Mapping of histone-binding sites in histone replacement-completed spermatozoa

Keisuke Yoshida, Masafumi Muratani, Hiromitsu Araki, Fumihito Miura, Takehiro Suzuki, Naoshi Dohmae, Yuki Katou, Katsuhiko Shirahige, Takashi Ito & Shunsuke Ishii

The majority of histones are replaced by protamines during spermatogenesis, but small amounts are retained in mammalian spermatozoa. Since nucleosomes in spermatozoa influence epigenetic inheritance, it is important to know how histones are distributed in the sperm genome. Conflicting data, which may result from different conditions used for micrococcal nuclease (MNase) digestion, have been reported: retention of nucleosomes at either gene promoter regions or within distal gene-poor regions. Here, we find that the swim-up sperm used in many studies contain about 10% population of sperm which have not yet completed the histone-to-protamine replacement. We develop a method to purify histone replacement-completed sperm (HRCS) and to completely solubilize histones from cross-linked HRCS without MNase digestion. Our results indicate that histones are retained at specific promoter regions in HRCS. This method allows the study of epigenetic status in mature sperm.
ncreasing evidences indicate that various paternal environmental factors affect the gene expression pattern and traits in an offspring. Paternal nutritional conditions such as a high-fat or low-protein diet induce gene expression changes in the pancreatic islets of rat or the liver of mouse offspring, respectively. Furthermore, in utero undernutrition perturbs the metabolism in offspring via sperm. Although the mechanism of paternal inheritance of environmentally induced changes remains elusive, some evidences suggested that histone modification changes play an important role. We previously demonstrated that Drosophila transcription factor ATF2 mediates the heat shock-induced or osmotic stress-induced reduction of histone H3 lysine 9 tri-methylation (H3K9me3), which is transmitted to the next generation. The other group also reported that overexpression of the histone H3 lysine 4 (H3K4) demethylase KDM1A (also known as LSD1) during spermatogenesis impairs the development and survivability of the offspring. However, the molecular mechanisms underlying paternal epigenetic transmission are still unclear.

To understand these mechanisms, it is important to know the distribution of histones in mature sperm. During mammalian spermatogenesis, most histones are replaced by protamines in spermatids. Histone-to-protamine replacement is essential for inducing sperm nuclear condensation, which is one of the most important parameters for estimating human sperm quality. Indeed, round spermatid nucleus injection into oocytes has a lower fertilization rate than intracytoplasmic sperm injection. In the clinical field of assisted reproduction techniques, mature sperm have traditionally been prepared from human semen mainly using two methods: Percoll gradient centrifugation or the swim-up procedure. The Percoll procedure to isolate normal sperm were the first innovative method to improve pregnancy rates in human in vitro fertilization. The swim-up method is now popularly used for isolating high-motility sperm and removing the abnormal sperm. Some studies suggest that the Percoll method is more effective than the swim-up method in producing samples with higher nuclear sperm condensation and lower populations of morphologically abnormal sperm. These results suggest the importance of sperm nuclear structure for fertility.

The sperm chromatin structure assay (SCSA) was developed for diagnostic assessment of human sperm quality by analyzing the nuclear structure of the sperm. This assay uses acid treatment to denature DNA at the sites of DNA strand breaks, followed by staining with acridine orange (AO). AO intercalates mainly into double-stranded DNA containing histones, but not protamines, which then generates green fluorescence. On the other hand, AO stacks on single-stranded DNA, which is formed by apoptosis-induced DNA double-strand breaks, generating red fluorescence. Thus, the fraction of sperm with high DNA stainability (HDS), as indicated by green fluorescence, has an abnormally high level of DNA staining due to lack of full proimation and consists of sperm population with incomplete histone replacement. Actually, the group with a larger proportion of HDS sperm in humans has a lower chance of pregnancy. The DNA fragmentation index (DFI), as measured by red fluorescence, indicates DNA fragmentation caused by apoptosis. The content of DFI fraction in mouse sperm samples is usually lower than in human samples, possibly because mouse sperm can be analyzed immediately after preparation and has less histones compared with human sperm.

A small fraction of histones remains associated with the sperm genome. By deep sequencing of chromatin DNA immuno-precipitated with anti-histone antibodies from mononucleosomes solubilized by micrococcal nuclease (MNase) without cross-linking, Hammoud et al. reported that the retained nucleosomes in human sperm are enriched at development-related genes, including imprinted gene clusters and the promoters of developmental transcription and signaling factors. Using a similar method, Erkek et al. also showed that mouse spermatozoa contain only 1% of residual histones, which are ten- and twofold more enriched at GC-rich promoters and exons, respectively, relative to genomic background.

However, two groups recently reported data inconsistent with those of the previous reports. By high-throughput sequencing of the mononucleosomal DNA released by MNase, Samans et al. showed that most nucleosomes from human and bovine sperm are enriched in distal intergenic regions and localized in repetitive DNA sequences. Using a recently developed nucleosome mapping method, Carone et al. also reported that the nucleosomes in mouse sperm are preferentially retained in gene-poor regions and are generally depleted from promoters, including developmental promoters such as the Hox promoters. Carone et al. speculated that such discrepancies might result from differences in the MNase digestion conditions of the isolated sperm, because they recovered mononucleosomes enriched over promoters of developmental regulators when a relatively high concentration of MNase was used. However, Samans et al. used relatively high concentrations of MNase similar to those used by Hammoud et al., but obtained contradictory results. Furthermore, the nucleosomal enrichments at repetitive sequences reported by Samans et al. resulted from the redundant use of sequencing reads that map to multiple locations in the genome, an inappropriate computational methodology. Thus, two inconsistent results have been reported, and a resolution of the conflicting data is needed for understanding the epigenetic status in sperm genome.

Here, we carefully examine the quality of sperm and find that the swim-up sperm fraction used by most research groups still contains the histone replacement-uncompleted sperm (HRunCS). By developing a method to purify the histone replacement-completed sperm (HCRCs) and to completely solubilize histones after cross-linking without MNase digestion, we demonstrate that histones are distributed on the promoters of specific genes in HCRCs. Results obtained by our method suggest that histones in mouse spermatozoa may contribute to expression regulation of the target genes during early embryonic development.

**Results**

**Isolation of the HCRCs.** To examine the quality of sperm, we first analyzed the sperm using SCSA. Spermatozoa develop from spermatogonia stem cells in the testis and then move to the epididymis, and into the vas deferens (Fig. 1a). To check the quality of the mouse sperm, we performed SCAs for total sperm fraction collected from different parts of the epididymis or vas deferens (Supplementary Fig. 2a, b). We found the HDS fractions of 26.8% and 20.3% in the total sperm from caput and corpus epididymis, respectively (Fig. 1a). HDS fractions of 11.6% and 8.7% were observed in the total sperm from the cauda epididymis and the vas deferens, respectively. These results indicated that the fraction of HDS, which has not completed the histone-to-protamine replacement, decreases during maturation in the epididymis, whereas the fraction of normal sperm, which has completed histone replacement, increases. Consistent with this, western blotting results showed that the amount of histone H3 decreased during movement from caput to cauda epididymis (Fig. 1b). These results suggest that the histone replacement by protamine is not completely finished when the sperm enters the epididymis, and is gradually completed during the movement of sperm from the caput to the cauda of the epididymis. Alternatively, the histone replacement-incomplete sperm might be removed by apoptosis during epididymal maturation. In addition,
**Fig. 1** Estimation of histone amount in each sperm fraction. 

**a** (Left) Picture of mouse epididymis consisting of caput, corpus, caudal epididymis, and vas deferens. (Middle) SCSA results for total sperm fractions collected from each part of the epididymis and vas deferens. Only cell population of singlet is shown here. Sperm cells are mainly separated into two fractions: normal sperm (normal) and the histone replacement-incomplete sperm with high DNA stainability (HDS) with AO. (Right) The mean ± s.d. of percentage of HDS population (n = 3).

**b** Immunoblot analysis of histone H3 in total lysates of sperm fraction collected from each part of the epididymis. Cell numbers applied to each lane are indicated by thousands. DNA was analyzed for verifying the cell number of each sample.

**c** (Left) SCSA results for total sperm and swim-up sperm fraction. (Right) The mean ± s.d. of percentage of HDS population (n = 6).

**d** Schematic procedure for preparation of the histone replacement-completed sperm (HRCS). Sperm cells with red or blue heads indicate HRunCS or mature sperm, respectively. High-density sperm heads generated by mild sonication were isolated by Percoll. Note that some portion of HRunCS could be sensitive to mild sonication.

**e** SCSA results for HRCS.

**f** Immunoblot analysis of H3 in total sperm, swim-up sperm, and HRCS preparation.
the sperm from cauda epididymis and the vas deferens still contain significant amounts of HRRunCS.

To examine the histone distribution in sperm chromatin, a relatively large amount of sperm is required, and the amount from the vas deferens is not sufficient. Consequently, in most of the experiments reported so far, sperm from the cauda epididymis, which contains ~tenfold more sperm than the vas deferens, were used. Furthermore, sperm with high motility selected by the swim-up method (swim-up sperm) were used in most of the studies. Therefore, we next analyzed the swim-up sperm using SCsASs. Unexpectedly, 6.1% of the swim-up sperm preparation showed HDS (Fig. 1c). To test whether the population of HDS varies depending on the procedure of the swim-up method used, the top 0.8 ml medium of 4.0-ml caudal epididymis cultivation (2.5 ml/4.0 ml in the standard procedure25) was also recovered and analyzed. About 8.7% of the prepared sperm showed HDS (Supplementary Fig. 2e), indicating that the population of HRRunCS does not vary depending on the procedure used.

To avoid contamination by histones derived from HRRunCS for mapping of H3-binding sites, we developed a method to isolate the HRCS from total sperm collected from caudal epididymis and vas deferens (in the following description, total sperm means total sperm fraction prepared from caudal epididymis and vas deferens), based on Percoll centrifugation for purification of high-density sperm heads. The total sperm were mildly sonicated to generate sperm heads by removing the tails and centrifuged in 82% Percoll solution. The high-density sperm were spun down, while the relatively low-density sperm and sperm tails remained in the supernatant (Fig. 1d and Supplementary Fig. 2d). SCsASs of the pelleted sperm indicated that this preparation contained almost no HDS fraction (Fig. 1e), and almost 100% purity of the HRCS fraction was obtained. The H3 antibody used in this study recognized H3.1 and H3.3 with the same degree of sensitivity (Supplementary Fig. 2e). Results of western blotting using this antibody indicated that the amount of histone H3 in the HRCS was about 1/5 that of the total sperm and that of the swim-up sperm (Fig. 1f). Target MS–MS quantification analysis also showed that the amount of histone H3 variants H3.1, H3.2, H3.3, or H3t in the HRCS was about 33, 14, 27, or 29% that of the total sperm, respectively (Supplementary Fig. 3). These results suggest that HRRunCS contains much more amount of histone proteins than HRCS. In total sperm, 10% of them was HRRunCS, and results of western blotting indicated that the ratio of H3 in total sperm (90% HRCS + 10% HRRunCS) and HRCS (100% HRCS) is 5:1. These results indicate that H3 protein contained in total and swim-up sperm samples is mainly derived from HRRunCS, and that the distribution of histone-binding sites in the swim-up sperm reflects the localization of histones in HRRunCS. Comparison of the histone H3 level by western blotting analysis showed that the amount of H3 in HRCS was about 0.3% of that in somatic cells of mouse embryonic fibroblasts (MEFs) (Supplementary Fig. 1f), which is lower than that previously reported in swim-up sperm23,24.

Complete solubilization of histones from cross-linked HRCS. Digestion of sperm chromatin using a high concentration of MNase might fail to detect specific regions of chromatin that are hypersensitive to MNase digestion. We observed that the tiny amount of nucleosomes contained in HRCS could not be completely solubilized even using a high concentration of MNase or sonication after cross-linking. Therefore, to avoid the effect of MNase digestion conditions, we developed a method to completely solubilize the nucleosomes of sperm without MNase digestion after cross-linking. Compared with the cross-linked somatic cells, cross-linked sperm are very hard and it is difficult to disrupt and solubilize nucleosomes by sonication. For these reasons, we developed a buffer to induce decondensation of the cross-linked sperm nuclei. Protamine is a strong basic protein and tightly binds to sperm DNA via ionic interactions between the positive charge on arginine residues of protamine and the negative charge on phosphate groups of DNA23. Furthermore, disulfide linkage between protamine molecules is required for compaction of sperm DNA26. We therefore treated sperm cells with dithiothreitol (DTT) for reducing the disulfide bonds and with heparin for neutralizing the positive charge of protamines (Fig. 2a). Heparin induces decondensation of sperm chromatin27. After this treatment, sperm heads were enlarged (Fig. 2b), indicating decondensation of sperm nuclei. When these cells were treated by sonication, all the sperm DNA and nucleosome histone H3/H4 were recovered in the supernatant after centrifugation (Fig. 2c), indicating that most of the nucleosomes were solubilized, and DNA was sheared to the appropriate size for ChiP experiments using this method.

Binding profiles of histone H3 in total sperm and HRCS. Using these conditions, we performed histone H3 ChiP-seq analysis in total sperm and HRCS. H3 signals in total sperm and HRCS were mainly found on gene promoter regions such as those of the Hoxa gene cluster, and the signals were considerably less intense or absent in HRCS (Fig. 3a and Supplementary Fig. 4). Genome-wide analysis indicated that 10,988 peaks in total sperm and 1320 peaks in HRCS were detected, and this difference in peak number may be consistent with the amount of H3 in total sperm and HRCS. Approximately 69% and 64% of peaks were localized to promoter regions in total sperm and HRCS, respectively (Fig. 3b), and H3 peaks were enriched in promoter regions and exon regions relative to background (Supplementary Fig. 5a). A minor population of peaks was found in intergenic repeat regions such as LINE and GC-rich repeats (Supplementary Fig. 5b). In the peak regions detected in total sperm, peak strength in HRCS was approximately half that in total sperm (Fig. 3c). About 73% of peaks (966/1320) in HRCS overlapped with those in total sperm, while most of the H3 target genes in HRCS (905/911) were also target genes in total sperm (Fig. 3d). About 64% of the H3 peaks (7067/10,988) detected in total sperm overlapped with the nucleosome regions detected by MNase-seq using swim-up sperm in the previous report16 (Supplementary Fig. 5c), suggesting similar histone-binding profiles in swim-up and total sperm. Consistent with this, both populations contained similar HDS fractions (Fig. 1c). About 22% (285/1320) of histone H3 peaks and 75% (686/911) of H3-binding genes in HRCS overlapped with those in the swim-up sperm, but the number of H3 peaks and H3-binding genes was much lower than in swim-up sperm16.

Around 76% (8351/10,988) and 67% (879/1320) of H3 peaks in total sperm and HRCS, respectively, were observed in CpG islands (CGIs) (Fig. 3e). The median observed/expected (o/e) ratios in HRCS peaks were higher than those in total sperm peaks and all CGIs, indicating that histone H3 is preferentially retained at CGIs with higher CG dinucleotide enrichment in mature sperm. Interestingly, H3 peaks were not uniformly distributed on each chromosome and were observed at lower frequencies only on sex chromosomes (Supplementary Fig. 5d). This tendency might be explained by the lower frequencies of CGIs with high o/e ratios (>1.0) on sex chromosomes relative to autosomal chromosomes. The DNA methylation profile of HRCS was not much different to that of the swim-up sperm (Supplementary Fig. 6a). Furthermore, H3-binding sites in total sperm and HRCS were predominantly in lower methylation regions, consistent with the previous report16 (Supplementary Fig. 6b). The DNA

[Referring to the original text for specific details and analysis related to the experimental procedures and results mentioned in the text.]

ARTICLE
methylation level was not significantly different between these H3-binding sites (Supplementary Fig. 6c).

We further analyzed the histone H3-binding genes, which we divided into two groups, HRCS target genes, where H3 peaks were detected in HRCS, and total sperm-specific target genes, where H3 peaks were detected only in total sperm samples, but not in HRCS. Motif analysis showed that repeated CGG sequence was extracted from binding sequences in HRCS target genes, whereas CC was extracted from those in total sperm-specific target genes (Supplementary Fig. 7a), supporting the observation that the o/e ratio of CpG in the H3-binding sites in HRCS was higher than that in total sperm (Fig. 3e). Using published data\textsuperscript{16}, we investigated the profiles of histone modifications around the transcription start sites (TSSs) of these H3-binding genes in RS and swim-up sperm. There was no specific difference in H3K4me3 and H3K27me3 in RS and swim-up sperm between HRCS target and total sperm-specific target genes (Supplementary Fig. 7b). Thus, these two target genes are not defined at least by H3K4me3 or H3K27me3 marks.

**Epigenetic feature of sperm H3-binding genes in the embryo.** The heatmap data of H3 signal showed that the HRCS target genes had a large amount of H3 also in the total sperm (Fig. 4a). A group of genes in total sperm-specific target genes also had a moderate level of H3 for the HRCS data. Therefore, we further divided the H3-binding genes into three categories: category_H corresponded to HRCS target genes, category_PH (potential HRCS target genes) to total sperm-specific target genes with moderate levels of H3 in the HRCS data, and category_TS (total sperm-specific target genes) to total sperm-specific target genes (Fig. 4a). Pathway analysis indicated that category_H genes included a group of genes related to developmental biology, which include the genes encoding Hox transcription factors and also related to neuronal cell differentiation (axon guidance and signaling by NGF), while category_TS involved genes related to gene expression and metabolism of proteins (Fig. 4b and Supplementary Data 1). Only development-related genes were shared between category_H genes and nucleosome target genes previously identified by MNase-seq\textsuperscript{16} (Supplementary Fig. 8a).

Using published data\textsuperscript{28,29}, we also investigated the chromatin state and epigenetic profiles of these target genes during early embryogenesis because they are important for understanding molecular mechanisms of epigenetic inheritance via sperm. At the four-cell stage, category_H and category_PH genes showed
**Fig. 3** Analysis for H3-binding regions in total sperm and HRCS. 

**a** H3 ChIP-seq analysis for total sperm and HRCS in *Hoxa* gene cluster regions. Blue signal indicates the number of raw reads, which is scaled to mapped regions in each sample. Light red boxes indicate the binding sites detected in both total sperm and HRCS. **b** Distribution of H3 peaks detected in total sperm and HRCS (10,988 and 1320 peaks, respectively) among various genomic regions. **c** Profile plot for signal intensity of H3 ChIP-seq around binding sites detected in total sperm and HRCS. **d** Overlap of H3-binding sites (top) or H3 target genes (bottom) between total sperm and HRCS. **e** Box plot for o/e ratio in all CpG islands or the CpG islands overlapping with H3-binding sites in total sperm or HRCS. The number of CpG islands overlapping with H3-binding sites in total sperm or HRCS is shown at the bottom. The elements in box plot indicate following value; center line, median; box limits, upper and lower quartiles 1.5x interquartile range; and points, outliers. P value was calculated by Wilcoxon rank-sum test.
lower ATAC-seq and H3K4me3 ChIP-seq signals than category_TS genes, suggesting that the chromatin structure of category_H/PH genes exhibits lower DNA accessibility and slightly weaker transcriptional activity than that of category_TS genes at this stage (Fig. 4a) (Supplementary Fig. 9). The profile plot of ATAC-seq signal also showed a sharp drop around the TSS of category_H/PH from the two-cell to eight-cell stage, but a clear peak around the TSS of category_TS and all genes from the two-cell to four-cell stage (Fig. 4c).
Fig. 4 Contribution of H3 binding in sperm to gene silencing during early embryogenesis. a Heatmap of signal intensity of H3 ChIP-seq in total sperm and HRCS, ATAC-seq at the four-cell stage, and H3K4me3 ChIP-seq at the four-cell stage, across H3 target genes in total sperm and HRCS. H3 target genes are categorized into three groups: category_H (target genes in HRCS), category_PH (Potential target genes in HRCS), and category_TS (specific target genes in total sperm), based on the signal strength of the H3 peak in HRCS. b Pathway analysis for category_H and category_TS genes using Reactome datasets. The five top pathway names are presented. c Profile plot for signal intensity of ATAC-seq around the TSS of all genes and category_H, category_PH, and category_TS genes. d Percent population of highly expressed genes (FPKM > 10), which are contained in all genes or category_H, category_PH, or category_TS target genes, from oocyte to ICM. The target genes in sperm and all genes were separated by expression level in the oocyte (FPKM > 10)

Finally, using published RNA-seq data, we analyzed the dynamics of expression profiles for the H3-binding genes in HRCS and total sperm during early embryogenesis. Interestingly, the proportion of highly expressed genes in category_H/PH was reduced drastically after the four-cell stage, but that in category_TS and total genes was moderate (Fig. 4d and Supplementary Fig. 10a, b). Consistent with this, H3K4me3 signal in the H3 target genes of HRCS was slightly lower than those of total sperm in early embryos (two-cell to ICM) (Supplementary Fig. 9). These data suggest that the maternal mRNAs of category_H/PH genes are degraded at the two-cell stage, and most of those mRNAs are not induced during zygotic gene activation (ZGA) in the four-cell embryo, while the genes of category_TS are transcriptionally induced during ZGA. To test whether this might be correlated with the presence of epigenetic silencing marks in HRCS, we have examined that these genes have the histone H3K9me2/3 marks. Actually, H3K9me2, not H3K9me3, was found in the promoter regions of category_H genes, whose expressions were suppressed at the stages between four-cell and ICM (Supplementary Fig. 11a, b). Thus, H3 binding and/or modified histone in sperm may silence the target genes during early embryogenesis.

**Discussion**

Here, we developed a method to analyze histone distribution in mouse sperm by using two modifications: the purification of histone-to-protamine replacement-completed sperm, HRCS, and the complete solubilization of nucleosomes after cross-linking, using DTT and heparin treatment and sonication without MNase digestion. The results obtained using this method indicate that histones are localized at the promoter regions of specific genes. The data using the swim-up sperm previously reported by Erkek et al. also indicated histone distribution on gene promoter regions, although fewer H3-binding genes were observed in HRCS than in the swim-up sperm. The data obtained using swim-up sperm are consistent with our results using total sperm from the cauda epididymis, which contain a similar population of HRunCs to the swim-up sperm. Our results did not show the enrichment of histone H3 in gene-poor intergenic regions and are therefore inconsistent with the results reported by the two groups. The reason for this discrepancy is unclear, but in the present study, histones were cross-linked to DNA before extraction, and the solubilization of histones by sonication was almost complete. In addition, MNase was not used in our method, so the results were not affected by the MNase digestion condition.

Recently, Barral et al. analyzed the process of the histone-to-protamine replacement and showed the nature of transitional structures. In condensing spermatids, the H2A.L2–TH2B dimer is first loaded onto the nucleosomes, which opens the nucleosomes and allows the invasion of nucleosomes by transition proteins. Nucleosome–transition proteins interact with the incoming protamines, which mediate pre-protamine-2 processing, and the protamines then bind to DNA in competition with histone–transition protein complexes. Since the displaced histones are unable to remain as octamers, protamine–DNA and displaced transition protein–histone complexes constitute a relatively stable transitional state, generating small subnucleosomal structures. HRCS might involve the small subnucleosomal structures, and our sample may contain them because MNase was not used in our method to discriminate the normal nucleosomal structures and the small subnucleosomal structures. Thus, further analyses might be needed to examine whether the small subnucleosomal structures might contribute to some H3-binding genes detected in HRCS.

We found HDS fractions, representing histone replacement-incomplete sperm, of 26.8%, 20.3%, and 11.6% of the total sperm from caput, corpus, and cauda epididymis, respectively. This suggests that sperm have not yet completely finished the histone-to-protamine replacement after leaving the testis, and histone replacement continues during movement in the epididymis. When sperm cells enter the epididymis, they have not yet acquired the capacity to move and are unable to fertilize the oocytes. They need the epididymal maturation process during movement in the epididymis to obtain fertilization capacity. The epididymal maturation involves a series of molecular events including the increased level of sialic acid residue, disulfide bond formation, membrane fluidity changes (caused by a decrease in the cholesterol/phospholipid ratio), phosphorylation of Izumo1, and tRNA fragmentation (see the review by Gervas and Viscott and references therein). However, the completion of histone-to-protamine replacement has not been recognized as an event of epididymal maturation, and the present study indicates that SCSA is a useful method to judge the completion of histone-to-protamine replacement in the study of mouse sperm. Our results also showed that the HDS fraction of the sperm from either cauda epididymis or vas deferens was approximately 10%. This may not be surprising because the HDS fraction of sperm from healthy men is also 5–10%. We have developed a method to purify HRCS using Percoll gradient centrifugation of high-density sperm. The amount of histone H3 in HRCS was about 1/5 that of swim-up and total sperm from cauda epididymis. Consistent with this, H3 binding in HRCS was observed in the promoter regions of fewer genes than in total sperm. These results suggest that preparation of high-density sperm using Percoll centrifugation is a more useful method than the widely used swim-up method, also in the clinical field of assisted reproduction techniques.

H3 binding mainly to the promoter region in HRCS suggests that its role might be transcriptional regulation after fertilization. The expression level of the H3-binding genes in HRCS tends to be high in oocytes and suppressed after the two-cell stage. Although further analysis is required to examine whether residual histones in sperm affect gene expression after fertilization, recent reports suggest that histones in gametes regulate epigenetic status in the fertilized egg. The previous study using swim-up sperm indicates that H3.3 is mainly incorporated into sperm chromatin at a higher rate than H3.1/H3.2. H3.3 contributes to gene silencing in ES cells by gain of H3K9me3 via ESET. These results suggest that inherited H3.3 from sperm to zygote might trigger heterochromatin formation in early embryogenesis,
despite the population of H3.3 derived from sperm being diluted during embryogenesis. Furthermore, H3K9me2 at the two paternally imprinted genes, H19 and Rasgfl1, in sperm is transmitted to the zygote, inhibits Tet3-dependent DNA demethylation, and maintains the DNA methylation status by recruiting PGC737. If a subset of H3 in HRCS has the H3K9me2 mark, such nucleosome regions may form silencing chromatin in early embryos by maintaining relatively higher DNA methylation or possibly recruiting gene silencing factors. Although our results indicate that H19 and Rasgfl1 are not involved in the H3-binding genes in HRCS (Supplementary Fig. 4d), H3K9me2 signal might be enriched in the small amount of histones on those genes. Similarly, H3K9me2 might be also enriched in the small amount of histone on the category_TS genes, which are suppressed during ZGA in the four-cell embryo. Alternatively, these epigenetic marks in the oocyte (or marks obtained in zygote during early embryonic development) on characterized promoter regions by CpG-enriched sequence may contribute to expresional regulation of these genes.

In the H3 target genes of HRCS, genes involved in neural cell differentiation (pathway of axon guidance and signaling by NGF) are enriched. ES cell differentiation is intrinsically directed to neural cells without external stimulus (the so-called default differentiation (pathway of axon guidance and signaling by NGF)). This pathway is essentially regulated by zinc-finger nuclear protein Zfp521 with the transcriptional co-activator p300, and antithetical expression of Zfp521 can convert ES cells to neural cells even in the presence of BMP4, an antagonistic factor of this pathway.40 Interestingly, Zfp521 and Ep300 (encoding p300) are included in the H3 target genes in HRCS (category_H). One possibility is that, in the early embryo, expression of these neural cell differentiation-related genes is silenced by some epigenetic mark derived from histones in the sperm, which might be important for suppressing the intrinsic capacity of ICM cells for neural differentiation and preventing ectopic differentiation of neural cells in the ICM. Further histone modification analysis in HRCS and in early embryos is required to understand the molecular mechanism of how sperm epigenetic information is transmitted to the embryo.

### Methods

#### Preparation of sperm fraction

Wild-type male mice on the C57BL/6 background were purchased from Japan SLC and used at 11–12 weeks of age. Sperm cells were collected from caput, corpus, or caudal epididymis or vas deferens. These tissues were incubated in 4 ml of M2 medium (M7167, Sigma) for 1 h at 37 °C, and then washed in PBS and centrifuged at 3000 × g for 5 min without brake to isolate the low-density somatic cells, and this step was repeated once. The cell pellet was then suspended into somatic cell lysis buffer (PBS containing 0.1% SDS and 0.5% Triton X-100) and centrifuged at 2500 × g for 5 min without brake to remove the low-density somatic cells, and this step was repeated once. The cell pellet was then suspended into somatic cell lysis buffer (PBS containing 0.1% SDS and 0.5% Triton X-100) and centrifuged at 2500 × g for 5 min without brake to remove the low-density somatic cells, and this step was repeated once. The cell pellet was then suspended into somatic cell lysis buffer (PBS containing 0.1% SDS and 0.5% Triton X-100) and centrifuged at 2500 × g for 5 min without brake to remove the low-density somatic cells, and this step was repeated once. The cell pellet was then suspended into somatic cell lysis buffer (PBS containing 0.1% SDS and 0.5% Triton X-100) and centrifuged at 2500 × g for 5 min without brake to remove the low-density somatic cells, and this step was repeated once. The cell pellet was then suspended into somatic cell lysis buffer (PBS containing 0.1% SDS and 0.5% Triton X-100) and centrifuged at 2500 × g for 5 min without brake to remove the low-density somatic cells, and this step was repeated once.
protease was added to TBE buffer before SDS. The input sample and eluted samples were incubated overnight at 65 °C for reversal of cross-links and then treated with RNAsase A and proteinase K. DNA was isolated using phenol: chloroform:isoamyl alcohol and ethanol precipitation, and purified with the QiAquick PCR Purification Kit (Qiagen).

Quantitative PCR was performed with a Quantifast SYBR Green PCR Kit (Qiagen) on a Quant Studio 3 Real-Time PCR System (Applied Biosystems) (primers, Supplementary Table 1). Each quantitative PCR analysis was run with technical triplicates. ChIP efficiency was presented as the percentage of the input sample used for the ChIP lsate.

Library preparation and sequencing for ChIP-seq analysis. Input and ChIP’d DNA samples were converted into sequencing library using NEBNext Ultra II DNA Library Prep Kit (New England Biolabs). Samples were end-repaired, adapter ligated, and size-selected using AMPure beads to obtain approximately 200–500 bp final library size. Indexed sequencing libraries were verified by DNA High-Sensitivity chip of Bioanalyzer (Agilent). Paired-end sequencing (2 x 151 cycles) was performed on Illumina NextSeq500 Mid Output Kit v2 (Illumina) at Tsukuba i-Laboratory LLP.

Processing read data and peak calling. Paired-end reads were mapped against mouse genome assembly mm10 using CLC Genomics Workbench (v10.1.1, Qiagen) with default settings. Mapped reads were exported as BAM files and converted to BED files by BED tool (v2.27.0) for downstream analysis. Only uniquely mapped reads were analyzed, and proper paired reads were used for peak calling. H3 peak positions were identified using MACS2 at FDR < 0.01 and FDR < 5.

Genome coordinates. Genomic compartments were defined with ncbiReSeq data obtained by UCSC table browser. The features were defined as follows: promoters (1 kb around the annotated TSS), exons, introns, upstream regions (from 5 kb upstream of the TSS), and downstream regions (from 5 kb downstream of the TSS). Intergenic regions were classified into intergenic repeat and intergenic unique regions by overlapping of repeat regions, defined with rmsk data obtained using UCSC table browser.

Quantification of o/e ratio in CpG. Genomic positions of CpG islands were defined with cgphlsandExt data obtained using UCSC table browser. Values of o/e ratios included in the above data were calculated.

Targeted bisulfite sequencing (TGBS). Targeted bisulfite sequencing (TGBS) libraries were prepared according to the post-bisulfite adapter tagging (PBAT) protocol with slight modifications. Briefly, 300 ng of genomic DNA was fragmented to ~500 bp with Covaris S220 and used for target enrichment with SureSelectXT Mouse Methyl-Seq kit (Agilent). The enriched DNA was subjected to Illumina HiSeq 2500 using HiSeq SR Rapid Cluster Kit V2 (Illumina) and HiSeq Rapid SBS Kit V2 (Illumina). Two indexed libraries were combined with a high concentration of spike-in PhiX control DNA (20%) and loaded on a single lane of the flow cell.

TGBS data analysis. TGBS reads were mapped to the mouse reference genome sequence (mm9) using Bmap (http://itolab.med.kyushu-u.ac.jp/BMap/index.html). Methylation levels were calculated for individual CG sites. Methylation levels of probe regions and H3 peak regions were determined by averaging the methylation levels of CG sites in individual features. The coordinates of H3 peak regions on mm10 assembly were converted to those on mm9 using the UCSC LiftOver tool for the integrative analyses of DNA methylation and histone deacetylation. This process filtered out seven out of 1320 H3 peak regions in HRCs and 41 out of 10,998 regions in total sperm. Correlation of methylation levels between HRCs and swim-up sperm was evaluated by calculating Pearson’s correlation coefficient and displayed as a scatter plot.

Comparison of our data with published datasets. Published ChIP-seq, ATAC-seq, and RNA-seq data were obtained from NCBI GEO to compare them with our results, and query codes of the datasets used are summarized in Supplementary Table 2.

Preparation of profile plot data and heatmap data. The number of raw reads mapped on 3 kbp around the TSS was counted, and read coverage was calculated after normalization by mapped read number. Matrix data of read coverage against each target promoter region in 50-bp bin size were prepared using deepTools244. Using this matrix, profile plots were prepared using average values for target regions. The degree of read coverage in each target promoter region is indicated by the heatmap.

Gene functional analysis. Functional analysis of target genes was performed for PANTHER classification system45. Using 820 category_H and S608 category.

Accepted: 14 August 2018
Published online: 24 September 2018

References
1. Ng, S. F. et al. Chronic high-fat diet in fathers programs β-cell dysfunction in female rat offspring. Nature 467, 963–966 (2010).
2. Carone, B. R. et al. Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. Cell 143, 1084–1096 (2010).
3. Radford, E. J. et al. In utero effects. In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. Science 345, 1255903 (2014).
4. Miska, E. A. & Ferguson-Smith, A. C. Transgenerational inheritance: models and mechanisms of non-DNA sequence-based inheritance. Science 354, 59–63 (2016).
5. Seong, K. H., Li, D., Shimizu, H., Nakamura, R. & Ishii, S. Inheritance of stress-induced, ATF-2-dependent epigenetic change. Cell 145, 1049–1061 (2011).
6. Siklenka, K. et al. Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. Science 350, aab2006 (2015).
7. Braun, R. E. Packaging paternal chromosomes with protamine. Nat. Genet. 28, 10–12 (2001).
8. Practice Committee of American Society for Reproductive Medicine. Practice Committee of Society for Assisted Reproductive Technology. Round spermatid nucleus injection (ROSI). Fertil. Steril. 90(Suppl. 3), S199–S201 (2008).
9. Mortimer, D. Sperm preparation methods. J. Androl. 21, 357–366 (2000).
10. Colleu, D., Lescoat, D. & Gouranton, J. Nuclear maturity of human spermatozoa selected by swim-up or by Percoll gradient centrifugation procedures. Fertil. Steril. 65, 160–164 (1996).
11. Sakkas, D. et al. The use of two density gradient centrifugation techniques and the swim-up method to separate spermatozoa from chromatin and nuclear DNA anomalies. Hum. Reprod. 15, 1112–1116 (2000).
12. Eveson, D. F. J. Sperm chromatin structure assay (SCSA)®. Methods Mol. Biol. 927, 147–164 (2013).
13. Virro, M. R., Larson-Cook, K. L. & Eveson, D. J. Sperm chromatin structure assay (SCSA) parameters are related to fertilization, blastocyst development, and ongoing pregnancy in in vitro fertilization and intracytoplasmic sperm injection cycles. Fertil. Steril. 81, 1289–1295 (2004).
14. Burgm, M. et al. The predictive value of sperm chromatin structure assay (SCSA) parameters for the outcome of intrauterine insemination, IVF and ICSI. Hum. Reprod. 19, 1401–1408 (2004).
15. Hammond, S. S. et al. Distinctive chromatin in human sperm packages genes for embryo development. Nature 460, 473–478 (2009).
16. Erkek, S. et al. Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. Nat. Struct. Mol. Biol. 20, 868–875 (2013).
17. Samans, B. et al. Uniformity of nucleosome preservation pattern in Mammalian sperm and its connection to repetitive DNA elements. Dev. Cell 30, 23–35 (2014).
18. Carone, B. R. et al. High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. Dev. Cell 30, 11–22 (2014).
19. Royo, H., Stadler, M. B. & Peters, A. H. Alternative computational analysis shows no evidence for nucleosyme enrichment at repetitive sequences in mammalian spermatozoa. Dev. Cell 37, 98–104 (2016).
20. Dansranjavin, T. & Schagdarsurengin, U. The rationale of the inevitable, or why is the consideration of repetitive DNA elements indispensable in studies of sperm nucleosomes. Dev. Cell 37, 13–14 (2016).

21. Saitou, M. & Kurimoto, K. Paternal nucleosomes: are they retained in human and mouse spermatozoa. Nat. Protoc. 8, 2449–2470 (2013).

22. Hisano, M. et al. Genome-wide chromatin analysis in mature mouse and human spermatozoa. Nat. Struct. Mol. Biol. 17, 679–687 (2010).

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

24. Hud, N. V., Milanovich, F. P. & Balhorn, R. Evidence of novel secondary structure of DNA-bound protamine is revealed by Raman spectroscopy. Biochemistry 33, 7528 (1994).

25. Vilfan, I. D., Conwell, C. C. & Hud, N. V. Formation of native-like mammalian sperm chromatin with folded bulb protamine. J. Biol. Chem. 279, 20088–20095 (2004).

26. Carrell, D. T., Emery, B. R. & Peterson, C. M. The correlation of sperm chromatin decondensation following in vitro exposure to heparin and sperm penetration rates. J. Assist. Reprod. Genet. 15, 560–564 (1998).

27. Wu, J. et al. The landscape of accessible chromatin in mammalian preimplantation embryos. Nature 534, 652–657 (2016).

28. Liu, X. et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature 537, 558–562 (2016).

29. Barral, S. et al. Histone variant H2A.L2 guides transition protein-dependent protamine assembly in male germ cells. Mol. Cell 66, 89–101 (2017).

30. Gervasi, M. G. & Visconti, P. E. Molecular changes and signaling events occurring in spermatozoa during epididymal maturation. Andrology 5, 204–218 (2017).

31. Wyrobek, J. I. et al. Advancing age has differential effects on DNA damage, chromatin integrity, gene mutations, and aneuploidy in sperm. Proc. Natl. Acad. Sci. USA 103, 9601–9606 (2006).

32. Campos, E. I., Stafford, J. M. & Reinberg, D. Epigenetic inheritance: histone mark acetylation and deposition are maintained across spermatogonial generations of young, adult, and aged mice. Trends Cell Biol. 24, 664–674 (2014).

33. Xu, Q. & Xie, W. Epigenome in early mammalian development: inheritance, reprogramming and establishment. Trends Cell Biol. 28, 237–253 (2018).

34. Voon, H. P. et al. ATRX Plays a key role in maintaining silencing at interstitial heterochromatic loci and imprinted genes. Cell Rep. 11, 405–418 (2015).

35. Nakamura, T. et al. PCG7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. Nature 486, 415–419 (2012).

36. Smukler, S. R. et al. Embryonic stem cell cultures maintain a primitive neural stem cell fate in the absence of extrinsic influences. J. Cell. Biol. 172, 79–90 (2006).

37. Kamiya, D. et al. Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. Nature 470, 503–509 (2011).

38. Yoshida, K. et al. The transcription factor ATF7 mediates lipopolysaccharide-induced epigenetic changes in macrophages involved in innate immunological memory. Nat. Immunol. 16, 1034–1043 (2015).

39. Miura, F. et al. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. Nucleic Acids Res. 40, e136 (2012).

40. Kamiya, D. et al. Large-scale gene function analysis with the PANTHER classification system. Nat. Protoc. 8, 1551–1566 (2013).

Acknowledgements
We thank K. Mochida and A. Ogura (RIKEN) for showing the unpublished data and useful discussion; K. Yamaguchi and Y. Okada (University of Tokyo) for useful discussions. This work was supported in part by a Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and by AMED under Grant Numbers JP16am0101059, JP16am0101060, JP18am0101103, and JP18am0101105.

Author contributions
K.Y. performed SCSA and immunoblotting analysis for each sperm sample, and did H3 ChIP-seq experiment. T.S. and N.D. performed quantitative protein analysis by LC-MS. M.M., R.N., and K.S. analyzed the data from ChIP-seq experiments. F.M., H.A., and T.I. analyzed the data from bisulfite-seq experiments. S.I. conceived and supervised the whole study, and S.I. and K.Y. wrote the paper.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-06243-9.

Competing interests: M.M. declares association with Tsukuba-i Laboratory as a technical consultant.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

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

© The Author(s) 2018