Maintenance of spatial gene expression by Polycomb-mediated repression after formation of a vertebrate body plan

Authors: Julien Rougeot[1,2,¥,*], Naomi D. Chrispijn[1], Marco Aben[1,2,£], Dei M. Elurbe[1,2,ß], Karolina M. Andralojc[1,€], Patrick J. Murphy[3,4], Pascal W.T.C. Jansen[5], Michiel Vermeulen[5], Bradley R. Cairns[3], Leonie M. Kamminga[1,2,$,*]

Affiliations: [1]Radboud University, Faculty of Science, Department of Molecular Biology, Radboud Institute for Molecular Life Sciences, Nijmegen, the Netherlands, [2]Radboud University Medical Center, Department of Molecular Biology, Nijmegen, the Netherlands, [3]Howard Hughes Medical Institute, Department of Oncological Sciences and Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, USA, [4]Wilmot Cancer Institute, Rochester Center for Biomedical Informatics, University of Rochester Medical Center, Rochester, NY, USA [5]Radboud University, Faculty of Science, Department of Molecular Biology, Radboud Institute for Molecular Life Sciences, Oncode Institute, Nijmegen, The Netherlands.

[¥]Present address: Laboratory of Entomology, Wageningen University, Wageningen, The Netherlands, [£]Present address: Radboud University Medical Center, Department of Human Genetics, Nijmegen, The Netherlands, [ß]Present address: Radboud University Medical Center, Centre for Molecular and Biomolecular Informatics, Radboud Institute for Molecular Life Sciences, Nijmegen, The Netherlands, [€]Present address: Radboud Institute for Molecular Life Sciences, Department of Biochemistry, Nijmegen, The Netherlands, [$]Present address: Wageningen University & Research – Library, Wageningen, the Netherlands.

*Corresponding authors

Corresponding author's email address: leonie.kamminga@wur.nl; julien.rougeot@wur.nl

Key words: Polycomb, Ezh2, zebrafish, ChIP-seq, transcriptomics, proteomics

Summary statement

Our unique zebrafish model of a maternal and zygotic mutant for the Polycomb group gene *ezh2* reveals major conserved and divergent mechanisms in epigenetic gene repression during vertebrate development.

Abstract

Polycomb group (PcG) proteins are transcriptional repressors that are important regulators of cell fate during embryonic development. Among them, Ezh2 is responsible for catalyzing the epigenetic repressive mark H3K27me3 and is essential for animal development. The ability of zebrafish embryos lacking both maternal and zygotic *ezh2* to form a normal body plan provides a unique model to comprehensively study Ezh2 function during early development in vertebrates. By using a multi-omics approach, we found that Ezh2 is required for the deposition of H3K27me3 and is essential for proper recruitment of Polycomb group protein Rnf2. However, despite the complete absence of PcG-associated epigenetic mark and proteins, only minor changes in H3K4me3 deposition and gene and protein expression occured. These changes were mainly due to local dysregulation of transcription factors outside their normal expression boundaries. Altogether, our results in zebrafish show that Polycomb-mediated gene repression is important right after the body plan is formed to maintain spatially restricted expression profiles of transcription factors and highlight the differences that exist in the timing of PcG protein action between vertebrate species.

Introduction

Development of multi-cellular organisms involves highly dynamic and controlled processes during which one single totipotent cell will multiply and differentiate into all the cells composing the adult individual. Specification of cell identity is controlled through the establishment of spatially and temporally restricted transcriptional profiles, which are subsequently maintained by epigenetic mechanisms (Brock and Fisher, 2005). Epigenetic maintenance of gene expression can act through modifications of the chromatin, the complex of DNA wrapped around an octamer of histones H2A, H2B, H3, and H4 and its associated proteins and non-coding RNAs, creating an epigenetic landscape, often referred to as the epigenome (Zhu and Li, 2016). These modifications can be propagated from mother to daughter cells and thereby maintain gene expression profiles by controlling the accessibility of the DNA to the transcriptional machinery (Li and Reinberg, 2011).

Polycomb Group (PcG) proteins are important regulators of the epigenome during development. First identified in *Drosophila melanogaster*, PcG proteins were found to maintain the pre-established pattern of *hox* gene expression (Kennison, 1995). Subsequent studies showed that PcG proteins are important for proper patterning during early embryonic development, tissue-specific development, and maintenance of the balance between pluripotency and differentiation of stem cells in multiple species (Schuettengruber et al., 2017). Two main PcG complexes have been described (Chittock et al., 2017). The Polycomb Repressive Complex 2 (PRC2) is composed of the core subunits EZH1/2 (Enhancer

of Zeste Homologue 1/2), SUZ12 (Suppressor of Zeste 12), and EED (Embryonic Ectoderm Development). EZH2 has a catalytically active SET domain that trimethylates lysine 27 of histone H3 (H3K27me3), an epigenetic mark associated with gene repression and found mainly at the transcriptional start sites of gene coding sequences (Mikkelsen et al., 2007). The catalytic subunits of PRC2 are mutually exclusive and EZH1 is postulated to complement the function of EZH2 in non-proliferative adult organs (Margueron et al., 2008; Shen et al., 2008). H3K27me3 can be recognized by the Polycomb Repressive Complex 1 (PRC1). A diversity of PRC1 compositions has been described and canonical PRC1 is composed of the core subunits RING1/RNF2 (Ring Finger Protein 2 a/b), PCGF1-6 (Polycomb Group RING fingers 1-6), PHC (Polyhomeotic), and CBX (Chromobox homolog) (Gao et al., 2012; Kloet et al., 2016). PRC1 catalyzes the ubiquitination of lysine 119 of histone H2A (H2AK1119ub) and promotes chromatin compaction and subsequent gene repression. In contrast to this canonical view, recent studies implicate that PRC1 is also active in the absence of PRC2 (He et al., 2013; Loubiere et al., 2016; Tavares et al., 2012). Trithorax Group (TrxG) proteins antagonize PcG protein function through the deposition of a trimethyl group on lysine 4 of histone H3 (H3K4me3) on promoters and enhancers from virtually all transcribed genes (Klymenko and Muller, 2004; Santos-Rosa et al., 2002; Schmitges et al., 2011).

In mice, loss of PRC2 genes *Ezh2*, *Eed*, or *Suz12* or PRC1 gene *Rnf2* leads to post-implantation embryonic lethality during early gastrulation (Faust et al., 1998; O'Carroll et al., 2001; Pasini et al., 2004; Voncken et al., 2003), making it difficult to study transcriptional regulation by PcG complexes during early development. Apart from the mouse model, very few studies have focused on characterization of PcG function during vertebrate development. Lately, the zebrafish embryo has emerged as a model of choice to study developmental epigenetics in vertebrates (Chrispijn et al., 2019; Lindeman et al., 2011; Murphy et al., 2018; Potok et al., 2013; Vastenhouw et al., 2010). We and others previously used loss-of-function mutants to show that *ezh2* is essential for zebrafish development (Dupret et al., 2017; San et al., 2018; San et al., 2016; Zhong et al., 2018). More particularly, our unique vertebrate model of zebrafish embryos mutant for both maternal and zygotic *ezh2*, referred to as *MZezh2* mutant embryos, develop seemingly normally until 1 dpf, forming a proper body plan. These mutants ultimately die at 2 dpf, exhibiting a 100% penetrant pleiotropic phenotype associated with a loss of tissue maintenance (San et al., 2016). This makes zebrafish *MZezh2* mutant embryos a valued model to study the function of Ezh2 during early development, from fertilization to tissue specification. Furthermore, these mutant embryos provide a unique context of a vertebrate in which trimethylation of H3K27 has never occurred, unlike cell culture, conditional, or zygotic mutant models.

We conducted a multi-omics approach in these *MZezh2* mutant embryos to study how PcG-mediated gene regulation controls axis formation and tissue specification. We focused our study on 24 hours post fertilization (hpf) embryos, when the first phenotypes become visible, and the anterior-posterior patterning of the embryos is properly established. Our results show conservation of basic PcG recruitment and silencing mechanisms and reveal that PRC2 function is essential for proper Rnf2 recruitment. However, very surprisingly, the transcriptional and proteomic profile of *MZezh2* mutant embryos remains largely unchanged compared to wildtype embryos despite the complete absence of Ezh2 protein and its associated epigenetic mark on the chromatin. The changes affect primarily a subset of PcG target genes. These genes are mainly transcription factors essential for developmental processes which present locally restricted aberrant gene expression. Our results show that zebrafish embryo development is initially independent of PcG repression until the stage of tissue maintenance and stress the differences that exist in the timing of PcG function requirement between vertebrate species.

Results

The repressive epigenetic mark H3K27me3 is absent in *MZezh2* embryos

To study the function of Ezh2 during development, we used the *ezh2* nonsense mutant allele *ezh2 (hu5670)* containing a premature stop codon within the catalytic SET domain, resulting in the absence of Ezh2 protein (San et al., 2016). Total elimination of both maternal and zygotic contribution of Ezh2 protein and mRNA, by using the germ cell transplantation technique described previously (Ciruna et al., 2002; San et al., 2016), allowed us to study the function of Ezh2 during early development. As previously shown, *MZezh2* mutant embryos display normal body plan formation and a mild phenotype at 24 hpf. They die at 48 hpf, at which point pleiotropic phenotypes are observed, such as smaller eyes, smaller brain, blood coagulation, and absence of pectoral fins (Fig. 1A). Western Blot analysis at 3.3 hpf and 24 hpf confirmed the absence of both maternal and zygotic Ezh2 in these mutants, respectively (Fig. 1B and Fig. S1). In addition, our previous study also reported that H3K27me3 was not detectable in *MZezh2* mutants by immunofluorescence (San et al., 2016).

To further confirm the absence of Ezh2 in *MZezh2* mutants and its effect on H3K27me3 deposition, we performed ChIP-sequencing (ChIP-seq) for Ezh2 and H3K27me3 at 24 hpf in both wildtype and *MZezh2* mutant embryos. ChIP-seq analyses for Ezh2 and H3K27me3 revealed 816 and 3,353 peaks in wildtype embryos, respectively (Fig. 1C, Table S1). Although the number of peaks differed between the two proteins, their binding profiles greatly overlap (Fig. 1E). Quantification showed that 85% of Ezh2 peaks also contain H3K27me3 (Fig. S2A). Known PcG target genes such as the *hoxab* gene cluster,

tbx genes, *isl1*, and *gsc* loci presented similar binding profiles for Ezh2 as for H3K27me3 (Fig. 1F,G, Fig. S2B), whereas the ubiquitously expressed genes *eif1ad* and *tbp* showed absence of both Ezh2 and H3K27me3 (Fig. S2B).

In *MZezh2* mutant embryos, the binding of Ezh2 and H3K27me3, as detected by ChIP-seq, was virtually absent, with 3 and 22 peaks detected for Ezh2 and H3K27me3, respectively (Fig. 1C). Manual inspection of these remaining peaks revealed that they are present in gene deserts and low complexity regions and are most probably artefacts (Fig. S2B). Ezh2 and H3K27me3 coverage was reduced to background levels in *MZezh2* mutants compared to wildtype (Fig. 1D). Finally, the *hoxab* gene cluster, *tbx3a, tbx5a*, *gsc,* and *isl1* loci, targeted by PcG repression in wildtypes, also showed a complete absence of Ezh2 and H3K27me3 binding in *MZezh2* mutants (Fig. 1F,G, Fig. S2B).

In order to verify that the absence of detection of Ezh2 and H3K27me3 in *MZezh2* mutant samples was not due to an inefficient ChIP-seq or a normalization artifact specific to mutant samples, the second ChIP-seq replicates for both Ezh2 and H3K27me3 were conducted with spike-in chromatin control. After normalization using the immunoprecipitated spike-in chromatin, the decreases in Ezh2 and H3K27me3 coverage in mutants compared to wildtypes appear even more pronounced than without spike-in normalization, both at the genome-wide level (Fig. S3A,B) as well as on target genes (Fig. S3C).

Altogether, these results demonstrate that in *MZezh2* mutants Ezh2 and H3K27me3 are absent from the chromatin.

Loss of PRC2-mediated repression results in impaired PRC1 recruitment during early development

It is postulated that PRC1 is recruited to chromatin by PRC2-deposited H3K27me3 but can also have a function independent of PRC2 (He et al., 2013; Loubiere et al., 2016; Tavares et al., 2012). As both Ezh2 and H3K27me3 are absent from *MZezh2* mutant embryos, we investigated whether PRC1 is still recruited to chromatin in these mutants. In zebrafish, Rnf2 is the only catalytic subunit of PRC1 (Le Faou et al., 2011). ChIP-seq for Rnf2 in wildtype embryos at 24 hpf reveals 837 peaks (Fig. 1C, Table S1) which are present at Ezh2 and H3K27me3 positive regions (Fig. 1E). We found that 70% of the Ezh2 peaks were also positive for Rnf2 in wildtype embryos (Fig. S2A).

In *MZezh2* mutant embryos, only 14 binding sites could be detected for Rnf2 (Fig. 1C) and Rnf2 average binding (measured in RPKM) was reduced to background level, as observed for Ezh2 and H3K27me3 binding (Fig. 1D). This loss of Rnf2 was observed at both gene clusters such as *hoxab* (Fig. 1F) and individual transcription factors such as *tbx3a*, *tbx5a*, *isl1*, and *gsc* (Fig. 1G, Fig. S2B). Similar to Ezh2 and H3K27me3, Rnf2 remaining peaks in *MZezh2* mutant embryos were detected in intergenic regions with repeat sequences and these are most probably also artefacts (Fig. S2B).

Furthermore, H2AK119ub was barely detectable in core histone extracts from *MZezh2* mutant embryos (Fig. S2C and Fig. S4), suggesting an impaired functional recruitment of canonical PRC1 to the chromatin in absence of Ezh2.

Loss of H3K27me3 in *MZezh2* mutant embryos induces gene specific gain of H3K4me3

As PcG and TrxG complexes are known to have an antagonistic effect on gene expression (Piunti and Shilatifard, 2016), we investigated whether the loss of H3K27me3 in *MZezh2* mutant embryos changed the deposition of H3K4me3, a mark associated with gene activation.

To this aim, we performed ChIP-seq for H3K4me3 in triplicates in both wildtype and *MZezh2* mutant embryos at 24 hpf. We observed a similar distribution of H3K4me3 peaks, with 10,556 peaks detected in wildtype embryos and 10,096 in *MZezh2* mutants (Fig. 2A, Table S1). A majority of 9,550 peaks were shared between wildtype and *MZezh2* mutant embryos (Fig. 2A), suggesting little to no differences in H3K4me3 deposition in absence of Ezh2.

We next assessed the differences in H3K4me3 peak intensity upon loss of Ezh2 by performing differential binding analysis using DiffBind. We identified 95 peaks with an enriched H3K4me3 deposition and 28 peaks with a decreased H3K4me3 intensity in *MZezh2* mutant compared with wildtype (Fig. 2B). Analysis of H3K4me3 coverage confirmed the increase of H3K4me3 binding of the sites detected by DiffBind whereas the decrease in H3K4me3 binding appeared less pronounced (Fig. 2C, upper left panel). Comparisons with Ezh2 and H3K27me3 ChIP-seq showed a clear enrichment in Ezh2 and H3K27me3 binding on the peaks enriched in H3K4me3 in *MZezh2* mutants (Fig. 2C upper right and lower left panels). A majority of the peaks enriched for H3K4me3 are PcG targets, with 74% (70 out of 95) targeted by Ezh2 or H3K27me3, which is more than expected by chance (p-*adj* < 0.001). Peaks with decreased H3K4me3 deposition show little enrichment in PcG targets (0.01 ≤ p-*adj* < 0.05, Fig. 2D) and coverage (Fig. 2C). This result shows that the targets of PcG repression in wildtype are more susceptible to present an altered H3K4me3 profile upon loss of Ezh2/H3K27me3.

We then searched for the closest genes from the regions with increased H3K4me3 peak coverage detected by DiffBind and identified 118 genes. For example, the transcription factors *hoxa13b*, *tbx5a*, and *gsc* showed enrichment for H3K4me3 close to their promoter (Fig. 2E). Gene ontology analysis revealed that these genes were mainly involved in transcriptional regulation and organismal development (Fig. 2F). Among these 118 identified genes, 51 encode for transcription factors, including members of the *hox*, *tbx*, *sox*, and *pax* gene families, and known targets of PcG complexes.

These results show that, on the whole embryo level, loss of PcG repression has an overall limited effect on the H3K4me3 active epigenetic mark at 24 hpf, and that the genes presenting an increase in H3K4me3 deposition are mainly transcription factors directly targeted by PcG repression.

Epigenetic changes in *MZezh2* mutant embryos have minor effects on the transcriptome and proteome

The *MZezh2* mutant embryos completely lack the H3K27me3 repressive mark and show a subtle yet selective increased deposition of H3K4me3 activating mark on genes coding for transcription factors. Therefore, we investigated the effect of loss of Ezh2 on the transcriptome and proteome of wildtype and *MZezh2* mutant embryos at 24 hpf.

Transcriptome analysis by RNA-seq in the two conditions revealed only 60 genes significantly upregulated ($log_2FC \geq 1$ and *P*-adj < 0.05) and 28 genes downregulated ($log_2FC \leq -1$ and *P*-adj < 0.05) in *MZezh2* mutant compared with wildtype embryos (Fig. 3A, Table S1). We also performed a proteome analysis on whole embryo extracts in both *MZezh2* mutant and wildtype conditions. This analysis identified 111 upregulated ($log_2FC \geq 1.5$ and *P-adj* < 0.05) and 110 downregulated ($log_2FC \leq -1.5$ and *P-adj* < 0.05) proteins in *MZezh2* mutants compared to wildtype controls (Fig. 3B, Table S1).

GO analysis showed that the dysregulated genes in the transcriptomic data are associated with control of organism development and regulation of transcription (Fig. 3C). The proteins dysregulated in the proteome analysis revealed anatomy terms associated with organs presenting clear phenotypes or even absence in the *MZezh2* mutant embryos, such as optic bud, heart tube, and fins (Fig. 3D). Therefore, it seems like the proteomic analysis reflects the observed phenotypes caused by the disturbed gene expression detected by the transcriptome analysis.

When comparing our RNA-seq results with our ChIP-seq data, we found that upregulated genes are preferentially associated with H3K27me3 (Fig. 3E, left panel), Ezh2, and Rnf2 (Fig. S5A,B) target genes. Quantification showed that 60% (36 out of 60) of the upregulated genes are targets of H3K27me3,

which is more than expected by chance (p-*adj* < 0.001) (Fig. S5C). Interestingly, genes with the higher overexpression are among genes with the higher H3K27me3 coverage (Fig. 3E, left panel). In contrast, upregulated genes show no association with H3K4me3 in wildtype condition, except for genes decorated by both H3K27me3 and H3K4me3 (Fig. 3E, right panel), but are associated with gain of H3K4me3 in *MZezh2* mutant condition (Fig. S5C). The downregulated genes also show significant association with H3K27me3, but did not show any correlation with gain or loss of H3K4me3 deposition (Fig. 3E and Fig. S5C). This could be explained, for example, by a secondary effect, such as overexpression of a repressor of these genes, or because H3K27me3 could be targeting these genes only in a subset of cells. We cannot distinguish between these potential causes, since experiments were done on whole embryos. In contrast, the proteomics data did not present any correlation with either the ChIP-seq or the RNA-seq results (Fig. 3F and Fig. S5D). It appears that proteomic analyses could not detect proteins encoded by H3K27me3 target genes, as demonstrated by the general low H3K27me3 coverage and absence of H3K27me3 targets among the proteins detected by the experiment (Fig. 3F).

Finally, proteome data indicate that, in addition to Ezh2, Suz12b is downregulated in *MZezh2* mutant embryos whereas other PRC2 core subunits were either not detected or not significantly downregulated (Fig. 3B and Fig. S6). Subunits of the canonical PRC1 complex were mostly not detected or not significantly overexpressed (Fig. S6).

Ezh2 controls maternal mRNA load in embryos

It is surprising that only a small number of genes is dysregulated upon loss of Ezh2 on whole embryo level. One could argue that gene expression levels are more dramatically changed when looking at specific cell populations. We therefore explored gene dysregulation at 0 hpf, before zygotic genome activation, and 3.3 hpf, when zygotic genome is activated and cell identity is more homogeneous than at 24 hpf.

We found 1,859 upregulated genes and 69 downregulated genes in *MZezh2* mutant embryos when compared with wildtype controls at 3.3 hpf (Fig. S7A). This distribution of dysregulated genes was similar to the results obtained in 1-cell stage embryos, when only maternal mRNAs are present, with 1,936 genes upregulated and 78 genes downregulated in *MZezh2* mutant embryos compared to wildtype controls (Fig. S7B).

Comparisons between time points show that genes overexpressed in *MZezh2* mutants at 0 and 3.3 hpf greatly overlap, whereas genes overexpressed at 24 hpf are more different (Fig. S7C). However, important transcription factors such as *gsc*, various *hox* genes, and *tbx5a* are dysregulated both at 24 hpf and 0 or 3.3 hpf. GO analysis on genes overexpressed upon loss of Ezh2 at 0 and 3.3 hpf identified specific terms clearly associated with late developmental processes and organogenesis and not with pre-gastrulation events. For example, axon guidance, neural crest cell development, and cardiac muscle cell differentiation were among the top terms identified (Fig. S7D).

These observations suggest that Ezh2 is important to control the load of maternal mRNAs and only later during development to maintain zygotic gene expression.

Loss of *ezh2* results in expression of *hox* genes outside their normal expression domains

We next carried out a spatial expression analysis on selected target genes to distinguish between the possibilities that absence of PcG-mediated repression leads to global but moderate gene dysregulation or to more severe gene dysregulation limited to specific cell types or tissues. We focused on embryos of 24 hpf, when mutants show the first phenotypes without lethality or apoptosis (San et al., 2016).

To start with, we concentrate on expression of different genes from the *hox* gene family. These genes are known targets of Polycomb-mediated repression (Mallo and Alonso, 2013). Every *hox* gene has an expression pattern that is restricted along the anterior-posterior axis (Prince et al., 1998). To obtain spatially resolved data along the anterior-posterior axis, we performed RT-qPCR on the anterior half and the posterior half of 24 hpf wildtype and *MZezh2* mutant embryos. We then compared the normalized relative expression levels between the different halves of the *MZezh2* mutant and wildtype embryos. The tested *hox* genes were selected based on their domain of expression along the anterior-posterior axis (Fig. 4A-D). The *hoxa9a* gene is expressed mostly in the posterior half of the embryo in wildtype embryos. In *MZezh2* mutant embryos *hoxa9a* expression increased only in the anterior part, to reach expression levels similar to the wildtype posterior expression (Fig. 4A). No significant differences were detected in the level of expression when comparing the posterior compartment of *MZezh2* mutant and wildtype embryos (Fig. 4A). Similar results were obtained for *hoxa9b*, where overexpression was detected in the anterior compartment of *MZezh2* mutant embryos compared to the anterior compartment of wildtype embryos (Fig. 4B). The *hoxa11b* and *hoxa13b* genes showed higher expression in the posterior half of the wildtype embryos compared to the anterior half (Fig. 4C,D). In the *MZezh2* mutant embryos, both *hox* genes were upregulated in the anterior half of the *MZezh2* mutant embryos compared to wildtypes but their expression level remained lower than in the posterior half of the wildtype embryos (Fig. 4C,D).

Our results are in agreement with previously published data where *hoxa9b*, *hoxd9a*, *hoxc8a*, and *hoxc6a* were shown to be ectopically expressed anteriorly in *MZezh2* mutant embryos (San et al., 2016). These comparative analyses of anterior and posterior parts of the embryo suggest that, upon loss of Ezh2, *hox* genes show ectopic anterior expression while keeping wildtype expression levels within their normal expression domains.

Different transcription factors show various profiles of dysregulation in the absence of Ezh2

To further pursue our investigation on the changes in gene expression patterns in absence of Ezh2, we performed *in situ* hybridization (ISH) on members from the *tbx* gene family of transcription factors. The *tbx2a*, *tbx2b*, *tbx3a*, and *tbx5a* genes have partial overlapping expression patterns in wildtype embryos, but also display gene specific expression domains (Fig. 5A). At 24 hpf, these *tbx* gene family members are expressed in the dorsal region of the retina, in the heart, and the pectoral fins (Ribeiro et al., 2007; Tamura et al., 1999). In addition, *tbx2a*, *tbx2b*, and *tbx3a* are expressed in the otic vesicle. The genes *tbx2b* and *tbx3a* are expressed in different ganglions and neurons in anterior and posterior regions of wildtype embryos (Ribeiro et al., 2007). Finally, expression of *tbx2b* can also be detected in part of pharyngeal arches 3-7 and the distal region of the pronephros and *tbx3a* expression can be detected in the branchial arches (Thisse, 2004). This spatial prevalence of *tbx* gene expression in the anterior half of the embryo was also detected by RT-qPCR at 24 hpf, where *tbx2a*, *tbx2b,* and *tbx5a* expression was significantly higher in the anterior than in the posterior part of wildtype embryos (Fig. 5B).

ISH for these *tbx genes* on *MZezh2* mutant embryos at 24 hpf suggest ectopic expression of these transcription factors around their normal expression pattern in the eye, the otic vesicle, and the heart, except for *tbx2b* (Fig. 5A). This scattering in gene expression was reflected in a trend towards a higher expression in the anterior half of *MZezh2* mutant embryos as detected by RT-qPCR, in which *tbx2a* and *tbx5a* expression showed significant upregulation upon the loss of Ezh2 (Fig. 5B). In addition, ISH for *tbx5a*, and to a lesser extent *tbx3a*, showed ubiquitous expression throughout the entire body of *MZezh2* mutants which was not visible in wildtypes (Fig. 5A). RT-qPCR results confirmed increased expression of *tbx5a* in both the anterior and posterior half of the *MZezh2* mutant embryos (Fig. 5B).

Besides the observed ectopic expression, all tested *tbx* genes showed absence of expression in specific structures upon Ezh2 loss. For example, in *MZezh2* mutant embryos there are no fin buds formed (San et al., 2016), and there is no expression of all four *tbx* genes in the region where the fin buds would normally be present (Fig. 5A). In *MZezh2* mutant embryos, the gene *tbx2b* showed no expression in

the pharyngeal arches 3-7 and the lateral line ganglions, and *tbx3a* was not observed in the branchial arches (Fig. 5A). This absence of expression was not detected by RT-qPCR (Fig. 5B) but a trend towards downregulation for *tbx2b* was observed in RNA-seq results on whole *MZezh2* mutant embryo lysates (Fig. 5C).

In addition, we tested transcription factors from other gene families which are targeted by H3K27me3 in wildtype embryos. The transcription factor *isl1*, expressed in all primary neurons (Dyer et al., 2014), showed a similar absence of expression in the fin bud and the cranial motor neurons in the midbrain (trigeminal, facial and vagal motor neurons), as observed for *tbx2a*. Its expression was also absent in the ventral region of the eye, the facial ganglia, and in the pronephros from *MZezh2* mutant embryos, where it is normally expressed in wildtype embryos (Heisenberg et al., 1999; Zhang et al., 2017) (Fig. 5A). This loss of expression in *MZezh2* mutant embryos was not detected by RT-qPCR but a clear tendency towards downregulation was detected by RNA-seq (Fig. 5B,C). Even more surprising was the expression pattern of *gsc* in the *MZezh2* mutant embryos. Wildtype embryos show highly specific *gsc* expression in the telencephalon and diencephalon nuclei, the branchial arches, and the otic vesicle (Thisse, 2004). This expression was lost in *MZezh2* mutant embryos and instead diffuse expression was observed (Fig. 5A). This observation was confirmed by RT-qPCR and RNA-seq where upregulation of *gsc* was clearly detected in *MZezh2* mutant embryos (Fig. 5B,C).

Taken together, these spatial expression analyses show that the tested transcription factors are expressed outside their normal wildtype expression boundaries in *MZezh2* mutant embryos at 24 hpf. Furthermore, expression of a subset of these genes is lost in specific tissues in the *MZezh2* mutant embryos.

Discussion

Here, we showed for the first time the genome-wide binding patterns of Ezh2 and Rnf2, the catalytic subunits of PRC2 and PRC1, respectively, in 24 hpf zebrafish embryos. The overall overlap between the binding patterns of Ezh2, Rnf2, and the PcG related epigenetic mark H3K27me3 suggests that the PcG-mediated gene repression mechanisms (Chittock et al., 2017) are evolutionary conserved in zebrafish development. The complete loss of H3K27me3 in *MZezh2* mutant embryos reveals that Ezh2 is the only methyltransferase involved in trimethylation of H3K27 during early zebrafish development. This result was expected as Ezh1, the only other H3K27me3 methyltransferase, was shown by a number of studies to be not maternally loaded nor expressed in the zebrafish embryo until at least after 1 dpf (Chrispijn et al., 2018; San et al., 2016; Sun et al., 2008; White et al., 2017). In addition, proteomic results showed decreased protein expression of most PRC2 subunits. This could indicate a

destabilization of PRC2 in absence of the catalytic subunit in *MZezh2* mutant embryos. We could therefore confirm that zebrafish embryos can form a normal body plan in the absence of PRC2-mediated gene repression.

The loss of Rnf2 binding in the *MZezh2* mutants suggests that only the canonical pathway, in which PRC2 is required for PRC1 recruitment, is active during this stage of development. This absence of PRC1 recruitment to the chromatin is not caused by an absence of the complex in the *MZezh2* mutants, since most of the PRC1 subunits were detectable and not dysregulated as shown by proteomic analysis. This is in contrast with studies in cultured mouse embryonic stem cells where non-canonical PRC1 complexes were shown to be recruited to developmental regulated genes independently of PRC2 (He et al., 2013; Tavares et al., 2012). This difference could be explained by the complete absence of H3K27me3 as from fertilization onwards in *MZezh2* mutant embryos, whereas other studies used conditional knockdown. Therefore, our model potentially suggests that the PRC2-independent recruitment of PRC1 during early development can occur if PRC1 recruitment was first primed by a PRC2-dependent mechanism happening earlier during development.

As repressive and activating marks are known to antagonize each other (Schmitges et al., 2011), one could expect an increase in the H3K4me3 level deposited by TrxG proteins in absence of H3K27me3 associated with an increase in gene activation. However, the effects on H3K4me3 deposition, gene expression, and protein expression are limited in *MZezh2* mutant embryos at 24 hpf. This observation is in agreement with the near complete absence of phenotype at this developmental time point. Thus, it appears that transcriptional regulation during zebrafish development is largely PRC2-independent until later stages of development, when maintenance of cellular identity is required. Ezh2, and thereby the PRC2 complex, could therefore be responsible for this maintenance, which seems crucial for development and growth. Yet, these defects were not associated with apoptosis (San et al., 2016). These results were unexpected, as PRC2 is described to be essential during mammalian development already during gastrulation (Faust et al., 1998; O'Carroll et al., 2001; Pasini et al., 2004). It implies, that even if PcG-mediated repression mechanisms are conserved, the developmental stages at which these mechanisms are required differ between species. Possibly, the external development of the zebrafish and its rapid early development could explain this difference in phenotype.

We also hypothesized that gene dysregulation in absence of Ezh2 is intense but limited to a subset of cells. To this aim, we performed transcriptome analyses during maternally controlled development (0 hpf) and after zygotic gene activation (3.3 hpf), time points at which the embryo contains one cell or a more homogenous population of cells. First, these transcriptome analyses revealed that Ezh2 is important to control the maternal mRNA load transmitted to the embryos. Indeed, in our germ cell

transplantation model, the parental females possess mostly wildtype somatic cells but a zygotic *ezh2* (*Zezh2*) mutant germ line. Thus, oogenesis occurs in absence of Ezh2 and leads to the production of oocytes with a modified maternal mRNA load, as reflected by the 0 hpf mutant transcriptome. GO analysis showed that the dysregulated genes belong to developmental pathways normally activated later during development, at the time of organogenesis. It is therefore surprising that the eggs containing a modified maternal mRNA load can mature properly and that the zygote can develop normally until long after the maternally controlled stage of development is over. We hypothesize that these ectopically expressed mRNAs are never translated or that other genes belonging to the same pathways are not expressed, preventing early activation of these late developmental processes.

Secondly, the comparison of the transcriptome analysis performed at 0 hpf with 3.3 hpf shows that mainly maternal mRNAs are dysregulated. This observation suggests that PRC2-dependent gene repression is not limited to a subset of cells during early development but is rather not required or only to a very limited extent until 24. hpf.

Although limited, genes that show a gain in H3K4me3 deposition or in expression upon loss of *ezh2* at 24 hpf are mainly transcription factors targeted by H3K27me3 in wildtype embryos. That only a minor fraction of all H3K27me3 target genes gained expression (36 out of 2610 = 1.2%, Fig 3E) suggests different mechanisms of regulation of PcG target genes at this time. Our hypothesis is that control of gene expression by signaling pathways and transcription factor networks (McGinnis and Tickle, 2005) is a robust mechanism and can be maintained until 1 dpf in absence of repression by PcG complexes. At 1 dpf, in absence of PcG-mediated repression, the first derepressed genes will be the genes subjected to the most fine-tuned transcriptional control, such as genes controlled by precise morphogen gradients. For example, it was shown that PRC2 attenuates expression of genes controlled by retinoic acid signaling (Laursen et al., 2013; Zhang et al., 2014). In vertebrates, and most particularly zebrafish, retinoic acid signaling is responsible for induction of formation of, among others, the forelimb field (Cunningham et al., 2013; Grandel and Brand, 2011), dorsoventral patterning of eyes (Lupo et al., 2005; Marsh-Armstrong et al., 1994), hindbrain patterning (Maves and Kimmel, 2005), *hox* gene expression (White et al., 2007), and the development of other organs (Samarut et al., 2015). All these processes are affected in *MZezh2* mutant embryos at 24 hpf and onwards and, therefore, could be explained by a defect in the response to retinoic acid signaling.

Spatial analysis of gene expression revealed different effects on gene expression patterns caused by loss of Ezh2. Anterior-posterior specific RT-qPCR showed that *hox* genes become abnormally expressed in the anterior half of the *MZezh2* mutant embryos; whereas expression levels in the posterior half remain unchanged. These results are supported by previous studies showing ectopic

expression of *hox* genes in PRC1 and PRC2 zebrafish mutants (San et al., 2016; van der Velden et al., 2012), but also in other animal models (Kennison, 1995). Other transcription factors, such as the *tbx* gene family members, showed more diverse patterns of dysregulation compared to *hox* genes. ISH and RT-qPCR showed that, among the *tbx* genes examined, some were overexpressed outside their normal expression domains (*tbx2a, tbx3a, and tbx5a*), whereas others were also ubiquitously upregulated (*tbx3a* and *tbx5a*). The case of eye patterning is a good example of the defect in control of gene expression pattern in *MZezh2* mutant embryos. In wildtype embryos, at 24 hpf, *tbx* genes are expressed in the dorsal part of the eye whereas *isl1* is expressed in the ventral part. Upon loss of Ezh2, our ISH results showed that the expression of the *tbx* genes expands to the whole eye whereas *isl1* disappears from the ventral region. We conclude that Polycomb-mediated repression is therefore responsible for maintenance of expression domains rather than control of expression levels at this time of development in the zebrafish embryo.

Expression analysis by ISH for *hox* and *tbx* genes as well as for *isl1* also showed loss of expression in specific structures in *MZezh2* mutant embryos. We reasoned that the absence of expression of *hox* and *tbx* genes in the fin bud is due to the absence of this structure in *MZezh2* mutants (San et al., 2016). The same phenomenon, absence of specific structures, could explain the lack of detection of *tbx2b* and *isl1* in pharyngeal arches, pronephros, and lateral line ganglions. The case of *gsc* expression is more striking, as its normal expression pattern is totally abolished and a diffuse expression pattern is detected. The *gsc* gene is known to be expressed in the Spemann organizer during gastrulation and therefore all cells will transiently express *gsc* when undergoing gastrulation (Joubin and Stern, 1999; Stachel et al., 1993). In absence of Ezh2, *gsc* expression could remain active in all cells after leaving the Spemann organizer, leading to a diffuse expression pattern and impaired tissue specific expression in 24 hpf *MZezh2* mutant embryos.

To conclude, our results show that major characteristics of PcG-mediated repression are conserved in zebrafish, including canonical recruitment or PcG complexes and their function in maintenance of pre-established gene expression patterns. Our use of a mutant depleted of both maternal and zygotic contribution of Ezh2 also reveals that no PRC2-independent recruitment of PRC1 occurs at this stage of development. Finally, we demonstrate that early embryonic development, including germ layer formation and cell fate specification, is independent of PcG-mediated gene repression until axes are formed and organs specified. PcG-mediated gene repression is then required to control precise spatial restricted expression of specific transcription factors. We hypothesize that subtle changes in expression of these important genes subsequently will lead to progressive and accumulating changes in gene network regulation and result in loss of tissue identity maintenance.

This surprising result highlights the fact that, despite the conservation of PcG-mediated repression mechanisms during evolution, the time frame within which PcG repression is required for proper development may vary greatly between species. Studying the PcG repression in additional species would improve our understanding of the importance of PcG biology during development.

Materials and methods

Zebrafish genetics and strains

Zebrafish (*Danio rerio*), were housed according to standard conditions (Westerfield, 2000) and staged according to Kimmel et al. (Kimmel et al., 1995). The *ezh2* nonsense mutant (*hu5670*) (San et al., 2016), *Tg (H2A::GFP)* (Pauls et al., 2001), and *Tg (vas::eGFP)* (Krovel and Olsen, 2002) zebrafish lines have been described before. Genotyping of the *ezh2* allele was performed as previously described (San et al., 2016) with following adaptations: different primer pairs were used for PCR and nested PCR (Table S2), of which the restriction profile is shown on Fig. S2D. All experiments were carried out in accordance with animal welfare laws, guidelines, and policies and were approved by the Radboud University Animal Experiments Committee.

Germ cell transplantation

Germ cell transplantation was performed as described previously (San et al., 2016). For all experiments below, *ezh2* germline mutant females were crossed with *ezh2* germline mutant males to obtain 100% *MZezh2* mutant progeny. The germline wild-type sibling males and females obtained during transplantation were used to obtain 100% wildtype progeny with similar genetic background and are referred to as wildtype. The embryos used were all from the first generation after germline transplantation.

Western blotting

At 3.3 hpf, 50 embryos were collected, resuspended in in 500 µl ½ Ringer solution (55 mM NaCl, 1.8 mM KCl, 1.25 mM NaHCO$_3$) and forced through a 21G needle and a cell strainer in order to remove the chorion and disrupt the yolk. At 24 hpf, 20 embryos were collected and resuspended by thorough pipetting in 500µl ½ Ringer solution in order to disrupt the yolk. The samples of 3.3 and 24 hpf were centrifuged for 5 minutes at 3,500 g at 4°C and washed two additional times with 500 µl ½ Ringer solution. The embryo pellet was frozen in liquid nitrogen and stored at -80°C. Whole protein extraction was performed by adding 40 µl of RIPA buffer (100 mM Tris-HCl pH 8, 300 mM NaCl, 2% NP-40, 1% Sodium Deoxycholate, 0.2% SDS, 20% glycerol, 1x cOmplete EDTA-free protease inhibitor cocktails from Sigma) and sonication for 2 cycles of 15s ON and 15s OFF on medium power at 4°C on a

PicoBioruptor (Diagenode). After 10 minutes incubation at 4°C, embryo lysates were centrifuged for 12 minutes at 16,000 g at 4°C and supernatant was transferred in a new tube. 20 µg protein was mixed with SDS containing sample loading buffer, denaturated at 95°C for 5 minutes and analyzed by Western blot analysis. Antibodies used for immunoblotting are described in Table S3 HRP-conjugated anti-rabbit secondary antibody was used (Table S3) and protein detection was performed with ECL Select Western Blotting Detection Reagent (GE Healthcare, RPN2235) on an ImageQuant LAS 4000 (GE Healthcare).The anti-H2A western-blot was performed on histone extracts, obtained according to van der Velden et al. (van der Velden et al., 2012), and detected on X-ray film. Full, uncropped, blots used for Fig. 1B and Fig. S2C are available in Fig. S1 and S4, respectively.

ChIP-sequencing

For chromatin preparation, embryos from a germline mutant or germline wildtype incross were collected at 24 hpf and processed per batches of 300 embryos. Embryos were first dechorionated by pronase (0.6 µg/µl) treatment and then extensively washed with E3 medium. Subsequently, embryos were fixed in 1% PFA (EMS, 15710) for 15 minutes at room temperature and fixation was terminated by adding 0.125M glycine and washed 3 times in cold PBS. Yolk from fixed embryos was disrupted by pipetting the fixed embryos 10 times with a 1 ml tip in 600 µl of ½ Ringer solution (55 mM NaCl, 1.8 mM KCl, 1.25 mM NaHCO$_3$) and incubated for 5 minutes at 4°C on a rotating wheel. Embryos were pelleted by centrifuging 30 seconds at 300 g and the supernatant was removed. De-yolked embryos were resuspended in 600 µl sonication buffer (20 mM Tris-HCl pH 7.5, 70 mM KCl, 1 mM EDTA, 10% glycerol, 0.125% NP40, 1x cOmplete EDTA-free protease inhibitor cocktails from Sigma) and homogenized with a Dounce homogenizer (6 strokes with pestle A, followed by 6 strokes with pestle B). Homogenates were sonicated for 6 cycles of 30 seconds ON/30 seconds OFF on a PicoBioruptor (Diagenode), centrifuged for 10 minutes at 16,000 g at 4°C, and the supernatant containing the chromatin was stored at -80°C. 20 µl of the supernatant was subjected to phenol-chloroform extraction and ran on an agarose gel to verify that a proper chromatin size of 200-400 bp was obtained. For ChIP, 100 µl of chromatin preparation (corresponding to 50 embryos) was mixed with 100 µl IP-buffer (50 mM Tris-HCL pH 7.5, 100 mM NaCl, 2 mM EDTA, 1% NP-40, 1x cOmplete EDTA-free protease inhibitor cocktails from Sigma) and antibody (for details on antibodies used see Table S3) and incubated overnight at 4°C on a rotating wheel. When relevant, *Drosophila* chromatin and anti-H2Av were used according to manufacturer's instructions were followed (Active Motif, 53093 and 61686). For immunoprecipitation, 20 µl of protein G magnetic beads (Invitrogen, 1003D) were washed in IP buffer and then incubated with the chromatin mix for 2 hours at 4°C on a rotating wheel. Samples were washed in 500 µl washing buffer 1 (IP-buffer + 0.1% Sodium Deoxycholate), followed by washing

in washing buffer 2 (washing buffer 1 + 400mM NaCl), washing buffer 3 (washing buffer 1 + 250mM LiCl), washing buffer 1 and a final wash in 250 µl of TE buffer. All washes were 5 minutes at 4°C on a rotating wheel. Chromatin was eluted from the beads by incubation in 100 µl of elution buffer (50 mM NaHCO$_3$ pH 8.8, 1% SDS) for 15 minutes at 65°C at 900 rpm in a thermomixer. The supernatant was transferred in a clean 1.5 ml tube. Elution was repeated a second time and both supernatants were pooled. The eluate was treated with 0.33 µg/µl RNaseA for 2 hours at 37°C. Samples were then decrosslinked by adding 10 µl of 4M NaCL and 1 µl of 10mg/ml proteinase K and incubated overnight at 65°C. DNA was then purified using MinElute Reaction Clean-Up kit (Qiagen, 28204).

1-5 ng of DNA was used to prepare libraries with the KAPA Hyper Prep Kit (KAPABiosystems, KK8504) and NEXTflex ChIP-Seq Barcodes for Illumina (Bioo Scientific, 514122) followed by paired-end 43bp sequencing on an Illumina NextSeq500 platform. All ChIP-seq were performed in two biological replicates, except for H3K4me3 in *MZezh2* mutant and wildtype embryos which were performed in triplicate.

RNA-sequencing

Ten to twenty manually dechorionated 24 hpf embryos of a germline mutant incross and a germline wildtype incross were homogenized in TRIzol (Ambion, 15596018). For 0 and 3.3 hpf, twenty non-dechorionated embryos were collected and homogenized in Trizol. Subsequently, the Quick RNA microprep kit (Zymo Research, R1051) was used to isolate RNA and treat the samples with DNAseI. Samples were depleted from rRNA using the Ribo-Zero rRNA Removal Kit (Illumina, MRZH11124), followed by fragmentation, cDNA synthesis, and libraries were generated using the KAPA Hyper Prep Kit (KAPABiosystems, KK8504). Sequencing libraries were paired-end sequenced (43 bp read-length) on an Illumina NextSeq500 platform. However, two samples per genotype at 24 hpf were generated with the TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero (Illumina, RS-122-2201) and single-end sequenced (50 bp read-length) on an Illumina HiSeq 2500. For wildtype and *MZezh2* mutant embryos, 6 and 7 biological replicates were used, respectively.

Mass spectrometry

At 24 hpf, 50 embryos were collected, dechorionated, and resuspended by gently pipetting in 500 µl deyolking buffer (1/2 Ginzburg Fish Ringer without Calcium: 55 mM NaCl, 1.8 mM KCl, 1.25 mM NaHCO3, 1x cOmplete EDTA-free protease inhibitor cocktail from Sigma) and incubated for 5 minutes in a Thermomixer at RT at 1,100 rpm to disrupt the yolk. The samples were then centrifuged for 30 seconds at 400 g and the pellet was washed two times in 0.5 ml wash buffer (110 mM NaCl, 3.5 mM KCl, 2.7 mM CaCl2, 10mM Tris/Cl pH8.5, 1x cOmplete EDTA-free protease inhibitor cocktail from

Sigma) for 2 minutes in a Thermomixer at RT and 1,100 rpm, followed by 30 seconds centrifugation at 400 g. Washed pellets were lysed in 100 μl RIPA buffer (50 mM Tris pH8.0, 150 mM NaCl, 0.1% SDS, 1% NP-40, 0.5% DOC, 20% glycerol, 1 mM Sodium Orthovanadate, 1x cOmplete EDTA-free protease inhibitor cocktails from Sigma) and sonicated for 2 cycles of 15s ON and 15s OFF on full power at 4°C on a Bioruptor (Diagenode). Samples were incubated for 1 hour on a rotating wheel at 4°C and centrifuged 10 minutes at 12,000 g and 4°C. Supernatant was flash frozen and stored at -80°C. After Bradford analysis, 100 μg protein lysate was used for FASP-SAX as previously described (Wisniewski et al., 2011). The peptide fractions were separated on an Easy nLC 1000 (Thermo Scientific) connected to a Thermo scientific Orbitrap Fusion Tribrid mass spectrometer. MS and MS/MS spectra were recorded in a top speed modus with a run cycle of 3s using Higher- energy Collision Dissociation (HCD) fragmentation. The raw mass spectrometry data were analyzed using the MAXQuant software version 1.6.0.1 (Cox and Mann, 2008) with default settings. Data was searched against the *Danio rerio* data base (UniProt June 2017). The experiment was performed with biological triplicates for each condition.

Bioinformatics analyses

For ChIP-seq analysis, fastq files were aligned to GRCz10 zebrafish genome version using BWA-MEM (version 0.7.10-r789) for paired-end reads (Li and Durbin, 2009). Statistics for all high-throughput sequencing samples generated for this study are presented in Table S4 and Fig. S8. Duplicated and multimapping reads were removed using samtools (Li et al., 2009) version 1.2 and Picard tools (http://broadinstitute.github.io/picard) version 2.14.1. When spike-in normalization was used, *Drosophila* reads were aligned to dm6 *Drosophila* genome version and a normalization factor was then applied to zebrafish reads according to manufacturer's protocol (Active Motif, 53093 and 61686). MACS2 (Zhang et al., 2008) version 2.1.1 was used to call peaks from each aligned bam files using an Input track from 24 hpf wild-type embryos as control sequence. Peaks separated by less than 1kb distance were merged, peaks that were called using Input alone were removed from all data sets using bedtools suit version 2.20.1, and the intersection between the replicates for each antibody in each condition was used to define the definitive peak sets. For visualization in heatmaps and genome browser snapshots, fastq files from duplicate ChIP-sequencing were merged, aligned as described above, subsampled to equalized read numbers between wildtype and *MZezh2* mutant conditions for each ChIP, and transformed into bigwig alignment files using bam2bw version 1.25. Peak lists were analyzed using bedtools and heatmaps were produced using deepTools plotHeatmap (Ramirez et al., 2016) version 2.5.3. Comparison between H3K4me3 peaks in *MZezh2* mutant and wildtype conditions was performed using DiffBind version 2.10.0 on the union between H3K4me3 peaks detected in both conditions.

For RNA-sequencing analysis, read counts per gene were retrieved using GeneCounts quantification method from STAR (Dobin et al., 2013) version 2.4.0 and the GRCz10 zebrafish genome version with Ensembl annotation version 87 as reference. Differential expression analysis was calculated with DESeq2 (Love et al., 2014) version 1.14.1.

For proteomics analysis, differential expression of protein between conditions was assessed with DEP (Zhang et al., 2018) version 1.2.0.

Gene Ontology analyses on selected genes were performed using DAVID bioinformatics resources (Huang da et al., 2009) version 6.8 and anatomical term enrichment was done using ZEOGS (Prykhozhij et al., 2013).

Whole mount *in situ* hybridization

Embryos at 24 hpf were dechorionated and fixed overnight at 4°C in 4% PFA in PBST (0.1% Tween), after which they were gradually transferred to 100% methanol. Prior to ISH, embryos were gradually transferred back to PBST and, subsequently, ISH was performed as described previously (Houwing et al., 2007). ISH were performed simultaneously for wildtype and *MZezh2* mutant embryos, with the same probe and chemical mixes and identical signal development time. The embryos were imaged by light microscopy on a Leica MZFLIII, equipped with a DFC450 camera.

RT-qPCR analyses

Total RNA was isolated using Trizol from 20 flash-frozen dechorionated 24 hpf wildtype and *MZezh2* mutant embryos cut in two with tweezers. Reverse transcription was achieved using Superscript III (Invitrogen, 18080093) and poly-dT primers. Standard qPCR using SYBR Green (iQ SYBR Green Supermix, BioRad, 1708880) was performed using the primers shown in Table S2. Relative expression was calculated based on expression of housekeeping genes *β-actin*. Comparable results were obtained using *eif1a* as reference gene (Fig. S9). Calculations were based on at least 3 independent replicates for both conditions.

Data availability

The sequencing data have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE119070. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository with the dataset identifier PXD010922.

References

Brock, H. W. and Fisher, C. L. (2005). Maintenance of gene expression patterns. *Developmental dynamics : an official publication of the American Association of Anatomists* **232**, 633-655.

Chittock, E. C., Latwiel, S., Miller, T. C. and Muller, C. W. (2017). Molecular architecture of polycomb repressive complexes. *Biochemical Society transactions* **45**, 193-205.

Chrispijn, N. D., Andralojc, K. M., Castenmiller, C. and Kamminga, L. M. (2018). Gene expression profile of a selection of Polycomb Group genes during zebrafish embryonic and germ line development. *PloS one* **13**, e0200316.

Chrispijn, N. D., Elurbe, D. M., Mickoleit, M., Aben, M., de Bakker, D. E. M., Andralojc, K. M., Huisken, J., Bakkers, J. and Kamminga, L. M. (2019). Loss of the Polycomb group protein Rnf2 results in derepression of tbx-transcription factors and defects in embryonic and cardiac development. *Scientific reports* **9**, 4327.

Ciruna, B., Weidinger, G., Knaut, H., Thisse, B., Thisse, C., Raz, E. and Schier, A. F. (2002). Production of maternal-zygotic mutant zebrafish by germ-line replacement. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 14919-14924.

Cox, J. and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372.

Cunningham, T. J., Zhao, X., Sandell, L. L., Evans, S. M., Trainor, P. A. and Duester, G. (2013). Antagonism between retinoic acid and fibroblast growth factor signaling during limb development. *Cell reports* **3**, 1503-1511.

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.

Dupret, B., Volkel, P., Vennin, C., Toillon, R. A., Le Bourhis, X. and Angrand, P. O. (2017). The histone lysine methyltransferase Ezh2 is required for maintenance of the intestine integrity and for caudal fin regeneration in zebrafish. *Biochimica et biophysica acta* **1860**, 1079-1093.

Dyer, C., Linker, C., Graham, A. and Knight, R. (2014). Specification of sensory neurons occurs through diverse developmental programs functioning in the brain and spinal cord. *Developmental dynamics : an official publication of the American Association of Anatomists* **243**, 1429-1439.

Faust, C., Lawson, K. A., Schork, N. J., Thiel, B. and Magnuson, T. (1998). The Polycomb-group gene eed is required for normal morphogenetic movements during gastrulation in the mouse embryo. *Development* **125**, 4495-4506.

Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y. and Reinberg, D. (2012). PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Molecular cell* **45**, 344-356.

Grandel, H. and Brand, M. (2011). Zebrafish limb development is triggered by a retinoic acid signal during gastrulation. *Developmental dynamics : an official publication of the American Association of Anatomists* **240**, 1116-1126.

He, J., Shen, L., Wan, M., Taranova, O., Wu, H. and Zhang, Y. (2013). Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nature cell biology* **15**, 373-384.

Heisenberg, C. P., Brennan, C. and Wilson, S. W. (1999). Zebrafish aussicht mutant embryos exhibit widespread overexpression of ace (fgf8) and coincident defects in CNS development. *Development* **126**, 2129-2140.

Houwing, S., Kamminga, L. M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D. V., Blaser, H., Raz, E., Moens, C. B., et al. (2007). A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* **129**, 69-82.

Huang da, W., Sherman, B. T. and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57.

Joubin, K. and Stern, C. D. (1999). Molecular interactions continuously define the organizer during the cell movements of gastrulation. *Cell* **98**, 559-571.

Kennison, J. A. (1995). The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function. *Annual review of genetics* **29**, 289-303.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists* **203**, 253-310.

Kloet, S. L., Makowski, M. M., Baymaz, H. I., van Voorthuijsen, L., Karemaker, I. D., Santanach, A., Jansen, P., Di Croce, L. and Vermeulen, M. (2016). The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation. *Nature structural & molecular biology* **23**, 682-690.

Klymenko, T. and Muller, J. (2004). The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO reports* **5**, 373-377.

Krovel, A. V. and Olsen, L. C. (2002). Expression of a vas::EGFP transgene in primordial germ cells of the zebrafish. *Mechanisms of development* **116**, 141-150.

Laursen, K. B., Mongan, N. P., Zhuang, Y., Ng, M. M., Benoit, Y. D. and Gudas, L. J. (2013). Polycomb recruitment attenuates retinoic acid-induced transcription of the bivalent NR2F1 gene. *Nucleic acids research* **41**, 6430-6443.

Le Faou, P., Volkel, P. and Angrand, P. O. (2011). The zebrafish genes encoding the Polycomb repressive complex (PRC) 1. *Gene* **475**, 10-21.

Li, G. and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. *Current opinion in genetics & development* **21**, 175-186.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079.

Lindeman, L. C., Andersen, I. S., Reiner, A. H., Li, N., Aanes, H., Ostrup, O., Winata, C., Mathavan, S., Muller, F., Alestrom, P., et al. (2011). Prepatterning of developmental gene expression by modified histones before zygotic genome activation. *Developmental cell* **21**, 993-1004.

Loubiere, V., Delest, A., Thomas, A., Bonev, B., Schuettengruber, B., Sati, S., Martinez, A. M. and Cavalli, G. (2016). Coordinate redeployment of PRC1 proteins suppresses tumor formation during Drosophila development. *Nature genetics* **48**, 1436-1442.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550.

Lupo, G., Liu, Y., Qiu, R., Chandraratna, R. A., Barsacchi, G., He, R. Q. and Harris, W. A. (2005). Dorsoventral patterning of the Xenopus eye: a collaboration of Retinoid, Hedgehog and FGF receptor signaling. *Development* **132**, 1737-1748.

Mallo, M. and Alonso, C. R. (2013). The regulation of Hox gene expression during animal development. *Development* **140**, 3951-3963.

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C. L., Dynlacht, B. D. and Reinberg, D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Molecular cell* **32**, 503-518.

Marsh-Armstrong, N., McCaffery, P., Gilbert, W., Dowling, J. E. and Drager, U. C. (1994). Retinoic acid is necessary for development of the ventral retina in zebrafish. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 7286-7290.

Maves, L. and Kimmel, C. B. (2005). Dynamic and sequential patterning of the zebrafish posterior hindbrain by retinoic acid. *Developmental biology* **285**, 593-605.

McGinnis, W. and Tickle, C. (2005). Pattern formation and developmental mechanisms: good mileage from comparative studies in cell biology, gene regulation, development and evolution. *Current opinion in genetics & development* **15**, 355-357.

Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-560.

Murphy, P. J., Wu, S. F., James, C. R., Wike, C. L. and Cairns, B. R. (2018). Placeholder Nucleosomes Underlie Germline-to-Embryo DNA Methylation Reprogramming. *Cell* **172**, 993-1006 e1013.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S. C., Surani, M. A. and Jenuwein, T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. *Molecular and cellular biology* **21**, 4330-4336.

Pasini, D., Bracken, A. P., Jensen, M. R., Lazzerini Denchi, E. and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *The EMBO journal* **23**, 4061-4071.

Pauls, S., Geldmacher-Voss, B. and Campos-Ortega, J. A. (2001). A zebrafish histone variant H2A.F/Z and a transgenic H2A.F/Z:GFP fusion protein for in vivo studies of embryonic development. *Development genes and evolution* **211**, 603-610.

Piunti, A. and Shilatifard, A. (2016). Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* **352**, aad9780.

Potok, M. E., Nix, D. A., Parnell, T. J. and Cairns, B. R. (2013). Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* **153**, 759-772.

Prince, V. E., Joly, L., Ekker, M. and Ho, R. K. (1998). Zebrafish hox genes: genomic organization and modified colinear expression patterns in the trunk. *Development* **125**, 407-420.

Prykhozhij, S. V., Marsico, A. and Meijsing, S. H. (2013). Zebrafish Expression Ontology of Gene Sets (ZEOGS): a tool to analyze enrichment of zebrafish anatomical terms in large gene sets. *Zebrafish* **10**, 303-315.

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 research* **44**, W160-165.

Ribeiro, I., Kawakami, Y., Buscher, D., Raya, A., Rodriguez-Leon, J., Morita, M., Rodriguez Esteban, C. and Izpisua Belmonte, J. C. (2007). Tbx2 and Tbx3 regulate the dynamics of cell proliferation during heart remodeling. *PloS one* **2**, e398.

Samarut, E., Fraher, D., Laudet, V. and Gibert, Y. (2015). ZebRA: An overview of retinoic acid signaling during zebrafish development. *Biochimica et biophysica acta* **1849**, 73-83.

San, B., Aben, M., Elurbe, D. M., Voeltzke, K., Den Broeder, M. J., Rougeot, J., Legler, J. and Kamminga, L. M. (2018). Genetic and Epigenetic Regulation of Zebrafish Intestinal Development. *Epigenomes* **2**, 19.

San, B., Chrispijn, N. D., Wittkopp, N., van Heeringen, S. J., Lagendijk, A. K., Aben, M., Bakkers, J., Ketting, R. F. and Kamminga, L. M. (2016). Normal formation of a vertebrate body plan and loss of tissue maintenance in the absence of ezh2. *Scientific reports* **6**, 24658.

Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J. and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407-411.

Schmitges, F. W., Prusty, A. B., Faty, M., Stutzer, A., Lingaraju, G. M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S., et al. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. *Molecular cell* **42**, 330-341.

Schuettengruber, B., Bourbon, H. M., Di Croce, L. and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**, 34-57.

Shen, X., Liu, Y., Hsu, Y. J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G. C. and Orkin, S. H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Molecular cell* **32**, 491-502.

Stachel, S. E., Grunwald, D. J. and Myers, P. Z. (1993). Lithium perturbation and goosecoid expression identify a dorsal specification pathway in the pregastrula zebrafish. *Development* **117**, 1261-1274.

Sun, X. J., Xu, P. F., Zhou, T., Hu, M., Fu, C. T., Zhang, Y., Jin, Y., Chen, Y., Chen, S. J., Huang, Q. H., et al. (2008). Genome-wide survey and developmental expression mapping of zebrafish SET domain-containing genes. *PloS one* **3**, e1499.

Tamura, K., Yonei-Tamura, S. and Izpisua Belmonte, J. C. (1999). Differential expression of Tbx4 and Tbx5 in Zebrafish fin buds. *Mechanisms of development* **87**, 181-184.

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., et al. (2012). RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* **148**, 664-678.

Thisse, B. T., C. (2004). Fast Release Clones: A High Throughput Expression Analysis. *ZFIN Direct Data Submission (http://zfin.org).*

van der Velden, Y. U., Wang, L., van Lohuizen, M. and Haramis, A. P. (2012). The Polycomb group protein Ring1b is essential for pectoral fin development. *Development* **139**, 2210-2220.

Vastenhouw, N. L., Zhang, Y., Woods, I. G., Imam, F., Regev, A., Liu, X. S., Rinn, J. and Schier, A. F. (2010). Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* **464**, 922-926.

Vizcaino, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., et al. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic acids research* **44**, D447-456.

Voncken, J. W., Roelen, B. A., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J. and van Lohuizen, M. (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 2468-2473.

Westerfield, M. (2000). The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio). 4th ed., 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**.

White, R. J., Nie, Q., Lander, A. D. and Schilling, T. F. (2007). Complex regulation of cyp26a1 creates a robust retinoic acid gradient in the zebrafish embryo. *PLoS biology* **5**, e304.

Wisniewski, J. R., Ostasiewicz, P. and Mann, M. (2011). High recovery FASP applied to the proteomic analysis of microdissected formalin fixed paraffin embedded cancer tissues retrieves known colon cancer markers. *Journal of proteome research* **10**, 3040-3049.

Zhang, R., Knapp, M., Suzuki, K., Kajioka, D., Schmidt, J. M., Winkler, J., Yilmaz, O., Pleschka, M., Cao, J., Kockum, C. C., et al. (2017). ISL1 is a major susceptibility gene for classic bladder exstrophy and a regulator of urinary tract development. *Scientific reports* **7**, 42170.

Zhang, X., Smits, A. H., van Tilburg, G. B., Ovaa, H., Huber, W. and Vermeulen, M. (2018). Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nature protocols* **13**, 530-550.

Zhang, Y., Liang, J. and Li, Q. (2014). Coordinated regulation of retinoic acid signaling pathway by KDM5B and polycomb repressive complex 2. *Journal of cellular biochemistry* **115**, 1528-1538.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, R137.

Zhong, Y., Ye, Q., Chen, C., Wang, M. and Wang, H. (2018). Ezh2 promotes clock function and hematopoiesis independent of histone methyltransferase activity in zebrafish. *Nucleic acids research* **46**, 3382-3399.

Zhu, P. and Li, G. (2016). Structural insights of nucleosome and the 30-nm chromatin fiber. *Current opinion in structural biology* **36**, 106-115.
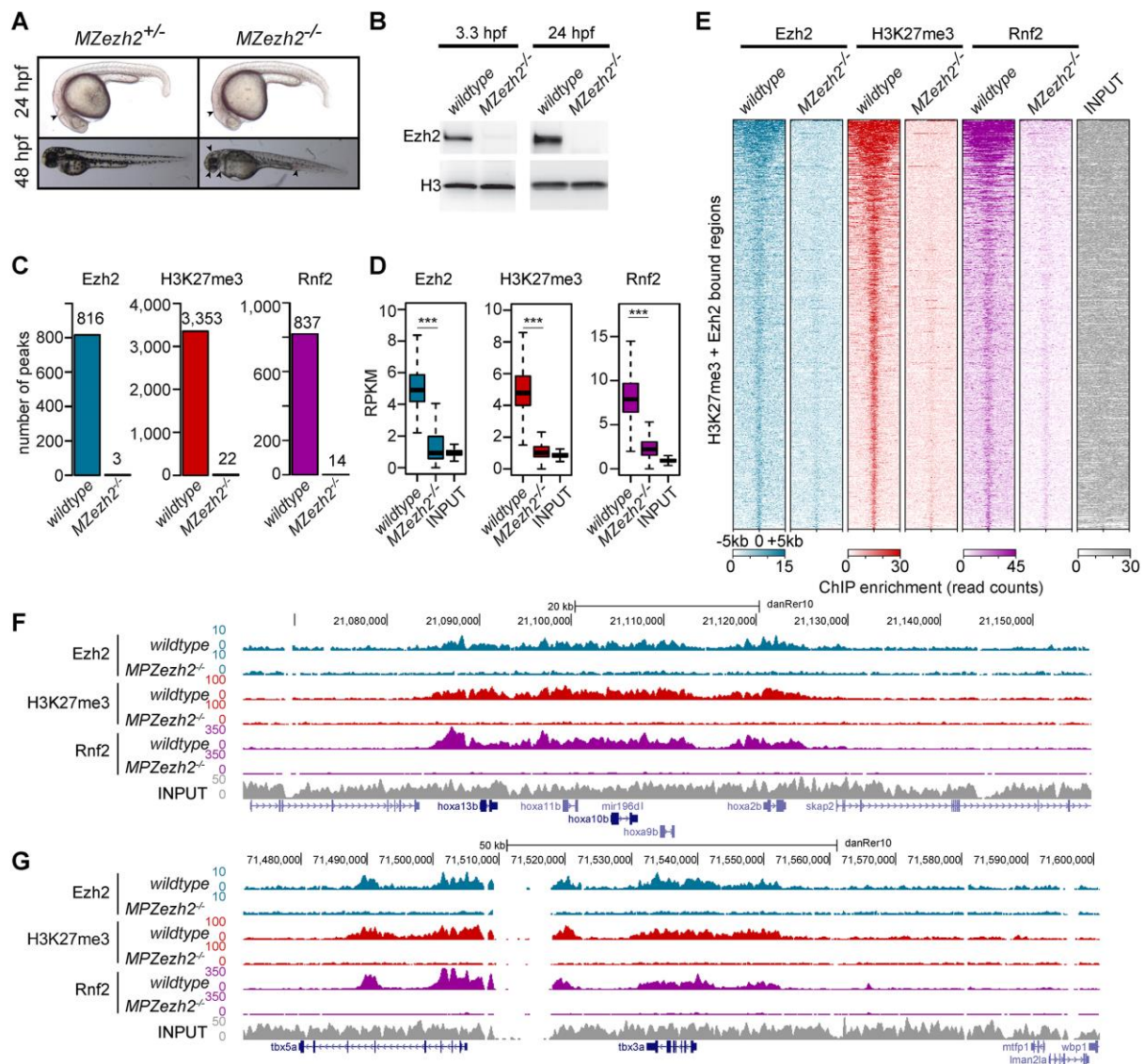
Figure 1. *MZezh2* **mutant (***MZezh2^-/-^***) embryos lack Ezh2, H3K27me3, and Rnf2 binding to the chromatin.** (**A**) *MZezh2^+/-^* (developing as wildtype embryos) and *MZezh2^-/-^* embryos at 24 and 48 hpf. At 24 hpf, *MZezh2^-/-^* embryos lack a clear mid-hindbrain boundary compared to heterozygous embryos (arrow head). At 48 hpf, *MZezh2^-/-^* embryos showed pleiotropic phenotypes compared to heterozygous embryos, such as small eyes, small brain, heart edema, and blood accumulation in the blood island (arrow heads). (**B**) Western blot analysis of Ezh2 at 3.3 hpf and 24 hpf of wildtype and *MZezh2^-/-^* embryos. Histone H3 was used as a loading control. Results presented are representative of three biological replicates. (**C**) Number of peaks called after Ezh2, H3K27me3, and Rnf2 ChIP-seq of wildtype and *MZezh2^-/-^* embryos at 24 hpf. Each peak set was obtained by the intersection of two independent biological replicates. (**D**) Box plots of Ezh2, H3K27me3, and Rnf2 RPKM-normalized

coverage after respective ChIP-seq in wildtype and in *MZezh2^-/-* embryos at 24 hpf. The input control was obtained from wildtype embryos at 24 hpf. Coverages were calculated based on peaks detected in wildtype embryos. t-test: *** *P*-value < 0.001. (**E**) Heatmaps for Ezh2, H3K27me3, and Rnf2 subsampled counts after ChIP-seq in 24 hpf wildtype and *MZezh2^-/-* embryos. Heatmap was ordered based on coverage intensity in Ezh2 and H3K23me3 ChIP-seq performed in wildtypes. Windows of 10 kb regions for all H3K27me3 o Ezh2 peaks in 24 hpf wildtype embryos are shown. The input track obtained from 24 hpf wildtype embryos was used as control and was not subsampled. (**F, G**) UCSC genome browser snapshot depicting the loss of Ezh2, H3K27me3, and Rnf2 after ChIP-seq in 24 hpf *MZezh2^-/-* embryos compared to wildtype embryos for (**F**) the *hoxab* gene cluster and (**G**) the *tbx5a* gene. Colors represent ChIP-seq for different proteins with blue: Ezh2, red: H3K27me3, purple: Rnf2, and grey: Input control.
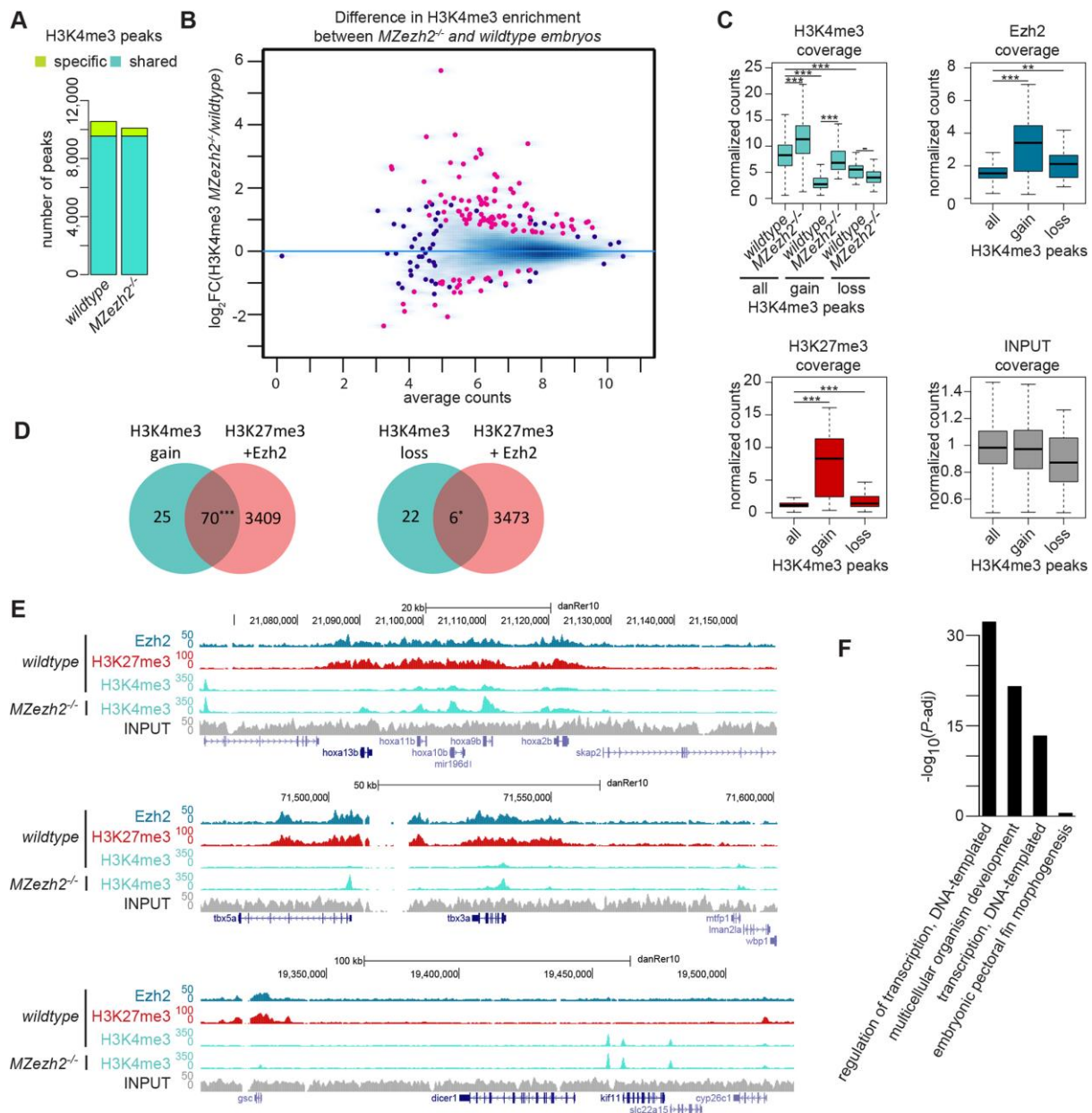
Figure 2. *MZezh2* mutant (*MZezh2⁻/⁻*) embryos show an increase in H3K4me3 preferentially on H3K27me3 targets. (**A**) Number of peaks called after H3K4me3 ChIP-seq in wildtype and *MZezh2* mutant (*MZezh2⁻/⁻*) embryos at 24 hpf. Turquoise and green represent peaks shared by the two conditions and peaks specific for one condition, respectively. Each peak set was obtained by the intersection of three independent biological replicates. (**B**) MA-plot showing the fold change (log$_2$-transformed) in H3K4me3 peak coverages in 24 hpf *MZezh2⁻/⁻* and wildtype embryos as a function of the normalized average count between the two conditions (log$_{10}$-transformed) as calculated with DiffBind on the union of H3K4me3 peaks detected in both wildtype and *MZezh2* mutant conditions. Red: log$_2$FC ≥ 1 or ≤-1 and *P*-adj < 0.05, blue: *P*-adj ≥ 0.05. When dot concentration is too high, dots are replaced by density for better visualization. (**C**) Box plots of subsampled counts after ChIP-seq for

H3K4me3 in wildtype and *MZezh2^{-/-}* embryos and for Ezh2 and H3K27me3 in wildtype embryos at 24 hpf. Box-plots display union of all H3K4me3 peaks detected in *MZezh2^{-/-}* or wildtype embryos (all) and H3K4me3 peaks enriched (gain) or decreased (loss) in *MZezh2^{-/-}* embryos compared with wildtypes as detected by DiffBind. Coverages are average of normalized counts between the triplicates for H3K4me3 and duplicates for Ezh2 and H3K27me3. The input track obtained from 24 hpf wildtype embryos was used as control. One-way ANOVA with post-tests: *** P-value < 0.001, ** P-value < 0.01. (**D**) Venn diagrams presenting the overlap between peaks with increased or decreased H3K4me3 levels (gain or loss) as detected by DiffBind with the presence of Ezh2 or H3K27me3 peaks within a +/-1 kb window. X^2: *** P-value < 0.001, ** P-value < 0.01, * P-value < 0.05. (**E**) UCSC browser snapshots of three genomic loci in wildtype and *MZezh2^{-/-}* embryos at 24 hpf. In C and E, blue, red, turquoise, and grey represent ChIP-seq for Ezh2, H3K27me3, H3K4me3, and Input control, respectively. (**F**) Gene Ontology analysis of the closest genes restricted two 2kb upstream or downstream from H3K4me3 peaks enriched in *MZezh2^{-/-}*.
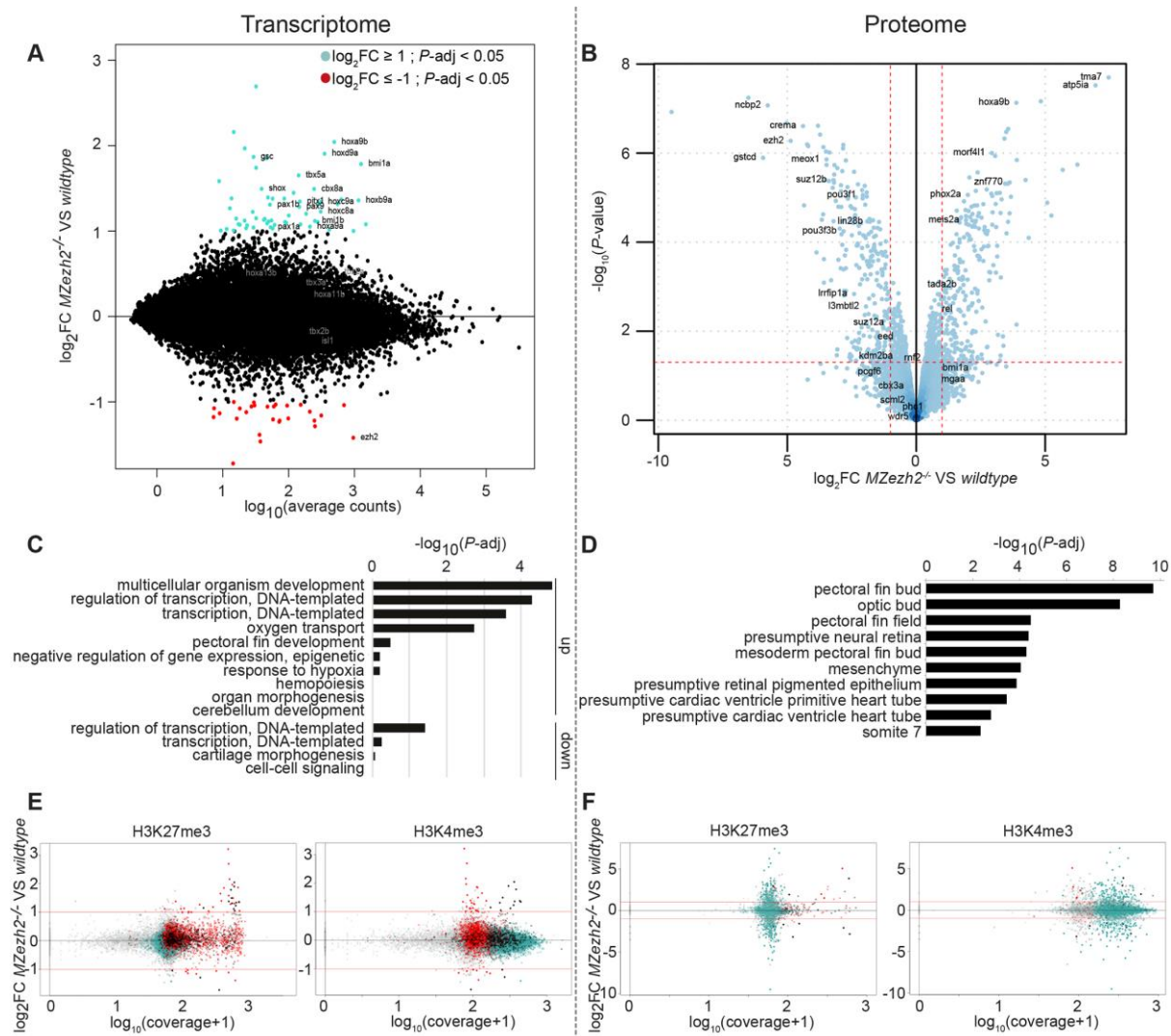
Figure 3. Loss of maternal zygotic *ezh2* results in overexpression of specific developmental genes.
(**A**) MA-plot showing the fold change (log$_2$-transformed) between gene expression in 24 hpf *MZezh2* mutant (*MZezh2$^{-/-}$*) and wildtype embryos as a function of the normalized average count between the two conditions (log$_{10}$-transformed) as calculated with DEseq2. Turquoise: log$_2$FC ≥ 1 and *P*-adj < 0.05, red: log$_2$FC ≤ -1 and *P*-adj < 0.05. For wildtype and *MZezh2$^{-/-}$* embryos, 6 and 7 biological replicates were used, respectively. (**B**) Volcano plot showing the *P*-value (-log$_{10}$-transformed) as a function of the fold-change (log$_2$-transformed) between protein expression level in *MZezh2$^{-/-}$* compared to wildtype embryos at 24 hpf. Data were obtained from biological triplicates for each condition. (**C**) Gene Ontology of biological processes associated with genes upregulated (up) or downregulated (down) in *MZezh2$^{-/-}$* embryos compared to wildtype embryos at 24 hpf. (**D**) Analysis of anatomical terms associated with proteins upregulated and downregulated in *MZezh2$^{-/-}$* embryos compared to wildtype embryos at 24 hpf. (**E-F**) Dot plot showing the fold change (log$_2$-transformed) between gene expression in 24 hpf *MZezh2$^{-/-}$* and wildtype embryos detected by RNA-seq (**E**) or proteome analysis

(**F**) as a function of the H3K27me3 (left panel) or H3K4me3 (right panel) coverage ($\log_{10}$(coverage+1) transformed). Red, turquoise, black, and gray dots represent genes associated with MACS2 detected peaks for H3K27me3, H3K4me3, both marks, or none.
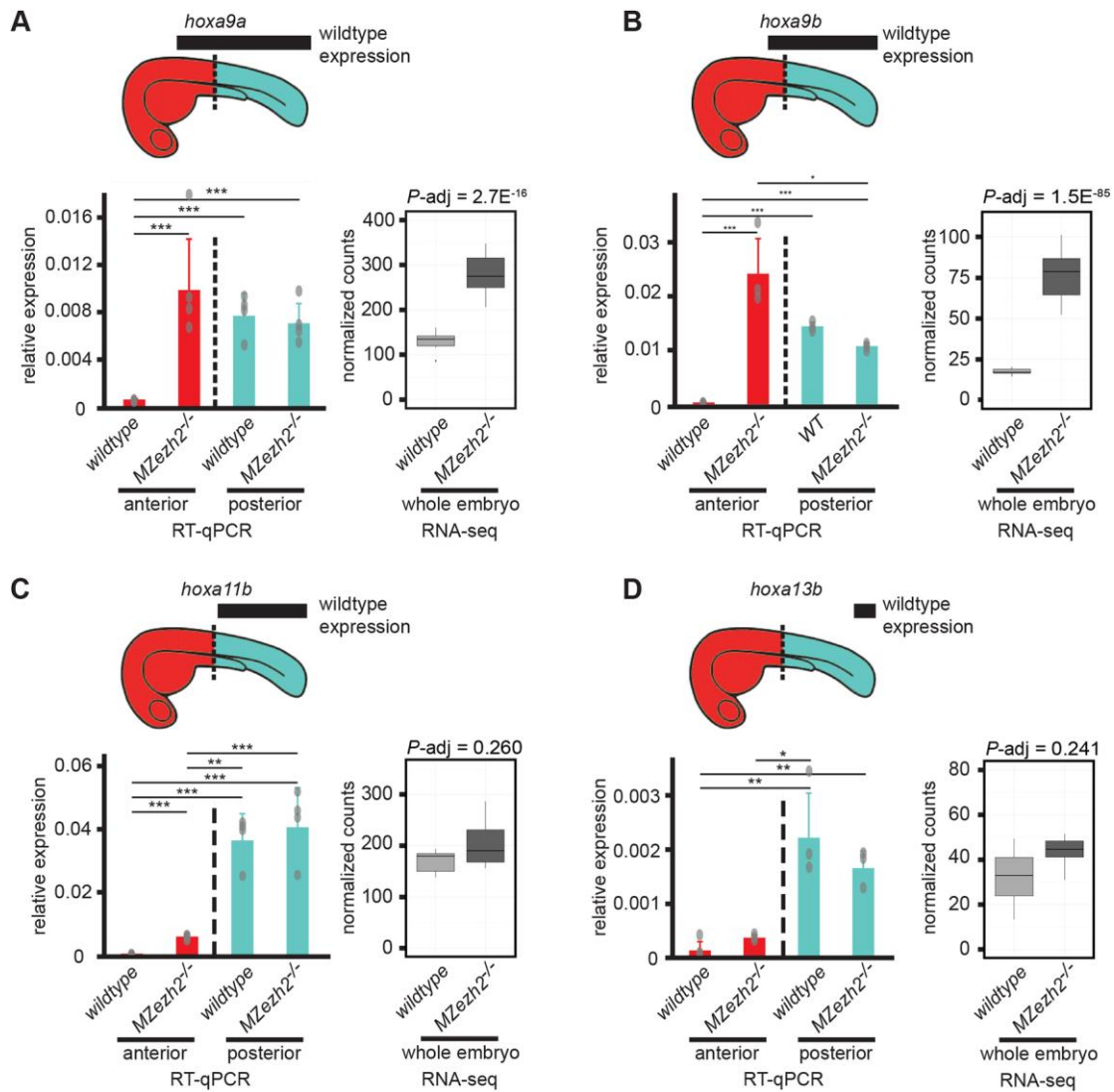
Figure 4. Loss of maternal and zygotic *ezh2* results in ectopic expression of *hox* genes. (**A, B, C, D**) Expression analysis of (**A**) *hoxa9a,* (**B**) *hoxa9b*, (**C**) *hoxa11b,* and (**D**) *hoxa13b* at 24 hpf. Bar plots on the left side of each panel represent relative expression of indicated *hox* genes in the anterior half (red) and posterior half (turquoise) of wildtype and *MZezh2* mutant (*MZezh2^-/-^*) embryos. Boxplots represent normalized counts from RNA-seq experiments in *MZezh2^-/-^* and wildtype whole embryo lysates at 24 hpf. Above is a schematic representation of 1 dpf embryos. Black boxes represent the expression domains of the *hox* genes in wildtype embryos based on published data (Thisse, 2004). Dashed lines represent the demarcation between anterior (red) and posterior (turquoise) parts of the embryo used for RT-qPCR analysis. Each experiment was performed at least in triplicates of 20 pooled anterior or posterior larval halves. For RT-qPCR, relative expression was calculated based on expression of housekeeping gene *actb1*. Bar plots represent mean±s.e.m. and overlaid dot plots represent individual RT-qPCR samples. Relative expression was compared between anterior or

posterior parts in *MZezh2*^{-/-} and wildtype embryos (one-way ANOVA with post-tests, *** *P*-value < 0.001, ** *P*-value < 0.01, * *P*-value < 0.05). For RNA-seq, adjusted *P*-values were extracted from Differential Expression analysis with DEseq2.
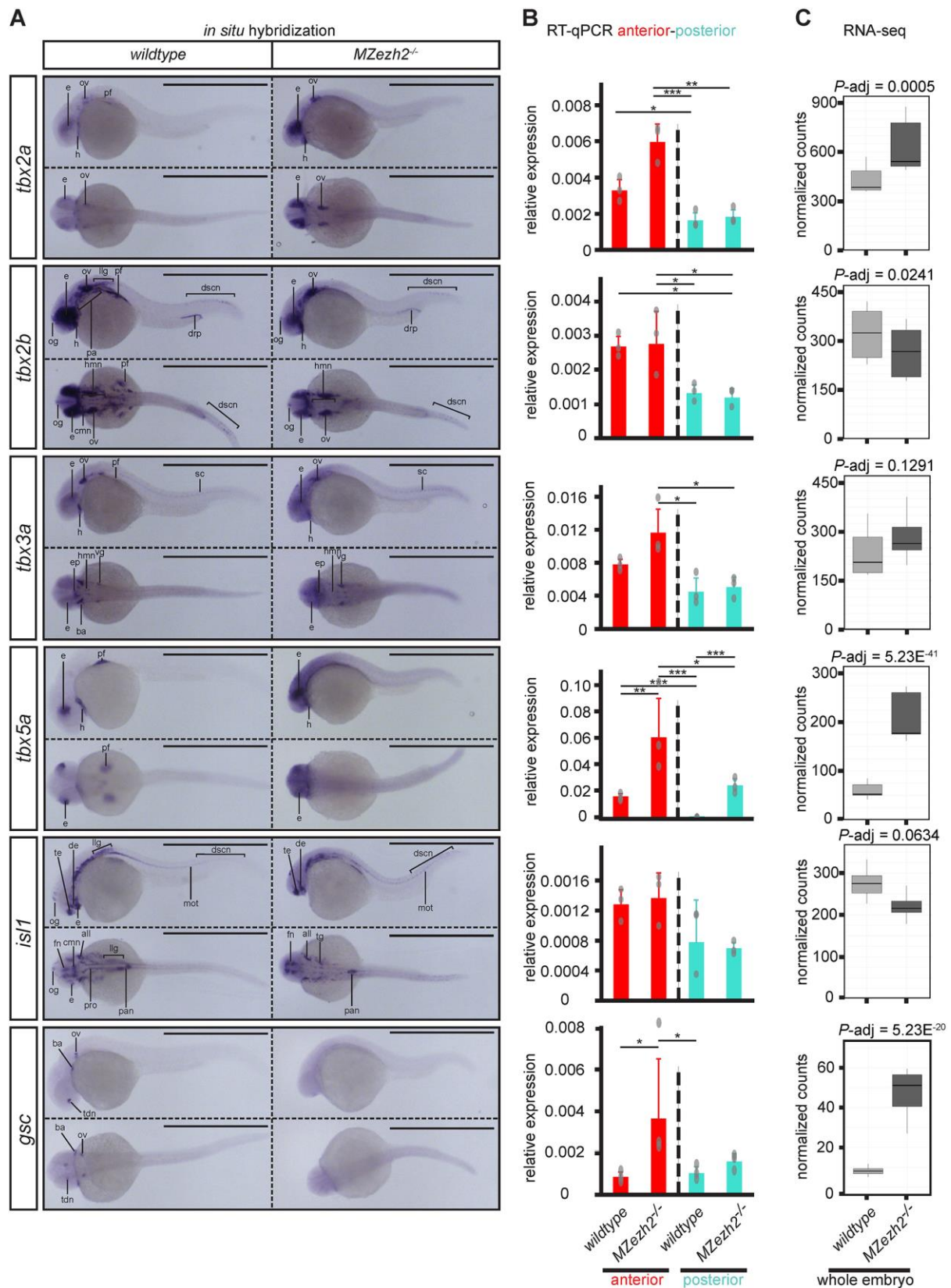
Figure 5. Transcription factor expression is spatially dysregulated in *MZezh2* mutant (*MZezh2⁻ᐟ⁻*) embryos. (**A**, **B**, **C**) Spatial expression analysis by (**A**) *in situ* hybridization (ISH), (**B**) RT-qPCR on anterior half and posterior half, and (**C**) RNA-seq results of transcription factors *tbx2a, tbx2b, tbx3a*, *tbx5a*, *isl1*,

and *gsc* in 24 hpf embryos. In ISH, scale bars represent 1 mm. Experiments were performed in biological duplicates of a least 15 pooled embryos for ISH and in triplicates or quadruplicates of 20 pooled larval halves for RT-qPCR. Relative expression was calculated based on expression of housekeeping gene *actb1*. Bar plots represent mean±s.e.m. and overlaid dot plots represent individual RT-qPCR samples. Relative expression was compared between anterior (red) or posterior (turquoise) parts in *MZezh2*$^{-/-}$ and wildtype embryos (one-way ANOVA with post-tests, *** *P*-value < 0.001, ** *P*-value < 0.01, * *P*-value < 0.05). Box plots represent normalized counts from RNA-seq experiments in whole *MZezh2*$^{-/-}$ and wildtype after Differential Expression analysis with DEseq2. all: anterior lateral lane ganglion, ba: branchial arch, cmn: cranial motor neurons, de: diencephalon, drp: distal region of the pronephros, dscn: dorsal spinal cord neurons, e: eye, ep: epiphysis, fn: forebrain nuclei, h: heart, hmn: hindbrain motor neurons, llg: lateral lane ganglion, mot: primary motor neurons, og: olfactory ganglion, ov: otic vesicle, pa: pharyngeal arches, pan: pancreas, pf: pectoral fin, pro: pronephros, sc: spinal cord, tdn: telencephalon and diencephalon nuclei, te: telencephalon, vg: ventral ganglion.
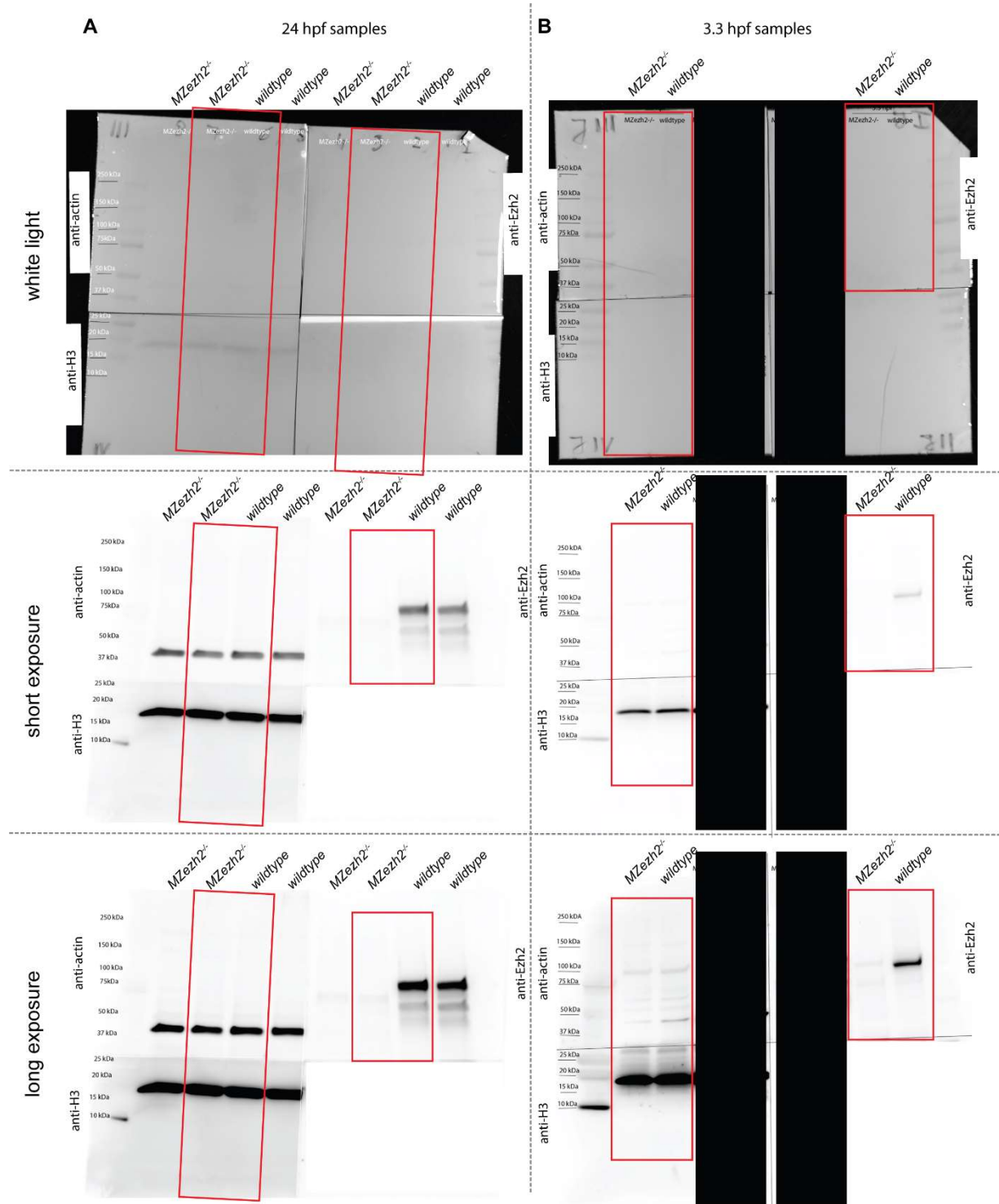
Rougeot_Supplemental_Fig.1

Fig. S1. Uncropped Western blots used for Figure 1B. (A) Images of uncropped Western blot taken for detection of Ezh2, H3K27me3, and Histone H3 in wildtype (*MZezh2+/+*) and *MZezh2* mutant (*MZezh2-/-*) embryos at 24 hpf obtained with white light illumination (top), chemical luminescence for a short exposure time (middle), and chemical luminescence for a longer exposure time

(bottom). (**B**) Images of uncropped western blot taken for detection of Ezh2, H3K27me3, and Histone H3 in wildtype (*MZezh2$^{+/+}$*) and *MZezh2* mutant (*MZezh2$^{-/-}$*) embryos at 3.3 hpf obtained with white light illumination (top), chemical luminescence for a short exposure time (middle), and chemical luminescence for a longer exposure time (bottom). Red framed lanes correspond to samples shown in Figure 1B. Black boxes cover lanes not used in this study.
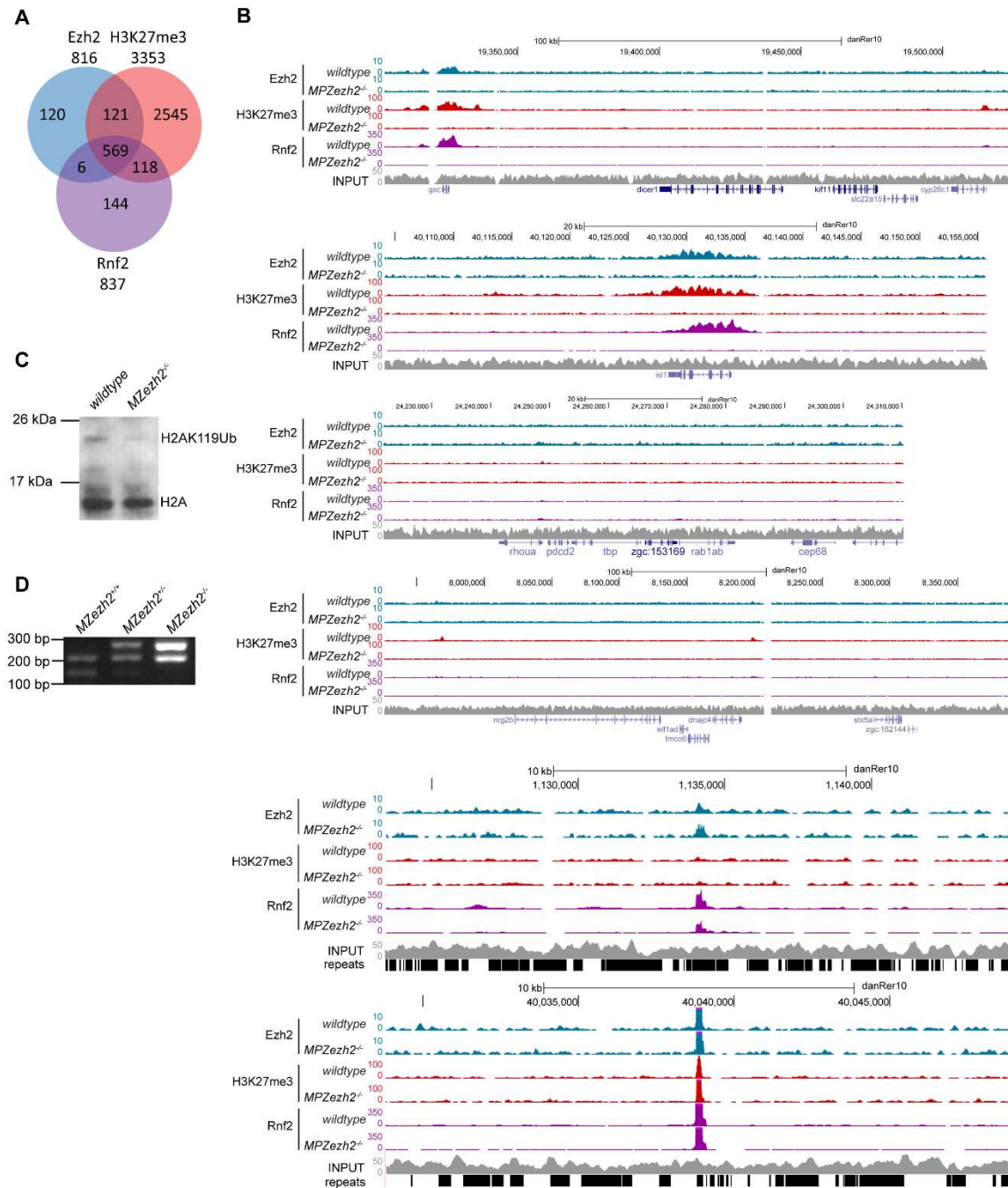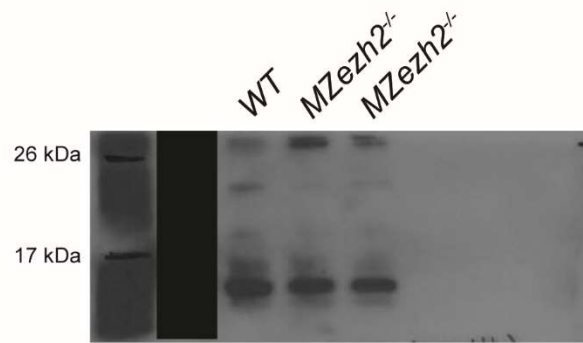
Rougeot_Supplemental_Fig.2

Fig. S2. Analysis of Ezh2, H3K27me3, and Rnf2 binding in wildtype and *MZezh2* mutant (*MZezh2-/-*) embryos at 24 hpf. (**A**) Venn diagrams presenting the overlap between Ezh2 (blue), H3K27me3 (red), and Rnf2 (purple) peaks detected in 24 hpf wildtype embryos. (**B**) UCSC browser snapshots of six

genomic loci depicting Ezh2, H3K27me3, and Rnf2 binding after ChIP-seq in *MZezh2⁻/⁻* embryos compared to wildtype embryos at 24 hpf. Colors represent ChIP-seq for different proteins with blue: Ezh2, red: H3K27me3, purple: Rnf2, and grey: Input control. (**C**) Western blot analysis of H2A on histone extracts at 24 hpf in wildtype and *MZezh2⁻/⁻* embryos. The presence of H2AK119 monoubiquitylation was visualized as a shift of the H2A band from 13 kDa to ≥20 kDa as showed by van der Velden et al. (2012). Experiment was performed in biological duplicates. (**D**) Example of *ezh2^hu5670* genotyping results after nested PCR, RsaI restriction, and gel electrophoresis in *MZezh2* wildtype (*MZezh2^{+/+}*), *MZezh2* heterozygous (*MZezh2^{+/-}*), and *MZezh2* mutant (*MZezh2^{-/-}*) embryos.
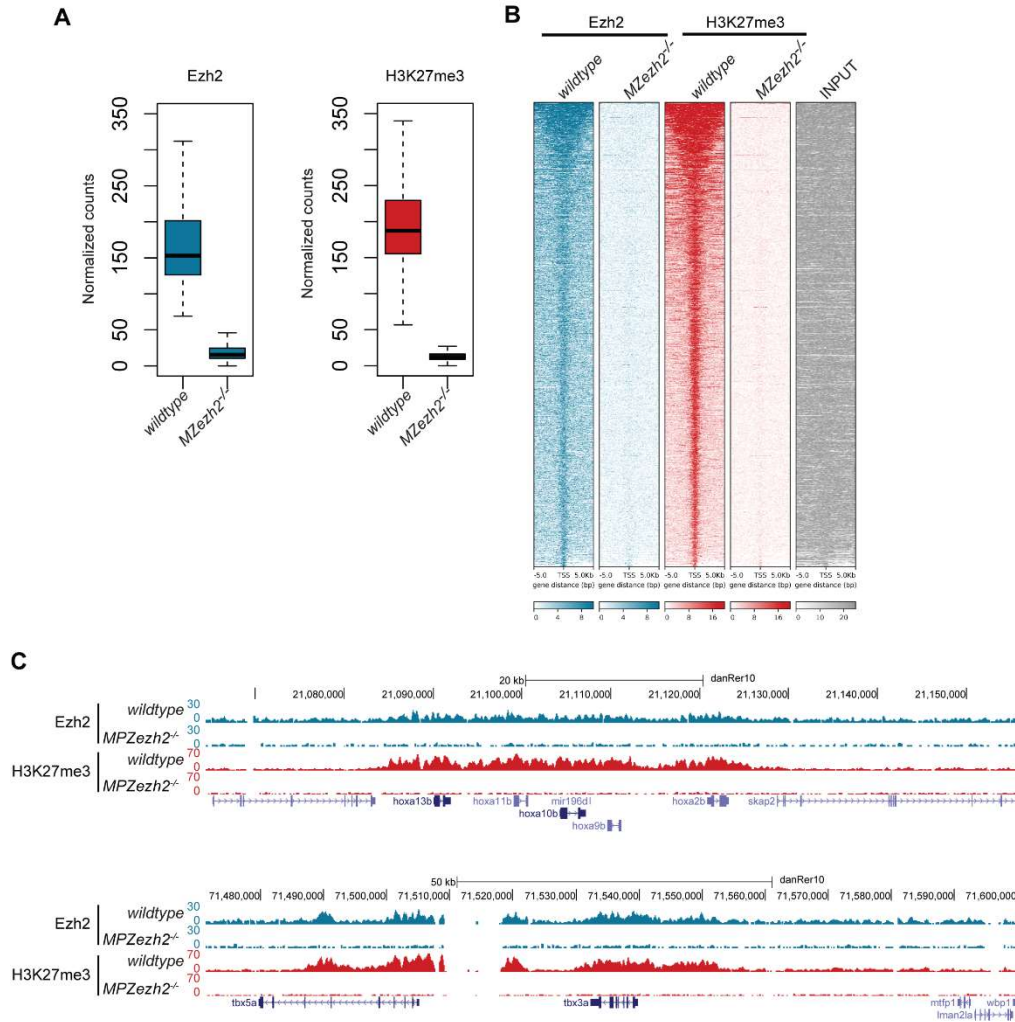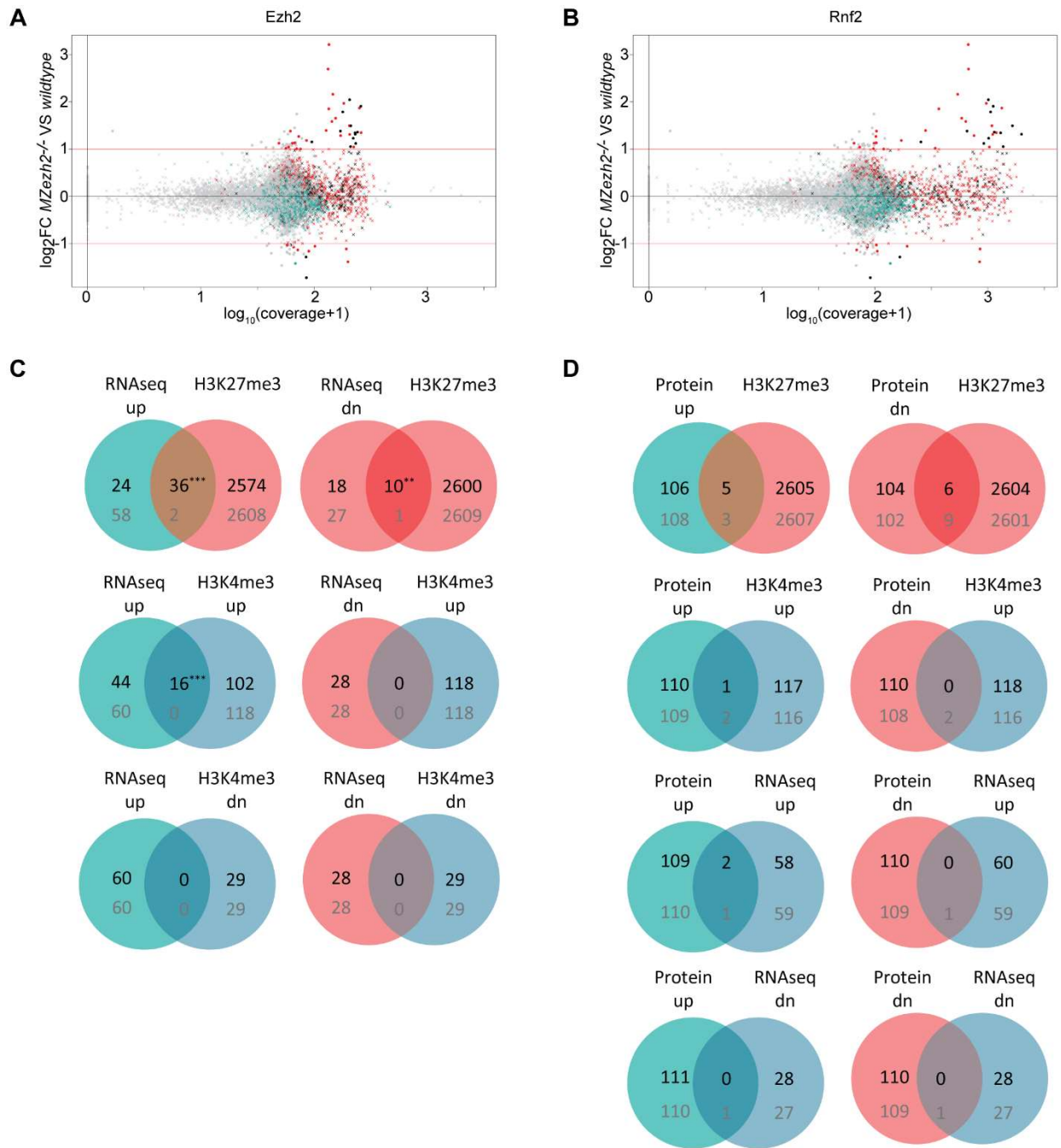
Rougeot_Supplemental_Fig.3

Fig. S3. ChIP-seq of Ezh2 and H3K27me3 using spike-in chromatin for normalization. (**A**) Box plots of Ezh2, H3K27me3, and Rnf2 coverage based on spike-in normalization after ChIP-seq in wildtype and in *MZezh2⁻/⁻* embryos at 24 hpf. Coverages were calculated based on positions of peaks detected in wildtype embryos. One replicate was performed with spike-in chromatin for each condition. (**B**) Heatmaps for Ezh2, H3K27me3, and Rnf2 counts normalized with spike-in chromatin after ChIP-seq in 24 hpf wildtype and *MZezh2⁻/⁻* embryos. Windows of 10 kb regions for all H3K27me3 or Ezh2 peaks in 24 hpf wildtype embryos are shown. The input track obtained from 24 hpf wildtype embryos was used as control and was not normalized. (**C**) UCSC genome browser snapshot depicting the loss of Ezh2 and H3K27me3 after ChIP-seq in 24 hpf *MZezh2⁻/⁻* embryos compared to wildtype embryos. Coverage were normalized with spike-in chromatin.

Rougeot_Supplemental_Fig.4

Fig. S4. Uncropped Western blot used for Figure S2C. Image of uncropped Western blot taken for detection of Histone H2A in wildtype (*WT*) and *MZezh2* mutant (*MZezh2*⁻ᐟ⁻) embryos at 24 hpf. The presence of H2AK119 monoubiquitylation was visualized as a shift of the H2A band from 13 kDa to ≥20 kDa as showed by van der Velden et al. (2012). Black box covers data not used in this study.

Rougeot_Supplemental_Fig.5

Fig. S5. Integration of ChIP-seq, RNA-seq, and proteomics data. (A) Dot plot showing the fold change (log$_2$-transformed) between gene expression in 24 hpf *MZezh2* mutant (*MZezh2$^{-/-}$*) and wildtype embryos detected by RNA-seq as a function of the Ezh2 coverage (log$_{10}$(coverage+1) transformed). **(B)** Dot plot showing the fold change (log$_2$-transformed) between gene expression in 24 hpf *MZezh2$^{-/-}$* and wildtype embryos detected by RNA-seq as a function of the Rnf2 coverage (log$_{10}$(coverage+1) transformed). In **A** and **B**, coverage was calculated on the gene region +/- 2 kb and

averaged between duplicates. (**C**) Venn diagrams presenting the overlap between genes upregulated (up) or downregulated (dn) in *MZezh2*$^{-/-}$ embryos compared to wildtype and presence of H3K27me3 or H3K4me3 peaks. The closest genes from H3K27me3 peaks in wildtype condition or H3K4me3 enriched (H3K4me3 up) and decreased (H3K4me3 dn) peaks according to DiffBind were used for this analysis. Black numbers represent comparison between actual DEseq2 identified genes and closest genes from peaks. Grey numbers represent comparisons between actual DEseq2 identified genes and random selected genes used as control. X^2: *** *P*-value < 0.001, ** *P*-value < 0.01, * *P*-value < 0.05. (**D**) Venn diagrams presenting the overlap between proteins overrepresented (Protein up) or underrepresented (Protein dn) in *MZezh2*$^{-/-}$ embryos compared to ChIP-seq and RNA-seq results. The closest genes from H3K27me3 peaks in wildtype condition or H3K4me3 enriched peaks according to DiffBind (H3K4me3 up) were used for this analysis. Black numbers represent comparison between actual dysregulated proteins and genes. Grey numbers represent comparisons between actual dysregulated proteins and random selected genes used as control. X^2-test did not provide any significant results.
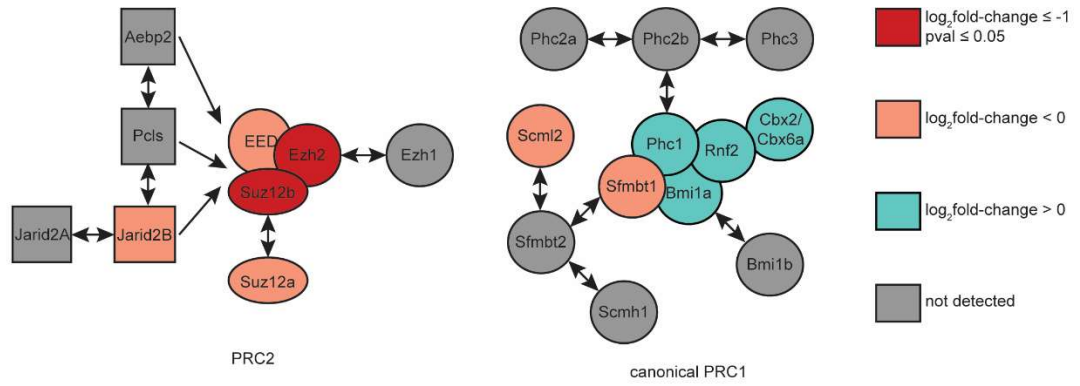
Rougeot_Supplemental_Fig.6

Fig. S6. Proteomic analysis in *MZezh2* mutant (*MZezh2^-/-^*) embryos at 24 hpf reveals downregulation of the core PRC2 components. Schematic representation of changes in protein expression level of PRC2 (left) and canonical PRC1 (right) subunits in *MZezh2^-/-^* compared to wildtype embryos at 24 hpf. Dark red: $\log_2$fold-change ≤ -1 and *P*-value ≤ 0.05, light red: $\log_2$fold-change < 0, turquoise: $\log_2$fold-change ≥ 0, grey: protein not detected.
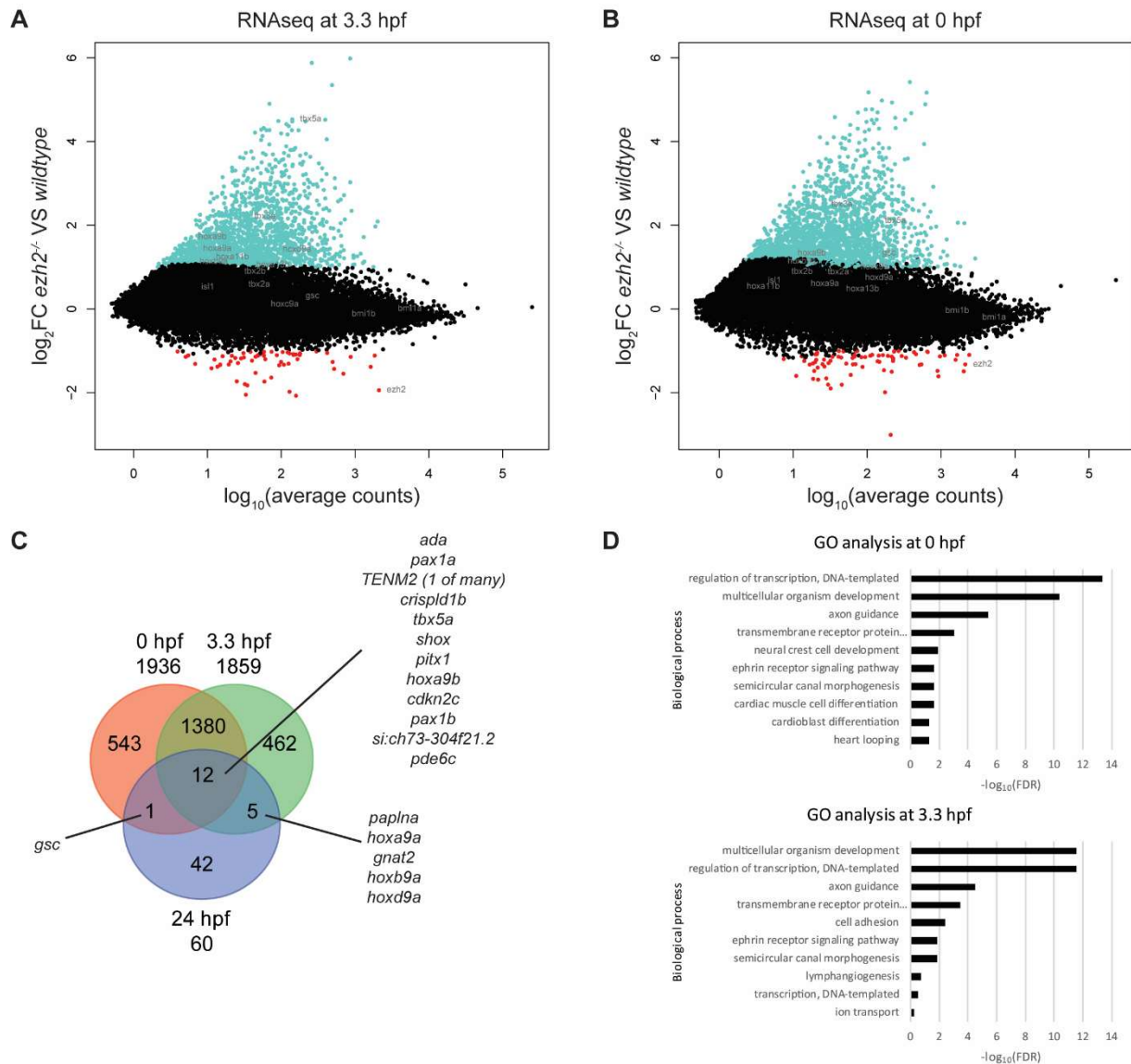
Rougeot_Supplemental_Fig.7

Fig. S7. Transcriptome analysis of *MZezh2* mutant at 3.3 and 0hpf. (**A**) MA-plot showing the fold change (log$_2$-transformed) between gene expression in 3.3 hpf *MZezh2* mutant (*MZezh2$^{-/-}$*) and wildtype embryos as a function of the normalized average count between the two conditions (log$_{10}$-transformed) as calculated with DEseq2. Turquoise: log$_2$FC ≥ 1 and *P*-adj < 0.05, red: log$_2$FC ≤ -1 and *P*-adj < 0.05. Experiments were performed in at least 5 biological replicates. (**B**) MA-plot showing the fold change (log$_2$-transformed) between gene expression in 0 hpf *MZezh2$^{-/-}$* and wildtype embryos as a function of the normalized average count between the two conditions (log$_{10}$-transformed) as calculated with DEseq2. Turquoise: log$_2$FC ≥ 1 and *P*-adj < 0.05, red: log$_2$FC ≤ -1 and *P*-adj < 0.05. (**C**) Venn diagram comparing genes overexpressed in *MZezh2$^{-/-}$* compared with wildtype embryos at 3 different time points. (**D**) Gene Ontology of biological processes associated with genes upregulated in *MZezh2$^{-/-}$* embryos compared to wildtype embryos at 3.3 and 0 hpf.
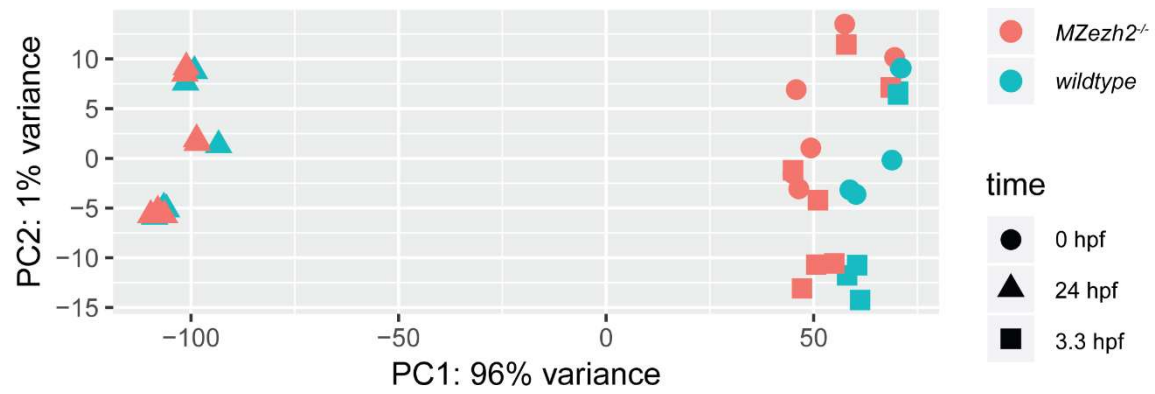
Rougeot_Supplemental_Fig.8

Fig. S8. Principal Component Analysis (PCA) for all RNA-sequencing samples generated for this study. Principal Component (PC) 1 explains most of the variation by clearly separating 24 hpf samples from 0 and 3.3 hpf samples.
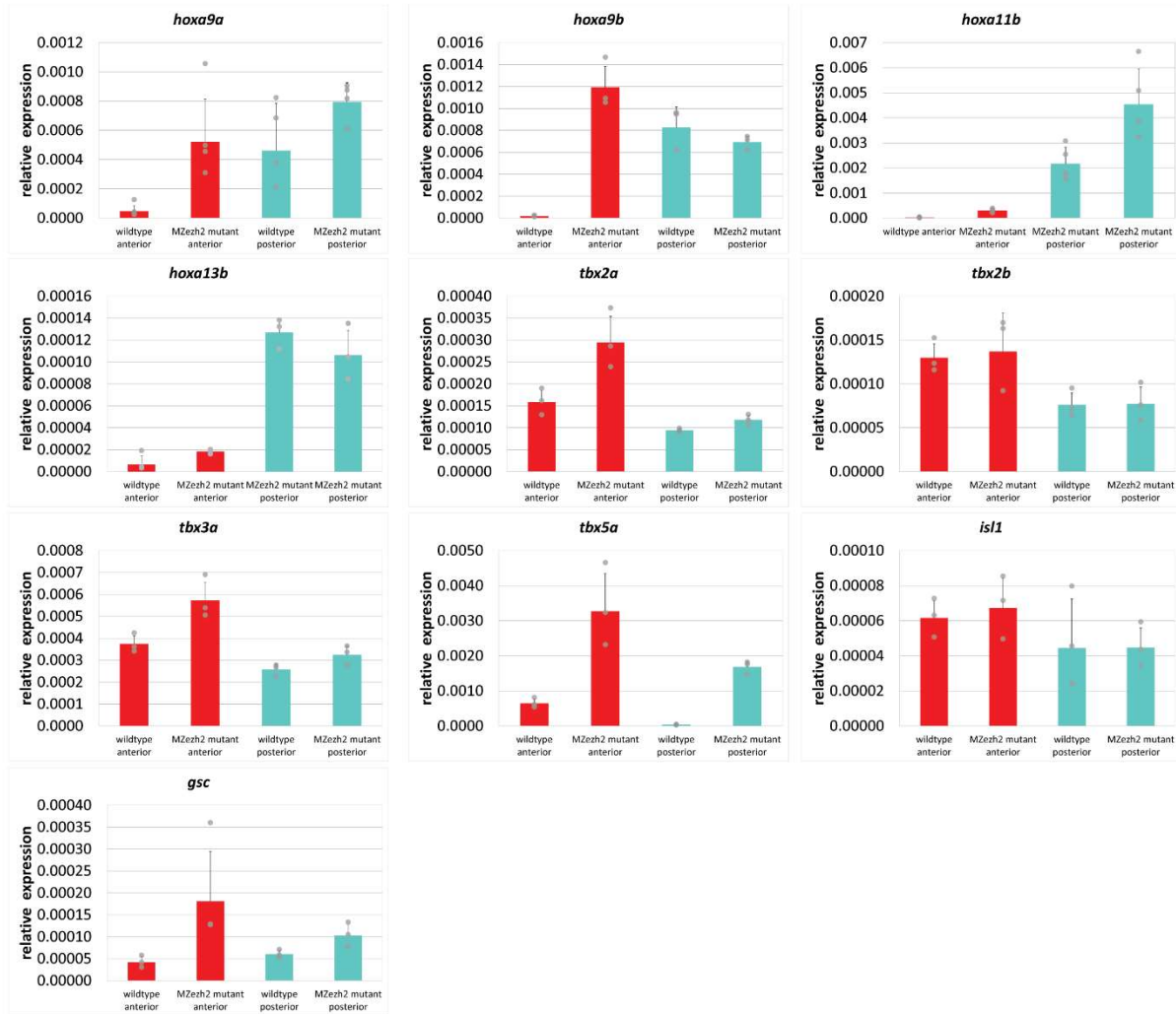
Rougeot_Supplemental_Fig.9

Fig. S9. Conclusions from RT-qPCR results of Figures 4 and 5 are similar when using alternative reference gene *eif1a*. Expression analysis of genes in Figures 4 and 5 when normalized against the reference gene *eif1a* instead of *actb1*. Bar plots represent relative expression of indicated genes in the anterior half (red) and posterior half (turquoise) of wildtype and *MZezh2* mutant (*MZezh2⁻ᐟ⁻*) embryos at 24 hpf. Bar plots represent mean±s.e.m. and experiments were performed with at least biological triplicates. Dot plots overlaid on bar plots represent results for individual RT-qPCR samples.

Table S1. Overview of RNA-seq, ChIP-seq, and proteomics results per gene

Click here to Download Table S1

Table S2. List of primers used in this study

name	sequence	experiment	target
p3_hu5670_ComFw	CAGAATCGGTTTCCAGGTTGCCG	genotyping	*ezh2* genomic PCR
p4_hu5670_ComRv	CAGTACTCTGAGATGAACTCATTC	genotyping	*ezh2* genomic PCR
LK_ezh2_exon_Fw	TGTAAAACGACGGCCAGTCAGAATCGGTTTCCAGGTTGCCG	genotyping	*ezh2* genomic nested PCR
LK_ezh2_exon_Rv	AGGAAACAGCTATGACCATTGCAGGAGACGTTTTTACTGTCCC	genotyping	*ezh2* genomic nested PCR
Hoxa9a_RTqPCR_Fw	AAGCAGAATCTAGCCGAACTG	RT-qPCR	*hoxa9a*
Hoxa9a_RTqPCR_Rv	CACAGGGTTTTCTGGATCAGC	RT-qPCR	*hoxa9a*
Hoxa9b_RTqPCR_Fw	CAACGGATCACATGATGAGAAAAT	RT-qPCR	*hoxa9b*
Hoxa9b_RTqPCR_Rv	CCAGTTGGACGAAGGGTTA	RT-qPCR	*hoxa9b*
Hoxa11b_RTqPCR_Fw	AGCAGCAATGGACAAAAGACAC	RT-qPCR	*hoxa11b*
Hoxa11b_RTqPCR_Rv	AAGAAAAATTCTCTCTCCAGCTCT	RT-qPCR	*hoxa11b*
Hoxa13b_RTqPCR_Fw	GTGTACTGCCCGAAAGATCA	RT-qPCR	*hoxa13b*
Hoxa13b_RTqPCR_Rv	ACCTGACACGGTATCTTGGA	RT-qPCR	*hoxa13b*
tbx2a_RTqPCR_Fw	GCTAAGGAGCTTTGGGATCA	RT-qPCR	*tbx2a*
tbx2a_RTqPCR_Rv	CACCTTGAACGGAGGAAACA	RT-qPCR	*tbx2a*
tbx2b_RTqPCR_Fw	TCTCAACACATGCTTGCCTC	RT-qPCR	*tbx2b*
tbx2b_RTqPCR_Rv	AAAAGTCCACCGAAGGTTGG	RT-qPCR	*tbx2b*
tbx3a_RTqPCR_Fw	CCCGATGCCGTTTCATCTG	RT-qPCR	*tbx3a*
tbx3a_RTqPCR_Rv	CCGAAAGGAGACATAGCCAG	RT-qPCR	*tbx3a*
tbx5a_RTqPCR_Fw	GGGAGCTGATACGAGCTTTT	RT-qPCR	*tbx5a*
tbx5a_RTqPCR_Rv	CGTGAGGCCTTAAATTCCGA	RT-qPCR	*tbx5a*
isl1_RTqPCR_Fw	TTTACAAATGGCAGCAGAGC	RT-qPCR	*isl1*
isl1_RTqPCR_Rv	CGGGTTGTTTTCTCAGGTTG	RT-qPCR	*isl1*
gsc_RTqPCR_Fw	CAACAGTGTCCGTGTATTCCT	RT-qPCR	*gsc*
gsc_RTqPCR_Rv	TCATTTGATGTGGGACTGGAG	RT-qPCR	*gsc*

Table S3. List of antibodies used in this study

antibody	brand	ref	Concentration µg/µl	ChIP (µl/IP)	WB (dilution)
anti-Ezh2	Cell Signaling	5246S	N/A	2	1:1,000
anti-Rnf2	Cell Signaling	5694S	N/A	4	N/A
anti-H3K27me3	Millipore	07-449	N/A	2	N/A
anti-H3K4me3	Millipore	04-745	N/A	2	N/A
anti-H2A	Millipore	07-146	N/A	N/A	1:1,000
anti-Histone H3	Sigma-Aldrich	H0164	N/A	N/A	1:2,000
HRP-conjugated anti-Rabit	Dako	P0217	N/A	N/A	1:3,000

Table S4. Statistics for all high throughput samples generated for this study

experiment	target	genotype	time	replicate	library type	M Seqs	%aligned	M Aligned	remarks
ChIP-seq	Ezh2	*MZezh2NULL*	24 hpf	1	paired-end	1.0	35.0%	0.4	
ChIP-seq	Ezh2	*MZezh2NULL*	24 hpf	2	paired-end	33.6	51.9%	17.5	
ChIP-seq	Ezh2	*wildtype*	24 hpf	1	paired-end	23.2	71.8%	16.7	
ChIP-seq	Ezh2	*wildtype*	24 hpf	2	paired-end	34.8	56.3%	19.6	
ChIP-seq	H3K27me3	*MZezh2NULL*	24 hpf	1	paired-end	2.9	60.3%	1.8	
ChIP-seq	H3K27me3	*MZezh2NULL*	24 hpf	2	paired-end	34.7	51.0%	17.7	
ChIP-seq	H3K27me3	*wildtype*	24 hpf	1	paired-end	23.0	79.6%	18.3	
ChIP-seq	H3K27me3	*wildtype*	24 hpf	2	paired-end	31.4	68.8%	21.6	
ChIP-seq	H3K4me3	*MZezh2NULL*	24 hpf	1	paired-end	24.3	78.8%	19.2	
ChIP-seq	H3K4me3	*MZezh2NULL*	24 hpf	2	paired-end	55.6	34.6%	19.3	
ChIP-seq	H3K4me3	*MZezh2NULL*	24 hpf	3	paired-end	24.9	77.1%	19.2	
ChIP-seq	H3K4me3	*wildtype*	24 hpf	1	paired-end	40.5	76.2%	30.9	
ChIP-seq	H3K4me3	*wildtype*	24 hpf	2	paired-end	82.4	15.0%	12.4	73.9% $(TA)_n$ contamination
ChIP-seq	H3K4me3	*wildtype*	24 hpf	3	paired-end	25.8	75.4%	19.5	
ChIP-seq	INPUT	*wildtype*	24 hpf	1	paired-end	83.4	72.7%	60.6	
ChIP-seq	Rnf2	*MZezh2NULL*	24 hpf	1	paired-end	33.4	47.5%	15.9	
ChIP-seq	Rnf2	*MZezh2NULL*	24 hpf	2	paired-end	20.5	74.6%	15.3	
ChIP-seq	Rnf2	*wildtype*	24 hpf	1	paired-end	53.3	68.3%	36.4	
ChIP-seq	Rnf2	*wildtype*	24 hpf	2	paired-end	18.9	74.9%	14.2	
RNA-seq		*MZezh2NULL*	0hpf	1	single-end	32.1	83.1%	26.7	
RNA-seq		*MZezh2NULL*	0hpf	2	single-end	30.9	81.8%	25.3	
RNA-seq		*MZezh2NULL*	0hpf	3	paired-end	14.4	55.5%	8.0	
RNA-seq		*MZezh2NULL*	0hpf	4	paired-end	15.7	57.5%	9.0	
RNA-seq		*MZezh2NULL*	0hpf	5	paired-end	20.1	55.3%	10.7	
RNA-seq		*MZezh2NULL*	0hpf	6	paired-end	21.7	55.7%	12.1	
RNA-seq		*wildtype*	0hpf	1	single-end	31.2	81.4%	25.4	
RNA-seq		*wildtype*	0hpf	2	single-end	32.0	86.6%	27.7	
RNA-seq		*wildtype*	0hpf	3	paired-end	25.9	23.2%	6.0	
RNA-seq		*wildtype*	0hpf	4	paired-end	21.5	68.7%	14.8	
RNA-seq		*wildtype*	0hpf	5	paired-end	23.1	67.1%	15.5	
RNA-seq		*MZezh2NULL*	24hpf	1	single-end	32.0	84.6%	27.1	
RNA-seq		*MZezh2NULL*	24hpf	2	single-end	30.1	74.1%	22.3	
RNA-seq		*MZezh2NULL*	24hpf	3	paired-end	8.8	60.4%	5.3	
RNA-seq		*MZezh2NULL*	24hpf	4	paired-end	15.7	75.8%	11.9	
RNA-seq		*MZezh2NULL*	24hpf	5	paired-end	15.9	73.6%	11.7	
RNA-seq		*MZezh2NULL*	24hpf	6	paired-end	24.5	73.1%	17.9	
RNA-seq		*MZezh2NULL*	24hpf	7	paired-end	16.2	68.0%	11.0	
RNA-seq		*wildtype*	24hpf	1	single-end	31.0	84.5%	26.2	
RNA-seq		*wildtype*	24hpf	2	single-end	36.7	81.5%	29.9	
RNA-seq		*wildtype*	24hpf	3	paired-end	17.5	65.8%	11.5	
RNA-seq		*wildtype*	24hpf	4	paired-end	20.0	62.9%	12.6	
RNA-seq		*wildtype*	24hpf	5	paired-end	18.1	72.3%	13.1	
RNA-seq		*wildtype*	24hpf	6	paired-end	15.3	70.7%	10.8	
RNA-seq		*MZezh2NULL*	3hpf	1	single-end	33.8	79.1%	26.7	
RNA-seq		*MZezh2NULL*	3hpf	2	single-end	28.2	84.3%	23.8	
RNA-seq		*MZezh2NULL*	3hpf	3	paired-end	25.0	58.0%	14.5	
RNA-seq		*MZezh2NULL*	3hpf	4	paired-end	23.6	72.4%	17.1	
RNA-seq		*MZezh2NULL*	3hpf	5	paired-end	20.6	53.9%	11.1	
RNA-seq		*MZezh2NULL*	3hpf	6	paired-end	20.1	51.8%	10.4	
RNA-seq		*MZezh2NULL*	3hpf	7	paired-end	28.4	52.4%	14.9	
RNA-seq		*wildtype*	3hpf	1	single-end	32.9	77.5%	25.5	
RNA-seq		*wildtype*	3hpf	2	single-end	30.5	83.0%	25.3	
RNA-seq		*wildtype*	3hpf	3	paired-end	18.9	69.4%	13.1	
RNA-seq		*wildtype*	3hpf	4	paired-end	18.3	53.5%	9.8	
RNA-seq		*wildtype*	3hpf	5	paired-end	16.4	57.3%	9.4	

M Seqs = number of sequenced reads in Million   M Aligned = number of aligned reads in Million