

Pre-marking chromatin with H3K4 methylation is required for accurate zygotic genome activation and development

Received: 31 January 2025

Accepted: 5 December 2025

Published online: 19 December 2025

 Check for updates

Meghana S. Oak  , Marco Stock  , Ana Janeva¹, Matthias Mezes¹, Antony M. Hynes-Allen⁵, Tobias Straub  , Ignasi Forné⁷, Andreas Ettinger  , Stephan Hamperl  , Axel Imhof^{7,8}, Jelle van den Ameele  , Antonio Scialdone^{1,2,3} & Eva Hörmanseder  

In vertebrate embryos, gene expression is first initiated at zygotic genome activation (ZGA). Maternally expressed transcription factors are essential for this process. However, it is unknown whether active chromatin modifications established in gametes are present in early embryos and contribute to ZGA and embryonic development. Here, we show that in *Xenopus laevis*, H3K4me3 occurs at common genomic loci in gametes, in transcriptionally quiescent pre-ZGA embryos, and in transcriptionally active ZGA embryos. These loci exhibit high H3K4me3 intensities and breadth, DNA hypomethylation, and elevated CpG content. We show that H3K4 methylation pre-marking is required for successful ZGA and development, including expression of the key ZGA transcription factor Pou5f3.2. We demonstrate that the H3K4 methyltransferase Cxxc1 ensures establishment of H3K4me3 and proper ZGA. These findings reveal a role for H3K4 methylation in defining active chromatin states in *Xenopus laevis* embryos and highlight its importance for accurate ZGA and embryonic development.

Cell fates are established through the activity of developmental cues, transcription factors (TFs) and epigenetic mechanisms that either permit or repress transcriptional programs. In vertebrates, the development of zygotes is supported by maternal transcripts and TFs until zygotic genome activation (ZGA), when gene expression patterns are established. The extent to which this depends on TFs or epigenetic mechanisms is not fully understood. It is unclear how parental chromatin modifications contribute to embryonic development via epigenetic memory.

Epigenetic memory refers to the maintenance of chromatin states across cell divisions independently of the developmental signal that induced the state¹. Most studies of epigenetic inheritance have confirmed the transmission of repressive histone modifications around transcriptionally inactive chromatin. However, the role of active histone modifications, traditionally associated with actively transcribed genes, in epigenetic memory is under debate². The maintenance of active chromatin states is primarily dependent on TF-coupled gene

¹Institute of Epigenetics and Stem Cells (IES), Helmholtz Zentrum München, München, Germany. ²Institute of Functional Epigenetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany. ³Institute of Computational Biology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany. ⁴TUM School of Life Sciences Weihenstephan, Technical University of Munich, Munich, Germany. ⁵Department of Clinical Neurosciences and MRC Mitochondrial Biology Unit, University of Cambridge, Cambridge, UK. ⁶Core Facility Bioinformatics, Biomedical Center Munich (BMC), Faculty of Medicine, Ludwig-Maximilians-Universität in Munich, Martinsried, Germany. ⁷Protein Analysis Unit, Biomedical Center (BMC), Faculty of Medicine, Ludwig-Maximilians-University (LMU) Munich, Martinsried, Germany. ⁸Department of Molecular Biology, Biomedical Center Munich, Ludwig-Maximilians University, Planegg-Martinsried, Germany. ⁹Research Institute of Molecular Pathology (IMP), Vienna Bio-Center, Vienna, Austria. ¹⁰These authors contributed equally: Meghana S. Oak, Marco Stock.  e-mail: meghana.oak@helmholtz-munich.de; eva.hoermanseder@imp.ac.at

expression³. Propagation of active histone marks is thought to be dependent on TF binding and continuous gene transcription, and not via direct transmission using autonomously operating read-and-write mechanisms. Therefore, the field does not unequivocally consider active histone marks as an epigenetic mechanism. Moreover, the tight link between active chromatin states and transcription complicates efforts to discern whether active histone marks are causal or consequential, and whether they reflect epigenetic mechanisms^{4,5}.

Embryos of species such as *Xenopus laevis* and zebrafish undergo a transcription-free window of at least 8 rapid cell divisions from fertilization to ZGA, challenging chromatin-modification propagation⁶. In zebrafish, the active chromatin mark H3K4me3 is present before ZGA at developmental gene promoters, and placeholder mechanisms driven by other chromatin modifications, H2A.Z and DNAm, were shown to set the stage for the first transcriptional program^{7,8}. In *Xenopus*, technical limitations hamper histone mark analyses of this dynamic, early embryonic chromatin, and thus, the presence of H3K4me3 is questioned⁹. Instead, it has been suggested that H3K4me3 is rapidly acquired de novo later around ZGA^{9,11}. This would exclude a role of H3K4me3 in transmitting active chromatin states from gametes to embryos, as well as overlook the contribution of such a mechanism to embryonic development.

Previous experiments revealed that H3K4me3 decorates specific genomic regions in *Xenopus* sperm nuclei and in blastula embryos prior to ZGA¹². Thus, chromatin marks could be present on the dynamic chromatin of pre-ZGA embryos. Cell-fate reprogramming studies using nuclear transfer to *Xenopus* eggs suggest that H3K4me3 may play a role in maintaining active transcriptional states^{13,14}. Demethylation of H3K4 in somatic donor nuclei rendered them more permissive to reprogramming, and transcriptional memory of active gene expression states was lost¹³. These experiments indicate that nuclear transfer embryos retain information encoded by H3K4me3 via epigenetic memory. It raises the question of whether this mechanism also propagates information from gametes to embryos during physiological development. Together, these observations support the involvement of H3K4me3 in maintaining active transcriptional states between fertilization and ZGA, but evidence is currently missing. It has not been tested whether maintenance of active chromatin states marked by H3K4 methylation is important for embryonic development.

Here, we made use of the transcriptionally silent pre-ZGA developmental period in *Xenopus laevis* embryos^{15–17} to investigate the dynamics of the active histone modification H3K4me3 and its role in embryonic development. We profiled histone modifications on chromatin in pre-ZGA embryos using mass spectrometry. We detected H3K4me3 together with other histone modifications associated with active chromatin on pre-ZGA chromatin, despite multiple rapid cell divisions in the absence of transcription. We traced the H3K4me3 mark on a large group of genes that are transcribed in the parental gametes, as well as the ZGA embryo, separated by transcriptional dormancy in cleavage-stage embryos. We show that H3K4me3 peaks are present around these genes at these selected time points prior to and during ZGA stages of development. Promoters of these genes are furthermore characterized by high CpG density and DNA accessibility, as well as DNA hypomethylation. Knockdown of the transcription-independent H3K4 methyltransferase Cxxc1 results in a reduction of H3K4me3 levels, improper gene expression of zygotic genes, and embryonic defects. H3K4 demethylation specifically during pre-ZGA cleavages shows defects of ZGA as well as increased embryonic defects. This suggests that H3K4me3 marking of chromatin during early embryonic cell cycles is essential for subsequent faithful ZGA. Together, this study demonstrates that H3K4 methylation is required for regulating embryonic transcription and development. The presence of H3K4me3 at common loci in gametes, prior to ZGA and during ZGA, is consistent with a potential role in epigenetic memory, suggesting that this modification contributes to information about chromatin states during early embryogenesis.

Results

Histone modifications linked with active chromatin are present in early embryos prior to ZGA

After fertilization, the *Xenopus* zygote divides rapidly 8 times before gene transcription is fully activated at ZGA. During these cleavages, the genome remains globally transcriptionally silent due to histone-mediated repression and inhibitory chromatin hindering TF binding until mid-blastula transition^{15–17}. We first tested whether chromatin marks normally associated with transcriptionally active chromatin states are present on this fast-replicating chromatin in the absence of gene transcription.

We established a method enriching chromatin proteins from yolk-rich embryos for chromatin profiling experiments. Using mass spectrometry analysis of histone tail post-translational modifications (see “methods”, Fig. 1a and Fig. S1a), we detected a wide range of active and repressive histone modifications on pre-ZGA embryo chromatin (256-cell stage). These included the active marks H3K4me1/2/3, H3K9ac and H3K27ac, but not H3K36me3, as well as the repressive marks H3K9me3 and H3K27me3 and the replication-associated histone modification H4K20me1 (Fig. 1b). As expected, we also observed these histone marks on transcriptionally active mid-ZGA embryo chromatin (4000-cell stage; Fig. 1b). Together, this reveals that highly dynamic chromatin of rapidly dividing pre-ZGA embryos is decorated by both repressive and active chromatin marks.

Western Blotting confirmed these results. We detected H3K4me3-modified chromatin in pre-ZGA embryos from as early as the 32-cell stage in development (Fig. 1c), and previous studies detected H3K4me3 in whole protein extracts from *Xenopus laevis* 1-cell zygotes and 4-cell embryos, at levels comparable to those observed in 256-cell embryos¹⁸. H3K4me3 levels relative to total H3 on chromatin remained constant between 32-cell and 256-cell stages, followed by a pronounced increase by mid-ZGA, consistent with our mass spectrometry data (Fig. S1b). Despite global transcriptional quiescence in embryos before ZGA, the active histone marks H3K4me3, H3K4me1, H3K9ac and H3K27ac were detectable in 256-cell stage embryos, as well as in post-ZGA stage embryos (Fig. 1d). Conversely, we did not detect the transcription-associated H3K36me3 and H3K79me3 modifications in pre-ZGA chromatin. These results support that active histone marks H3K4me1/3, H3K9ac and H3K27ac are present on chromatin prior to ZGA.

In *Xenopus*, as well as in *Drosophila*, zebrafish and mouse, a defined cluster of microRNAs and β-catenin target genes is transcribed before the major wave of ZGA^{19–25}. We hypothesized that this early transcription may reinforce H3K4 promoter methylation through transcription-coupled mechanisms^{4,26} mediated by RNA Polymerase II and III, as RNA Polymerase I is absent in pre-ZGA embryos^{15,27}. We tested whether transcription inhibition via the RNA Polymerase II and III inhibitor α-amanitin results in acute depletion of H3K4me3 in embryos during ZGA (Fig. S1c). At such high doses, α-amanitin is shown to not only inhibit elongation, but also destabilize the RPB1 subunit, causing rapid RNA Polymerase II loss^{28,29}. We co-injected the uridine analog 5-EU, which is incorporated into nascent RNAs³⁰ to reveal effective inhibition of transcription (Fig. S1d). We observed that both H3K4me3 and H4 are detected in transcription-inhibited, 5-EU negative embryos and transcriptionally active, 5-EU positive embryos at developmental stages corresponding to mid-ZGA (Fig. 1e), in line with previous reports¹¹. We noted that H3K4me3 levels are only reduced by 26% in α-amanitin-injected embryos (Fig. 1e), indicating that while a subset of H3K4me3 is transcription-dependent, the majority of H3K4me3 marks on mid-ZGA embryo chromatin is independent of RNA Pol II and III-mediated transcription.

This demonstrates that the dynamic chromatin of rapidly dividing, transcription-inhibited *Xenopus laevis* embryos is decorated by H3K4me3, together with several other histone modifications. It also indicates that during developmental stages preceding ZGA, the presence of active chromatin modifications may be supported by mechanisms operating independently of feedback from ongoing gene transcription.

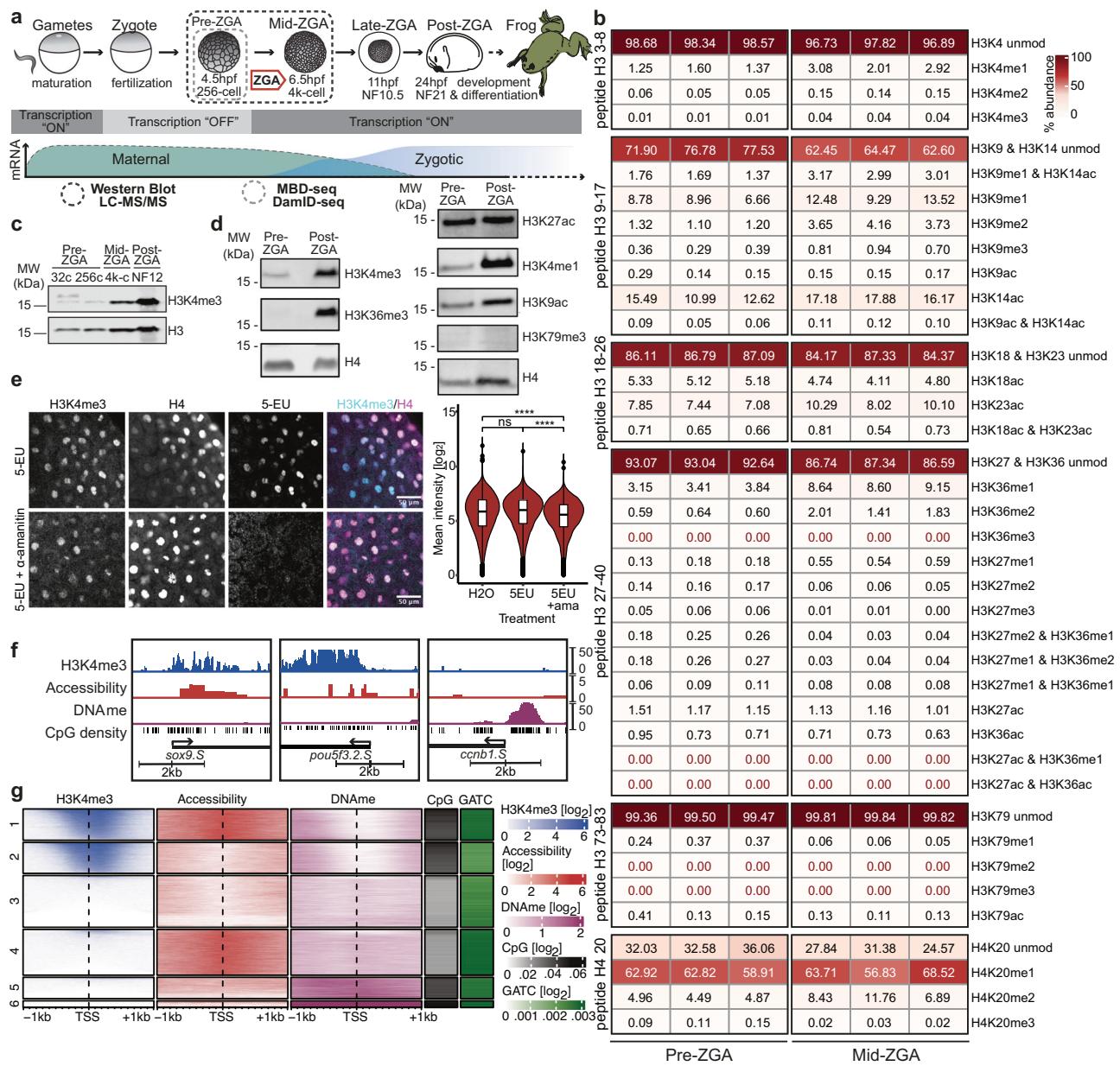


Fig. 1 | Profiling histone modifications and epigenetic factors present on pre-ZGA embryo chromatin. **a** A schematic highlighting early developmental stages of interest in this study and their transcriptional states with a focus on ZGA and the methods used in this figure. **b** Heatmap representing the percent abundance of histone modifications on each respective peptide relative to the total abundance of the corresponding peptide, obtained by PTM analysis of LC-MS/MS. Undetectable modifications are highlighted in red text. Each column represents a single biological replicate of pre-ZGA or mid-ZGA embryos. **c** Time-series Western blot showing the occurrence of H3K4me3 and H3 in embryos. **d** Western blots of histone modifications in pre-ZGA and post-ZGA embryos using chromatin-bound protein isolation. Western blots were independently repeated 2–3 times with consistent results. **e** Representative images of H3K4me3 levels in control and transcription-inhibited mid-ZGA embryos. Nascent transcripts are labeled using 5-EU. Overlay of H3K4me3 (cyan) and H4 (magenta) is represented. Scale bars: 50 μ m. Nuclei are segmented using H4, and

H3K4me3 values are measured. (Right) Quantification of H3K4me3 immunofluorescence in H4-segmented nuclei across Z-stacks ($N=3$ embryos per condition, $n=2$ independent experiments). Box plots show the median (center line), 25–75 percentiles (box), whiskers extending to 1.5 \times the interquartile range, and outliers as individual points. Statistical test: one-sided Wilcoxon rank-sum test with the alternative hypothesis that the values from the amaranth-injected condition are lower than controls; p : (****) ≤ 0.0001 ; n.s. are $p > 0.05$. Exact p : H₂O vs 5EU: 0.2; H₂O vs 5EU + Ama: 2.2×10^{-16} ; 5EU + Ama vs Ama: 2.2×10^{-16} . **f** IGV tracks displaying H3K4me3 ChIP-seq, CATA_{Da}, and MBD-seq data for pre-ZGA embryos with CpG regions at representative promoters (TSS \pm 1kb). **g** Heatmap showing H3K4me3, accessibility, DNAme, CpG and GATC density across promoters (TSS \pm 1kb) in pre-ZGA embryos. Accessibility enrichment is normalized to GATC frequency within the promoter region. Clusters are identified by k-means clustering on the z-scored data, excluding GATC density ($k=6$; Hartigan-Wong algorithm).

H3K4me3-marked promoters in pre-ZGA embryos are in distinct chromatin configurations

After establishing that pre-ZGA chromatin is marked with active histone modifications, we addressed the genomic context of H3K4me3. In transcriptionally active cells, H3K4me3 is enriched at promoters of expressed genes and associated with DNA hypomethylation, high CpG

density and accessible chromatin^{31–35}. We tested whether these associations also apply to *Xenopus laevis* embryos prior to ZGA.

We profiled H3K4me3, DNA methylation (DNAme) and chromatin accessibility in pre-ZGA embryos using ChIP-seq, MBD-seq and CATA_{Da} (Chromatin Accessibility profiling using Targeted Dam-ID), respectively (Fig. 1a, f)^{36,37} and calculated their occurrence at promoters

(1 kb \pm TSS), normalizing accessibility to local GATC motif density (Fig. S1e). In pre-ZGA embryos, DNAm was depleted at ~95% of H3K4me3-marked promoters. Similarly, a prior study in *Xenopus tropicalis* similarly found that at mid-ZGA, ~90% H3K4me3-marked promoters were depleted of DNAm³⁸ (Fig. S1e). Promoters with high H3K4me3 levels and DNA hypomethylation also showed high CpG densities (Fig. 1g clusters 1 and 2). Furthermore, approximately a third of the H3K4me3-marked promoters were in an accessible configuration, DNA hypomethylated and CpG-rich (Fig. 1g cluster 1), resembling promoters of actively transcribed genes. The remaining H3K4me3-marked promoters were less accessible, but DNA hypomethylated and CpG-rich (Fig. 1g cluster 2). Promoters lacking H3K4me3 were in three different configurations: DNA methylated, CpG-rich with low accessibility, suggesting a repressive chromatin state (Fig. 1g clusters 5 and 6), DNA hypomethylated, CpG-poor and relatively accessible (Fig. 1g cluster 4), or DNA hypomethylated, CpG-poor and less accessible (Fig. 1g cluster 3).

This shows that in pre-ZGA *Xenopus laevis* embryos, characterized by widespread transcriptional quiescence, H3K4me3-marked promoters exhibit distinct chromatin configurations considering CpG content, DNAm and chromatin accessibility.

Genes with high H3K4me3 intensity and breadth are shared across gametes and early embryos

We then investigated the origins and fate of promoter-associated H3K4me3 identified in pre-ZGA embryos by comparing its genomic landscape from the last transcriptionally active stage in gametes to the next transcriptionally active stage at ZGA (Fig. 2a). For this, we utilized published H3K4me3 ChIP-seq datasets for spermatid, sperm, pre- and post-ZGA embryonic stages^{12,39,40}.

We found that H3K4me3 peaks predominantly associated with promoters in the gametes and early embryos (Fig. S2a). Of all identified H3K4me3 promoter peaks, 75% were shared across all four time points and made up 39% of all promoters (“SHARED” group; Fig. 2b, c). This suggests that H3K4me3 is established during the spermatid stage and subsequently present at the promoters of these genes until ZGA. In contrast, for 48% of all promoters, H3K4me3 peaks were not detected at any of the four time points (“ABSENT” group, Fig. 2c, d). Only the remaining 13% of promoters displayed dynamic H3K4me3 peak changes across stages (Figs. 2b and S2b) and were subclassified by pattern (Fig. S2c). We refer to genes with a promoter H3K4me3 peak in spermatids that is absent at any subsequent stage as the “LOST” group (Figs. 2b and S2d), and those acquiring a peak after the spermatid stage as the “GAINED” group (Figs. 2b and S2d). These results suggest that most promoters marked by H3K4me3 in pre-ZGA embryos are already marked in the male gametes and also in ZGA stage embryos, although whether this reflects a continuous transmission remains unclear.

Next, we characterized the genes within these H3K4me3 dynamics groups. Gene ontology (GO) analyses revealed an enrichment of terms associated with housekeeping functions in the SHARED group, whereas terms related to embryo development, gastrulation and pattern specification were enriched in the GAINED group, and to a lesser extent in the SHARED group (Fig. S2e). No significant enrichment for GO-terms was found in the LOST group. Thus, in embryos, promoter H3K4me3 present before ZGA may support housekeeping functions, whereas acquiring H3K4me3 at ZGA may support developmental and differentiation.

We then evaluated H3K4me3 peak characteristics of the H3K4me3 dynamics groups. We found that the SHARED group showed significantly higher H3K4me3 peak intensities and breadths around the promoter regions than all other groups across all analyzed time points (Figs. 2e and S2f). Relative H3K4me3 promoter enrichment and breadth in the spermatid indicated how long H3K4me3 remains detectable during subsequent developmental stages: promoters of the SHARED group had the highest H3K4me3 enrichment and breadth

scores in spermatid (Fig. S2g), followed by promoters of the LOST subgroup with progressively lower scores. At ZGA, promoters that gained H3K4me3 peaks at pre-ZGA show greater H3K4me3 enrichment and breadth compared to peaks gained later at the ZGA stage (Fig. S2h). Consistent with reports of H3K4me3 extending into the coding region of actively transcribed genes in somatic cells⁴¹, we observed an asymmetrical extension of the peak into the gene body for GAINED genes during ZGA onset (Fig. S2f). These analyses revealed that H3K4me3 peak intensities and breadth closely correlate with how long H3K4me3 marking is detected in early development.

Given the connections between CpG content, DNA hypomethylation and H3K4me3 deposition in embryonic stem cells (ESCs)⁴², we explored how these features relate to H3K4me3 dynamics in *Xenopus*. We found that promoters of the SHARED group had significantly increased DNA accessibility (Fig. 2f), lower DNAm (Fig. 2h), and higher CpG content (Fig. 2i) than all other groups in pre-ZGA embryos. Interestingly, promoters gaining H3K4me3 during ZGA had lower chromatin accessibility (Fig. 2f), CpG content (Fig. 2h), and higher DNAm levels (Fig. 2g) than SHARED gene promoters in pre-ZGA embryos. This indicates that H3K4me3 presence around CpG-rich and accessible promoters during early development anticorrelates with DNAm before ZGA (Fig. 2i).

Together, we observe that H3K4me3 domains around promoters displaying high intensities and breadth are present at the same genomic loci in the last transcriptionally active stages in gametes, across transcriptionally dormant pre-ZGA stages, and in the next transcriptionally active ZGA stage in embryos.

H3K4me3 marks genes transcribed in gametes and at ZGA during the interjacent silent window

We next asked whether H3K4me3 promoter enrichment across a window of transcriptional silence contributes to inheriting memory of past active transcriptional states. We first hypothesized that promoters marked by H3K4me3 in gametes and embryos were once actively transcribed in the gametes and then again during ZGA.

We used datasets of nascent RNA-seq spanning pre- and mid-ZGA stages (5–9 hpf), and total RNA-seq datasets from egg and spermatid, to characterize gene expression groups based on their transcriptional dynamics^{39,40,43}. Genes whose transcripts were detected in either maternal or paternal gametes, and in embryos during these ZGA time points, were defined as “Gamete+Zygotic” genes (GZ-genes, Fig. 3a, b, c). Additionally, we grouped genes exclusively expressed in the maternal and paternal gametes (“Gamete-Specific”, GS-genes), and genes expressed during ZGA and not detected in either gamete (“Zygote-Specific”, ZS-genes; Fig. 3a, b, c). Genes with undetectable transcripts at all specified time points were named “Not-Detected” (ND) genes. GO analyses revealed that “embryo development” was associated with both GZ- and ZS-genes (Fig. S3a). In agreement with our analysis of H3K4me3 dynamics, terms related to housekeeping functions were primarily linked to GZ-genes, while terms associated with gastrulation, morphogenesis and symmetry formation were specifically enriched in ZS-genes (Fig. S3b). Together, this suggests that GZ-genes may propagate their active transcriptional state from the gametes to embryos and have potential housekeeping or developmental functions.

To address whether GZ gene promoters are marked by H3K4me3 across developmental stages, we compared H3K4me3 peak characteristics of GS-, ZS- and ND- genes at four time points—spermatid, sperm, pre-ZGA and ZGA (Fig. 3d). We observed that GZ genes had significantly higher H3K4me3 peak intensities and breadth than all other gene groups at all time points, notably also during the transcriptionally quiescent stages (Figs. 3d and S3c). GS-genes initially showed lower mean peak intensities than GZ-genes, dropping below those of ZS-genes during ZGA (Fig. 3d). Genes expressed only in male gamete progenitors (Spermatid-Specific or SS) displayed even lower average H3K4me3 levels than the overarching GS group (Fig. S3d).

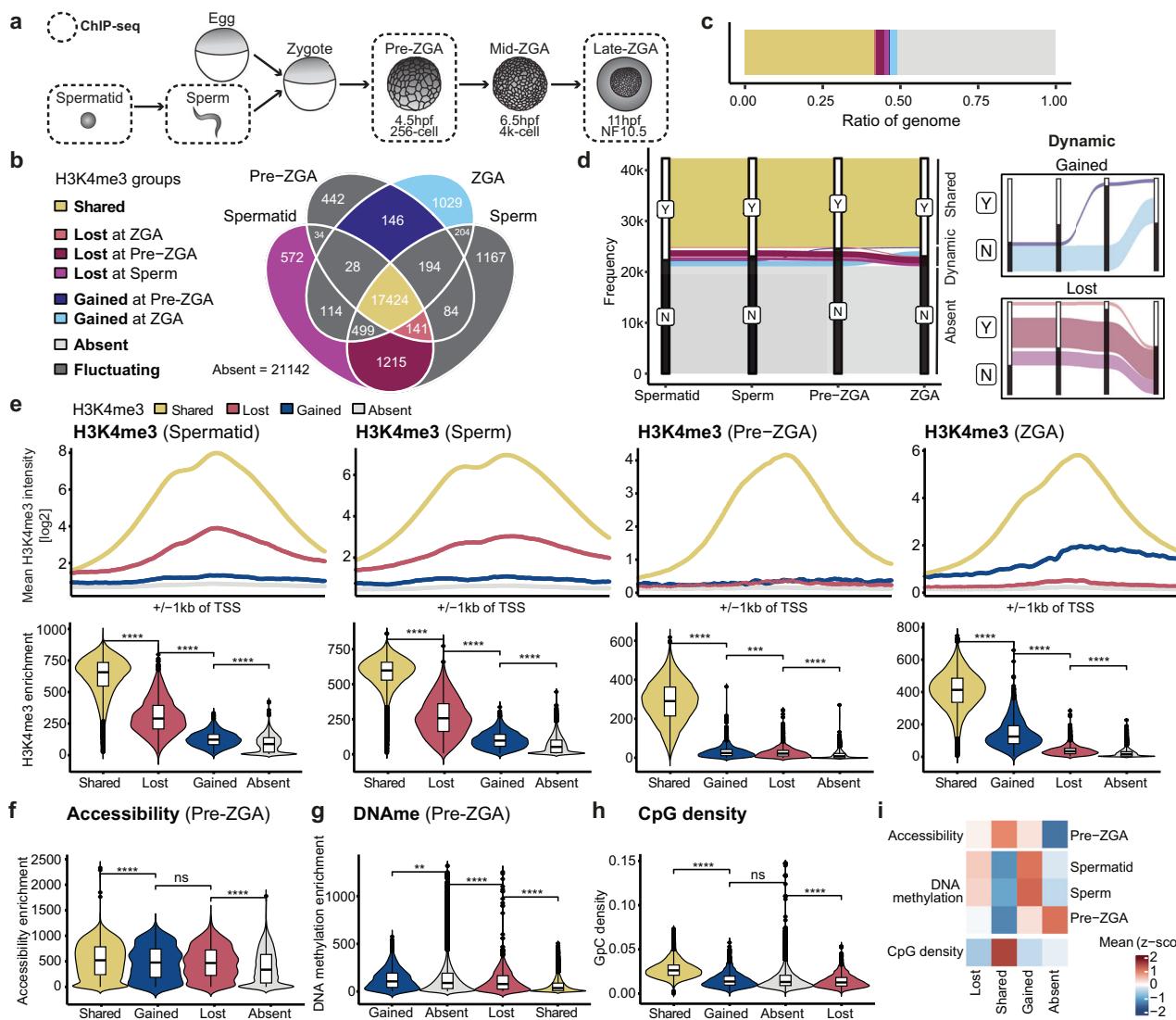


Fig. 2 | H3K4me3 peak intensity closely correlates with the dynamics of H3K4me3 localization. **a** Illustration highlighting stages of interest for H3K4me3 ChIP-seq. **b** Venn diagram showing the binarized overlap of H3K4me3 peaks at all promoter regions (TSS \pm 1kb) across spermatid, sperm, pre-ZGA and post-ZGA timepoints. Colors are based on the dynamics of H3K4me3 across the stages: Yellow: genes that “share” H3K4me3 (“SHARED”, $n = 17424$), blue: genes that “gain” H3K4me3 (“GAINED”, $n = 1175$), purple: genes that “lose” H3K4me3 peaks across timepoints (“LOST”, $n = 1928$). Promoters with fluctuating H3K4me3 dynamics are excluded from further analysis and marked in dark gray. **c** Ratio of genes that fall into each group of H3K4me3 dynamics. **d** Alluvial plots representing H3K4me3 dynamics dynamics of each group. **e** (Top) H3K4me3 around TSS (\pm 1kb) for H3K4me3 dynamics groups at each time point-spermatid, sperm, pre-ZGA, post-ZGA,

respectively. (Bottom) H3K4me3 promoter enrichment in each group for each time point. Statistical test: one-sided Wilcoxon rank-sum test with alternative hypothesis that values of left distribution tend to be larger than right distribution; p : (****) ≤ 0.0001 , (***) ≤ 0.001 , (**) ≤ 0.01 , (*) ≤ 0.05 ; n.s. are $p > 0.05$. Box plots show the median (center line), 25–75th percentiles (box), whiskers extending to 1.5× the interquartile range, and outliers as individual points. **f** Accessibility promoter enrichment around TSS (\pm 1kb) in each group for pre-ZGA stage. **g** DNAme promoter enrichment around TSS (\pm 1kb) in each group for pre-ZGA stage. **h** Comparison of promoter CpG density around TSS (\pm 1kb) in each group for pre-ZGA stage. **i** Heatmap showing the z-scored mean of accessibility, CpG density and DNAme at promoters for available timepoints for the major H3K4me3 dynamics groups.

Groups with zygotically expressed genes, i.e., GZ and ZS groups, gained asymmetrically broader peaks during ZGA, likely due to active transcription^{11,41} (Fig. S3c). Together, this indicates that, in agreement with our hypothesis, genes expressed in the gametes and embryos (GZ group) display high intensities and breadth of H3K4me3 around their promoters across the interadjacent transcriptionally silent phase.

Furthermore, H3K4me3-enriched GZ-gene promoters also showed high DNA accessibility (Fig. S3f), DNA hypomethylation (Fig. S3g) and high CpG content (Fig. S3h) at pre-ZGA compared to other groups (Fig. 3e). ZS-gene promoters showed lower chromatin accessibility (Fig. S3f), CpG content (Fig. S3h) and increased DNAme (Fig. S3g) compared to GZ-gene promoters at pre-ZGA (Fig. 3e).

Since GZ-genes, expressed in both gametes and embryos, showed similar characteristics to the SHARED group, we addressed their overlap (Fig. 3f). Indeed, over 80% of GZ-genes overlapped and were significantly associated with the SHARED group, but not with any of the other groups (Fig. 3g, h and Table S1). GS-genes were significantly associated with the LOST group, fitting with their loss of gene expression during development, while ZS-genes were associated with the GAINED group of genes, matching their de novo activation at ZGA stage (Fig. 3h). Consistently, genes characterized as both SHARED and GZ had higher levels of accessibility, CpG density and displayed low levels of DNAme also in spermatid and sperm (Fig. S3i).

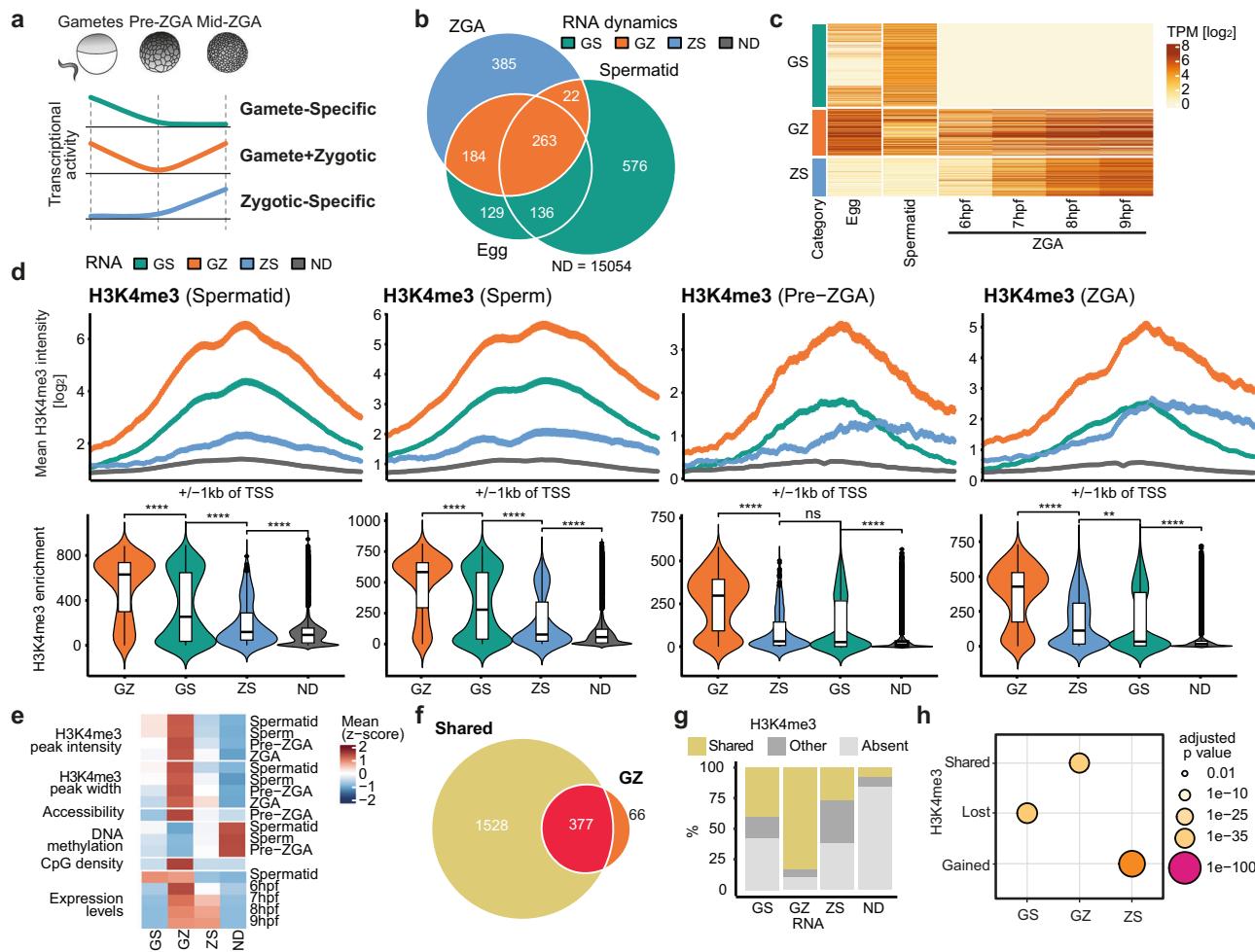


Fig. 3 | H3K4me3 marks gene promoters independently of transcription between two transcriptionally active time points. **a** Schematic representing transcriptional activity in gametes, pre-ZGA and ZGA embryos in each RNA group. Gamete-Specific (GS, $n = 841$): genes that are only expressed in the gametes, Gamete+Zygotic (GZ, $n = 469$): genes that are expressed both in the gametes and zygotically, Zygotic-Specific (ZS, $n = 385$): genes that are expressed only zygotically. **b** Venn diagram denoting overlap of genes detected in the egg, spermatid (total RNA) and at ZGA stages (nascent). Colors are based on the dynamics of RNA expression across stages. **c** Heatmap of RNA expression (log-scaled TPM) of the different RNA dynamics groups in the egg, spermatid and different timepoints during ZGA (6-9hpf). **d** (Top) H3K4me3 around TSS (± 1 kb) for RNA dynamics groups at each timepoint - spermatid, sperm, pre-ZGA, post-ZGA, respectively. Values displayed as mean \pm SE. (Bottom) H3K4me3 promoter enrichment in each group for each time point. Statistical test: one-sided Wilcoxon rank-sum test with

alternative hypothesis that values of left distribution tend to be larger than right distribution; p : (****) ≤ 0.0001 , (***) ≤ 0.001 , (**) ≤ 0.01 , (*) ≤ 0.05 ; n.s. are $p > 0.05$. Box plots show the median (center line), 25–75th percentiles (box), whiskers extending to $1.5 \times$ the interquartile range, and outliers as individual points.

e Heatmap showing the z-scored mean of H3K4me3 peak intensity and peak width, accessibility, CpG density and DNAm at promoters and expression levels for available timepoints for the different RNA dynamics groups. **f** Overlap of SHARED H3K4me3 dynamics group and the GZ RNA dynamics group. **g** Composition of each RNA dynamics group with respect to the H3K4me3 dynamics groups. LOST and GAINED are summarized as OTHER. **h** Results of statistical tests for association between the H3K4me3 dynamics groups and RNA dynamics groups. P -values are calculated with one-sided Fisher's Exact Test for positive association and FDR adjusted.

While almost all GZ-genes also belong to the SHARED group of genes, the reverse is not true—not all genes of the SHARED group are also GZ-genes (Fig. 3f, g). This suggests that promoter H3K4me3 at pre-ZGA is not strictly indicative of gene expression at ZGA. SHARED genes whose transcripts were not detected at ZGA (Fig. 3f, g, GS and ND groups) may either be expressed below the detection limit of nascent RNA-seq or later in development. This suggests that promoters of genes actively expressed in gametes and embryos were also marked by H3K4me3.

Our results reveal a close correlation between H3K4me3 dynamics and gene expression dynamics in the early embryo. We uncover an association between genes that are pre-marked by H3K4me3 during early development and genes that are transcribed in both spermatid and ZGA, a finding compatible with H3K4me3 driving the maintenance of active chromatin states.

Genes pre-marked with H3K4me3 are expressed early during ZGA

H3K4me3 at promoters and in gene bodies is reported to be important for high processivity of RNA pol II and, consequently, for fast, efficient gene transcription²⁶. Thus, we hypothesized that pre-marking of genes with H3K4me3 prior to ZGA could facilitate their early and fast expression during ZGA.

We analyzed nascent RNA-seq data⁴³ and classified all zygotically expressed genes into four groups based on the timepoint at which their transcripts were first detected (Fig. 4a). Indeed, the earliest identified transcripts at 6 hours post-fertilization (hpf) were from genes significantly associated with the SHARED group, whereas later-expressed genes (7–9 hpf) were associated with both SHARED and GAINED groups (Fig. 4b). Similarly, the earliest expressed genes (6 hpf) were associated with the GZ group, and the significance of association

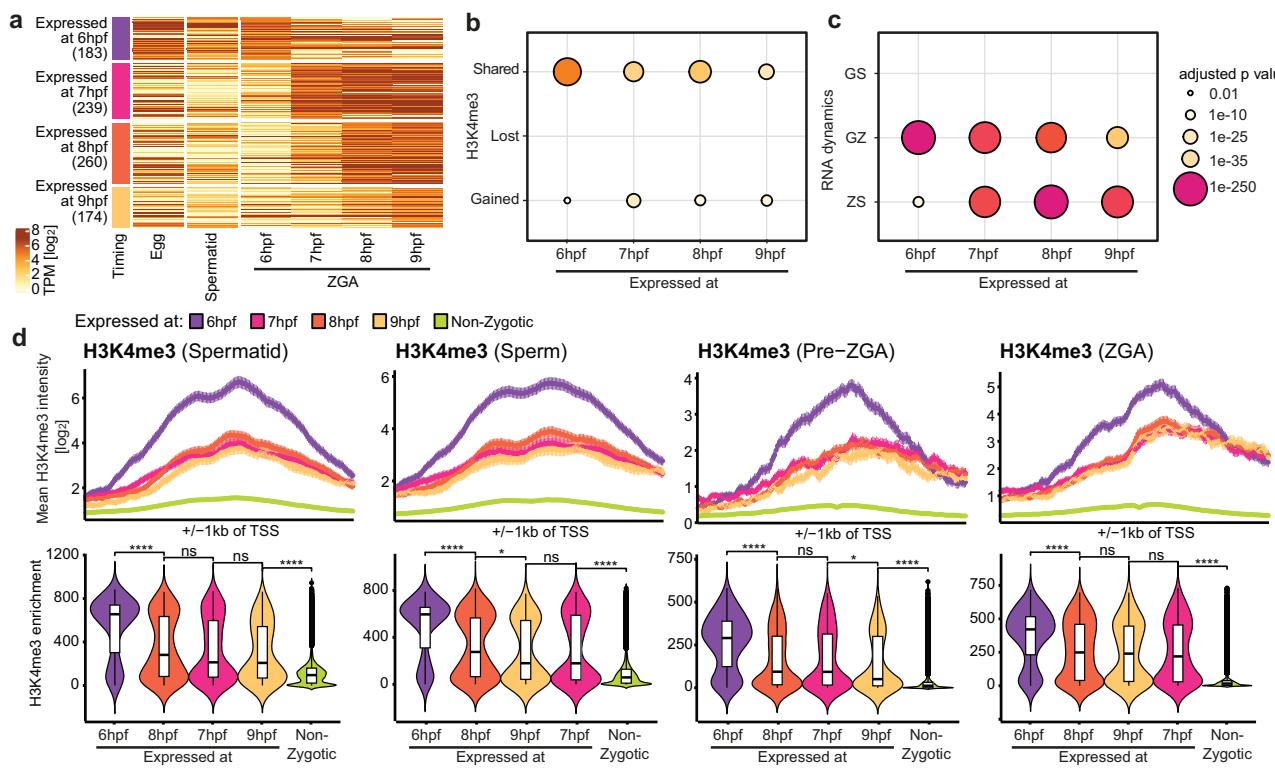


Fig. 4 | Genes pre-marked by H3K4me3 are expressed early during zygotic genome activation. **a** Heatmap of expression of zygotic genes in egg, spermatid, and different timepoints during ZGA grouped by their first detected expression timepoint during ZGA (6 hpf: $n = 183$, 7 hpf: $n = 239$, 8 hpf: $n = 260$, 9 hpf: $n = 174$). Zygotic genes = Gamete+Zygotic (GZ) + Zygotic-Specific (ZS). **b** Results of statistical tests for association between H3K4me3 dynamics groups and gene groups based on their first-detected expression timepoint during ZGA (6–9 hpf). P -values are calculated with one-sided Fisher's Exact Test for positive association and FDR adjusted. **c** Results of statistical tests for association between RNA dynamics groups

and gene groups based on their first-detected expression timepoint during ZGA (6–9 hpf). **d** (Top) H3K4me3 around TSS (± 1 kb) for expression timing gene groups at each timepoint - spermatid, sperm, pre-ZGA, post-ZGA. Values displayed as mean \pm SE. (Bottom) H3K4me3 promoter enrichment in each group for each time point. Statistical test: one-sided Wilcoxon rank-sum test with alternative hypothesis that values of left distribution tend to be larger than right distribution; p : (****) ≤ 0.0001 , (**) ≤ 0.001 , (*) ≤ 0.05 ; n.s. are $p > 0.05$. Box plots show the median (center line), 25–75th percentiles (box), whiskers extending to $1.5 \times$ the interquartile range, and outliers as individual points.

progressively decreased over time (Fig. 4c). Conversely, genes with later-detected transcripts (7–9 hpf) were associated with the ZS group, with the significance of association increasing with time (Fig. 4c).

We further found that genes expressed at 6 hpf had stronger promoter H3K4me3 peak intensities than other groups at gamete, pre-ZGA and ZGA stages (Fig. 4d). No significant differences in peak intensities were observed among later-expressed genes, while non-zygotic genes remained significantly depleted (Fig. 4d). Later-expressed genes (7–9 hpf) gained H3K4me3 asymmetrically from their promoter region towards the 3' end as ZGA began, with peak breadth increasing in parallel (Fig. S4a), reminiscent of the trend seen for gene promoters in the GAINED group and reported previously¹¹ (Fig. S2f).

Promoters of early-expressed genes had high CpG density, in contrast to the low enrichment at non-zygotic promoters (Fig. S4b). At pre-ZGA, DNAme was reduced at early-expressed gene promoters but remained high at non-zygotic and later-expressed genes, suggesting that methylation is present at loci not yet poised for expression (Fig. S4c). This was consistent with the negative correlation between DNAme, nascent expression levels, and H3K4me3 abundance at pre-ZGA (Fig. S4d). Chromatin accessibility distinguished between expressed from non-zygotic genes, but did not correlate with gene expression timing (Fig. S4e). Together, these features highlight that early-expressed genes largely overlap with genes from the SHARED group, and are associated with CpG-rich, accessible, hypomethylated promoters marked by H3K4me3, facilitating timely transcriptional onset (Fig. S4f).

In summary, we uncovered a correlation between H3K4me3 dynamics and gene expression dynamics in early embryos. We find that H3K4me3 marks promoter regions of genes transcribed in the gametes, silenced across multiple cell divisions, and reactivated early during ZGA.

H3K4me3 marking of chromatin during early embryonic divisions is required for proper ZGA

Having established that H3K4me3 is detectable at promoter regions of genes after multiple cell divisions when transcription is globally paused, we next investigated whether this pre-marking is also functionally linked to gene expression dynamics at ZGA. We aimed to determine whether disrupting H3K4 methylation in the cell cycles preceding ZGA, but not during ZGA itself, affects the proper activation of gene expression at ZGA.

We established an assay to reduce H3K4 methylation levels in embryos between fertilization and ZGA. We used a combination of (1) overexpressing the H3K4-specific demethylase Kdm5b^{wt} to remove the methyl residues and (2) expression of the H3.3 variant with a lysine-to-methionine mutation H3.3^{K4M}, sequestering H3K4 methyltransferases and preventing deposition of new methyl residues in a dominant negative manner^{44–46}. The catalytically inactive demethylase Kdm5b^{ci}, combined with wild-type H3.3^{wt}, served as a control treatment. All mRNA transcripts were equipped with a C-terminal 3xHA tag and an auxin-inducible degron (AID) sequence to enable their degradation via TIR1 prior to the ZGA stage.

A combination of *Kdm5b*^{wt} and *H3.3*^{K4M} mRNAs, or a combination of their catalytic inactive counterparts *Kdm5b*^{ci} and *H3.3*^{wt} as the control, was co-injected into one-cell embryos together with *TIR1* (Fig. 5a). We confirmed their expression before ZGA (Fig. 5b), resulting in a decrease of H3K4me3 in pre-ZGA embryos expressing *Kdm5b*^{wt} and *H3.3*^{K4M} (Fig. 5c) compared to the control conditions. After 8 cell divisions, auxin-induced degradation of overexpressed proteins restricted the disruption of H3K4 methylation within a “window” from the zygote to the pre-ZGA time-point (Fig. 5a). We observed complete degradation of the ectopic proteins within 2 h (Fig. 5b).

To validate this system, we tested whether H3K4me3 restoration in the “window” treatment condition occurs in a genome-wide localized manner. We performed CUT&RUN against H3K4me3 at 6.5 hpf in window-treated and control embryos, alongside wild-type embryos (Fig. 5d). To compare the genome-wide H3K4me3 signal in control and treated embryos, we defined a consensus peak set as regions present in at least two of three independent CUT&RUN replicates in wild-type embryos. Principal component analysis of spike-in-normalized H3K4me3 counts at these peaks revealed that window-treated samples clustered more closely with their respective controls and with wild-type samples than with the persistently treated samples in the PC space (Fig. S5a). This indicates that H3K4me3 levels in the window-treated samples have recovered to those of wild-type and control samples at mid-ZGA. To test this, we performed differential peak analysis within the consensus peak set and found very few differentially marked promoters between window-treated and control embryos (67 up, 59 down; $p < 0.05$) (Fig. 5e). In contrast, differential analysis revealed substantially more altered peaks in persistent-treated embryos (393 up, 358 down; $p < 0.05$) (Fig. 5e). Taken together, these analyses suggest a widespread recovery of promoter-associated H3K4me3 in the window- compared to the persistent-treatment condition.

Next, we tested whether temporary disruption of H3K4me3 during early embryonic cell division affects the onset of ZGA. We treated embryos as described above (Fig. 5a) and analyzed them at three stages spanning the onset of ZGA (5.5 hpf, 6.5 hpf, 7.5 hpf) by RNA-seq. When comparing embryos with perturbed H3K4me3 propagation against control embryos, we detected significant up- and down-regulation of transcripts, with the number of differentially expressed genes (DEGs) increasing with time (Fig. S5b). Downregulated genes from all timepoints overlapped well between biological replicates (Fig. S5c). Using our CUT&RUN data, we confirmed that at 6.5 hpf, downregulated zygotic genes in the window condition had recovered their H3K4me3 levels, whereas those in the persistent condition remained significantly depleted, while unaffected genes showed no significant change in promoter H3K4me3 enrichment (Fig. S5d). Notably, we observed a strong bias of GZ- and ZS-genes among these downregulated genes in both “window” and “persistent” depletion conditions (Figs. 5f and S5e). Hence, we focused further analysis on all genes found to be expressed at ZGA (GZ+ZS groups of genes).

Although none of our injection conditions fully disrupted the overall dynamics of global transcriptional activation during ZGA, as indicated by the comparable increase in mean transcript levels of ZGA genes between control-injected and non-injected wild-type embryos (Fig. S5f), we observed that, on average, 25% of these zygotic genes were significantly downregulated at 7.5 hpf. In contrast, very few zygotic genes were upregulated in treatments compared to controls (Fig. S5g). Interestingly, downregulated genes from all timepoints together consistently overlapped across “window” and “persistent” depletion conditions (Fig. S5h). The expression of zygotic genes displaying promoter H3K4me3 in normal embryos at all investigated timepoints (SHARED group) and genes gaining promoter H3K4me3 during ZGA (GAINED group) were strongly affected by both treatments, with expression reductions approaching those observed upon transcriptional inhibition with triptolide⁴⁷ (Figs. 5g and S5i). Mean

expression levels of the SHARED group were significantly reduced in treated embryos of the “window” depletion at 7.5 hpf compared to controls, suggesting a direct link between H3K4me3 pre-marking prior to ZGA and accurate gene activation during ZGA (Figs. 5h and S5j). Also, genes gaining promoter H3K4me3 around ZGA (GAINED group) were affected by both “persistent” and “window” treatments. Although the mean expression levels of these genes at 7.5 hpf remained significantly lower in treated embryos compared to controls (Figs. 5h and S5j), H3K4me3 levels at promoters were largely restored in the window condition (Fig. S5k). In contrast, promoter H3K4me3 levels remained significantly depleted in the persistent condition (Fig. S5k), consistent with the observed downregulation of zygotic gene expression (Figs. 5h and S5j). Together, this indicates that a temporary disruption of H3K4 methylation affects the reliable expression of genes accompanied by de novo H3K4me3 deposition at promoters during ZGA. We also observed some downregulated zygotic genes from the ABSENT group, possibly due to secondary effects of the treatment (Figs. 5g and S5i).

Interestingly, among the downregulated genes in the SHARED group under the window treatment is *pou5f3.2*, a TF that, together with *Sox3*, facilitates ZGA in *Xenopus laevis* (Fig. 5i)^{47,48}. *Pou5f3.2* downregulation by the window treatment could potentially affect the expression of zygotic genes that are normally activated by this TF. In contrast, we found that the fold change in *sox3* expression at 7.5 hpf is only mildly reduced in window-treated embryos relative to controls, which likely stems from *sox3.S* and *sox3.L* being GZ genes with strong maternal contribution (Fig. 5h). Overall, persistent reduction of H3K4me3 produced only a slightly stronger phenotype than that of the shorter window treatment, suggesting that presence of pre-ZGA H3K4 methylation, in addition to H3K4me3 acquisition during ZGA, is important for accurate activation of many zygotic genes.

We then tested whether H3K4 methylation in pre-ZGA embryos is crucial for embryonic development. Injection of mRNA at the concentrations determined above (Fig. S6a, b; 1x mRNA dosage) led to the survival failure of both treated and control embryos (Fig. S6c, d), likely due to mRNA overload or roles of the injected proteins other than H3K4 demethylation. Upon halving the dosage (0.5x mRNA dosage), we still observed a reduction in H3K4me3 in both treatment conditions compared to controls (Fig. S6a, b). Importantly, under these conditions, control embryos showed improved survival, while embryos with either window or persistent H3K4me3 reduction continued to exhibit developmental failure (Fig. S6c, d). Together, these results indicate that loss of H3K4 methylation contributes to the observed developmental defects. Notably, even transient reduction of H3K4me3 during the pre-ZGA window led to increased embryonic lethality, indicating that H3K4me3 marking of chromatin prior to ZGA contributes to successful development (Fig. S6e).

In summary, our experiments show that interfering with H3K4 methylation in the globally transcriptionally silent window affects proper ZGA. Additionally, we show that H3K4 methylation in pre-ZGA stages supports successful embryonic development.

Cxxc1 assists transcription-independent deposition of H3K4me3 and zygotic gene expression

We next examined the mechanisms underlying H3K4me3 marking of promoters during the transcriptionally quiescent pre-ZGA window, focusing on enzymes reported to act independently of transcription. In mouse ESCs and oocytes, CFP1 and MLL2 function as transcription-independent H3K4 methyltransferases^{42,49–51} and their *Xenopus* orthologs, *Cxxc1* and *Kmt2b*, are expressed during pre-ZGA³⁹. Both enzymes contain a CXXC domain that recognizes unmodified CpGs and recruits the methyltransferase. Consistently, promoters displaying H3K4me3 at all investigated stages are characterized by high CpG density and low DNAme (Figs. 2i and 3e). These observations support a model in which *Cxxc1* and *Kmt2b* mediate transcription-independent H3K4

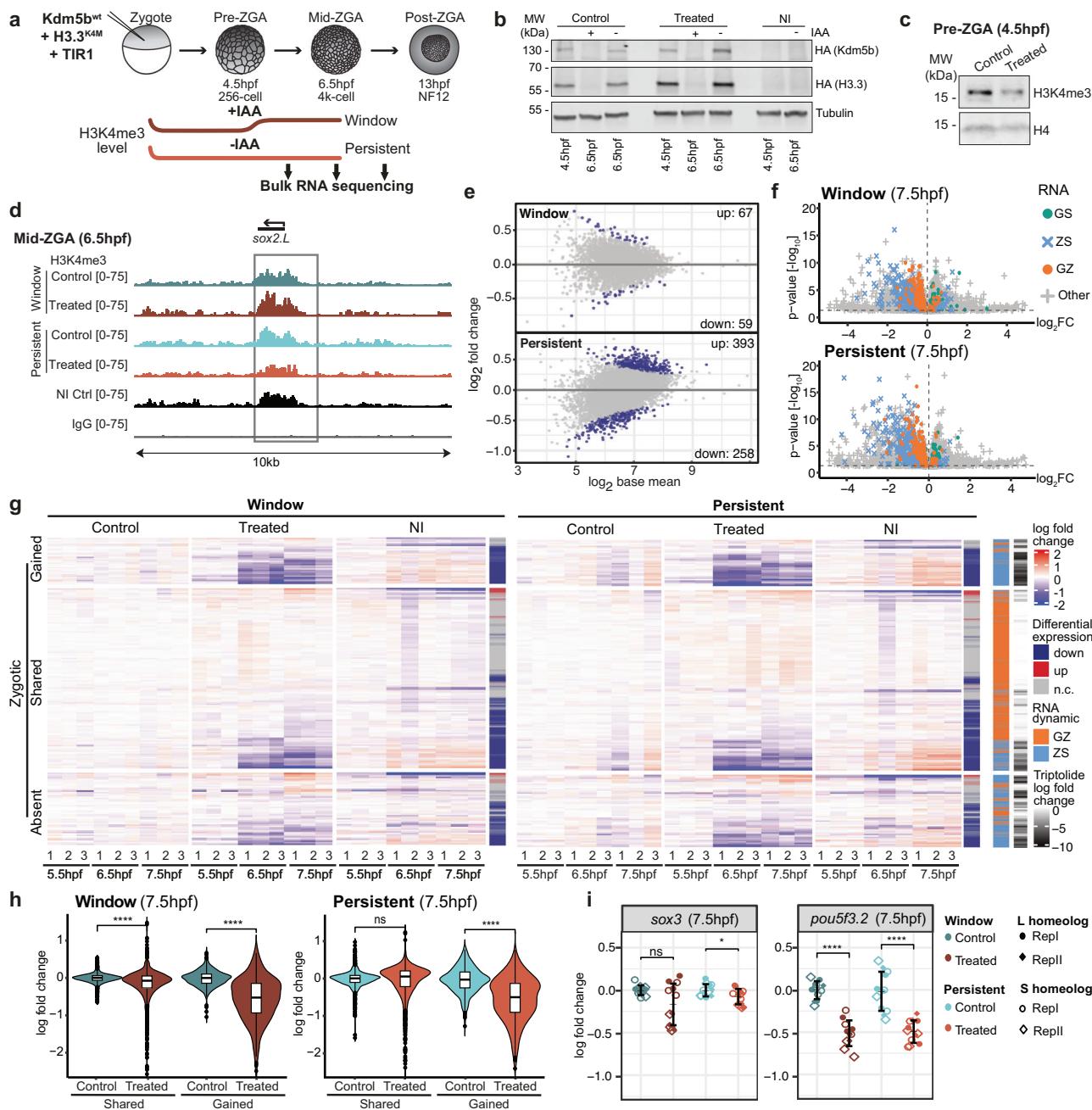


Fig. 5 | Promoter H3K4me3 during early cell divisions is required for proper ZGA. **a** Experimental setup for window-specific and persistent H3K4me3 depletion. **b** Western blot of cytoplasmic fractions from pre- and mid-ZGA embryos showing ectopic HA-tagged Kdm5bwt + H3.3K4M (treatment) or Kdm5bci + H3.3 wt (control) at pre-ZGA, and their auxin-induced (IAA) degradation at mid-ZGA. Tubulin represents the loading control; "NI" denotes non-injected. Blots were independently repeated 2–3 times with consistent results. **c** H3K4me3 intensity relative to H4 in treated and control pre-ZGA embryos after chromatin-bound protein isolation. **d** Genome browser snapshots for *sox2L* showing *E. coli* spike-in normalized coverage for control and treated samples under window and persistent conditions, and wildtype and IgG controls (biological replicate 2). **e** MA-plot comparing H3K4me3 levels in treated versus control samples at promoter peaks for window (top) and persistent (bottom) conditions. Gray: all promoter peaks subset from the consensus peak set in (b). Blue: significantly changed ($p < 0.05$). **f** Volcano plots displaying log2 fold change and adj. p -value (cutoff: 0.05) at 7.5 hpf, colored by RNA dynamics groups for (top) window and (bottom) persistent conditions (biological replicate 1). **g** Zygotic gene expression (three technical replicates per sample), clustered by H3K4me3 dynamics groups (SHARED, GAINED, ABSENT) (biological replicate 1). Log2 fold changes are relative to the mean control TPM at each time point. Genes are annotated for differential expression in each condition and by RNA dynamics group. The rightmost column displays triptolide-induced log2 fold changes at 7 hpf (adapted from Phelps et al., 2022), representing the expected maximal transcriptional downregulation at this stage. **h** Log2 fold change of zygotic SHARED ($n = 473$) and GAINED ($n = 121$) genes over the mean TPM of respective control technical replicates in window (left) and persistent (right) conditions at 7.5 hpf in biological replicate 1. One-sided Wilcoxon rank-sum test with alternative hypothesis that values of the left distribution tend to be larger than the right distribution; p : (****) ≤ 0.0001 , (**) ≤ 0.001 , (*) ≤ 0.01 ; n.s. > 0.05 . Box plots show the median (center line), 25–75th percentiles (box), whiskers extending to 1.5 \times the interquartile range, and outliers as individual points. **i** Log2 fold change expression of *pou5f3.2* and *sox3* homeologs over mean TPM of respective controls at 7.5 hpf ($n = 12$); both biological replicates shown (distinct point styles). Error bars indicate mean \pm SD.

sample), clustered by H3K4me3 dynamics groups (SHARED, GAINED, ABSENT) (biological replicate 1). Log2 fold changes are relative to the mean control TPM at each time point. Genes are annotated for differential expression in each condition and by RNA dynamics group. The rightmost column displays triptolide-induced log2 fold changes at 7 hpf (adapted from Phelps et al., 2022), representing the expected maximal transcriptional downregulation at this stage. **h** Log2 fold change of zygotic SHARED ($n = 473$) and GAINED ($n = 121$) genes over the mean TPM of respective control technical replicates in window (left) and persistent (right) conditions at 7.5 hpf in biological replicate 1. One-sided Wilcoxon rank-sum test with alternative hypothesis that values of the left distribution tend to be larger than the right distribution; p : (****) ≤ 0.0001 , (**) ≤ 0.001 , (*) ≤ 0.01 ; n.s. > 0.05 . Box plots show the median (center line), 25–75th percentiles (box), whiskers extending to 1.5 \times the interquartile range, and outliers as individual points. **i** Log2 fold change expression of *pou5f3.2* and *sox3* homeologs over mean TPM of respective controls at 7.5 hpf ($n = 12$); both biological replicates shown (distinct point styles). Error bars indicate mean \pm SD.

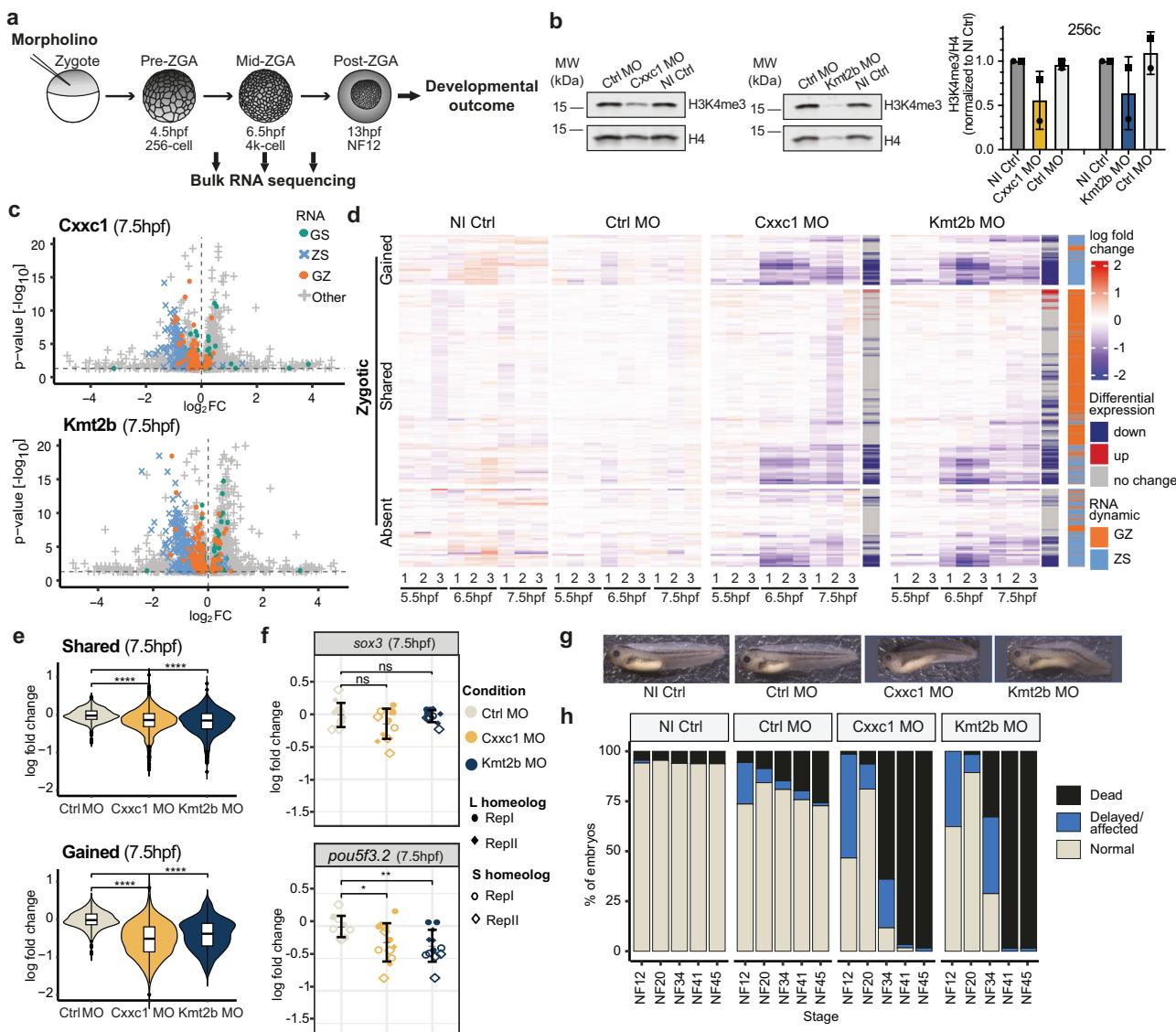


Fig. 6 | Cxxc1 and Kmt2b are required for proper ZGA and embryonic development. **a** Illustration of experimental setup. **b** (Left) Western blot showing H3K4me3 levels in Cxxc1 and Kmt2b knockdown embryos compared to non-injected control and control morpholino-injected embryos at pre-ZGA. (Right) Quantification of H3K4me3 signal intensity relative to H4 for each condition for three independent experiments. Data are presented as mean values \pm SD. **c** Volcano plots displaying expression log2 fold change and adj. p-value (cutoff: 0.05) in knockdown conditions compared to control morpholino for biological replicate 1 at 7.5 hpf, colored by RNA dynamics groups. **d** Expression of zygotic genes in 3 technical replicates of every sample for biological replicate 1, clustered by H3K4me3 dynamics group (SHARED, GAINED and ABSENT). Log2 fold change of every gene is calculated over the mean TPM of three technical replicates of the control morpholino of the respective time point. Genes are annotated with differential expression status separately for the Cxxc1 and Kmt2b knockdown conditions, as well as by RNA dynamics group. **e** Log2 fold change of gene expression

levels of all zygotic genes calculated over mean TPM of control morpholino technical replicates for biological replicate 1 for SHARED (top, $n = 473$) and GAINED (bottom, $n = 121$) groups at 7.5 hpf. Statistical test: one-sided Wilcoxon rank-sum test with alternative hypothesis that values of left distribution tend to be larger than right distribution; p : (****) ≤ 0.0001 , (**) ≤ 0.001 , (*) ≤ 0.01 ; n.s. are $p > 0.05$. Box plots show the median (center line), 25–75th percentiles (box), whiskers extending to 1.5 \times the interquartile range, and outliers as individual points. **f** Log2 fold change expression of *pou5f3.2* and *sox3* calculated over mean TPM of control morpholino replicates at 7.5 hpf ($n = 12$). Both biological replicates are represented, each by a different point style. Error bars indicate mean \pm SD. **g** Representative phenotype images of embryos for each condition at NF41. $n = 2$ experiments, $N > 35$ per condition. **h** Quantification of survival phenotype. $n = 2$ experiments, $N > 35$ embryos per condition. Source data are provided as a Source Data file.

trimethylation via sequence-specific hypomethylation-related reader-writer mechanisms.

We tested whether Cxxc1 and Kmt2b contribute to H3K4 methylation in early *Xenopus* embryos and proper gene activation during ZGA. We injected translation-blocking antisense morpholinos (asMOs) against both homeologs of *kmt2b* and *cxxc1*, or control morpholinos (ctrlMO), into one-cell embryos and allowed them to develop (Fig. 6a). We observed global depletion of Cxxc1 and Kmt2b protein levels in asMO- (Fig. S7a) and correspondingly reduced H3K4me3 levels in asMO-

injected embryos compared to ctrlMO-injected embryos at both pre-ZGA (Fig. 6b) and post-ZGA (Fig. S7b).

We then investigated the effect of this treatment on ZGA. We performed RNA-seq on Cxxc1 and Kmt2b knockdown embryos at three time points spanning the onset of ZGA (Fig. 6a). Differential gene expression analyses revealed significant up- and downregulation of transcripts in both Cxxc1 and Kmt2b knockdown embryos, compared to the controls, with the number of DEGs increasing with developmental time (Fig. S7c). We identified overlap of downregulated genes

across biological replicates in both Cxxc1 and Kmt2b knockdown conditions (Fig. S7d). We observed a bias of GZ- and ZS-genes among the downregulated genes (Figs. 6c and S7e) and focused further analysis on all zygotically expressed genes.

We found that an average of 10% and 16% of zygotic genes were downregulated in Cxxc1 and Kmt2b knockdown embryos, respectively, at 7.5 hpf (Fig. S7f), with an overlap across both conditions (Fig. S7g). Out of the downregulated zygotic genes, those belonging to the groups that keep H3K4me3 around their promoters during early embryonic development (SHARED) and those that gain H3K4me3 during ZGA stage (GAINED) were affected (Figs. 6d and S7h). At 7.5 hpf, we observed significantly decreased mean expression levels of genes in the SHARED and GAINED groups in Cxxc1 or Kmt2b knockdown versus control embryos (Figs. 6e and S7i). Both conditions affected the expression of genes involved in development or in transcription regulation (Fig. S7j). In both Cxxc1 and Kmt2b asMO-injected embryos, we observed a significantly reduced fold change in *pou5f3.2* expression compared to controls, but only a slight reduction in *sox3* expression at 7.5 hpf, explained by its maternal contribution (Fig. 6f). These observations link Cxxc1 and Kmt2b to H3K4me3 propagation and acquisition during *Xenopus* development and demonstrate a role for these enzymes in proper ZGA. To link early chromatin regulation to later transcriptional changes, we compared the genes affected upon Cxxc1 and Kmt2b knockdown with those sensitive to H3K4me3 loss (Fig. 5), revealing an overlap across all four conditions (Fig. S7o). This supports an impact of early H3K4me3 marking on gene expression programs.

We then addressed whether Kmt2b and Cxxc1 depletion affected development. Both Cxxc1 and Kmt2b knockdown embryos showed delayed development compared to control and wildtype embryos, as well as morphological defects, including aberrant gut formation, an overall twisted, shorter body axis, and a “bent tail” phenotype (Fig. 6g). Tadpoles with methyltransferase knockdown did not survive past the feeding tadpole stage (4 days post fertilization), suggesting that Cxxc1 and Kmt2b are each essential for development (Fig. 6h). It is yet unclear if their function is important for successful development only during pre ZGA stages or also during later stages.

To test phenotype specificity, we generated and validated MO-resistant *cxxc1* mRNAs (Fig. S7k). Injection of regular *cxxc1* mRNAs or Cxxc1 asMOs alone caused poor embryonic survival (Fig. S7l–n). However, co-injection of MO-resistant *cxxc1* mRNA with the Cxxc1 asMO resulted in partial rescue of the phenotype, indicating asMO specificity for Cxxc1 (Fig. S7l–n). Rescue experiments for Kmt2b could not be performed due to technical problems. Nevertheless, a substantial overlap between genes downregulated upon Cxxc1 and Kmt2b knockdown and those sensitive to H3K4me3 loss (Fig. S7o) supports a shared role of these enzymes in regulating gene expression and accentuates the importance of H3K4me3 marking on proper embryonic gene expression.

In summary, our results reveal that the methyltransferases Cxxc1 and likely Kmt2b regulate accurate activation of zygotic genes marked by H3K4me3 during embryogenesis and are important for embryonic development. These findings underscore the importance of H3K4me3 reader–writer enzymes in ensuring proper H3K4me3 patterns, facilitating accurate zygotic gene expression, and supporting successful development.

Discussion

This study shows that early embryonic chromatin is pre-marked with H3K4me3 and reveals its crucial role in proper ZGA and development. We show that H3K4me3 is found at common promoter loci in gametes, transcriptionally silent pre-ZGA embryos, and during ZGA when transcription begins. We show that pre-marking chromatin with H3K4 methylation supports correct ZGA and is facilitated by H3K4 methyltransferases Cxxc1 and likely Kmt2b, whose disruption results in reduced H3K4me3 levels, aberrant ZGA, and developmental failure.

Thus, these findings support a role of H3K4 methylation in mediating epigenetic memory of active chromatin states.

We find that numerous histone modifications are present on pre-ZGA chromatin, including active modifications. This is unexpected given the dynamic nature of early embryonic chromatin, which replicates approximately every 30 min with repressed transcription. We observe that transcription-independent Cxxc1 contributes to H3K4me3 deposition^{42,50} and that Kmt2b may also deposit canonical promoter H3K4me3 in pre-and post-ZGA *Xenopus* embryos, in addition to its known role in establishing broad non-canonical domains in mouse oocytes⁵¹. These enzymes likely recognize CpG-rich, hypomethylated promoters and reinstate H3K4me3 after DNA replication, which may be enhanced by broad H3K4me3 domains across several nucleosomes. However, evaluating H3K4me3 at discrete stages cannot resolve continuous transmission at SHARED promoters, and additional contributions from TFs or DNA sequence remain possible.

We find that H3K4me3 occurs at pre-ZGA despite global transcription repression and the absence of the original inducing signal. While current interference methods cannot distinguish between H3K4me2/3, loss of H3K4 methylation impairs gene activation and embryonic development, consistent with H3K4 methylation contributing to epigenetic memory of active chromatin states^{13,14,52}. Whether other active histone modifications detected at pre-ZGA show similar dynamics remains unexplored.

H3K4me3 and DNAme are mutually antagonistic, as H3K4me3 impedes de novo DNAme by inhibiting DNMT3 binding^{34,53}. Conversely, hypomethylated DNA recruits H3K4 methyltransferases via the CXXC domain^{50,51,54}. This raises the question whether early H3K4me3 reflects underlying DNAme patterns, or vice versa. Previous work in *Xenopus tropicalis* hypothesized that H3K4me3 is absent in cleavage-stage embryos and acquired during ZGA based on DNAme, favoring hypomethylated, CpG-rich regions¹¹. Our analyses support a distinct model for H3K4me3 in *Xenopus laevis*, consistent with recent hybrid *Xenopus* findings¹⁸. We demonstrate that H3K4me3 is present in gametes and at both pre- and post-ZGA. These findings suggest that the two frog species have diverged in ZGA regulatory mechanisms. Nevertheless, promoter hypomethylation in *Xenopus laevis* pre-ZGA embryos mirrors trends reported in *Xenopus tropicalis* mid-ZGA embryos³⁸.

What is the role of H3K4me3 during early development? In previous experiments in zebrafish and mouse addressing the roles of histone modifications and variants such as H3K4me3, H3K27ac, H3K9me3, H3K27me3, H2AK119ub and H2A.Z in supporting ZGA, embryos were typically subjected to interference from fertilization onwards^{24,44,55–68}. Unfortunately, this fails to distinguish whether the propagation of the histone mark before ZGA or the accumulation of the mark throughout early development is important for establishing accurate gene expression patterns at ZGA. We show that H3K4 methylation perturbation within a time window before ZGA has a gene expression phenotype almost as strong as persistent depletion, suggesting that pre-ZGA H3K4me3 marking is important for proper onset of ZGA and successful development.

Canonical pre-ZGA H3K4me3 patterns in *Xenopus* closely resemble those of human oocytes and embryos⁶⁹, in contrast to non-canonical patterns in mouse oocytes⁵⁸. However, *Xenopus* and human embryos may transiently exhibit mouse-like H3K4me3 patterns at an earlier developmental stage. Mll2 is expressed in early human embryos and depleted as ZGA begins, much like the expression patterns of its ortholog Kmt2b in *Xenopus*⁶⁹. In context with other species-similarities, such as the higher degree of paternal histone retention after fertilization in both species and multiple transcriptionally silent cell divisions before ZGA^{12,70–72}, *Xenopus* may provide unique insights into mechanisms supporting epigenetic memory of active states in pre-ZGA human embryos.

We find that early gene reactivation during ZGA correlates with H3K4me3 pre-marking. Because promoter and gene-body H3K4me3 are important for high RNA Pol II processivity²⁶, pre-ZGA H3K4me3 at promoters may aid in rapid gene expression early during ZGA. Interestingly, TFs reported to facilitate ZGA in *Xenopus*, Sox3 and Pou5f3^{2,47,48}, are amongst those genes. Their combinatorial knock-down causes developmental arrest at gastrulation^{47,48} in *Xenopus*. Both persistent and window disruption of H3K4 methylation before ZGA result in significant downregulation of *pou5f3.2* and, to a lesser extent, *sox3*, likely causing decreased protein levels and downstream ZGA effects.

In summary, our findings highlight that H3K4me3 pre-marking of promoter regions in early embryos is essential for subsequent accurate ZGA and development, with Cxxc1 ensuring proper deposition at hypomethylated, CpG-rich loci.

Methods

Xenopus laevis culture

Adult *Xenopus Laevis* were obtained from Nasco (901 Janesville Avenue, P.O. Box 901, Fort Atkinson, WI 53538-0901, USA) and Xenopus1 (Xenopus1, Corp. 5654 Merkel Rd. Dexter, MI. 48130). All frog maintenance and care were conducted according to the German Animal Welfare Act. Research animals were used following guidelines approved and licensed by ROB-55.2-2532.Vet_02-23-126.

Embryo collection

Sexually mature females were pre-primed for ovulation by injection of 150 IU of Ovogest 300 I.E./ml into their dorsal lymph sac 4–5 days prior to experiments. 500 IU of the same hormone at 1000 I.E./ml was injected on the day before experiments, and the females were kept overnight at 14 °C in 1x MMR (Marc's Modified Ringer's buffer: 100 mM NaCl, 2 mM KCl, 1 mM MgSO₄, 2 mM CaCl₂, 0.1 mM EDTA, 5 mM HEPES pH 7.8). Females were then moved to room temperature to lay eggs. Eggs were collected in a timely manner, and in vitro fertilization was performed.

In vitro fertilization

In vitro fertilization was performed on eggs by shredding a piece of freshly dissected testes between the two arms of a rifled pair of forceps to release the sperm in a glass dish containing eggs in a minimal amount of 1x MMR. After a 2-min incubation, the 1x MMR was diluted approximately 10-fold with distilled water, and the fertilized eggs were kept at room temperature for 20–30 min. Subsequently, the jelly coating on the fertilized eggs was chemically removed using a dejellying solution (2% L-Cysteine in milliQ H₂O, pH 7.8; adjusted by NaOH), followed by washes in 0.1x MMR.

Microinjections

For labeling nascent RNA, embryos were injected with 9.2 nl of 50 μM 5-EU at the 1-cell stage using a NANONINJECT II microinjector (Drummond Scientific Company). For transcription inhibition, embryos were co-injected with 4.6 ng α-amanitin. For all injection experiments, embryos were kept in 0.5x MMR during and up to 2 h post-injections, after which they were transferred to 0.1x MMR for further development.

Chromatin-bound protein isolation of pre-ZGA embryos

Two hundred 256-cell stage embryos per sample were collected at the desired stage. Buffer was removed, and embryos were gently washed thrice using embryo extraction buffer (10 mM HEPES pH 7.7, 100 mM KCl, 50 mM Sucrose, 1 mM MgCl₂, 0.1 mM CaCl₂) and then spun down for 1 min at 700 g. All excess liquid was removed, and samples were stored at -80 °C until further processing. Frozen embryos were thawed on ice and spun at 17,000 × g for 10 min at 4 °C. This results in the formation of three layers: the upper lipid layer and the bottom

layer of cellular debris are avoided, and the middle gray layer of nuclei is collected using a P200 pipette. The nuclear fraction is washed twice with SunASP buffer (250 mM Sucrose, 75 mM NaCl, 0.5 mM Spermidine, 0.15 mM Spermine) with a 5-min spin at 4 °C at 3500 g and then twice with RIPA buffer (50 mM Tris-HCl pH 7.4, 1% NP40, 0.25% Na-deoxycholate, 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 0.5 mM DTT, 5 mM Na-butyrate) supplemented with protease inhibitors, spun for 10 min at 14,000 × g at 4 °C. A Kimwipe is used to aid the removal of yolk and debris caught on the walls of the tube while removing the supernatant between washes. For SDS-PAGE, the pellet was resuspended in 2x E1 buffer (50 mM HEPES-KOH, pH 7.5, 140 mM NaCl, 0.1 mM EDTA, pH 8.0, 10% Glycerol, 0.5% Igepal CA-630, 0.25% Triton X-100, 1% β-mercaptoethanol) and 4x Laemmli sample buffer (BIO-RAD, #1610747) supplemented with 10% β-mercaptoethanol. An equivalent of 50 embryos was loaded per well.

Western blotting

For testing histone modifications in gastrula (NF stage 12) embryos, 1–3 embryos were collected and resuspended in E1 buffer, followed by centrifugation at 4 °C for 2 min at 3500 g. The supernatant containing the cytoplasmic fraction was collected, and the pellet was dislodged with a pipette tip, resuspended in Tris/SDS solution (50 mM Tris-HCl, pH 6.8 + 2% SDS) supplemented by protease inhibitors and vortexed well. 4x Laemmli buffer was added to both chromatin and non-chromatin bound fractions, chromatin was fragmented by manual disruption and heated for 15 min at 95 °C. 1/4th of an embryo was loaded per well.

For testing cytoplasmic proteins in early cleavage stage embryos, embryos were resuspended in RIPA buffer (50 mM Tris-HCl, pH 7.4, 1% NP40, 0.25% Na-deoxycholate, 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 0.5 mM DTT, 5 mM Na-butyrate) supplemented with protease inhibitors and spun down at 14,000 g for 10 min at 4 °C. The supernatant containing the cytosolic fraction was collected, and 4x Laemmli buffer was added. The samples were manually disrupted and vortexed well, followed by heating at 95 °C for 15 min. 1/4th of an embryo was loaded per well.

SDS-PAGE was performed on 4–20% gradient gels from (Bio-Rad, Mini-PROTEAN TGX #4561096) followed by transfer on PVDF membranes and blocking in 5% BSA in TBS. Depending on the experiment, antibodies were used at 1:1000 dilutions and were incubated overnight at 4 °C (see Table S4 for antibody details). Membranes were imaged using the Licor Odyssey CLx system, and image analysis was performed using ImageJ.

Histone tail mass spectrometry and data analysis

Pre-ZGA (256-cell stage i.e., 4.5 hpf) and mid-ZGA (4000-cell stage i.e., 6.5 hpf) embryos were prepared using the chromatin-bound protein sample preparation method described above (*N*=150 embryos per time point, *n*=3). An equivalent of 50x embryos was loaded per well on a 4–20% gradient gel for SDS-PAGE to resolve histones. Gels were stained using a Coomassie R-250 staining solution (0.2% Coomassie R-250 (w/v), 30% ethanol, 10% glacial acetic acid) at room temperature for 30–40 min. Gels were then destained (30% ethanol, 10% glacial acetic acid) at room temperature every few hours for approximately 24 h until distinct histone bands were achieved against a transparent background. Gels were excised between the 10–20 kDa range, and in-gel sample preparation was performed for mass spectrometry.

Gel pieces containing histones were washed with 100 mM ammonium bicarbonate, dehydrated with acetonitrile, chemically propionylated with propionic anhydride and digested overnight with trypsin. Tryptic peptides were extracted sequentially with 70% acetonitrile/0.25% TFA and acetonitrile, filtered using C8-StageTips, vacuum concentrated and reconstituted in 15 μl of 0.1% FA.

For LC-MS/MS purposes, desalted peptides were injected in an Ultimate 3000 RSLC nano system (Thermo) and separated in a 25 cm

analytical column (75 μ m ID, 1.6 μ m C18, IonOpticks) with a 50 min gradient from 2 to 37% acetonitrile in 0.1% formic acid. The effluent from the HPLC was directly electrosprayed into an Orbitrap Exploris 480 HF (Thermo) operated in data-dependent mode to automatically switch between full scan MS and MS/MS acquisition with the following parameters: survey full scan MS spectra (from m/z 250–1200) were acquired with resolution $R = 60,000$ at m/z 400 (AGC target of 3×10^6). The 15 most intense peptide ions with charge states between 2 and 6 were sequentially isolated to a target value of 2×10^5 ($R = 15,000$) and fragmented at 27% normalized collision energy. Typical mass spectrometric conditions were: spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 250 °C; ion selection threshold, 33,000 counts.

Data analysis was performed with the Skyline (version 21.2) by using doubly and triply charged peptide masses for extracted ion chromatograms. Automatic selection of peaks was manually curated based on the relative retention times and fragmentation spectra with results from Proteome Discoverer 1.4. Integrated peak values were exported for further calculations. The relative abundance of an observed modified peptide was calculated as a percentage of the overall peptide.

Immunofluorescence and image analysis

Immunofluorescence and confocal imaging of nascent transcripts in whole-mount embryos was performed as previously described^{47,43} with minor modifications. Embryos were fixed at mid-ZGA, stained for 5-EU incorporation and/or histone marks, cleared, and imaged by confocal microscopy. Nuclear segmentation was performed using a custom-trained Cellpose model applied to the DNA channel. Mean nuclear intensities were extracted for downstream quantification. Detailed experimental and imaging parameters and computational workflow are provided in the Supplementary Methods.

MBD-seq sample collection, library preparation and sequencing
250x pre-ZGA (256-cell stage, i.e., 4.5 hpf) embryos were collected per replicate and prepared using the chromatin-bound protein sample preparation method as described above. Methylated DNA capture was performed using the Methylated DNA Capture kit (Diagenode: #C02020010) following the manufacturer's instructions. The experimental protocol, library preparation and sequencing parameters are explained in detail in the Supplementary Methods.

MBD-seq and ChIP-seq data processing

Sequencing reads were aligned to the *Xenopus laevis* genome (v10.1) with bowtie (v4.2). The alignment process excluded discordant and mixed reads, with paired-end reads mapped to the reference genome. Aligned reads were converted to BAM format and filtered to retain only properly paired alignments. HOMER (v4.11) was used for peak identification against corresponding input controls. Broad peaks were called using histone-style parameters, while narrow peaks were called with default settings. Identified peaks were converted to BED format for downstream analysis. Unless stated otherwise, all analysis on both raw signal intensities and peak calling data was performed on gene promoters. This was done by quantifying the signal intensities, or called peaks, in ± 1 kb window around the TSS of each gene.

CATaDa sample collection, library preparation and sequencing
Chromatin accessibility of pre-ZGA embryos was profiled using CATaDa (Chromatin Accessibility Profiling Using Targeted DamID) as described previously³⁶. Embryos were injected at the 1-cell stage with 2.3 ng pRN3P_AID-Dam-only construct (a kind gift by Maria-Elena Torres-Padilla; Addgene plasmid # 136065)⁷³. 500x injected or non-injected control 256-cell stage embryos were collected per replicate and prepared using the chromatin-bound protein sample preparation method described above. Pellets were resuspended in 400 μ l gDNA

Tissue Lysis Buffer from the Monarch Genomic DNA Purification Kit (#T3010S); instructions for lysis of animal tissue and genomic DNA binding steps were followed as described. Elution was performed with 100 μ l preheated gDNA Elution Buffer. Sample volumes were decreased to 15 μ l using a SpeedVac instrument for 25 min.

All samples were processed for CATaDa as described previously⁷⁴. Extracted genomic DNA was digested overnight at 37 °C with DpnI (NEB, R0176S), and column-purified with the QIAquick PCR Purification Kit (Qiagen, 28104). Adapters were blunt-end ligated to DpnI-digested fragments using T4 DNA ligase (NEB, M0202S; 2 h at 16 °C, heat inactivation at 65 °C for 20 min). Fragments were digested with DpnII (NEB, R0543S), purified with a 1:1.5 ratio of Seramag beads (Fisher Scientific, 65152105050250), and PCR amplified using DreamTaq HS DNA polymerase (Thermo Fisher Scientific, EP1703). PCR product was column-purified with the QIAquick PCR Purification Kit (Qiagen, 28104), sonicated and digested with AlwI (NEB, R0513S) before sequencing library preparation. CATaDa fragments were prepared for Illumina sequencing according to a modified TruSeq protocol⁷⁴. Sequencing was performed as paired-end 50 bp reads by the CRUK Genomics Core Sequencing facility on a NovaSeq 6000.

CATaDa data processing

Analysis of fastq files from CATaDa experiments was performed with the damidseq pipeline script⁷⁵. Preprocessed fastq-files were mapped to the XI.v10.1 genome assembly using bowtie2, and mapped reads were binned into fragments delineated by 5'-GATC-3' motifs. Files were converted to the bigwig file format with bedGraphToBigWig (v4) for visualization with the Integrative Genomics Viewer (v2.13.0). MACS (v3.0.1)⁷⁶ was used to call broad peaks on the mapped read (.bam) files generated by the damidseq_pipeline.

Signal intensity calculation for ChIP-seq, MBD-seq, and CATaDa

The resulting coverages from the pre-processing explained above were log2-transformed after adding a pseudo-count of 1. Final signal intensities for Meta plots and Heatmaps were calculated by adding up the coverage of all replicates.

Meta plots (line plots)—Aggregate line plots, depicting the mean signal intensity of gene sets of interest with bands representing standard error of the mean, were generated using ggplot2 (v3.4.4).

Heatmaps—Heatmaps were generated using the EnrichedHeatmap package (v1.32.0) in R. Each row represents the signal intensity of one gene.

Violin plots—Violin plots show the enrichment calculated with the enriched_score function of the EnrichedHeatmap package within the promoter regions of the genes. Promoter regions were defined as ± 1 kb windows centered on the TSS. The violins are scaled to the same maximum width to improve readability of the density distributions. A one-sided Wilcoxon rank-sum test was used to calculate statistical significance. Four stars (****) denote a p -value smaller or equal than 0.0001. Three stars (***) denote a p -value smaller than or equal to 0.001. Two stars (**) denote a p -value smaller than or equal to 0.01. One star (*) denotes a p -value smaller than or equal to 0.05. P -values larger than 0.05 are denoted with n.s.

Statistical testing for association of gene groups

To quantify the significance of the association between multiple gene lists, such as H3K4me3 dynamics groups, RNA dynamics groups and ZGA timing groups, Fisher's exact test was used. The results are visualized in Balloon plots. Each dot represents the result of a Fisher's exact test on a 2×2 contingency table. The table is created by taking one group of the first list of groups of interest and merging all other groups of the list into a comparison group. The same is repeated for the second list of groups of interest. The one-sided Fisher's test is performed for the positive association of the two gene lists of interest.

The dot size and color show the FDR-adjusted *p*-value to account for multiple testing.

Gene ontology analysis

GO enrichment was calculated with the clusterProfiler package (v4.10.0) in R. The visualization of the similarities of the enriched GO terms (Benjamini-Hochberg adjusted *p* < 0.01) was done with the simplifyEnrichment package (v1.12.0).

mRNA production

mRNAs for all overexpression, AID, and rescue experiments were produced by in vitro transcription from linearized plasmids (full cloning and transcription details in Supplementary Methods).

Auxin experimental setup

For the “treated” condition, embryos were co-injected at the 1-cell stage with 11 ng Kdm5b^{wt}, 345 pg H3.3^{K4M} and 5 ng TIR1 mRNA. “Control” condition embryos were co-injected at the 1-cell stage with 4 ng Kdm5b^{ci}, 345 pg H3.3^{wt} and 5 ng TIR1 mRNA. mRNA amounts were determined on the basis of equal expression levels of the over-expressed proteins in pre-ZGA stage embryos. For “window” depletion treatment, embryos were grown until the pre-ZGA stage (4.5 hpf/256-cell stage) and transferred to an auxin solution (500 μ M auxin in 0.1 \times MMR) to develop further. “Persistent” depletion treatment embryos were grown in 0.1x MMR throughout development. To validate protein expression, pre-ZGA embryos were collected for Western blotting. To validate auxin-induced protein degradation, embryos were collected at the mid-ZGA stage (6.5 hpf/4k-cell stage) and collected for Western blot and immunofluorescence. To assess the extent of H3K4me3 depletion, embryos were collected either at the pre-ZGA stage for chromatin-bound protein isolation and Western blot, or at NF12 for Western blot.

Morpholino-mediated knockdown experimental setup

Custom antisense morpholinos (asMOs) targeting the translation start sites of both L and S homeologs of *Xenopus laevis* *kmt2b* and *cxxc1* were designed with the help of Gene Tools, LLC. The recommended standard control morpholino from Gene Tools LLC was used as a control. The morpholino sequences are shown in Supplementary Table 3.

1-cell stage embryos were injected with 9.2 nl of 1 mM morpholino solution and developed in 0.5x MMR solution for 2 h before moving to 0.1x MMR. Embryos were grown at 18 °C until the desired stage. For analysis of developmental phenotypes, embryos were imaged and counted at periodic intervals with daily replacement of 0.1x MMR.

CUT&RUN

CUT&RUN was performed as described previously^{47,77} with minor modifications. 45 whole embryos were collected per sample at mid-ZGA stage (6.5 hpf). The complete CUT&RUN experimental protocol, conditions, library preparation and sequencing parameters are provided in detail in the Supplementary Methods.

CUT&RUN data processing and visualization

Sequencing reads were aligned to the *Xenopus laevis* genome (v10.1) and a spike-in genome (*Escherichia coli*, GCA_001606525) using Bowtie2 (v2.4). The alignment process excluded discordant and mixed reads, with paired-end reads mapped to the reference genome. Aligned reads were converted to BAM format and filtered to retain only properly paired alignments. Normalization scaling factors were computed using deepTools (v3.5.3) based on the read distribution across genomic regions. HOMER (v4.11) was used for peak identification. Narrow peaks were called using histone-style parameters and

converted to BED format for downstream analysis. Differential peak analysis for H3K4me3 signal was performed on a consensus peak set defined based on narrow H3K4me3 peaks called in at least two out of three independent biological replicates from wildtype embryos, using the findOverlaps function from GenomicRanges (v1.58.0) with maxgap = 150 bp.

Meta plots (line plots) depicting the mean signal intensity of gene sets of interest with bands representing the standard error of the mean were generated using normalizeToMatrix from EnrichedHeatmap (v1.36.0) and *E. coli* spike-in normalized signal. For statistical comparison of H3K4me3 signal across gene sets of interest, the Kolmogorov-Smirnov test was used.

RNA-seq experimental setup

For RNA-sequencing experiments, independent experiments were performed on two different frogs, serving as two biological replicates. Within each biological replicate, three sets of three embryos each were collected per time point, corresponding to three replicates within each independent experiment. Embryos were collected as described at hourly intervals after the pre-ZGA stage (i.e., 4.5 hpf/256-cell stage): at 5.5 hpf, 6.5 hpf and 7.5 hpf. All liquid was removed from the tube, and samples were snap-frozen on dry ice and then stored at -80 °C until further processing. Gene expression was analyzed for morpholino knockdown experiments and H3K4me3 window-depletion experiments.

Total RNA extraction and ribosomal depletion

Three embryos were collected per sample and stored at -80 °C. Total RNA was isolated using the RNeasy kit (Qiagen, #74106) according to the manufacturer’s instructions and ribosomal RNA was depleted using custom-made oligomer mixes for *Xenopus laevis* rRNA as described previously⁷⁸ with minor modifications. Complete protocols are provided in the Supplementary Methods.

RNA-seq library preparation and sequencing

Sequencing libraries were generated using the NEBNext Ultra II Directional RNA Library Preparation Kit for Illumina (NEB, #E7760) as per the manufacturer’s instructions using 12–13 PCR amplification cycles. The quality of cDNA libraries was assessed using the Agilent High Sensitivity D5000 ScreenTape System (Agilent, #5067-5592) in the Agilent 4150 TapeStation System or using the Agilent DNF-474 HS NGS Fragment Kit (Agilent, DNF-474-1000) in the Fragment Analyzer System (Agilent). Libraries were multiplexed, and sequenced was performed as paired-end 50 bp reads on the Illumina NovaSeqX+ by the Helmholtz Core Facility Genomics (CF-GEN).

RNA-seq data processing

Paired sequencing reads were processed using Kallisto (v0.48) for pseudoalignment and quantification of transcript abundance. Transcript and annotation files were downloaded from Xenbase (v10.1) for *Xenopus laevis* (transcripts and annotation). After quantification, transcript-level abundances were imported using “tximport” (v1.32.0) and converted to a “SummarizedExperiment” (v1.34.0) object in R (v4.3.2). Transcript isoforms are summed up for the respective gene count. Datasets of two independent batches were subjected to differential expression analysis on each of their three respective replicates performed with DESeq2 (v1.42.1). Reported *p*-values are FDR adjusted for multiple testing. Transcript abundances were normalized to transcripts per million and log2-transformed with a pseudo count of 1 for further analysis.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Original data produced in the experiments have been deposited on GEO. The MBD-seq data are available under the accession number [GSE286525](#). The CATAcDa data is available under the accession number [GSE286524](#). The RNA-seq data from the auxin experiment are available under the accession number [GSE286887](#). The RNA-seq data from the morpholino treatments are available under the accession number [GSE286886](#). The CUT&RUN data is available under accession number [GSE303651](#). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE⁷⁹ partner repository with the dataset identifier [PXD058937](#). Publicly available data were obtained from GEO with the following accession numbers: [GSE125982](#), [GSE75164](#), [GSE76059](#), [GSE201835](#). All other relevant source data supporting the key findings of this study are provided in this paper. Source data are provided with this paper.

Code availability

All original code can be accessed at Zenodo (<https://doi.org/10.5281/zenodo.1766712>) or GitHub (https://github.com/ScialdoneLab/h3k4me3_maintenance)⁸⁰. The preprocessed data for the analysis can be downloaded from Zenodo (<https://doi.org/10.5281/zenodo.14648394>).

References

1. Steffen, P. A. & Ringrose, L. What are memories made of? How polycomb and trithorax proteins mediate epigenetic memory. *Nat. Rev. Mol. Cell Biol.* **15**, 340–356 (2014).
2. Stewart-Morgan, K. R., Petryk, N. & Groth, A. Chromatin replication and epigenetic cell memory. *Nat. Cell Biol.* **22**, 361–371 (2020).
3. Escobar, T. M. et al. Active and repressed chromatin domains exhibit distinct nucleosome segregation during DNA replication. *Cell* **179**, 953–963.e11 (2019).
4. Howe, F. S., Fischl, H., Murray, S. C. & Mellor, J. Is H3K4me3 instructive for transcription activation? *Bioessays* **39**, 1–12 (2017).
5. Murray, S. C. et al. H3K4me3 is neither instructive for, nor informed by, transcription. Preprint at <https://doi.org/10.1101/709014> (2019).
6. Jukam, D., Shariati, S. A. M. & Skotheim, J. M. Zygotic genome activation in vertebrates. *Dev. Cell* **42**, 316–332 (2017).
7. Lindeman, L. C. et al. Prepatternning of developmental gene expression by modified histones before zygotic genome activation. *Dev. Cell* **21**, 993–1004 (2011).
8. Murphy, P. J., Wu, S. F., James, C. R., Wike, C. L. & Cairns, B. R. Placeholder nucleosomes underlie germline-to-embryo DNA methylation reprogramming. *Cell* **172**, 993–1006.e13 (2018).
9. Veenstra, G. J. C. Dynamics of chromatin remodeling during embryonic development. In *Xenopus* 173–184 (CRC Press, 2022). <https://doi.org/10.1201/9781003050230-13>
10. Akkers, R. C. et al. A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Dev. Cell* **17**, 425–434 (2009).
11. Hontelez, S. et al. Embryonic transcription is controlled by maternally defined chromatin state. *Nat. Commun.* **6**, 10148 (2015).
12. Oikawa, M. et al. Epigenetic homogeneity in histone methylation underlies sperm programming for embryonic transcription. *Nat. Commun.* **11**, 3491 (2020).
13. Hörmanseder, E. et al. H3K4 methylation-dependent memory of somatic cell identity inhibits reprogramming and development of nuclear transfer embryos. *Cell Stem Cell* **21**, 135–143.e6 (2017).
14. Ng, R. K. & Gurdon, J. B. Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. *Nat. Cell Biol.* **10**, 102–109 (2008).
15. Newport, J. & Kirschner, M. A major developmental transition in early *Xenopus* embryos: II. control of the onset of transcription. *Cell* **30**, 687–696 (1982).
16. Almouzni, G. & Wolffe, A. P. Constraints on transcriptional activator function contribute to transcriptional quiescence during early *Xenopus* embryogenesis. *EMBO J.* **14**, 1752–1765 (1995).
17. Chen, H., Einstein, L. C., Little, S. C. & Good, M. C. Spatiotemporal patterning of zygotic genome activation in a model vertebrate embryo. *Dev. Cell* **49**, 852–866.e7 (2019).
18. Long, Q. et al. Modification of maternally defined H3K4me3 regulates the inviability of interspecific *Xenopus* hybrids. *Sci. Adv.* **9**, eadd8343 (2023).
19. Collart, C. et al. High-resolution analysis of gene activity during the *Xenopus* mid-blastula transition. *Development* **141**, 1927–1939 (2014).
20. Yang, J., Tan, C., Darken, R. S., Wilson, P. A. & Klein, P. S. -Catenin/Tcf-regulated transcription prior to the midblastula transition. *Development* **129**, 5743–5752 (2002).
21. Owens, N. D. L. et al. Measuring absolute RNA copy numbers at high temporal resolution reveals transcriptome kinetics in development. *Cell Rep.* **14**, 632–647 (2016).
22. Yasuda, G. Temporal regulation in the early embryo: is MBT too good to be true? *Trends Genet.* **8**, 124–127 (1992).
23. Giraldez, A. J. et al. Zebrafish MIR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**, 75–79 (2006).
24. Chan, S. H. et al. Brd4 and P300 confer transcriptional competency during zygotic genome activation. *Dev. Cell* **49**, 867–881.e8 (2019).
25. Hadzhiev, Y. et al. The miR-430 locus with extreme promoter density forms a transcription body during the minor wave of zygotic genome activation. *Dev. Cell* **58**, 155–170.e8 (2023).
26. Wang, H. et al. H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. *Nature* **615**, 339–348 (2023).
27. Shiokawa, K., Misumi, Y. & Yamana, K. Demonstration of rRNA synthesis in pre-gastrular embryos of *Xenopus Laevis*. *Dev. Growth Differ.* **23**, 579–587 (1981).
28. Nguyen, V. In vivo degradation of RNA polymerase II largest subunit triggered by alpha-amanitin. *Nucleic Acids Res.* **24**, 2924–2929 (1996).
29. Lee, K.-B., Wang, D., Lippard, S. J. & Sharp, P. A. Transcription-coupled and DNA damage-dependent ubiquitination of RNA polymerase II in vitro. *Proc. Natl. Acad. Sci. USA* **99**, 4239–4244 (2002).
30. Jao, C. Y. & Salic, A. Exploring RNA transcription and turnover in vivo by using click chemistry. *Proc. Natl. Acad. Sci. USA* **105**, 15779–15784 (2008).
31. Bernstein, B. E. et al. A Bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
32. Gross, D. S. & Garrard, W. T. Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* **57**, 159–197 (1988).
33. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).
34. Ooi, S. K. T. et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714–717 (2007).
35. Hu, J.-L. et al. The N-terminus of histone H3 is required for de novo DNA methylation in chromatin. *Proc. Natl. Acad. Sci. USA* **106**, 22187–22192 (2009).
36. Aughey, G. N., Estacio Gomez, A., Thomson, J., Yin, H. & Southall, T. D. CATAcDa reveals global remodelling of chromatin accessibility during stem cell differentiation in vivo. *eLife* **7**, e32341 (2018).
37. Van Den Ameele, J. et al. Targeted DamID detects cell-type-specific histone modifications in intact tissues or organisms. *PLoS biology*, **23**, e3002944 (2025).
38. Bogdanović, O. et al. Temporal uncoupling of the DNA methylome and transcriptional repression during embryogenesis. *Genome Res.* **21**, 1313–1327 (2011).
39. Session, A. M. et al. Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature* **538**, 336–343 (2016).

40. Teperek, M. et al. Sperm is epigenetically programmed to regulate gene transcription in embryos. *Genome Res.* **26**, 1034–1046 (2016).

41. Ng, H. H., Robert, F., Young, R. A. & Struhl, K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* **11**, 709–719 (2003).

42. Thomson, J. P. et al. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* **464**, 1082–1086 (2010).

43. Chen, H. & Good, M. C. Nascent transcriptome reveals orchestration of zygotic genome activation in early embryogenesis. *Curr. Biol.* **32**, 4314–4324.e7 (2022).

44. Aoshima, K., Inoue, E., Sawa, H. & Okada, Y. Paternal H3K4 methylation is required for minor zygotic gene activation and early mouse embryonic development. *EMBO Rep.* **16**, 803–812 (2015).

45. Jayaram, H. et al. S-adenosyl methionine is necessary for inhibition of the methyltransferase G9a by the lysine 9 to methionine mutation on histone H3. *Proc. Natl. Acad. Sci. USA* **113**, 6182–6187 (2016).

46. Lewis, P. W. et al. Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* **340**, 857–861 (2013).

47. Phelps, W. A. et al. Hybridization led to a rewired pluripotency network in the allotetraploid *Xenopus laevis*. *eLife* **12**, e83952 (2023).

48. Gentsch, G. E., Spruce, T., Owens, N. D. L. & Smith, J. C. Maternal pluripotency factors initiate extensive chromatin remodelling to predefine first response to inductive signals. *Nat. Commun.* **10**, 4269 (2019).

49. Wang, H. & Helin, K. Roles of H3K4 methylation in biology and disease. *Trends Cell Biol.* **35**, 115–128 (2025).

50. Clouaire, T. et al. Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes Dev.* **26**, 1714–1728 (2012).

51. Hanna, C. W. et al. MLL2 conveys transcription-independent H3K4 trimethylation in oocytes. *Nat. Struct. Mol. Biol.* **25**, 73–82 (2018).

52. Ng, R. K. & Gurdon, J. B. Epigenetic memory of active gene transcription is inherited through somatic cell nuclear transfer. *Proc. Natl. Acad. Sci. USA* **102**, 1957–1962 (2005).

53. Hu, D. et al. Not all H3K4 methylations are created equal: MLL2/COMPASS dependency in primordial germ cell specification. *Mol. Cell* **65**, 460–475.e6 (2017).

54. Hashimoto, H., Vertino, P. M. & Cheng, X. Molecular coupling of DNA methylation and histone methylation. *Epigenomics* **2**, 657–669 (2010).

55. Andreu-Vieyra, C. V. et al. MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing. *PLoS Biol.* **8**, e1000453 (2010).

56. Burton, A. et al. Heterochromatin establishment during early mammalian development is regulated by pericentromeric RNA and characterized by non-repressive H3K9me3. *Nat. Cell Biol.* **22**, 767–778 (2020).

57. Chen, Z., Djekidel, M. N. & Zhang, Y. Distinct dynamics and functions of H2AK119ub1 and H3K27me3 in mouse preimplantation embryos. *Nat. Genet.* **53**, 551–563 (2021).

58. Dahl, J. A. et al. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**, 548–552 (2016).

59. Dang, Y. et al. The lysine deacetylase activity of histone deacetylases 1 and 2 is required to safeguard zygotic genome activation in mice and cattle. *Development* **149**, dev200854 (2022).

60. François-Campion, V. et al. Sperm derived H2AK119ub1 is required for embryonic development in *Xenopus Laevis*. *Nat Commun.* **16**, 3268 (2024).

61. Ibarra-Morales, D. et al. Histone variant H2A.Z regulates zygotic genome activation. *Nat. Commun.* **12**, 7002 (2021).

62. Liu, X. et al. Hierarchical accumulation of histone variant H2A.Z regulates transcriptional states and histone modifications in early mammalian embryos. *Adv. Sci.* **9**, 2200057 (2022).

63. Rong, Y. et al. USP16-mediated histone H2A lysine-119 deubiquitination during oocyte maturation is a prerequisite for zygotic genome activation. *Nucleic Acids Res.* **50**, 5599–5616 (2022).

64. Sankar, A. et al. KDM4A regulates the maternal-to-zygotic transition by protecting broad H3K4me3 domains from H3K9me3 invasion in oocytes. *Nat. Cell Biol.* **22**, 380–388 (2020).

65. Wu, K. et al. Dynamics of histone acetylation during human early embryogenesis. *Cell Discov.* **9**, 29 (2023).

66. Zhang, B. et al. Widespread enhancer dememorization and promoter priming during parental-to-zygotic transition. *Mol. Cell* **72**, 673–686.e6 (2018).

67. Zou, Z., Wang, Q., Wu, X., Schultz, R. M. & Xie, W. Kick-starting the zygotic genome: licensors, specifiers, and beyond. *EMBO Rep.* **25**, 4113–4130 (2024).

68. Fukushima, H. S. & Takeda, H. Coordinated action of multiple active histone modifications shapes the zygotic genome activation in teleost embryos. *Nat. Commun.* **16**, 5222 (2025).

69. Xia, W. et al. Resetting histone modifications during human parental-to-zygotic transition. *Science* **365**, 353–360 (2019).

70. Erkek, S. et al. Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nat. Struct. Mol. Biol.* **20**, 868–875 (2013).

71. Jung, Y. H. et al. Chromatin states in mouse sperm correlate with embryonic and adult regulatory landscapes. *Cell Rep.* **18**, 1366–1382 (2017).

72. Yamaguchi, K. et al. Re-evaluating the localization of sperm-retained histones revealed the modification-dependent accumulation in specific genome regions. *Cell Rep.* **23**, 3920–3932 (2018).

73. Borsos, M. et al. Genome–lamina interactions are established de novo in the early mouse embryo. *Nature* **569**, 729–733 (2019).

74. Marshall, O. J., Southall, T. D., Cheetham, S. W. & Brand, A. H. Cell-type-specific profiling of protein–DNA interactions without cell isolation using targeted DamID with next-generation sequencing. *Nat. Protoc.* **11**, 1586–1598 (2016).

75. Marshall, O. J. & Brand, A. H. damidseq_pipeline: an automated pipeline for processing DamID sequencing datasets. *Bioinformatics* **31**, 3371–3373 (2015).

76. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

77. Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019 (2018).

78. Phelps, W. A., Carlson, A. E. & Lee, M. T. Optimized design of antisense oligomers for targeted rRNA depletion. *Nucleic Acids Res.* **49**, e5–e5 (2021).

79. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidence. *Nucleic Acids Res.* **50**, D543–D552 (2022).

80. Stock, M. et al. Pre-marking chromatin with H3K4 methylation is required for accurate zygotic genome activation and development. *ScialdoneLab/h3k4me3_maintenance: Revision code*. Zenodo <https://doi.org/10.5281/zenodo.17667122> (2025).

Acknowledgements

M.S.O. and E.H. were supported by the HO 6864/2-1 project grant and the Chromatin Dynamics CRC1064, Project-ID 213249687, both DFG; Recognition Award of the Helmholtz Association; Project grant of the MRC MR/P00479/1; S.H. by ERC starting grant 852798. Work in the labs of E.H. and A.S. was funded by the Helmholtz Association. A.E. received support from the Helmholtz Association, awarded to Maria-Elena Torres-Padilla. M.S. was supported by the “Joachim Herz Stiftung” fellowship and the “Munich School for Data Science—MUDS”. JvdA is supported by a Wellcome Clinical Research Career Development Fellowship (219615/Z/19/Z), Wellcome Discovery Award (226653/Z/22/Z), a UKRI BBSRC Grant (BB/X00256X/1) and Medical Research Council (MC_UU_00028/

8). The authors are grateful to the members of the CAM LMU Munich and G. Eckstein and I. de la Rosa Velazquez from the Helmholtz Genomics Core Facility (CF-GEN) for technical support. We thank all present and past lab and institute members, especially Nemanja Vasovic, as well as Ralph Rupp and Kikue Tachibana, for reagents, their critical input and support. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission.

Author contributions

M.S.O. performed experimental work with assistance from M.M., A.J. analyzed the CUT&RUN experiments, I.F. and A.I. performed and analyzed the mass spectrometry experiments, A.H.M. and J.v.d.A performed and analyzed the CATaDa experiments, and A.E. provided expertise in confocal imaging. S.H. provided guidance on experimental design through scientific discussions. M.S., T.S. and M.S.O. performed computational analyses with input from A.S. and E.H. The manuscript was written by M.S.O. and E.H. with input from A.S. and M.S., and E.H. designed and supervised the study.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-67692-7>.

Correspondence and requests for materials should be addressed to Meghana S. Oak or Eva Hörmanseder.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025