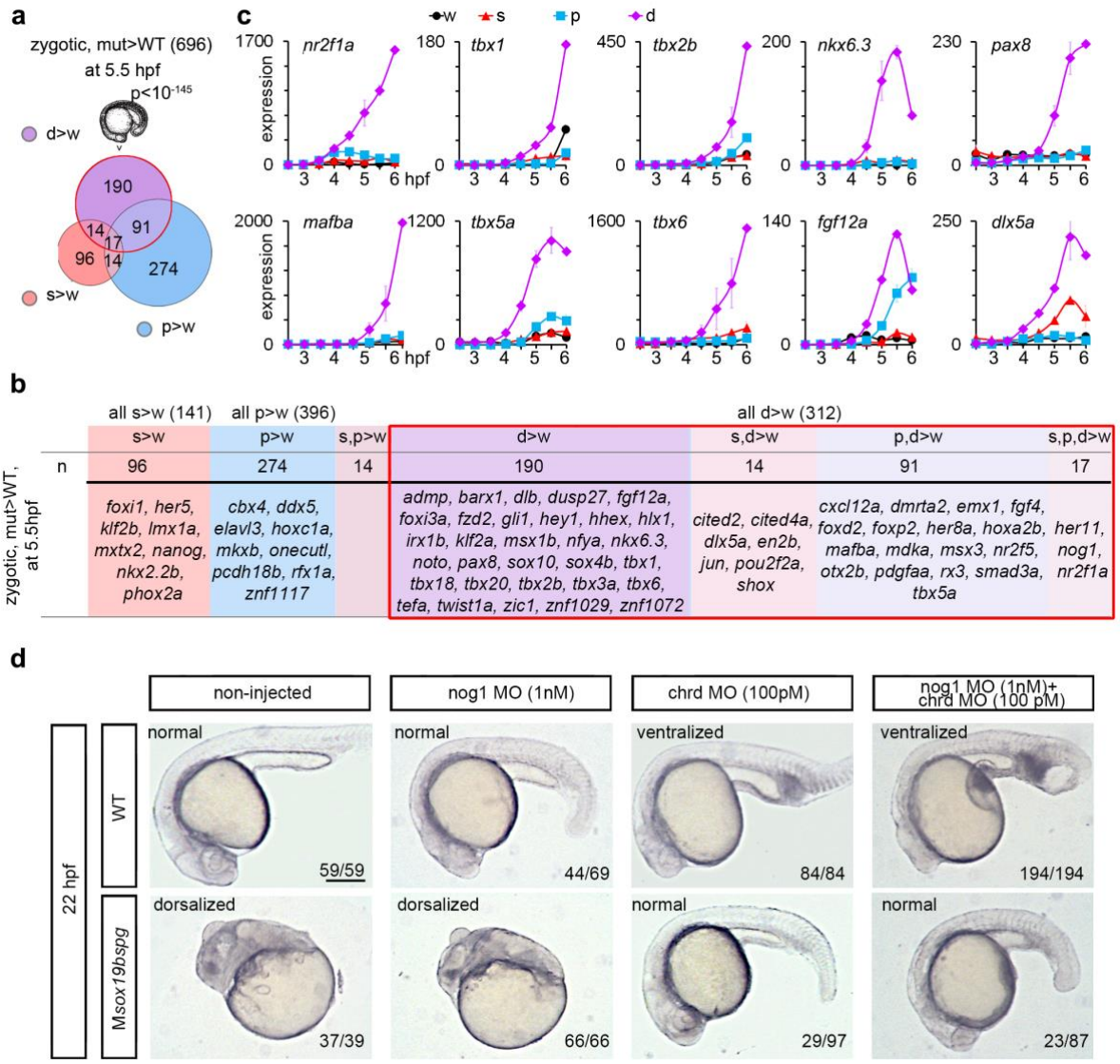

Figure 6: Pou5f3 and Sox19b synergistically repress differentiation regulators and dorsal genes. a, b. Groups of zygotic transcripts upregulated in one, two or three mutant backgrounds (s>w, p>w, d>w, red frame) at 5.5 hpf. p - binomial p-value for enrichment in transcripts expressed at 14-16 hpf in the wild-type (a) Example genes (b). **c.** Transcriptional regulators of organogenesis and dorsal genes are induced in the double mutant. **d.** Normal development can be rescued by reducing Chordin, but not Noggin levels in M*sox19bspg*+/- mutants. The wild-type or M*sox19bspg*+/- embryos were injected with the indicated morpholinos or non-injected. The numbers show the ratio of embryos with indicated phenotype/ all embryos alive at 22 hpf. The arrows show abnormally expanded blood progenitor cells in the ventralized wild-type embryos. Anterior to the left, dorsal up, scale bar 200µm.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

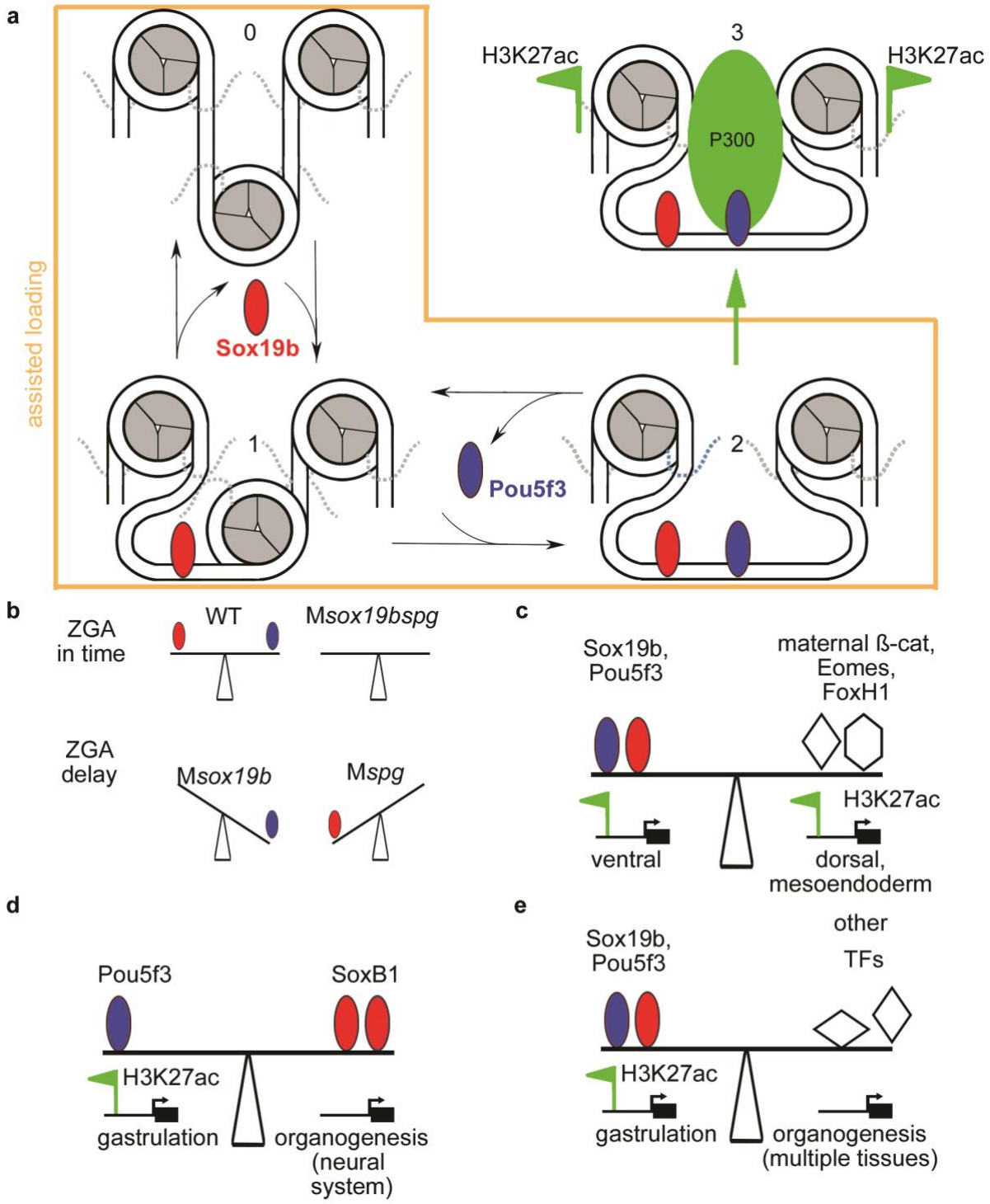

Figure 7: Summary of Pou5f3 and Sox19b activities. a. Sox19b-assisted Pou5f3 loading on co-dependent enhancers type 2. Sox19b competes with histones on *sox* motif (0-1), which allows Pou5f3 binding (2) to *pou:sox* motif nearby. (3): DNA-bound Pou5f3 promotes H3K27 acetylation of the neighboring nucleosomes**.** Green oval – P300. Nucleosome drawings are adapted from[68]. **b.** The balance of Pou5f3 and Sox19b is important for ZGA timing. **c.** The dorso-ventral balance at ZGA: Pou5f3 and Sox19b prime ventral genes for activation, dorsal and mesendodermal genes are primed independently on Pou5f3 and Sox19b (by maternal ß-cat, Eomesodermin and FoxH1). **d.** Pou5f3 activates gastrulation genes acting on enhancer types 1 and 2 and indirectly suppresses premature expression of organogenesis and neural

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

differentiation genes by SoxB1 on enhancers type 3**. e.** The absence of both Pou5f3 and Sox19b results in activation of enhancers type 4 and premature expression of differentiation genes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Experimental model and subject details

Wild-type fish of AB/TL and mutant sox19b$_{m1434}$ strains were raised, maintained and crossed under standard conditions as described by Westerfield[69]. The mutant spg$_{m793}$ line were maintained as described previously[20]. Embryos obtained from crosses were collected within 10 minutes and raised in egg water at 28.5°C. Staging was performed following the Kimmel staging series[70]. 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 mutants was approved by the Ethics Committee for Animal Research of the Koltzov Institute of Developmental Biology RAS, protocol 26 from 14.02.2019.

sox19b TALEN mutagenesis

To generate a *sox19b* zebrafish mutant, we used the TALEN technique and targeted the first exon for the mutation. The square brackets indicate the spacer sequence, flanking sequences should bind to TAL1 and TAL2 [38]: 5'-GATGGAGCACGAGCT[GAAGACCGCTGGTCCA]CCCCACACCCTCCAGC-3' (according to assembly Jul. 2010, zv9/danRer7). For restriction digest the enzyme BbsI was selected with the corresponding restriction site 5'-GAAGAC-3'. After injecting the TALENs (100 ng/µl each) into one-cell stage wild-type embryos we tested the proper activity of TALENs. We extracted genomic DNA from 20 of 24 hpf old WT and 20 injected embryos by lysing the cells with lysis buffer (10 mM Tris pH 8, 50 mM KCl, 0.3% Tween20, 0.3% NP-40, 1mM EDTA) and incubated the embryos at 98°C for 10 min. After cooling down Proteinase K solution (20 mg/ml, A3830, AppliChem) was added and incubated over night or at least 2 hours. Subsequently the Proteinase K was destroyed by heating up to 98°C for 10 min. We used the following primers for PCR: Sox19b-f1 5'-ATTTGGGGTGCTTTCTTCAGC-3' and Sox19b-r1 5'-GTTCTCCTGGGCCATCTTCC-3'. This gives a product of 362 bp length. This product contains two restriction sites for BbsI, so we got after digestion overnight three bands with sizes of 40 bp, 132 bp and 190 bp in which the 40 bp band also appears in wild-type embryos. After successful TALEN injection we let the fish grow and found through an outcross with wild-type two founders out of 19 tested fishes. From these two founders we chose the one with an 8 bp deletion resulting in a frameshift and a stop codon after 62 amino acids (5'-GATGGAGCACGAGCTGAAGA|CCACCCCACACCCTCCAGC-3', the line shows the position where the deletion occurred). For further experiments we used then the offspring with this mutation.

Morpholino knockdown

Two translation-blocking morpholinos for each of *sox2*, *sox3*, *sox19a* and *sox19b* were designed by Okuda [21]. All morpholinos were provided by Gene Tools. LLC (Philomath, USA). To generate SoxB1 knockdown embryos, 1.8 ng of morpholino mix for SoxB1 gene was microinjected into 1-cell stage embryos. For TKD 5.4 ng of standard control morpholino was taken, while 7.2 ng of standard morpholino was injected as QKD control. Noggin and Chordin Morpholinos were designed by Dal-Pra et al (2006)[59]. To block Noggin translation, 1 nM Noggin MO was injected. To moderately reduce Chordin translation, 100 pM Chordin Morpholinos were injected, as recommended by Dal-Pra et al. (2006).

Rescue of MZ*sox19b* TKD with *sox19b* mRNA

Sox19b open reading frame was amplified from zebrafish total cDNA (4.3 hpf), using primers designed according to *sox19b* mRNA sequence in UCSC. The PCR product was cloned in pCRII-TOPO vector with TOPO TA Cloning® Kit (Invitrogen) following the manual, and then sub-cloned into PCS2+ vector. mRNA was *in-vitro* transcribed with mMESSAGE mMACHINE® SP6 Kit (Ambion) according to the user's manual. Sox19b mRNAs were cleaned up with QIAGEN RNeasy® Mini Kit and 20 pg per embryo was co-injected into 1-cell stage embryo together with TKD Morpholinos.

Cell division counting in living embryos

SYTOX Green (0.2 mM, ThermoFisher SCIENTIFIC) mixed with 0.2 M KCl and 1 nl was injected in wild-type and *MZsox19b* mutant embryos at the 1-cell stage. Embryos which showed a proper SYTOX Green fluorescence were selected under the fluorescence stereomicroscope and placed in a furrow of an agarose plate. They were covered with a 24 x 60 mm cover slide which was fixed with 2% low

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

melting agarose and overlaid with 0.3x Danieau's buffer to prevent drying out of the embryos. Subsequently the development of the embryos was documented by using Leica MZ16 FA Fluorescence Stereomicroscope and Leica Microsystems LAS AF software which took every 3 min a bright field image and a fluorescent image (GFP filter, 395 nm). The temperature during documentation was 24°C. Resulting images were analyzed with ImageJ by counting the single images between mitoses. The obtained number of images was multiplied with 3 min to achieve the time span in minutes between the cell divisions.

Live imaging

Zebrafish embryos were imaged according to the wild-type control stages under Leica MZ APO stereomicroscope using AxioVision SE64 Rel. 4.9.1 software, with 50x magnification for single embryo and 12.5 x magnifications for the overview. Photos were arranged in Adobe Photoshop CS5.

Time-lapse imaging

Wild type and mutant embryos were collected at the same time and dechorionated manually in 0.3x Danieau's buffer. Then one embryo from each genotype was picked and placed on a 1.5% agarose chamber filled with 0.3x Danieau's buffer. Images were taken with 3 min interval for 24 hours by either Leica MZ APO stereomicroscope using AxioVision SE64 Rel. 4.9.1 software, or The IMAGINGSOURCE DFK21F04 with IC capture ver 2.3.394.1917. Room temperature was kept at 26-28°C during the documentation. Images were further processed in ImageJ 1.50i.

Whole mount *in situ* hybridization (WISH)

To visualize the expression pattern of some chosen genes we performed whole mount in situ hybridization as previously described[71]. The plasmids for anti-sense RNA probe synthesis were kind gifts of Wolfgang Driever, Matthias Hammerschmidt, Yusuke Kamachi and Liliana Solnica-Krezel. Embryos were fixed at proper stages with 4% paraformaldehyde at 4°C overnight, 3 times washed in PBST for 5 min, dechorionated manually in PBST and dehydrated with ascending series of methanol to 100%. Embryos in 100% methanol were stored at -20°C. After rehydrating with descending series of methanol to PBST, embryos were washed 3 times in PBST, pre-hybridized in 300µl Hyb-Mix for 3 hours at 65°C. 2 µl of prepared probes were added to 300µl Hyb-Mix for hybridization overnight at 65°C. Emrbyos were placed into 24-well plates and washed with in situ robot BioLane™ HTI by series washing steps: firstly embryos were washed three times for 20min in 300 µl 50% Formamide at 65°C, in 500 µl 2x SSCT for 15 times at 65°C, three times in 500 µl 0.2x SSCT for 20 min at 65°C, twice in 1 ml PBST for 10 min at room temperature. Then embryos were incubated in blocking solution (2% goat serum (heat inactivated) in PBST/BSA (2 mg/ml) for 2 hours at RT and incubated overnight with anti-DIG (1:5000 diluted in PBST) at 4°C. After washing 6 times for 20 min in 1ml PBST and once for 10min in 100Mm Tris-HCl, (pH=9.5), embryos were incubated with staining buffer for 15 min and the robot program is finished. We replaced staining buffer with 500 µl staining solution and stained for a proper time on a shaker. To stop the staining process embryos were washed with stop solution. Stained embryos were fixed in an increasing series of glycerol (25%, 50% and 75% in PBST), finally stored in 100% glycerol at 4°C. These embryos were imaged with Leica MZ APO stereomicroscope using AxioVision SE64 Rel. 4.9.1 software and the images were processed in Adobe Photoshop CS4.

MNase-seq, data processing and visualization.

MNase-seq and data processing was performed as previously described[47]. Briefly, 200-400 MZ*sox19b* embryos were dechorionated and fixed 10 minutes in 1% Formaldehyde at dome (4.3 hours post-fertilization) stage. The nuclei were isolated and digested with MNase. The yield of and degree of digestion was controlled using Agilent High Sensitivity DNA Kit on Agilent Bioanalyzer, according to manufacturer instructions. Chromatin was digested so that it contained 80% mono-nucleosomes. Libraries were prepared using the Illumina sequencing library preparation protocol and single-end sequenced on an Illumina HiSeq 2500 by Eurofins Company (Germany). All further data processing was done using European Galaxy server usegalaxy.eu[72]. Sequenced reads were mapped to the zebrafish danrer7/Zv9 assembly[73] using Bowtie2[74]. Resulting BAM files were converted to BED format, all mapped reads were extended to 147 bp in their 3' direction, truncated to the middle 61 bp and converted back to BAM format using BED Tools[75]. To create Bigwig files for the visualization of nucleosome density, BAM files for MZ*sox19b* (this work), MZ*spg* and WT[47] were normalized to rpkm (reads per million reads per one kilobase) and converted to bigwig format using BAM coverage program in deepTools[76]. To create Bigwig files for the visualization of nucleosome displacement by Sox19b and Pou5f3, the log2 ratio of nucleosome occupancy (in rpkm) between each mutant and the wild-type was obtained using BAM compare program in deepTools. The heatmaps or profiles of

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

selected genomic regions were plotted using plotheatmap or plotprofile programs in deepTools (main Fig. 3cd, Fig. 4b, Fig. S1, Fig. S2).

ChIP-seq for chromatin marks

The freshly laid eggs of MZ*sox19b*, MZ*spg* mutants and wild-type were collected in 10-15 min intervals. Unfertilized eggs were removed at 2-4 cell stage. Collected embryos were transferred to 0.5x Danieau's solution (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; dilute 60X before use)[69] followed by enzymatic dechorionation with pronase E (0.3 mg/ml). The reaction was stopped by adding 1 % BSA (final conc. 0.04 %) followed by two to three washing steps with 0.5x x Danieau's. The eggs were cultured in glass petri dishes to prevent the embryos from adhere to the dish and thus eventually rip. They were incubated at 28 °C until the 4.3 hpf stage was reached. In order to fix the chromatin state at developmental stage 4.3 hpf (dome) and avoid nucleosome shifts, the dechorionated embryos were homogenized in 10 ml 0.5 % Danieau's containing 1x protease inhibitor coctail (PIC) using a Dounce tissue grinder and immediately treated with 1 % (v/v) Methanol-free Formaldehyde (Pierce) for exactly 10 min at room temperature. The homogenizate was transferred into a 15 ml falcon tube and shaken on a rotating platform for the rest of the 10 min. The fixation was stopped with 0.125 M Glycine by shaking for 5 min on a rotating platform. Subsequently the homogenizate was centrifuged for 5 min, 4700 rpm at 4 °C, whereupon a white pellet formed. The supernatant was discarded, and the pellet was resolved in Wardle cell lysis buffer (10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 0.5 % NP-40, 1-4 ml/1000 embryos). The lysate was distributed upon 2 ml eppendorf tubes, followed by 5 min incubation on ice with subsequent 1 min centrifugation, 2700 g at 4 °C. The supernatant was discarded again and 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% Twen20, pH 7.5). In order to count the obtained nuclei, the pellet was resolved in 1 ml ice cold 1x PBST, of which 10 µl were diluted 1:1 with Sytox® green (1:400) and fluorescence microscope examined by using the Neubauer counting chamber. The amount of nuclei was calculated using the following equation:

Nuclei/(µl)=(number of counted nuclei)/(〖counted area (mm〗^2)* chamber depth (mm)*dilution)

The residual nuclei were again pelleted by 1 min centrifugation at 2700 g and 4 °C, subsequently snap frozen in liquid nitrogen and stored at -80 °C. 2.5 million nuclei were used to perform one ChIP-Seq experiment.The chromatin was thawed and resolved in 2 ml of Wardle 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 ensure that the sonication was successful, 30 µl of the sheared chromatin was de-crosslinked with 250 mM NaCl over night 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 at 14,000 rpm and 4 °C. 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. 3 µg of referring antibody were 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 Dynal 10003D) were transferred into a 2 ml eppendorf tube and placed on a magnetic rack in order to remove the liquid from the beads. Subsequently 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 in order to discard the supernatant. The beads were resuspended in 333 µl RIPA buffer containing PIC. The Protein G-antibody-chromatin complex was washed 4x5 min on a rotating platform with 1 ml of RIPA buffer (10 mM Tris-HCl (pH 7.6), 1 mM EDTA, 1 % NP-40, 0.7 % sodium deoxycholate, 0.5M LiCL), 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 and the supernatant was removed. In order 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. Afterwards the supernatant was transferred to a fresh eppendorf tube by pulling the beads down, using the magnetic rack. 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 as control in parallel (230 µl elution buffer per 30 µl input).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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 validate if ChIP experiments were successful, we performed quantitative PCR in ChIP and Input control material, using the primers for the positive and negative reference genomic regions, enriched in or devoid of chromatin marks (see KEY RESOURCES TABLE). According to previously published data[48], the chromatin region near *tiparp* gene was highly enriched in H3K27ac and H3K4me3 histone marks at 4.3 hpf, while genomic region near *igsf2* gene did not bear any of these mark (Bogdanovic et al., 2012). QPCR was carried out using the SsoAdvanced™ Universal SYBR® Green Supermix from BIO-RAD. ChIP and input were normalized 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. 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'.
In order to convert a small amount of DNA into indexed libraries for Next Generation Sequencing (NGS) on the Illumina platform we used the NEBNext® Ultra™ DNA Library Prep Kit according to manufacturer instructions, with the modifications indicated below. The library preparation follows a 4-step protocol including end-repair (5' phosphorylation, dA-tailing), adaptor ligation, PCR enrichment including barcoding and clean up. 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 subsequently eluted by adding 33 µl of 0.1x TE (pH 8) and incubation 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. In order to reduce the peak of residual unligated Adaptors, the reaction was re-purified, by adding H2O up to 50 µl and 45 µl of AMPure XP Beads. The concentration was determined using the Qubit™ Fluorometer and Quanti-iT™ PicroGreen® dsDNA Kit. The 7 ChIP-seq libraries were sequenced at 70 mln paired end 150bp reads each: WT K27ac Chip, MZ*spg* K27ac Chip, MZ*sox19b* K27ac Chip1, MZ*sox19b* K27ac Chip2, WT K4*me3* Chip, MZ*spg* K4me3 Chip, MZ*sox19b* K4me3 Chip. The 7 corresponding input libraries were sequenced to 30 mln reads in the Novogene company (China).

Chromatin ChIP-seq data analysis

H3K27ac and H3K4me3 ChIP-seq data processing was done using european Galaxy server usegalaxy.eu[72]. Sequenced reads were mapped to the zebrafish danrer7/Zv9 assembly[73] using Bowtie2[74]. Peak calling was performed using MACS algorithm [77]. To reduce the number of false positives, peak calling for each H3K27ac ChIP-seq experiment (WT K27ac Chip, MZ*spg* K27ac Chip, MZ*sox19b* K27ac Chip1, MZ*sox19b* K27ac Chip2) was performed five times: peaks were called using four different background files (with the application MACS2 callpeak), or without background (with the application MACS2 bgpeakcall). Only H3K27ac peaks present in all five peak calls were kept for further analysis. For MZ*sox19b* ChIP-seq, only the peaks present in both biological replicates were taken. H3K4me3 peaks were called once with the application MACS2 bgpeakcall. On the next step, H3K27ac or H3K4me3 peaks from different genotypes were piled together and the overlapping regions were merged. To obtain the common list of putative H3K27 acetylated enhancers in all genotypes, the H3K27ac peaks which overlapped with H3K4me3 peaks were removed. The resulting list in BED format was sorted according to descending peak strength. To create Bigwig files for the visualization of histone mark levels and for k-means-clustering, the log2 ratio between each ChIP and corresponding input (in rpkm) was obtained using BAM compare program in deepTools[76]. For MZ*sox19b*, one of the two biological replicates (rep.1) was used. Unsupervised clustering was performed with k-means algorithm within plotheatmap program in deepTools. The heatmaps or profiles of H3K27ac levels on selected genomic were plotted using plotheatmap or plotprofile programs in deepTools (Fig.4a,b, Fig.S3). Histone modification profiles in single genes (Fig. 4d-g, Fig. S4, Fig.S7) were visualized using UCSC browser.

GREAT analysis, "top enhancer type" genes

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GREAT analysis [49] of genomic regions matching to H3K27ac ChIP-seq peaks of clusters 1 to 5 was performed at http://great.stanford.edu/great/public-3.0.0/html/. Genomic regions were associated with genes using 20 kb single nearest gene association rule. The genes, associated with the highest number of peaks of clusters 1-4 (Table S2), were visually checked in UCSC browser and selected as " top enhancer type 1-4" genes.

Motif analysis.

Position-weight matrices for *sox* and *pou:sox* motifs from the Supplementary Table 2 in Veil et al., 2019[47], were used. Genomic coordinates of the individual motif occurrences in +/- 1.5 kb from the center of selected genomic regions were obtained using FIMO[78] with p-value threshold 10^{-4}. The genomic coordinates of motifs were saved as a BED file, converted to Bigwig format with BEDTools[75], and used for visualization of the motif density on the selected genomic regions (Fig.4b, Fig.S1, Fig.S2) using deepTools2[76]. To score the number of motifs, overlapping *sox* or *pou:sox* motifs were merged and trimmed to 20 bp. The numbers of non-redundant matches for *pou:sox* and *sox* per 320 bp genomic region was scored. The results are presented on Fig. 3g and Fig.4c.

RNA-seq, data processing and visualization.

The embryos were synchronized at the four cell stage and collected every 30 minutes starting from 2.5 hpf (pre-ZGA, 256 cell stage, 8th cell cycle) till midgastrula (6 hpf). 40-45 WT, MZ*spg*, MZ*sox19b* and MZ*sox19bspg* zebrafish embryos were collected, sorted and fixed in liquid nitrogen for each time point. 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 single-end 50bp sequencing (35M reads per sample) on Illumina platform was performed by NOVOGENE company (China) for each sample. In total, we sequenced 78 samples in four biological replicates (Fig.5a) and 23 technical replicates for WT and single mutants. Further data processing was performed on european Galaxy server usegalaxy.eu[72]. All sequenced clean reads data were trimmed on 5' end for 5 bp by Trim Galore! program according to the sequencing quality control. Trimmed reads were mapped to danRer 11 using RNA STAR[79] and ENSEMBL gene models. Number of reads for each gene was counted in Feature count[80]. Feature counts were cross-normalized using DEseq2[81]. Processed RNA-seq data table deposited in GEO: GSE137424 was used for plotting expression profiles for single genes in Excel and as an input for RNA-sense. In our time series data, single 6 hpf data point was missing in one of the two biological replicates for MZ*sox19bspg* and in parallel WT control sample (Fig.5a). These missing 6 hpf time points for WT and MZ*sox19bspg* were duplicated from the other replicate to get an input for RNA-sense program. RNA-sense analysis was run with the following user-defined parameters (see sub-chapter "RNA-sense program" below for explanations): pVal switch=0.15, pVal FC=0.01, FC=2, threshold=100. To identify zygotic genes downregulated in the mutants (Fig.c,d, Table S3), Step 1 of RNA-sense was performed for the wild-type condition. To identify zygotic genes upregulated in the mutants at 5.5 hpf (Fig.6a-c, Table S4), Step 1 of RNA-sense was performed for each of the mutants.

RNA-sense program

In order to facilitate the biological interpretation of time-resolved RNA-seq data, we developed a 3-step procedure called RNA-sense (see Fig. S6). In principle, the usage of RNA-sense is not only restricted to RNA time series. RNA-sense can be applied to compare two groups of data series to capture the differences in the dynamic changes between the groups. The series data could be temporal, spatial, or any other continuous condition like series concentration of drug treatment. The data itself could be any sequencing data, including DNA, RNA and protein, or any other comparable large datasets.

In step one, time-resolved RNA-seq data in one of two conditions, e.g. the wild-type (user-defined parameter ExperimentStepDetection="WT"), are analyzed with respect to their temporal profile. The transcripts expressed below a user-defined threshold are excluded from the analysis. First, for each gene and for each measurement time point t, dynamic data is split into two groups before and after (after and equal to) time point t. The data is fitted by both a one-step model (two different means before and after time point t) and by a constant model (mean over all data points) and models are compared pairwise by means of likelihood ratio tests for each time point t. If the one-step model is significantly better (with user-defined p-value cutoff, pVal switch) than the constant model, a switch is detected for this time point. The difference of the means before and after the time point defines the direction of the switch "up" or "down". If switches were detected at different time points for each gene, the first possible time point is chosen.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In step two, fold changes between wild-type and mutant data are analyzed. For each gene and for each time point, Robinson and Smyth exact negative binomial test (with user-defined p-value cutoff, pVal FC) is performed to determine whether genes are significantly up- or downregulated in the mutant with respect to wild-type. The function exact.nb.test from the R package NBPSeq is used for analysis (https://cran.rstudio.com/web/).
In step three, results of step one and two are combined. Genes are grouped in a matrix form with respect to switch time (y-coordinate) and mutant fold change (x-coordinate). Genes for which fold change was detected at several time points appear several times in the matrix (Fig.S6A, B). For each tile of the matrix, Fisher's exact test for non-randomness is performed to analyze the correlation between the two properties switch time and fold change detection. Tiles with a low p-value in the Fisher test show a high correlation between switch time and fold change. This can be interpreted as a high number of genes for which the switch point is shifted in time in the mutant condition.
RNA-sense is a flexible tool with several user-defined parameters:
-ExperimentStepDetection: tells which of two conditions (i.e. WT or mutant) should be used for switch detection in Step 1
-threshold – the transcript is included in the analysis, if the expression value in at least one data point reaches the threshold.
-pVal switch – p-value threshold for switch detection in Step 1
-FC – fold change value threshold (optional) for Step 2
-pVal FC – p-value threshold for fold change analysis at Step 2
The code and example files for automatically performing the 3-step procedure are available in the R-package RNAsense that was developed jointly with the paper and which is available on Bioconductor https://bioconductor.org/.

In-vitro nucleosome predictions and HNARs.

Nucleosome prediction program from Kaplan et al. 2009[66] was integrated into the Galaxy platform using the Galaxy tool SDK planemo (https://github.com/galaxyproject/planemo) and following the best practices for Galaxy tool development (http://galaxy-iuc-standards.readthedocs.io/en/latest/best_practices.html). The tool was uploaded into the european Galaxy ToolShed (ref. https://www.ncbi.nlm.nih.gov/pubmed/25001293) and is available at the Galaxy instance. The sequences around 24310 TF-binding and control genomic regions (Veil et al., 2019, Table S1)[47] were extended to 10 kb to account for the edge effects, and *in-vitro* nucleosome prediction value was derived for every bp. For visualization, nucleosome predictions were converted to BigWig files in DeepTools2 and used for plotting (Fig. 3b, Fig. S1, Fig. S2) or in the UCSC browser (Fig. 3f). The maximal and minimal nucleosome prediction values within 320 bp around TF-binding and control regions and their genomic positions were taken from in Table S1[47]. To orient the genomic regions aligned on [nucmax] along ascending nucleosome prediction values, we searched for the min. nucleosome prediction at +/- 160 bp around [nucmax]. If the min. prediction was downstream of [nucmax], we reversed the strand from + to -. The strand for oriented plots is listed in Table S1[82]. To rank 24310 TF-binding and control regions to quartiles according to *in-vitro* predicted nucleosome occupancy, the average nucleosome prediction value within the region 0-+150 bp was taken (Fig. S2).

Propeller Twist shape

Propeller twist values for *pou5f3* regulatory region (Fig.3f) and aligned genomic regions (Fig.S2) were calculated on TFBS shape server at http://rohslab.cmb.usc.edu/TFBSshape/[83]

Quantification and statistical analysis

The sequencing coverage between the samples was normalized as rkpm (reads per million reads per one kilobase). Normalized difference between the mutant and wild-type (ΔmutMNase) was calculated as Δmut=((rpkm(mut)-rpkm(wt))/(rkpm(mut)+rpkm(wt)). Average ΔmutMNase values per 320 bp around HNAR were taken. For H3K27ac ChIP-seq, normalized difference between the mutant and wild-type (Δmut K27ac) was calculated as ((log2(ChIP/input)mut-log2(ChIP/input)wt)/(log2(ChIP/input)mut+log2(ChIP/input)wt)). Average Δmut K27ac values per 1kb around HNAR were taken. Data were analyzed using JMP (SAS Institute 2012 version 10) using one-way ANOVA and linear regression (Table S1).

Data and Software Availability

Software

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RNA-sense software developed during this study is available at https://bioconductor.org/packages/release/bioc/html/RNAsense.html

Data Resources (published and generated during this work)

Data set	reference	source
ATAC-seq of the WT 4.3 hpf zebrafish embryos	45	GEO: GSE101779
ChIP-seq for H3K27ac and H3K4me3 in three genotypes, 4.3 hpf	this work	GEO: GSE143306
ChIP-seq for Pou5f3 and SoxB1	12	GEO: GSE39780
ChIP-seq for Nanog	50	GEO: GSE34683
ChIP-seq for Eomesodermin	53	GEO: GSE51894
ChIP-seq for FoxH1	55	GEO: GSE67648
Consensus *pou:sox* and *sox* motifs	47	Supplementary Table 2 in Veil et al., 2019
List of Pou5f3, SoxB1 and Nanog -bound and control genomic regions centered on HNARs (on max. nucleosome prediction within 320 bp around ChIP-seq peak) and oriented with min. nucleosome prediction to the left	47	Supplementary Table 1 in Veil et al., 2019
MNase-seq of MZ*sox19b* embryos, 4.3 hpf	this work	GEO: GSE125945
MNase-seq of WT and MZ*spg* embryos, 4.3 hpf	47	GEO: GSE109410
Time-resolved RNA-seq at 8 time points in 4 genotypes	this work	GEO: GSE137424

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1 Schulz, K. N. & Harrison, M. M. Mechanisms regulating zygotic genome activation. *Nature Reviews Genetics*, doi:10.1038/s41576-018-0087-x (2018).

2 Vastenhouw, N. L., Cao, W. X. & Lipshitz, H. D. The maternal-to-zygotic transition revisited. *Development* **146**, doi:10.1242/dev.161471 (2019).

3 Amodeo, A. A., Jukam, D., Straight, A. F. & Skotheim, J. M. Histone titration against the genome sets the DNA-to-cytoplasm threshold for the Xenopus midblastula transition. *Proc Natl Acad Sci U S A* **112**, E1086-1095, doi:10.1073/pnas.1413990112 (2015).

4 Joseph, S. R. *et al.* Competition between histone and transcription factor binding regulates the onset of transcription in zebrafish embryos. *Elife* **6**, doi:10.7554/eLife.23326 (2017).

5 Wilky, H., Chari, S., Govindan, J. & Amodeo, A. A. Histone concentration regulates the cell cycle and transcription in early development. *Development*, doi:10.1242/dev.177402 (2019).

6 Ferg, M. *et al.* The TATA-binding protein regulates maternal mRNA degradation and differential zygotic transcription in zebrafish. *EMBO J* **26**, 3945-3956, doi:10.1038/sj.emboj.7601821 (2007).

7 Veenstra, G. J., Destree, O. H. & Wolffe, A. P. Translation of maternal TATA-binding protein mRNA potentiates basal but not activated transcription in Xenopus embryos at the midblastula transition. *Mol Cell Biol* **19**, 7972-7982, doi:10.1128/mcb.19.12.7972 (1999).

8 Chan, S. H. *et al.* Brd4 and P300 Confer Transcriptional Competency during Zygotic Genome Activation. *Dev Cell* **49**, 867-881 e868, doi:10.1016/j.devcel.2019.05.037 (2019).

9 Sato, Y. *et al.* Histone H3K27 acetylation precedes active transcription during zebrafish zygotic genome activation as revealed by live-cell analysis. *Development* **146**, doi:10.1242/dev.179127 (2019).

10 Liang, H. L. *et al.* The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. *Nature* **456**, 400-403, doi:10.1038/nature07388 (2008).

11 Lee, M. T. *et al.* Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. *Nature*, 360-364., doi:10.1038/nature12632 (2013).

12 Leichsenring, M., Maes, J., Mossner, R., Driever, W. & Onichtchouk, D. Pou5f1 transcription factor controls zygotic gene activation in vertebrates. *Science* **341**, 1005-1009, doi:10.1126/science.1242527 (2013).

13 Gentsch, G. E., Spruce, T., Owens, N. D. L. & Smith, J. C. The role of maternal pioneer factors in predefining first zygotic responses to inductive signals. *bioRxiv*, 306803, doi:10.1101/306803 (2019).

14 Gao, L. *et al.* Chromatin Accessibility Landscape in Human Early Embryos and Its Association with Evolution. *Cell* **173**, 248-259 e215, doi:10.1016/j.cell.2018.02.028 (2018).

15 De Iaco, A. *et al.* DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat Genet* **49**, 941-945, doi:10.1038/ng.3858 (2017).

16 Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet* **20**, 207-220, doi:10.1038/s41576-018-0089-8 (2019).

17 McDaniel, S. L. *et al.* Continued Activity of the Pioneer Factor Zelda Is Required to Drive Zygotic Genome Activation. *Mol Cell* **74**, 185-195 e184, doi:10.1016/j.molcel.2019.01.014 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

18 Schulz, K. N. *et al.* Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early Drosophila embryo. *Genome Res* **25**, 1715-1726, doi:10.1101/gr.192682.115 (2015).

19 Sun, Y. *et al.* Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic genome activation. *Genome Res*, 1703–1714, doi:10.1101/gr.192542.115 (2015).

20 Lunde, K., Belting, H. G. & Driever, W. Zebrafish pou5f1/pou2, homolog of mammalian Oct4, functions in the endoderm specification cascade. *Current biology : CB* **14**, 48-55 (2004).

21 Okuda, Y., Ogura, E., Kondoh, H. & Kamachi, Y. B1 SOX coordinate cell specification with patterning and morphogenesis in the early zebrafish embryo. *PLoS Genetics* **6**, e1000936, doi:10.1371/journal.pgen.1000936 (2010).

22 Huangfu, D. *et al.* Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol* **26**, 1269-1275, doi:10.1038/nbt.1502 (2008).

23 Montserrat, N. *et al.* Generation of induced pluripotent stem cells from human renal proximal tubular cells with only two transcription factors, OCT4 and SOX2. *J Biol Chem* **287**, 24131-24138, doi:10.1074/jbc.M112.350413 (2012).

24 Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell.* **122**, 947-956. (2005).

25 Loh, Y. H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet.* **38**, 431-440 (2006).

26 Chen, J. *et al.* Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**, 1274-1285, doi:10.1016/j.cell.2014.01.062 (2014).

27 Biddle, J. W., Nguyen, M. & Gunawardena, J. Negative reciprocity, not ordered assembly, underlies the interaction of Sox2 and Oct4 on DNA. *Elife* **8**, doi:10.7554/eLife.41017 (2019).

28 Li, S., Zheng, E. B., Zhao, L. & Liu, S. Nonreciprocal and Conditional Cooperativity Directs the Pioneer Activity of Pluripotency Transcription Factors. *Cell Rep* **28**, 2689-2703 e2684, doi:10.1016/j.celrep.2019.07.103 (2019).

29 Friman, E. T. *et al.* Dynamic regulation of chromatin accessibility by pluripotency transcription factors across the cell cycle. *Elife* **8**, doi:10.7554/eLife.50087 (2019).

30 Frum, T. *et al.* Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Developmental cell* **25**, 610-622, doi:10.1016/j.devcel.2013.05.004 (2013).

31 Wicklow, E. *et al.* HIPPO pathway members restrict SOX2 to the inner cell mass where it promotes ICM fates in the mouse blastocyst. *PLoS Genet* **10**, e1004618, doi:10.1371/journal.pgen.1004618 (2014).

32 Gerhart, J. in *Biological Regulation and Development* Vol. 2 (ed RF Goldberg) 133–315 (Plenum Press, 1980).

33 Newport, J. & Kirschner, M. A major developmental transition in early Xenopus embryos: II. Control of the onset of transcription. *Cell* **30**, 687-696 (1982).

34 Signoret, J. & Lefresne, J. Contribution a L'etude de la segmentation de l'oeuf d'axolotl: I. Definition de la transition blastuleenne. *Ann. Embryol. Morphog.* **4**, 113-123 (1971).

35 Kafiani, C. A., Timofeeva, M. J., Neyfakh, A. A., Melnikova, N. L. & Rachkus, J. A. RNA synthesis in te early embryogenesis of a fish (Misgurnus fossilis). *J Embryol Exp Morphol* **21**, 295-308 (1969).

36 Neyfakh, A. A. Radiation Investigation of Nucleo-Cytoplasmic Interrelations in Morphogenesis and Biochemical Differentiation. *Nature* **201**, 880-884 (1964).

37 Shih, Y. H. *et al.* SoxB1 transcription factors restrict organizer gene expression by repressing multiple events downstream of Wnt signalling. *Development* **137**, 2671-2681, doi:10.1242/dev.054130 (2010).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

38 Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* **39**, e82, doi:10.1093/nar/gkr218 (2011).

39 Kane, D. A. & Kimmel, C. B. The zebrafish midblastula transition. *Development* **119**, 447-456 (1993).

40 Keller, P. J., Schmidt, A. D., Wittbrodt, J. & Stelzer, E. H. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science* **322**, 1065-1069, doi:10.1126/science.1162493 (2008).

41 Eckerle, S., Ringler, M., Lecaudey, V., Nitschke, R. & Driever, W. Progesterone modulates microtubule dynamics and epiboly progression during zebrafish gastrulation. *Dev Biol* **434**, 249-266, doi:10.1016/j.ydbio.2017.12.016 (2018).

42 Song, S. *et al.* Pou5f1-dependent EGF expression controls E-cadherin endocytosis, cell adhesion, and zebrafish epiboly movements. *Developmental cell* **24**, 486-501, doi:10.1016/j.devcel.2013.01.016 (2013).

43 Belting, H. G. *et al.* Pou5f1 contributes to dorsoventral patterning by positive regulation of vox and modulation of fgf8a expression. *Dev biol* **356(2):323-36**, 323-336, doi:10.1016/j.ydbio.2011.05.660 (2011).

44 Reim, G. & Brand, M. Maternal control of vertebrate dorsoventral axis formation and epiboly by the POU domain protein Spg/Pou2/Oct4. *Development* **133**, 2757-2770, doi:dev.02391 [pii]10.1242/dev.02391 (2006).

45 Liu, G., Wang, W., Hu, S., Wang, X. & Zhang, Y. Inherited DNA methylation primes the establishment of accessible chromatin during genome activation. *Genome Res* **28**, 998-1007, doi:10.1101/gr.228833.117 (2018).

46 Palfy, M., Schulze, G., Valen, E. & Vastenhouw, N. L. Chromatin accessibility established by Pou5f3, Sox19b and Nanog primes genes for activity during zebrafish genome activation. *PLoS Genet* **16**, e1008546, doi:10.1371/journal.pgen.1008546 (2020).

47 Veil, M., Yampolsky, L., Gruening, B. & Onichtchouk, D. Pou5f3, SoxB1, and Nanog remodel chromatin on High Nucleosome Affinity Regions at Zygotic Genome Activation. *Genome Res*, doi:10.1101/gr.240572.118 (2019).

48 Bogdanovic, O. *et al.* Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. *Genome Res* **22**, 2043-2053, doi:10.1101/gr.134833.111 (2012).

49 Hiller, M. *et al.* Computational methods to detect conserved non-genic elements in phylogenetically isolated genomes: application to zebrafish. *Nucleic Acids Res* **41**, e151, doi:10.1093/nar/gkt557 (2013).

50 Xu, C. *et al.* Nanog-like Regulates Endoderm Formation through the Mxtx2-Nodal Pathway. *Developmental cell* **22**, 625-638, doi:10.1016/j.devcel.2012.01.003 (2012).

51 Bruce, A. E., Howley, C., Dixon Fox, M. & Ho, R. K. T-box gene eomesodermin and the homeobox-containing Mix/Bix gene mtx2 regulate epiboly movements in the zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists* **233**, 105-114, doi:10.1002/dvdy.20305 (2005).

52 Du, S., Draper, B. W., Mione, M., Moens, C. B. & Bruce, A. Differential regulation of epiboly initiation and progression by zebrafish Eomesodermin A. *Developmental biology* **362**, 11-23, doi:10.1016/j.ydbio.2011.10.036 (2012).

53 Nelson, A. C. *et al.* Global identification of Smad2 and Eomesodermin targets in zebrafish identifies a conserved transcriptional network in mesendoderm and a novel role for Eomesodermin in repression of ectodermal gene expression. *BMC Biol* **12**, 81, doi:10.1186/s12915-014-0081-5 (2014).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.16.949362; this version posted May 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

54 Slagle, C. E., Aoki, T. & Burdine, R. D. Nodal-dependent mesendoderm specification requires the combinatorial activities of FoxH1 and Eomesodermin. *PLoS Genet* **7**, e1002072, doi:10.1371/journal.pgen.1002072 (2011).

55 Dubrulle, J. *et al.* Response to Nodal morphogen gradient is determined by the kinetics of target gene induction. *eLife* **4**, doi:10.7554/eLife.05042 (2015).

56 Harvey, S. A. *et al.* Identification of the zebrafish maternal and paternal transcriptomes. *Development* **140**, 2703-2710, doi:10.1242/dev.095091 (2013).

57 Haberle, V. *et al.* Two independent transcription initiation codes overlap on vertebrate core promoters. *Nature* **507**, 381-385, doi:10.1038/nature12974 (2014).

58 De Robertis, E. M. Spemann's organizer and the self-regulation of embryonic fields. *Mech Develop* **126**, 925-941, doi:10.1016/j.mod.2009.08.004 (2009).

59 Dal-Pra, S., Furthauer, M., Van-Celst, J., Thisse, B. & Thisse, C. Noggin1 and Follistatin-like2 function redundantly to Chordin to antagonize BMP activity. *Developmental biology* **298**, 514-526, doi:10.1016/j.ydbio.2006.07.002 (2006).

60 Furthauer, M., Thisse, B. & Thisse, C. Three different noggin genes antagonize the activity of bone morphogenetic proteins in the zebrafish embryo. *Developmental biology* **214**, 181-196, doi:10.1006/dbio.1999.9401 (1999).

61 Tosic, J. *et al.* Eomes and Brachyury control pluripotency exit and germ-layer segregation by changing the chromatin state. *Nat Cell Biol* **21**, 1518-1531, doi:10.1038/s41556-019-0423-1 (2019).

62 Argelaguet, R. *et al.* Single cell multi-omics profiling reveals a hierarchical epigenetic landscape during mammalian germ layer specification. *bioRxiv*, 519207, doi:10.1101/519207 (2019).

63 Charney, R. M. *et al.* Foxh1 Occupies cis-Regulatory Modules Prior to Dynamic Transcription Factor Interactions Controlling the Mesendoderm Gene Program. *Dev Cell* **40**, 595-607 e594, doi:10.1016/j.devcel.2017.02.017 (2017).

64 Paraiso, K. D. *et al.* Endodermal Maternal Transcription Factors Establish Super-Enhancers during Zygotic Genome Activation. *Cell Rep* **27**, 2962-2977 e2965, doi:10.1016/j.celrep.2019.05.013 (2019).

65 Chronis, C. *et al.* Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell* **168**, 442-459 e420, doi:10.1016/j.cell.2016.12.016 (2017).

66 Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362-366, doi:10.1038/nature07667 (2009).

67 Hou, L., Srivastava, Y. & Jauch, R. Molecular basis for the genome engagement by Sox proteins. *Semin Cell Dev Biol* **63**, 2-12, doi:10.1016/j.semcdb.2016.08.005 (2017).

68 Bowman, G. D. & Poirier, M. G. Post-translational modifications of histones that influence nucleosome dynamics. *Chem Rev* **115**, 2274-2295, doi:10.1021/cr500350x (2015).