Sperm derived H2AK119ub1 is required for embryonic development in *Xenopus Laevis*

Valentin Francois--Campion[1,^], Florian Berger[1,^], Mami Oikawa[2], Maissa Goumeidane[1], Romain Gibeaux[3], and Jérôme Jullien[1,*]

AFFILIATIONS

[1] Nantes Université, CHU Nantes, INSERM, Center for Research in Transplantation and Translational Immunology, UMR 1064, F-44000 Nantes, France

[2] Laboratory of Regenerative Medicine, Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392, Japan

[3] Univ Rennes, CNRS, IGDR (Institut de Génétique et Développement de Rennes) - UMR 6290, F-35000 Rennes, France

^equal contribution

*Corresponding author email: jerome.jullien@inserm.fr

ABSTRACT [140words]

Deposition of 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 retained 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.

INTRODUCTION

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 [1–5]. 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[6–9]. 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[10]. Lastly, interference with histone variant or post-translational modifications during the formation of the gametes, or after fertilisation, lead to embryonic gene misregulation[9,11,12]. 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 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 points towards a key role of polycomb repressive complexe-1 (PRC1), upstream of PRC2, in the establishment of facultative heterochromatin, both in cultured cells and in embryos[13]. The H2AK119 ubiquitylation activity of the PRC1 complex appears to be integral part of this mechanism[13,14]. 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 activiy [15–18] 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

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[19] while it appears absent in that of zebrafish[20]. 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**)[10]. Western Blot (WB) analysis captures this loss of H2A and reveal an overall decrease in H2AK119ub1 in sperm compared to spermatid (**Fig. 1B**, see H2A/H4 and H2AK119ub/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 throughout 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.

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 Fig.1A-B**). 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. 1D**), similar to what is observed in somatic cells[21] and embryos[14,20]. Noticeably, we observed in the top 10% TSS harbouring H2AK119ub1 an enrichment for genes associated with developmental function (**Fig.1D, Table S1**), and reminiscent of the association reported with H3K27me3 in many species[6–9]. 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[22,23] and up to ~40% in germ cells[19]. In *Xenopus* sperm, it accumulates over region covering larger part of the genome than other modified histones with broad distribution such as H3K27me3 (**Suppl Fig.1C**). To identify region of H2AK119ub1 genomic enrichment we therefore used RECOGNICER, a coarse-graining approach designed to capture enrichment over broad scale[24] (**Suppl Fig.1C-F, Table S1**). Although we observed peaks of accumulation around genes, such as illustrated for Wnt1.L, Pax6.L or the Hoxa.S cluster (**Fig. 1E)**, a large proportion of H2AK119ub1 also accumulates on the sperm genome in distal region away from genes (**Fig. 1F**). Noticeably we observed that H2AK119ub1 is significantly associated with sperm gene regulatory elements (TSSs and enhancers (**Fig. 1G**)). Epigenetic control of gene expression often entails a combination of epigenetic cues such as those observed in X inactivation[25] or in transcriptional reprogramming[26]. We therefore investigated the relationship between genomic regions enriched in sperm for H2AK119ub1 and for other histone post-translational modifications. We focused on H3K27me3[10], associated with the activity of Polycomb Repressive Complex 1, and H3K4me3[10], a histone modification associated with transcription (Peaks, *e.g.* around Pax6.L and HoxA loci, **Fig. 1E**). 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. 1H**). We observed a very strong promoter bias for peaks of H2AK119ub1 co-occurring with methylated H3K4me3, an expected trend given that H3K4me3 is known to strongly associate with gene promoter (**Fig. 1I**). By contrast, when occurring in the absence of H3K4me3, H2AK119ub1 is located away from promoter and in intergenic regions. Importantly, in sperm chromatin on the whole, H2AK119ub1 is associated to different type of nucleoprotein complexes than that associated with H3K27me3 and H3K4me3 (**Suppl Fig.1G**). 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.

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 fertilization 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 fertilization[10]. We carried out H2AK119ub1 and H3K4me3 ChIP-seq on untreated sperm, or sperm having gone through egg extract mediated replication (**Fig. 2A**). Combining this dataset with existing spermatid and gastrulae ChIP-seq[9,10,27] we obtain a global view of H2AK119ub1, H3K4me3, and H3K27me3 dynamic during spermiogenesis and early embryogenesis (**Fig.2B**). Remarkably, while H3K4me3 and H3K27me3 are mostly stable, H2AK119ub1 exhibits drastic redistribution throughout this developmental period. Indeed, a large set of genomic location acquires H2AK119ub1 following sperm chromatin replication (cluster 4, **Fig.2B**), and includes a large proportion of gene TSSs (**Suppl Fig. 2A**). 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 51% stable peaks overall (19345/37747)), and sperm H2AK119ub1, mostly changing (23.5% of sperm peaks maintained after replication (15796/67233) and 15% stable peaks overall (15796/104956)) (**Fig. 2C, Suppl Fig. 2B)**. 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. 2D**). 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. 2C**). Broader H2AK119ub1peaks with higher level of H2A retention discriminate location where H2AK119ub1/H3K4me3 is maintained through replication from those where it is lost (**Suppl Fig. 2C**). We conclude that sperm exposure to egg factors profoundly remodels paternal H2AK119ub1 genomic distribution.

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[15,28] (**Fig.2 D**). 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. 2E, Suppl Fig. 2E**). Importantly, MNAse digestion of oocyte extract USP21-treated sperm yields nucleosomal and subnucleosomal size DNA fragments typical to that of naïve *Xenopus* sperm[9] (**Fig.2F**, lane 1&2; **Suppl Fig. 2F**). 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[29] (**Fig.2F,** lane 3&4**; Suppl Fig. 2F)**. We next asked to what extend 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

188 with control or USP21 extract oocyte with or without a subsequent incubation in egg
189 extract (**Suppl Fig. 2G**). Interestingly, WB analysis indicates that H2AK119ub1,
190 erased from sperm chromatin by USP21 treatment, is at least partially reinstated by
191 subsequent exposure to egg extract (**Fig. 2G**), indicating the existence of a maternal
192 H2AK119 ubiquitylation activity independent of pre-existing sperm H2AK119ub1.
193 Using ChIP-seq, we next ask if sperm H2AK119ub1 erasure prior to exposure to egg
194 factors affects the post-replication genomic location of H2AK119ub1 on paternal
195 chromatin. We first focused on sperm genomic sites where we previously detected
196 accumulation of H3K4me3, H3K27me3, or H2AK119ub1 (as in **Fig. 2B)**. This analysis
197 confirmed that USP21 efficiently erases H2AK119ub1 from its genomic site in sperm
198 (**Suppl Fig. 2H**). We observed an overall reduction in egg factor mediated
199 H2AK119ub1 deposition on chromatin derived from USP21+ sperm compared to that
200 from USP21-sperm (**Suppl Fig. 2H**). We then compared H2AK119ub1 peaks in
201 samples originating from replicated USP21-treated sperm to that from replicated
202 control-treated sperm (**Fig. 2H**). Surprisingly, we observed that sperm USP21 treatment
203 leads not only to loss of H2AK119ub1 peaks post-replication (lost peaks) but also
204 triggers the appearance of a large set of peaks that is not deposited by maternal factors
205 in control sperm samples (gained peaks) (**Fig. 2H**). Comparatively, USP21 treatment
206 impact on postreplication H3K4me3 is much milder (**Suppl Fig. 2I**). To better
207 characterize this shift, we first evaluated if H2AK119 ubiquitylation sites detected
208 following egg extract exposure of sperm harbour genomic features associated with
209 vPRC1 targeting in somatic cell (reviewed in[30]). H2AK119ub1 sites that are specific to
210 egg extract treated control sperm (lost peaks, **Fig. 2H**) are indeed enriched for TF motif
211 associated with vPRC1 targeting (*i.e.* USF1/2 E2F, E-box, **Fig. 2I, Table S1**). We also
212 detected a strong association of these sites with unmethylated CpG island (**Fig. 2J-K**),
213 a genomic feature also shown to attract vPRC1 activity. By contrast, sperm treatment
214 with USP21 prior to egg factor exposure redirects H2AK119ub1 accumulation towards
215 repeat elements and away from such targets (**Fig. 2L-M**). We conclude that maternal
216 factors mediated redistribution of sperm H2AK119ub1 is instructed by pre-existing
217 H2AK119ub1 derived from paternal chromatin. In the absence of sperm H2AK119ub1,
218 maternal factor mediated H2A ubiquitylation activity is shifted from CpG islands to
219 repeat elements.

220 To then determine the functional relevance of sperm derived H2AK119ub1 to
221 embryo development, we took advantage of the availability of paternally derived
222 embryos in frog. As described previously, such embryos develop solely on chromatin
223 derived from sperm and provide a direct measure of the paternal contribution to early
224 development[9]. We generated paternally derived embryos from sperm pre-incubated in
225 control or USP21 containing oocyte extracts and monitored their development
226 (**Fig.3A**)[31]. From gastrulation onwards, we observed a decrease in the developmental
227 potential of USP21 treated- compared to control treated- sperm derived embryos
228 (**Fig.3B**). We conclude that besides its DNA content, sperm also contributes to the
229 embryo an USP21 sensitive signal that is necessary for development. The observed
230 defects occur around the time of zygotic gene activation suggesting that USP21
231 treatment associated developmental failure might relate to defective regulation of
232 embryonic gene expression. To evaluate this possibility, we carried out transcriptome
233 analysis of early gastrulae generated from pool of untreated, control extract-treated or
234 USP21 extract-treated haploid embryos. We performed RNA-seq analysis using sperm,
235 oocyte extract, and egg originating from different preparations for each of the biological
236 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 (**Fig.3C**, left; **Suppl Fig. 3A-B**). By contrast, USP21 extract treatment leads to embryonic misregulation of 269 genes (**Fig.3C**, 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.3D, Suppl Fig. 3C, Table S2**). 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 (**Table S2**). 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. 3D**). 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.3B)** and is reminiscent of the transcriptome heterogeneity detected in cloned embryos[32]. 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. 3E**). During development of embryos generated by *in vitro* fertilisation[33], 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[34,35]. 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. 3E-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.

To evaluate this possibility further, we instigated 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. 3G**). 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 histone PTMs, namely H3K4me3 and H3K27me3 (as in **Fig.1G**). 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. 3E**). 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

286 (**Fig. 3F**). Importantly, the significant enrichment of USP21 sensitive genes promoter
287 and gene body for a bivalent H2AK119ub1/H3K4me3 chromatin configuration
288 reported in sperm holds true after sperm replication (**Suppl Fig. 3H**). We next
289 considered the possibility that while a local accumulation of H2AK119ub1 could be
290 maintained post-replication, sperm modified histones could have undergone replicative
291 dilution. In other word, during replication the local accumulation of histone along the
292 genome could be maintained albeit at a reduced level, for example through
293 redistribution of sperm derived modified histones between the replicated strands of
294 DNA. Mechanistically, this is important as in frog ~8 rounds of replication separate
295 fertilisation form the main wave of ZGA and could therefore leads to complete loss of
296 sperm derived modified histones. Therefore, to compare H2AK119ub1 level pre- and
297 post- replication we calibrated the signal *via* semi-synthetic H2AK119ub1 nucleosome
298 spiked-in used in the ChIP-seq (see M&M) (**Fig. 3G**). We observed a global decrease
299 of H2AK119ub1 signal on TSSs upon egg extract treatment, suggesting that sperm
300 chromatin replication induces redistribution of this modified histones at these locations
301 (black line, **Fig. 3G**). When considering USP21 sensitive gene however, we observed
302 a different trend. While an overall decrease of H2AK119ub1 signal is also detected
303 upon replication, H2AK119ub1 loss is clearly less pronounced on the TSSs of this
304 subset of gene than seen genome wide (**Fig. 3G),** a feature shared with the broader set
305 of gene that have maternal-zygotic contribution (**Suppl Fig. 3I**). These analyses
306 indicate that USP21 sensitive genes are indeed associated with a sperm H2AK119ub1
307 configuration that persists through egg factor mediated remodelling of the paternal
308 epigenome. Lastly, we asked how sperm USP21 treatment induced disruption of
309 paternal epigenome remodelling by egg factors affects USP21 sensitive genes. We
310 found that the shift of H2AK119ub1 away from CpG towards repeat elements induced
311 by USP21 treatment (as in **Fig. 2H**) significantly depleted post replication association
312 of USP21 sensitive genes from this histone modification (**fig. 3H**). Analysis of gene
313 TSSs also showed that the preferential post replication retention of H2AK119ub1 on
314 USP21 sensitive genes is lost in response to H2AK119ub1 sperm editing (**Fig. 3I**).

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

318

319 **DISCUSSION**

320 In this work, we found that in *Xenopus Laevis*, paternal H2AK119ub1 genomic
321 distribution is extensively remodelled upon exposure to egg factors, reminiscent of the
322 situation found in the mouse[19]. This remodelling is characterized by an accumulation
323 of H2AK119ub1 around unmethylated CpG island and/or TF binding sites (*i.e.*
324 USF1/2), genomic features that are known target of PRC1 activity. In mouse embryonic
325 stem cell, the establishment of polycomb repressive domain occurs in a stepwise
326 manner with a variant PRC1 complex initiating their formation through ubiquitylation
327 of H2AK119, followed by recruitment of the PRC2 complex and subsequent
328 methylation of H3K27me3. Such sequential events have also been described during
329 Neural Progenitor Cell differentiation[36] and X chromosome inactivation[37]. Recently,
330 this sequence of events have been investigated in the mouse and zebrafish
331 embryos[19,20,38]. In these embryos H2AK119ub1 is also required for the deposition of
332 H3K27me3 and transcriptional repression around the time of zygotic gene activation.

333 Our observation that H2AK119ub1 is present in pre-ZGA embryo in the absence of
334 H3K27me3 argues for the existence of a similar mechanism in frog[39,40]. Interference
335 with H2AK119ub1 in the mouse, either by knocking down PRC1 component[38] in
336 oocyte, or erasing H2AK119ub1 from embryos[19,20] leads to mis-regulation of genes at
337 ZGA and is associated with embryonic lethality. While clearly demonstrating a role for
338 H2AK119ub1 in regulation of transcription around ZGA, these approaches entail
339 prolonged epigenetic interference during either gametogenesis or early embryogenesis
340 and, as a consequence, do not enable to determine the extent to which embryonic
341 H2AK119ub1 is instructed/inherited from sperm and/or oocyte
342 H2AK119ub1[9,11,12,19,20,38]. Here, by editing the sperm epigenome, after completion of
343 spermiogenesis and prior to exposure to egg factors, we circumvent these limitations.
344 We demonstrate that sperm contribute to the embryo a USP21 sensitive cue required
345 for paternal H2AK119ub1 remodelling by egg factor, embryonic gene expression, and
346 developmental competence (**Fig. 3J**). We find that egg factor mediated accumulation
347 on paternal CpG island depends on H2AK119ub1 delivered by the sperm. Strikingly,
348 we observe that erasure of H2AK119ub1 from sperm leads to a global shift in genomic
349 accumulation from CpG islands to repeat elements. Given the propensity of vPRC1 to
350 be targeted and to spread from existing H2AK119ub1 location[41] it raises the possibility
351 that pre-existing sperm derived ''docking sites" could limit chromatin sampling by the
352 polycomb complex and prevent H2AK119ub1 deposition at abnormal locations.
353 Embryos derived from H2AK119ub1 edited sperm exhibit transcriptional
354 misregulation, mostly upregulation, compatible with the repressive function of this
355 histone PTM on transcription. A similar observation is observed in mouse when
356 H2AK119ub1 deubiquitylase is expressed at fertiization[19], or in zebrafish when Rnf2
357 is chemically inhibited[20]. Given the extent of egg factor mediated H2AK119ub1
358 accumulation on gene promoters observed in frog, mouse or zebrafish, it is somewhat
359 surprising that only a limited number of genes are mis-regulated when its distribution
360 is massively altered. We identify an association of misregulated genes with
361 H2AK119ub1 in frog sperm and find that this association is preferentially maintained
362 on these genes upon sperm chromatin replication, further supporting a direct role for
363 this histone post-translational modification in the epigenetic programming of sperm for
364 regulation of embryonic transcription. We cannot rule out that a part of the genes is
365 indirectly affected, for example as downstream target of a "primary" H2AK119ub1
366 target. Alternatively, we provide evidence that gene sensitivity to H2AK119ub1 erasure
367 from sperm occurs when genes are associated with H3K4me3, a histone PTM known
368 to act as a vector of intergenerational information[9–11,42,43]. It is likely that additional
369 epigenetic features besides H3K4me3 contribute to gene sensitivity to H2AK119ub1
370 erasure from sperm. Additionally, we observe that most affected genes are normally
371 expressed late in embryogenesis (tadpole stage) a feature also seen in mouse with a
372 large proportion of gene sensitive to H2AK119ub1 depletion expressed post-
373 implantation[8]. It is therefore possible that the effect of H2AK119ub1 expression is only
374 partially revealed at early embryonic stage since transcription factors targeting these
375 genes might not be present.

376 Our work advance knowledge form previous studies by showing that part of
377 H2AK119ub1 found in embryos is paternally instructed. This necessary function of
378 sperm derived modified histones has important bearing on our understanding of
379 inter/trans generational transmission of epigenetic information. One could indeed
380 envisage that such a naturally occurring mechanism of epigenetic information
381 transmission to the embryos could be hijacked to transmit environmentally induced

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

387

388 **FIGURE LEGENDS**

389 **<u>Figure 1:</u> *H2AK119ub1 associates with gene regulatory regions in sperm chromatin***

390 (**A**) Schematic representation of *Xenopus laevis* nucleosome remodelling during
391 spermiogenesis (**B**) Left, Western blot analysis of H2AK119ub1, H2A and H4 histones.
392 Right quantitation of signal ratio of H2A to H4, H2AK119ub1 to H4, and H2AK119ub1
393 to H2A, in spermatid and sperm. * $p<0.05$, ratio paired t test (n=4). (**C**) Western blot
394 analysis of H2AK119ub1, H3K27me3, H3K9me3, H3K4me3, and H4 in *Xenopus*
395 *laevis* embryo pre- (blastulae) and post- (gastrulae, and neurula) ZGA. **(D)** Heatmap of
396 H2AK119ub1 signal (log (ChIP/Input)) on all *Xenopus Laevis* genes TSS +/-5kB. GO
397 enrichment is shown for the top 10% genes. **(E)** Genome browser view of log ratio of
398 ChIP to input signal of sperm H2AK119ub1, H3K4me3, and H3K27me3 around *wt1.L,*
399 *pax6.L*, snx10.s and *hoxa.S* loci. Boxed area underneath the gene track indicates regions
400 of enrichment (peaks). **(F)** Distribution of sperm H2AK119ub1 peaks on genomic
401 features. **(G)** Fold enrichment of H2AK119Ub1 peaks on different genomic features,
402 *** $p<0.001$, permutation test. **(H)** Pie chart indicating the proportion of H2AK119ub1
403 peaks either alone, or overlapping H3K4me3 and/or H3K27me3 peaks in sperm. **(I)**
404 Genomic distribution of sperm H2AK119ub1 peaks according to overlap with
405 H3K4me3 and/or H3K27me3 peaks.

406 **<u>Figure 2:</u> *Sperm H2AK119ub1 is required for the remodelling of paternal chromatin***
407 ***by egg factors.***

408 **(A)** Experimental design for monitoring modified histones fate following sperm
409 exposure to egg factors. (**B**) Heatmap of H2AK119ub1, H3K4me3, or H3K27me3
410 signal in spermatid, sperm, replicated sperm, and gastrula. All genomic locations
411 enriched for one of this histone modification (peaks) in at least one of these
412 developmental stages are represented. **(C)** Venn diagram indicating the changes to
413 H3K4me3, H2AK119ub1, or overlapping H3K4me3/H2AK119ub1 peaks during
414 sperm chromatin replication. Left, all H2AK119ub1 peaks; Middle all H3K4me3
415 peaks; Right, overlapping H2AK119ub1/H3K4me3 peaks. Percentages of sperm peaks
416 that are maintained after replication are indicated underneath. (**D**) Schematic
417 representation of sperm treatment with oocyte extract containing USP21. (**E**) Left,
418 Western blot of H2AK119ub1, H3K4me3, H3K9me3, H3K27me3 and histone H4 on
419 oocyte extract treated sperm, USP21 treated sperm, or untreated sperm. Right, signal
420 ratio of H2AK119ub1 to H4 in oocyte extract treated sperm, USP21 treated sperm, or
421 untreated sperm. ** $p<0.01$, *** $p<0.005$, ratio paired t test (n=4). (n=6) (**F**) DNA
422 fragments generated by MNase digestion of *Xenopus laevis* sperm, oocyte extract
423 treated sperm, egg extract treated sperm and somatic cell.
424 (**G**) Western blot analysis of H2AK119ub1 and histone H4 on sperm after treatment
425 with control- or USP21- oocyte extract without (left) or with (right) subsequent
426 replication in egg extract. (**H**) Heatmap of H2AK119ub1 and H3K4me3 signal in sperm
427 after treatment with control- or USP21- oocyte extract without or with subsequent

replication in egg extract. All genomic locations enriched for H2AK119ub1 (peaks) in replicated samples are represented. (**I**) Transcription factor motifs enriched in peaks that are lost, maintained, or gained upon replication of control versus USP21 treated sperm. (**J-M**) DNAme and H2AK119ub1signal on *X.Laevis* genome CpG island (J) and repeat element (L). Fold enrichment of H2AK119ub1 peaks on CpG island (K) and repeat element (M).

<u>Figure 3:</u> ***Sperm H2AK119ub1 is required for optimal embryonic development and gene expression.***

(**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 RRID:SCR_003280). (**B**) Developmental progression to blastula, gastrula or swimming tadpole stage of haploid embryos derived from control or a USP21 oocyte extract treated sperm. Chi-squared test, * p-value<0.05. (**C**) MA-Plot showing as y-axis the log foldchange (LogFC) of gene expression between control oocyte extract treated sperm *versus* untreated sperm (left) and between USP21 oocyte extract treated sperm derived embryo and untreated sperm (right). X-axis values are 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. **(E)** Fold enrichment of different H2AK119ub1 peaks categories on promoter, enhancer, and gene body of USP21 sensitive genes. * p-value<0.05; ** p-value < 0.01, permutation test. **(F)** H2AK119ub1 and H3K4me3 signal on TSS +/- 5kB for all genes (black), USP21sensitive genes (red) before and after sperm replication. *** p-value<0.001, Wilcoxon-Mann-Whitney test on TSS+/1kb interval (shaded orange area). **(G)** Calibrated H2AK119ub1 ChIP-seq signal on TSS +/- 5kB for all genes (black), USP21 genes (red) before and after sperm replication. *** p-value<0.001, Wilcoxon-Mann-Whitney test. **(H)** Fold enrichment of lost, maintained, or gained H2AK119ub1 peaks on promoter, enhancer, and gene body of maintained, loss or conserved peaks in replicated USP21 treated sperm *versus* control sperm. * p-value<0.05; ** p-value<0.01, ***p-value<0.001, permutation test. **(I)** Calibrated H2AK119ub1 ChIP-seq signal on TSS +/- 5kB for all genes (black), USP21 genes (red) before and after replication using a WT sperm or USP21 treated sperm. *** p-value<0.001, Wilcoxon-Mann-Whitney test. **(J)** Model of H2AK119ub1 epigenetic programming of sperm for embryonic development in *Xenopus Laevis.*

<u>Supplemental Figure S1:</u> ***ChIP-seq analysis of H2AK119ub1 distribution of H2AK119ub1in sperm chromatin***

(**A**) Genome browser view of the log ratio of H2AK119ub1 ChIP to input signal around the *hoxa*.L cluster on 3 sperm biological replicates. (**B**) Spearman correlation of sperm and replicated sperm H2AK119ub1, H3K27me3 and H3K4me3 ChIP-seq replicates. (**C**) Fingerprint plot of input, H3K4me3, H3K27me3, and H2AK119ub1 ChIP samples. (**D**) Overlap between RECOGNICER and MACS2 peaks for H2AK119ub1, H3K7me3, and H3K4me3 peaks. **(E)** Heatmap displaying H2AK119ub1 log(ChIP/Input) signal +/- 5kb from RECOGNICER peaks center (middle) *versus* an equivalent number of genomic regions randomly selected (left). On the right RECOGNICER peaks are spitted between those identified by RECOGNICER only versus those detected in both RECOGNICER and MACS2. (**F**) Genome track showing logratio of ChIP to input signal for H2AK119ub1 and H3K27me3 as well as the region where peaks are detected

475 by MACS2 and RECOGNICER (boxed area, red for H2AK119ub1, yellow for
476 H3K27me3) at different scale view. **(G)** ChIP-seq fragment size distribution for
477 H2AK119ub1 (red) in sperm and spermatid, and for H3K4me3 (green) and H3K27me3
478 (yellow) in sperm. Matching inputs are shown in grey.

479 **Supplemental Figure S2:** ***Sperm H2AK119ub1 fate following exposure to egg factors***

480 (**A**) Heatmap of H2AK119ub1 (red), H3K4me3 (green), and H3K27me3 (yellow)
481 log(ChIP/Input) signal ratio on genes TSS from spermatid, sperm prior and after egg
482 extract treatment, and gastrula stage 12 embryo. (**B**) Scaled Venn diagram of
483 H2AK119ub1, H3K4me3 or overlapping H2AK119ub1/H3K4me3 peaks in spermatid,
484 sperm, and sperm exposed to egg factors. Percentage of peaks that are maintained at
485 each developmental transition are indicated underneath. (**C**) Top: boxplot showing
486 calibrated H2AK119ub1 on H2AK119ub1/H3K4me3 peaks lost (375), maintained
487 (5222) or gained (10442) following exposure to egg factors (top). Bottom: peak size
488 (blue dots) as well as peak H2A enrichment particle (red dots) in each peak category.
489 **(D)** Genomic distribution of sperm peaks before or after replication, as well as the
490 subset of peaks lost, maintained, or gained during replication. **(E)** Quantitation of
491 H3K4me3, H3K9me3 and H3K27me3 to H4 WB signal ratio in untreated sperm,
492 oocyte extract treated sperm or USP21 treated sperm. **(F)** Capillary electrophoresis
493 profiles of DNA size obtained following MNAse digestion of sperm (blue), oocyte
494 extract treated sperm (cyan), egg extract treated sperm (green) or somatic cell (purple).
495 **(G)** Experimental design for monitoring modified histones fate from control or USP21
496 treated sperm. **(H)** Heatmap of H2AK119ub1 (red), H3K4me3 (green) and H3K27me3
497 (green) log(ChIP/Input) signal ratio in spermatid, control- and USP21- treated sperm
498 prior and after egg extract treatment, and gastrula stage 12 embryo. Heatmap includes
499 the same genomic location as in figure 2B. **(I)** Heatmap of H2AK119ub1 and H3K4me3
500 signal in sperm after treatment with control- or USP21- oocyte extract without or with
501 subsequent replication in egg extract. All genomic locations enriched for H3K4me3
502 (peaks) in replicated samples are shown.

503 **Supplemental Figure S3:** ***Gene expression in gastrulae derived from control or***
504 ***USP21 treated sperm.***

505 **(A)** Principal Component analysis of replicate RNA-seq samples used in this study
506 (USP21, USP21 extract treated sperm derived gastrulae; Control, control extract treated
507 sperm derived gastrulae; Untreated, Untreated sperm derived gastrulae. **(B)**
508 Hierarchical clustering of replicate samples used in this analysis. **(C)** Heatmap
509 depicting expression of USP21 sensitive genes in replicates from USP21 extract treated
510 sperm derived gastrulae, control extract treated sperm derived gastrulae, as well as
511 gastrulae derived from untreated sperm. **(D)** Heat map of gene expression in single
512 haploid gastrula derived from control- or USP21- treated sperm. Expression is shown
513 for the USP21 sensitive gene set identified on embryos pool (fig 3D). **(E)** USP21
514 sensitive gene set expression during early Xenopus embryo development (data from).
515 Expression is shown for oocyte, egg, blastulae (stage 8-9), gastrulae (stage 10-12),
516 neurula (stage 15-20), as well as early to late tailbud (stage 25-40) are indicated. The
517 right-hand side panel shows the expression of USP21 sensitive gene in gastrulae
518 derived from control- and USP21-treated sperm. **(F)** Violin plot of gene expression in
519 egg (log of normalised gene count) for all genes (black), USP21 sensitive genes (red),
520 maternal genes (cyan), maternal and zygotic genes (green) and zygotic only gene
521 (orange). USP21 sensitive genes are enriched for genes with maternal and zygotic

522 expression in *Xenopus Laevis*. The table below indicates p-Value resulting from an
523 hypergeometric test to evaluate enrichment of USP21 sensitive gene for either
524 maternal-zygotic (MZ) or zygotic only (Z) genes. MZ and Z dataset taken from[34,35].
525 **(G)** Fold enrichment of H2AK119ub1 peaks on promoter, enhancer, and gene body of
526 USP21 sensitive genes. *p-value<0.05; permutation test. **(H)** Fold enrichment of
527 different H2AK119ub1 peaks categories on promoter, enhancer, and gene body of
528 USP21 sensitive genes. *p-value<0.05; **p-value < 0.01, ***p-value<0.001,
529 permutation test. **(I)** Calibrated H2AK119ub1 ChIP signal on TSS +/- 5kb for all genes
530 (black), USP21 sensitive genes (red), maternal genes (cyan), maternal zygotic genes
531 (green) and zygotic genes (orange) before and after sperm replication. *** p-
532 value<0.001, Wilcoxon-Mann-Whitney test.

533

534 **Table S1***: Germ cell histone PTMs ChIP-seq data*

535 **Table S2**: ***Control- and epigenetically edited- sperm derived embryos RNA-seq data***

536

537 **MATERIALS AND METHODS**

538 **Contact for Reagent and Resource Sharing.** Further information and requests for
539 resources and reagents should be directed to and will be fulfilled by the Lead Contact,
540 Dr J. Jullien (jerome.jullien@inserm.fr).

541 **Experimental Model and Subject Details.** Mature *Xenopus laevis* females were
542 obtained from Tefor Paris-saclays (151 Rte de la Rotonde, 91400 Saclay, France,
543 https://tefor.net/tefor-paris-saclay-nos-equipes/tefor-paris-saclay-aqua/). Our work
544 with Xenopus complied with all relevant ethical regulations for animal testing and
545 research. Work with frog was covered under the licence #202212131604754/APAFIS
546 39794V2 and frog husbandry and all experiments were performed according to the
547 relevant regulatory standards. Animals were kept in a recirculating system (ALS,
548 https://www.aquariumlocationservices.fr/professionnels-elevages-aquatique-
549 aquariophilie/) at a density of one adult/3l, with ~10% water change per day. Water
550 was sequentially filtered with mechanical pad sump filter, nitrifying bacteria filter, and
551 UV sterilized. Water quality parameters were as follow: temperature 17-22°C; PH 6-8.
552 Photoperiod was set to 12h ON/12h OFF. Frogs are fed twice per week with NEO
553 MARIN coul 4 - X4 - S20 (Le Gouessant, https://www.legouessant.com/).
554 Unconsumed food was removed 30 min after the start of feeding.

555 **Oocyte extract.** Fresh ovaries were manually dissociated in 15-25 oocytes clumps. 3-
556 5 mL of dissociated ovary fragments were transferred into 12.5mL 1X Modified Barth's
557 Saline (MBS, 88mM NaCl, 1 mM KCl, 10 mM Hepes, 2.4 mM NaHCO3, 1.7 mM
558 MgSO4, 1.3 mM Ca(NO3)2, 0.8mM CaCl2) supplemented with 250uL of 5mg/mL
559 Liberase (Roche, 05 401 127 001) and incubated for 2h30-3h at room temperature (RT)
560 with gentle rocking. Defolliculated oocytes were washed three times with 1X MBS and
561 healthy stage 5-6 oocytes manually sorted. Oocytes were then injected in the vegetal
562 pole with 25ng of USP21-NLS-HA mRNA, or control GFP mRNA and incubated for
563 3 days at 18°C in 1X MBS, renewing medium and discarding dead cells every day.
564 ~500-1000 oocytes were collected in an Eppendorf tube on ice, the excess 1X MBS
565 discarded and replaced with 1X MBS supplemented with Protease Inhibitor Cocktail

566 (Sigma, P2714-1BTL). The oocytes were then stacked up by centrifugation at 100g for
567 1min at RT and excess medium carefully discarded. Oocytes were then thoroughly
568 vortexed and centrifuged at 21 000g for 10min at 4°C. The clear intermediate phase
569 between yolk (top) and pigment (bottom) was retrieved. The collected fraction was
570 further cleared by an additional centrifugation at 21 000g for 10min at 4°C. The oocyte
571 extract was then aliquoted and stored at -80°C.

572 **Oocyte extract treatment.** Permeabilized sperm stored in SSB at -80°C were thawed,
573 washed with two volumes of ice cold 1XPBS and centrifuged at 800g for 8min at 4°C.
574 2M sperm in 1X PBS were centrifuged at 21 000g for 1min at 4°C. The supernatant
575 was discarded, and the pelleted sperm resuspended in 10uL of oocyte extract on ice.
576 The sperm incubated in the oocyte extract for 20-40min at 23°C on a heating block.
577 Treatment was stopped by addition of 90uL of ice cold 1X PBS and centrifugation at
578 21 000g for 1min at 4°C. The supernatant discarded and the remaining pellets
579 resuspended in SSB for -80°C long term storage.

580 **Protein extraction and Western Blot analysis**. Sperm samples stored in SSB were
581 washed in 1X PBS and 1M sperm pellets were resuspended in 30uL 2X Blue Eye
582 (600mM NaCl, 1% Triton X100, 0.5% Na-Deoxycholate, 4% Sodium Dodecyl Sulfate,
583 20% Glycerol, 0.0004% Bromophenol blue, 175mM Tris-HCl, 10% Betamercapto
584 ethanol), boiled at 96°C for 5min and centrifuged at 21 000g for 5min at 4°C. The
585 supernatant was transferred to a new tube and directly use for analysis or stored at -
586 20°C. 4-20% mini-PROTEAN® TGX (Biorad, #4561094) were loaded with 10uL of
587 sample and western blot analysis were performed according to standard protocols. For
588 blotting nitrocellulose membranes, a semi-dry transfer was used (Biorad, Trans-Blot®
589 SD Semi-Dry Transfer Cell) 30-40min at 25V. The following primary antibodies were
590 used: anti-H2A (1:2500, Cell Signalling, 3636), anti-H2AK119ub1 (1:2500, Cell
591 Signalling, 8240), anti-H3K4me3 (1:2500, Abcam, ab8580), anti H3K9me3 (1:2500,
592 Abcam, ab8898), anti-H3K27me3(1:2500, Cell Signalling, 9735S), anti-H4 (1:2500,
593 Abcam, ab31830) and anti-HA (1:2500, Sigma, H9658). Membranes were incubated
594 with primary antibodies 2h at RT or overnight (ON) at 4 °C. Secondary antibody are
595 anti-rabbit (1:25000, ThermoFisher, A21076) , anti-mouse IgG Alexa Fluor 680
596 (1:25000, ThermoFisher ,A21058), Goat anti-rabbit Alexa 800 (1:25,000,
597 ThermoFisher, A32735), or IRDye 800CW Goat Anti-Mouse IgG (H+L) (LI-COR,
598 926-32210) were used. The membranes were imaged either on a ChemiDoc MP
599 Imaging System (biorad) or an Odyssey scanner (LI-COR).

600 **Sperm Chromatin Immunoprecipitation (ChIP).** Chromatin fractionation and
601 chromatin immunoprecipitation (ChIP) were performed as described before[10,44] with
602 slight modifications. 80 µl of Magnetic beads (M280 Sheep anti-rabbit IgG, Invitrogen,
603 #11204D) were used per reaction and all wash steps were carried out with a magnet.
604 Beads were washed with 1 ml of TE pH 8.0, lysis buffer (LB)(1:1 mixture of Buffer 1
605 and MNase buffer, see below), and incubated in 1 ml of LB with 100 µl of 10 mg/ml
606 BSA at 4C for 30 min ("pre-blocking"). Pre-blocked beads were washed with 1 ml of
607 LB twice and resuspend in 80 µl of LB. Half of beads were incubated with antibody
608 overnight and leftovers were stored in 4C for "pre-clear chromatin" step. For ChIP, the
609 following antibodies were used: anti-H3K4me3 (Abcam, ab8580), anti-H2AK119ub1
610 (Cell Signaling, 8240). One million digitonin permeabilised sperm nuclei or egg extract
611 treated sperm nuclei were resuspended in 50 µl of Buffer 1 (0.3 M Sucrose, 15 mM Tris
612 pH 7.5, 60 mM KCl, 15 mM NaCl, 5 mM MgCl2, 0.1 mM EGTA, 0.5 mM DTT)

613 containing 0.125ng H2AK119ub biotinyltated nucleosome (Epicypher #16-0020),
614 followed by addition of 50 µl of Buffer 1 with detergent (Buffer 1 including 0.5% NP40
615 and 1% NaDOC). Samples were incubated for 10 min on ice. 100 µl of MNase buffer
616 (0.3 M Sucrose, 85 mM Tris, 3 mM $MgCl_2$, 2 mM $CaCl_2$, 2.5U of micrococcal
617 nuclease: Roche 10107921001) was added in to each tube (0.5 millions of cells per
618 tube). Tubes were incubated at 37C for 30 min in pre-warmed water bath. Reaction was
619 stopped by adding 2 µl of 0.5 M EDTA pH8.0 in the same order as started. Tubes were
620 vortexed and placed on ice for at least 5 min. Supernatant and pellet were separated by
621 centrifugation at 13,000 rpm for 10 min at room temperature. Beads (stored at 4C for
622 pre-clear chromatin) were washed with 1 ml of LB twice. Supernatant and EDTA-free
623 Protease Inhibitor Cocktail were added into the beads and rotating on the wheel at 4C
624 for 60 min. 10% of volume of pre-cleared chromatin was taken as "input" and stored at
625 4C. Antibody conjugated beads were washed with 1 ml of LB twice. Pre-cleared
626 chromatin was added to washed beads and incubated at 4C for 6 h. After incubation,
627 beads were washed with 1 ml of Washing buffer A (50 mM Tris-HCl pH 7.5, 10 mM
628 EDTA and 75 mM NaCl) on the wheel at 4C for 5 min and buffer was discarded
629 carefully. Subsequently, beads were washed wish 1 ml of Washing buffer B (50 mM
630 Tris-HCl pH 7.5, 10 mM EDTA and 125 mM NaCl) and buffer was discarded carefully.
631 Another 1 ml of Washing buffer B was added and transferred to 1.5 ml Protein lo-bind
632 tubes (Eppendorf, 0030108116) with beads. Buffer was discarded carefully. 150 µl of
633 elution buffer (1:9 mixture of 10% SDS and TE) was added. Tubes were placed at 25C
634 on the Eppendorf thermomixer and shaken at 800 rpm for 15 min. Supernatant were
635 taken and transferred to the new tube using magnetic stand. Another 150 µl of elution
636 buffer was added and repeated this step (in total 300 □l). Input samples were diluted to
637 300 µl total by adding TE. 2 µl of 2 mg/ml RNase A (heat-inactivated or DNase-free)
638 was added and incubated at 37C for 30 min. 4 µl of 10 mg/ml Proteinase K was added
639 and incubated at 55C overnight. After incubation, tubes were shortly spun. 300 µl of
640 Phenol/chloroform were added and vortexed for 30 sec. Samples were centrifuged at
641 13,000 rpm for 10 min at room temperature. 1 ul of Glycogen, 1/10 volume of
642 NaAcetate and 2 volumes of 100% EtOH were added and vortexed. Tubes were stored
643 on dry ice for 20-30 min. Once frozen, tubes were centrifuged at 13,000 rpm for 30 min
644 at 4C. Supernatant was discarded carefully and 500 µl of 70% EtOH was added. Tubes
645 were centrifuged at 13,000 rpm for 10 min at 4C. Pellets were dried and suspended in
646 30 μl of ddH_2O.

647 **Egg extract treatment.** Interphasic extracts were generated as previously
648 described[45] .For the quantification of DNA replication in individual nuclei in egg
649 extract, 33 µM rhodamine dUTP (Roche, 11534378910) was added at the indicated
650 times (20, 40, 60, 80 mins after incubation) for 5 min as indicated, and the reaction was
651 stopped with 1xPBS and fixed in 4% paraformaldehyde. The reaction was spun through
652 a 30% sucrose cushion in PBS. The sperm cells were rinsed twice in 1x PBS for 10 min
653 and mounted in Vectashield mounting medium with DAPI (Vector laboratories, H-
654 1200). Counting of rhodamine positive cells were performed on an epifluorescence
655 microscope.

656 ***Xenopus* sperm collection *Xenopus* sperm collection.** *Xenopus* sperm collection was
657 performed as described before[10,44]. For each round of sperm purification, testes from
658 six adult *Xenopus laevis* males were isolated and manually cleaned from blood vessels
659 and fat bodies in 1 × Marc's Modified Ringers (MMR, 100 mM NaCl, 2 mM KCl, 1
660 mM $MgSO_4$, 2 mM $CaCl_2$, 3 mM HEPES, pH 7.4) using forceps and paper tissues.
661 Subsequently, testes were torn into small pieces with forceps and homogenized with 2–

662 3 strokes of a Dounce homogenizer (one testis at a time). The cell suspension was then
663 filtered to remove tissue debris and cell clumps using a 60µm Mesh (CellTrics, cat. 04-
664 0042-2317) and spun down at 800 rcf, 4°C, for 20 min. Supernatant was discarded and
665 the cell pellet was resuspended in 12 mL of 1 × MMR. If any red blood cells were
666 visible at the bottom of the pellet (as a result of incomplete removal of blood vessels),
667 only the uncontaminated part of the pellet was recovered, taking extreme care not to
668 disturb the red blood cells. Subsequently, step gradients of iodixanol (Optiprep; Sigma,
669 D1556; 60% iodixanol in water) in 1×MMR final were manually prepared in pre-chilled
670 14 mL ultra- clear centrifuge tubes (Beckman Coulter, #344060) in the following order
671 from the bottom to the top of the tube: 4 mL of 30% iodixanol, 1 mL of 20% iodixanol,
672 5 mL of 12% iodixanol (all in 1 × MMR), and 2 mL of cell suspension in 1 × MMR on
673 top. Gradients were spun down in a pre-chilled SW40Ti rotor at 7500 rpm (10,000g),
674 4°C, for 15 min, deceleration without brake (Beckman Coulter Ultracentrifuge, Optima
675 L-100XP). The pelleted fraction, containing mature sperm, was collected. Collected
676 fractions were diluted six times with 1× MMR and collected by spinning first at 805
677 rcf, 4°C, for 20 min and then at 3220 rcf, 4°C, for 20 min to pellet remaining cells.
678 Pelleted cells were subsequently permeabilized with Digitonin (10 mg/mL as a final
679 concentration, Sigma, D141) for 5 min at room temperature as detailed before (Smith
680 et al., 2006).

681 **Library preparation and sequencing.** ChIP samples were subjected to ChIP-seq
682 library preparation using QIAseq Ultralow Input Lib UDI-A Kit (Qiagen, 180497)
683 and Agencourt AMPure XP beads (Beckman coulter, A63881). Samples pre- and
684 post-library protocol were analyzed on a LabChip GX (PerkinElmer) using DNA
685 1KHS and 1K Assay LabChIP, respectively. Libraries were quantitated using
686 NEBNext Library Quant Kit (NEB, E7630,). For RNA sequencing, RNAs were
687 purified using Qiagen RNeasy Mini Kit (#74004) and used to generate cDNA
688 sequencing libraries using an Illumina TruSeq Kit (#RS-122-2001), according to the
689 manufacturer's specifications. Paired-end sequencing was carried out either on a
690 NOVA-seq 6000 or HiSeq 1500 sequencer (Illumina).
691

692 **RNA-Seq analysis.** Adapters were trimmed and low-quality reads filtered using
693 cutadapt (v1.18)[46]. Sample quality was confirmed with fastqc (v0.11.8)[47]. fastq were
694 aligned to *Xenopus Laevis* genome (v9.2) using STAR aligner (v2.7.3.a)[48]. Count table
695 was generated using HTSeq-count (v0.12.3)[49] and exon as genes reference. MultiQC
696 (v1.9)[50] summarized data pre-processing. Differential gene expression analysis of the
697 count table was performed on R (v3.6)[51] using DESeq2tools[52]. Details of the pipeline
698 used can be found at https://gitlab.univ-nantes.fr/E114424Z/BulkRNAseq. Gene
699 ontology analysis were carried out with gprofiler[53] using human orthologs of *Xenopus*
700 *laevis* genes recovered from Xenbase (https://www.xenbase.org/xenbase/static-
701 xenbase/ftpDatafiles.jsp)[54]. *Xenopus laevis* development RNA-seq data were from
702 published dataset GSE73430.

703

704 **ChIP-Seq analysis.** Adapters were trimmed and low-quality reads filtered using
705 cutadapt (v1.18)[46]. Sample quality was confirmed with fastqc (v0.11.9)[47]. Alignment
706 was performed with bwa mem[55] against *Xenopus laevis* (v9.2) or against a *Xenopus*
707 *laevis* (v9.2) - 601DNA hybrid genome generated with bwa index (v0.7.17-r1188).
708 Reads extraction and analysis was performed with samtools view and samtools flagstat

709 (v1.12)[56]. ChIP-Seq sample quality was assessed by Deeptools (v3.5.1)[57] using
710 plotFingerprint function and bamPEFragmentSize. Replicates quality was visualised
711 with plotCorrelation using spearman correlation after genome fragmentation into bins
712 of 10Kb. PCR duplicates were marked by Picard tools using MarkDuplicates command
713 (https://broadinstitute.github.io/picard) and then removed with samtools view -F 0x400.
714 Replicates were pooled by subsampling the same amount of reads per replicates using
715 sambamba (v0.8.1)[58] and samtools view (-s option). Peak calling was processed with
716 MACS2 (https://github.com/macs3-project/MACS)[59] (-q 0.01 --nomodel --extsize 147
717 --gsize=2.6e9) for H3K4me3, with MACS2 broad option for H3K27me3 (--broad --
718 broad-cutoff 0.1 -q 0.01 --nomodel --extsize 147 --gsize=2.6e9) and RECOGNICER[24]
719 (qvalue 0.001) /MACS2 (--broad --broad-cutoff 0.1 -q 0.01 --nomodel --extsize 147)
720 for H2AK119ub1. Peaks overlap was defined using bedtools intersect having at least 1
721 bp of overlap between peaks. H2AK119ub1 signal was calibrated using reads mapping
722 to the 601DNA and corresponding to the fully ubiquitylated H2AK119 nucleosome
723 spiked-in in samples prior to ChIP (see ChIP M&M section). To evaluate difference in
724 modifed histone ChIP-seq signal around the TSS of gene sets, Wilcoxon Mann-
725 Whitney test was performed on the +/-1 kb window around the TSSs. ChIPSeeker[60] and
726 ChIPPeakAnno[61,62] were used to annotate peaks on R[51]. RegioneR[63] was used to
727 perform peaks permutation to assess enrichment on USP21 sensitive gene features (*i.e.*
728 TSS, enhancer, or gene body) compared to random set of genes (resampleRegions,
729 10000 iterations) and to assess peak enrichment on genomic features
730 (randomizeRegions, 1000 iterations) Log10 chip over input was computed using
731 Deeptools with BamCompare using 50bp binsize. Heatmap and TSS plot were
732 produced with computeMatrix and plotHeatmap or plotProfile. IGV was used for peaks
733 and log10(chip/input) signal visualisation. Heatmap were produced with
734 computeMatrix using 50bp binsize +/-5kb of TSS or peak center. Heatmaps were
735 generated with plotHeatmap function using unsupervised clustering (--hclust option),
736 and metaplot using plotProfile function. IGV[64] was used to visualize output signal
737 tracks. Bedtools getfasta[65] was used to extract FASTA sequences corresponding to
738 peaks from a selected heatmap cluster. To identify transcription factors motif in the
739 selected peak set, Homer[66] was then used with the findMotifs.pl function and all other
740 peaks of the heatmap as a background set. CpG islands and repeats elements were
741 recovered from UCSC genome annotation database
742 (https://hgdownload.soe.ucsc.edu/goldenPath/xenLae2/database/). All scripts used for
743 these analysis are available on Github (https://github.com/FCValentin/H2AubPaper).

744

745 **Data Availability.** The *Xenopus laevis* ChIP-seq and RNA-seq data generated for this
746 study will be referenced as PRJEB56442 at the European Genome-phenome Archive
747 (EGA). Additional dataset used are from GSE73430 (Xenopus embryo RNA-seq),
748 GSE125982 (sperm and replicated sperm H3K27me3 ChIP-seq), GSE126599
749 (Gastrulae H3K27me3, H2AK119ub1, and H3K4me3 ChIP-seq), and GSE75164
750 (sperm MBD-seq).

751

752 **AUTHORS CONTRIBUTION**

753 Conceptualization: JJ; Methodology: JJ, VF-C, FB; Validation: JJ, VF-C, FB; Formal
754 analysis: VF-C, FB, MO, MG; Investigation: JJ, VF-C, FB, RG, MO, BH; Writing

original draft: JJ; Writing review and editing: JJ, FB, VFC. Visualisation: FB, VF-C; Supervision: JJ; Project administration: JJ; Funding acquisition: JJ.

COMPETING INTERESTS

The authors declare no competing interests

ACKOWLEDGEMENTS

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) form the Human Frontier Science Program. We would like to acknowledge the support of Xenbase.

REFERENCE

1. Öst, A. *et al.* Paternal diet defines offspring chromatin state and intergenerational obesity. *Cell* (2014) doi:10.1016/j.cell.2014.11.005.

2. Dias, B. G. & Ressler, K. J. Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat. Neurosci.* (2014) doi:10.1038/nn.3594.

3. Radford, E. J. *et al.* In utero undernourishment perturbs the adult sperm methylome and is linked to metabolic disease transmission. *Science (80-.).* **345**, 1–26 (2015).

4. Ben Maamar, M., Beck, D., Nilsson, E. E., Kubsad, D. & Skinner, M. K. Epigenome-wide association study for glyphosate induced transgenerational sperm DNA methylation and histone retention epigenetic biomarkers for disease. *Epigenetics* **16**, (2021).

5. 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**, (2020).

6. U., B. *et al.* Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nature Structural and Molecular Biology* (2010).

7. Hammoud, S. S. *et al.* Distinctive chromatin in human sperm packages genes for embryo development. *Nature* (2009) doi:10.1038/nature08162.

8. Wu, S. F., Zhang, H. & Cairns, B. R. Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. *Genome Res.* (2011) doi:10.1101/gr.113167.110.

794 9. Teperek, M. *et al.* Sperm is epigenetically programmed to regulate gene
795 transcription in embryos. *Genome Res.* **26**, (2016).

796 10. Oikawa, M. *et al.* Epigenetic homogeneity in histone methylation underlies
797 sperm programming for embryonic transcription. *Nat. Commun.* (2020)
798 doi:10.1038/s41467-020-17238-w.

799 11. Siklenka, K. *et al.* Disruption of histone methylation in developing sperm
800 impairs offspring health transgenerationally. *Science (80-.).* (2015)
801 doi:10.1126/science.aab2006.

802 12. Murphy, P. J., Wu, S. F., James, C. R., Wike, C. L. & Cairns, B. R. Placeholder
803 Nucleosomes Underlie Germline-to-Embryo DNA Methylation
804 Reprogramming. *Cell* (2018) doi:10.1016/j.cell.2018.01.022.

805 13. Blackledge, N. P. *et al.* Variant PRC1 complex-dependent H2A ubiquitylation
806 drives PRC2 recruitment and polycomb domain formation. *Cell* **157**, (2014).

807 14. Tamburri, S. *et al.* Histone H2AK119 Mono-Ubiquitination Is Essential for
808 Polycomb-Mediated Transcriptional Repression. *Mol. Cell* **77**, (2020).

809 15. Nakagawa, T. *et al.* Deubiquitylation of histone H2A activates transcriptional
810 initiation via trans-histone cross-talk with H3K4 di- and trimethylation. *Genes*
811 *Dev.* (2008) doi:10.1101/gad.1609708.

812 16. Joo, H. Y. *et al.* Regulation of cell cycle progression and gene expression by
813 H2A deubiquitination. *Nature* (2007) doi:10.1038/nature06256.

814 17. Cao, J. & Yan, Q. Histone ubiquitination and deubiquitination in transcription,
815 DNA damage response, and cancer. *Frontiers in Oncology* (2012)
816 doi:10.3389/fonc.2012.00026.

817 18. Chen, D., Dai, C. & Jiang, Y. Histone H2A and H2B Deubiquitinase in
818 Developmental Disease and Cancer. *Cancer Transl. Med.* (2015)
819 doi:10.4103/2395-3977.168578.

820 19. Chen, Z., Djekidel, M. N. & Zhang, Y. Distinct dynamics and functions of
821 H2AK119ub1 and H3K27me3 in mouse preimplantation embryos. *Nat. Genet.*
822 **53**, (2021).

823 20. Hickey, G. J. M. *et al.* Establishment of developmental gene silencing by
824 ordered polycomb complex recruitment in early zebrafish embryos. *Elife* **11**,
825 (2022).

826 21. Conway, E. *et al.* BAP1 enhances Polycomb repression by counteracting
827 widespread H2AK119ub1 deposition and chromatin condensation. *Mol. Cell*
828 **81**, (2021).

829 22. Albright, S. C., Nelson, P. P. & Garrard, W. T. Histone molar ratios among
830 different electrophoretic forms of mono- and dinucleosomes. *J. Biol. Chem.*
831 **254**, (1979).

832 23. Matsui, S., Seon, B. K. & Sandberg, A. A. Disappearance of a structural
833 chromatin protein A24 in mitosis: Implications for molecular basis of

834 chromatin condensation. *Proc. Natl. Acad. Sci. U. S. A.* **76**, (1979).

835 24. Zang, C., Wang, Y. & Peng, W. RECOGNICER: A coarse-graining approach
836 for identifying broad domains from ChIP-seq data. *Quant. Biol.* **8**, (2020).

837 25. Heard, E. & Disteche, C. M. Dosage compensation in mammals: Fine-tuning
838 the expression of the X chromosome. *Genes and Development* vol. 20 (2006).

839 26. Jullien, J. *et al.* Gene Resistance to Transcriptional Reprogramming following
840 Nuclear Transfer Is Directly Mediated by Multiple Chromatin-Repressive
841 Pathways. *Mol. Cell* **65**, (2017).

842 27. Kuznetsov, J. N. *et al.* BAP1 regulates epigenetic switch from pluripotency to
843 differentiation in developmental lineages giving rise to BAP1-mutant cancers.
844 *Sci. Adv.* **5**, (2019).

845 28. Jin, J. *et al.* The deubiquitinase USP21 maintains the stemness of mouse
846 embryonic stem cells via stabilization of Nanog. *Nat. Commun.* (2016)
847 doi:10.1038/ncomms13594.

848 29. Gaggioli, V., Le Viet, B., Germe, T. & Hyrien, O. DNA topoisomerase IIα
849 controls replication origin cluster licensing and firing time in Xenopus egg
850 extracts. *Nucleic Acids Res.* **41**, (2013).

851 30. Blackledge, N. P. & Klose, R. J. The molecular principles of gene regulation
852 by Polycomb repressive complexes. *Nature Reviews Molecular Cell Biology*
853 vol. 22 (2021).

854 31. Zahn, N., Levin, M. & Adams, D. S. The zahn drawings: New illustrations of
855 xenopus embryo and tadpole stages for studies of craniofacial development.
856 *Dev.* **144**, (2017).

857 32. Hörmanseder, E. *et al.* H3K4 Methylation-Dependent Memory of Somatic Cell
858 Identity Inhibits Reprogramming and Development of Nuclear Transfer
859 Embryos. *Cell Stem Cell* **21**, (2017).

860 33. Session, A. M. *et al.* Genome evolution in the allotetraploid frog Xenopus
861 laevis. *Nature* **538**, (2016).

862 34. Chen, H. & Good, M. C. Imaging nascent transcription in wholemount
863 vertebrate embryos to characterize zygotic genome activation. in *Methods in*
864 *Enzymology* vol. 638 (2020).

865 35. Chen, H., Einstein, L. C., Little, S. C. & Good, M. C. Spatiotemporal
866 Patterning of Zygotic Genome Activation in a Model Vertebrate Embryo. *Dev.*
867 *Cell* **49**, (2019).

868 36. Tsuboi, M. *et al.* Ubiquitination-Independent Repression of PRC1 Targets
869 during Neuronal Fate Restriction in the Developing Mouse Neocortex. *Dev.*
870 *Cell* **47**, (2018).

871 37. Żylicz, J. J. *et al.* The Implication of Early Chromatin Changes in X
872 Chromosome Inactivation. *Cell* **176**, (2019).

873 38. Mei, H. *et al.* H2AK119ub1 guides maternal inheritance and zygotic deposition
874 of H3K27me3 in mouse embryos. *Nat. Genet.* **53**, (2021).

875 39. Van Heeringen, S. J. *et al.* Principles of nucleation of h3k27 methylation
876 during embryonic development. *Genome Res.* (2014)
877 doi:10.1101/gr.159608.113.

878 40. Akkers, R. C. *et al.* A Hierarchy of H3K4me3 and H3K27me3 Acquisition in
879 Spatial Gene Regulation in Xenopus Embryos. *Dev. Cell* (2009)
880 doi:10.1016/j.devcel.2009.08.005.

881 41. Zhao, J. *et al.* RYBP/YAF2-PRC1 complexes and histone H1-dependent
882 chromatin compaction mediate propagation of H2AK119ub1 during cell
883 division. *Nat. Cell Biol.* **22**, (2020).

884 42. Lismer, A. *et al.* Histone H3 lysine 4 trimethylation in sperm is transmitted to
885 the embryo and associated with diet-induced phenotypes in the offspring. *Dev.*
886 *Cell* **56**, (2021).

887 43. Pepin, A. S., Lafleur, C., Lambrot, R., Dumeaux, V. & Kimmins, S. Sperm
888 histone H3 lysine 4 tri-methylation serves as a metabolic sensor of paternal
889 obesity and is associated with the inheritance of metabolic dysfunction. *Mol.*
890 *Metab.* **59**, (2022).

891 44. Teperek, M. *et al.* Sperm and spermatids contain different proteins and bind
892 distinct egg factors. *Int. J. Mol. Sci.* **15**, (2014).

893 45. Hannak, E. & Heald, R. Investigating mitotic spindle assembly and function in
894 vitro using Xenopus laevis egg extracts. *Nat. Protoc.* **1**, (2006).

895 46. Martin, M. Cutadapt removes adapter sequences from high-throughput
896 sequencing reads. *EMBnet.journal* **17**, (2011).

897 47. Andrews, S. & others. FastQC: a quality control tool for high throughput
898 sequence data. 2010.
899 *Https://Www.Bioinformatics.Babraham.Ac.Uk/Projects/Fastqc/* (2019).

900 48. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*
901 **29**, (2013).

902 49. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with
903 high-throughput sequencing data. *Bioinformatics* **31**, (2015).

904 50. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: Summarize
905 analysis results for multiple tools and samples in a single report. *Bioinformatics*
906 **32**, (2016).

907 51. Dessau, R. B. & Pipper, C. B. [''R"--project for statistical computing]. *Ugeskr.*
908 *Laeger* **170**, (2008).

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

911 53. Raudvere, U. *et al.* G:Profiler: A web server for functional enrichment analysis

912 and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, (2019).

913 54. Fisher, M. *et al.* Xenbase: Key features and resources of the Xenopus model
914 organism knowledgebase. *Genetics* **224**, (2023).

915 55. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-
916 Wheeler transform. *Bioinformatics* **25**, (2009).

917 56. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**,
918 (2021).

919 57. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-
920 sequencing data analysis. *Nucleic Acids Res.* **44**, (2016).

921 58. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba:
922 Fast processing of NGS alignment formats. *Bioinformatics* **31**, (2015).

923 59. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**,
924 (2008).

925 60. Wang, Q. *et al.* Exploring Epigenomic Datasets by ChIPseeker. *Curr. Protoc.*
926 **2**, (2022).

927 61. Zhu, L. J. *et al.* ChIPpeakAnno: A Bioconductor package to annotate ChIP-seq
928 and ChIP-chip data. *BMC Bioinformatics* **11**, (2010).

929 62. Zhu, L. J. Integrative analysis of ChIP-chip and ChIP-seq dataset. *Methods*
930 *Mol. Biol.* **1067**, (2013).

931 63. Gel, B. *et al.* RegioneR: An R/Bioconductor package for the association
932 analysis of genomic regions based on permutation tests. *Bioinformatics* **32**,
933 (2016).

934 64. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics
935 Viewer (IGV): High-performance genomics data visualization and exploration.
936 *Brief. Bioinform.* **14**, (2013).

937 65. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for
938 comparing genomic features. *Bioinformatics* **26**, (2010).

939 66. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription
940 Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell
941 Identities. *Mol. Cell* **38**, (2010).

942

943

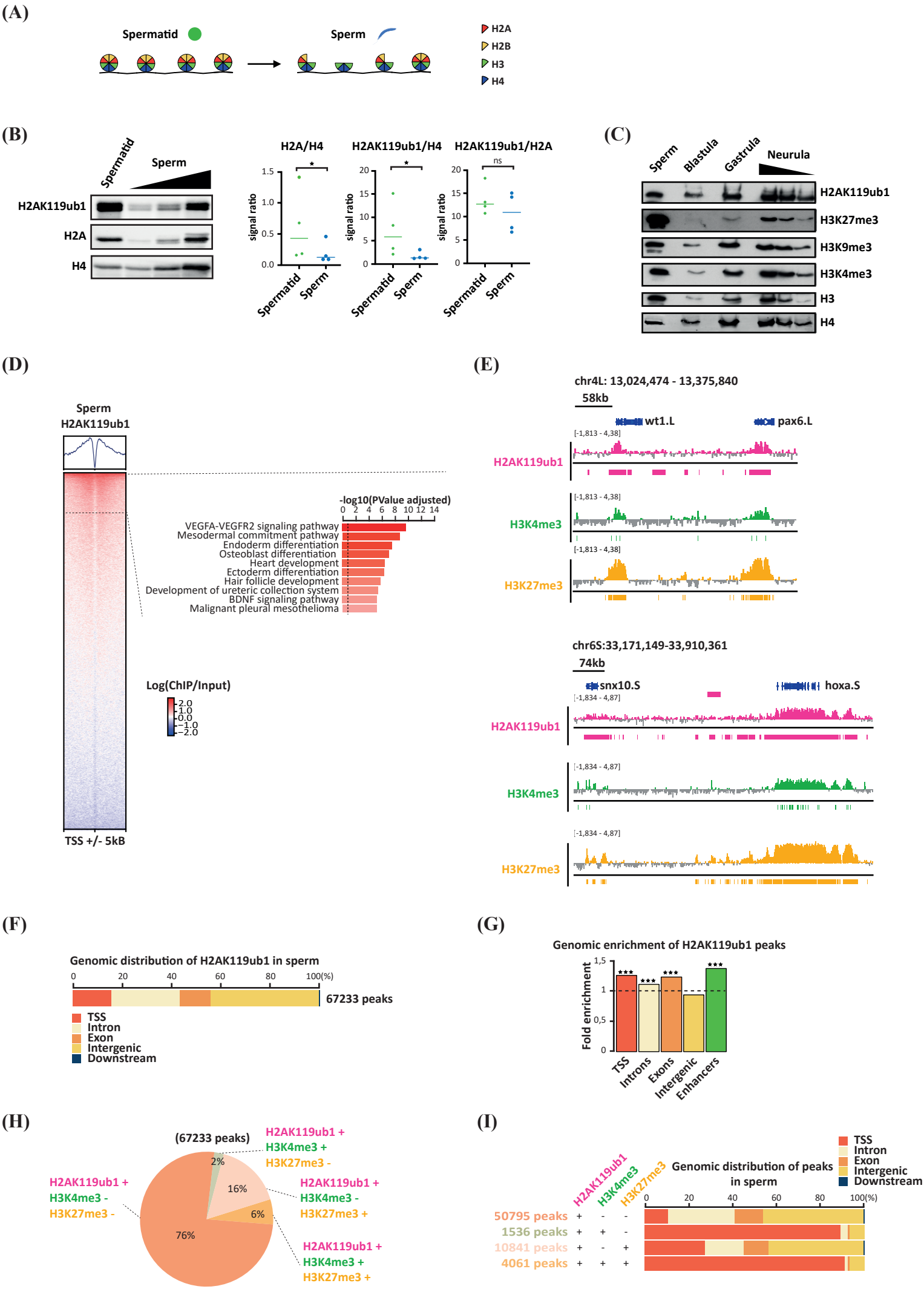
(A)
Spermatid
Sperm
H2A
H2B
H3
H4
(B)
Spermatid
Sperm
H2AK119ub1
H2A
H4
H2A/H4
H2AK119ub1/H4
H2AK119ub1/H2A
signal ratio
Spermatid
Sperm
(C)
Sperm
Blastula
Gastrula
Neurula
H2AK119ub1
H3K27me3
H3K9me3
H3K4me3
H3
H4
(D)
Sperm
H2AK119ub1
-log10(PValue adjusted)
VEGFA-VEGFR2 signaling pathway
Mesodermal commitment pathway
Endoderm differentiation
Osteoblast differentiation
Heart development
Ectoderm differentiation
Hair follicle development
Development of ureteric collection system
BDNF signaling pathway
Malignant pleural mesothelioma
Log(ChIP/Input)
TSS +/- 5kB
(E)
chr4L: 13,024,474 - 13,375,840
58kb
wt1.L
pax6.L
H2AK119ub1
H3K4me3
H3K27me3
chr6S:33,171,149-33,910,361
74kb
snx10.S
hoxa.S
(F)
Genomic distribution of H2AK119ub1 in sperm
67233 peaks
TSS
Intron
Exon
Intergenic
Downstream
(G)
Genomic enrichment of H2AK119ub1 peaks
Fold enrichment
TSS
Introns
Exons
Intergenic
Enhancers
(H)
(67233 peaks)
H2AK119ub1 +
H3K4me3 +
H3K27me3 -
H2AK119ub1 +
H3K4me3 -
H3K27me3 -
H2AK119ub1 +
H3K4me3 -
H3K27me3 +
H2AK119ub1 +
H3K4me3 +
H3K27me3 +
2%
16%
6%
76%
(I)
H2AK119ub1
H3K4me3
H3K27me3
Genomic distribution of peaks in sperm
TSS
Intron
Exon
Intergenic
Downstream
50795 peaks
1536 peaks
10841 peaks
4061 peaks

FIGURE 1

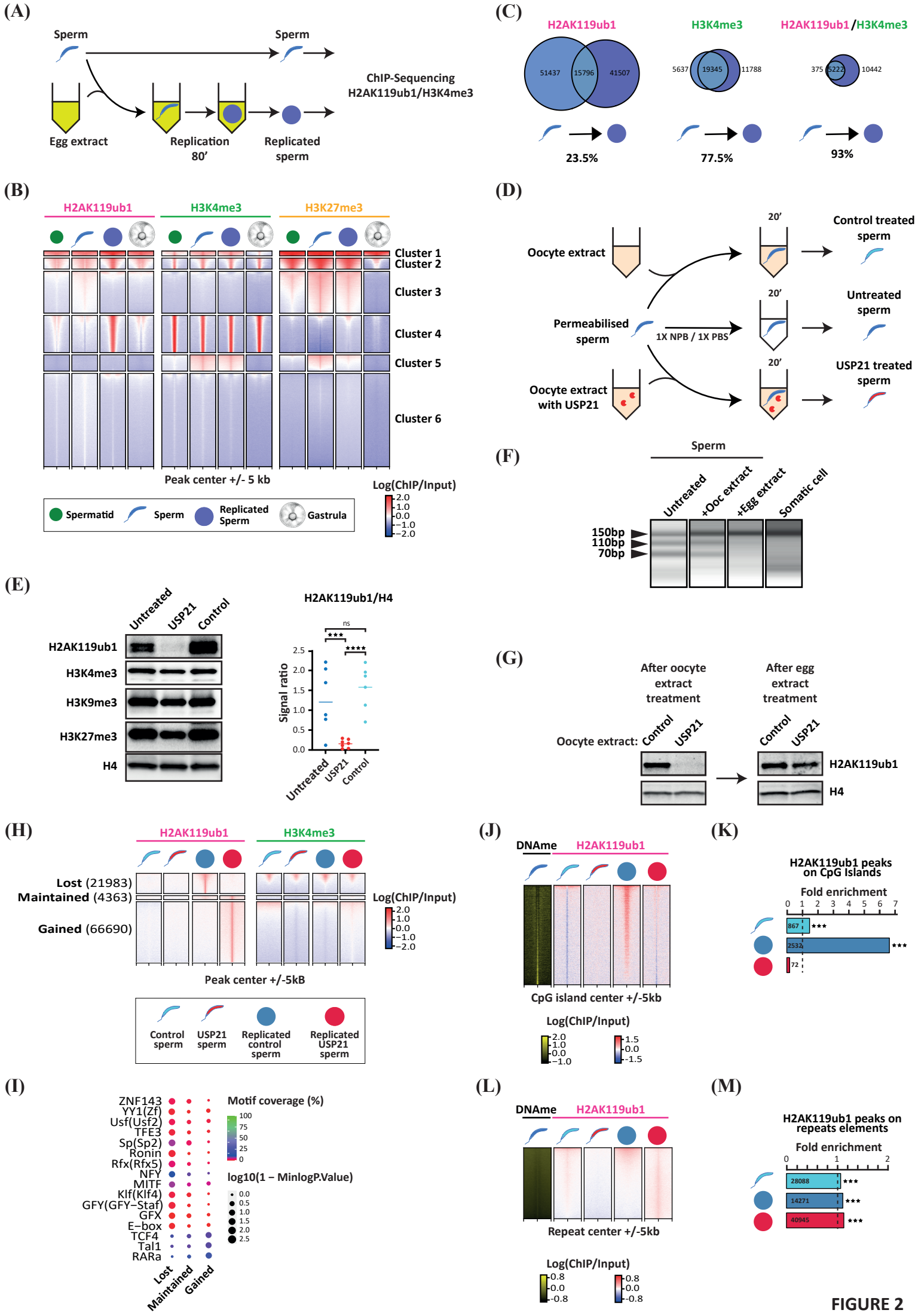
(A)
Sperm
Sperm
Egg extract
Replication 80'
Replicated sperm
ChIP-Sequencing H2AK119ub1/H3K4me3
(B)
H2AK119ub1
H3K4me3
H3K27me3
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Peak center +/- 5 kb
Log(ChIP/Input)
2.0
1.0
0.0
−1.0
−2.0
Spermatid
Sperm
Replicated Sperm
Gastrula
(C)
H2AK119ub1
H3K4me3
H2AK119ub1 / H3K4me3
51437
15796
41507
5637
19345
11788
375
5222
10442
23.5%
77.5%
93%
(D)
Oocyte extract
Permeabilised sperm
1X NPB / 1X PBS
Oocyte extract with USP21
20'
20'
20'
Control treated sperm
Untreated sperm
USP21 treated sperm
(E)
Untreated
USP21
Control
H2AK119ub1
H3K4me3
H3K9me3
H3K27me3
H4
H2AK119ub1/H4
Signal ratio
2.5
2.0
1.5
1.0
0.5
0.0
***
****
ns
Untreated
USP21
Control
(F)
Sperm
Untreated
+Ooc extract
+Egg extract
Somatic cell
150bp
110bp
70bp
(G)
After oocyte extract treatment
After egg extract treatment
Oocyte extract:
Control
USP21
Control
USP21
H2AK119ub1
H4
(H)
H2AK119ub1
H3K4me3
Lost (21983)
Maintained (4363)
Gained (66690)
Log(ChIP/Input)
2.0
1.0
0.0
−1.0
−2.0
Peak center +/-5kB
Control sperm
USP21 sperm
Replicated control sperm
Replicated USP21 sperm
(I)
ZNF143
YY1(Zf)
Usf(Usf2)
TFE3
Sp(Sp2)
Ronin
Rfx(Rfx5)
NFY
MITF
Klf(Klf4)
GFY(GFY−Staf)
GFX
E−box
TCF4
Tal1
RARa
Lost
Maintained
Gained
Motif coverage (%)
100
75
50
25
0
log10(1 − MinlogP.Value)
0.0
0.5
1.0
1.5
2.0
2.5
(J)
DNAme
H2AK119ub1
CpG island center +/-5kb
Log(ChIP/Input)
2.0
1.0
0.0
−1.0
1.5
0.0
-1.5
(K)
H2AK119ub1 peaks on CpG Islands
Fold enrichment
0
1
2
3
4
5
6
7
867
2532
72
***
***
(L)
DNAme
H2AK119ub1
Repeat center +/-5kb
Log(ChIP/Input)
0.8
0.0
-0.8
0.8
0.0
-0.8
(M)
H2AK119ub1 peaks on repeats elements
Fold enrichment
0
1
2
28088
14271
40945
***
***
***

FIGURE 2

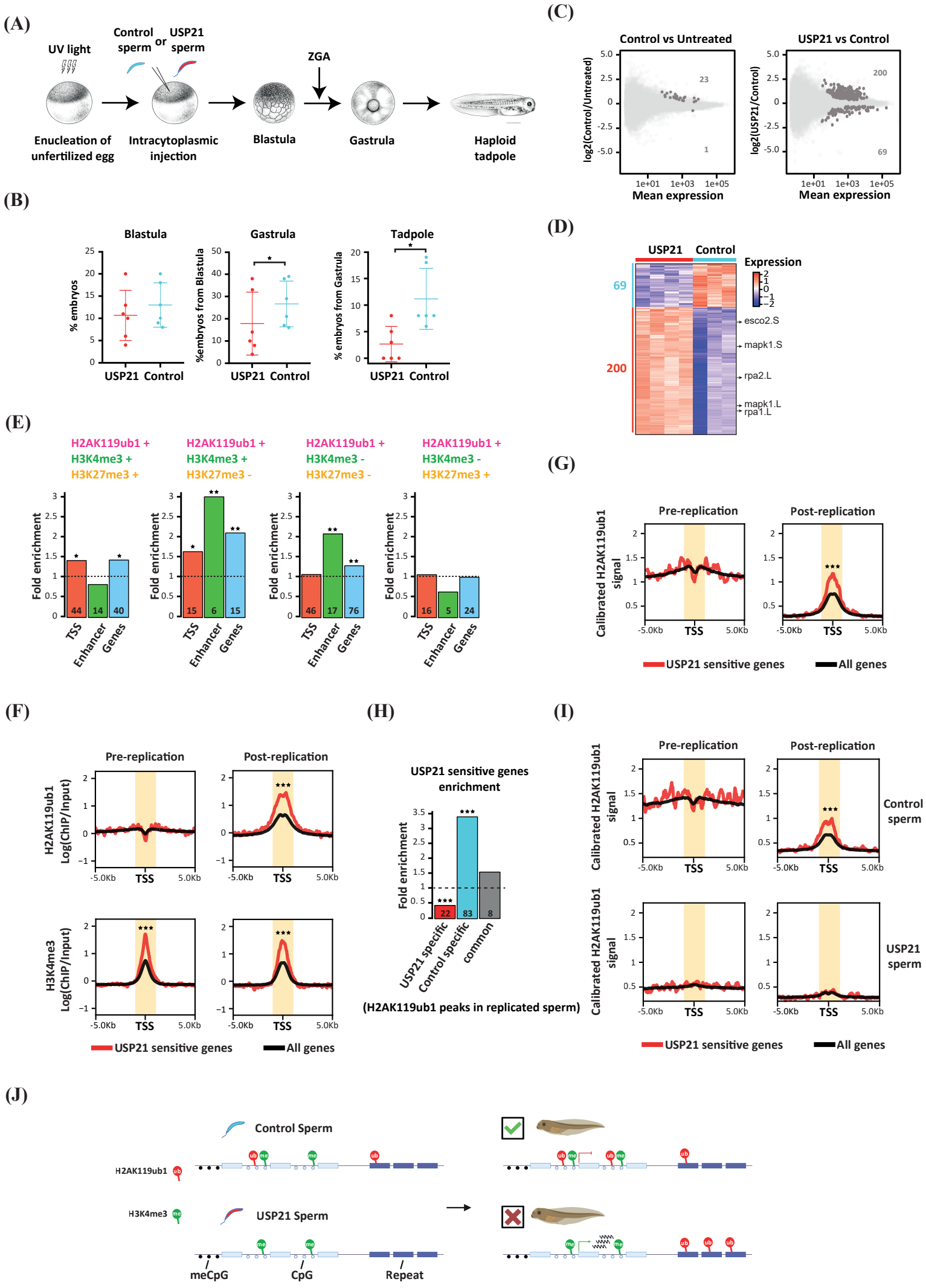
(A)
UV light
Control sperm
or
USP21 sperm
ZGA
Enucleation of unfertilized egg
Intracytoplasmic injection
Blastula
Gastrula
Haploid tadpole
(B)
Blastula
Gastrula
Tadpole
% embryos
%embryos from Blastula
% embryos from Gastrula
USP21
Control
(C)
Control vs Untreated
USP21 vs Control
log2(Control/Untreated)
log2(USP21/Control)
Mean expression
(D)
USP21
Control
Expression
69
200
esco2.S
mapk1.S
rpa2.L
mapk1.L
rpa1.L
(E)
H2AK119ub1 +
H3K4me3 +
H3K27me3 +
H2AK119ub1 +
H3K4me3 +
H3K27me3 -
H2AK119ub1 +
H3K4me3 -
H3K27me3 -
H2AK119ub1 +
H3K4me3 -
H3K27me3 +
Fold enrichment
TSS
Enhancer
Genes
(F)
Pre-replication
Post-replication
H2AK119ub1 Log(ChIP/Input)
H3K4me3 Log(ChIP/Input)
-5.0Kb
TSS
5.0Kb
USP21 sensitive genes
All genes
(G)
Pre-replication
Post-replication
Calibrated H2AK119ub1 signal
USP21 sensitive genes
All genes
(H)
USP21 sensitive genes enrichment
Fold enrichment
USP21 specific
Control specific
common
(H2AK119ub1 peaks in replicated sperm)
(I)
Pre-replication
Post-replication
Calibrated H2AK119ub1 signal
Control sperm
USP21 sperm
USP21 sensitive genes
All genes
(J)
Control Sperm
H2AK119ub1
H3K4me3
USP21 Sperm
meCpG
CpG
Repeat

FIGURE 3