Correction

DEVELOPMENTAL BIOLOGY

The authors note that the name Junko Odahima should have appeared as Junko Odajima. The corrected author line appears below. The online version has been corrected.

Norikatsu Miyoshi, Jente M. Stel, Keiko Shioda, Na Qu, Junko Odajima, Shino Mitsunaga, Xiangfan Zhang, Makoto Nagano, Konrad Hochedlinger, Kurt J. Isselbacher, and Toshi Shioda

www.pnas.org/cgi/doi/10.1073/pnas.1613505113
Erasure of DNA methylation, genomic imprints, and epimutations in a primordial germ-cell model derived from mouse pluripotent stem cells

Norikatsu Miyoshi,1,2 Jente M. Stel,1a,3 Keiko Shioda,4 Na Qu,5 Junko Odajima,6 Shino Mitsunaga,6 Xiangfan Zhang,5 Makoto Nagano,4 Konrad Hochdelinger,4,6,d Kurt J. Isselbacher,2,a and Toshi Shioda,a,2

1Massachusetts General Hospital Center for Cancer Research, Charlestown, MA 02129; 2Institute for Environmental Studies Vrije Universiteit, Amsterdam 1081 HV, The Netherlands; 3Department of Obstetrics and Gynecology, McGill University and Research Institute of McGill University Health Centre, Montreal, QC, Canada H4A 3J1; and 4Department 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 5mCs were also significantly diminished in PGCLCs, they retained greater amounts of 5mCs than intragonadal PGCs. The genomes of both PGCLCs and PGCs selectively retained both 5mCs and 5hmc at a small number of repeat sequences such as SAT-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 Gtll2 gene. These observations support the usefulness of PGCLCs in studying the genome-wide epigenetic erasure including imprinted genes, epimutations, and erasure-resistant loci, which may be involved in transgenerational epigenetic inheritance.


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).

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 germ line 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 genome 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.

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

PNAS | August 23, 2016 | vol. 113 | no. 34 | 9545–9550

www.pnas.org/cgi/doi/10.1073/pnas.1610259113

*Corresponding author.
†To whom correspondence may be addressed. Email: kisselbacher@mgh.harvard.edu or tshioda@mgh.harvard.edu.

1N.M. and J.M.S. contributed equally to this work.
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). The three gene clusters indicated in the Top heatmap are enlarged in the Bottom heatmaps. (b) PCA of transcriptional changes during differentiation of PSCs to PGCLCs via EpiLCs.

**Results**

The SSEA1\(^{+}\)Integrin \(\beta^{3}\)/c-Kit\(^{+}\) Triple-Positive Mouse PGCs Resemble Early Stage PGCs in Marker mRNA Expression. Mouse E12.5 intragonadal PGCs characterized by germine-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 \(\beta^{3}\)/c-Kit\(^{+}\) triple-positive status (Fig. S1 C 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 \(\beta^{3}\) double-positive cells also strongly expressed c-Kit (Fig. SIF, Top row) whereas only 36% of SSEA1\(^{+}\)/c-Kit\(^{+}\) double-positive cells were Integrin \(\beta^{3}\) positive (Fig. SIF, Bottom row). In the present study, the SSEA1\(^{+}\)/Integrin \(\beta^{3}\) 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 spermatogenic stem cells (20) or mCherry expressed by the human EF1\(\alpha\) 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 transriptomes 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 homogenous upon differentiation to EpiLCs, but diverged 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 genes were 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 PGC.

Expression of Fgf5 [an early stage EpiLC maker (17)] was strong in EpiLCs but reduced in PGCLCs whereas expression of Wnt3 [an intermediate 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 Tnfpc2, 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 Snai1 was transiently expressed during EpiLC differentiation to PGCLCs but later suppressed when intragonadal PGC markers were induced (17), our PGCLCs expressed Snai1 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 (Fig. S2F) whereas intragonadal PGC markers (Dazl or Ddx4, Fig. S2G) did not express the Uhrf1/Np95 cofactor of Dnmt1. Agreed 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 transcriptional 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 5meCs and 5hmeCs in PGCLCs.** To examine the epigenetic status of PGCLCs, we determined distributions of 5meCs and 5hmecs in the genomes of mouse iPSCs, EpiLCs, PGCLCs, and E12.5 intragonadal PGCs by deep sequencing of gDNA fragments enriched for 5meCs using biotin-conjugated methylcytosine-binding protein 2 [MBD-Seq. (21)], which has no significant affinity to 5hmecs (22), and gDNA fragments enriched for 5hmecs by...
of PGCLCs and PGCs show significant similarities in both relative distributions across genomic features and total numbers of the 5meC-enriched regions (2,178 in PGCLCs vs. 2,791 in PGCs), and 91% of the 5meC-enriched regions detected in PGCLCs were also found in PGCs (Fig. S6). The 5meC enrichment profiles of iPSCs and EpiLcs were similar to 5meCs except that only 251 5hmC-enriched regions (assigned mostly to repetitive elements) were found in PGCs (Fig. 3A and Fig. S5A). The majority of the repeat-containing, 5meC-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 5meC-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 5meC 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 5meCs and 5hmCs at a region containing IAPs in PGCLCs and PGCs. Fig. S7B shows the WGBS data corresponding to a part of the IAP-related 5meC/5hmC-enriched region indicated in Fig. S7A, demonstrating significant retention of bisulfite-resistant cytosines (i.e., the sum of 5meCs and 5hmCs) at two CpG sites in the gDNA of E13.5 male PGCs. Fig. 3 C and D shows similar analyses for a region rich in GSAT_MM and SYNREP_MM repeats. Although some 5meC/5hmC peaks in the deep-sequencing tracks were not informative, as they were also evident in the nonenriched mouse genome resequencing track (peak c), several informative peaks (a, b, c) supported the presence of 5meC- and 5hmC-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 5meC/5hmC peaks at SYNREP_MM in Fig. 3C (peak c on the nonenriched track) left the 5meC/5hmC retention in this element possibly uninformative, consistent with the issues stemming from its up to 75% nucleotide base identity to GSAT MM. Two additional examples of GSAT_MM retention of 5meC/5hmC peaks and bisulfite-resistant cytosines are shown in Fig. S7 C-F. Agreeing with the retention of 5meCs and 5hmCs 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. S7 G 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 5meC peaks to be closely associated with 5hmC peaks (Fig. 3C and Fig. S7 A, C, E, G, and H) although the enrichment-based deep-sequencing approach did not provide relative amounts of 5hmCs to 5meCs.

DNA Demethylation at the ICRs in PGCLCs. Demethylation of the ICRs is a hallmark of intragenadal 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 intragenadal PGCs (25). For all of the six ICRs examined, our deep-sequencing analysis showed progressive loss of 5meCs upon iPSC differentiation to EpiLcs and then to PGCLCs (Fig. 4A and Fig. S8 A–E). Expression of the
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. S2f). 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, 5mCcs and 5hmCcs were retained in the genomes of all types of PGCLCs as well as E12.5 embryonic PGCs (Fig. S8g). Note that no 5mCc or 5hmCc 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. S8f).

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-Gt12-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 Gt12-DMR (26, 27). Taking advantage of this epimutation that is experimentally inducible in iPSC, we examined whether aberrant region-specific DNA methylation 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 Gt12-DMR in mouse iPSCs (Fig. 4a and Fig. S9a). The accuracy of our 5mCc profiling is supported by the nearly identical MBD-seq tracks of normal [Gt12(+)] and silenced [Gt12(−)] iPSCs except for the IG- and Gt12-DMRs. Whereas these aberrant 5mCc peaks were still observed in EpiLCs, they were not detected in PGCLCs. Concomitantly, Gt12 mRNA expression, which was suppressed in Gt12(−) iPSCs, was restored in PGCLCs to a level comparable to PGCLCs derived from Gt12(+) iPSCs (Fig. 4b). Interestingly, the IG-DMR and the region between the IG- and the Gt12-DMRs of Gt12(−) iPSCs showed aberrant reduction in 5hmCc 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-Gt12-Dio3 imprinting cluster in iPSCs was erased upon differentiation to PGCLCs.

Discussion

Transcriptional and Epigenomic Characteristics of Mouse PGCLCs. Following the protocol described by Hayashi et al. (17) with slight modifications, we generated SSEA1^Integrin β3^/c-Kit^ triple-positive PGCLCs from mouse PSCs (Fig. S1). Transcriptional 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, Supn, H19, and Kenp1 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 Supn 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]Gt12-Dio3, H19, Igf2r, Kenp1, Nespas-Gnas, Meg1) (Fig. 4 and Fig. S8) as well as global loss of 5mCcs (Figs. 2a and C and 3d). The progressive increase in mRNA expression of imprinted genes during PSC differentiation to PGCLC via EpiLC (Fig. S2f) may reflect release from monoallelic suppression by DNA methylation. The restoration of Gt12 mRNA expression in PGCLCs derived from Gt12(−) iPSCs

---

**Fig. 3.** Genomic feature distributions of 5meCs and 5hmCcs in the genomic DNA of mouse iPSCs, EpiLCs, PGCLCs, and in vivo PGCs. (A) 5meC and 5hmC distributions across genomic features. (B) Distributions of 5meCs across repeat sequences. RepM, genome-wide RepeatMasker-masked elements. Small elements (≤5%) are left blank in pie charts. *Sa, satellite repeats; **GS, GSAT_MM; **SY, SYNRPEP_MM. Other keys of pie charts are defined in Fig. 5. (C) An example of deep-sequencing tracks showing 5meC and 5hmCc peaks at GSAT_MM and SYNRPEP_MM satellite repeats. Height of peaks reflects relative strength of DNA methylation across the four 5meC tracks (linearly scaled 0–1 between the baseline and the maximal methylation, red bar), DNA hydroxymethylation (four 5hmCc tracks, green bar), or nonenriched genome regions (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 nonenriched mouse genome sequence tracks or changes between different cell types. 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 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, Supn, H19, and Kenp1 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 Supn 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]Gt12-Dio3, H19, Igf2r, Kenp1, Nespas-Gnas, Meg1) (Fig. 4 and Fig. S8) as well as global loss of 5mCcs (Figs. 2a and C and 3d). The progressive increase in mRNA expression of imprinted genes during PSC differentiation to PGCLC via EpiLC (Fig. S2f) may reflect release from monoallelic suppression by DNA methylation. The restoration of Gt12 mRNA expression in PGCLCs derived from Gt12(−) iPSCs
from silencing due to the aberrant hypermethylation of the ICR of the Dlk1-Gtl2-Dio3 imprinting cluster [Fig. 4 and Fig. S9] indicating 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 ± SEM).

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

**Erasure of DNA Methylation in PGCLCs and PGCs.** The DNA methy- lomes of PSCs, EpiLCs, PGCLCs, and E12.5 PGCs using MBD-seq (Fig. 2) largely agreed with the gDNA methylation dynamics in mouse embryonic germline cells determined by Seisenberger et al. using WGBS (9), reproducing significant retention of 5meCs at IAPs or nonpromoter CpG islands (CGIs) in PGCLCs and PGCs (Fig. 24 and Fig. S7A). MBD-seq also detected germline retention of 5meCs at repeat sequences GSAT_MM, LSU_rRNA_Hsa, and SSU_rRNA_Hsa (Fig. 3A–C and Figs. S5 and S7 C, E, G, and H). Recalibration of the WGBS data of Seisenberger et al. validated germline retention of 5meCs at GSAT_MMs (Fig. 3D and Fig. S7 D and F) as well as IAPs (Fig. S7B) although 5meC retention at other repeat elements was not validated due to insufficient bisulfite conversion of the WGBS data.

The importance of 5meCs 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 5meCs in mouse iPSCs were dramatically lost during differentiation to PGCLCs via EpiLCs (Figs. 2B and C and 3A). The 5meC 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 5meC-enriched gDNA segments compared to PGCs (Figs. 2 and 3A and Fig. S5). Interestingly, 5meC-enriched gDNA fragments detected in the genomes of PGCs and PGCLCs were often coenriched with 5meCs (Fig. 3C and Fig. S7 A, C, E, G, and H). Genomic DNA regions strongly enriched with 5meCs in PSCs were typically enriched with 5meCs as well, and these 5meCs were often retained after differentiation to PGCLCs even when 5meCs were erased (Fig. S8H, orange shading). However, ICR of the Kctr1 imprinting cluster (IG-DMR) was strongly methylated in ESCs without coenrichment of 5meCs (Fig. S8H, a and c) whereas its 5meC content was augmented in PGCLCs and 5meCs 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 5meCs whereas aberrantly hypermethylated ICRs were deficient in 5meCs (Fig. 4A and Fig. S9). Taken together, these observations suggest that gDNA regions coenriched with 5meCs and 5meCs may be prone to demethylation, including 5meC-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 methylene 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. 4D 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 induced epimutations will provide unique future opportunities to examine erasure, and possible retention, of various types of epimutations at specific gDNA locations during germline

**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 5meCs (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 ± SEM).
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 nucleosome base level because experience-induced changes in genomic 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 5meC retention at IAPs and nonpromoter CGIs in PGCs, we have identified additional 5meC-retaining genomic elements, including the GSAT/MM repeats. Deep-sequencing techniques that distinguish 5meCs and 5hmeCs have revealed coretention and dynamics of these epigenetic marks at ICRs and 5meC-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 HD098013 (to K.H.) and E002316 and E0024861 (to T.S.).
Supporting Information

Miyoshi et al. 10.1073/pnas.1610259113

SI Materials and Methods

Cell Culture. Transcriptional and epigenomic characteristics as well as pluriptotency of ESCs and iPSCs [Gtli2(+) and Gtli2(–)] were described in our previous study (26). These PSCs [v6.5, ES12, ES15 (ESCs); Ptf1159, Ptf1159, R21-2, R21-4, R21-5, and R21-6 (iPSCs; asterisks indicate Gtli2-negative clones)] were maintained in ESC medium [DMEM supplemented with 15% EScult-grade FCS and 1,000 U/mL recombinant mouse leukemia inhibitory factor (Stemcell Technologies), 2 mM l-glutamine, 1× penicillin–streptomycin, and 1× nonessential amino acids (Gibco), 40 nM 2-mercaptoethanol, and 50 μg/mL l-ascorbic acid (Sigma)]. Except for the ESC clone v6.5, which was provided by N. Bardeesy, Massachusetts General Hospital Center for Cancer Research, Boston, all PSCs examined in the present study shared the same genetic background of the male reprogrammable mouse strain harboring an inducible OKSM expression cassette [Jackson Laboratory stock no. 011001, B6;129S4-Cellشاركت in the FACSAria II cells or sorted SSEA1, anti c-Kit, and/or anti-Integrin β3 antibodies for FACS enrichment of SSEA1+/c-Kit+/Integrin β3+ PGCs.

Differentiation of Mouse PSCs to EpilCs and PGCLCs. EpilCs and PGCLCs were generated from PSCs following the protocol described by Hayashi et al. [Fig. S1L (17, 34)] with modifications. PSCs [v6.5, ES12, R21-2, R21-4, R21-5, and R21-6] were differentiated to EpilCs on fibronectin-coated dishes in modified N2B27 medium containing 40% (vol/vol) DMEM/F12 medium, 40% (vol/vol) Neurobasal medium, 0.8% B27 supplement, and 1% KSR (Thermo Fisher); 20 ng/mL activin A, and 12 ng/mL FGF2 (R&D Systems); 1× insulin–transferrin–selenium mix, 1 mM l-glutamine, and 1× penicillin–streptomycin (Gibco); and 8 μg/mL progesterone, 6 μg/mL putrescine, 40 nM 2-mercaptoethanol, 50 μg/mL l-ascorbic acid, and 20 μg/mL BSA (Sigma). Medium was changed once after 24 h of incubation, and EpilCs were harvested after 48 h of induction. EpilC cultures (derived from PSC clones v6.5, ES12, R21-2, R21-4, R21-5, and R21-6) were dissociated to single cells by TrypLE Express, and 30 million cells were casted into the Aggrewell 400 microwell plates (Stemcell Technologies) in PGCLC induction medium [Glasgow’s MEM supplemented with 15% KSR, 500 ng/mL recombinant human BMP4, 100 ng/mL stem cell factor, 50 ng/mL epidermal growth factor (R&D Systems), 1,000 U/mL recombinant mouse leukemia inhibitory factor, 1× nonessential amino acids, 2 mM l-glutamine, 50 μg/mL l-ascorbic acid, 100 nM 2-mercaptoethanol, and 1× penicillin–streptomycin] to form EBs at the density of 3,000 cells per microwell and 9,600 microwells per plate. After EBs were maintained in the PGCLC medium under a floating condition for 6 d, PGCLCs were isolated from TrypLE Express-dissociated cell strainer-filtered EB suspension. The transgenic mouse B6;129S4–Lleukemia inhibitory factor (Stemcell Technologies), 2 mM l-glutamine, and R21-6 (iPSCs; asterisks indicate Gtli2-negative clones)] were maintained in ESC medium [DMEM supplemented with 15% EScult-grade FCS and 1,000 U/mL recombinant mouse leukemia inhibitory factor (Stemcell Technologies), 2 mM l-glutamine, 1× penicillin–streptomycin, and 1× nonessential amino acids (Gibco), 40 nM 2-mercaptoethanol, and 50 μg/mL l-ascorbic acid (Sigma)]. Except for the ESC clone v6.5, which was provided by N. Bardeesy, Massachusetts General Hospital Center for Cancer Research, Boston, all PSCs examined in the present study shared the same genetic background of the male reprogrammable mouse strain harboring an inducible OKSM expression cassette [Jackson Laboratory stock no. 011001, B6;129S4–Cellشاركت in the FACSAria II cells or sorted SSEA1, anti c-Kit, and/or anti-Integrin β3 antibodies for FACS enrichment of SSEA1+/c-Kit+/Integrin β3+ PGCs. The typical yield of live PGCs after FACS was ∼50,000 per batch, each of which consisted of 6–10 embryos. Batches with lower PGC yields (<30,000) were excluded from the study.

Transplantation of PGCLCs into Seminiferous Tubules of Neonatal Mice. PGCLCs were generated from ESCs and iPSCs expressing EGFP driven by the Pou5f1 promoter/distal enhancer (19) or the mCherry red fluorescence protein driven by a truncated human EF1α promoter [introduced by the lentiviral vector plVX-EF1α-mCherry-C1 (Clontech)], Pou5f1 is expressed in mouse spermatogonial stem cells (20), and the truncated human EF1α promoter is selectively active in mouse germline cells (35). The fluorescence protein-labeled PGCLCs were transplanted into testes of 129/SvEv x C57BL/6 F1 hybrid male pups the endogenous germline cells of which were depleted with exposure to busulfan at 8 days post partum (dpp) (5.4 × 10^5 PGCLCs per testis) using the technique described in our previous study (36). The recipient animals were returned to their littersmates after surgery, and PGCLC colonization was evaluated using fluorescence microscope examination of dissected seminiferous tubules 4–5 mo after injection.

Transcriptional Analysis. Total RNA and genomic DNA were simultaneously prepared from the same batch of cells using the QIAGEN AllPrep kits (miniscale kit for PSCs and EpilCs, and microscale kit for FACS-enriched PGCLCs and PGCs). RNA concentration and integrity were determined using Nanodrop (Thermo Fisher) or Agilent Bioanalyzer or Tapestation. Low-quality RNA (RNA integrity number < 9.0) was eliminated from the study.

Expression of specific mRNA transcripts was determined by real-time qPCR using the SuperScript double-stranded cDNA synthesis kit and TaqMan assays with Universal Master Mix II containing Uracll-N-glycosylase (Thermo Fisher). qPCR amplification and data collection were performed using ABI7500 qPCR equipment (Thermo Fisher) and normalized with Gapdh mRNA TaqMan assay. The TaqMan Assay targets and probe IDs were as follows: Pou5f1, Mm0053917_g1; Sox2, Mm03053810_s1; Nanog, Mm0219550_s1; Klf4, Mm00516104_m1; Myc, Mm00487804_m1; Pdml, Mm00476128_m1; Prdm14, Mm01237814_m1; Wnt3, Mm00437336_m1; Dppa3, Mm01184918_g1; Dmnt1, Mm01150563_m1; Dmnt3a, Mm00432881_m1; Dmnt3b, Mm01240113_m1; Gapdh, 99999915_g1. Quantification of the mouse Gtli2 mRNA transcripts was described in our previous study (26).

Transcriptional profiling and data analysis were performed as we previously described (26). Transcriptome data were obtained using Affymetrix GeneChip microarray mouse 430 2.0 chips. Hybridization probes were synthesized using NutGEN Applause
Deep-Sequencing Determination of 5mCcs and 5hmCcs. Genome-wide gDNA methylation and hydroxymethylation profiles of mouse iPSCs, EpiLCs, PGCLCs, and E12.5 embryonic PGCs were determined by deep sequencing of 5mC- or 5hmC-enriched DNA fragments. Embryonic DNA was sonicated to 200–300 bp fragments using a Covaris S2 sonicator in AFA microtubes (10–200 ng DNA in 80 μL 0.1x Tris/HCl-EDTA buffer) at 4 °C, and DNA size distribution was determined using Bioanalyzer or Tapestation.

DNA fragments were subjected to enrichment for 5mCcs and 5hmCcs using biotin-conjugated methyl-CpG-binding domain (MBD2) of the recombinant human methyl-CpG-binding domain protein 2 (MBD2) [MBD-seq (21)] and selective labeling of 5hmCcs with a biotinylated tag by two-step chemical reactions involving β-glucosyltransferase (23), respectively. Reagents and protocols for enrichment of 5mCcs and 5hmCcs-rich gDNA fragments were provided in the MethylMiner kit (Thermo Fisher) or the Hydroxymethyl Collector kit (Active Motif), respectively. To avoid sensitivity bias of enrichment reactions resulting from different amounts of input DNA, an equal amount (10 ng) of fragmented DNA was put into each enrichment reaction for both 5mCcs and 5hmCcs for all cell types, and eluted DNA from multiple (typically 8–24) 10-ng scale reactions was pooled for ethanol precipitation to yield at least 0.5 ng 5mC-enriched gDNA fragments. For both the 5mC and 5hmC enrichment kits, the manufacturer-recommended minimum amount of input DNA was 5 ng. Ten-nanogram input DNA has been shown to have sufficient molecular complexity for genome-wide Chip-seq profiling of the mouse genome for common histone modifications (24). Successful enrichment was monitored by simultaneously performing positive control reactions using kit-provided reagents.

The 5mC-rich DNA fragments were eluted in 200 μL per reaction with 2 M NaCl and then desalted/concentrated to 20 μL by ethanol precipitation with Pellet Paint coprecipitant fluorescent dye polymer (Millipore). The 5hmC-rich DNA fragments were eluted in 50 μL per reaction elution buffer provided in the kit and subjected to ethanol precipitation.

The 5mC- or 5hmC-enriched gDNA fragments were subjected to deep-sequencing library construction using the SOLiD 5500 Fragment Library Core Kit and the SOLiD EZ Bead emulsion PCR system (Thermo Fisher). Deep sequencing was performed using the SOLiD 5500XL deep sequencer (50 + 50 nt, paired-end), and the XSQ-format raw data were converted to the csfasta/QV.qual format using the XSQ tool provided by Life Technologies. The csfasta/QV.qual sequences were aligned to the NCBI37/mm9 mouse genome reference sequence to obtain the robust multiarray average (RMA) normalization were performed with special precautions because of their far smaller numbers of the 5mC or 5hmC peaks in the whole genome compared with PSCs or EpiLCs. To this end, results of analyses using MEDIPS were verified using the MACS deep-sequencing peak detection tool, which was originally developed for analysis of Chip-seq data and adopted for the ENCODE project of Chip-seq Roadmap database construction and manual inspection of the visualized 5mC and 5hmC peaks using the Integrated Genomics Viewer after conversion to normalized wig or bigwig data format. The read peaks shown in the figures are normalized tdf file images unless specified otherwise. Peak detection cutoff was set as 10 times of the mean of all of the 100-bp windows assigned to the whole-genome sequence. Association of the detected peaks to known genomic features or RepeatMasker families/classes/functions was performed using MEDIPS functions, the ChiPseeker R/Bioconductor package, and the Ruby scripts developed in our laboratory. The aligned sequence data (bam format) of 5mC- or 5hmC-enriched segments in the genomes of each cell type (i.e., iPSCs, EpiLCs, PGCLCs, and PGCs) were merged using samtools and subjected to differential (hydroxymethyl)analysis using the MEDIPS.meth function. An aligned, whole-genome resequencing data file (bam format) for the genome of the C57BL/6 mouse was provided to the MEDIPS.meth function as no-enrichment input DNA. The locations of DNA (hydroxymethyl)methylation peaks determined by MEDIPS were compared with the mouse mm9 RepeatMasker database-registered locations of repeat sequences using Ruby scripts.

Reanalysis of WGBS Data. The fastq-format Illumina deep-sequencing raw data of mouse WGBS deposited by Seisenberger et al. (9) were downloaded from the European Nucleotide Archive (ERX167039 and ERX167050 for E6.5 epiblasts (total 350 million reads) and ERX167042 for E13.5 male PGCs (244 million reads)) and subjected to alignment using the Bismark bisulfite mapper v0.14.4 with Bowtie2 as the backend aligner. The resulting bam files were subjected directly to manual inspection using the Integrative Genomics Viewer for counting bisulfite-converted and -resistant cytosines. To detect bisulfite-resistant CpG cytosines in the areas of interest determined by the MBD-seq analysis, we set the following criteria:
1. For simplicity, we examined only the Watson strand of the double-stranded genomic DNA.

2. Read coverage at a region of interest (~150-bp window) is sufficient (greater than 30× at each base).

3. Non-CpG cytosines (i.e., CpA, CpC, and CpT) within a region of interest are efficiently converted to thymidines by bisulfite (>80%).

4. Non-CpG cytosines in a region of interest show no differences in bisulfite conversion rate between the E6.5 and E13.5 WGBS data unless there is a reasonable indication that they are also methylated or hydroxymethylated.

5. CpG cytosines show significantly greater resistance to bisulfite conversion compared with the surrounding non-CpG cytosines. We expect strong resistance in the E6.5 WGBS data, and the question is whether these CpG cytosines are also significantly resistant to bisulfite conversion.

Statistical significance of the bisulfite-resistant CpG cytosines over the background of the bisulfite-sensitive CpA, CpT, and CpC cytosines in the genome of E13.5 maple PGCs was examined using Student's t test.

---

**Fig. S1.** Characterization of PGCs and PGCLCs. (A–D) PGC isolation from E12.5 fetal male mouse gonads. (A) Microscopic images of PGCs expressing EGFP driven by the Pou5f1 promoter/distal enhancer (Left, phase contrast; Right, fluorescence). (B) Alkaline phosphatase staining of intragonadal PGCs. (C and D) FACS profiling of PGCs for cell-surface expression of SSEA1, Integrin β3, and c-Kit. (C) Expression of Integrin β3 in SSEA1+/c-Kit+ PGCs. (D) Expression of c-Kit in SSEA1+/Integrin β3+ PGCs. (E–G) PGCLC production from mouse PSCs. (E) Overall protocol. PSCs were first differentiated to EpiLCs as adherent cell culture for 48 h and then to PGCLCs in EBs for 6 d. (F) FACS profiling of mouse iPSCs, EpiLCs, and PGCLCs for expression of SSEA1, Integrin β3, and c-Kit. iPSCs express EGFP driven by the germline-active Pou5f1 distal enhancer/promoter. Boxes on the Right show expression of EGFP and c-Kit in Integrin β3+/SSEA1+ PGCLCs (Top) and Integrin β3 expression in c-Kit+/SSEA1+ PGCLCs (Bottom). (G) Colonization of PGCLCs in mouse testes. Mouse PGCLCs were labeled with germline-active human EF1α promoter (mCherry) or Pou5f1 distal enhancer/promoter (Pou5f1ΔPE-EGFP). The white light image shows seminiferous tubules, and the epifluorescence image shows an EGFP-positive segment of the tubules (pointed by arrowheads).
Fig. S2. mRNA expression of marker genes in mouse PSCs, EpiLCs, PGCLCs, and E12.5 in vivo PGCs. Box plots represent median, first and third quartile, and 95% confidence intervals of mRNA expression (normalized intensity values) determined using Affymetrix microarray (n > 3). (A) Epiblast markers. (B–D) PGC markers for (B) early, (C) mid, and (D) late stages. (E) Tet enzymes. (F) DNA methyltransferases and the Uhrf1 cofactor of DNMT1. (G) Pluripotency genes. (H) Enzymes involved in active demethylation of DNA. (I) Imprinted genes.
**Fig. S3.** Microarray and deep-sequencing quality control data. (A) RMA-normalized signal intensities of Affymetrix microarray data shown in Fig. 1. Box plots are shown with the same color codes as in Fig. 1. The evenly distributed signal intensities across samples do not show signs of batch effects. The degrees of transcriptomal heterogeneity shown in the PCA analysis (Fig. 1B) are not directly explained by different intensities of normalized microarray data. (B) Coverage of CpG sites in the mouse mm9 reference genome sequences by deep-sequencing data for DNA methylation and hydroxymethylation (the seqCoverage function of the MEDIPS Bioconductor package). (C) Saturation analysis of deep-sequencing data (the saturation function of MEDIPS). Sufficient CpG coverage and saturation by deep-sequencing reads ensures appropriate interrogation of DNA methylation and hydroxymethylation. Note that strong CpG coverage and saturation do not immediately result in detection of DNA methylation and hydroxymethylation, which are dependent on formation of peaks over the background. The low CpG coverage or saturation of deep-sequencing reads for DNA hydroxymethylation in PGCs was due to the strong depletion of 5hmeCs in this type of cells. (D and E) Numbers of uniquely mapped deep-sequencing reads obtained for different types of cells for (D) DNA methylation and (E) DNA hydroxymethylation.
Fig. S4. Distributions of aligned deep-sequencing quality scores along the read positions for (A) DNA methylation and (B) DNA hydroxymethylation (outputs of the FastQC quality control tool). The dark green areas were automatically truncated by the analytical software.
**A** Genomic Features (number of regions)

<table>
<thead>
<tr>
<th>Feature</th>
<th>iPSC</th>
<th>EpiLC</th>
<th>PGCLC</th>
<th>PGC(E12.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5meC</td>
<td>763</td>
<td>227k</td>
<td>217k</td>
<td>27%</td>
</tr>
<tr>
<td>5hmec</td>
<td>628k</td>
<td>8567</td>
<td>1008</td>
<td>251</td>
</tr>
</tbody>
</table>

**B** All Methylated Regions

<table>
<thead>
<tr>
<th>Class</th>
<th>RepM</th>
<th>PGCLC</th>
<th>PGC</th>
<th>PGC-PGCLC Common</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>RepM (Top 20)</th>
<th>Satellites</th>
<th>PGCLC</th>
<th>PGC</th>
<th>PGC-PGCLC Common</th>
</tr>
</thead>
<tbody>
<tr>
<td>RepM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satellite</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. S5. Distributions of 5meCs and 5hmecs across genomic features in mouse PSCs, PGCLCs, and in vivo PGCs. Numbers of total features are shown at the center of each pie chart. (A) 5meC and 5hmec distributions in all genomic features (left; reflecting numbers of regions only) and in repeat sequences (right; reflecting products of numbers and sizes of regions). (B) Detailed distributions of 5meC in repeat sequences based on numbers of regions. RepM, profiles of the entire RepeatMasker elements. RepeatMasker features are integrated into families, which are further integrated into classes. Some of the small RepeatMasker features (<5%) are shown as blanks without labels.
Fig. S6. Venn diagrams representing overlapping locations of 5meCs and 5hmeCs between mouse PGCLCs and in vivo PGCs at distinct genomic features.
Fig. S7. Examples of 5meC and 5hmeC retention in PGCLCs and PGCs. A, C, E, G, and H show deep-sequencing tracks at (A) an IAP, (C and E) GSAT_MM, (G) LSU_rRNA_Hsa, and (H) SSU_rRNA_Hsa. Informative peaks are indicated by blue arrows. Height of peaks reflects relative strength of DNA methylation across the four 5meC tracks (linearly scaled 0–1 between the baseline and the maximal methylation in the displayed area indicated by a red vertical bar at the right), DNA hydroxymethylation (four 5hmeC tracks, green vertical bar), or nonenriched genome resequencing coverage (nonenriched track, blue vertical bar); note that scaled value 1 is not equal to 100% methylation. (B, D, and F) Reanalysis of the whole-genome bisulfite sequencing data generated by Seisenberger et al. (9) for E6.5 mouse epiblasts and E13.5 male PGCs corresponding to sequences shown in A, C, and E, respectively. Yellow shading 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).
Fig. S8. Demethylation at the ICRs in PGCLCs. Deep-sequencing tracks of 5meCs at the ICRs in the genomes of Gtl2(+) and Gtl2(-) mouse iPS cells and their EpiLC and PGCLC derivatives are shown. (A) Igf2-H19, (B) Igf2r, (C) Kcnq1, (D) Gnas, and (E) Meg1/Grb10. (F) 5meC deep-sequencing tracks of the individual cell clones at the H19, Igf2r, and Gtl2/Meg3 ICRs (shaded in red). Asterisks indicate the Gtl2(+) iPS clones, which showed abnormal hypomethylation of the Gtl2/Meg3 ICRs, and their derivatives. Height of peaks in A–F reflects relative strength of DNA methylation across all of the tracks in each panel (or each subpanel of F), linearly scaled between the baseline and the maximal methylation in the displayed area shown with vertical bars at right. (G) Retention of DNA methylation and hydroxymethylation at an IAP location in PGCLCs and PGCs. Height of peaks reflects relative strength of DNA methylation across the seven 5meC tracks (linearly scaled between the baseline and the maximal methylation in the displayed area indicated with red vertical bar at right), seven DNA hydroxymethylation (5hmeC tracks, green vertical bar), or nonenriched genome resequencing coverage (Nonenriched track, blue vertical bar). IAP location is shown by horizontal bar at the top. (H) DNA methylation and hydroxymethylation profiles of the v6.5 mouse ESCs and their PGCLC derivatives at the ICR of the Kcnq1 imprinting cluster. Height of peaks reflects relative strength of DNA methylation across the two 5meC tracks (linearly scaled between the baseline and the maximal methylation in the displayed area indicated with red vertical bar at right), two 5hmeC tracks (cyan bar), or nonenriched genome resequencing track (blue bar). The ICR (KvDMR1) is shown with green shading, in which strong DNA methylation in the ESCs (a) was lost in PGCLCs (b) whereas DNA hydroxymethylation was enhanced upon ESC differentiation to PGCLC (compare peaks c and d). Two regions shaded with orange show loss of DNA methylation upon ESC differentiation to PGCLC without significant changes in DNA hydroxymethylation.
**Fig. S9.** Erasure of DNA hypermethylation at the IG-DMR and Gtl2-DMR of Gtl2(−) iPSCs during differentiation to PGCLCs. Deep-sequencing tracks of 5meCs (A) and 5hmCs (B) in the genomes of Gtl2(+) and Gtl2(−) mouse iPS cells and their EpiLC and PGCLC derivatives are shown. Height of peaks reflects relative strength of DNA methylation (A) or hydroxymethylation (B) across all seven tracks in each panel, linearly scaled between the baseline and the maximal methylation in the displayed area shown with vertical bars at right. Orange and cyan lines 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. Note that regions b and d locate between the two DMRs.