Title: Sperm is epigenetically programmed to regulate gene transcription in embryos

Authors: Marta Teperek[1,2]*[†], Angela Simeone[1,2][†], Vincent Gaggioli[1,2], Kei Miyamoto[1,2], George E. Allen[1], Serap Erkek[3,4,5], Taejoon Kwon[6], Edward M. Marcotte[6], Philip Zegerman[1,2], Charles R. Bradshaw[1], Antoine H.F.M. Peters[3,4], John B. Gurdon[1,2], Jerome Jullien[1,2]*

Affiliations:

[1]Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QN, United Kingdom

[2]Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, United Kingdom

[3]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland

[4]Faculty of Sciences, University of Basel, Basel, Switzerland

[5]present address: EMBL Heidelberg, Meyerhofstrasse 1, 69117 Heidelberg, Germany

[6] Department of Molecular Bioscience, Center for Systems and Synthetic Biology, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, Texas, 78712, The United States of America

[†] Equal contribution

*Corresponding authors. E-mail: j.jullien@cam.ac.uk; mt446@cam.ac.uk

1   **Running title:** Sperm is programmed for embryonic transcription

2

3   **Abstract**

4   For a long time it has been assumed that the only role of sperm at fertilization is to

5   introduce the male genome into the egg. Recently, ideas have emerged that the epigenetic

6   state of the sperm nucleus could influence transcription in the embryo. However, conflicting

7   reports have challenged the existence of epigenetic marks on sperm genes, and there are no

8   functional tests supporting the role of sperm epigenetic marking on embryonic gene

9   expression. Here we show that sperm is epigenetically programmed to regulate embryonic

10 gene expression.

11 By comparing the development of sperm- and spermatid-derived frog embryos we

12 show that the programming of sperm for successful development relates to its ability to

13 regulate transcription of a set of developmentally important genes. During spermatid

14 maturation into sperm, these genes lose H3K4me2/3 and retain H3K27me3 marks.

15 Experimental removal of these epigenetic marks at fertilization deregulates gene expression in

16 the resulting embryos in a paternal chromatin dependent manner. This demonstrates that

17 epigenetic instructions delivered by the sperm at fertilization are required for correct

18 regulation of gene expression in the future embryos.

19 The epigenetic mechanisms of developmental programming revealed here are likely to

20 relate to the mechanisms involved in transgenerational transmission of acquired traits.

21 Understanding how parental experience can influence development of the progeny has broad

22 potential for improving human health.

23

24

1   **Keywords:** *Xenopus*, programming, sperm, spermatid, histone post-translational

2   modifications, transcription, H3K27, H3K4, epigenetics, embryonic development

3

4   **Introduction**

5

6   Embryos obtained by fertilization develop to adulthood more frequently than those

7   obtained by other methods, such as nuclear transfer (Gurdon et al. 1958; Kimura and Yanagimachi

8   1995), suggesting that sperm is specially programmed to support normal embryonic

9   development. Several hypotheses were proposed to explain the nature of this programming,

10   including the idea that sperm is programmed for efficient replication after fertilization

11   (Lemaitre et al. 2005) or for supporting proper embryonic transcription (Suzuki et al. 2007; Hammoud

12   et al. 2009; Zheng et al. 2012). The latter hypothesis was proposed following the observation that

13   promoters of developmentally-important genes escape global replacement of histones by

14   protamines in mature sperm. In fact, these promoters retain post-translationally modified

15   histones, suggesting that epigenetic marks on sperm chromatin may be transmitted to the

16   embryo at fertilization and could subsequently pattern transcription of embryonic genes

17   (Suzuki et al. 2007; Hammoud et al. 2009; Brykczynska et al. 2010; Wu et al. 2011; Paradowska et al. 2012;

18   Zheng et al. 2012; Ihara et al. 2014; Siklenka et al. 2015). However, the validity of this hypothesis was

19   recently questioned (Carone et al. 2014; Samans et al. 2014).

20   In this work we compared the developmental potential of sperm to that of spermatids

21   in order to understand the nature of sperm programming for development in *Xenopus laevis*.

22   The use of spermatids, immediate precursors of sperm, suits such comparisons because: 1.

23   spermatids have completed meiosis and have the same DNA content as sperm, 2. spermatid

24   chromatin structure resembles that of somatic cells (Gaucher et al. 2010). Furthermore, in the

25   mouse, embryos derived from injection of spermatids into unfertilized oocytes develop to

adulthood less frequently than embryos derived from injection of sperm, suggesting that developmentally important information is acquired during spermatid to sperm maturation (Kimura and Yanagimachi 1995; Kishigami et al. 2004). Here, we demonstrate that sperm is epigenetically programmed to regulate transcription of several developmentally important embryonic genes.

Results

Sperm-derived embryos develop better than spermatid-derived embryos

We first compared the development of embryos obtained by transplanting somatic cell nuclei (SCNT) with the development of sperm-derived embryos. To minimize the experimental difference in the way the embryos were generated, both types of embryos were generated by nuclear injection: cloned embryos were obtained by transplanting nuclei of embryonic cells to enucleated eggs (Fig. 1A), and sperm-derived embryos were produced by intracytoplasmic injection of sperm (ICSI) to eggs (Fig. 1C). We observed that cloned embryos developed less efficiently to the swimming tadpole stage than sperm-derived embryos (Fig. 1B). This illustrates the better developmental potential of sperm over that of a somatic cell. In this experimental set up however, the way the embryos are generated is quite different: the maternal genome is present in sperm-derived embryos whereas it has been removed in the SCNT embryos. In order to better assess the developmental potential of sperm, we aimed to compare embryos produced in the same way. For that purpose, we generated embryos by injecting permeabilized purified sperm or spermatids (Supplemental Fig. S1) to the cytoplasm of unfertilized eggs (ICSI, Fig. 1C). In that way both types of embryos are obtained in the same manner and their development can be compared. The two types of embryos reached the gastrula stage with a similar frequency. However, sperm-

1   derived embryos developed significantly better to the swimming tadpole stage than

2   spermatid-derived ones (p-value < 0.05) (Fig. 1D&1E). This is in agreement with

3   observations made previously in mouse (Kimura and Yanagimachi 1995; Kishigami et al. 2004).

4   Spermatid-derived embryos and cloned embryos exhibit a similar reduction in developmental

5   potential when compared to sperm-derived embryos (Fig. 1B-1E), suggesting that the

6   spermatid is as severely impaired to support development as a somatic cell.

7   　　　　In conclusion, embryos can be generated in the same way from sperm or spermatids,

8   and spermatids show reduced developmental potential compared to sperm. Therefore, the

9   comparison of sperm and spermatids was subsequently used to investigate why sperm support

10   better development than other cell types.

11

12   ***Spermatids replicate their DNA as efficiently as sperm***

13   　　　　Since it has been shown that sperm as opposed to other cells replicate DNA more

14   efficiently (Lemaitre et al. 2005), we hypothesized that the poor embryonic development of

15   spermatid-derived embryos is due to inefficient DNA replication. Egg extracts from *Xenopus*

16   have been widely used to investigate mechanisms of replication (Hutchison et al. 1989; Lemaitre et

17   al. 2001). By incubating nuclei in egg extracts one can mimic replication events that occur

18   prior to the first embryonic cell division. We measured DNA replication efficiency in sperm

19   and spermatids incubated in egg extracts {Hutchison, 1989 #16) by molecular combing analysis

20   (Gaggioli et al. 2013). In this assay the egg extract is supplemented with modified nucleotide

21   precursors that will be incorporated into replicating DNA (Fig. 2A). Following replication,

22   DNA is stretched on a slide, and both the total (green) and replicated (red) DNA fibers are

23   fluorescently labeled (Fig. 2B). By measuring the extent of replication on several hundreds of

24   DNA fibers, we observed equally efficient DNA replication in both sperm and spermatids

25   (Fig. 2C). We concluded from this analysis that the nature of sperm programming is not

1   related to replication. We then tested whether spermatid-derived embryos are capable of

2   correctly initiating embryonic transcription.

3

4   *Haploid sperm-derived embryos develop better than haploid spermatid-derived embryos*

5   To rigorously assess the developmental potential and transcriptional ability of sperm

6   and spermatids, and to eliminate the risk of any potential interference from the maternal

7   nucleus, we used haploid paternally-derived embryos generated by injection of permeabilized

8   sperm or spermatids into enucleated eggs (Narbonne et al. 2011) (Fig. 3A and Supplemental Fig.

9   S2). We first confirmed that haploid sperm-derived embryos developed significantly better to

10   the swimming tadpole stage than haploid spermatid-derived embryos (p-value < 0.05) (Fig.

11   3B), recapitulating the results from diploid embryos (Fig. 1E). Previous mammalian

12   experiments comparing developmental potential of sperm and spermatids used diploid,

13   biparental embryos (Kimura and Yanagimachi 1995; Kishigami et al. 2004). Our results with paternal

14   haploid embryos directly demonstrate that the sperm nucleus supports better development

15   than the spermatid nucleus, with or without the maternal genome. Furthermore, this

16   experimental setup allows a specific assessment of transcription originating exclusively from

17   sperm- or spermatid-derived chromatin at the time of embryonic gene activation.

18

19   *Developmentally important genes are misregulated in spermatid-derived embryos*

20   Since experiments in the mouse suggested that sperm might be better than other cell

21   types at supporting mRNA transcription (Ziyyat and Lefevre 2001; Vassena et al. 2007; Ihara et al.

22   2014), we tested this hypothesis using haploid sperm- and spermatid-derived embryos.

23   Embryos were rigorously staged (see Supplemental experimental procedures) and collected at

24   gastrulation, before the onset of developmental defects. Gene expression was then analyzed

25   by RNA-seq, using a set of 39,384 transcripts (provided as Supplemental gene

annotation). Out of 18340 expressed genes, 255 were differentially expressed in spermatid-derived embryos compared to sperm-derived embryos (FDR < 0.05) (Supplemental Table S1). 100 of these transcripts showed consistent changes in at least 6 out of 7 experiments and we will refer to them as 'misregulated' (Fig. 3C, Supplemental Table S1). The majority (82/100) were upregulated in spermatid-derived embryos and they include transcriptional regulators (e.g. *gata2*, *gata3*, *hes1* and *fos)* as well as morphogens (e.g. *bmp2*, *bmp7* or *dhh*) essential for embryonic development. Accordingly, gene enrichment analysis revealed that several development-related terms were significantly enriched in the list of misregulated genes (p-value < 0.05) (Fig. 3D).

The fact that most of the misregulated genes were upregulated in spermatid-derived embryos raised the possibility that these genes were actively transcribed in spermatids and continued to be transcribed in embryos. Indeed the spermatid, as opposed to the sperm, is a transcriptionally active cell and this difference might explain why genes are upregulated in embryos originating from spermatids rather than sperm. To test this possibility we compared the expression level of genes in spermatids and in spermatid-derived embryos. We did not observe any correlation between the expression of misregulated genes in spermatid-derived embryos and their expression in spermatids (Fig. 3E, r = - 0.17, p-value < 0.05). This suggests that the upregulation of these genes in spermatid-derived embryos is not the result of transcript carry-over (or ongoing transcription) from spermatid chromatin. Because permeabilised spermatids used to generate embryos are likely to contain additional RNAs, we performed an additional control for the potential effect of spermatid-derived RNAs on embryonic development. We purified total RNAs from testis, and injected a quantity corresponding to the amount found in one spermatid (50pg) to embryos generated with sperm. TRIzol® was used to isolate RNAs as it allows the recovery of RNA in a broad range of sizes (El-Khoury et al. 2016). All embryos generated in that way developed normally,

1   indicating that testicular RNAs are not detrimental to embryonic development (Supplemental

2   Fig. S3). Lastly we have verified that the synthesis of rRNAs is not affected in spermatid-

3   derived embryos. Indeed, in mouse, defect in rRNAs synthesis has been proposed to explain

4   the developmental defect of nuclear transfer embryos (Zheng et al. 2012) (Suzuki et al. 2007) , and

5   could explain the difference in the bulk of RNA synthesis observed between sperm and

6   spermatid derived embryos (Bui et al. 2011). We observed that newly synthesized 18S and 28S

7   rRNA are equally well produced in sperm and spermatid derived embryos (Supplemental Fig.

8   S4).

9   We conclude that the developmental failure of spermatid-derived embryos is not

10 associated with carried over spermatid RNAs or with defects in rRNAs expression. Rather,

11 we observe a correlation between developmental defects and misexpression of a set of

12 developmentally important genes in spermatid-derived embryos. We hypothesised that

13 differences in gene expression between sperm- and spermatid-derived embryos might result

14 from epigenetic differences of sperm/spermatid chromatin.

15

16 *Epigenetic differences between sperm and spermatid chromatin*

17   To investigate potential links between the epigenetic marking of paternal chromatin

18 and embryonic gene expression, we carried out epigenetic profiling of mononucleosomal

19 chromatin from sperm and spermatid using an extensive MNase digestion protocol applied by

20 others to probe for histones stably associated with chromatin in mature sperm in mouse and

21 human (Hammoud et al. 2009; Brykczynska et al. 2010) (Supplemental Table S2 and S3,

22 Supplemental Fig. S5). In *Xenopus*, the transition from spermatid to sperm is characterized by

23 histone H3 and H4 retention and partial loss of H2A and H2B (Risley and Eckhardt 1981). We

24 first compared nucleosome occupancy profiles in sperm and spermatids. Similarly to what

25 was observed in other vertebrates (Hammoud et al. 2009; Brykczynska et al. 2010), we observed

8

1   higher nucleosomal occupancy (MNase-seq, Fig. 4A) around TSSs (transcriptional start

2   sites) in sperm when compared to spermatids. In this context, the positioned nucleosomes

3   show a similar distribution in sperm and spermatids (Supplemental Fig. S6). Secondly, we

4   analysed DNA methylation profiles in sperm and spermatids by MBD-seq. DNA methylation

5   was higher around sperm TSSs than spermatid TSSs (Fig. 4B). These differences were

6   observed at the genome-wide level between sperm and spermatids, as well as on the set of

7   misregulated genes (Supplemental Fig. S6 and S7). A lower level of nucleosome occupancy

8   and DNA methylation in spermatids compared to sperm could therefore explain the

9   upregulation of genes in spermatid- compared to sperm-derived embryos.

10   To further characterize sperm and spermatid chromatin, we performed ChIP-seq

11   analysis of histone marks associated with activation (H3K4me2, H3K4me3) or repression

12   (H3K27me3, H3K9me3) of transcription. We looked for peaks (Fig. 4C, Supplemental Table

13   S4) as well as for the overall methylation levels (Fig. 4D, Supplemental Fig. S8,

14   Supplemental Table S5) around TSSs for each of these modifications. When compared to all

15   genes, the set of misregulated genes was significantly enriched for H3K27me3 in both sperm

16   and spermatids (Fig. 4C). However, since H3K27me3 was present in both sperm and

17   spermatids, it means that this repressive mark alone cannot explain the difference in gene

18   expression observed in sperm- and spermatid-derived embryos.

19   Interestingly, histone marks associated with active transcription (H3K4me2/3)

20   showed an enrichment at promoters of misregulated genes in spermatids but not in sperm

21   (Fig. 4C, Fig. 4D), providing a plausible explanation for the upregulation of these genes in

22   spermatid-derived embryos.

23

24   ***Co-existence of H3K4me2/3 and H3K27me3 in spermatids correlates with embryonic gene***

25   ***upregulation***

9

1   The epigenetic features analysed (histone marks, DNA methylation and nucleosome

2 occupancy) can individually provide a possible explanation for the observed differences in

3 expression in sperm- and spermatid-derived embryos. However, complex interactions

4 involving more than one epigenetic feature might better explain this differential embryonic

5 gene expression. In order to identify such interactions we have performed a partial correlation

6 analysis. In this analysis all the measured parameters are assessed simultaneously to produce

7 maps indicating the way epigenetic features associate with each other and with gene

8 expression. The aim of such analysis is to extract more general principles describing the

9 paternal epigenetic program underlying gene expression and identify which epigenetic

10 features are likely to have the strongest contribution to embryonic gene expression. We

11 applied the partial correlation analysis to identify links between the measured epigenetic

12 features in sperm and the expression of the misregulated genes in sperm-derived embryos

13 (Fig. 5A), or between the measured epigenetic features in spermatid and their expression in

14 spermatid-derived embryos (Fig. 5B). We observed that sperm H3K4me2/3 and embryonic

15 gene expression were positively associated, whereas sperm H3K27me3, H3K9me3 and DNA

16 methylation are negatively linked to embryonic gene expression (Fig. 5A). Interestingly,

17 activating H3K4me2/3 and repressive H3K27me3 marks in spermatids were positively linked

18 to gene expression in spermatid-derived embryos and at the same time they were also

19 strongly associated with each other (Fig. 5B). These associations were also observed when

20 performing a similar analysis using an extended set of misregulated genes obtained by

21 relaxing the selection parameters from FDR≤ 0.05 (255 genes) to FDR ≤ 0.4 and |logFC| ≥

22 0.2 (1116 genes). The use of this extended set increased the predictive power of the analysis

23 and showed stronger links between the features tested within an overall similar network

24 (Supplemental Fig. S9). Therefore, the difference between sperm- and spermatid-derived

25 embryos is best explained by the fact that in contrast to sperm, where H3K27me3 is

10

overrepresented at genes differentially expressed in haploid embryos, in spermatids H3K27me3 co-exists with H3K4me2/3 on these genes, thereby contributing to their upregulation in spermatid-derived embryos.

We checked whether the observed distribution of H3K4me2/3 and H3K27me3 on the misregulated genes of *Xenopus* sperm was a conserved feature across species. For that purpose we investigated how these histone marks were distributed in human sperm (Hammoud et al. 2009), on human orthologues of the *Xenopus* misregulated genes. We observed that similarly to *Xenopus* sperm, human sperm showed an enrichment for H3K27me3 on misregulated genes (Fig. 5C). Additionally, misregulated genes did not exhibit any enrichment over the genome-wide distribution for H3K4me2 in both species. This indicates a conservation of sperm epigenetic features on these genes in the two species.

We next tested if paternally-derived H3K4me2/3 and H3K27me3 were indeed involved in patterning embryonic gene expression.

Paternal H3K4me2/3 and H3K27me3 influence embryonic gene expression

To test the function of epigenetic marks from sperm or spermatids chromatin on the regulation of embryonic transcription, we experimentally removed these marks in embryos (Fig. 6A and Supplemental Fig. S10). mRNAs encoding histone demethylases or control mRNAs were first injected into immature oocytes. After allowing 24h for the enzymes to be expressed, the oocytes were *in vitro* matured into eggs (IVM) and injected with sperm or spermatids (ICSI). The resulting embryos were collected at gastrulation stage for RNA-seq analysis. In this protocol histones from both maternal and paternal chromatins are demethylated when the embryo is generated. By comparing embryos produced with different paternal chromatins (sperm or spermatid) we can evaluate the effect of paternal epigenetic mark removal on embryonic gene expression.

1   We first expressed the H3K4me2/3 demethylase, KDM5B. As expected, removal of

2   the activating H3K4me2/3 marks led to gene downregulation: 68% (1893 genes) and 80%

3   (1392 genes) of all differentially expressed (FDR < 0.05) genes were downregulated in

4   sperm- and spermatid-derived embryos, respectively (Fig. 6B, Supplemental Table S6).

5   Importantly, genes downregulated in sperm-derived embryos showed only limited overlap

6   with those downregulated in spermatid-derived embryos (Fig. 6C). This indicates a paternal

7   chromatin dependent effect of H3K4me2/3 removal on embryonic gene expression.

8   Additionally, among the genes affected by H3K4me2/3 removal in sperm- and spermatid-

9   derived embryos, the misregulated genes are overrepresented, indicating that paternal

10   H3K4me2/3 specifically regulates this set of genes (Fig. 6E). Interestingly, the genes

11   downregulated in spermatid-derived embryos are enriched for H3K4me2/3 in spermatids

12   (Fig. 6D). These observations are in agreement with the hypothesis that the loss of

13   H3K4me2/3 from H3K27me3 marked genes during the spermatid to sperm maturation is

14   necessary for their proper expression in embryos.

15   To further validate this hypothesis we tested the influence of paternal H3K27me3 by

16   overexpressing the H3K27me3 demethylase KDM6B (Fig. 6A). In accordance with its

17   repressive function, removal of H3K27me3 at fertilization led to upregulation of genes at

18   gastrulation in both sperm- and spermatid-derived embryos. 87% (487 genes) and 76% (173

19   genes) of differentially expressed genes (FDR < 0.05) were upregulated in sperm- and

20   spermatid-derived embryos, respectively (Fig. 6F, Supplemental Table S6). Again, there was

21   only a partial overlap between genes affected by KDM6B in sperm- and spermatid-derived

22   embryos (Fig. 6G), indicating that this effect is paternal chromatin dependent. The affected

23   genes were marked by H3K27me3 in the corresponding paternal cells (Fig. 6H).

24   Additionally, upon H3K27me3 demethylation about five times more genes were specifically

25   upregulated in sperm- than in spermatid-derived embryos (402 versus 88 genes). This

12

1   suggests that the programming of genes for embryonic expression in the paternal chromatin

2   relies on the establishment of an effective H3K27me3-mediated repression at the spermatid to

3   sperm transition. Lastly, the misregulated genes are enriched among the genes affected by the

4   H3K27me3 removal indicating that paternal H3K27me3 specifically regulates this set of

5   genes (Fig. 6E).

6

7

8   **Discussion**

9   Previous works characterizing the epigenetic features of sperm in zebrafish, mouse

10   and human have revealed the presence of modified histones around genes involved in

11   embryonic development (Hammoud et al. 2009; Brykczynska et al. 2010; Wu et al. 2011). In these

12   species, the presence of activating (H3K4me3) and repressive (H3K27me3) histone marks in

13   sperm correlated with gene expression in the early embryos (Hammoud et al. 2009; Brykczynska et

14   al. 2010; Wu et al. 2011). In this work we have used a comparison of sperm and its immediate

15   precursor, the spermatid, to investigate the functional relationship between histone marks and

16   gene expression.

17   Our analysis shows that, similarly to what has been observed in mouse (Kimura and

18   Yanagimachi 1995; Kishigami et al. 2004), spermatids are not as good at supporting development as

19   sperm. Second, we tested several hypotheses proposed to explain the developmental

20   advantage of sperm over spermatids. We have ruled out the hypothesis that spermatids are

21   less efficient than sperm at supporting replication. Instead, we found evidence supporting the

22   hypothesis that sperm is programmed to support proper embryonic expression of genes

23   encoding important embryonic regulators (Fig. 3). Importantly, overexpression and knock

24   down studies of several of these genes have shown embryonic developmental defects

25   reminiscent of what is observed in spermatid-derived embryos *(sfrp2* (Lee et al. 2006); *tbx3*

13

(Weidgang et al. 2013), *foxa2* (Suri et al. 2004), *otx2* (Yasuoka et al. 2014)). These observations suggests that misexpression of this set of genes is the cause of the developmental defect observed in spermatid-derived embryos. We also showed that the developmental advantage of sperm over spermatids is maintained in haploid, paternally derived embryos, indicating that the effect observed is independent of the presence of the maternal genome. To our knowledge this is the first time that these two hypotheses, developmental advantage related to ability to support replication *versus* transcription, have been rigorously tested.

These analyses allowed us to conclude that sperm is not merely a carrier of DNA, but that it also contributes epigenetic information required for proper embryonic gene expression. We then focused our analysis on the sperm chromatin as it represents the most likely vector of such epigenetic information.

During spermiogenesis in *Xenopus laevis* core histones H3 and H4 are retained whereas ~90% of core histones H2A and H2B are lost(Risley and Eckhardt 1981). This leaves *Xenopus* sperm with about 10% of the amount of nucleosomal content of a spermatid. This level of histone retention in sperm is higher than that of mouse (~1%) (Brykczynska et al. 2010), lower than that of zebrafish (~100%) (Wu et al. 2011), and similar to that of human (~10%) (Brykczynska et al. 2010). Nucleosome retention in vertebrates therefore seems to show a degree of variation among species. The epigenetic analysis of *Xenopus* sperm provided here extends the repertoire of characterized higher vertebrate sperm chromatin and identifies conserved chromatin features relevant to developmental programming. In that respect we observed that the programming of sperm for embryonic gene expression entails a loss of H3K4me2/3 marking at H3K27me3 target genes during spermatid to sperm maturation (Fig. 6I). We showed that the set of genes programmed for embryonic expression during *Xenopus* sperm maturation had similar epigenetic features in *Xenopus* and human sperm (Fig. 4C and Fig. 5D). So, despite the existence of hugely variable degree of histone retention in sperm among

14

species, this points towards the existence of universal mechanisms preparing sperm for participation in the normal development of embryos in vertebrates.

To functionally test the role of sperm epigenetic marks on embryonic gene expression one would ideally like to erase these marks from the sperm nucleus immediately prior to the generation of embryos. Chromatin of mature sperm is highly condensed and inaccessible, making enzymatic treatments to alter the epigenetic marks inefficient. Alternative strategies have been developed to use such enzymes either during the process of spermiogenesis, prior to full maturation of sperm (Siklenka et al. 2015), or at fertilization when the sperm chromatin becomes again accessible (used in this study). In mouse, the former strategy has been used to overexpress the H3K4/H3K9 demethylase KDM1A in germ cells. Embryos generated with sperm from animals overexpressing KDM1A exhibited developmental defects which were transmitted for several generations in the absence of exogenous KDM1A. This analysis demonstrated the existence of epigenetic instruction delivered by sperm to the embryos and transmitted through several generations. However, the overexpression of KDM1A very early in the process of sperm differentiation leads to numerous abnormalities in sperm, for example the presence of additional mRNAs (Siklenka et al. 2015). These abnormalities are indirect effects of the overexpression of KDM1A early in the process of sperm differentiation. For that reason, it has been difficult to link the difference in gene expression and associated developmental defects to particular epigenetic changes in sperm. The approach we describe here complements and extends previous analyses. First, by comparing the epigenetic profiles of sperm and spermatids to differential gene expression in sperm- and spermatid-derived embryos we identified H3K4me2/3 and H3K27me3 as candidate marks responsible for the programming of genes. We then tested this hypothesis by demethylating the chromatin using KDM6B (H3K27 demethylase) or KDM5B (H3K4 demethylase) in embryos generated with sperm or spermatids. Importantly, in our experimental setup both sperm and spermatids used

15

had been through a normal differentiation process. Removal of H3K4me2/3 at fertilization affects different set of genes in sperm- and spermatid-derived embryos. Genes affected in sperm-derived embryos are enriched for H3K27me3 whereas genes affected in spermatid-derived embryos are enriched for H3K4me2/3. This indicates the importance of H3K4me2/3 dynamics at the transition from spermatid to sperm for patterning of the future embryonic gene expression. One hypothesis to explain the sensitivity of sperm H3K27me3 marked genes to H3K4me2/3 removal would be that these genes acquire H3K4 methylation following fertilization. Our analysis also suggests a conserved role for these marks in *Xenopus* and mouse (Siklenka et al. 2015). Additionally, we also demonstrated that removal of H3K27me3 at fertilization affects the embryonic expression of genes that are marked by H3K27me3 in sperm/spermatids. Recent reports probing histone modifications distribution in mouse and human sperm suggested that these epigenetic marks occurred mostly on repetitive regions of the genome rather than genes (Carone et al. 2014; Samans et al. 2014). These observations put into question the possibility that such marks would influence gene expression in embryos. By providing functional test of the need for histone modifications for embryonic gene expression, our analysis, together with that of Siklenka *et al*, clearly shows that, regardless of their genomic location, sperm delivered modified histones are important regulators of expression in future embryos (Siklenka et al. 2015).

Further investigations into the nature of sperm programming, especially the requirement of other epigenetic marks and their cross-talk in the patterning of embryonic expression, will provide a better understanding of the transgenerational inheritance of epigenetic traits via gametes, and could shed light on the mechanisms underlying male infertility and other diseases in humans.

Materials and Methods

All the experiments involving the use of animals were conducted according to the regulatory standards of the funding bodies.

Separation of sperm and spermatids

For each round of sperm and spermatid purification, testes from 6 adult *Xenopus laevis* males were isolated and manually cleaned from blood vessels and fat body in 1×MMR (100 mM NaCl, 2 mM KCl, 2 mM CaCl2, 1 mM MgCl2, 5 mM HEPES pH 7.4) using forceps and paper tissues. It is crucial to clean the testes well from any non-testicular tissues, as otherwise the cells released from the tissues may negatively affect the final purity of isolated cells. Subsequently, testes were torn into small pieces with forceps and homogenized with 2-3 strokes of a Dounce homogenizer (tissue from 1 testis at a time). The cell suspension was then filtered to remove tissue debris and cell clumps (CellTrics, cat. 04-0042-2317) and spun down at 800rcf, 4°C, 20 minutes. Supernatant was discarded and cell pellet resuspended in 12ml of 1×MMR. If any red blood cells were visible at the bottom of the pellet (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, is 60% iodixanol in water) in 1×MMR final were manually prepared in pre-chilled 14ml ultra-clear centrifuge tubes (Beckman Coulter, 344060) in the following order from the bottom to the top of the tube: 4ml of 30% iodixanol, 1ml of 20% iodixanol, 5ml of 12% iodixanol (all in 1×MMR) and 2ml of cell suspension in 1×MMR on top. Gradients were spun down in a pre-chilled SW40Ti rotor at 7500rpm (10000g), 4°C, 15 minutes, deceleration without break (Beckman Coulter Ultra-centrifuge, Optima L-100XP). The top interface fraction (between 1×MMR and 12% iodixanol), containing spermatids, and the pelleted fraction, containing mature sperm, were collected.

17

1   Collected fractions were diluted six times with 1×MMR and collected by spinning first at 805

2   rcf, 4°C, 20 minutes and re-spinning at 3220 rcf, 4°C, 20 minutes to pellet remaining cells.

3   Pelleted cells were subjected to nuclei preparation (see below).

4

5   *Sperm and spermatid nuclei preparation, intra-cytoplasmic sperm injections (ICSI) to non-*

6   *enucleated and to enucleated eggs and embryo culture*

7   Sperm and spermatids nuclei were permeabilized as described before (Smith et al. 2006)

8   and stored at -80°C. Injections were performed using a Drummond Nanoject microinjector

9   (NanojectII Auto Nanolitre Injector, Biohit, 3-00-206A) and glass capillaries (Biohit, 3-00-

10   203-G/XL) pulled using a Flaming-Brown micropipette puller (settings: heat 700, pull 100,

11   velocity 100, time 10). Cell suspension was sucked into the injection needle filled with

12   mineral oil. Cells were injected in sperm dilution buffer (SDB) (Smith et al. 2006) and cell

13   concentration was adjusted by doing mock injections on a microscope slide to deliver 1 cell

14   per 4.6nl injection. The eggs were placed in batches of 20-25 on a blotting paper. If they were

15   to be enucleated, they were placed with animal pole facing upwards, whereas if they were not

16   subjected to enucleation, they were placed on a side (with the marginal zone upwards). For

17   enucleation, eggs were treated for 30s with a UV mineralite lamp (Gurdon 1960) (this step was

18   omitted for non-enucleated eggs). Jelly was removed by a 5s Hanovia lamp treatment. The

19   eggs were immediately injected with sperm or spermatid solution and moved to 1×MBS

20   (Gurdon 1976) supplemented with 0.2% bovine serum albumin (BSA). Cell suspension in the

21   needle was replaced every 20-25 batch of eggs injected. At 4-cell stage embryos were sorted

22   (all the non-cleaved embryos or those with irregular cleavage furrows were discarded) and

23   the culture media replaced with 0.1MBS, 0.2% BSA. Embryos were cultured in 0.1MBS,

24   0.2% BSA (changed daily) at 16-18°C incubator. Assessment of developmental stages was

25   performed according to Nieuwkoop and Faber (Nieuwkoop 1994). Using this table, matching

1   gastrula embryos from the various experimental groups were collected at stage $10^{1/2}$-11, and

2   processed for gene expression analysis (see supplemental data procedures for details).

3

4   *Interphase egg extract preparation*

5   Eggs were collected in 1×MMR, dejellied with 0.2×MBS, 2% cysteine (pH 7.8-7.9)

6   (Sigma, W326305) and washed with 0.2×MMR. Subsequently, eggs were activated for 3

7   minutes at RT with 0.2×MMR supplemented with 0.2ug/ml calcium ionophore (Sigma,

8   C7522). Eggs were rinsed with 0.2×MMR and subsequently all abnormal or not activated

9   eggs were removed. Eggs were washed with 50ml of ice-cold extraction buffer (EB) (5mM

10   KCl, 0.5mM $MgCl_2$, 0.2mM DTT, 5mM Hepes pH 7.5) supplemented with protease

11   inhibtiors (PI) (Roche, 11873580001), transferred into centrifugation tube (Thinwall, Ultra-

12   Clear™, 5 mL, 13 × 51 mm tubes, Beckman, 344057) and supplemented with 1ml of EB

13   buffer with PI and 100ug/ml of cytochalasin B (Sigma, C2743) and placed on ice for 10

14   minutes. Subsequently, eggs were spun briefly at 350g for 1 minute at 4°C (SW55Ti rotor,

15   Beckman Coulter Ultra-centrifuge, Optima L-100XP) and excess buffer was discarded. Eggs

16   were then spun at 18000g for 10 minutes at 1°C, the extract was collected with a needle,

17   transferred to a fresh, pre-chilled tube, supplemented with PI and 10ug/ml of cytochalasin B

18   and re-spun using the same conditions. Extract was collected with a needle and used fresh for

19   replication assay (see below).

20

21   *Replication in egg extracts and sample preparation for analysis of DNA fibres*

22   Replication on single DNA fibres was performed as described before (Gaggioli et al.

23   2013) with slight modifications. Freshly prepared egg extracts were supplemented with 20×

24   energy regeneration mix: 2mg/ml Creatine Kinase (Roche, 10127566001), 150mM Creatine

25   Phosphate (Roche, 10621714001), 20mM ATP (Roche, 10519979001), 2mM EGTA, 20mM

19

1   MgCl2) and with 20μM biotin-16-dUTP (Roche, 11093070910). Permeabilized cells were

2   added to a final concentration of 200 nuclei/ul of extract and incubated at RT for 2h (tapping

3   every 10 minutes). Reaction was stopped by adding 10 volumes of ice-cold 1×PBS

4   (Phosphate Buffer Saline: 137mM NaCl, 2.7mM KCl, 10mM $Na_2HPO_4 \times 2H_2O$, 2mM

5   KH_2PO_4) and cells were spun down at 1000g, 4°C, 7 minutes. Cells were resuspended in 50ul

6   of 1×PBS and mixed immediately with 50ul of melted (at 65°C) 2% low melting point

7   agarose (Invitrogen, 16520050) in 1×PBS. After solidification, the agarose plug was

8   incubated overnight (O/N) at 50°C with 1mL 0.5M EDTA pH8, 100uL 10% sarkosyl (Sigma,

9   L5125), 1mg/mL Proteinase K (New England Biolabs, P8102S) followed by three washes in

10   TE pH 6.5. Subsequently, the plug was incubated twice in TE supplemented with 0.1mM

11   PMSF (Sigma, 93482) for 30 minutes at 50°C and washed four times with 1ml of 50mM

12   MES (Sigma, 69889) pH 6.35, 1mM EDTA (1h at RT each wash). Then the solution was

13   removed; the plug was melted in 400ul of MES pH 6.35, 1mM EDTA at 68°C for 20 minutes

14   and the agarose was digested with 2 units of β-agarase (New England Biolabs, M0392S) O/N

15   at 42°C.

16   *Analysis of replication on single DNA fibres*

17   Silanized coverslips were prepared as described before (Labit et al. 2008). Thirty

18   microliters of replicated DNA solution was pipetted onto a silanized coverslip, covered with a

19   non-silanized coverslip and incubated for 5 minutes at RT. Subsequently, the top coverslip

20   was slid away to stretch DNA fibres and the silanized coverslip with stretched fibres was

21   fixed in 3:1 solution of Methanol:Glacial Acetid Acid for 10 minutes, RT. The fibres were

22   then denatured with 2.5M HCl (1h, RT) and dehydrated by washes in 70% ethanol, 90%

23   ethanol and 100% ethanol (1 minute for each wash). Subsequently, the coverslip was dried,

24   washed 3 times in PBS, 0.1% Tween (Sigma, P5927) (5 minutes for each wash) and blocked

in 3%BSA in PBS (1h, RT). All antibodies were diluted in PBS, 3%BSA, 0.1% Tween. Total DNA was detected simultaneously to replicated DNA with primary antibodies: anti-DNA antibody (Millipore, MAB3034) 1:300 dilution, and streptavidin-Alexa 594 antibody 1:50 to detect biotin (Invitrogen, S-11227) for 30 minutes at 37°C. Primary antibodies were washed away with PBS, 0.1% Tween (4 washes) and detected with secondary antibodies diluted 1:50: chicken anti-mouse Alexa 488 (Invitrogen, A-21200) and biotinylated antibody anti-streptavidin (Vector Labs, BA-0500) for 30 minutes, 37°C. After four washes in PBS, 0.1% Tween, a tertiary detection was performed with antibodies diluted 1:50: goat anti-chicken Alexa 488 (Invitrogen, A-11039) and streptavidin-Alexa 594 for 30 minutes, 37°C. The coverslip was washed three times with PBS 0.1% Tween, three times in PBS and mounted on a microscope slide with a mounting medium (50% glycerol in PBS) and sealed with a nail polish. Images were acquired with a Zeiss 510 META confocal LSM microscope. Image analysis was performed in ImageJ; the amount of replicated DNA and total DNA was measured individually on single DNA fibres.

RNA extraction and preparation of cDNA library for sequencing

Spermatid (1 millions) or pool of 5 stage 10.5-11.5 embryos were collected and frozen at -80°C. RNA extractions were performed using Qiagen RNeasy Mini kit according to the manufacturer's protocol. RNA was eluted in 50ul of DEPC H2O and used to generate cDNA sequencing libraries using Illumina TrueSeq kit (RS-122-2001), according to the manufacturer's protocol.

mRNA Injection to 1-cell embryos

Mouse KDM6B (aa1025-1642) or KDM5B (aa1-770) were cloned using p-Entry cloning system (Invitrogen, K2400-20 and 11791-020) into pCS2+ plasmid with a C-terminal

1   HA-tag and NLS-tag. mRNA was synthesized *in vitro* using MEGAscript® SP6 Kit

2   (Ambion, AM1330M) following the manufacturer's instructions. Eggs were *in vitro* fertilized

3   and dejellied using 2% Cysteine solution in 0.1×MMR. Injections into 1-cell stage embryos

4   were performed in injection solution (Smith et al. 2006) using a Drummond Nanoject

5   microinjector, delivering 9.2ng of mRNA per injection (mRNA at 1mg/ml in DEPC H_2O).

6   Embryos were cultured at 18°C and collected for Western Blot analysis at stage 21 (Nieuwkoop

7   1994). Western blot analysis were performed on 12% polyacrylamide gels using antibodies

8   against H3K27me3 (Cell Signalling, 9733), H3K9me3/2 (Cell Signalling, 5327), H3K4me2/3

9   (Abcam 8580), H4 (Abcam 31830), and against H3 (Abcam 18521)

10

11   *Preparation of ChIP-seq samples*

12   Sperm and spermatids were separated as described above. Chromatin fractionation

13   and chromatin immunoprecipitation (ChIP) were performed as described before (Erkek et al.

14   2013; Hisano et al. 2013) with slight modifications. Pre-treatment of sperm cells with DTT

15   was omitted and chromatin was digested with 2.5U of MNase/1 million of cells (Roche,

16   12533700) for 30 minutes at 37°C. The following antibodies against histone marks were used

17   in the study: anti-H3K4me2 (Millipore, 07-030), anti-H3K4me3 (Abcam, ab8580), anti-

18   H3K4me3 (Millipore, CS200580), anti-H3K27me3 (Millipore, 07-449), anti-H3K27me3

19   (kind gift from Dr Thomas Jenuwein), and anti-H3K9me3 (Abcam, ab8898). Before ChIP,

20   primary antibodies were bound to magnetic beads conjugated with secondary antibody

21   (Invitrogen, 11204D) according to manufacturer's protocol and all wash steps in the protocol

22   were carried with a magnet, instead of centrifugation. Bound DNA was isolated, separated by

23   electrophoresis and mononucleosomal bands from sperm and spermatids were excised and

24   subjected to library preparation with TruSeq DNA kit (Illumina, FC-121-2001). For the

25   generation of the input sample, 5-10% of the MNase digested chromatin was collected and

22

1   the same purification scheme was followed as with the immunoprecipitated chromatin prior

2   to library preparation with TruSeq DNA kit (Illumina, FC-121-2001).

3

4   *Preparation of MBD-seq samples*

5   Sperm and spermatids chromatin were separated as described above and 200ng of

6   digested genomic DNA was used to purify methylated DNA using the Methyl Collector TM

7   Ultra kit (Active Motif 55005). The purification was carried out according to the

8   manufacturer's instructions, using the low salt buffer to wash the bead-methyl DNA

9   complexes. The purified methylated DNA and the input DNA were then subjected to library

10 preparation with TruSeq DNA kit (Illumina, FC-121-2001).

11

12 *Sequencing data analysis*

13 Details of the sequencing data analysis methods used in this study are described in

14 Supplemental information section.

15 **Data access**

16 All ChIP-seq, RNA-seq , MBD-seq, and MNase-seq data sets have been submitted to the

17 NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession

18 number GSE75164.

19

11 **Author Contributions**:

12 M.T. and J.J. conceived the study, M.T., A.S. and J.J. analysed the experiments and wrote the

13 paper, M.T. performed the experiments, V.G. performed and analysed replication

14 experiments, AS. performed all bioinformatics analysis, G.A. and A.S. analysed RNA-seq

15 data, K.M. and J.J. performed *in vitro* maturation/ICSI experiments, S.E. and A.H.F.M.P.

16 helped with ChIP-seq experiments, T.K. and E.M. helped to acquire RNA-seq data and

17 provided *Xenopus laevis* transcriptome and genome, C.R.B. customized *Xenopus laevis*

18 genome and transcriptome and performed homology searches, A.S., C.R.B. and G.A. wrote

19 the experimental procedures section for ChIP-seq and RNA-seq. K.M., V.G., A.S., P.Z. and

20 J.B.G. helped with the experimental design and paper writing. J.J. and J.B.G. supervised the

21 research.

22

23

24

1   **Figure legends:**

2

3   **Figure 1.** *Xenopus sperm is better at supporting development than a spermatid or a somatic*

4   *cell.*

5   **(A)** Experimental design for the generation of cloned embryos. The somatic nucleus of a

6   gastrula cell is transplanted to a UV-enucleated egg. The resulting embryos are scored at

7   gastrulation and tadpole stage. **(B)** Scoring of embryos as % of gastrulae and as % of

8   swimming tadpoles to the total number of cleaved embryos (average of n=6 independent

9   experiments (sperm ICSI), and n=3 independent experiments (embryo cell NT). The total

10   number of embryos analyzed is shown above the graph. Error bars: sem. * p-value < 0.05

11   (chi-square test). **(C)** Experimental design for the generation of sperm- and spermatid-derived

12   embryos. Permeabilized sperm or spermatids are injected to the cytoplasm (ICSI) of

13   unfertilized egg. The resulting embryos are scored at gastrulation and tadpole stage. **(D)**

14   Representative images of sperm- and spermatid-embryos. Scale bars = 1mm. **(E)** Scoring of

15   embryos as % of gastrula and as % of swimming tadpoles to the total number of cleaved

16   embryos (average of n=6 independent experiments). The total number of embryos analyzed is

17   shown above the graph. Error bars: sem. * p-value < 0.05 (chi-square test).

18

19   **Figure 2.** *Spermatids are as good as sperm at DNA replication.*

20   **(A)** Sperm and spermatids are separately incubated with egg extracts supplemented with

21   biotin-dUTP. Subsequently, DNA fibers are isolated and subjected to molecular combing,

22   which reveals replication on single DNA fibers. **(B)** Examples of DNA fibers after

23   immunostaining procedure. Antibody staining against DNA reveals the total length of the

24   fibre (green) and antibody staining against biotin reveals the replicated DNA (red). The

25   bottom panels show representative examples of replication staining from sperm and from

1   spermatids incubated in egg extracts. **(C)** Replication extent measured as the proportion of

2   DNA that incorporated biotin-dUTP to the total fibre length. Results are from at least 125

3   independent DNA fibres (22000 kb of DNA for each sample). Error bars: sem. Samples were

4   not significantly different (p-value=0.41, KS-test).

5

6   **Figure 3.** *Transcription of developmentally important genes is misregulated in spermatid-*

7   *derived embryos compared to sperm-derived embryos*

8   **(A)** Schematic representation of paternally-derived haploid embryos generation by UV

9   enucleation of eggs followed by intra-cytoplamic sperm injection (ICSI). **(B)** Developmental

10   advantage of sperm over spermatid is maintained in haploid embryos. Embryos were scored

11   as the % of embryos reaching a gastrula stage and a swimming tadpole stage to the total

12   number of cleaved embryos (average of n=3 independent experiments). Numbers of embryos

13   analyzed are indicated above the bars. Error bars: sem. * indicates p-value < 0.05 (chi-square

14   test). **(C)** Genes important for development are misregulated (mostly upregulated) in

15   spermatid derived embryos. Heatmap representing log fold-change in expression levels of the

16   100 genes (rows) misregulated in spermatid versus sperm gastrula embryos (FDR < 0.05, red

17   – upregulated; blue – downregulated in spermatid) across seven independent experiments

18   (columns). **(D)** Developmentally-important gene ontology terms enriched in the list of

19   misregulated genes (p-value < 0.05). **(E)** Upregulation of genes in spermatid-derived

20   embryos does not correlate with their transcription in spermatid. Density scatter plot showing

21   gene expression in spermatid-derived embryos versus that in spermatids. No correlation is

22   observed between the two parameters for all genes (r=0.06) as well as for the misregulated

23   genes (red dots, r=-0.17).

24

26

1   **Figure 4.** *Genes that are misregulated in spermatid-derived embryos have different*

2   *epigenetic features in sperm and spermatid.*

3   **(A)** Genome-wide average nucleosome occupancy at the TSS of sperm (blue) and spermatid

4   (green) genes. **(B)** Boxplots showing genome-wide DNA methylation levels at the TSS±1kb

5   of sperm (blue) and spermatid (green) genes. Inset shows correlation between the DNA

6   methylation levels of sperm and spermatid (R = 0.8, p-value < 0.05), red line: regression,

7   dotted line: diagonal. **(C)** Percentage of genes harboring H3K27me3, H3K4me3, H3K4me2

8   or H3K9me3 peaks genome-wide (GW) and at misregulated genes (Mis). *: p-value < 0.05

9   (chi-square test). **(D)** Heatmaps representing H3K27me3, H3K4me3, H3K4me2 and

10   H3K9me3 overall levels (see Supplemental Information and Supplemental Fig. S8) at

11   misregulated genes in sperm (first column) and spermatid (second column). Each map is

12   sorted according to the signal in spermatid. Boxplots show the distribution of methylation

13   levels across misregulated genes. *: p-value < 0.05 (KS-test, Supplemental Table S7).

14

15   **Figure 5.** *H3K27me3 target genes that lose H3K4me2/3 in sperm compared to spermatids*

16   *are misregulated in spermatid-derived embryos.*

17   **(A, B)** Differential gene expression between sperm- and spermatid-derived embryos best

18   correlates with differential H3K4me2/3 and H3K27me3 marking in sperm and spermatids.

19   Partial correlation network between all tested epigenetic features of the paternal chromatin

20   (**A**- sperm, **B**- spermatid) and gene expression in the corresponding embryos. Edges (lines)

21   represent positive (red) or negative (blue) partial correlations. Edges thickness: strength of

22   the partial correlations. **(C)** H3K4me2/3 and H3K27me3 marking on misregulated genes is

23   conserved between *Xenopus* and human sperm. As compared to all orthologues, the

24   misregulated orthologues are enriched for H3K27me3 marks over the genome-wide average

27

1   in human sperm (Chi-square test, *: p-value < 0.05). No statistical enrichment for H3K4me2

2   on misregulated genes as compared to the genome-wide average is observed in human sperm.

3

4

5   **Figure 6.** *Paternal genome marking by H3K4me2/3 and H3K27me3 is required for gene*

6   *expression in the embryos.*

7   **(A)** Histone demethylase expression assay. **(B)** MA plot showing log fold-change (logFC, y-

8   axis) in gene expression between *Kdm5b* (H3K4me2/3 demethylase) - versus control-mRNA

9   injected embryos, against log counts per million (logCPM, x-axis). Red dots: genes

10 differentially expressed (FDR < 0.05) ; N = 4 independent experiments. **(C)** Venn-diagram of

11 downregulated genes upon KDM5B expression in sperm- (blue) and spermatid-derived

12 (green) embryos. **(D)** Percentages of genes downregulated upon KDM5B expression in

13 embryos that show H3K4me2/3 and H3K27me3 promoter peaks in the paternal cell. *: p-

14 value < 0.05 (chi-square test), ↑: over-represented when compared to genome-wide

15 distribution. (**E**) Proportion of misregulated genes affected in each demethylase expression

16 assay. *: p-value < 0.05 (chi-square test). **(F)** same as (**B**) for KDM6B (H3K27me3

17 demethylase) expression. **(G)** same as (**C**) with genes upregulated upon KDM6B expression.

18 **(H)** same as (**D**) for genes upregulated upon KDM6B expression. **(I)** Model of epigenetic

19 programming of sperm for the regulation of embryonic transcription.

20

References

Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, Beisel C, Schubeler D, Stadler MB, Peters AH. 2010. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nature structural & molecular biology* **17**: 679-687.

Bui HT, Wakayama S, Mizutani E, Park KK, Kim JH, Van Thuan N, Wakayama T. 2011. Essential role of paternal chromatin in the regulation of transcriptional activity during mouse preimplantation development. *Reproduction* **141**: 67-77.

Carone BR, Hung JH, Hainer SJ, Chou MT, Carone DM, Weng Z, Fazzio TG, Rando OJ. 2014. High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. *Developmental cell* **30**: 11-22.

El-Khoury V, Pierson S, Kaoma T, Bernardin F, Berchem G. 2016. Assessing cellular and circulating miRNA recovery: the impact of the RNA isolation method and the quantity of input material. *Sci Rep* **6**: 19529.

Erkek S, Hisano M, Liang CY, Gill M, Murr R, Dieker J, Schubeler D, Vlag J, Stadler MB, Peters AH. 2013. Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nat Struct Mol Biol* **20**: 868-875.

Gaggioli V, Le Viet B, Germe T, Hyrien O. 2013. DNA topoisomerase IIalpha controls replication origin cluster licensing and firing time in Xenopus egg extracts. *Nucleic acids research* **41**: 7313-7331.

Gaucher J, Reynoird N, Montellier E, Boussouar F, Rousseaux S, Khochbin S. 2010. From meiosis to postmeiotic events: the secrets of histone disappearance. *The FEBS journal* **277**: 599-604.

Gurdon JB. 1960. The Effects of Ultraviolet Irradiation on Uncleaved Eggs of Xenopus laevis. *Quarterly Journal of Microscopical Science* **101**: 299-311.

Gurdon JB. 1976. Injected nuclei in frog oocytes: fate, enlargement, and chromatin dispersal. *Journal of embryology and experimental morphology* **36**: 523-540.

Gurdon JB, Elsdale TR, Fischberg M. 1958. Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. *Nature* **182**: 64-65.

Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. 2009. Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460**: 473-478.

Hisano M, Erkek S, Dessus-Babus S, Ramos L, Stadler MB, Peters AH. 2013. Genome-wide chromatin analysis in mature mouse and human spermatozoa. *Nature protocols* **8**: 2449-2470.

Hutchison CJ, Brill D, Cox R, Gilbert J, Kill I, Ford CC. 1989. DNA replication and cell cycle control in Xenopus egg extracts. *Journal of cell science Supplement* **12**: 197-212.

Ihara M, Meyer-Ficca ML, Leu NA, Rao S, Li F, Gregory BD, Zalenskaya IA, Schultz RM, Meyer RG. 2014. Paternal poly (ADP-ribose) metabolism modulates retention of inheritable sperm histones and early embryonic gene expression. *PLoS genetics* **10**: e1004317.

Kimura Y, Yanagimachi R. 1995. Mouse oocytes injected with testicular spermatozoa or round spermatids can develop into normal offspring. *Development* **121**: 2397-2405.

Kishigami S, Wakayama S, Nguyen VT, Wakayama T. 2004. Similar time restriction for intracytoplasmic sperm injection and round spermatid injection into activated oocytes for efficient offspring production. *Biology of reproduction* **70**: 1863-1869.

Labit H, Goldar A, Guilbaud G, Douarche C, Hyrien O, Marheineke K. 2008. A simple and optimized method of producing silanized surfaces for FISH and replication mapping on combed DNA fibers. *BioTechniques* **45**: 649-652, 654, 656-648.

Lee HX, Ambrosio AL, Reversade B, De Robertis EM. 2006. Embryonic dorsal-ventral signaling: secreted frizzled-related proteins as inhibitors of tolloid proteinases. *Cell* **124**: 147-159.

Lemaitre JM, Danis E, Pasero P, Vassetzky Y, Mechali M. 2005. Mitotic remodeling of the replicon and chromosome structure. *Cell* **123**: 787-801.

Lemaitre JM, Vassetzky Y, Mechali M. 2001. Analysis of Chromatin Assembly, Chromatin Domains, and DNA Replication Using Xenopus Systems. In *Mapping Protein/DNA Interactions by Cross-Linking*, Paris.

Narbonne P, Simpson DE, Gurdon JB. 2011. Deficient induction response in a Xenopus nucleocytoplasmic hybrid. *PLoS biology* **9**: e1001197.

Nieuwkoop PD, and Faber, J. . 1994. Normal table of Xenopus laevis (Daudin) : a systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis In *New York, Garland Pub*.

29

Paradowska AS, Miller D, Spiess AN, Vieweg M, Cerna M, Dvorakova-Hortova K, Bartkuhn M, Schuppe HC, Weidner W, Steger K. 2012. Genome-wide identification of promoter binding sites for H4K12ac in human sperm and its relevance for early embryonic development. *Epigenetics : official journal of the DNA Methylation Society* **7**: 1057-1070.

Risley MS, Eckhardt RA. 1981. H1 histone variants in Xenopus laevis. *Developmental biology* **84**: 79-87.

Samans B, Yang Y, Krebs S, Sarode GV, Blum H, Reichenbach M, Wolf E, Steger K, Dansranjavin T, Schagdarsurengin U. 2014. Uniformity of nucleosome preservation pattern in Mammalian sperm and its connection to repetitive DNA elements. *Developmental cell* **30**: 23-35.

Siklenka K, Erkek S, Godmann M, Lambrot R, McGraw S, Lafleur C, Cohen T, Xia J, Suderman M, Hallett M et al. 2015. Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. *Science* doi:10.1126/science.aab2006.

Smith SJ, Fairclough L, Latinkic BV, Sparrow DB, Mohun TJ. 2006. Xenopus laevis transgenesis by sperm nuclear injection. *Nature protocols* **1**: 2195-2203.

Suri C, Haremaki T, Weinstein DC. 2004. Inhibition of mesodermal fate by Xenopus HNF3beta/FoxA2. *Developmental biology* **265**: 90-104.

Suzuki T, Minami N, Kono T, Imai H. 2007. Comparison of the RNA polymerase I-, II- and III-dependent transcript levels between nuclear transfer and in vitro fertilized embryos at the blastocyst stage. *The Journal of reproduction and development* **53**: 663-671.

Vassena R, Han Z, Gao S, Baldwin DA, Schultz RM, Latham KE. 2007. Tough beginnings: alterations in the transcriptome of cloned embryos during the first two cell cycles. *Developmental biology* **304**: 75-89.

Weidgang CE, Russell R, Tata PR, Kuhl SJ, Illing A, Muller M, Lin Q, Brunner C, Boeckers TM, Bauer K et al. 2013. TBX3 Directs Cell-Fate Decision toward Mesendoderm. *Stem Cell Reports* **1**: 248-265.

Wu SF, Zhang H, Cairns BR. 2011. Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. *Genome research* **21**: 578-589.

Yasuoka Y, Suzuki Y, Takahashi S, Someya H, Sudou N, Haramoto Y, Cho KW, Asashima M, Sugano S, Taira M. 2014. Occupancy of tissue-specific cis-regulatory modules by Otx2 and TLE/Groucho for embryonic head specification. *Nat Commun* **5**: 4322.

Zheng Z, Jia JL, Bou G, Hu LL, Wang ZD, Shen XH, Shan ZY, Shen JL, Liu ZH, Lei L. 2012. rRNA genes are not fully activated in mouse somatic cell nuclear transfer embryos. *The Journal of biological chemistry* **287**: 19949-19960.

Ziyyat A, Lefevre A. 2001. Differential gene expression in pre-implantation embryos from mouse oocytes injected with round spermatids or spermatozoa. *Human reproduction* **16**: 1449-1456.
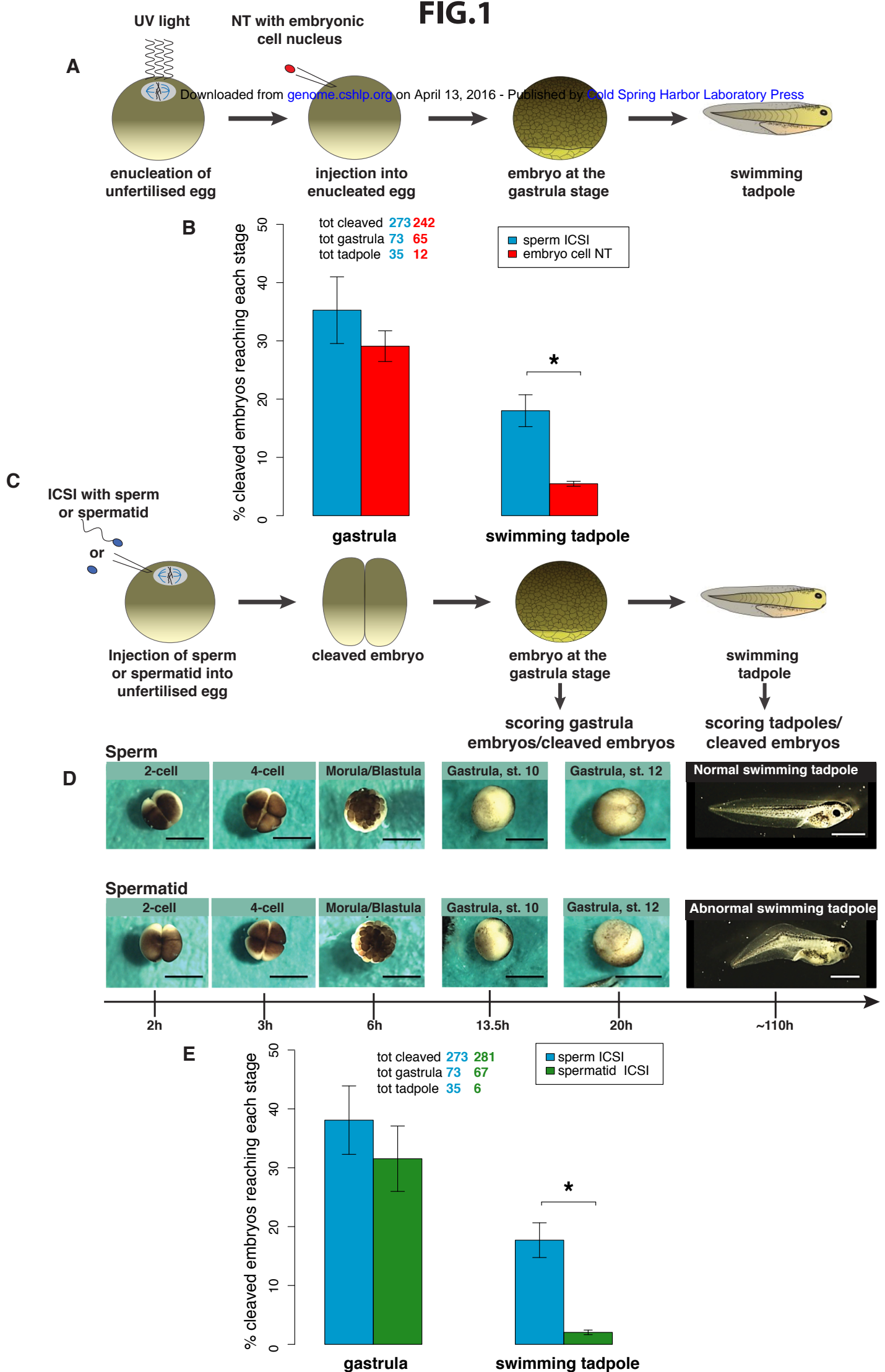
FIG.1

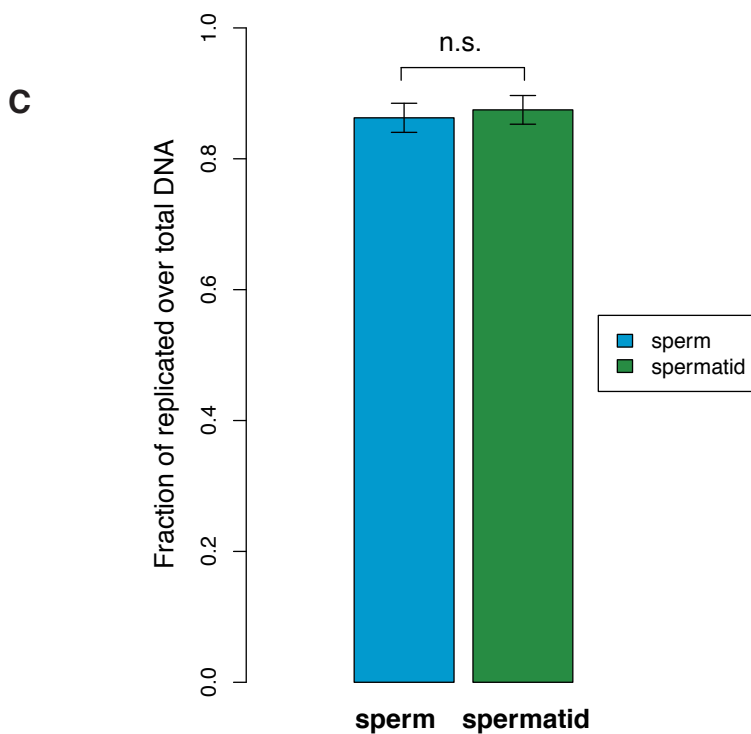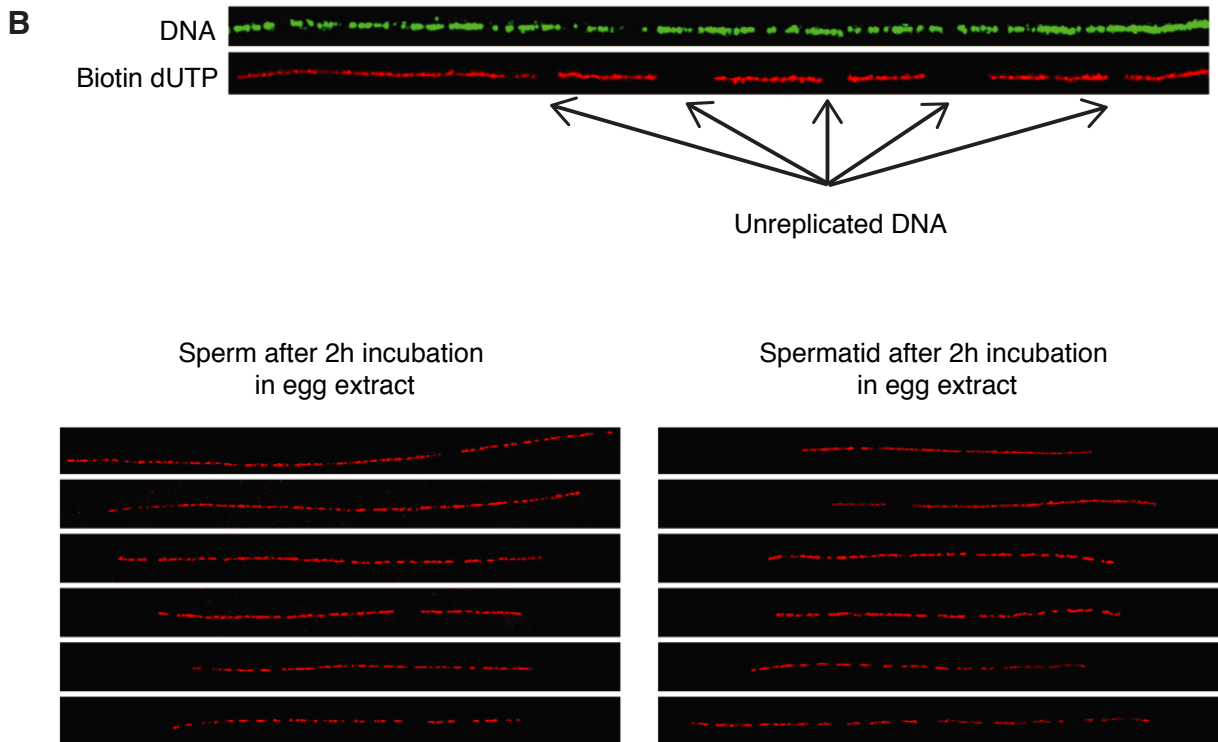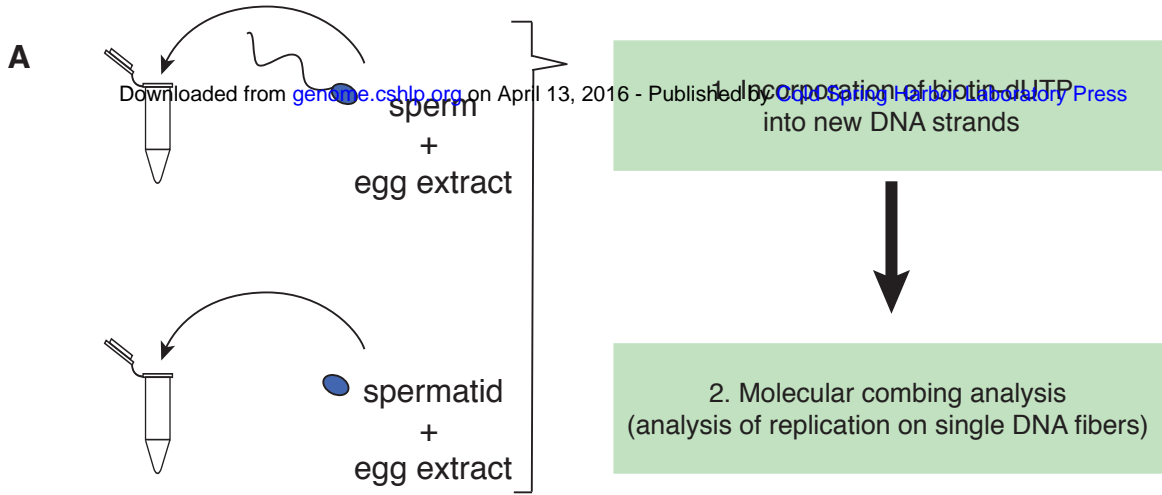**A** enucleation of unfmcleused egg

FIG.2

FIG.3

A enucleation of unfertilised egg → injection into enucleated egg → embryo at the gastrula stage → haploid swimming tadpole

B

tot cleaved	76	75
tot gastrula	31	27
tot tadpole	21	8

sperm ICSI
spermatid ICSI

C Experiments 1 2 3 4 5 6 7 avg logFC

upregulated in 7/7 spermatid-derived embryos — gjb1, plod2, gata3, hes1, dhh

upregulated 6/7 spermatid-derived embryos — fos, hes1, bmp7, mn1

downregulated in 6/7 spermatid-derived emryos

downregulated in 7/7 spermatid-derived embryos — c19orf26

Low — High
logFC (spermatid/sperm)

D GO terms enriched in 'Misregulated Genes'

Multicellular organismal development

Tissue development

System development

Developmental process

Anatomical structure development

Mesoderm development

E R = -0.17, p-value < 0.01

FIG.4

A sperm

B spermatid

C

Number of X. laevis orthologues of human genes with H3K27me3 and H3K4me2 marks in sperm

	Number of genes	Number (%) of H3K27me3-positive genes	Number (%) of H3K4me2-positive genes
All orthologs (*X. laevis* to human)	8812	1140 (13%)	2469 (28%)
Misregulated genes orthologs	54	18 (33%) *	12 (22%)

FIG.5
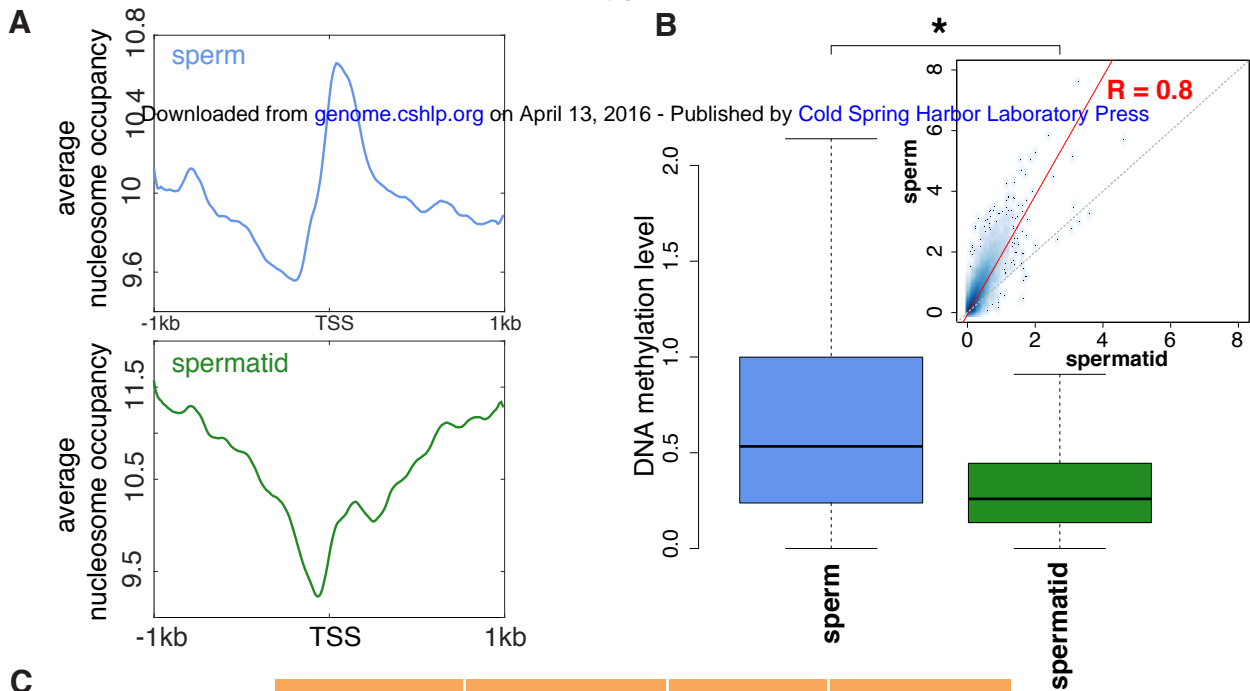
FIG.6

A

mRNA Demethylase (Kdm5b,Kdm6b)/control

in vitro maturation *(IVM)*

Injection of mRNA into GV state oocyte

Injection of sperm or spermatid into *in vitro* matured egg (ICSI)

Embryo at the gastrula stage

RNA-seq

Transcriptome profiling

B Kdm5b IVM/ICSI

sperm: 902 / 1893
spermatid: 344 / 1392

F Kdm6b IVM/ICSI

sperm: 487 / 71
spermatid: 173 / 54

C
S 1893 DE DOWN / P 1392 DE DOWN
1539 | 354 | 1038

G
S 487 DE UP / P 173 DE UP
402 | 85 | 88
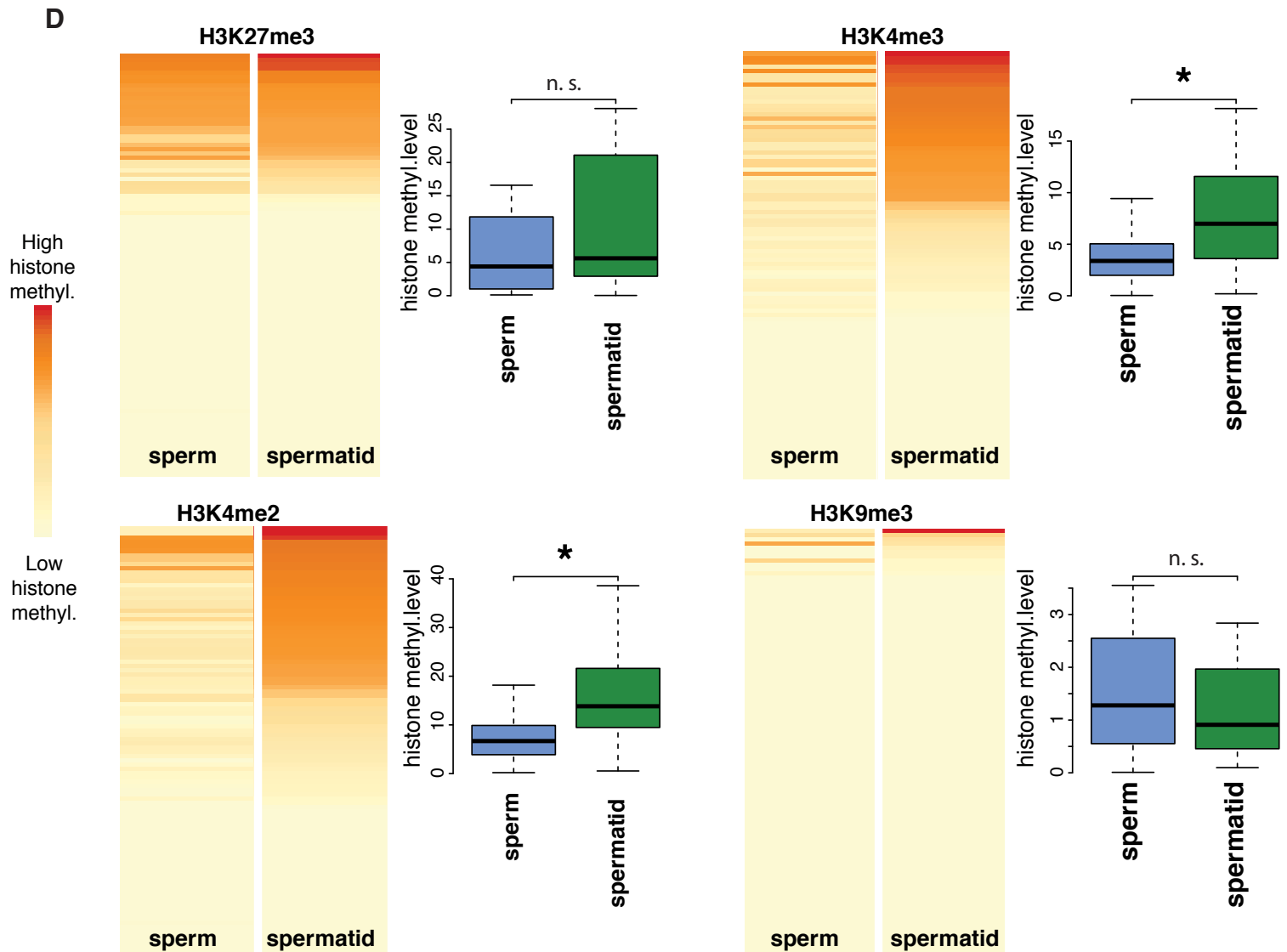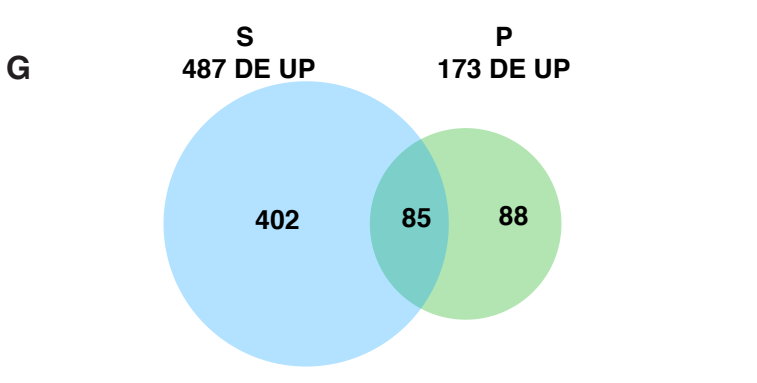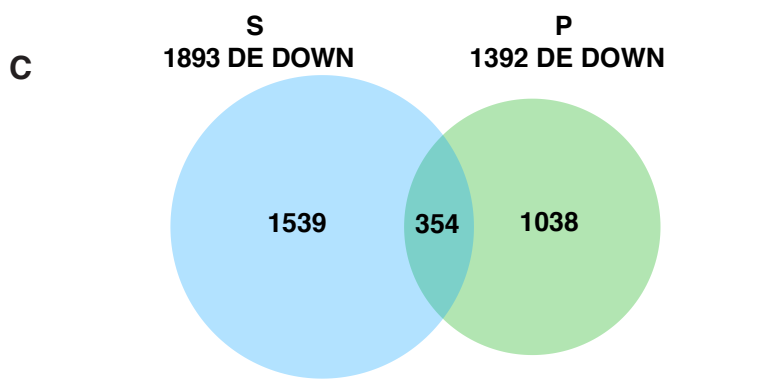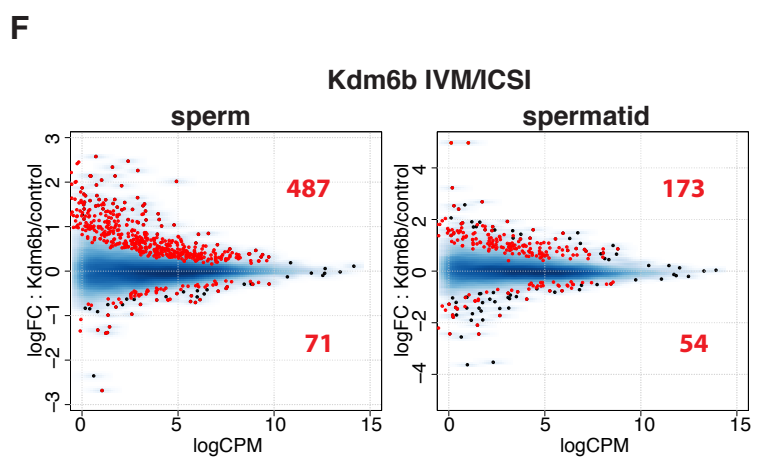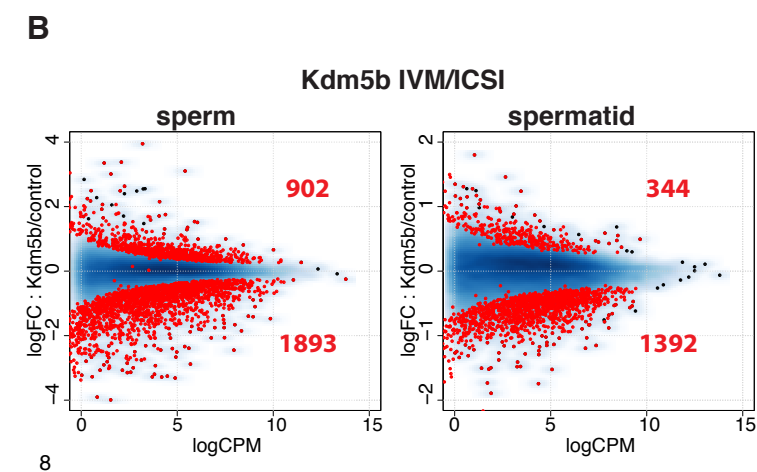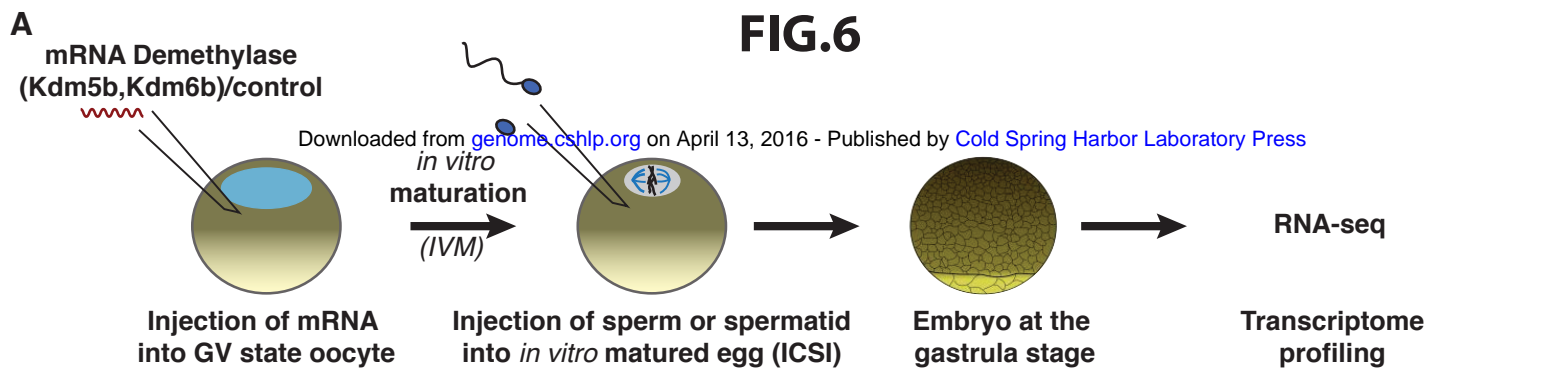
D

Promoter mark	sperm (DOWN)	spermatid (DOWN)
H3K4me3	3.4%	71% * ↑
H3K4me2	58.1%	82.2% * ↑
H3K27me3	21.3% * ↑	23%

H

Promoter mark	sperm (UP)	spermatid (UP)
H3K4me3	1.2%	50.3%
H3K4me2	50.7%	73.4%
H3K27me3	46.8% * ↑	46.82% * ↑

E

Kdm5b sensitive in sperm embryo
Kdm5b sensitive in spermatid embryo
Kdm6b sensitive in sperm embryo
Kdm6b sensitive in spermatid embryo

I

Spermatid to Sperm transition

e.g. gata3, hes1, plod2
Spermatid ICSI

gata3 hes1 plod2 ↑ upregulation
Spermatid-derived embryo

H3K27me3
H3K4me2
H3K4me3

e.g. gata3, hes1, plod2
Sperm ICSI

gata3 hes1 plod2
Sperm-derived embryo
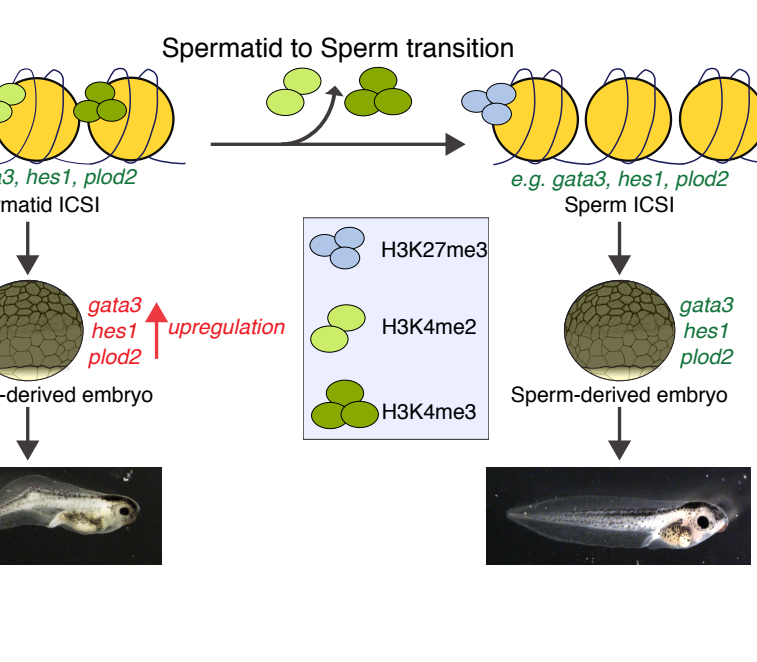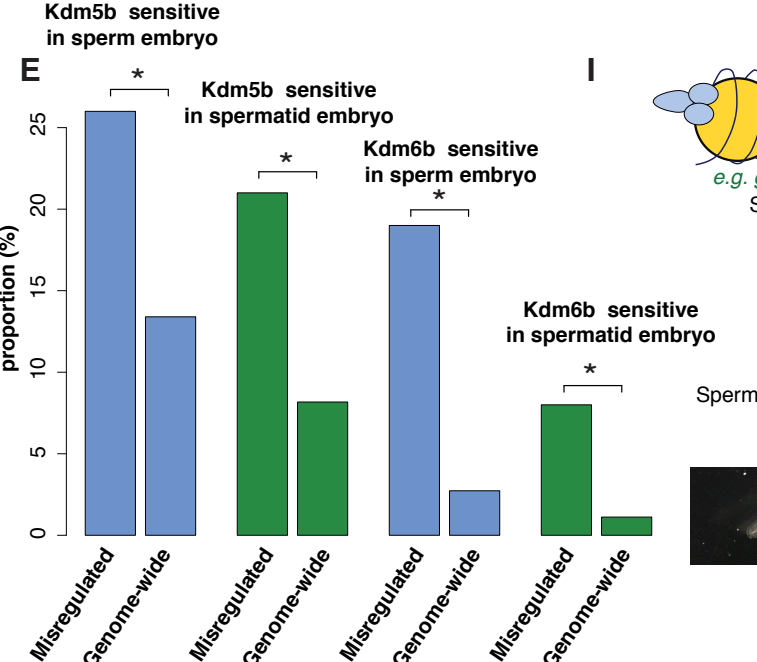
Sperm is epigenetically programmed to regulate gene transcription in embryos

Marta Teperek, Angela Simeone, Vincent Gaggioli, et al.

P<P	Published online March 31, 2016 in advance of the print journal.
Accepted Manuscript	Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.
Open Access	Freely available online through the *Genome Research* Open Access option.
Creative Commons License	This manuscript is Open Access.This article, published in *Genome Research*, is available under a Creative Commons License (Attribution 4.0 International license), as described at http://creativecommons.org/licenses/by/4.0/.
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Genome Research* go to:
http://genome.cshlp.org/subscriptions