

Erasure of DNA methylation, genomic imprints, and epimutations in a primordial germ-cell model derived from mouse pluripotent stem cells

Norikatsu Miyoshi^{a,1}, Jente M. Stel^{a,b,1}, Keiko Shioda^a, Na Qu^a, Junko Odajima^a, Shino Mitsunaga^a, Xiangfan Zhang^c, Makoto Nagano^c, Konrad Hochedlinger^{a,d}, Kurt J. Isselbacher^{a,2}, and Toshi Shioda^{a,2}

^aMassachusetts General Hospital Center for Cancer Research, Charlestown, MA 02129; ^bInstitute for Environmental Studies Vrije Universiteit, Amsterdam 1081 HV, The Netherlands; ^cDepartment of Obstetrics and Gynecology, McGill University and Research Institute of McGill University Health Centre, Montreal, QC, Canada H4A 3J1; and ^dDepartment of Stem Cell and Regenerative Biology, Harvard Stem Cell Institute, Cambridge, MA 02138

Contributed by Kurt J. Isselbacher, June 28, 2016 (sent for review February 22, 2016; reviewed by Piroska Szábo and Moshe Szyf)

The genome-wide depletion of 5-methylcytosines (5mCs) caused by passive dilution through DNA synthesis without daughter strand methylation and active enzymatic processes resulting in replacement of 5mCs with unmethylated cytosines is a hallmark of primordial germ cells (PGCs). Although recent studies have shown that *in vitro* differentiation of pluripotent stem cells (PSCs) to PGC-like cells (PGCLCs) mimics the *in vivo* differentiation of epiblast cells to PGCs, how DNA methylation status of PGCLCs resembles the dynamics of 5mC erasure in embryonic PGCs remains controversial. Here, by differential detection of genome-wide 5mC and 5-hydroxymethylcytosine (5hmC) distributions by deep sequencing, we show that PGCLCs derived from mouse PSCs recapitulated the process of genome-wide DNA demethylation in embryonic PGCs, including significant demethylation of imprint control regions (ICRs) associated with increased mRNA expression of the corresponding imprinted genes. Although 5hmCs were also significantly diminished in PGCLCs, they retained greater amounts of 5hmCs than intragonadal PGCs. The genomes of both PGCLCs and PGCs selectively retained both 5mCs and 5hmCs at a small number of repeat sequences such as GSAT_MM, of which the significant retention of bisulfite-resistant cytosines was corroborated by reanalysis of previously published whole-genome bisulfite sequencing data for intragonadal PGCs. PSCs harboring abnormal hypermethylation at ICRs of the *Dlk1-Gtl2-Dio3* imprinting cluster diminished these 5mCs upon differentiation to PGCLCs, resulting in transcriptional reactivation of the *Gtl2* gene. These observations support the usefulness of PGCLCs in studying the germline epigenetic erasure including imprinted genes, epimutations, and erasure-resistant loci, which may be involved in transgenerational epigenetic inheritance.

PGCLC | epigenetic reprogramming | genetic imprinting | epimutation

Evidence is accumulating that parental experiences such as pain, nutritional restrictions, or exposure to toxic chemicals can be transmitted to subsequent generations via epigenetic alterations without mutations in the genomic DNA (gDNA) (1–3). Multigenerational transmission of a nongenetic phenotype is considered *transgenerational* when it is persistent beyond the epigenetic reprogramming in primordial germ cells (PGCs) (1, 2), potentially conveying illness including metabolic diseases, malignancies, reproductive defects, or behavioral alterations (2, 4, 5). However, this is still a controversial subject due partly to the lack of direct experimental demonstration of transgenerational epigenetic alterations escaping the epigenetic erasure in mammalian PGCs (2, 6, 7).

In early stage mouse embryos, a small cluster of Prdm1-positive PGCs consisting of about 40 cells arise in epiblast at embryonic day 7.25 (E7.25), and PGCs migrate toward the genital ridges while they are rapidly proliferating. By E12.5, about 25,000 PGCs settle in the genital ridges and cease cell division (8). Genome-wide gDNA demethylation is initiated in the migrating PGCs and completed in the intragonadal PGCs, decreasing the global CpG methylation

level from 70% in E6.5 epiblast to about 10% in E13.5 PGCs (9). This massive genome-wide gDNA demethylation is critical for “resetting” the sex-specific epigenetic status of imprinted genes, which is important for normal development of fetuses in the subsequent generation, and it is achieved through passive dilution of 5-methylcytosines (5mCs) in the absence of the Dnmt1/Np95-dependent maintenance methylation of the daughter strands during DNA replication as well as multistep enzymatic processes resulting in replacement of 5mCs with unmethylated cytosines, which may involve 5-hydroxymethylcytosines (5hmCs) as intermediates (9–14). A small fraction of genomic elements such as mouse intracisternal A particles (IAP) was reported to escape this global gDNA demethylation, and their possible roles in the transgenerational epigenetic inheritance have been proposed (2, 9, 15). On the other hand, a recent study detected aberrant 5mC distributions in the spermatogonial gDNA of mice prenatally exposed to endocrine disruptors, but these epimutations were not persistent in the subsequent generation beyond the germline epigenetic reprogramming (6). The fate of epimutations introduced in the reprogramming-resistant genomic elements still remains to be documented.

Significance

Whether acquired epigenetic changes can escape the genome-wide epigenetic erasure in the primordial germ cells, which are the embryonic precursors of all types of germline cells and gametes, resulting in transgenerational transfer has been under debate. We have shown that an *in vitro* cell culture model of mouse primordial germ cells effectively recapitulates the process of germline epigenetic erasure, including DNA demethylation at both physiologically methylated and abnormally hypermethylated imprinting control regions. We also have identified examples of genomic repetitive sequences characterized by significant resistance to the genome-wide DNA demethylation process in mouse primordial germ cells and their cell culture models. Our study paves the way for mechanistic studies of transgenerational epigenetic inheritance using a cell culture model.

Author contributions: N.M., M.N., K.H., K.J.I., and T.S. designed research; N.M., K.S., N.Q., J.O., S.M., X.Z., M.N., and T.S. performed research; J.M.S. and T.S. analyzed data; and N.M., M.N., K.H., K.J.I., and T.S. wrote the paper.

Reviewers: P.S., Van Andel Research Institute; and M.S., McGill University.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: Affymetrix microarray and deep-sequencing data have been deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (accession nos. GSE80983 and GSE81175) and Sequence Read Archive (SRA) database (accession no. SRP074457).

¹N.M. and J.M.S. contributed equally to this work.

²To whom correspondence may be addressed. Email: kiselbacher@mgh.harvard.edu or tshioda@mgh.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1610259113/-DCSupplemental.

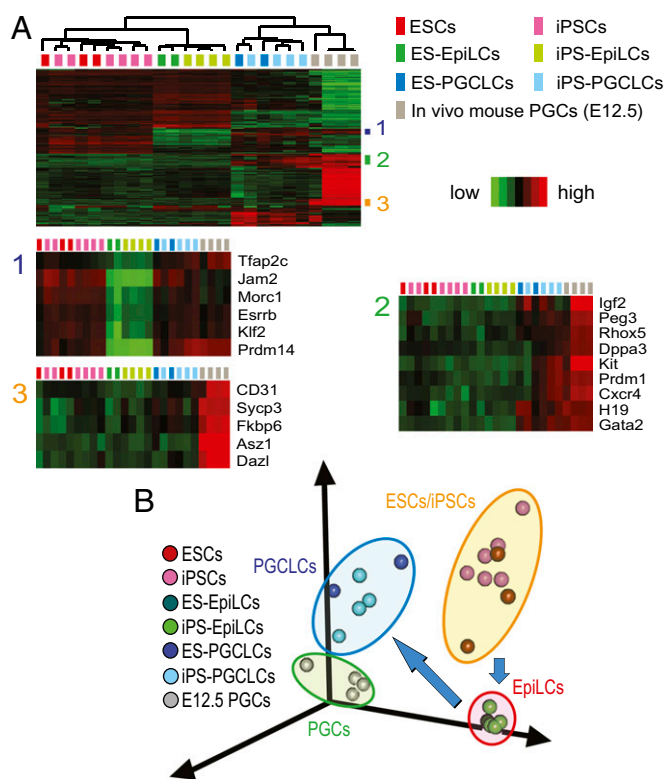


Fig. 1. Transcriptomes of mouse PSCs, EpiLCs, PGCLCs, and in vivo PGCs. (A) Hierarchical clustering heatmap of differentially expressed genes. EpiLCs and PGCLCs are indicated with their precursor PSCs (e.g., ES-EpiLCs are EpiLCs derived from ESCs). The three gene clusters indicated in the *Top* heatmap are enlarged in the *Bottom* heatmaps. (B) PCA of transcriptomal changes during differentiation of PSCs to PGCLCs via EpiLCs.

Recently, it has been shown that pluripotent stem cells (PSCs) such as embryonic stem cells (ESCs) or induced pluripotent stem cells (iPSCs) can be differentiated into PGC-like cells (PGCLCs) in vitro (16). For example, Hayashi et al. produced PGCLCs from mouse PSCs via the generation of epiblast-like cells (EpiLCs) as intermediates (17, 18). To examine advantages and limitations of mouse PGCLCs as a cell culture model for studies on transgenerational epigenomics, we performed microarray-based transcriptomal profiling and deep-sequencing analyses of genomic 5mC and 5hmC distributions in PGCLCs and compared these genomic characteristics with those of E12.5 mouse intragonadal PGCs. We show genome-wide dynamics of 5mC and 5hmC erasure during PSC differentiation to PGCLCs via EpiLCs, demonstrating precise recapitulation of the DNA methylome, including previously known and unknown gDNA elements resistant to the global erasure of 5mCs and 5hmCs. We also demonstrate that transcription-suppressing abnormal hypermethylation at the imprinting control region (ICR) of the *Dlk1-Gtl2-Dio3* imprinting cluster in iPSCs was erased upon differentiation to PGCLCs to regain mRNA expression. These observations support the use of mouse PGCLCs for mechanistic studies of germline epigenetic reprogramming and transgenerational epigenetic inheritance as a valid model of embryonic PGCs.

Results

The SSEA1⁺/Integrin β ³⁺/c-Kit⁺ Triple-Positive Mouse PGCLCs Resemble Early Stage PGCs in Marker mRNA Expression. Mouse E12.5 intragonadal PGCs characterized by germline-specific transcriptional activation driven by the *Pou5f1* distal enhancer/promoter (Fig. S1A) (19) and alkaline phosphatase activity (Fig. S1B) were examined for their surface-marker protein expression by FACS, which revealed their SSEA1⁺/Integrin β ³⁺/c-Kit⁺ triple-positive

status (Fig. S1C and D). Following the protocol described by Hayashi et al. (17), we produced mouse EpiLCs and the day-6 PGCLCs from PSCs (Fig. S1E). More than 98% of PGCLCs enriched by FACS as SSEA1⁺/Integrin β ³⁺ double-positive cells also strongly expressed c-Kit (Fig. S1F, *Top* row) whereas only 36% of SSEA1⁺/c-Kit⁺ double-positive cells were Integrin β ³⁺-positive (Fig. S1F, *Bottom* row). In the present study, the SSEA1⁺/Integrin β ³⁺ double-positive day-6 PGCLCs, which were almost triple-positive including c-Kit, were subjected to further analyses. When transplanted into mouse seminiferous tubules, PGCLCs visualized by EGFP expressed by the *Pou5f1* distal enhancer/promoter [which is active in PGCLCs/PGCs (19) and spermatogonial stem cells (20)] or mCherry expressed by the human EF1 α promoter (also active in mouse germline cells) colonized in the lumen of the tubules (Fig. S1G), agreeing with the original report by Hayashi et al. about the capacity of PGCLCs to develop spermatogenic colonies as transplants in the tubules (17).

Unsupervised hierarchical clustering (Fig. 1A) and principal component analysis (PCA) (Fig. 1B) clearly separated transcriptomes along cell types—namely, PSCs, EpiLCs, PGCLCs, and intragonadal PGCs. Transcriptomes of PGCLCs were not separated along the types of PSCs from which they were derived (i.e., ESCs or iPSCs). The transcriptomes among the individual PSC clones showed significant heterogeneity and became remarkably homogeneous upon differentiation to EpiLCs, but diversified again among PGCLCs (Fig. 1B), suggesting that differentiation to EpiLC was a nearly deterministic process whereas commitment to PGCLC seemed stochastic. Among the genes induced upon EpiLC differentiation to PGCLCs, those belonging to clusters 1 and 2 in Fig. 1A were enriched with early markers of PGCs. Cluster 2 was also enriched with imprinted genes. Cluster 3 genes were more strongly expressed in intragonadal PGCs than in PGCLCs and enriched with markers of late-stage PGCs.

Expression of *Fgf5* [an early stage EpiLC maker (17)] was strong in EpiLCs but reduced in PGCLCs whereas expression of *Wnt3* [a late-stage EpiLC maker (17)] was maintained in both EpiLCs and PGCLCs (Fig. S2A). PGCLCs strongly expressed mRNA markers of committed and/or migrating PGCs (e.g., *Prdm1*, *Prdm14*, *c-Kit*, and *Tfap2c*, Fig. S2B). Induction of *Dppa3* and suppression of *c-Myc*, which were reported to occur in PGCLCs after expression of the migrating PGC markers (17), were observed in our PGCLCs (Fig. S2C) whereas intragonadal PGC markers (*Dazl* or *Ddx4*, Fig. S2D) were not induced. Agreeing with a previous report that *Snail1* was transiently expressed during EpiLC differentiation to PGCLCs but later suppressed when intragonadal PGC markers were induced (17), our PGCLCs expressed *Snail1* but intragonadal PGCs did not (Fig. S2B). Our PGCLCs expressed all of the three Tet enzymes (Fig. S2E). Compared with EpiLCs, expression of the *Dnmt3a* and *Dnmt3b* de novo DNA methyltransferases as well as the *Uhrf1/Np95* cofactor of *Dnmt1* was reduced in PGCLCs whereas expression of the *Dnmt1* maintenance DNA methyltransferase was maintained (Fig. S2F), agreeing with a previous study (17). Expression of the pluripotency genes *Pou5f1*, *Klf4*, *Sox2*, and *Nanog* (Fig. S2G) as well as *Tdg* and *Aicda* encoding thymine-DNA glycosylase and activation-induced cytidine deaminase, respectively, was stronger in PGCLCs than in intragonadal PGCs (Fig. S2H). Quality control analysis of microarray signal intensities confirmed the absence of significant batch effects that could have affected the above observations (Fig. S3A). Taken together, our transcriptomal profiling suggests that the differentiation status of our PGCLCs was comparable to the PGCLCs described by Hayashi et al. (17), presumably close to the migrating E8.5–E9.5 PGCs.

Erasure of 5mCs and 5hmCs in PGCLCs. To examine the epigenetic status of PGCLCs, we determined distributions of 5mCs and 5hmCs in the genomes of mouse iPSCs, EpiLCs, PGCLCs, and E12.5 intragonadal PGCs by deep sequencing of gDNA fragments enriched for 5mCs using biotin-conjugated methylcytosine-binding protein 2 [MBD-sEq. (21)], which has no significant affinity to 5hmCs (22), and gDNA fragments enriched for 5hmCs by

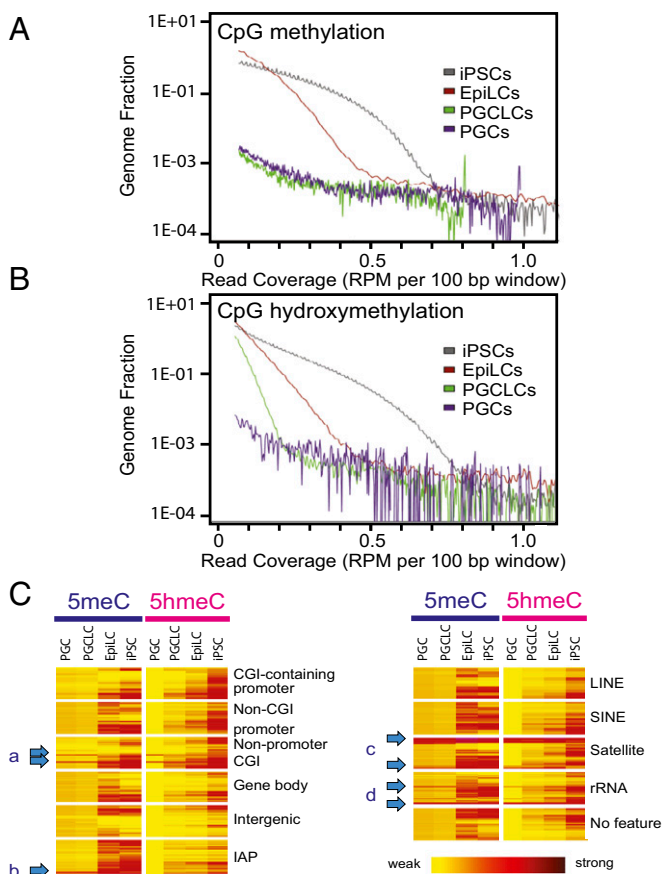


Fig. 2. Global reduction in gDNA 5mCs and 5hmeCs during mouse iPSC differentiation to PGCLCs. (A and B) Density distributions of (A) 5mCs and (B) 5hmeCs. The x axes represent densities of 5mCs or 5hmeCs in 100-bp windows, and the y axes indicate genome-wide frequencies. (C) Heatmaps of 5mC and 5hmeC densities across genomic features. Arrows *a–d* point to elements retaining 5mCs and/or 5hmeCs in PGCLCs/PGCs.

chemical labeling with no reactivity to 5mCs (23). Thus, in contrast to the bisulfite sequencing that cannot distinguish 5mCs and 5hmeCs (24), our approach permitted differential detection of gDNA fragments enriched with these two types of cytosine modifications. Deep-sequencing quality control assessments confirmed sufficient CpG site coverage and saturation in our analyses (Figs. S3 B–E and S4).

Distribution plot analyses revealed significant reduction in both 5mCs and 5hmeCs during differentiation of iPSCs to EpiLCs. EpiLC differentiation to PGCLCs further reduced 5mCs to a level that appeared comparable to E12.5 PGCs with the sensitivity of our 5mC detection method (Fig. 2A) whereas PGCLCs retained weak but significant amounts of 5hmeC-enriched gDNA segments compared with PGCs (Fig. 2B). Heatmaps of 5mC and 5hmeC distributions across the functional gDNA features revealed that a small fraction of gDNA elements at the nonpromoter CpG islands (Fig. 2C, *a*), IAPs (Fig. 2C, *b*), satellite repeats (Fig. 2C, *c*), and rRNA genes (Fig. 2C, *d*) concomitantly retained these epigenetic marks in both PGCLCs and PGCs. The contents of 5mCs detected by MBD-seq were indistinguishable between PGCLCs and PGCs across the gDNA features, whereas the contents of 5hmeCs were more significantly diminished in PGCs compared with PGCLCs (Fig. 2C). Detailed classification of 5mC-enriched gDNA fragments across genomic features revealed their strong enrichment in gene bodies and intergenic regions outside repetitive sequences in iPSCs and EpiLCs, whereas enrichment of these features was remarkably diminished in PGCLCs and PGCs (Fig. 3A and Fig. S5A). The 5mC enrichment profiles

of PGCLCs and PGCs show significant similarities in both relative distributions across genomic features and total numbers of the 5mC-enriched regions (2,178 in PGCLCs vs. 2,791 in PGCs), and 91% of the 5mC-enriched regions detected in PGCLCs were also found in PGCs (Fig. S6). The 5hmeC enrichment profiles of iPSCs and EpiLCs were similar to 5mCs except that only 251 5hmeC-enriched regions (assigned mostly to repetitive elements) were found in PGCs (Fig. 3A and Fig. S5A). The majority of the repeat-containing, 5mC-enriched gDNA regions in PGCLCs and PGCs were found within the interspersed repeat classes such as SINES, LINES (short- and long-interspersed nuclear elements), or LTRs (which include the IAPs), approximately reflecting the genome-wide RepeatMasker registration profile of the mouse NCBI37/mm9 reference genome sequence (Fig. 3B and Fig. S5B). Interestingly, the satellite repeats (shown as *Sa) were overrepresented in all 5mC-enriched regions, and their proportion was increased further in the 50 regions with the highest relative methylation scores. Among the satellite sequences, the closely related GSAT_MM (shown as **GS) and SYNREP_MM (#SY) repeats were overrepresented.

To obtain further evidence of 5mC retention at the repetitive elements, we performed visual inspections of deep-sequencing data generated in our present study, as well as the whole-genome bisulfite sequencing (WGBS) data of mouse E6.5 epiblasts and E13.5 male PGCs published by Seisenberger et al. (9). Fig. S7A shows an example of deep-sequencing tracks demonstrating significant retention of both 5mCs and 5hmeCs at a region containing IAPs in PGCLCs and PGCs. Fig. S7B shows the WGBS data corresponding to a part of the IAP-related 5mC/5hmeC-enriched region indicated in Fig. S7A, demonstrating significant retention of bisulfite-resistant cytosines (i.e., the sum of 5mCs and 5hmeCs) at two CpG sites in the gDNA of E13.5 male PGCs. Fig. 3C and D shows similar analyses for a region rich in GSAT_MM and SYNREP_MM repeats. Although some 5mC/5hmeC peaks in the deep-sequencing tracks were not informative, as they were also evident in the nonenriched mouse genome resequencing track (peak *e*), several informative peaks (*a*, *b*, *d*) supported the presence of 5mC- and 5hmeC-enriched gDNA regions within GSAT_MM repeats (Fig. 3C). Inspection of the WGBS data for GSAT_MM repeats in the corresponding region identified three instances of an identical 74-nt sequence containing three CpG sites with significant retention of bisulfite-resistant cytosines in the gDNA of E13.5 male PGCs (Fig. 3D). On the other hand, the apparent lack of 5mC/5hmeC peaks at SYNREP_MM in Fig. 3C (peak *c* on the nonenriched track) left the 5mC/5hmeC retention in this element unconfirmed, possibly due to technical issues stemming from its up to 75% nucleotide base identity to GSAT_MM. Two additional examples of GSAT_MM retention of 5mC/5hmeC peaks and bisulfite-resistant cytosines are shown in Fig. S7C–F. Agreeing with the retention of 5mCs and 5hmeCs at the ribosomal RNA gene shown in Fig. 2C (arrow *d*), visual inspection of deep-sequencing tracks at regions containing LSU_rRNA_Hsa and SSU_rRNA_Hsa ribosomal RNA genes revealed the presence of informative peaks (Fig. S7G and H) although insufficient bisulfite conversion of the WGBS data for these regions precluded nucleotide base-resolution analysis. Interestingly, we observed a strong tendency for 5hmeC peaks to be closely associated with 5mC peaks (Fig. 3C and Fig. S7A, C, E, G, and H) although the enrichment-based deep-sequencing approach did not provide relative amounts of 5hmeCs to 5mCs.

DNA Demethylation at the ICRs in PGCLCs. Demethylation of the ICRs is a hallmark of intragonadal PGCs (9, 12). Hayashi et al. reported highly limited ICR demethylation in their PGCLCs, the epigenetic status of which was hence presumed by the authors to be similar to E8.5–E9.5 migrating PGCs before initiation of the imprinting erasure (17, 18). In contrast, Zhou et al. recently reported more advanced ICR demethylation in PGCLCs, placing their epigenetic status close to E12.5 intragonadal PGCs (25). For all of the six ICRs examined, our deep-sequencing analysis showed progressive loss of 5mCs upon iPSC differentiation to EpiLCs and then to PGCLCs (Fig. 4A and Fig. S8A–E). Expression of the

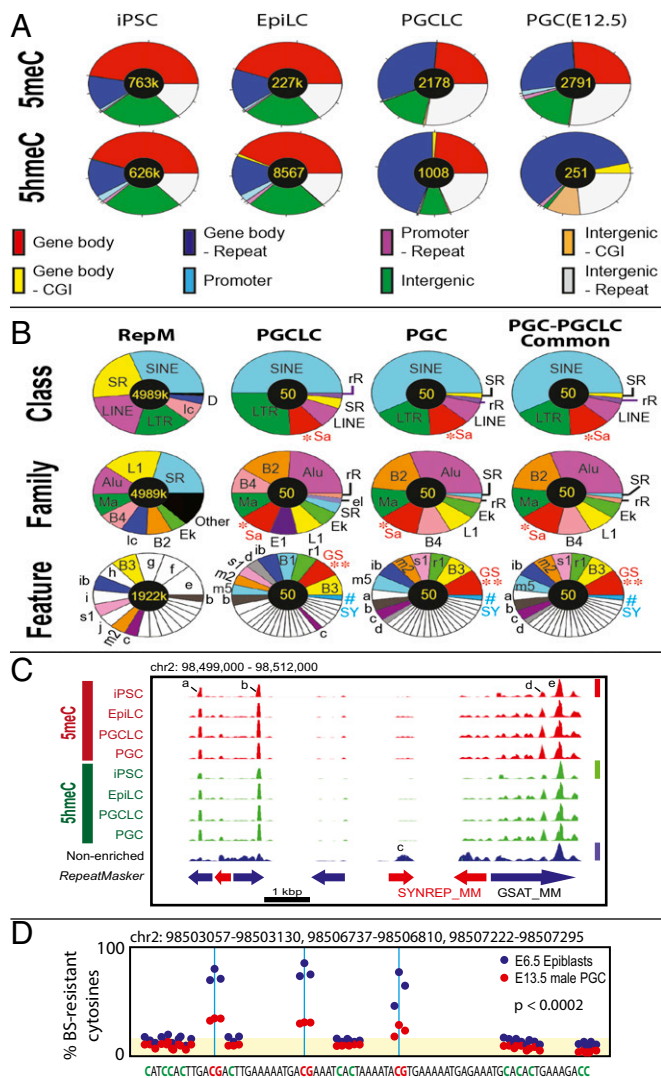


Fig. 3. Genomic feature distributions of 5mCs and 5hmCs in the genomic DNA of mouse iPSCs, EpiLCs, PGCLCs, and in vivo PGCs. (A) 5mC and 5hmC distributions across repeat sequences. (B) Distributions of 5mCs across repeat sequences. RepM, genome-wide RepeatMasker-registered elements. Small elements (<5%) are left blank in pie charts. *Sa, satellite repeats; **GS, GSAT_MM; #SY, SYNREP_MM. Other keys of pie charts are defined in Fig. S5. (C) An example of deep-sequencing tracks showing 5mC and 5hmC peaks at GSAT_MM and SYNREP_MM satellite repeats. Height of peaks reflects relative strength of DNA methylation across the four 5mC tracks (linearly scaled 0–1 between the baseline and the maximal methylation, red bar), DNA hydroxymethylation (four 5hmC tracks, green bar), or nonenriched genome resequencing (blue bar); note that scaled value 1 is not equal to 100% methylation. Peaks a, b, and d are “informative” based on their enrichment over the non-enriched mouse genome resequencing track or changes between different types of cells. Peaks c and e are present in the nonenriched track and so are uninformative. (D) Reanalysis of the whole-genome bisulfite sequencing data generated by Seisenberger et al. (9) for a 74-nt sequence repeated three times in the GSAT_MM regions shown in C. Blue and red dots show percentage of bisulfite-resistant cytosines in E6.5 epiblasts and E13.5 PGCs, respectively. Yellow shade indicates the background levels of bisulfite-resistant cytosines in CpA, CpT, and CpC dinucleotides. The P values represent statistical significance between the CpG-context bisulfite-resistant cytosines over the background (t test).

mRNA transcripts for the corresponding imprinted genes increased in PGCLCs compared with PSCs or EpiLCs but still more weakly than in E12.5 intragonadal PGCs, suggesting that the epigenetic status of PGCLCs produced in our present study may be between E9.5 and E12.5 PGCs (Fig. S2J). Significant and

progressive ICR demethylation was observed in all individual PSC-EpiLC-PGCLC differentiation experiments with no apparent differences among the PSC precursor clones (Fig. S8F). On the other hand, at the location of an IAP shown in Fig. S7A, 5mCs and 5hmCs were retained in the genomes of all types of PGCLCs as well as E12.5 embryonic PGCs (Fig. S8G). Note that no 5mC or 5hmC peak was detected in the genomes of PGCLCs or PGCs around the ICRs shown in Fig. S8A–F due to the absence of IAP, GSAT_MM, LSU_rRNA_Hsa, or SSU_rRNA_Hsa repeat sequences. Interestingly, the ICR demethylation observed upon differentiation of PSCs to PGCLCs was often accompanied by increased DNA hydroxymethylation at the same region, whereas DNA hydroxymethylation outside the ICRs was typically diminished or unchanged upon PSC differentiation to PGCLC (Fig. S8H).

Erasure of Region-Specific Epimutations During iPSC Differentiation to PGCLCs. We previously showed that generation of iPSCs by somatic cell reprogramming in the absence of sufficient vitamin C caused silencing of the *Dlk1-Gtl2-Dio3* imprinting cluster, resulting in diminished pluripotency (26, 27). This silencing was associated with aberrant DNA hypermethylation of maternal IG-DMR (differentially methylated region) and *Gtl2*-DMR (26, 27). Taking advantage of this epimutation that is experimentally inducible in iPSCs, we examined whether aberrant, region-specific hypermethylation can be erased during iPSC differentiation to PGCLC. Reproducing our previously published bisulfite-pyrosequencing analysis (26), MBD-seq detected aberrant DNA hypermethylation at the IG-DMR and the *Gtl2*-DMR in mouse iPSCs (Fig. 4A and Fig. S9A). The accuracy of our 5mC profiling is supported by the nearly identical MBD-seq tracks of normal [*Gtl2*(+)] and silenced [*Gtl2*(–)] iPSCs except for the IG- and *Gtl2*-DMRs. Whereas these aberrant 5mC peaks were still observed in EpiLCs, they were not detected in PGCLCs. Concomitantly, *Gtl2* mRNA expression, which was suppressed in *Gtl2*(–) iPSCs, was restored in PGCLCs to a level comparable to PGCLCs derived from *Gtl2*(+) iPSCs (Fig. 4B). Interestingly, the IG-DMR and the region between the IG- and the *Gtl2*-DMRs of *Gtl2*(–) iPSCs showed aberrant reduction in 5hmC peaks (Fig. 4A and Fig. S9B), which were erased during iPSC differentiation to PGCLC. Thus, the aberrant DNA hypermethylation at the ICRs of the *Dlk1-Gtl2-Dio3* imprinting cluster in iPSCs was erased upon differentiation to PGCLCs.

Discussion

Transcriptomal and Epigenomic Characteristics of Mouse PGCLCs. Following the protocol described by Hayashi et al. (17) with slight modifications, we generated SSEA1⁺/Integrin β 3⁺/c-Kit⁺ triple-positive PGCLCs from mouse PSCs (Fig. S1). Transcriptomal profiling (Fig. 1 and Fig. S2) placed our PGCLCs isolated from 6-d culture embryoid bodies (EBs) in a status similar to the PGCLCs that Hayashi et al. obtained from EBs earlier than the 6-d culture but later than the 2-d culture (17). In a recent study, Zhou et al. generated mouse PGCLCs from 6-d culture EBs using a similar protocol (25) and observed a marker gene expression profile similar to the 6-d EB PGCLCs of Hayashi et al. (17). On the other hand, whereas Hayashi et al. observed only limited DNA demethylation at ICRs of the *Igf2r*, *Snmpn*, *H19*, and *Kcnq1* imprinting clusters and so placed their PGCLCs at a stage corresponding to E8.5–E9.5 migrating PGCs in mouse embryos [when the ICR demethylation in PGCs is not yet significant (9, 17, 18)], Zhou et al. reported more advanced ICR demethylation at the *Snmpn* and *H19* imprinting clusters, placing their PGCLCs at a stage similar to E12.5 intragonadal mouse embryonic PGCs (25). In our present study, PGCLCs showed significant demethylation at all six ICRs examined (*Dlk1-Meg3/Gtl2-Dio3*, *H19*, *Igf2r*, *Kcnq1*, *Nespas-Gnas*, *Meg1/Grb10*) (Fig. 4 and Fig. S8) as well as global loss of 5mCs (Figs. 2A and C and 3A). The progressive increase in mRNA expression of imprinted genes during PSC differentiation to PGCLC via EpiLC (Fig. S2J) may reflect release from monoallelic suppression by DNA methylation. The restoration of *Gtl2* mRNA expression in PGCLCs derived from *Gtl2*(–) iPSCs

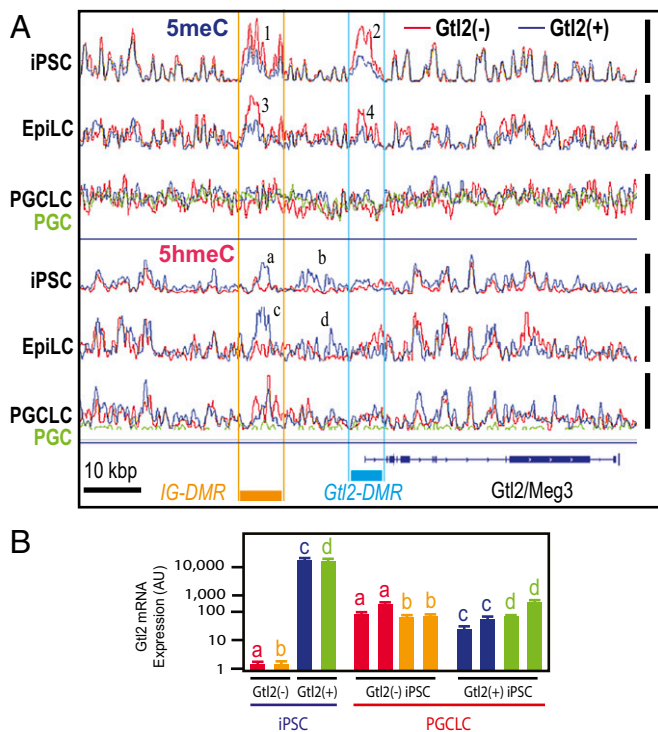


Fig. 4. Erasure of DNA hypermethylation at the IG-DMR and Gtl2-DMR of Gtl2(-) iPSCs during differentiation to PGCLCs. (A) Superimposed deep-sequencing tracks of 5mCs (Top three tracks) and 5hmeCs (Bottom three tracks). Blue, red, and green traces represent Gtl2(+), Gtl2(-), and in vivo PGC, respectively, and all traces in each track are adjusted in a track-specific linear scale between the minimal and maximal methylation or hydroxymethylation in the displayed area shown with vertical bars at the right. The same data are displayed with fixed scales across tracks in Fig. S9. Orange and cyan bars indicate locations of IG-DMR and Gtl2-DMR, respectively. Numbers 1–4 show differential methylation between Gtl2(+) and Gtl2(-) iPSCs and EpiLCs at the DMRs. (a–d) Differential hydroxymethylation. (B) Expression of Gtl2 mRNA in independent clones of mouse Gtl2(-) iPSCs (a and b), Gtl2(+) iPSCs (c and d), and PGCLCs produced from them. Bars indicate qPCR data for Gtl2 mRNA expression normalized with Gapdh mRNA expression ($n = 3$, mean \pm SEM).

from silencing due to the aberrant hypermethylation of the ICR of the *Dlk1-Gtl2-Dio3* imprinting cluster [Fig. 4 and Fig. S9 (26)] further supports ICR demethylation in our PGCLCs. Although our results suggest the usefulness of mouse PGCLCs for mechanistic studies of the germline DNA demethylation including ICRs, it remains to be determined whether this in vitro model accurately represents a particular physiological status of embryonic PGCs. To achieve this goal, future studies should consider sensitivity, quantitativeness, and specificity of the analytical methods. For example, PCR-based bisulfite sequencing may be insufficient for quantitative evaluation of ICR methylation (17, 18, 25). Specificity of bisulfite conversion (9, 17, 18, 25) is incomplete because it does not distinguish 5mCs from 5hmeCs. Whereas MBD-seq distinguishes 5mCs from 5hmeCs, in our present study this method did not robustly detect low levels of DNA methylation at the IG-DMR, Gtl2-DMR, or the Rtl1 in E12.5 PGCs, which was detected by Singh et al. using the Methylated CpG Island Recovery Assay (MIRA) and a custom-design microarray that targeted imprinted genes and IAP flanking regions (28). In contrast to MBD-seq using the 5mC-binding domain of human MBD2 for enrichment, MIRA uses heterodimers of MBD2b and MBD3L1, which has a significantly stronger affinity to 5mCs than MBD2 (29). Thus, the absence of 5mC in our study should be interpreted that DNA methylation was diminished to a level below the detection limit rather than complete depletion of 5mCs. Although the 5mC profiles of PGCLCs observed in the present study were indistinguishable

from the profile of E12.5 embryonic PGCs, it remains to be determined whether weak DNA methylation in PGCLCs could be similar to earlier stage of PGCs.

Erasure of DNA Methylation in PGCLCs and PGCs. The DNA methylomes of PSCs, EpiLCs, PGCLCs, and E12.5 PGCs using MBD-seq (Fig. 2) largely agreed with the gDNA demethylation dynamics in mouse embryonic germline cells determined by Seisenberger et al. using WGBS (9), reproducing significant retention of 5mCs at IAPs or nonpromoter CpG islands (CGIs) in PGCLCs and PGCs (Fig. 2A and Fig. S7A). MBD-seq also detected germline retention of 5mCs at repeat sequences GSAT_MM, LSU_rRNA_Hsa, and SSU_rRNA_Hsa (Fig. 3A–C and Figs. S5 and S7C, E, G, and H). Reanalysis of the WGBS data of Seisenberger et al. validated germline retention of 5mCs at GSAT_MMs (Fig. 3D and Fig. S7D and F) as well as IAPs (Fig. S7B) although 5mC retention at other repeat elements was not validated due to insufficient bisulfite conversion of the WGBS data.

The importance of 5hmeCs in the active DNA demethylation and imprinting erasure in germline cells has been well recognized (12–14, 30, 31). In our present study, the abundant 5hmeCs in mouse iPSCs were dramatically lost during differentiation to PGCLCs via EpiLCs (Figs. 2B and C and 3A). The 5mC content in PGCLCs and E12.5 intragonadal PGCs detected with the sensitivity of MBD-seq was largely comparable. However, PGCLCs retained about a four times greater number but relatively weak 5hmeC-enriched gDNA segments compared to PGCs (Figs. 2 and 3A and Fig. S5). Interestingly, 5mC-enriched gDNA fragments detected in the genomes of PGCs and PGCLCs were often coenriched with 5hmeCs (Fig. 3C and Fig. S7A, C, E, G, and H). Genomic DNA regions strongly enriched with 5mCs in PSCs were typically enriched with 5hmeCs as well, and these 5hmeCs were often retained after differentiation to PGCLCs even when 5mCs were erased (Fig. S8H, orange shading). However, ICR of the *Kncq1* imprinting cluster (KvDMR1) was strongly methylated in ESCs without coenrichment of 5hmeCs (Fig. S8H, a and c) whereas its 5hmeC content was augmented in PGCLCs and 5mCs were lost (Fig. S8H, b and d). In contrast, in iPSCs and EpiLCs, the normal ICRs of the *Dlk1-Gtl2-Dio3* imprinting cluster (IG-DMR and Gtl2-DMR) were significantly enriched with 5hmeCs whereas aberrantly hypermethylated ICRs were deficient in 5hmeCs (Fig. 4A and Fig. S9). Taken together, these observations suggest that gDNA regions coenriched with 5mCs and 5hmeCs may be prone to demethylation, including 5mC-retaining regions in PGCLCs/PGCs.

Germline Epigenetic Erasure as a Barrier to Nongenetic Transgenerational Inheritance. It has been proposed that a small fraction of genomic elements that escape the epigenetic erasure (such as IAPs or nonpromoter CGIs) may serve as vehicles of the transgenerational epigenetic inheritance (2, 9). However, a systematic examination recently reported by Iqbal et al. showed that transcriptional and DNA methylome aberrations introduced in spermatogonia of fetuses by in utero exposure to endocrine-disrupting chemicals were not persistent beyond the germline epigenetic erasure in a statistically significant manner even when the analysis was extended to IAPs (6). This negative but insightful observation may suggest the ability of PGCs to effectively repair epimutations or perhaps reflect technical challenges of identifying transgenerational epimutations that might occur stochastically within repetitive sequences. Taking advantage of the experimentally reproducible DNA hypermethylation at the otherwise demethylated maternal IG-DMR and Gtl2-DMR of the *Dlk1-Gtl2-Dio3* imprinting cluster in mouse iPSCs (26, 27), our present study directly demonstrates significant reduction in this abnormal hypermethylation during iPSC differentiation to PGCLC (Fig. 4A and Fig. S9), which resulted in functional restoration of the Gtl2 imprinted mRNA expression (Fig. 4B). The ability of the PGCLC cell culture model to erase experimentally introduced epimutations will provide unique future opportunities to examine erasure, and possible retention, of various types of epimutations at specific gDNA locations during germline

differentiation. It remains to be determined whether this PGCLC model can also be used to examine erasure of epimutations introduced outside ICRs and/or within repetitive elements, and the resolution power of this approach should be improved at the nucleotide base level because experience-induced changes in gametic gDNA methylation were reported to be specific to CpG sites, thus critically affecting gDNA binding to transcription factors (32, 33). It is also an interesting question as to whether or not apparently physiological epigenetic changes resulting from specific and regulated mechanisms (vs. stochastic, nonphysiological epimutations) are erased in the PGCLCs. The development of epigenome editing methods to introduce specific epimutations at targeted loci in the genome of iPSCs should provide unique opportunities to systematically evaluate the capabilities of PGCLCs to erase various types and locations of epigenetic changes or epimutations.

In summary, our present study has shown that mouse PGCLCs effectively recapitulate the genome-wide DNA demethylation events occurring in the intragonadal PGCs, including demethylation of ICRs. Reproducing previously reported 5mC retention at IAPs and nonpromoter CGIs in PGCs, we have identified additional 5mC-retaining genomic elements, including the GSAT MM repeats. Deep-sequencing techniques that distinguish 5mCs and 5hmCs have revealed core retention and dynamics of these epigenetic marks at ICRs and 5mC-retaining

elements during PSC differentiation to PGCLCs. Finally, taking advantage of a region-specific epimutation experimentally introduced in iPSCs, our study has provided direct evidence that aberrant DNA hypermethylation at an ICR was diminished during the germline epigenetic reprogramming, resulting in functional restoration of the epigenetically silenced gene expression. These observations support the usefulness of mouse PGCLCs as a valuable cell culture model of embryonic PGCs for mechanistic studies of germline epigenetic reprogramming.

Materials and Methods

Experimental methods are described in *SI Materials and Methods*. The animal experiment protocol for the above procedures was reviewed and approved by the Institutional Animal Care and Use Committee of the Massachusetts General Hospital. The animal experiment protocol for the PGCLC transplantation was reviewed and approved by the Institutional Animal Care and Use Committee of the McGill University. Affymetrix microarray and deep-sequencing data have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus and Sequence Read Archive databases (accession nos. GSE80983 and GSE81175).

ACKNOWLEDGMENTS. We thank Haley Ellis and Shiomi (Misa) Yawata for technical assistance. This work was supported by Canadian Institutes of Health Research Grant MOP-130467 (to M.N.) and National Institutes of Health Grants HD058013 (to K.H.) and ES023316 and ES024861 (to T.S.).

1. Prokopyuk L, Western PS, Stringer JM (2015) Transgenerational epigenetic inheritance: Adaptation through the germline epigenome? *Epigenomics* 7(5):829–846.
2. Szyf M (2015) Nongenetic inheritance and transgenerational epigenetics. *Trends Mol Med* 21(2):134–144.
3. Alvarado S, et al. (2015) An epigenetic hypothesis for the genomic memory of pain. *Front Cell Neurosci* 9:88.
4. Nilsson EE, Skinner MK (2015) Environmentally induced epigenetic transgenerational inheritance of reproductive disease. *Biol Reprod* 93(6):145.
5. Bohacek J, Mansuy IM (2015) Molecular insights into transgenerational non-genetic inheritance of acquired behaviours. *Nat Rev Genet* 16(11):641–652.
6. Iqbal K, et al. (2015) Deleterious effects of endocrine disruptors are corrected in the mammalian germline by epigenome reprogramming. *Genome Biol* 16(1):59.
7. Heard E, Martienssen RA (2014) Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell* 157(1):95–109.
8. Saitou M, Yamaji M (2012) Primordial germ cells in mice. *Cold Spring Harb Perspect Biol* 4(11):a008375.
9. Seisenberger S, et al. (2012) The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol Cell* 48(6):849–862.
10. Ohno R, et al. (2013) A replication-dependent passive mechanism modulates DNA demethylation in mouse primordial germ cells. *Development* 140(14):2892–2903.
11. Kagiwada S, Kurimoto K, Hirota T, Yamaji M, Saitou M (2013) Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J* 32(3):340–353.
12. Hackett JA, Zyliz JJ, Surani MA (2012) Parallel mechanisms of epigenetic reprogramming in the germline. *Trends Genet* 28(4):164–174.
13. Kawasaki Y, et al. (2014) Active DNA demethylation is required for complete imprint erasure in primordial germ cells. *Sci Rep* 4:3658.
14. Yamaguchi S, Shen L, Liu Y, Sandler D, Zhang Y (2013) Role of Tet1 in erasure of genomic imprinting. *Nature* 504(7480):460–464.
15. Guibert S, Forné T, Weber M (2012) Global profiling of DNA methylation erasure in mouse primordial germ cells. *Genome Res* 22(4):633–641.
16. Ge W, Chen C, De Felici M, Shen W (2015) In vitro differentiation of germ cells from stem cells: A comparison between primordial germ cells and in vitro derived primordial germ cell-like cells. *Cell Death Dis* 6(10):e1906.
17. Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M (2011) Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* 146(4):519–532.
18. Nakaki F, et al. (2013) Induction of mouse germ-cell fate by transcription factors in vitro. *Nature* 501(7466):222–226.
19. Szabó PE, Hübner K, Schöler H, Mann JR (2002) Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. *Mech Dev* 115(1-2):157–160.
20. Dann CT, et al. (2008) Spermatogonial stem cell self-renewal requires OCT4, a factor downregulated during retinoic acid-induced differentiation. *Stem Cells* 26(11):2928–2937.
21. Harris RA, et al. (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat Biotechnol* 28(10):1097–1105.
22. Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N (2012) MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* 151(7):1417–1430.
23. Song C-X, et al. (2011) Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol* 29(1):68–72.
24. Huang Y, et al. (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* 5(1):e8888.
25. Zhou Q, et al. (2016) Complete meiosis from embryonic stem cell-derived germ cells in vitro. *Cell Stem Cell* 18(3):330–340.
26. Stadtfeld M, et al. (2010) Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* 465(7295):175–181.
27. Stadtfeld M, et al. (2012) Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. *Nat Genet* 44(4):398–405.
28. Singh P, et al. (2013) De novo DNA methylation in the male germ line occurs by default but is excluded at sites of H3K4 methylation. *Cell Reports* 4(1):205–219.
29. Rauch T, Li H, Wu X, Pfeifer GP (2006) MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells. *Cancer Res* 66(16):7939–7947.
30. Hackett JA, et al. (2013) Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* 339(6118):448–452.
31. Yamaguchi S, et al. (2013) Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. *Cell Res* 23(3):329–339.
32. Szyf M, Tang Y-Y, Hill KG, Musci R (2016) The dynamic epigenome and its implications for behavioral interventions: A role for epigenetics to inform disorder prevention and health promotion. *Transl Behav Med* 6(1):55–62.
33. Weaver ICG, et al. (2004) Epigenetic programming by maternal behavior. *Nat Neurosci* 7(8):847–854.
34. Hayashi K, Saitou M (2013) Generation of eggs from mouse embryonic stem cells and induced pluripotent stem cells. *Nat Protoc* 8(8):1513–1524.
35. Furuchi T, Masuko K, Nishimune Y, Obinata M, Matsui Y (1996) Inhibition of testicular germ cell apoptosis and differentiation in mice misexpressing Bcl-2 in spermatogonia. *Development* 122(6):1703–1709.
36. Zohni K, Zhang X, Tan SL, Chan P, Nagano MC (2012) The efficiency of male fertility restoration is dependent on the recovery kinetics of spermatogonial stem cells after cytotoxic treatment with busulfan in mice. *Hum Reprod* 27(1):44–53.
37. Brind'Amour J, et al. (2015) An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat Commun* 6:6033.
38. Lienhard M, Grimm C, Morkel M, Herwig R, Chavez L (2014) MEDIPS: Genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. *Bioinformatics* 30(2):284–286.