


# Sperm derived H2AK119ub1 is required for embryonic development in *Xenopus laevis*

Received: 12 January 2024

Accepted: 28 March 2025

Published online: 05 April 2025

 Check for updates


Valentin Francois--Campion <sup>1,6</sup>, Florian Berger <sup>1,6</sup>, Mami Oikawa<sup>2</sup>, Maissa Goumeidane<sup>1</sup>, Nolwenn Mouni  e <sup>1</sup>, Vanessa Chenouard <sup>1</sup>, Kseniya Petrova<sup>3</sup>, Jose G. Abreu<sup>3,4</sup>, Cynthia Fourgeux <sup>1</sup>, Jeremie Poschmann <sup>1</sup>, Leonid Peshkin <sup>3</sup>, Romain Gibeaux <sup>5</sup> & J  r  me Jullien <sup>1</sup> 

Ubiquitylation of H2A (H2AK119ub1) by the polycomb repressive complexe-1 plays a key role in the initiation of facultative heterochromatin formation in somatic cells. Here we evaluate the contribution of sperm derived H2AK119ub1 to embryo development. In *Xenopus laevis* we found that H2AK119ub1 is present during spermiogenesis and into early embryonic development, highlighting its credential for a role in the transmission of epigenetic information from the sperm to the embryo. In vitro treatment of sperm with USP21, a H2AK119ub1 deubiquitylase, just prior to injection to egg, results in developmental defects associated with gene upregulation. Sperm H2AK119ub1 editing disrupts egg factor mediated paternal chromatin remodelling processes. It leads to post-replication accumulation of H2AK119ub1 on repeat element of the genome instead of CpG islands. This shift in post-replication H2AK119ub1 distribution triggered by sperm epigenome editing entails a loss of H2AK119ub1 from genes misregulated in embryos derived from USP21 treated sperm. We conclude that sperm derived H2AK119ub1 instructs egg factor mediated epigenetic remodelling of paternal chromatin and is required for embryonic development.

Several studies point towards an epigenetic contribution of sperm for embryonic development. First, in several species paternal exposure to various environmental insults results in changes to the offspring<sup>1–5</sup>. In addition to such environmentally induced epigenetic changes, sperm is also proposed to naturally harbour epigenetic cues that are required for development. Histones have been considered as possible vectors for such epigenetic information. Indeed, in several species including human, modified histones such as H3K4me3 and H3K27me3 are found in the sperm chromatin, in particular around developmental genes and in a pattern reminiscent of that found in embryonic stem cells<sup>6–9</sup>. In a previous work we provided evidence that in the frog, *Xenopus laevis*,

some of these modified histones are homogenously distributed in a population of sperm cells, hence indicating that they have the required attributes for a role in the regulation of embryonic development<sup>10</sup>. Lastly, interference with histone variant or post-translational modifications during the formation of the gametes, or after fertilisation, lead to embryonic gene misregulation<sup>9,11,12</sup>. Although these later experiments support a functional role for sperm-derived modified histone in embryonic development, indirect effect of the interference on the process of spermiogenesis or early embryos development cannot be ruled out. A direct assay to test for the requirement of sperm modified histones in embryo development is still lacking. This is due in

<sup>1</sup>Nantes Universit  , CHU Nantes, INSERM, Center for Research in Transplantation and Translational Immunology, UMR 1064, Nantes, France. <sup>2</sup>Laboratory of Regenerative Medicine, Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, Japan. <sup>3</sup>Systems Biology, Harvard Medical School, Boston, MA, USA. <sup>4</sup>Instituto de Ci  ncias Biom  dicas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil. <sup>5</sup>Univ Rennes, CNRS, IGDR (Institut de G  n  tique et D  veloppement de Rennes) - UMR 6290, Rennes, France. <sup>6</sup>These authors contributed equally: Valentin Francois--Campion, Florian Berger.

 e-mail: [jerome.jullien@inserm.fr](mailto:jerome.jullien@inserm.fr)

part to the difficulty to access the chromatin in the highly compacted sperm nucleus. In this work, we aim to fill this gap and devised a strategy to alter the sperm epigenome just prior to its introduction to the egg. We specifically considered H2AK119ub1 as a possible carrier of necessary epigenetic information from the sperm to the embryos. Indeed, recent works point towards a key role of polycomb repressive complex-1 (PRC1), upstream of PRC2, in the establishment of facultative heterochromatin, both in cultured cells and in embryos<sup>13</sup>. The H2AK119 ubiquitylation activity of the PRC1 complex appears to be integral part of this mechanism<sup>13,14</sup>. In *Xenopus laevis* we found that H2AK119ub1 dynamic during spermiogenesis and early embryonic development is compatible with a role in the epigenetic programming of sperm for embryonic development. We applied an in vitro epigenetic editing assay that uses USP21 deubiquitylase activity<sup>15–18</sup> to erase H2AK119ub1 from sperm chromatin. Embryos derived from such epigenetically edited sperm are developmentally impaired. Transcriptome and epigenome analysis indicates that H2AK119ub1, in combination with H3K4me3, is enriched on a set of embryonic genes misregulated in embryos derived from USP21 treated sperm. Such USP21 sensitive genes preferentially retain H2AK119ub1 upon sperm chromatin exposure to maternal factors. Lastly, we provide evidence that egg factor mediated deposition of H2AK119ub1 on paternal chromatin is influenced by pre-existing sperm H2AK119ub1. We conclude that sperm derived H2AK119ub1 instructs egg factor mediated epigenetic remodelling of paternal chromatin and is required for embryonic development.

## Results

### H2AK119ub1 is present in sperm and early embryo

We first evaluated the overall levels of H2AK119ub1 in chromatin during spermiogenesis and early embryonic development. Indeed, the presence of such modified histone in sperm seems to differ between species, its presence documented in that of mouse<sup>19</sup> while it appears absent in that of zebrafish<sup>20</sup>. In frogs the last stage of spermatogenesis is associated with a remodelling of nucleosomes resulting in a partial loss of H2A/H2B while a full complement of H3/H4 is retained on sperm chromatin (Fig. 1A)<sup>10</sup>. Western Blot (WB) analysis captures this loss of H2A and reveal an overall decrease in H2AK119ub1 in sperm compared to spermatid (Figs. 1B and Suppl Fig. S1-A, see H2A/H4 and H2AK119ub1/H4 ratio, respectively). Importantly however, the H2AK119ub1 to H2A ratio is similar in sperm compared to spermatid, indicating that in this context of spermiogenesis-associated H2A loss, a significant level of H2AK119ub1 is maintained in sperm (H2AK119ub1/H2A, Fig. 1B). We next assessed H2AK119ub1 occurrence after fertilisation: in early frog embryos around ZGA (blastulae stage) and later during gastrulation and neurulation. Interestingly, contrary to what is seen with H3K27me3, a PRC2 associated modification, H2AK119ub1 is detected in early embryogenesis (Fig. 1C). Altogether, these observations highlight H2AK119ub1 credentials for a potential role in the transmission of epigenetic information from sperm to the embryos.

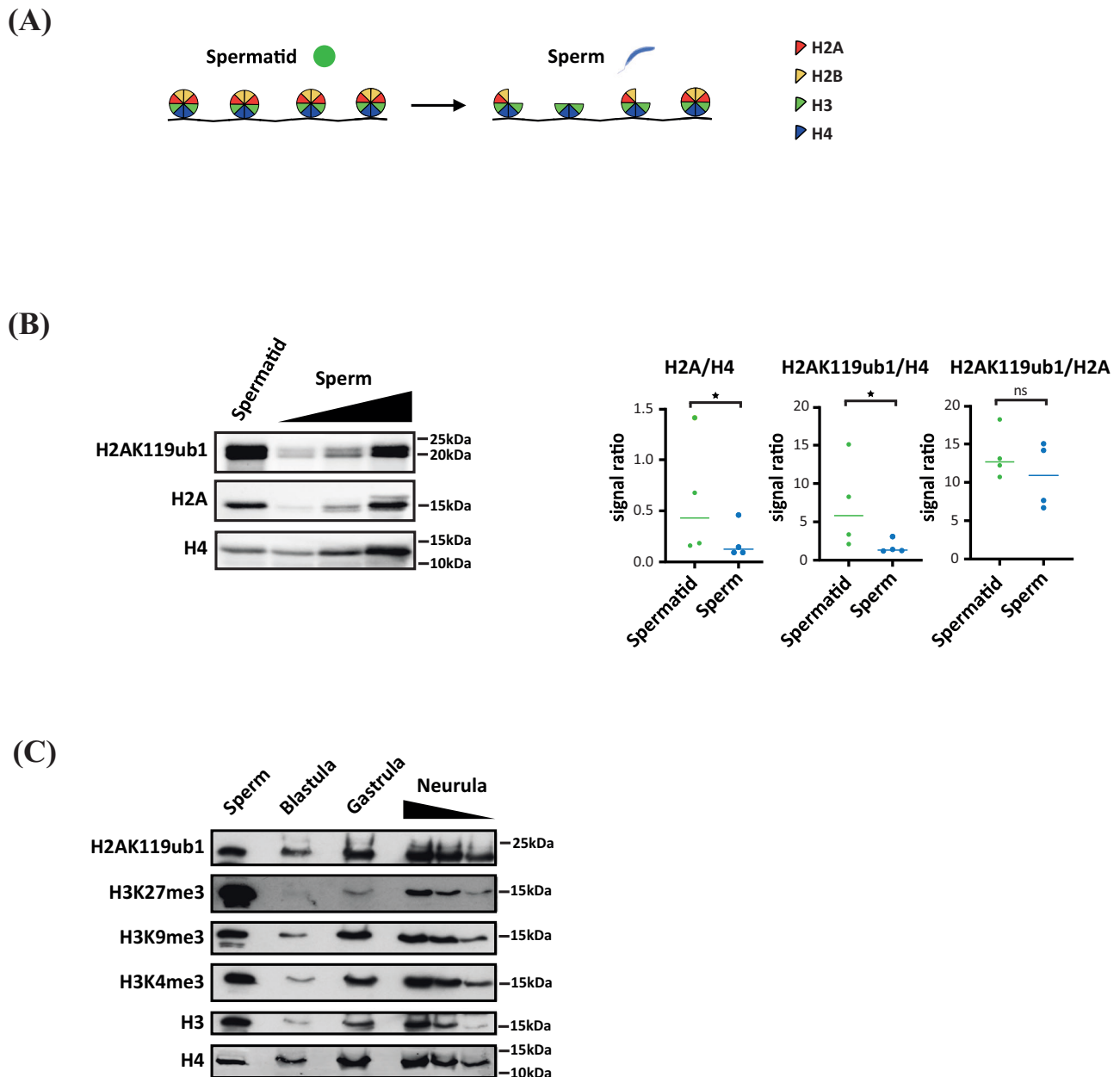
### H2AK119ub1 associates with gene regulatory regions in sperm

We then carried out H2AK119ub1 ChIP to map the distribution of this histone post-translational modification (PTM) on the sperm genome. Biological triplicates were well correlated and subsequently pooled for downstream analysis (Suppl Figs. 2A–D). We first assessed the distribution of H2AK119ub1 around the transcription start of all sperm genes. H2AK119ub1 accumulates around the TSS of a fraction of sperm genes (Fig. 2A), similar to what is observed in somatic cells<sup>21</sup> and embryos<sup>14,20</sup>. Noticeably, we observed in the top 10% TSS harbouring H2AK119ub1 an enrichment for genes associated with developmental function (Fig. 2A, Suppl Data 1), and reminiscent of the association reported with H3K27me3 in many species<sup>6–9</sup>. We next sought to more generally identify regions associated with H2AK119ub1 genome wide. H2AK119ub1 is a pervasive histone modification typically present on

~10% of nuclear H2A in somatic cell<sup>22,23</sup> and up to ~40% in germ cells<sup>19</sup>. In *Xenopus laevis* sperm, it accumulates over region covering larger part of the genome than other modified histones with broad distribution such as H3K27me3 (Suppl Fig. 2E). To identify region of H2AK119ub1 genomic enrichment we therefore used RECOGNICER, a coarse-graining approach designed to capture enrichment over broad scale<sup>24</sup> (Suppl Fig. 3, Data 1). Although we observed peaks of accumulation around genes, such as illustrated for *wnt11L*, *pax6.L* or the *hoxa.S* cluster (Fig. 2B), a large proportion of H2AK119ub1 also accumulates on the sperm genome in distal region away from genes (Fig. 2C). Noticeably we observed that H2AK119ub1 is significantly associated with sperm gene regulatory elements (promoters and enhancers, Fig. 2D). Epigenetic control of gene expression often entails a combination of epigenetic cues such as those observed in X inactivation<sup>25</sup> or in transcriptional reprogramming<sup>26</sup>. We therefore investigated the relationship between genomic regions enriched in sperm for H2AK119ub1 and for other histone post-translational modifications. We focused on H3K27me3<sup>10</sup>, associated with the activity of Polycomb Repressive Complex 2, and H3K4me3<sup>10</sup>, a histone modification associated with transcription (Peaks, e.g. around *pax6.L* and *hoxa.S* loci, Fig. 2B). We then determined the genome-wide association of H2AK119ub1 with methylated K4 and/or K27 on histone H3. While most H2AK119ub1 peaks exist independently of these methylated histones, ~24% of H2AK119ub1 peaks overlap with either H3K4me3 and/or H3K27me3 peaks (Fig. 2E). We observed a very strong promoter bias for peaks of H2AK119ub1 co-occurring with H3K4me3, an expected trend given that H3K4me3 is known to strongly associate with gene promoter (Fig. 2F). By contrast, when occurring in the absence of H3K4me3, H2AK119ub1 is located away from promoter and in intergenic regions. Analysis of fragment length obtained following sequencing of ChIP and Input fractions allow the discrimination of DNA fragment associated with nucleosomes (full complement of histones protecting 150 bp of DNA) from those associated with sub-nucleosomes (nucleosomes lacking one or two H2A-H2B dimer and protecting 110 or 70 bp of DNA) in sperm chromatin (Suppl Fig. 4). Noticeably, in sperm chromatin on the whole, H2AK119ub1 is mostly associated to nucleosomes while H3K27me3 and H3K4me3 tend to be associated with subnucleosomes. So, while we observe co-enrichment of H2AK119ub1 with methylated histone H3 at a subset of genomic loci, those histone modifications are possibly present at those sites on adjacent nucleosomal/subnucleosomal nucleoprotein complexes rather than on the same particle. Altogether, this survey of H2AK119ub1 distribution in sperm reveals its enrichment on gene regulatory region further strengthening its potential for gene regulation across generation.

### Sperm H2AK119ub1 is remodelled by egg factors

To further evaluate how sperm H2AK119ub1 could have an impact on embryonic development, we next sought to follow its fate after fertilisation. Following delivery to the egg, protamines are removed from sperm chromatin and a canonical nucleosomal structure is re-instated. In conjunction with these events the sperm chromatin is replicated. Therefore, the extensive changes to sperm chromatin that follow fertilisation have the potential to profoundly alter the sperm derived modified histone landscape. To determine the fate of sperm modified histones we turned to interphase egg extracts that can mimic the sperm chromatin remodelling events that follow fertilisation, including protamine to histone transition and DNA replication<sup>10</sup>. We carried out H2AK119ub1 and H3K4me3 ChIP-seq on untreated sperm, or sperm having gone through egg extract mediated replication (Fig. 3A). Combining this dataset with existing spermatid and gastrulae ChIP-seq<sup>9,10,27</sup> we obtain a global view of H2AK119ub1, H3K4me3, and H3K27me3 dynamic during spermiogenesis and early embryogenesis (Fig. 3B). Remarkably, while H3K4me3 and H3K27me3 are mostly stable, H2AK119ub1 exhibits drastic redistribution throughout this

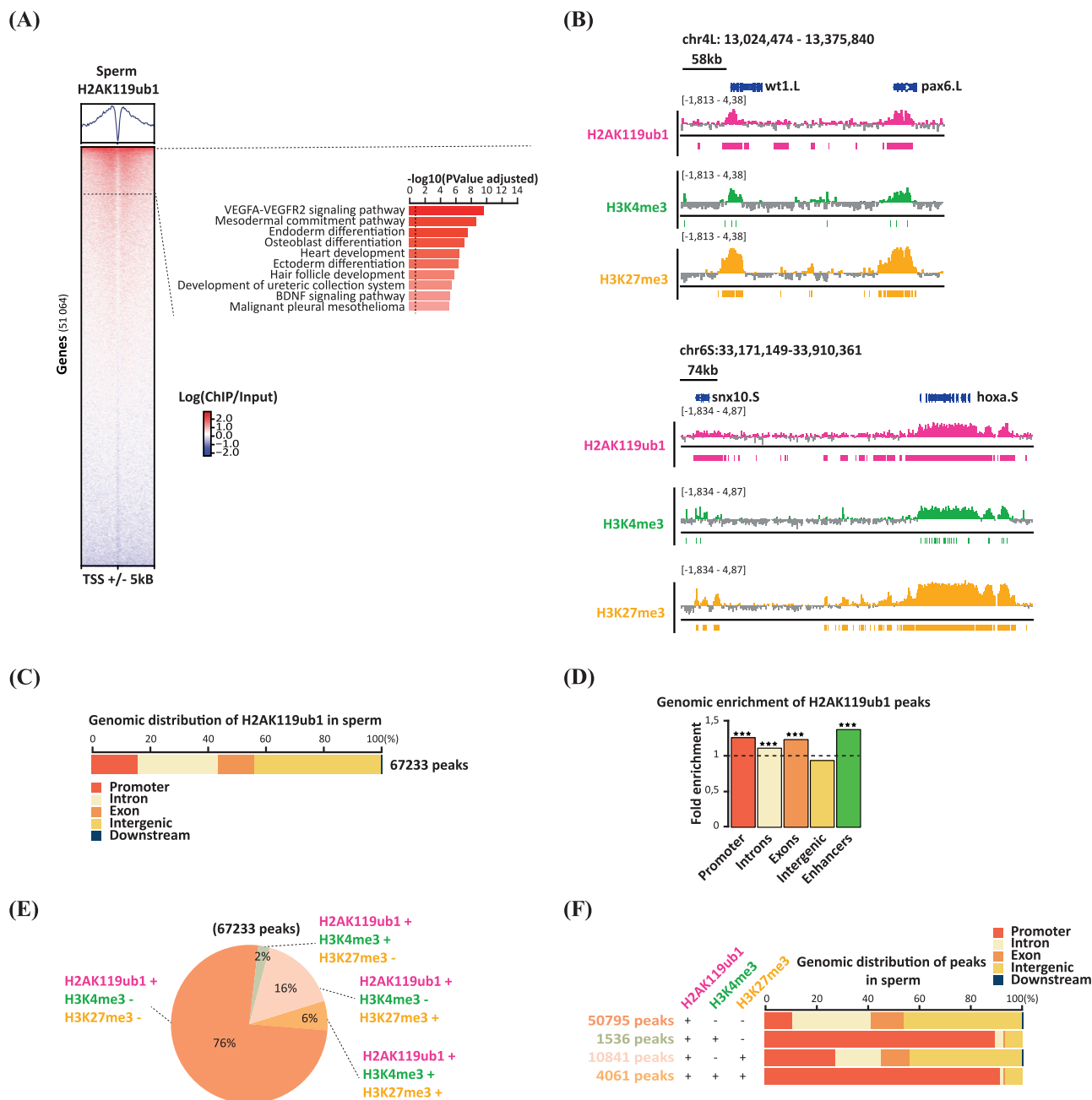


**Fig. 1 | H2AK119ub1 dynamics during spermiogenesis and early embryonic development.** **A** Schematic representation of *Xenopus laevis* nucleosome remodelling during spermiogenesis **(B)** Left, Western blot analysis of H2AK119ub1, H2A and H4 histones. Right quantitation of signal ratio of H2A to H4, H2AK119ub1 to H4, and H2AK119ub1 to H2A, in spermatid and sperm. \*  $p < 0.05$ , two tailed ratio paired t

test ( $n = 4$  independent sperm/spermatid preparations). Median is shown. (H2A/H4,  $p$ -value = 0.0178; H2AK119ub1/H4,  $p$ -value = 0.0354; H2AK119ub1/H2A,  $p$ -value = 0.1634). **C** Western blot analysis of H2AK119ub1, H3K27me3, H3K9me3, H3K4me3, H3 and H4 in *Xenopus laevis* embryo pre- (blastulae) and post- (gastrulae, and neurula) ZGA. Data representative of two independent experiments.

developmental period. Indeed, a large set of genomic location acquires H2AK119ub1 following sperm chromatin replication (cluster 4, Fig. 3B, Suppl Data 2), and includes a large proportion of gene TSSs (Suppl Fig. 5A). To better delineate the principle of this developmental dynamic, we then compared H2AK119ub1 and H3K4me3 peaks in untreated (sperm) *versus* egg extract treated sperm (replicated sperm). We confirm a marked disparity in replication dynamics between sperm H3K4me3, mostly stable (77% of sperm peaks conserved after replication (19345/24982) and 52% stable peaks overall (19345/36770)), and sperm H2AK119ub1, mostly changing (23.5% of sperm peaks maintained after replication (15796/67233) and 15% stable peaks overall (15796/108740)) (Fig. 3C, Suppl Fig. 5B). Most of the replication induced redistribution of H2AK119ub1 happens in

intergenic region, while H2AK119ub1 peaks stable through this process are mostly located on promoter (Suppl Fig. 5D). In this dynamic landscape, the subset of H2AK119ub1 genomic location associated with H3K4me3 in sperm appears selectively conserved (93% of sperm bivalent peaks are maintained after replication (5222/5597) (Fig. 3C). Broader H2AK119ub1 peaks with higher level of H2A retention discriminate location where H2AK119ub1/ H3K4me3 is maintained following exposure to egg factors from those where it is lost (Suppl Fig. 5C). Lastly, we confirmed that the H2AK119ub1 configuration observed following in vitro treatment of sperm with egg extract is also detected in blastulae stage embryo at the time of ZGA (Suppl Fig. 6A-C). We conclude that sperm exposure to egg factors profoundly remodels paternal H2AK119ub1 genomic distribution.



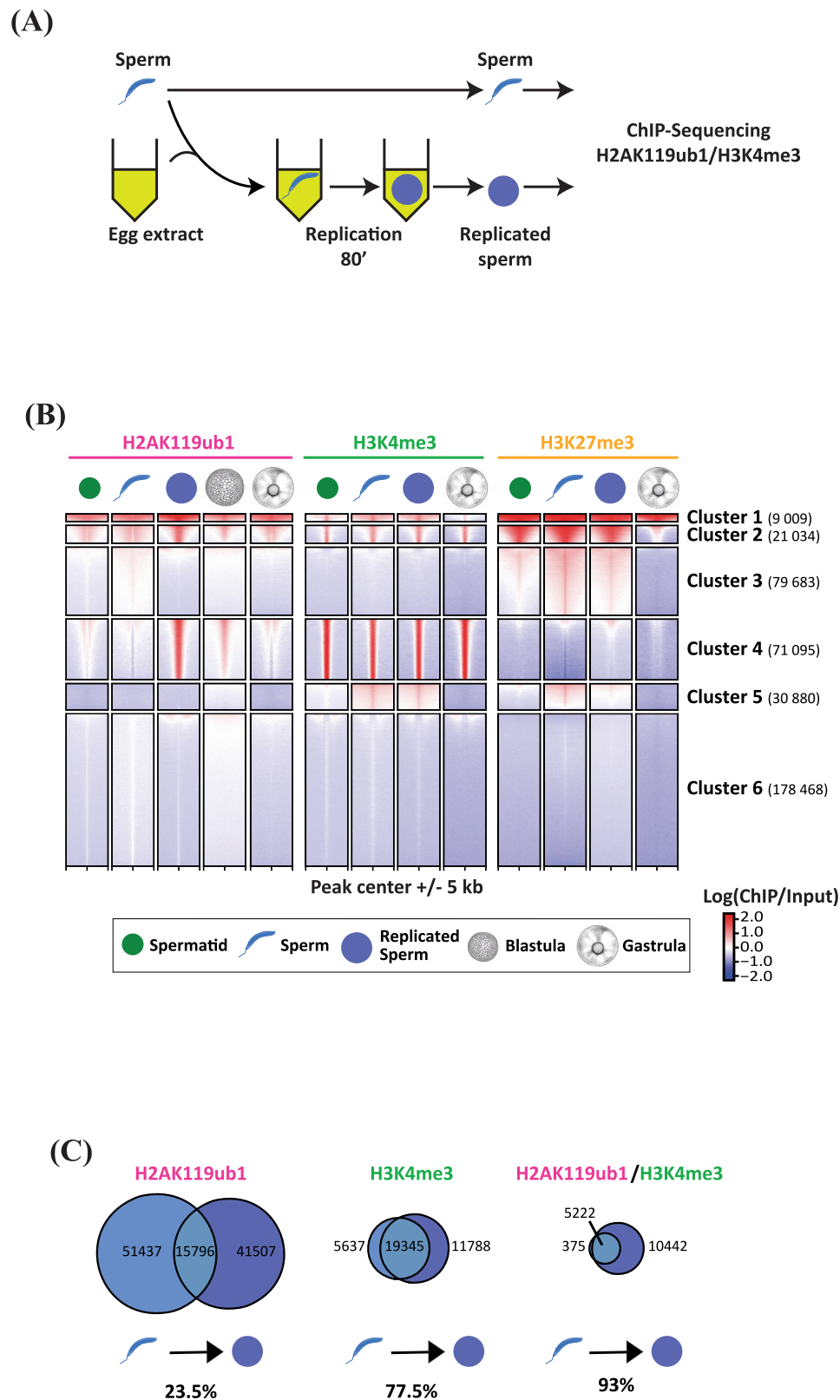
**Fig. 2 | H2AK119ub1 associates with gene regulatory regions in sperm chromatin.** **A** Heatmap of H2AK119ub1 signal (log(ChIP/Input)) on all *Xenopus laevis* genes TSS +/- 5kb. GO enrichment is shown for the top 10% genes. **B** Genome browser view of log(ChIP/Input) of sperm H2AK119ub1, H3K4me3, and H3K27me3 around *wt1.L*, *pax6.L*, *snx10.S* and *hoxa.S* loci. Boxed area underneath the gene track indicates regions of enrichment (peaks). **C** Distribution of sperm H2AK119ub1 peaks

on genomic features. **D** Fold enrichment of H2AK119ub1 peaks on different genomic features, \*\*\*  $p < 0.001$ , two-sided permutation test. **E** Pie chart indicating the proportion of H2AK119ub1 peaks either alone, or overlapping H3K4me3 and/or H3K27me3 peaks in sperm. **F** Genomic distribution of sperm H2AK119ub1 peaks according to overlap with H3K4me3 and/or H3K27me3 peaks.

### Sperm H2AK119ub1 impacts chromatin remodelling by egg factors

We next evaluated to which extent such egg factor-mediated H2AK119ub1 remodelling is dependent on incoming sperm chromatin component. To directly evaluate sperm H2AK119ub1 functional relevance to this process, we set out to erase this histone modification, after completion of spermatogenesis, and prior to exposure to egg factors. To that end, we incubated mature sperm in an oocyte extract containing USP21, a H2A deubiquitylase<sup>45,28</sup> (Fig. 4A). In such extract, the sperm nucleus is partially decondensed, allowing efficient deubiquitylation of H2AK119ub1 while the levels of trimethylated H3K4,

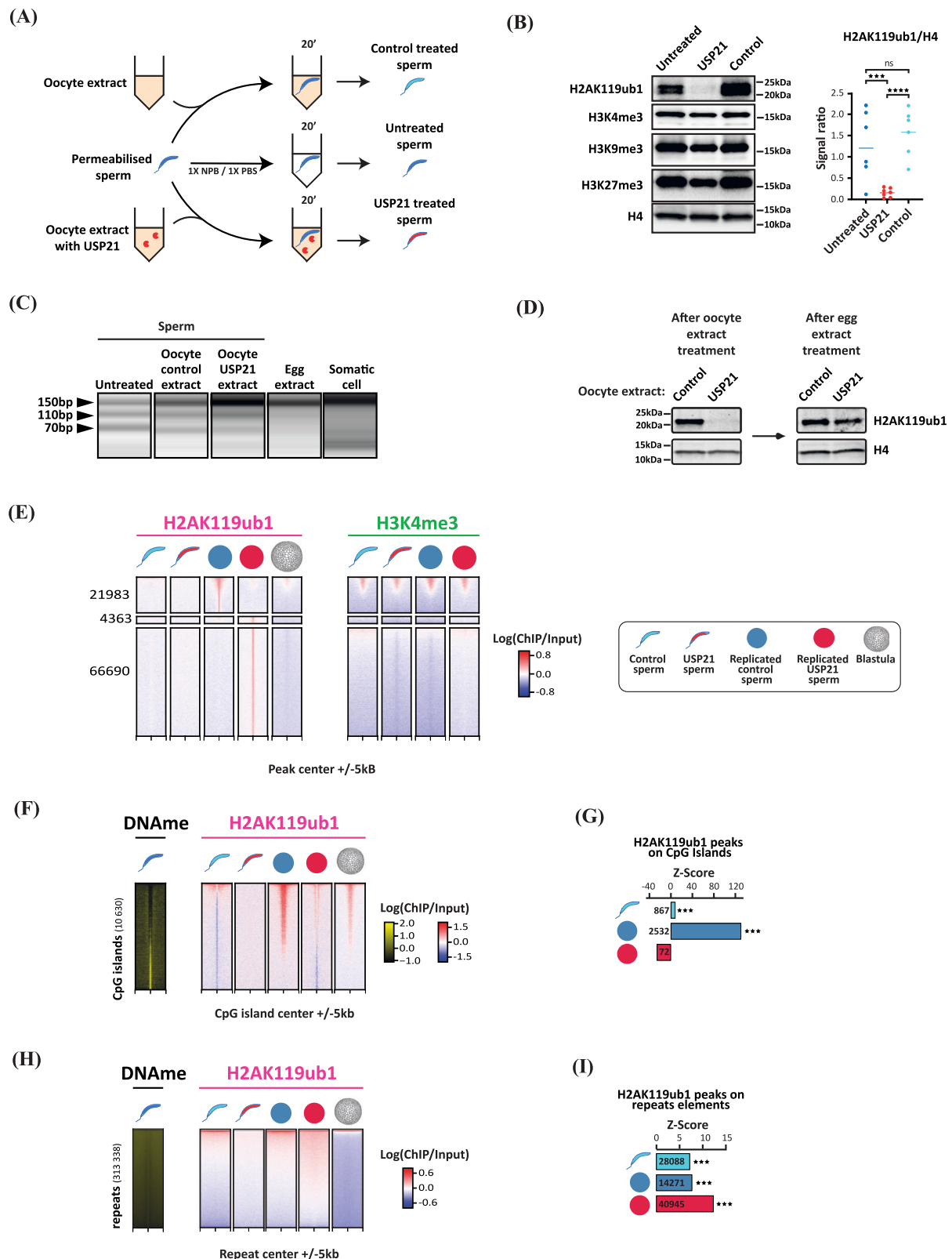
H3K9 and H3K27 remain unchanged (Fig. 4B, Suppl Fig. 7A, Suppl Fig. 1B). Importantly, MNase digestion of USP21 extract-treated sperm yields nucleosomal and subnucleosomal size DNA fragments typical to that of naïve *Xenopus* sperm<sup>9</sup> (Fig. 4C, Suppl Figs. 7B, and 7C **left**). This indicates that the overall sperm chromatin structure is maintained after oocyte extract treatment and does not convert to a somatic type of chromatin organisation as triggered by egg extract treatment<sup>29</sup> (Fig. 4C, Suppl Fig. 7). We next asked to what extent USP21 mediated erasure of H2AK119ub1 from sperm chromatin influences its egg factor mediated redistribution. To that end, we carried out modified histone analysis in sperm treated with control or USP21 extract oocyte with or



**Fig. 3 | Sperm H2AK119ub1 is remodelled following exposure to egg factors.** **A** Experimental design for monitoring modified histones fate following sperm exposure to egg factors. **B** Heatmap of H2AK119ub1, H3K4me3, and H3K27me3 signal (log(ChIP/Input)) in spermatid, sperm, replicated sperm, blastula, and gastrula. All genomic locations enriched for one of these histone modifications

(peaks) in at least one of these developmental stages are represented. **C** Venn diagram indicating the changes to H2AK119ub1 (left), H3K4me3 (middle) or overlapping H3K4me3/H2AK119ub1 (right) peaks during sperm chromatin replication in egg extract. Percentages of sperm peaks that are maintained after replication are indicated underneath.





without a subsequent incubation in egg extract (Suppl Fig. 8A). Interestingly, WB analysis indicates that H2AK119ub1, erased from sperm chromatin by USP21 treatment, is at least partially reinstated by subsequent exposure to egg extract (Fig. 4D), indicating the existence of a maternal H2AK119 ubiquitylation activity independent of pre-existing sperm H2AK119ub1. Using ChIP-seq, we next ask if sperm H2AK119ub1 erasure prior to exposure to egg components affects the maternal

factor mediated remodelling of H2AK119ub1 on paternal chromatin. We first focused on sperm genomic sites where we previously detected accumulation of H3K4me3, H3K27me3, or H2AK119ub1 (as in Fig. 3B). This analysis confirmed that USP21 efficiently erases H2AK119ub1 from its genomic site in sperm (Suppl Figs. 8B, Suppl Data 3). We observed an overall reduction in egg factor mediated H2AK119ub1 deposition on chromatin derived from USP21 treated sperm compared to that from

**Fig. 4 | H2AK119ub1 remodelling by egg factors is altered by sperm pre-treatment with a H2A deubiquitylase.** **A** Schematic representation of sperm treatment with oocyte extract containing USP21. **B** Left, Western blot of H2AK119ub1, H3K4me3, H3K9me3, H3K27me3 and histone H4 on untreated sperm, USP21 extract treated sperm and control extract treated sperm. Right, signal ratio of H2AK119ub1 to H4 in untreated sperm, USP21 treated sperm, or oocyte extract treated sperm. \*\*\*  $p < 0.005$ , \*\*\*\*  $p < 0.001$ , two-sided ratio paired t test ( $n = 6$  independent experiments). Median is shown. (USP1 versus untreated,  $p$ -value = 0.0007; Control versus untreated,  $p$ -value = 0.1519; USP21 versus control  $p$ -value = 0.0001). **C** DNA fragments generated by MNase digestion of *Xenopus laevis* untreated sperm, control extract treated sperm, USP21 extract treated sperm, egg extract treated sperm and somatic cell. Data representative of two independent experiments. **D** Western blot analysis of H2AK119ub1 and histone H4 on sperm after

treatment with control- or USP21- oocyte extract without (left) or with (right) subsequent replication in egg extract. Data representative of two independent experiments. **E** Heatmap of H2AK119ub1 and H3K4me3 signal ( $\log(\text{ChIP}/\text{Input})$ ) in sperm after treatment with control- or USP21- oocyte extract without or with subsequent replication in egg extract. H2AK119ub1 signal is also shown in blastula stage embryo. All genomic locations enriched for H2AK119ub1 (peaks) in replicated samples are represented. **F–I** DNase and H2AK119ub1 signal ( $\log(\text{ChIP}/\text{Input})$ ) on *Xenopus laevis* genome CpG island (**F**) and repeat element (**H**). Z-score of H2AK119ub1 enrichment on CpG island (**G**) and repeat element (**I**) in sperm, replicated sperm and replicated USP21 sperm. H2AK119ub1 signal is also shown in blastula stage embryo. \*\*\*  $p = 0.0001$ , two-sided permutation test. Numbers of H2AK119ub1 peak overlapping the genomic feature are indicated for each cell type.

control sperm (Suppl Fig. 8B). We then compared H2AK119ub1 peaks in samples originating from replicated USP21-treated sperm to that from replicated control-treated sperm (Fig. 4E). Surprisingly, we observed that sperm USP21 treatment leads not only to loss of H2AK119ub1 peaks post-replication (lost peaks) but also triggers the appearance of a large set of peaks that is not deposited by maternal factors in control sperm samples (gained peaks) (Fig. 4E). Comparatively, USP21 treatment impact on post-replication H3K4me3 is much milder (Suppl Fig. 8C). To better characterise this shift, we first evaluated if H2AK119 ubiquitylation sites detected following egg extract exposure of sperm harbour genomic features associated with variant PRC1 (vPRC1) targeting in somatic cell (reviewed in ref. 30). Those H2AK119ub1 sites that are specific to egg extract treated control sperm (lost peaks, Fig. 4E) are associated with cluster 4 in Fig. 3B and are indeed enriched for TF motifs associated with vPRC1 targeting (i.e. USF1/2 E2F, E-box, Suppl Fig. 8D, Suppl Data 2). We also detected a strong association of these sites with unmethylated CpG island (Fig. 4F, G), a genomic feature also shown to attract vPRC1 activity. By contrast, sperm treatment with USP21 prior to egg factor exposure redirects H2AK119ub1 accumulation towards repeat elements and away from such targets (Fig. 4H, I, Suppl Fig. 9, Suppl Data 4). We conclude that maternal factors mediated redistribution of sperm H2AK119ub1 is instructed by pre-existing H2AK119ub1 derived from paternal chromatin. In the absence of sperm H2AK119ub1, maternal factor mediated H2A ubiquitylation activity is shifted from CpG islands to repeat elements.

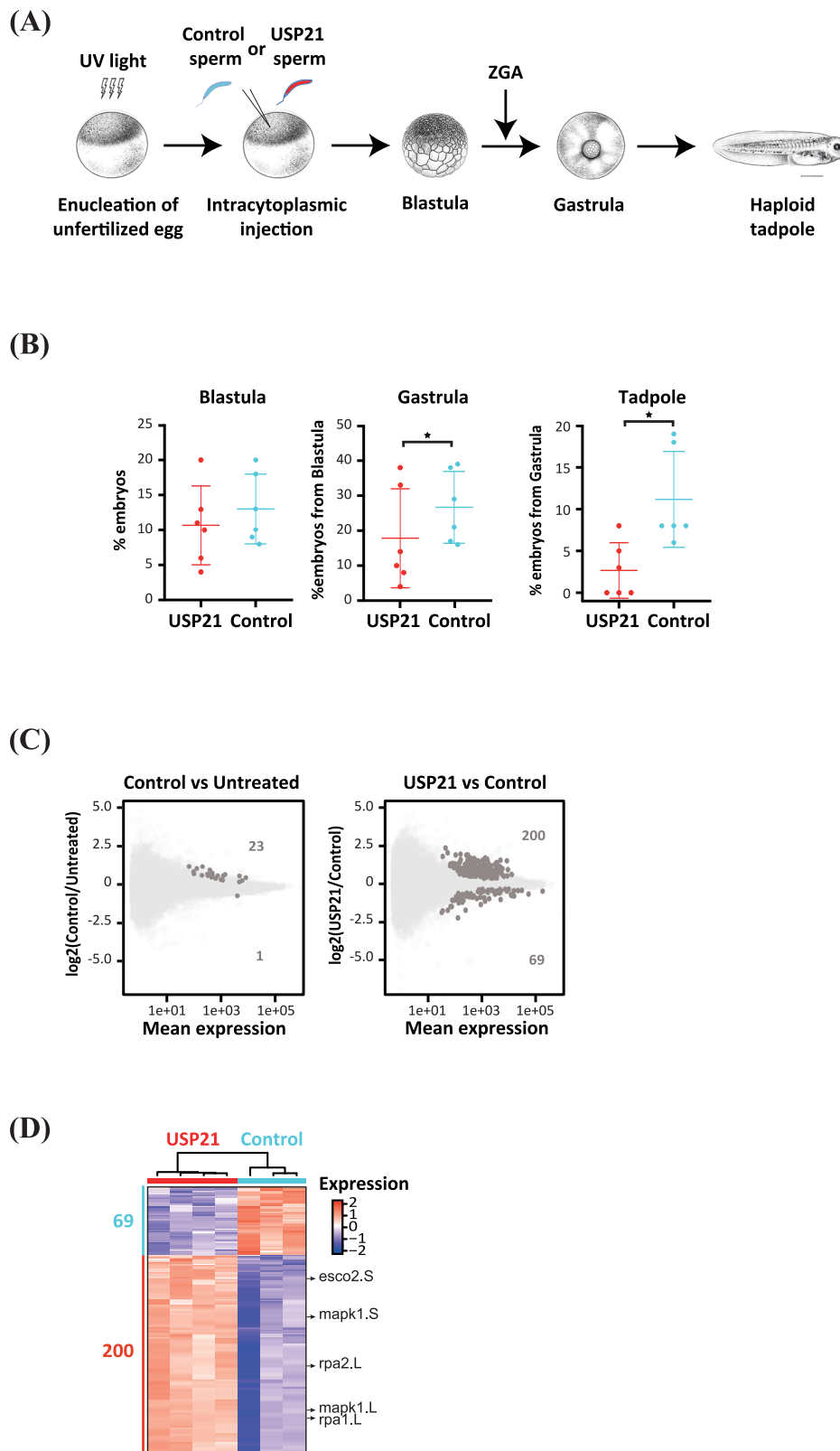
### Sperm H2AK119ub1 erasure alters embryonic gene expression

To then determine the functional relevance of sperm derived H2AK119ub1 to embryo development, we took advantage of the availability of paternally derived embryos in frog. As described previously, such embryos develop solely on chromatin derived from sperm and provide a direct measure of the paternal contribution to early development<sup>9</sup>. We generated paternally derived embryos from sperm pre-incubated in control or USP21 containing oocyte extracts and monitored their development (Fig. 5A)<sup>31</sup>. From gastrulation onwards, we observed a decrease in the developmental potential of USP21 treated- compared to control treated- sperm derived embryos (Fig. 5B). We conclude that besides its DNA content, sperm also contributes to the embryo an USP21 sensitive signal that is necessary for development. The observed defects occur around the time of zygotic gene activation suggesting that USP21 treatment associated developmental failure might relate to defective regulation of embryonic gene expression. To evaluate this possibility, we carried out transcriptome analysis of early gastrulae generated from pool of untreated, control extract-treated or USP21 extract-treated haploid embryos. We performed RNA-seq analysis using sperm, oocyte extract, and egg originating from different preparations for each of the biological replicates. Differential gene expression analysis indicates that the treatment of sperm with a control oocyte extract has minor effect on the resulting embryo transcriptome when compared to embryos generated from untreated sperm (Figs. 5C, left; Suppl Fig. 10A, B). By contrast, USP21

extract treatment leads to embryonic misregulation of 269 genes (Fig. 5C, right), the majority of which are upregulated in embryos derived from USP21 extract treated sperm compared to embryos derived from control extract treated sperm (74%, 200/269, Fig. 5D, Suppl Fig. 10C, Suppl Data 5). Thereafter, this set of genes found differentially expressed in embryos derived from sperm treated with USP21 oocyte will be referred to as USP21 sensitive genes. GO enrichment analysis highlight genes related to cell cycle processes in this gene set (Suppl Data 5). We next asked to which extend the expression pattern of USP21 sensitive genes identified on pooled embryos reflect the behaviour in single embryos. To that end, we carried out transcriptome analysis in single paternally derived gastrulae. We find that the expression pattern of USP21 sensitive gene observed in pooled embryos is replicated in half of the single embryos derived from USP21-treated sperm (4/8 USP21 treated sperm derived embryos, Suppl Fig. 10D). This heterogeneous transcriptome pattern matches the observed developmental outcome, with a fraction of gastrulae derived from USP21-treated sperm developing to the tadpole stage (Fig. 5B) and is reminiscent of the transcriptome heterogeneity detected in cloned embryos<sup>32</sup>. We then interrogated the developmental dynamic of USP21 sensitive genes. We observed a clear distinction between up- and down regulated USP21 sensitive genes (Suppl Fig. 10E). During development of embryos generated by in vitro fertilisation<sup>33</sup>, upregulated genes exhibit a strong maternal contribution followed by clearing prior to gastrulation (stage 10–12) and re-expression at tadpole stage (stage 25–40) while downregulated genes show increased expression around gastrulation. A recent study indicates that at ZGA, transcription occurs on genes with (maternal-zygotic, “MZ genes”) or without (zygotic, “Z genes”) a maternal contribution of transcripts<sup>34,35</sup>. Interestingly, we observed that USP21 sensitive genes are enriched for MZ genes, indicating that these genes exhibit zygotic transcription in the context of a maternally derived transcripts (Suppl Fig. 10E–F). Taken together, these analyses support the hypothesis whereby USP21 mediated erasure of the repressive histone modifications H2AK119ub1 from sperm leads to precocious embryonic expression of a subset of genes that could in turn impact critical embryonic process such as cell division.

### Misregulated genes are H2AK119ub1/H3K4me3 targets in sperm

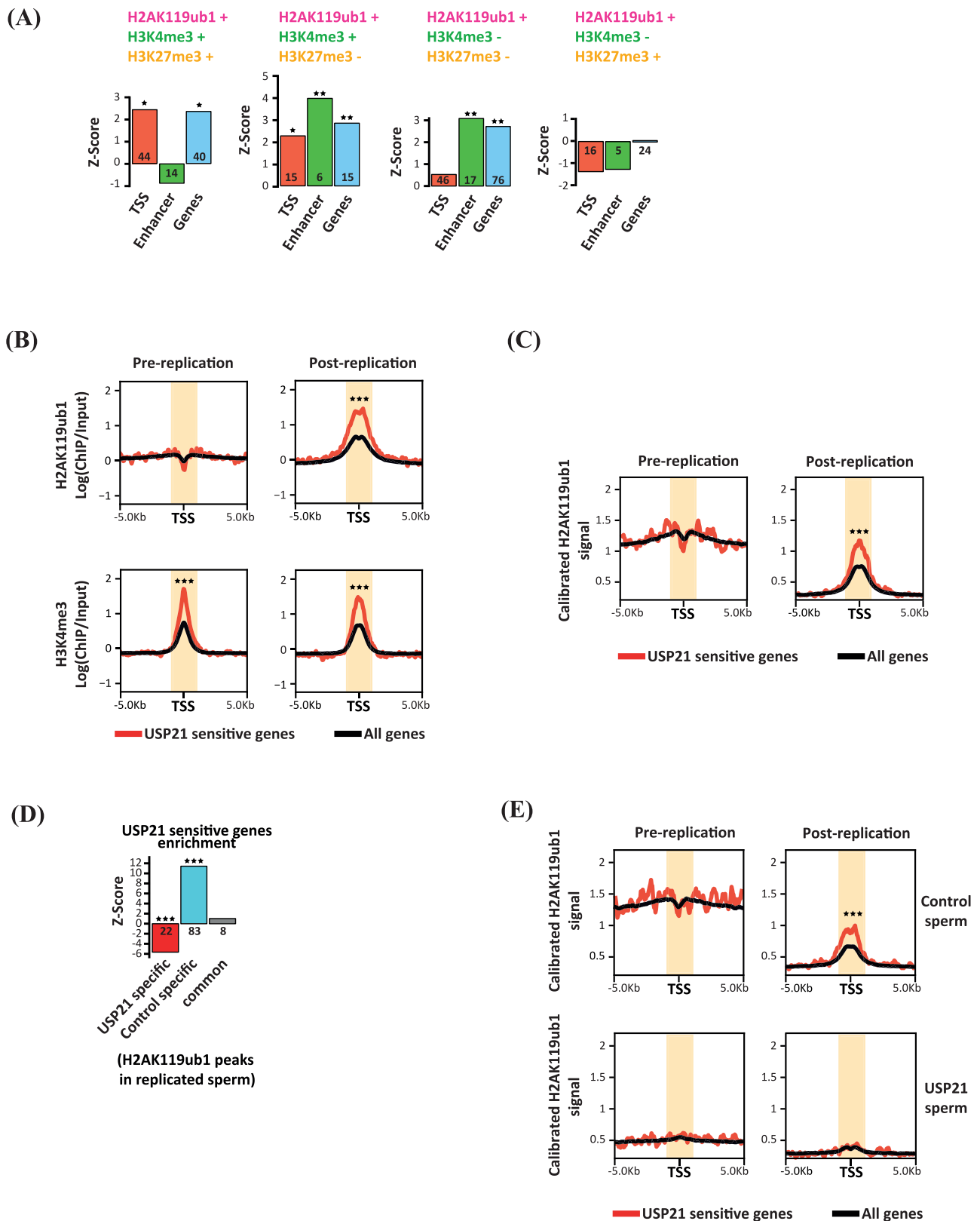
To evaluate this possibility further, we investigated to which extent H2AK119ub1 genomic distribution associate with embryonic gene sensitivity to sperm USP21 treatment. We observed an enrichment of sperm H2AK119ub1 peaks on USP21 sensitive gene body (Suppl Fig. 10G). Additionally, using a genome partitioning approach (hidden Markov model) we also observe that USP21 sensitive genes are associated with genomic regions exhibiting H2AK119ub1 either in sperm or in sperm following exposure to egg extract (Suppl Fig. 10H). However, many more genes harbour this modification in sperm, without being affected by USP21 treatment. We therefore next interrogated the association of USP21 sensitive genes with different types of H2AK119ub1 peaks stratified according to association with other



**Fig. 5 | Sperm pre-treatment with a H2A deubiquitylase decreases developmental potential and leads to embryonic gene misregulation. A** Experimental design for the generation of haploid embryos from sperm or USP21 treated sperm. *Xenopus* illustrations © Natalya Zahn (2022) from Xenbase ([www.xenbase.org](http://www.xenbase.org) RRID:SCR\_003280). **B** Developmental progression to blastula ( $p$ -value = 0.3055), gastrula ( $p$ -value = 0.0454) or swimming tadpole ( $p$ -value = 0.01747) stage of haploid embryos derived from a USP21 or control oocyte extract treated sperm ( $n$  = 6 independent experiments). Two sided Chi-squared test, \*  $p$ -value < 0.05. Median

and SD values are shown. **C** MA-Plot showing as y-axis the log foldchange (Log2FC) of gene expression between control oocyte extract treated sperm *versus* untreated sperm (left) and between USP21 oocyte extract treated sperm derived embryo and control oocyte extract treated sperm (right). x-axis is the averaged gene expression across samples. **D** Heatmap showing expression of differentially expressed gene in gastrulae derived from USP21 oocyte extract treated sperm *versus* that derived from control oocyte extract treated sperm.





histone PTMs, namely H3K4me3 and H3K27me3 (as in Fig. 2E, F). We detect an association of USP21 sensitive gene features with either H2AK119ub1 only peaks (H2AK119ub1+; H3K4me3-; H3K27me3-) or trivalent H2AK119ub1 peaks (H2AK119ub1+; H3K4me3+; H3K27me3+) (Fig. 6A). Strikingly however, USP21 sensitive genes show a strong enrichment for bivalent H2AK119ub1/H3K4me3 and depletion for bivalent H2AK119ub1/H3K27me3 chromatin

configuration. This analysis suggests that the effect on embryonic gene expression of sperm USP21 treatment arise from erasure of H2AK119ub1 in the context of a bivalent H2AK119ub1/H3K4me3 chromatin status. We also observed that, post-replication, H2AK119ub1 and H3K4me3 signals around USP21 sensitive gene TSSs are higher than around the bulk of *X. laevis* genes TSSs (Fig. 6B). Importantly, the significant enrichment of USP21 sensitive genes promoter and gene

**Fig. 6 | Genes sensitive to sperm deubiquitylase treatment are H2AK119ub1 target. A** Z-score of different H2AK119ub1 peaks categories enrichment on TSS, enhancer, and gene body of USP21 sensitive genes. \*  $p$ -value < 0.05; \*\*  $p$ -value < 0.01, two-sided permutation test. Number of H2AK119ub1 peak overlapping the USP21 sensitive gene feature are indicated for each category. **B** H2AK119ub1 and H3K4me3 signal on TSS  $\pm$  5 kb for all genes (black), USP21 sensitive genes (red) before and after sperm replication. \*\*\*  $p$ -value < 0.001, two-sided Wilcoxon-Mann-Whitney test on TSS  $\pm$  1 kb interval (shaded orange area). **C** Calibrated H2AK119ub1 ChIP-seq signal on TSS  $\pm$  5 kb for all genes (black), USP21 sensitive genes (red)

before and after sperm replication. \*\*\*  $p$ -value < 0.001, two-sided Wilcoxon-Mann-Whitney test. **D** Z-score of USP21 sensitive gene enrichment on H2AK119ub1 peaks gained (USP21 specific), lost (Control specific) or maintained (common) in replicated USP21 treated sperm versus replicated control sperm. \*\*\*  $p$ -value < 0.001, two-sided permutation test. Number of USP21 sensitive gene overlapping each category is shown. **E** Calibrated H2AK119ub1 ChIP-seq signal on TSS  $\pm$  5 kb for all genes (black), USP21 sensitive genes (red) before and after replication using a control oocyte extract treated sperm or USP21 treated sperm. \*\*\*  $p$ -value < 0.001, two-sided Wilcoxon-Mann-Whitney test.

body for a bivalent H2AK119ub1/H3K4me3 chromatin configuration reported in sperm holds true after sperm replication (Suppl Fig. 10I). This association of USP21 sensitive genes with H2AK119ub1 is also observed in embryos at ZGA, with such trend still present but less pronounced later during embryogenesis (Suppl Fig. 6D). We then took advantage of a published dataset describing the consequences of expressing the H2A deubiquitylase BAP1 in early mouse embryos to investigate if such a relationship is conserved across species<sup>19</sup>. Using a hidden Markov model, we indeed found that gene identified as BAP1 sensitive in mouse two cell embryos are enriched in genome domains exhibiting H2AK119ub1 marking in sperm (Suppl Fig. 10J, K). Altogether these observations support a role for sperm derived H2AK119ub1 in the regulation of the expression of a subset of embryonic genes.

We next considered the possibility that while a local accumulation of H2AK119ub1 could be maintained post-replication, sperm modified histones could have undergone replicative dilution. In other word, during replication the local accumulation of histone along the genome could be maintained albeit at a reduced level, for example through redistribution of sperm derived modified histones between the replicated strands of DNA. Mechanistically, this is important as in frog -8 rounds of replication separate fertilisation from the main wave of ZGA and could therefore leads to complete loss of sperm derived modified histones. Therefore, to compare H2AK119ub1 level pre- and post-replication we calibrated the signal via semi-synthetic H2AK119ub1 nucleosome spiked-in used in the ChIP-seq (see M&M) (Fig. 6C). We observed a global decrease of H2AK119ub1 signal on TSSs upon egg extract treatment, suggesting that sperm chromatin replication induces redistribution of this modified histones at these locations (black line, Fig. 6C). When considering USP21 sensitive gene however, we observed a different trend. While an overall decrease of H2AK119ub1 signal is also detected upon replication, H2AK119ub1 loss is clearly less pronounced on the TSSs of this subset of gene than seen genome wide (red line, Fig. 6C), a feature shared with the broader set of gene that have maternal-zygotic contribution (green line, Suppl Fig. 10L). These analyses indicate that USP21 sensitive genes are indeed associated with a sperm H2AK119ub1 configuration that persists through egg factor mediated remodelling of the paternal epigenome. Lastly, we asked how sperm USP21 treatment induced disruption of paternal epigenome remodelling by egg factors affects USP21 sensitive genes. We found that the shift of H2AK119ub1 away from CpG towards repeat elements induced by USP21 treatment (as in Fig. 4E) significantly deplete post replication association of USP21 sensitive genes from this histone modification (Fig. 6D). Analysis of gene TSSs also showed that the preferential post replication retention of H2AK119ub1 on USP21 sensitive genes is lost in response to H2AK119ub1 sperm editing (Fig. 6E).

Taken together our results reveal that the setup of paternally derived zygotic chromatin is dictated by component of the sperm epigenome and that this process is required for proper expression of a subset of embryonic genes.

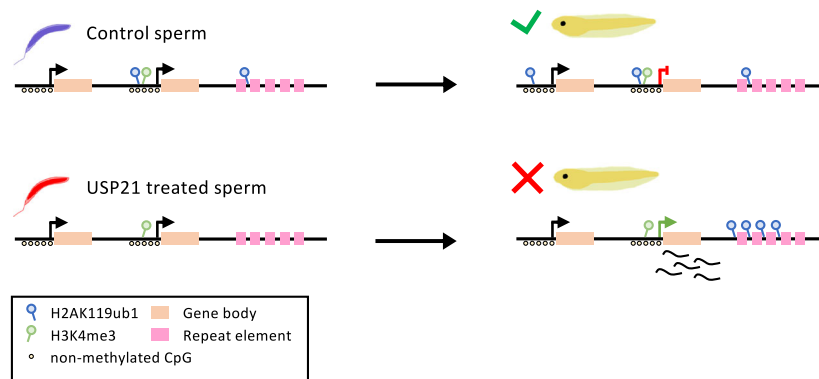
## Discussion

In this work, we found that in *Xenopus laevis*, paternal H2AK119ub1 genomic distribution is extensively remodelled upon exposure to egg

factors, reminiscent of the situation found in the mouse<sup>19</sup>. In this dynamic epigenetic landscape, a subset of genomic location is nevertheless found to maintain H2AK119ub1 following sperm replication in extract as well as following early embryonic division. This suggests that the mechanism recently described in cultured cells and demonstrating recycling of H2AK119ub1 between daughter cells<sup>36</sup> could apply to the transmission of H2A associated epigenetic cue from the gametes to the dividing cell of the frog embryos. Besides those stable sites of H2AK119ub1 accumulation, sperm incubation in egg extract also triggers H2AK119ub1 deposition on new genomic sites. Whether this remodelling requires replication is not known since sperm treatment with egg extract triggers many processes besides replication. Once established those new sites appear stable through replication as they could be detected in late cleavage stage embryos. However, our study is limited by the lack of knowledge about H2AK119ub1 genomic distribution during *xenopus laevis* pre-ZGA developmental period. Particularly, we are lacking in vivo data describing the H2AK119ub1 landscape immediately after fertilisation and during the first round of embryonic divisions. Therefore, mechanistically we don't know if the remodelled H2AK119ub1 distribution found after in vitro sperm replication, and in vivo after 8 round of cell division, is faithfully transmitted throughout cleavage stage or transiently adopt a different configuration during that period.

Egg factor dependent accumulation of H2AK119ub1 occurs around unmethylated CpG island and/or TF binding sites (*i.e.* USF1/2), genomic features that are known target of PRC1 activity in other systems<sup>30</sup>. Besides transcription factors, Kdm2b a component of vPRC1 is involved in more broadly targeting the H2AK119 ubiquitylation activity of this complex to unmethylated CpG island in ESC<sup>30</sup>. Kdm2b is a likely candidate for a role in the observed H2AK119ub1 remodelling following exposure to egg components as it is maternally provided as protein in frog (<https://www.xenbase.org/xenbase/gene/geneExpressionChart.do?method=drawProtein>). In mouse embryonic stem cell, the establishment of polycomb repressive domain occurs in a stepwise manner with a variant PRC1 complex initiating their formation through ubiquitylation of H2AK119, followed by recruitment of the PRC2 complex and subsequent methylation of H3K27me3. Such sequential events have also been described during Neural Progenitor Cell differentiation<sup>37</sup> and X chromosome inactivation<sup>38</sup>. Recently, this sequence of events have been investigated in the mouse and zebrafish embryos<sup>19,20,39</sup>. In these embryos H2AK119ub1 is also required for the deposition of H3K27me3 and transcriptional repression around the time of zygotic gene activation. Our observation that H2AK119ub1 is present in pre-ZGA embryo in the absence of H3K27me3 argues for the existence of a similar mechanism in frog<sup>40,41</sup>.

Interference with H2AK119ub1 in the mouse, either by knocking down PRC1 component<sup>39</sup> in oocyte, or erasing H2AK119ub1 from embryos<sup>19,20</sup> leads to mis-regulation of genes at ZGA and is associated with embryonic lethality. While clearly demonstrating a role for H2AK119ub1 in regulation of transcription around ZGA, these approaches entail prolonged epigenetic interference during either gametogenesis or early embryogenesis and, as a consequence, do not enable to determine the extent to which embryonic H2AK119ub1 is instructed/inherited from sperm and/or oocyte H2AK119ub1<sup>9,11,12,19,20,39</sup>. Here, by editing the sperm epigenome, after completion of



**Fig. 7** | Model of H2AK119ub1 epigenetic programming of sperm for embryonic development in *Xenopus laevis*.

spermiogenesis and prior to exposure to egg factors, we circumvent these limitations. We demonstrate that sperm contribute to the embryo a USP21 sensitive cue required for paternal H2AK119ub1 remodelling by egg factor, embryonic gene expression, and developmental competence (Fig. 7). We find that egg factor mediated accumulation on paternal CpG island depends on H2AK119ub1 delivered by the sperm. Strikingly, we observe that erasure of H2AK119ub1 from sperm leads to a global shift in genomic accumulation from CpG islands to repeat elements. Given the propensity of vPRC1 to be targeted to and to spread from existing H2AK119ub1 location<sup>42</sup> it raises the possibility that pre-existing sperm derived “docking sites” could limit chromatin sampling by the polycomb complex and prevent H2AK119ub1 deposition at genomic location where targeting is usually limited such as the repeat elements. Embryos derived from H2AK119ub1 edited sperm exhibit transcriptional misregulation, mostly upregulation, compatible with the repressive function of this histone PTM on transcription. A similar observation is observed in mouse when H2AK119ub1 deubiquitylase is expressed at fertilisation<sup>19</sup>, or in zebrafish when Rnf2 is chemically inhibited<sup>20</sup>. Given the extent of egg factor mediated H2AK119ub1 accumulation on gene promoters observed in frog, mouse or zebrafish, it is somewhat surprising that only a limited number of genes are mis-regulated when its distribution is massively altered. We identify an association of misregulated genes with H2AK119ub1 in frog sperm and find that this association is preferentially maintained on these genes upon sperm chromatin replication, further supporting a direct role for this histone post-translational modification in the epigenetic programming of sperm for regulation of embryonic transcription. We cannot rule out that a part of the genes is indirectly affected, for example as downstream target of a “primary” H2AK119ub1 target. Alternatively, we provide evidence that gene sensitivity to H2AK119ub1 erasure from sperm occurs when genes are associated with H3K4me3, a histone PTM known to act as a vector of intergenerational information<sup>9–11,43,44</sup>. It is likely that additional epigenetic features besides H3K4me3 contribute to gene sensitivity to H2AK119ub1 erasure from sperm. Additionally, we observe that most affected genes are normally expressed late in embryogenesis (tadpole stage) a feature also seen in mouse with a large proportion of gene sensitive to H2AK119ub1 depletion expressed post-implantation<sup>8</sup>. It is therefore possible that the effect of H2AK119ub1 expression is only partially revealed at early embryonic stage since transcription factors targeting these genes might not be present.

Our work advance knowledge from previous studies by showing that part of H2AK119ub1 found in embryos is paternally instructed. This necessary function of sperm derived modified histones has important bearing on our understanding of inter/trans generational transmission of epigenetic information. One could indeed envisage that such a naturally occurring mechanism of epigenetic information

transmission to the embryos could be hijacked to transmit environmentally induced sperm epigenetic cues. Based on the very dynamic nature of histone PTMs around fertilisation, it is likely that transmission to the embryos will be context- (combination of marks) and location- dependent. Future studies will aim at evaluating to which extent sperm histones PTMs differ between individuals and to which extent they are maintained in embryos.

## Methods

### Contact for reagent and resource sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr J. Jullien ([jerome.jullien@inserm.fr](mailto:jerome.jullien@inserm.fr)).

### Experimental model and subject details

Mature *Xenopus laevis* females were obtained from Tefor Paris-saclays (151 Rte de la Rotonde, 91400 Saclay, France, <https://tefor.net/tefor-paris-saclay-nos-equipes/tefor-paris-saclay-aqua/>). Our work with *Xenopus* complied with all relevant ethical regulations for animal testing and research. Work with frog was approved by the by Ministère de l'Enseignement Supérieur et de la Recherche under the licence #202212131604754/APAFIS 39794V2 and frog husbandry and all experiments were performed according to the relevant regulatory standards. Animals were kept in a recirculating system (ALS, <https://www.aquariumlocationservices.fr/professionnels-elevages-aquatique-aquariophilie/>) at a density of one adult/3 l, with ~10% water change per day. Water was sequentially filtered with mechanical pad sump filter, nitrifying bacteria filter, and UV sterilised. Water quality parameters were as follow: temperature 17–22 °C; PH 6–8. Photoperiod was set to 12 h ON/12 h OFF. Frogs are fed twice per week with NEO MARIN coul 4 - X4 - S20 (Le Gouessant, <https://www.legouessant.com/>). Unconsumed food was removed 30 min after the start of feeding. Animals were used as a source of gametes (oocytes, eggs, and sperm). Embryos generated from those gametes were not sexed.

### Oocyte extract

Fresh ovaries were manually dissociated in 15–25 oocytes clumps. 3–5 mL of dissociated ovary fragments were transferred into 12.5 mL 1X Modified Barth's Saline (MBS, 88 mM NaCl, 1 mM KCl, 10 mM Hepes, 2.4 mM NaHCO<sub>3</sub>, 1.7 mM MgSO<sub>4</sub>, 1.3 mM Ca(NO<sub>3</sub>)<sub>2</sub>, 0.8 mM CaCl<sub>2</sub>) supplemented with 250 µL of 5 mg/mL Liberase (Roche, 05 401 127 001) and incubated for 2h30–3h at room temperature (RT) with gentle rocking. Defolliculated oocytes were washed three times with 1X MBS and healthy stage 5–6 oocytes manually sorted. Oocytes were then injected in the vegetal pole with 25 ng of USP21-NLS-HA mRNA, or control GFP mRNA and incubated for 3 days at 18 °C in 1X MBS, renewing medium and discarding dead cells every day. ~500–1000

oocytes were collected in an Eppendorf tube on ice, the excess 1X MBS discarded and replaced with 1X MBS supplemented with Protease Inhibitor Cocktail (Sigma, P2714-1BTL). The oocytes were then stacked up by centrifugation at 100 g for 1 min at RT and excess medium carefully discarded. Oocytes were then thoroughly vortexed and centrifuged at 21 000 g for 10 min at 4 °C. The clear intermediate phase between yolk (top) and pigment (bottom) was retrieved. The collected fraction was further cleared by an additional centrifugation at 21,000 g for 10 min at 4 °C. The oocyte extract was then aliquoted and stored at -80 °C.

### Oocyte extract treatment

Permeabilized sperm stored in SSB<sup>9</sup> at -80 °C were thawed, washed with two volumes of ice cold 1X PBS and centrifuged at 800 g for 8 min at 4 °C. 2 M sperm in 1X PBS were centrifuged at 21000 × g for 1 min at 4 °C. The supernatant was discarded, and the pelleted sperm resuspended in 10 µL of oocyte extract on ice. The sperm incubated in the oocyte extract for 20–40 min at 23 °C on a heating block. Treatment was stopped by addition of 90 µL of ice cold 1X PBS and centrifugation at 21000 × g for 1 min at 4 °C. The supernatant discarded and the remaining pellets resuspended in SSB for -80 °C long term storage.

### Egg extract treatment

Interphasic extracts were generated as previously described<sup>45</sup>. For the quantification of DNA replication in individual nuclei in egg extract, 33 µM rhodamine dUTP (Roche, 11534378910) was added at the indicated times (20, 40, 60, 80 mins after incubation) for 5 min as indicated, and the reaction was stopped with 1X PBS and fixed in 4% paraformaldehyde. The reaction was spun through a 30% sucrose cushion in PBS. The sperm cells were rinsed twice in 1X PBS for 10 min and mounted in Vectashield mounting medium with DAPI (Vector laboratories, H-1200). Counting of rhodamine positive cells were performed on an epifluorescence microscope.

### Protein extraction and Western Blot analysis

Sperm samples stored in SSB were washed in 1X PBS and 1 M sperm pellets were resuspended in 30 µL 2X Blue Eye (600 mM NaCl, 1 % Triton X100, 0.5 % Na-Deoxycholate, 4 % Sodium Dodecyl Sulphate, 20% Glycerol, 0.0004 % Bromophenol blue, 175 mM Tris-HCl, 10 % Betamercapto ethanol), boiled at 96 °C for 5 min and centrifuged at 21000 × g for 5 min at 4 °C. The supernatant was transferred to a new tube and directly use for analysis or stored at -20 °C. 4–20% mini-PROTEAN<sup>®</sup> TGX (Biorad, #4561094) were loaded with 10 µL of sample and western blot analysis were performed according to standard protocols. For blotting nitrocellulose membranes, a semi-dry transfer was used (Biorad, Trans-Blot<sup>®</sup> SD Semi-Dry Transfer Cell) 30–40 min at 25 V. The following primary antibodies were used: anti-H2A (1:2500, Cell Signalling, 3636), anti-H2AK119ub1 (1:2500, Cell Signalling, 8240), anti-H3K4me3 (1:2500, Abcam, ab8580), anti H3K9me3 (1:2500, Abcam, ab8898), anti-H3K27me3 (1:2500, Cell Signalling, 9735S), anti-H4 (1:2500, Abcam, ab31830) and anti-HA (1:2500, Sigma, H9658). Membranes were incubated with primary antibodies 2 h at RT or overnight (ON) at 4 °C. Secondary antibody are anti-rabbit (1:25000, ThermoFisher, A21076), anti-mouse IgG Alexa Fluor 680 (1:25000, ThermoFisher, A21058), Goat anti-rabbit Alexa 800 (1:25000, ThermoFisher, A32735), or IRDye 800CW Goat Anti-Mouse IgG (H + L) (LI-COR, 926-32210) were used. The membranes were imaged either on a ChemiDoc MP Imaging System (Biorad) or an Odyssey scanner (LI-COR).

### Sperm chromatin immunoprecipitation (ChIP)

Chromatin fractionation and chromatin immunoprecipitation (ChIP) were performed as described before<sup>10,46</sup> with slight modifications. 80 µL of Magnetic beads (M280 Sheep anti-rabbit IgG, Invitrogen,

#1204D) were used per reaction and all wash steps were carried out with a magnet. Beads were washed with 1 mL of TE pH 8.0, lysis buffer (LB) (1:1 mixture of Buffer 1 and MNase buffer, see below), and incubated in 1 mL of LB with 100 µL of 10 mg/mL BSA at 4 °C for 30 min ("pre-blocking"). Pre-blocked beads were washed with 1 mL of LB twice and resuspend in 80 µL of LB. Half of beads were incubated with antibody overnight and leftovers were stored in 4 °C for "pre-clear chromatin" step. For ChIP, the following antibodies were used: anti-H3K4me3 (Abcam, ab8580), anti-H2AK119ub1 (Cell Signalling, 8240). One million digitonin permeabilised sperm nuclei or egg extract treated sperm nuclei were resuspended in 50 µL of Buffer 1 (0.3 M Sucrose, 15 mM Tris pH 7.5, 60 mM KCl, 15 mM NaCl, 5 mM MgCl<sub>2</sub>, 0.1 mM EGTA, 0.5 mM DTT) containing 0.125 ng H2AK119ub biotinylated nucleosome (Epiccypher #16-0020), followed by addition of 50 µL of Buffer 1 with detergent (Buffer 1 including 0.5 % NP40 and 1 % NaDOC). Samples were incubated for 10 min on ice. 100 µL of MNase buffer (0.3 M Sucrose, 85 mM Tris, 3 mM MgCl<sub>2</sub>, 2 mM CaCl<sub>2</sub>, 2.5 U of micrococcal nuclease: Roche 10107921001) was added in to each tube (0.5 million of cells per tube). Tubes were incubated at 37 °C for 30 min in pre-warmed water bath. Reaction was stopped by adding 2 µL of 0.5 M EDTA pH 8.0 in the same order as started. Tubes were vortexed and placed on ice for at least 5 min. Supernatant and pellet were separated by centrifugation at 13000 × g for 10 min at room temperature. Beads (stored at 4 °C for pre-clear chromatin) were washed with 1 mL of LB twice. Supernatant and EDTA-free Protease Inhibitor Cocktail were added into the beads and rotating on the wheel at 4 °C for 60 min. 10% of volume of pre-cleared chromatin was taken as "Input" and stored at 4 °C. Antibody conjugated beads were washed with 1 mL of LB twice. Pre-cleared chromatin was added to washed beads and incubated at 4 °C for 6 h. After incubation, beads were washed with 1 mL of Washing buffer A (50 mM Tris-HCl pH 7.5, 10 mM EDTA and 75 mM NaCl) on the wheel at 4 °C for 5 min and buffer was discarded carefully. Subsequently, beads were washed with 1 mL of Washing buffer B (50 mM Tris-HCl pH 7.5, 10 mM EDTA and 125 mM NaCl) and buffer was discarded carefully. Another 1 mL of Washing buffer B was added and transferred to 1.5 mL Protein lo-bind tubes (Eppendorf, 0030108116) with beads. Buffer was discarded carefully. 150 µL of elution buffer (1:9 mixture of 10 % SDS and TE) was added. Tubes were placed at 25 °C on the Eppendorf thermomixer and shaken at 800 rpm for 15 min. Supernatant were taken and transferred to the new tube using magnetic stand. Another 150 µL of elution buffer was added and repeated this step (in total 300 µL). Input samples were diluted to 300 µL total by adding TE. 2 µL of 2 mg/mL RNase A (heat-inactivated or DNase-free) was added and incubated at 37 °C for 30 min. 4 µL of 10 mg/mL Proteinase K was added and incubated at 55 °C overnight. After incubation, tubes were shortly spun. 300 µL of Phenol/chloroform were added and vortexed for 30 sec. Samples were centrifuged at 16000 × g for 10 min at room temperature and aqueous phase collected. 1 µL of Glycogen, 1/10 volume of 3 M NaAcetate pH 5.2 and 2 volumes of 100% EtOH were added and vortexed. Tubes were stored on dry ice for 20–30 min. Once frozen, tubes were centrifuged at 16000 × g for 30 min at 4 °C. Supernatant was discarded carefully and 500 µL of 70% EtOH was added. Tubes were centrifuged at 16000 × g for 10 min at 4 °C. Pellets were dried and suspended in 30 µL of ddH<sub>2</sub>O.

### Embryo chromatin immunoprecipitation (ChIP)

Batches of 200 in vitro fertilised embryos were fixed in 15 ml 1 % Paraformaldehyde in 1X MMR for 20 min at room temperature, followed by 4 washes with 1X MMR. Embryos were split in two and equilibrated at 4 °C overnight in 500 µL HEG solution (50 mM HEPES-KOH pH 7.5, 1 mM EDTA, 20 % Glycerol), then excess buffer was discarded, and samples were transferred to -80 °C. For chromatin extraction embryos were resuspended in 700 µL of sonication buffer (20 mM Tris-HCl pH 8.0, 70 mM KCl, 1 mM EDTA pH 8.0, 10% Glycerol, 5 mM DTT, 0.125% NP40, 1X complete protease inhibitors) and homogenised by pipetting



up and down. 200  $\mu$ L embryo extract were distributed in 1.5 mL tubes and sonicated in two sets of 15 min cycles (with 30 s on/off cycles). Sonicated samples were centrifuged at 16000  $\times g$  for 5 min at 4 °C. Supernatants were retrieved and the centrifugation was repeated. Before ChIP, 180  $\mu$ L (2  $\times$  10%) of the chromatin extract was taken as an Input control and stored at -20 °C. For each ChIP, 720  $\mu$ L of chromatin extract was mixed with 2  $\mu$ L of anti-H2AK119ub1 antibodies (Cell Signalling, 8240) and incubated overnight on a rotating wheel. The next day, 40  $\mu$ L (20  $\mu$ L per IP) of Magnetic beads (M280 Sheep anti-rabbit IgG, Invitrogen, #11204D) were washed with 1X PBS using a magnet and incubated with 1 mL of 1X PBS containing 1 mg/mL BSA for 2 h at 4 °C on a rotating wheel. Blocking solution was discarded from the beads using a magnet and beads were resuspended in 200  $\mu$ L (100  $\mu$ L per IP) of incubation buffer (50 mM Tris-HCl pH 8, 100 mM NaCl, 2 mM EDTA, 1% NP40, 1 mM DTT, 1X Protease inhibitor). 100  $\mu$ L of incubation buffer containing beads is transferred to chromatin/antibody mix and incubated for 4 h at 4 °C. Samples were washed with Buffer 1 (50 mM Tris-HCl pH 8, 100 mM NaCl, 2 mM EDTA, 1 mM DTT, 1 % NP40, 0.1 % sodium deoxycholate, 1X complete protease inhibitors) for 5 min at 4 °C on rotating wheel. Using a magnet, supernatant was discarded and beads subsequently washed with Buffer 2 (50 mM Tris-HCl pH 8, 500 mM NaCl, 2 mM EDTA, 1 mM DTT, 1 % NP40, 0.1 % sodium deoxycholate, 1X complete protease inhibitors), Buffer 3 (50 mM Tris-HCl pH 8, 100 mM NaCl, 250 mM LiCl, 2 mM EDTA, 1 mM DTT, 1 % NP40, 0.1 % sodium deoxycholate, 1X complete protease inhibitors), Buffer 1 after which immunoprecipitated chromatin was eluted using 400  $\mu$ L (310  $\mu$ L for Input fraction) of Elution buffer (0.1 M NaHCO<sub>3</sub>, 1 % SDS) overnight at 65 °C. The next day, 4  $\mu$ L of 20 mg/mL RNase A was added to the chromatin solution and incubated for 2 h at 37 °C. After incubation, tubes were shortly spun. 300  $\mu$ L of Phenol/chloroform were added and vortexed for 30 sec. Samples were centrifuged at 16000  $\times g$  for 10 min at room temperature and aqueous phase collected. 25  $\mu$ g of Glycogen, 1/10 volume of NaAcetate and 2 volumes of 100% EtOH were added and vortexed. Tubes were stored on dry ice for 20–30 min. Once frozen, tubes were centrifuged at 16000  $\times g$  for 30 min at 4 °C. Supernatant was discarded carefully and 500  $\mu$ L of 70% EtOH was added. Tubes were centrifuged at 16000  $\times g$  for 10 min at 4 °C. Pellets were dried and suspended in 30  $\mu$ L of ddH<sub>2</sub>O.

#### **Xenopus sperm collection**

*Xenopus* sperm collection was performed as described before<sup>10,46</sup>. For each round of sperm purification, testes from six adult *Xenopus laevis* males were isolated and manually cleaned from blood vessels and fat bodies in 1X Marc's Modified Ringers (MMR, 100 mM NaCl, 2 mM KCl, 1 mM MgSO<sub>4</sub>, 2 mM CaCl<sub>2</sub>, 3 mM HEPES, pH 7.4) using forceps and paper tissues. Subsequently, testes were torn into small pieces with forceps and homogenised with 2–3 strokes of a Dounce homogeniser (one testis at a time). The cell suspension was then filtered to remove tissue debris and cell clumps using a 60  $\mu$ m Mesh (CellTrics, cat. 04-0042-2317) and spun down at 800 rcf, 4 °C, for 20 min. Supernatant was discarded and the cell pellet was resuspended in 12 mL of 1  $\times$  MMR. If any red blood cells were visible at the bottom of the pellet (as a result of incomplete removal of blood vessels), only the uncontaminated part of the pellet was recovered, taking extreme care not to disturb the red blood cells. Subsequently, step gradients of iodixanol (Optiprep; Sigma, D1556; 60% iodixanol in water) in 1X MMR final were manually prepared in pre-chilled 14 mL ultra-clear centrifuge tubes (Beckman Coulter, #344060) in the following order from the bottom to the top of the tube: 4 mL of 30 % iodixanol, 1 mL of 20 % iodixanol, 5 mL of 12 % iodixanol (all in 1  $\times$  MMR), and 2 mL of cell suspension in 1X MMR on top. Gradients were spun down in a pre-chilled SW40Ti rotor at 10000  $\times g$  4 °C, for 15 min, deceleration without brake (Beckman Coulter Ultracentrifuge, Optima L-100XP). The pelleted fraction, containing mature sperm, was collected. Collected fractions were diluted six times with 1X MMR and collected by spinning first at 805 rcf, 4 °C,

for 20 min and then at 3220 rcf, 4 °C, for 20 min to pellet remaining cells. Pelleted cells were subsequently permeabilized with Digitonin (10 mg/mL as a final concentration, Sigma, D141) for 5 min at room temperature as detailed before (Smith et al., 2006).

#### **Library preparation and sequencing**

ChIP samples were subjected to ChIP-seq library preparation using QIAseq Ultralow Input Lib UDI-A Kit (Qiagen, 180497) and Agencourt AMPure XP beads (Beckman coulter, A63881). Samples pre- and post-library protocol were analysed on a LabChip GX (PerkinElmer) using DNA 1 KHS and 1K Assay LabChIP, respectively. Libraries were quantitated using NEBNext Library Quant Kit (NEB, E7630). For RNA sequencing, RNAs were purified using Qiagen RNeasy Mini Kit (#74004) and used to generate cDNA sequencing libraries using an Illumina TruSeq Kit (#RS-122-2001), according to the manufacturer's specifications. Paired-end sequencing was carried out either on a NOVA-seq 6000 or HiSeq 1500 sequencer (Illumina). ChIP sample sequencing data are summarised in Suppl Data 6.

#### **RNA-seq analysis**

Adaptors were trimmed and low-quality reads filtered using cutadapt (v1.18)<sup>47</sup>. Sample quality was confirmed with fastqc (v0.11.8)<sup>48</sup>. fastqc were aligned to *Xenopus laevis* genome (v9.2) using STAR aligner (v2.7.3.a)<sup>49</sup>. Count table was generated using HTSeq-count (v0.12.3)<sup>50</sup> and exon as genes reference. MultiQC (v1.9)<sup>51</sup> summarised data pre-processing. Differential gene expression analysis of the count table was performed on R (v3.6)<sup>52</sup> using DESeq2tools<sup>53</sup>. Details of the pipeline used can be found at <https://gitlab.univ-nantes.fr/E114424Z/BulkRNAseq>. Gene ontology analysis were carried out with gprofiler<sup>54</sup> using human orthologs of *Xenopus laevis* genes recovered from Xenbase (<https://www.xenbase.org/xenbase/static-xenbase/ftpDatafiles.jsp>)<sup>55</sup>. *Xenopus laevis* development RNA-seq data were from published dataset GSE73430.

#### **ChIP-seq analysis**

Adaptors were trimmed and low-quality reads filtered using cutadapt (v1.18)<sup>47</sup>. Sample quality was confirmed with fastqc (v0.11.9)<sup>48</sup>. Alignment was performed with bwa mem<sup>56</sup> against *Xenopus laevis* (v9.2) or against a *Xenopus laevis* (v9.2) - 601DNA hybrid genome generated with bwa index (v0.7.17-r1188). Reads extraction and analysis was performed with samtools view and samtools flagstat (v1.12)<sup>57</sup>. ChIP-Seq sample quality was assessed by Deeptools (v3.5.1)<sup>58</sup> using plotFingerprint function and bamPEFragmentSize. Replicates quality was visualised with plotCorrelation using spearman correlation after genome fragmentation into bins of 10Kb. PCR duplicates were marked by Picard tools using MarkDuplicates command (<https://broadinstitute.github.io/picard>) and then removed with samtools view -F 0  $\times$  400. Replicates were pooled by subsampling the same amount of reads per replicates using sambamba (v0.8.1)<sup>59</sup> and samtools view (-s option). Peak calling was processed with MACS2 (<https://github.com/macs3-project/MACS>)<sup>60</sup> (-q 0.01 --nomodel --extsize 147 --gsize=2.6e9) for H3K4me3, with MACS2 broad option for H3K27me3 (--broad --broad-cutoff 0.1 -q 0.01 --nomodel --extsize 147 --gsize=2.6e9) and RECOGNICER<sup>24</sup> (qvalue 0.001) /MACS2 (--broad --broad-cutoff 0.1 -q 0.01 --nomodel --extsize 147) for H2AK119ub1. SICER peak calling was performed using the following parameters -s XL -w 200 -rt 1 -f 150 -egf 0.74 -fdr 0.001 -g 600 -e 1000. Peaks overlap was defined using bedtools intersect having at least 1bp of overlap between peaks. H2AK119ub1 signal was calibrated using reads mapping to the 601DNA and corresponding to the fully ubiquitylated H2AK119 nucleosome spiked-in in samples prior to ChIP (see ChIP M&M section). To evaluate difference in modified histone ChIP-seq signal around the TSS of gene sets, Wilcoxon Mann-Whitney test was performed on the  $\pm$ 1 kb window around the TSSs. ChIPSeeker<sup>61</sup> and ChIPPeakAnno<sup>62,63</sup> were used to annotate peaks on R<sup>52</sup>. The RegioneR<sup>64</sup> package was used to perform



peaks permutation and assess enrichment on USP21 sensitive gene features (*i.e.* TSS, enhancer, or gene body) compared to a random set of genes (using 10000 iterations of the `resampleRegions` function for Figs. 6A, D, S3C two right panels, S3E, S10G, S10I) and to assess peak enrichment on genomic features (using 1000 iterations of the `randomizeRegions` function, for Figs. 2D, 4G, I, S3C two left panels, S10H, S10K). `Log10 (ChIP/Input)` was computed using `Deeptools` with `Bam-Compare` using 50 bp binsize. Heatmap and TSS plot were produced with `computeMatrix` and `plotHeatmap` or `plotProfile`. IGV was used for peaks and `log10(ChIP/Input)` signal visualisation. Heatmap were produced with `computeMatrix` using 50 bp bin size  $\pm 5$  kb of TSS or peak centre. Heatmaps were generated with `plotHeatmap` function using unsupervised clustering (`-hclust` option), and `metaplot` using `plotProfile` function. IGV<sup>65</sup> was used to visualise output signal tracks. `Bedtools` `getfasta`<sup>66</sup> was used to extract FASTA sequences corresponding to peaks from a selected heatmap cluster. To identify transcription factors motif in the selected peak set, `Homer`<sup>67</sup> was then used with the `findMotifs.pl` function and all other peaks of the heatmap as a background set. CpG islands and repeats elements were recovered from UCSC genome annotation database (<https://hgdownload.soe.ucsc.edu/goldenPath/xenLae2/database/>). For genome partitioning approach, `chromHMM` model was initialised using `BinarizeBam` on ChIP and Input sample file, using binsize of 150 bp (`-b 150`). The model was trained using a random initiation (`-init random`), over binsize of 150 bp (`-b 150`) and a maximum of 5000 iterations (`-r 5000`). `SNPsplit` was used for reads assignment to paternal and maternal copies as described in ref. 19.

## Reporting summary

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

## Data availability

The *Xenopus laevis* ChIP-seq and RNA-seq data generated for this study are referenced as [PRJEB56442](https://www.ebi.ac.uk/ena/browser/view/PRJEB56442) at the European Nucleotide Archive (ENA). Additional dataset used are from GSE73430 (*Xenopus* embryo RNA-seq), GSE125982 (sperm and replicated sperm H3K27me3 ChIP-seq), GSE126599 (Gastrulae H3K27me3 and H3K4me3 ChIP-seq), and GSE75164 (sperm MBD-seq, spermatid H3K27me3), and GSE153531 (H2AK119ub1Cut&Run mouse embryos and gametes). Source data are provided with this paper.

## Code availability

All scripts used for these analyses are available on Github (<https://github.com/FCValentin/H2AubPaper>) and linked to Zenodo (<https://doi.org/10.5281/zenodo.14725112>).

## References

- Öst, A. et al. Paternal diet defines offspring chromatin state and intergenerational obesity. *Cell* <https://doi.org/10.1016/j.cell.2014.11.005> (2014).
- Dias, B. G. & Ressler, K. J. Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat. Neurosci.* <https://doi.org/10.1038/nn.3594> (2014).
- Radford, E. J. et al. In utero undernourishment perturbs the adult sperm methylome and is linked to metabolic disease transmission. *Sci. (80-)* **345**, 1–26 (2015).
- Ben Maamar, M., Beck, D., Nilsson, E. E., Kubsad, D. & Skinner, M. K. Epigenome-wide association study for glyphosate induced trans-generational sperm DNA methylation and histone retention epigenetic biomarkers for disease. *Epigenetics* **16**, 1150–1167 (2021).
- Lismer, A., Siklenka, K., Lafleur, C., Dumeaux, V. & Kimmins, S. Sperm histone H3 lysine 4 trimethylation is altered in a genetic mouse model of transgenerational epigenetic inheritance. *Nucleic Acids Res* **48**, 11380–11393 (2020).
- U., B. et al. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat. Struct. Mol. Biol.* **17**, 679–687 (2010).
- Hammoud, S. S. et al. Distinctive chromatin in human sperm packages genes for embryo development. *Nature* <https://doi.org/10.1038/nature08162> (2009).
- Wu, S. F., Zhang, H. & Cairns, B. R. Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. *Genome Res.* <https://doi.org/10.1101/gr.113167.110> (2011).
- Teperek, M. et al. Sperm is epigenetically programmed to regulate gene transcription in embryos. *Genome Res* **26**, 1034–1046 (2016).
- Oikawa, M. et al. Epigenetic homogeneity in histone methylation underlies sperm programming for embryonic transcription. *Nat. Commun.* <https://doi.org/10.1038/s41467-020-17238-w> (2020).
- Siklenka, K. et al. Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. *Sci. (80-)* <https://doi.org/10.1126/science.aab2006> (2015).
- 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* <https://doi.org/10.1016/j.cell.2018.01.022> (2018).
- Blackledge, N. P. et al. Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **157**, 1445–1459 (2014).
- Tamburri, S. et al. Histone H2AK119 mono-ubiquitination is essential for polycomb-mediated transcriptional repression. *Mol. Cell* **77**, 840–856.e5 (2020).
- Nakagawa, T. et al. Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation. *Genes Dev.* <https://doi.org/10.1101/gad.1609708> (2008).
- Joo, H. Y. et al. Regulation of cell cycle progression and gene expression by H2A deubiquitination. *Nature* <https://doi.org/10.1038/nature06256> (2007).
- Cao, J. & Yan, Q. Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. *Front. Oncol.* <https://doi.org/10.3389/fonc.2012.00026> (2012).
- Chen, D., Dai, C. & Jiang, Y. Histone H2A and H2B deubiquitinase in developmental disease and cancer. *Cancer Transl. Med.* <https://doi.org/10.4103/2395-3977.168578> (2015).
- 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).
- Hickey, G. J. M. et al. Establishment of developmental gene silencing by ordered polycomb complex recruitment in early zebrafish embryos. *Elife* **11**, e67738 (2022).
- Conway, E. et al. BAP1 enhances Polycomb repression by counter-acting widespread H2AK119ub1 deposition and chromatin condensation. *Mol. Cell* **81**, 3526–3541.e8 (2021).
- Albright, S. C., Nelson, P. P. & Garrard, W. T. Histone molar ratios among different electrophoretic forms of mono- and dinucleosomes. *J. Biol. Chem.* **254**, 1065–1073 (1979).
- Matsui, S., Seon, B. K. & Sandberg, A. A. Disappearance of a structural chromatin protein A24 in mitosis: Implications for molecular basis of chromatin condensation. *Proc. Natl Acad. Sci. USA* **76**, 6386–6390 (1979).
- Zang, C., Wang, Y. & Peng, W. RECOGNICER: a coarse-graining approach for identifying broad domains from ChIP-seq data. *Quant. Biol.* **8**, 359–368 (2020).
- Heard, E. & Distech, C. M. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev.* **20**, 1848–1867 (2006).
- Jullien, J. et al. Gene resistance to transcriptional reprogramming following nuclear transfer is directly mediated by multiple chromatin-repressive pathways. *Mol. Cell* **65**, 873–884.e8 (2017).

27. Kuznetsov, J. N. et al. BAP1 regulates epigenetic switch from pluripotency to differentiation in developmental lineages giving rise to BAP1-mutant cancers. *Sci. Adv.* **5**, eaax1738 (2019).
28. Jin, J. et al. The deubiquitinase USP21 maintains the stemness of mouse embryonic stem cells via stabilization of Nanog. *Nat. Commun.* <https://doi.org/10.1038/ncomms13594> (2016).
29. Gaggioli, V., Le Viet, B., Germe, T. & Hyrien, O. DNA topoisomerase II $\alpha$  controls replication origin cluster licensing and firing time in *Xenopus* egg extracts. *Nucleic Acids Res.* **41**, 7313–7331 (2013).
30. Blackledge, N. P. & Klose, R. J. The molecular principles of gene regulation by Polycomb repressive complexes. *Nat. Rev. Mol. Cell Biol.* **22**, 815–833 (2021).
31. Zahn, N., Levin, M. & Adams, D. S. The zahn drawings: new illustrations of *xenopus* embryo and tadpole stages for studies of craniofacial development. *Dev.* **144**, 2708–2713 (2017).
32. 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).
33. Session, A. M. et al. Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature* **538**, 336–343 (2016).
34. Chen, H. & Good, M. C. Imaging nascent transcription in whole-mount vertebrate embryos to characterize zygotic genome activation. in *Methods in Enzymol.* **638**, 139–165 (2020).
35. 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).
36. Flury, V. et al. Recycling of modified H2A-H2B provides short-term memory of chromatin states. *Cell* **186**, 1050–1065.e19 (2023).
37. Tsuboi, M. et al. Ubiquitination-independent repression of PRC1 targets during neuronal fate restriction in the developing mouse neocortex. *Dev. Cell* **47**, 758–772 (2018).
38. Żylicz, J. J. et al. The implication of early chromatin changes in X chromosome inactivation. *Cell* **176**, 182–197.e23 (2019).
39. Mei, H. et al. H2AK119ub1 guides maternal inheritance and zygotic deposition of H3K27me3 in mouse embryos. *Nat. Genet.* **53**, 539–550 (2021).
40. Van Heeringen, S. J. et al. Principles of nucleation of h3k27 methylation during embryonic development. *Genome Res.* <https://doi.org/10.1101/gr.159608.113> (2014).
41. Akkers, R. C. et al. A Hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Dev. Cell* <https://doi.org/10.1016/j.devcel.2009.08.005> (2009).
42. Zhao, J. et al. RYBP/YAF2-PRC1 complexes and histone H1-dependent chromatin compaction mediate propagation of H2AK119ub1 during cell division. *Nat. Cell Biol.* **22**, 439–452 (2020).
43. Lismer, A. et al. Histone H3 lysine 4 trimethylation in sperm is transmitted to the embryo and associated with diet-induced phenotypes in the offspring. *Dev. Cell* **56**, 671–686.e6 (2021).
44. Pepin, A. S., Lafleur, C., Lambrot, R., Dumeaux, V. & Kimmins, S. Sperm histone H3 lysine 4 tri-methylation serves as a metabolic sensor of paternal obesity and is associated with the inheritance of metabolic dysfunction. *Mol. Metab.* **59**, 101463 (2022).
45. Hannak, E. & Heald, R. Investigating mitotic spindle assembly and function in vitro using *Xenopus laevis* egg extracts. *Nat. Protoc.* **1**, 2305–2314 (2006).
46. Teperek, M. et al. Sperm and spermatids contain different proteins and bind distinct egg factors. *Int. J. Mol. Sci.* **15**, 16719–16740 (2014).
47. Martin, M. et al. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, (2011).
48. Andrews, S. & others. FastQC: a quality control tool for high throughput sequence data. 2010. <https://www.Bioinformatics.Babraham.ac.uk/Projects/Fastqc/> (2019).
49. Dobin, A. et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
50. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
51. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
52. Dessau, R. B. & Phipper, C. B. [“R”-project for statistical computing]. *Ugeskr. Laeger* **170**, (2008).
53. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
54. Raudvere, U. et al. G:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).
55. Fisher, M. et al. Xenbase: Key features and resources of the *Xenopus* model organism knowledgebase. *Genetics* **224**, iyad018 (2023).
56. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
57. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Giga-science* **10**, giab008 (2021).
58. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
59. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
60. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
61. Wang, Q. et al. Exploring epigenomic datasets by ChIPseeker. *Curr. Protoc.* **2**, e585 (2022).
62. Zhu, L. J. et al. ChIPpeakAnno: a bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinform.* **11**, 237 (2010).
63. Zhu, L. J. Integrative analysis of ChIP-chip and ChIP-seq dataset. *Methods Mol. Biol.* **1067**, 105–124 (2013).
64. Gel, B. et al. RegioneR: an R/bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics* **32**, 289–291 (2016).
65. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
66. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
67. Heinz, S. et al. simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).

## Acknowledgements

J.J. group is funded by ANR-21-CE13-0031-01 grant from the Agence Nationale de la Recherche and a Connect Talent grant from Ville de Nantes and Region pays de la Loire. M.O. is funded by a postdoctoral fellowship from the Japan Society for the Promotion of Science (JSPS). R.G. was supported for this work by a Career Development Award grant (CDA00019/2019-C) from the Human Frontier Science Programme. L.P. group is supported by NIH award R24 OD031956. We would like to acknowledge the support of Xenbase.

## Author contributions

Conceptualisation: J.J.; Methodology: J.J., V.F.-C., F.B., L.P.; Validation: J.J., V.F.-C., F.B.; Formal analysis: V.F.-C., F.B., M.O., M.G., V.C.; Investigation: J.J., V.F.-C., F.B., N.M., C.F., J.P., R.G., M.O., B.H., K.P., J.G.A.; Writing original draft: J.J.; Writing review and editing: J.J., F.B., V.F.C. Visualisation: F.B., V.F.-C.; Supervision: J.J.; Project administration: J.J.; Funding acquisition: J.J.

## 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/s41467-025-58615-7>.

**Correspondence** and requests for materials should be addressed to Jérôme Jullien.

**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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2025