Stem Cell Reports
Article

ISSCR

OPEN ACCESS

Expanding homogeneous culture of human primordial germ cell-like cells maintaining germline features without serum or feeder layers

Mutsumi Kobayashi,[1,*] Misato Kobayashi,[1] Junko Odajima,[1] Keiko Shioda,[1] Young Sun Hwang,[2] Kotaro Sasaki,[2] Pranam Chatterjee,[3,4,5] Christian Kramme,[4,5] Richie E. Kohman,[4,5] George M. Church,[4,5] Amanda R. Loehr,[6] Robert S. Weiss,[6] Harald Jüppner,[7] Joanna J. Gell,[8,9,10] Ching C. Lau,[8,9,10] and Toshi Shioda[1,9,*]

[1]Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA 02129, USA
[2]Institute for Regenerative Medicine, Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA
[3]MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA
[5]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA
[6]Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
[7]Endocrine Unit and Pediatric Nephrology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
[8]The Jackson Laboratory for Genomic Medicine, Farmington, CT 06032, USA
[9]Connecticut Children's Medical Center, Hartford, CT 06106, USA
[10]University of Connecticut School of Medicine, Farmington, CT 06030, USA
*Correspondence: m.kobayashi310@gmail.com (M.K.), shioda@helix.mgh.harvard.edu (T.S.)
https://doi.org/10.1016/j.stemcr.2022.01.012

SUMMARY

In vitro expansion of human primordial germ cell-like cells (hPGCLCs), a pluripotent stem cell-derived PGC model, has proved challenging due to rapid loss of primordial germ cell (PGC)-like identity and limited cell survival/proliferation. Here, we describe long-term culture hPGCLCs (LTC-hPGCLCs), which actively proliferate in a serum-free, feeder-free condition without apparent limit as highly homogeneous diploid cell populations maintaining transcriptomic and epigenomic characteristics of hPGCLCs. Histone proteomics confirmed reduced H3K9me2 and increased H3K27me3 marks in LTC-hPGCLCs compared with induced pluripotent stem cells (iPSCs). LTC-hPGCLCs established from multiple human iPSC clones of both sexes were telomerase positive, senescence-free cells readily passaged with minimal cell death or deviation from the PGC-like identity. LTC-hPGCLCs are capable of differentiating to DAZL-positive M-spermatogonia-like cells in the xenogeneic reconstituted testis (xrTestis) organ culture milieu as well as efficiently producing fully pluripotent embryonic germ cell-like cells in the presence of stem cell factor and fibroblast growth factor 2. Thus, LTC-hPGCLCs provide convenient access to unlimited amounts of high-quality and homogeneous hPGCLCs.

INTRODUCTION

Primordial germ cells (PGCs), the common embryonic precursors of gametes, emerge in peri-implantation mammalian embryos (Sasaki et al., 2016). PGC-like cells (PGCLCs) are pluripotent stem cell-derived cell culture models of PGCs (Saitou, 2021). Mouse PGCLCs reflecting post-migration PGCs with advanced global genomic DNA (gDNA) demethylation can generate fertilization-competent sperm and oocytes (Hayashi et al., 2012; Ishikura et al., 2021; Miyoshi et al., 2016). In contrast, human PGCLCs (hPGCLCs) resemble earlier stages of embryonic PGCs with limited global gDNA demethylation (Irie et al., 2015; Mitsunaga et al., 2017; Sasaki et al., 2015; Tang et al., 2015). In organoid culture milieux, hPGCLCs can differentiate to more advanced stages of germline cells mimicking oogonia (Yamashiro et al., 2018) or prospermatogonia (Hwang et al., 2020).

Although the use of PGCLCs as a convenient and reliable surrogate model of embryonic PGCs is becoming increasingly popular (Saitou, 2021), the technical difficulty in cell culture expansion of PGCLCs poses a hurdle to its application for experiments, requiring large amounts of homogeneous cells such as chemical or genetic screening. Ohta et al. reported extended culture of mouse PGCLCs for 9 days on m220 feeders using a medium containing 10 μM each of rolipram and forskolin (FR10), stem cell factor (SCF), and 2.5% FCS (Ohta et al., 2017). Our preceding study maintained hPGCLCs for up to 21 days on STO feeders in a medium containing FR10, SCF, and 2.5% FCS (Gell et al., 2020). Even when hPGCLCs survive and proliferate *in vitro*, they tend to diverge from the PGC-like identity to other types of cells. By repeatedly eliminating such unwanted cells from the culture by FACS at every passaging step, Murase et al. provided evidence that hPGCLCs can maintain their proliferation potential long term *in vitro* (Murase et al., 2020). However, it remains a major challenge to expand hPGCLC culture to a large number of homogeneous cells maintaining their PGC-like characteristics.

Here, we describe long-term expansion of hPGCLCs homogenously preserving the PGC-like characteristics in cell culture without serum or feeder layer cells. Even after

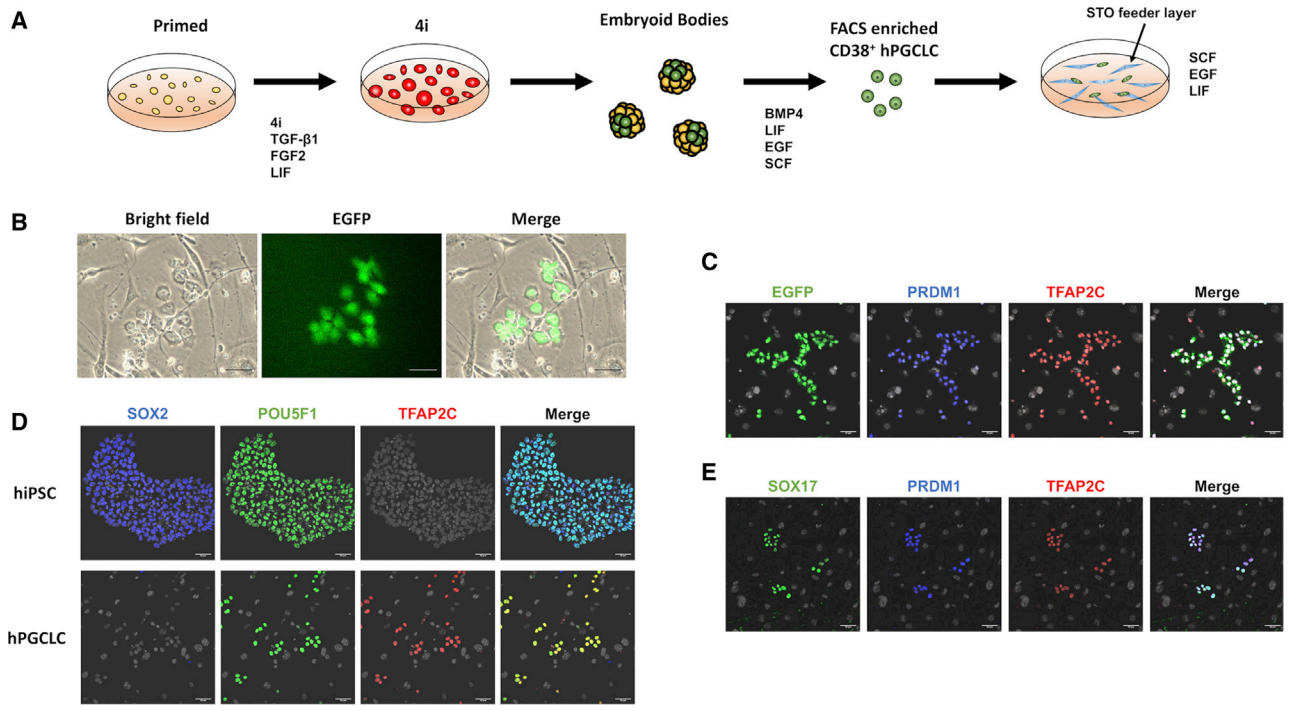


Figure 1. Initial expansion of hPGCLCs on STO feeder layer

(A) Steps to initiate long-term expansion of hPGCLCs. EGFP-labeled, primed pluripotency hiPSCs were exposed to the 4i/LIF medium for 48 hs and subjected to formation of EBs in microwells. hPGCLCs were induced in a medium containing BMP4, LIF, EGF, and SCF, enriched by FACS as CD38$^+$ cells, and inoculated on STO feeder layer in a medium containing LIF, EGF, and SCF. (B) A small cluster of hPGCLCs formed after 7-day culture. (C–E) IF with nuclei visualized with Hoechst (gray); Scale bars, 50 μm. (C) Expression of PRDM1 and TFAP2C in hPGCLCs maintained in culture for 4 weeks (D) Expression of SOX2, POU5F1, and TFAP2C in hiPSCs and hPGCLCs (maintained for 4 weeks on STO feeder layer). (E) Expression of SOX17, PRDM1, and TFAP2C in hPGCLCs maintained in culture for 4 weeks.

150 days of expansion *in vitro*, the telomerase-positive hPGCLCs maintain their PGC-like morphology and gene expression profile with no signs of senescence or deviation to other types of cells while their global gDNA is modestly demethylated upon increasing passage number. Upon changing cell culture condition, hPGCLCs are effectively converted to pluripotent embryonic germ cell-like cells (EGCLCs), restoring the induced pluripotent stem cell (iPSC)-like transcriptome and gDNA methylome but losing the imprinted allelic restriction of H19 RNA expression. Thus, long-term culture (LTC) hPGCLCs represent a novel type of stably maintained, pluripotency-relevant cell culture resource that permanently retains biological characteristics of hPGCLCs and are easily expanded *in vitro* as a stable cell line.

RESULTS

Expansion of hPGCLCs on STO feeder layer

Our preceding study showed an ~7.6-fold increase in yield of the CD38$^+$ hPGCLCs FACS-isolated from embryoid bodies (EBs) during the last 3 days of the 8-day EB culture in hPGCLC derivation medium containing bone morphogenetic protein 4 (BMP4) (Mitsunaga et al., 2017). However, our subsequent experiments showed that removal of BMP4 from days 5–8 EB culture did not reduce hPGCLC yield (Figure S1), leading us to speculate that other growth factors in hPGCLC derivation medium—namely, stem cell factor (SCF), epidermal growth factor (EGF), and leukemia inhibitory factor (LIF)—may support proliferation of hPGCLCs. To explore this possibility, we inoculated EGFP-labeled CD38$^+$ hPGCLCs freshly isolated from day-8 EBs onto the STO feeder layer in a medium containing SCF, EGF, and LIF (Figure 1A). Seven days later, hPGCLCs formed grape-like loose clusters, reproducing our preceding study (Gell et al., 2020) (Figure 1B). In this condition, hPGCLCs actively proliferated and readily passaged for 4 weeks on STO feeder layer. Immunofluorescence (IF) cell staining showed that these EGFP$^+$ hPGCLCs strongly expressed hPGCLC markers PRDM1, TFAP2C, and SOX17 (Figures 1C–1E) but not the pluripotency-associated marker SOX2 (Figure 1D). The PRDM1$^+$/TFAP2C$^+$ double-positive cells also strongly expressed POU5F1, a marker of

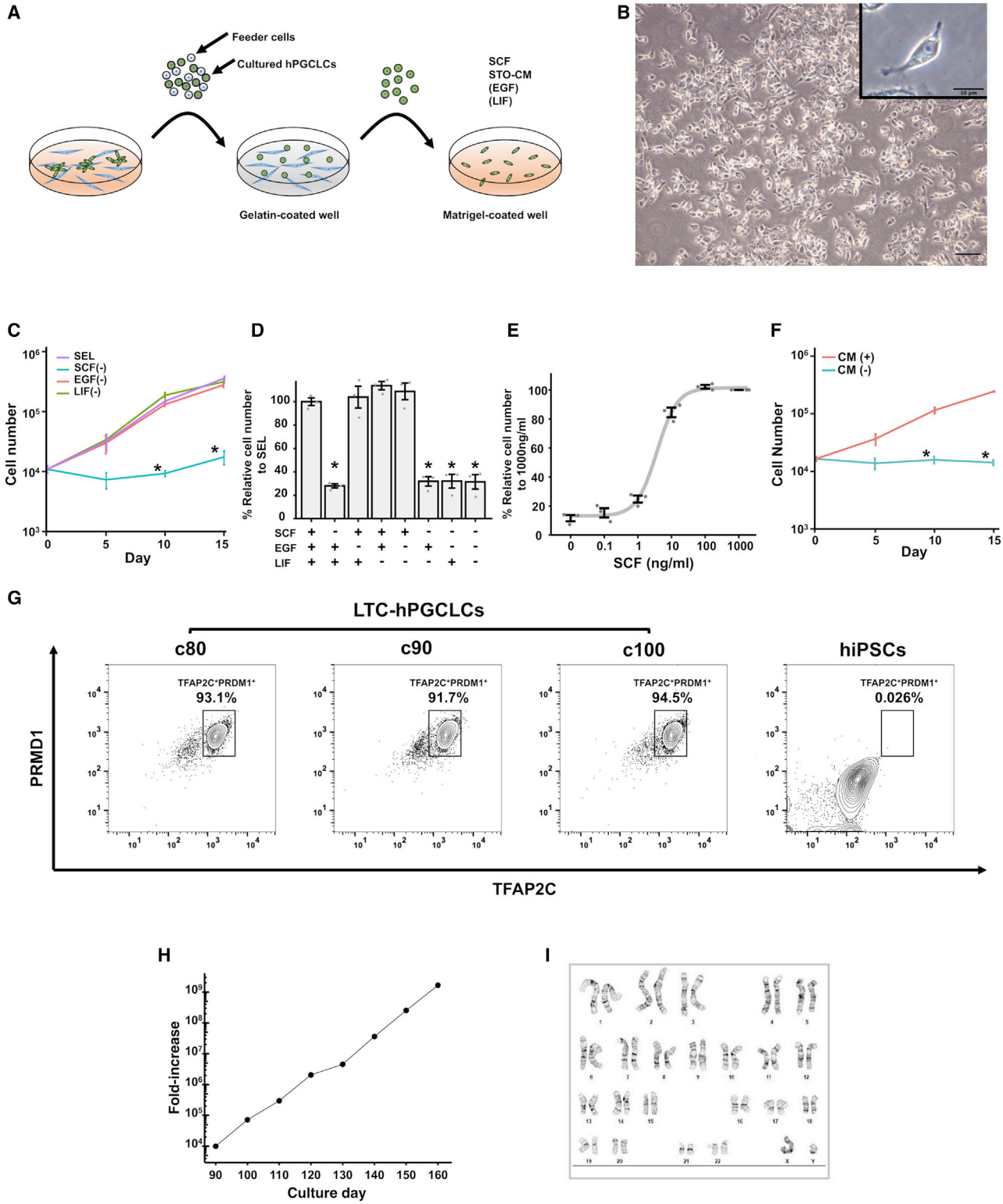


Figure 2. Feeder-free, long-term expansion of hPGCLCs

(A) Steps to initiate feeder-free culture hPGCLCs. After STO feeder cells were removed by differential adhesion to gelatin-coated surface, hPGCLCs were inoculated on Matrigel-coated wells in STO-CM containing SCF, EGF, and/or LIF. (B) Dispersed appearance of LTC-hPGCLC

(legend continued on next page)

both human induced pluripotent stem cells (hiPSCs) and PGCLCs (Figure 1D). These results indicate that hPGCLCs expanded on STO feeder layer in the presence of SCF, EGF, and LIF, preserving expression of hPGCLC markers. Hereafter, we refer to the long-term culture hPGCLCs as LTC-hPGCLCs.

Feeder-free expansion of LTC-hPGCLCs

We next attempted to expand LTC-hPGCLCs in a feeder-free condition. When LTC-hPGCLCs reached ~20% confluence, we removed STO cells by differential adhesion to gelatin-coated dishes, and LTC-hPGCLCs were inoculated on Matrigel-coated dishes in STO-conditioned medium (STO-CM) containing SCF, EGF, and LIF (Figure 2A). In this condition, which is referred to as the SEL condition, LTC-hPGCLCs scattered on the Matrigel-coated surface and randomly migrated (Figure 2B). LTC-hPGCLCs harbored cytoplasmic granules that were not found in hiPSCs (Figure 2B, inset) and still expressed PGC markers PRDM1, TFAP2C, SOX17, CD38, and POU5F1 but not the pluripotency-associated marker SOX2 (Figures S2A–S2C). All LTC-hPGCLCs strongly expressed CD38, and all $CD38^+$ hPGCLCs also expressed both PRDM1 and TFAP2C (Figures S2D–S2F). LTC-hPGCLCs actively proliferated in STO-CM supplemented with SCF, whereas EGF or LIF was dispensable (Figures 2C, 2D). A significant growth-supporting effect of SCF was observed at 10 ng/mL and saturated at 100 ng/mL (Figure 2E). The size of the caspase3/7-activated apoptotic cells in rapidly growing LTC-hPGCLC cell population was minimal (2.28%), whereas cells exposed to staurosporine, a potent apoptosis inducer, increased the apoptotic population size to 10.1% (Figure S2G). The small size of the apoptotic cell population agrees with our microscopic observations supporting the apparent absence of significant cell death during culture and passaging of LTC-hPGCLCs. Whereas LTC-hPGCLCs did not proliferate on non-coated plastic wells, coating with laminin or vitronectin supported short-term proliferation as efficiently as Matrigel for up to 7 days (Figures S3A, S3B). Whether coating with laminin and/or vitronectin supports long-term expansion of a homogeneous population of LTC-hPGCLCs needs to be determined in future studies. The SCF-supported, feeder-free expansion of LTC-hPGCLCs was dependent on STO-CM (Figure 2F). Hereafter, we refer to to the LTC-hPGCLC maintenance medium containing both SCF and STO-CM as S-CM.

Previous studies attempted to maintain PGCLCs in media containing 2.5% fetal calf serum (FCS) (Gell et al., 2020; Murase et al., 2020; Ohta et al., 2017). However, addition of 2.5% hESC-grade FCS to S-CM resulted in rapid emergence of cells morphologically distinct from hPGCLCs and lost expression of both PRDM1 and TFAP2C (Figures S3C and S3D). These $PRDM1^-/TFAP2C^-$ double-negative cells were morphologically heterogeneous (Figure S3C), with no characteristics of pluripotent stem cells, and many of them rapidly proliferated to dominate the cultures. In contrast, at least 90% of LTC-hPGCLCs expanded in S-CM maintained expression of both PRDM1 and TFAP2C after 80–100 days of culture (Figures 2G, S3C–S3E).

LTC-hPGCLCs actively proliferated for at least 160 days with a 4.5-day population doubling time, expanding an initial population of 10,000 hPGCLCs to more than one billion cells in 70 days (Figure 2H). A G-banding assay (Figure 2I) and digital karyotyping (Figure S4) demonstrated the normal male diploid karyotype of LTC-hPGCLCs after 100 days of feeder-free expansion. The above-mentioned experiments performed with LTC-hPGCLCs derived from the A4 male hiPSCs were reproduced using a female hiPSC clone (ATCC-BXS0115/ACS-1029; referred to to as F2 in this study) as well as two more male hiPSC clones, A5 and 9A13 (Figures S2F and S4). Note that the hiPSC clones A4 and A5 were derived from the same donor, whereas 9A13 was from a different donor.

STO-CM inhibits expression of pluripotency-associated markers in LTC-hPGCLCs

Mammalian PGCs become committed to produce only gametes (i.e., unipotent) after they colonize in the nascent gonads, whereas earlier stages of PGCs are poised to de-differentiate into the pluripotent state embryonic germ

feeder-free culture. Phase contrast images show high-density cells (scale bar, 100 μm) and a magnified single cell (inset; scale bar, 10 μm). (C–E) SCF dependency of LTC-hPGCLC expansion in STO-CM. Data were pooled from three independent experiments. (C) Growth curves of LTC-hPGCLCs in STO-CM containing three growth factors: SCF, EGF, and LIF (SEL; purple line, full condition), or SEL lacking SCF (blue), EGF (orange), or LIF (green). Each data point represents mean ± SEM *Statistically significant reduction ($p < 0.05$) in growth compared with the SEL condition. (D) Relative numbers of LTC-hPGCLCs expanded in a 7-day culture in STO-CM in the presence of indicated growth factors (mean ± SEM). Cell numbers of the condition containing SEL is defined as 100%; * statistically significant reduction from it ($p < 0.05$). (E) A dose-response curve of SCF for 7-day growth of LTC-hPGCLCs in STO-SCM without EGF or LIF. Each data point represents mean ± SEM (F) Growth curves of LTC-hPGCLCs in the presence of SCF with or without STO-CM (mean ± SEM; *$p < 0.05$). Data were pooled from three independent experiments. (G) Expression of PRDM1 and TFAP2C in hiPSCs and LTC-hPGCLCs expanded for 80–100 days in culture. Fixed and permealized cells were immunostained and analyzed with FACS. Rectangular gates indicate $PRDM1^+/TFAP2C^+$ cells. (H) Growth curve of feeder-free LTC-hPGCLCs from 90 to 160 days supported by SCF and STO-CM. (I) Normal diploid karyotype (G-banding) of LTC-hPGCLCs expanded in the SCF-supported STO-CM for 100 days.

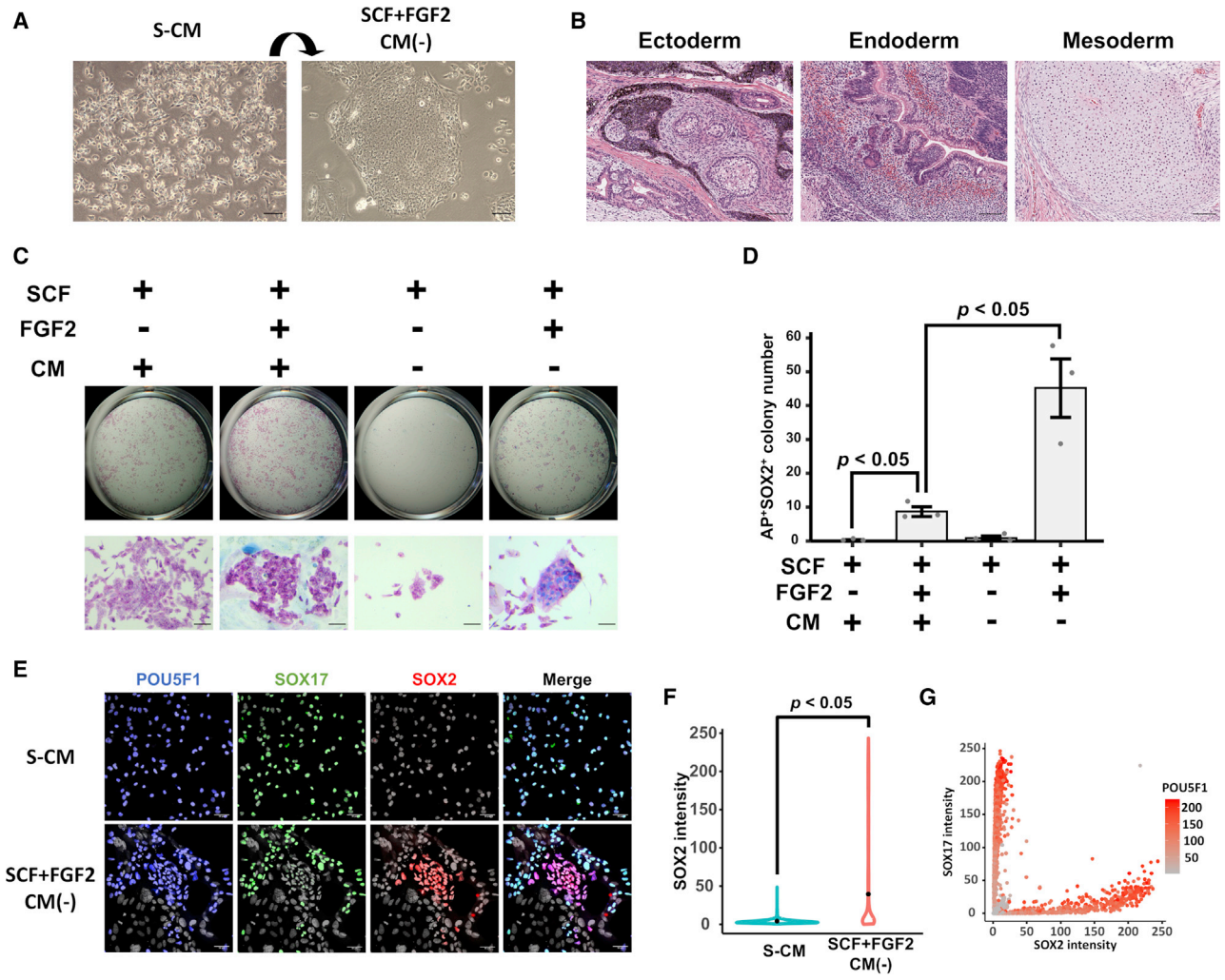
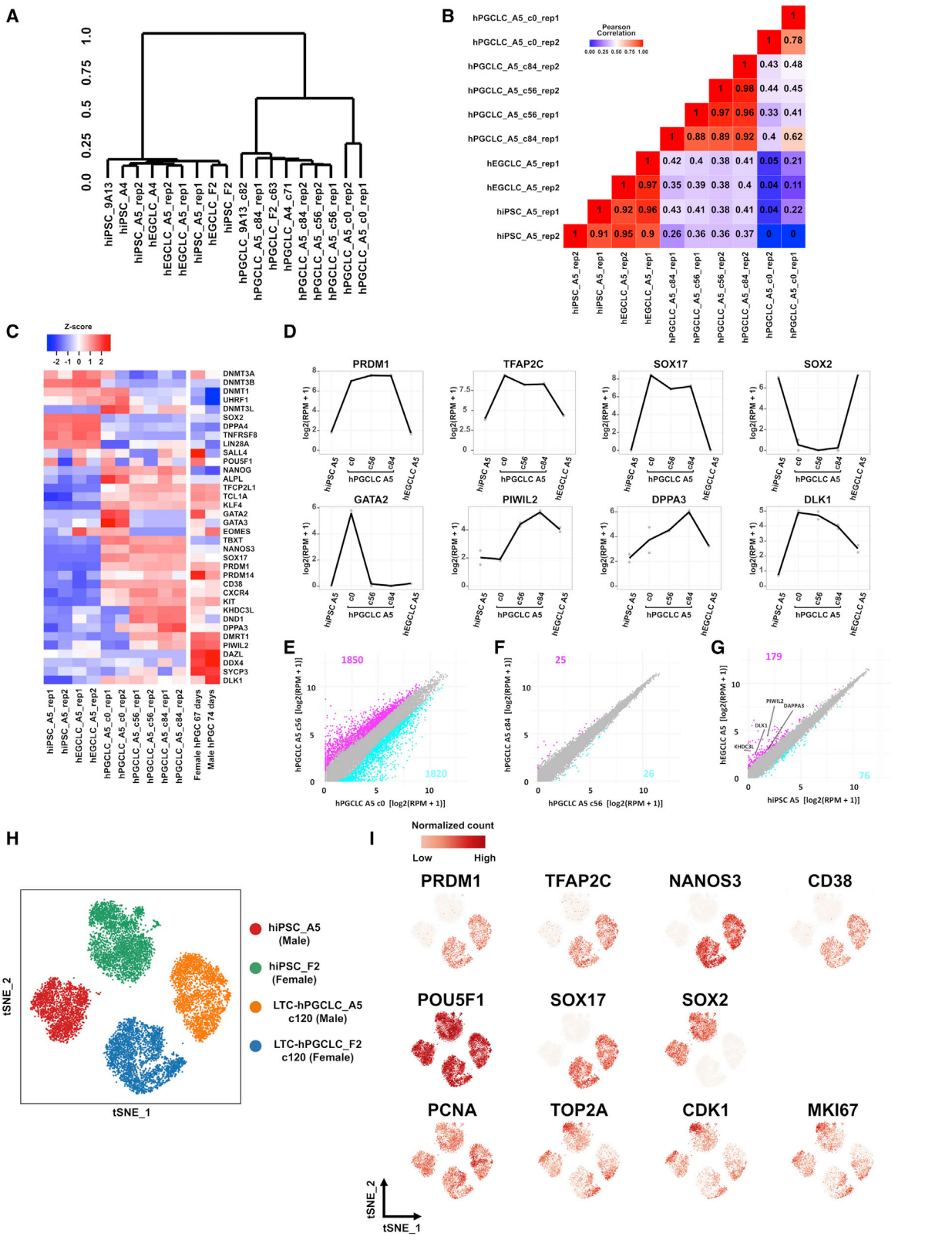


Figure 3. De-differentiation of LTC-hPGCLCs to hEGCLCs

(A) Phase contrast images of hPGCLCs cultured in S-CM (left) and hEGCLCs emerged after 10-day culture containing SCF and FGF2 but lacking STO-CM (right). (B) H&E staining of a teratoma formed from hEGCLCs in NSG mice. (C) hPGCLCs cultured in the indicated conditions were stained for alkaline phosphatase (AP, red) and SOX2 (blue). Top row, global appearance; bottom row, microscopic images (scale bar, 50 μm). (D) Number of colonies including $AP^+/SOX2^+$ cells in each well (mean ± SEM). Data were pooled from three independent experiments. (E) IF of POU5F1 (blue), SOX17 (green), and SOX2 (red) in LTC-hPGCLCs maintained for 10 days in S-CM or hEGCLC induction medium containing SCF and FGF2 but lacking CM. Nuclei were visualized by Hoechst (gray). (F) Nuclear expression of SOX2 in LTC-hPGCLCs cultured in S-CM or hEGCLC induction medium for 10 days (violin plots; dots show average intensities). Data were pooled from three independent experiments. (G) Quantitative image analysis on nuclear expression of SOX2, SOX17, and POU5F1 in LTC-hPGCLCs cultured for 10 days in hEGCLC induction medium. Data were pooled from three independent experiments.

cells (Nicholls et al., 2019). To test whether STO-CM prevents hPGCLCs from de-differentiation to pluripotency, we cultured feeder-free LTC-hPGCLCs in the SCF-containing maintenance medium lacking STO-CM but supplemented with fibroblast growth factor 2 (FGF2), which supports the growth of human primed pluripotent stem cells, and we observed that many large colonies morphologically similar to hiPSC colonies emerged in 10 days (Figure 3A). Those hiPSC-like cells formed teratomas when transplanted into NOD scid-γ (NSG) mice (Figure 3B). Differentiation to three germ layers in these teratomas confirmed the pluripotency of the iPSC-like cells. Whereas LTC-hPGCLCs were alkaline phosphatase (AP) positive but SOX2 negative (Figure 3C, leftmost), the hiPSC-like colonies were $AP^+/SOX2^+$ double positive (Figure 3C, rightmost). These $AP^+/SOX2^+$ cells readily expanded in the mTeSR medium for primed human pluripotent stem cells and were transcriptionally indistinguishable from *bona*



(legend on next page)

fide primed hiPSCs (Figure 4). On the basis of the above-mentioned observations, we concluded that the AP^+/$SOX2^+$ hiPSC-like cells resemble human embryonic germ cells; thus, we refer to to them as human embryonic germ cell-like cells (hEGCLCs). In the absence of both STO-CM and FGF2, hPGCLCs were unable to proliferate, and hEGCLCs did not emerge in the culture either (Figures 3C and 3D). Finally, in the presence of STO-CM, derivation of hEGCLCs was strongly suppressed even in the presence of FGF2 (Figures 3C and 3D). These results support the notion that STO-CM suppresses de-differentiation of hPGCLCs to hEGCLCs. SOX17 and SOX2 were identified as specific markers of LTC-hPGCLCs and hEGCLCs, respectively (Figure 1D; also, see, Figures S2B, and S5). During LTC-hPGCLC conversion to hEGCLC, a small portion of cells lost expression of POU5F1, a common marker of these two cell types (Figure 3E). SOX2 was expressed in LTC-hPGCLCs only when they were exposed to the hEGCLC derivation medium (SCF^+/$FGF2^+$/CM^-) but not when maintained in the S-CM medium (Figure 3F). In the $POU5F1^+$ LTC-hPGCLCs maintained in the hEGCLC derivation condition, expressions of SOX17 and SOX2 were mutually exclusive, which was confirmed by quantitative image analysis of the IF photographs (Figure 3G). Eventually, all cells maintained in the hEGCLC derivation medium expressed SOX2 (Figure 3C).

LTC-hPGCLCs maintain transcriptomic characteristics of early-stage human PGCs

We determined transcriptomic profiles of feeder-free LTC-hPGCLCs by bulk RNA-seq. Transcriptomes of freshly isolated, pre-expansion (c0) hPGCLCs and long-term-expanded (c56–c84) hPGCLCs were similar but distinguishable (Figures 4A, 4B, and 4E), agreeing with the observation made by Murase et al. that transcriptome of the pre-expansion hPGCLCs significantly changed during a short-period expansion culture for 10 days (Murase et al., 2020). Both Murase et al. and we observed that expression of GATA2 and GATA3 strongly diminished from pre- to post-expansion of hPGCLCs (Figure 4C and Murase et al., 2020). Both the pre- and post-expansion hPGCLCs maintained strong expression of representative hPGC markers PRDM1, TFAP2C, SOX17, NANOS3, KIT, CD38, and CXCR4 (Figures 4C, 4D, and S5), supporting their legitimate germline identity. Transcriptomes of LTC-hPGCLCs derived from different iPSC clones and maintained for 56–84 days (c56–c84) showed strong similarities among one another (Figures 4A, 4B, and 4F). Interestingly, LTC-hPGCLCs expressed later-stage hPGC markers DPPA3, PIWIL2, and DND1 more strongly than pre-expansion hPGCLCs (Figures 4C, 4D, and S5); however, expression of germline markers reflecting more advanced stages of development (e.g., DDX4 and DAZL) in LTC-hPGCLCs remained weak (Figures 4C, 4D, and S5). Tyser et al. recently presented single-cell RNA-seq data of *in vivo* hPGCs in very early stages of development at 16–18 gestational days (Tyser et al., 2021). Their dataset identified eight $NANOG^+$/$POU5F1^+$/$NANOS3^+$ hPGCs, which sporadically expressed hPGC markers TFAP2C, PRDM1, or CD38 and were negative for GATA2, GATA3, DAZL, DDX4, or SYCP3 (Figures S6A and S6B). On the other hand, single-cell RNA-seq data presented by Li et al. for *in vivo* $NANOG^+$/$NANOS3^+$ hPGCs in a later stage at 4 weeks of gestation (Li et al., 2017) expressed the late-PGC markers DAZL and DDX4 weakly but at unignorable levels (Figures S6C–S6E). In either of these single-cell RNA-seq data, expression of GATA2 or GATA3 was not detected in hPGCs (Figures S6B, S6D, and S6E), placing our c0 hPGCLCs to a unique $GATA2^+$/$GATA3^+$ stage in hPGC development as described by Kojima et al. (2021). Taken together, these results indicate that hPGCLCs experience significant transcriptomic changes during the initial transition from the GATA2/3-positive pre-expansion state to the GATA2/3-negative post-expansion state, whereas both the pre- and post-expansion cells are legitimate hPGCLCs. Once entering the post-expansion state, transcriptomes of LTC-hPGCLCs remain stable.

Homogeneity of LTC-hPGCLCs determined by single-cell RNA-seq

To evaluate cellular heterogeneity in LTC-hPGCLC cultures, we examined their single-cell RNA-seq profiles. The t-distributed stochastic neighbor embedding (t-SNE) plot analysis demonstrated the highly homogeneous nature of LTC-hPGCLC cultures comparable to that of hiPSCs (Figure 4H). Expression of hPGCLC marker genes (*PRDM1*,

Figure 4. Transcriptomic profiles of hiPSCs, hPGCLCs, and hEGCLCs

(A–D) Bulk RNA-seq data. (A) Unsupervised hierarchical clustering. hPGCLCs generated from hiPSCs (clones A4, A5, F2, and 9A13) were long-term expanded (culture days are shown as c#). hEGCLCs were derived from hPGCLCs. Rep1 and rep2 are independently performed replicates. (B) Pearson's correlation coefficients between transcriptomes. (C) Heatmap representation of marker gene expression. hPGC data (gestational days 67–74) are from another study (Gkountela et al., 2015). (D) Gene expression dynamics in hiPSCs, c0-hPGCLCs, LTC-hPGCLCs, and hEGCLCs. (E–G) Changes in gene expression in hiPSCs, hPGCLCs, and hEGCLCs. Scatterplots compare (E) hPGCLCs (c0) versus LTC-hPGCLCs (feeder free, 56 days); (F) LTC-hPGCLCs cultured feeder free for 56 and 84 days; and (G) hiPSCs and hEGCLCs. Axes are in $\log_2$ (reads per million + 1), and the numbers and data points of differentially expressed genes (log2 [fold change] $\geq$ 1) are shown in magenta or cyan. (H and I) Single cell RNA-seq tSNE plot (H) and superimposed expression of marker genes (I).

TFAP2C, *SOX17*, *NANOS3*, and *CD38*) and a pluripotency-associated marker gene (*SOX2*) showed strong homogeneity within each cluster of cells, whereas expression of *POU5F1* was homogenously ubiquitous in all cells (Figure 4I). The tSNE clusters of LTC-hPGCLCs (as well as those of hiPSCs) were clearly separated by sex, which is consistent with significantly stronger expression of genes located on the X chromosome in female cells than in males', although expression of autosomal genes is largely unaffected by sex (Figure S5E). Degrees of the female-augmented expression of the X chromosomal genes were largely unaffected by cell types (Figure S5E). The sensitivity of our single-cell RNA-seq analysis was sufficient to detect cellular heterogeneity in expression of cell cycle genes (*PCNA* for G1-S; *TOP2A*, *CDK1*, and *MKI67* for G2/M) within each cluster (Figure 4I). PCA analysis separated the four tSNE clusters into two groups by cell types (hiPSC versus hPGCLC), with small degrees of heterogeneity in each group along by sex (Figure S5C), largely agreeing with the tSNE plots (Figure S5D). Thus, the single-cell RNA-seq data confirm that LTC-hPGCLCs maintain highly homogeneous populations of hPGCLCs consisting of actively proliferating cells even after long-term expansion without accumulating other types of cells.

Epigenetic characteristics of LTC-hPGCLCs

Global deposition of H3K27me3 increases upon differentiation of human pluripotent stem cells to hPGCLCs, whereas deposition of H3K9me2 decreases (Sasaki et al., 2015). IF and western blotting confirmed that these changes are preserved in LTC-hPGCLCs (Figures 5A and 5B), providing further evidence that LTC-hPGCLCs maintain their hPGC-like characteristics.

Although proteomics of histone modifications is a powerful approach to quantitative determination of epigenetic characteristics of mammalian cells, it requires large amounts of homogeneous cells that are not readily obtained using previous methods of hPGCLC expansion (Gell et al., 2020; Murase et al., 2020). Taking advantage of our LTC-hPGCLCs, we performed proteomics determination of 84 histone modifications in their genome (Figure S7A) and were able to confirm a reduction in H3K9me2 and an increase in H3K27me3 as differences between LTC-hPGCLCs and hiPSCs, agreeing with IF and western blotting data (Figures 5A and, 5B). Various other changes in histone modifications between LTC-hPGCLCs and hiPSCs (e.g., H4K12Ac) were also evident, although detailed confirmations of such modifications await future studies.

We determined the global gDNA methylome of LTC-hPGCLCs by whole-genome bisulfite sequencing (WGBS; Figure 5C). The global 5-methylcytosine (5mC) level decreased from hiPSCs (mean = 73.3%) to c0-hPGCLCs (64.2%) and then further in LTC-hPGCLCs (54.2%) across all major genomic features (Figure 5D). A small decrease from c56 (56.2%) to c84 (54.2%) LTC-hPGCLCs was statistically insignificant. The global 5meC levels of hiPSCs and hEGCLCs were comparable, indicating that the ~20% global gDNA demethylation in LTC-hPGCLCs from their pluripotent precursor cells was largely reversible. The differentially methylated regions (DMRs) of the imprinting genes tended to be demethylated in LTC-hPGCLCs more strongly than the global average (Figure 5E). Among the DMRs, the intergenic-DMR (IG-DMR) of the *DLK1-DIO3* imprinting cluster and *H19* DMR showed relatively robust demethylation in (LTC-)hPGCLCs compared with hiPSCs (IG-DMR: 85.0% in hiPSCs versus 28.1% in LTC-hPGCLCs; *H19*: 70.3% in hiPSCs versus 18.6% in LTC-hPGCLCs). On the other hand, *PEG3*-DMR and *MEG3*-DMR maintained their high levels of CpG methylation (*PEG3*: 90.0% in hiPSC versus 82.6% in LTC-hPGCLCs; *MEG3*: 91.0% in hiPSCs versus 82.0% in LTC-hPGCLCs). Thus, DMR demethylation in LTC-hPGCLCs was heterogeneous and partial. The DMRs regulating the imprinted expression of the H19 RNA were almost completed demethylated in LTC-hPGCLCs, whereas these regions are ~50% methylated in hiPSCs, as confirmed by WGBS (Figure 5E) and methylation-specific multiplexed ligation-dependent probe amplification (MS-MLPA; Figures S7B–S7D). Agreeing with this strong demethylation, H19 RNA expression was bi-allelic in LTC-hPGCLCs of both sexes, whereas it is monoallelic in the precursor hiPSCs (Figures S7E and S7F). Whether other imprinting genes are bi-allelically expressed in LTC-hPGCLCs, and possibly in hEGCLCs as well, needs to be determined in future studies.

The long-terminal repeats (LTRs) of the human endogenous retroviruses (HERVs) generally reduced their 5mC contents in LTR-hPGCLCs, although the degree of reduction was relatively small in the HERV-IP family (Figure 5F). In contrast, LTR5 and LTR5_Hs (Hs, human specific), which are the youngest classes of HERVs, showed the most remarkable LTR demethylation in both c0-hPGCLCs and LTC-hPGCLCs (Figures 5F and 5G).

Transcriptomic and epigenomic characteristics of hEGCLCs

Transcriptomic profiles of hiPSCs and hEGCLCs showed striking similarities (Figures 4A and, 4C) with Pearson's correlation coefficients greater than 0.90 (Figure 4B). Between them, we identified 179 differentially expressed genes (DEGs; Figure 4G), including *KHDC3L*, *DPPA3*, *PIWIL2*, and *DLK1* (Figures 4C, 4D, and 4G). Because these four DEGs were also expressed more strongly in LTC-hPGCLCs than in hiPSCs (Figures 4C and 4D), we speculate that they were incompletely re-suppressed during re-programming of LTC-hPGCLCs to the pluripotent hEGCLCs.

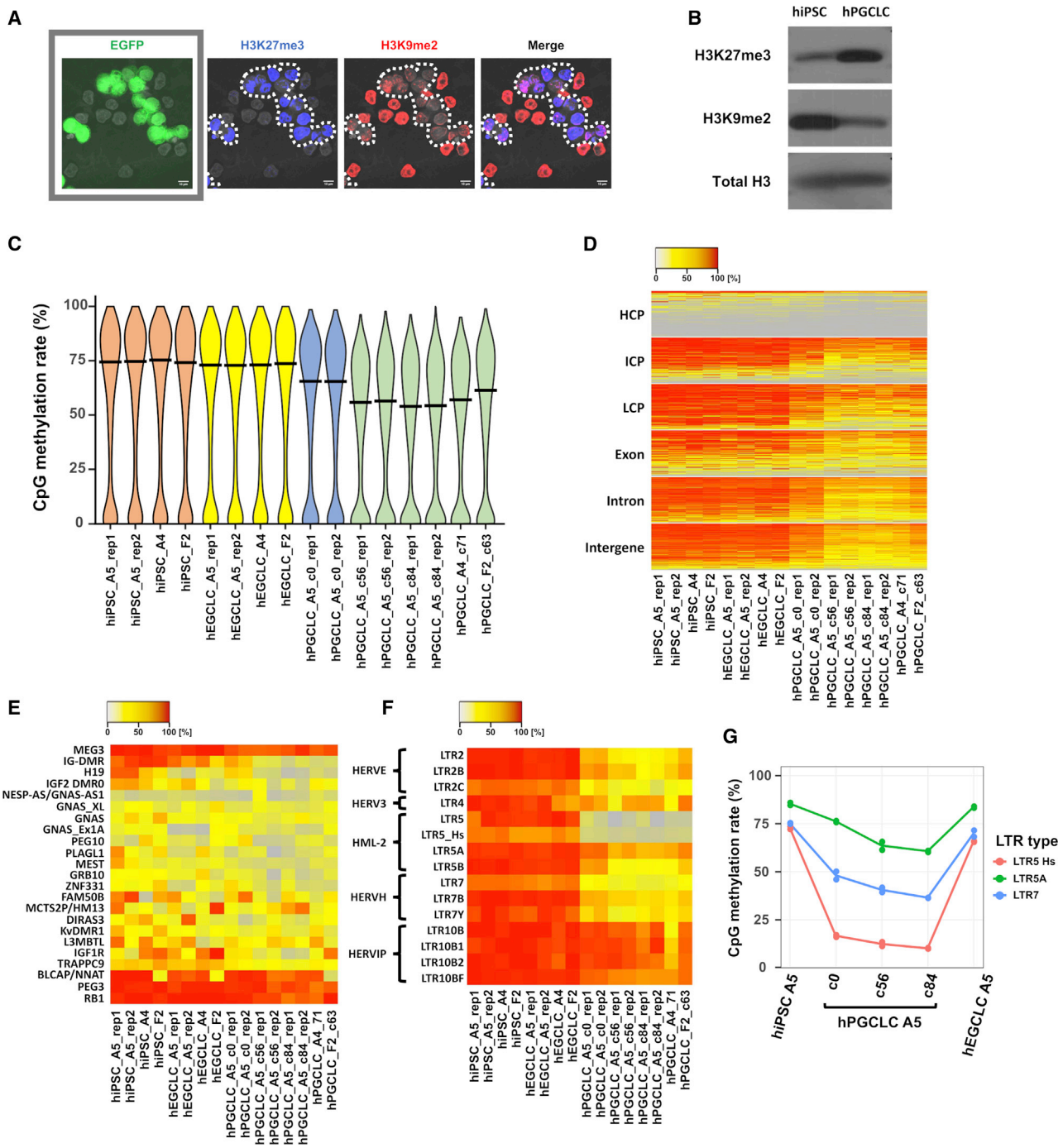


Figure 5. Epigenetic profiles of hiPSCs, hPGCLCs, and hEGCLCs

(A) IF comparison of H3K9me2 and H3K27me3 expression between hiPSCs and LTC-hPGCLCs. Mixtures of hiPSCs (EGFP negative) and LTC-hPGCLCs (EGFP positive) were subjected to multiplex staining for H3K27me3 (blue), and H3K9me2 (red). Nuclei were visualized by Hoechst (gray). Dotted lines delineate LTC-hPGCLCs. (B) Western blotting detection of H3K27me3 and H3K9me2 in hiPSCs and LTC-hPGCLCs. (C) Violin plots of DNA methylomes (whole-genome bisulfite sequencing). Horizontal bars indicate global averages. (D–F) Heatmaps of CpG methylation at (D) various genomic features (HCP, ICP, and LCP are gDNA regions with high, intermediate, and low densities of CpG sites, respectively); (E) imprinting-controlling differentially methylated regions (DMRs), and (F) representative human endogenous retroviruses. (G) Line plot of CpG methylation rate of LTR5_Hs, LTR5A, and LTR7 in the indicated cells.

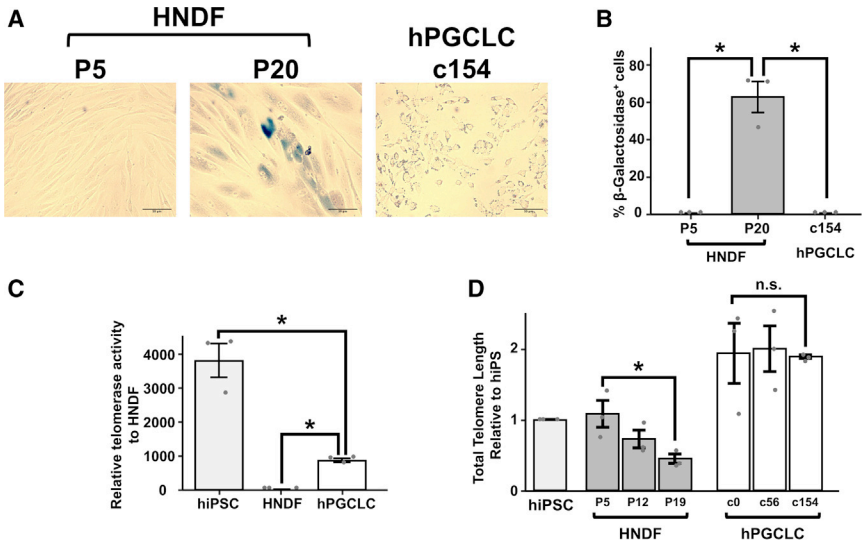


Figure 6. Senescence assessment of LTC-hPGCLCs

(A and B) Enzymatic detection of senescence-associated β-galactosidase (βGal) activity. Human neonatal dermal fibroblasts (HNDF) were maintained for five or 20 passages and subjected to βGal staining. (A) Phase contrast microscopic images (scale bar, 50 μm). (B) Quantification of $βGal^+$ cells (mean ± SEM). Data were pooled from three independent experiments. (C) Telomerase enzymatic activity. Data were normalized to HNDF and shown as mean ± SEM Data were pooled from three independent experiments. (D) Whole-genome telomere length. hiPSCs and LTC-hPGCLCs and HNDF (passages 5–19) were subjected to a telomere length assay, and data are presented as mean ± SEM normalized to hiPSCs. Data were pooled from three independent experiments. *Statistical significance ($p < 0.05$, t test); n.s., not significant.

However, degrees of CpG methylation in promoter or gene body of these four DEGs determined by WGBS did not show significant differences between hiPSCs and hEGCLCs. Expression of DLK1 protein in LTC-hPGCLCs or hEGCLCs was not detected by IF, apparently due to technical issues. Whereas the imprinting genes differentially expressed between hiPSCs and hEGCLCs may significantly contribute to biological differences between these two pluripotent types of cells, non-imprinting DEGs would certainly affect the distinct characteristics of these cells as well. Future studies are necessary for more detailed comparisons between hiPSCs and hEGCLCs.

Whereas DMRs of several imprinting genes were demethylated in LTR-hPGCLCs compared with hiPSCs, many of them were re-methylated in hEGCLCs but to weaker degrees than in hiPSCs (Figure 5E). For example, the average 5mC level of the IG-DMR, which is involved in regulation of *DLK1* expression, was high in hiPSCs (84%), reduced in the c0-hPGCLCs (64%), and decreased further in LTC-hPGCLCs (28%) but partially re-methylated to 70% in hEGCLCs (Figure 5E). These results suggest that de-differentiation of LTC-hPGCLCs to hEGCLCs is accompanied by re-methylation of several imprinting-controlling DMRs although degrees of their re-methylation often do not reach the levels in the precursor hiPSCs. The mechanism of DNA re-methylation during the transition from LTC-hPGCLC to hEGCLC is unknown. During this transition, expression of the *de novo* DNA methyltransferases DNMT3A and DNMT3B as well as the maintenance methyltransferase DNMT1 (along with its cofactor UHRF1) was upregulated to the level observed in hiPSCs (Figure 4C). The increased expression of these methyltransferases/cofactor may support global DNA re-methylation in hEGCLCs. Although expression of DNMT3L mRNA was rather decreased, DNMT3L protein might have been able to support functions of the *de novo* methyltransferases during the transition.

LTC-hPGCLCs are telomerase-positive cells whose telomere length is preserved during long-term expansion

Because the growth curve of LTC-hPGCLCs was linear during their long-term expansion (Figures 2H, S4B, and S4C), we examined whether LTC-hPGCLCs experience cellular senescence. Senescence-associated β-galactosidase activity was detected in human neonatal dermal fibroblasts (HNDFs), from which the A4 and A5 hiPSC clones were generated (Mitsunaga et al., 2017), after 20 passages (Figures 6A, 6B). In contrast, our attempts to detect β-galactosidase activity in c154 LTC-hPGCLCs under the same condition was conclusively negative (Figure 6B). LTC-hPGCLCs had significant telomerase activity that was ~25% the level of hiPSCs (Figure 6C) and showed no sign of telomere length shortening during long-term expansion in cell culture (Figure 6D). These results indicate that the LTC-hPGCLC is a telomerase-positive cell culture model whose proliferation is not limited by cellular senescence, reminiscent of the ability of human pluripotent stem cells or malignant cells to proliferate without apparent limits of passaging.

Differentiation of LTC-hPGCLCs to DAZL-positive cells resembling M-prospermatogonia in the xrTestis organ culture

Hwang et al. recently developed the xenogeneic reconstituted testis (xrTestis) organ culture system consisting of

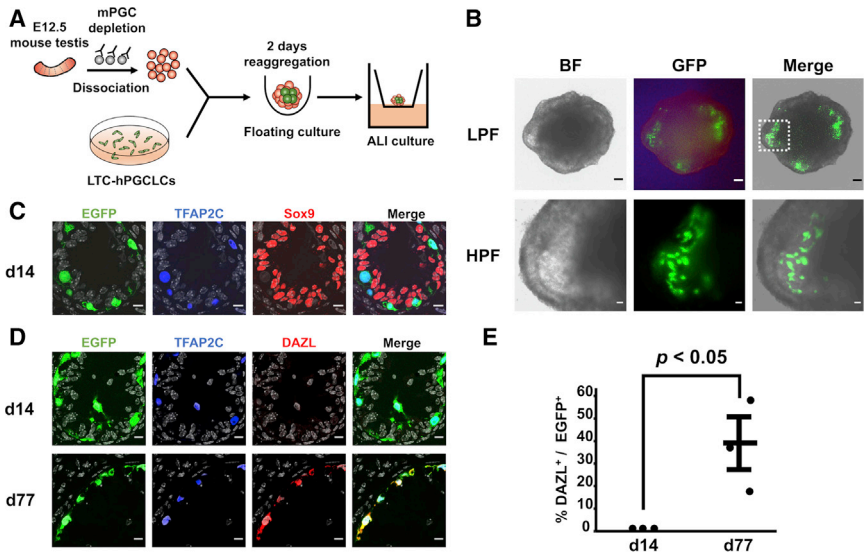


Figure 7. Differentiation of LTC-hPGCLCs to DAZL-positive gonocyte-like cells in xrTestis

(A) Schematic representation of xrTestis. Single-cell suspensions of E12.5 mouse fetal testes were depleted of mouse PGCs (mPGCs) by MACS and co-cultured with EGFP-labeled LTC-hPGCLCs in a floating culture condition. Resulting cell aggregates were maintained in an air-liquid interphase (ALI) culture condition. (B) Microscopic images of day-77 xrTestis. BF, bright field. The area shown by a dotted rectangle in the low-power field (LPF; scale bar, 100 μm) is enlarged in the high-power field (HPF; scale bar, 20 μm). (C) IF of human TFAP2C and mouse Sox9 in day-14 xrTestis. Scale bars, 10 μm. (D) IF of human TFAP2C and DAZL in day-14 and day-77 xrTestis. Scale bars, 10 μm. (E) Ratio of $DAZL^+$ cells in $EGFP^+$ cells in day-14 and day-77 xrTestis (mean ± SEM of three independently formed xrTestis).

mouse fetal testicular somatic cells and hPGCLCs and demonstrated that freshly isolated, DAZL-negative hPGCLCs were capable of differentiating to DAZL-positive cells resembling M-prospermatogonia (Hwang et al., 2020). We constructed the xrTestis organ culture aggregates using EGFP-labeled LTC-hPGCLCs expanded for 120 days in cell culture (Figure 7A). The xrTestis spontaneously formed tubular structures reminiscent of seminiferous tubules, and LTC-hPGCLCs were localized exclusively within these tubules (Figure 7B). IF of the cross sections of the tubules revealed central hollow spaces surrounded by Sox9-positive mouse Sertoli cells, and the $EGFP^+/TFAP2C^+$ LTC-hPGCLCs were integrated within the tubules outside the Sertoli cells (Figure 7C). After 77-day of xrTestis culture, approximately 40% of the $EGFP^+/TFAP2C^+$ cells expressed DAZL, whereas no DAZL expression was detected at 14 days (Figures 7D, 7E). These data indicate that LTC-hPGCLCs retained their capability of differentiating to "M-prospermatogonia-like cells" (MLCs) in xrTestis after 120 days of expansion in cell culture.

DISCUSSION

Our current study established LTC-hPGCLCs, which can be expanded in cell culture long term in a serum-free and feeder-free condition without losing their homogeneous, PGC-like cellular identity. FACS analyses showed no accumulation of cells that lost expression of the hPGCLC markers PRDM1 or TFAP2C in the LTC-hPGCLC cultures after *in vitro* expansion for 70–153 days (Figures 2G and S2F). This remarkable stability is in sharp contrast to a preceding study, in which hPGCLCs had to be re-enriched by FACS at every passage since a large portion of them continuously lost their germline identity over time (Murase et al., 2020). Our RNA-seq data demonstrate very strong transcriptomal similarities between LTC-hPGCLCs maintained in culture for 56 and 84 days (Figure 4B) as well as robust and persistent expression of a number of hPGC marker genes (Figure 4C). These results confirmed that our protocol supports FACS-free expansion of LTC-hPGCLCs retaining their hPGC-like identity. Single-cell RNA-seq data demonstrate the highly homogeneous aspects of LTC-hPGCLCs after 120-day expansion in cell culture, whereas heterogeneous expression of cell cycle genes was evident in the homogeneous clusters of cells (Figures 4H, 4I, S5C, and S5D).

During cell culture expansion or passaging, LTC-hPGCLCs show a minimal level of apoptotic cell loss, whereas they are still sensitive to apoptotic induction by staurosporine (Figure S2G). This novel resource would significantly lower a hurdle to using the hPGCLC model for human germline cell research. Successful maintenance of LTC-hPGCLCs in feeder-free cell culture is dependent on four critical factors, namely, the use of SCF, STO-CM, and Matrigel, and the absence of FCS. STO-CM is required to stimulate hPGCLC proliferation (Figures 2F, 3C, and 3D) and to prevent de-differentiation of hPGCLCs to pluripotent hEGCLCs (Figure 3). During the conversion of the $POU5F1^+/SOX2^-/SOX17^+$ LTC-hPGCLCs to $POU5F1^+/SOX2^+/SOX17^-$ hEGCLCs, expressions of SOX2 and SOX17 were mutually exclusive, with no evidence of $SOX2^+/SOX17^+$ double-positive cells (Figures 3E–3G).

In the presence of SCF and FGF2 but the absence of STO-CM, LTC-hPGCLCs were readily converted to hEGCLCs

(Figure 3). We observed no difference in the efficiency of hEGCLC generation between freshly isolated hPGCLCs and LTC-hPGCLCs (data not shown). Nicholls et al. showed that early-stage, migrating mouse embryonic PGCs (mPGCs) can generate EGCs but late-stage mPGCs that have settled in the genital ridges rapidly lose this capability (Nicholls et al., 2019). The capability of (LTC-) hPGCLCs to generate hEGCLCs may reflect that hPGCLCs resemble early-stage human embryonic PGCs (Mitsunaga et al., 2017). It is interesting to speculate that late-stage human embryonic PGCs in the genital ridges may be incapable of generating embryonic germ cells and that the xrTestis culture of (LTC-)hPGCLCs might be useful for exploring mechanisms of the retention and/or loss of the reinstatement of the pluripotency; this speculation should be addressed in future studies.

Addition of FCS to our optimized hPGCLC long-term maintenance medium (S-CM; SCF + STO-CM) turned disruptive rather than supportive (Figures S3A–S3C), which may partly explain why preceding efforts of long-term expansion of hPGCLCs in media containing FCS have proved challenging (Gell et al., 2020; Murase et al., 2020). Our attempts to replace the STO-CM with growth factors secreted from STO cells into the medium such as CXCL12, TGFβs, and IGFs (Talbot et al., 2012) have been unsuccessful so far.

Evidence is accumulating that specification of human and monkey PGCs in early-stage embryos require the pioneer transcription factors such as GATA3 and TFAP2A, which bind their target DNA sequences in tightly wrapped heterochromatin regions to induce chromatin remodeling and enhance their access to other transcription factors (Chen et al., 2019; Kojima et al., 2021; Murase et al., 2020; Sasaki et al., 2016). *In vivo* expression of GATA3 in the PGCs is strong soon after PGC specification but rapidly suppressed during their migration far before they settle in the genital ridges (Sasaki et al., 2016). *In vitro* expression of GATA3 in hPGCLCs is rapidly induced when precursor cell aggregates are exposed to the hPGCLC induction medium containing BMP4 (Chen et al., 2019; Kojima et al., 2021) but ceased during the course of hPGCLC induction (Chen et al., 2019) or during the initial, short-term (~10 days) expansion of hPGCLCs in cell culture (Murase et al., 2020). Prolonged expression of GATA3 from an inducible vector in hPGCLCs in the BMP4-containing hPGCLC induction medium resulted in major cell death (Kojima et al., 2021). Thus, our GATA2/3-positive, freshly isolated (c0) hPGCLCs correspond with the GATA2/3-positive hPGCLCs observed by others (Chen et al., 2019; Kojima et al., 2021; Murase et al., 2020), and our GATA2/3-negative LTC-hPGCLCs are equivalent to the post-expansion hPGCLCs described by Murase et al. (2020). Importantly, both the pre- and the post-expansion hPGCLCs obtained by us and others expressed the key hPGC (LC) markers, indicating that the two states of cells are both legitimate hPGCLCs. Thus, our LTC-hPGCLCs stably maintain the GATA2/3-negative, post-expansion state of hPGCLCs, which still resemble the DDX4/DAZL-negative, early-stage hPGCs (Chen et al., 2017). It is unknown whether long-term maintenance of the GATA2/3-positive state of hPGCLCs, which are presumably close to the germline specification stage of very early hPGCs *in vivo*, is possible or not. Applicability of our method to cell culture expansion of human embryonic PGCs in the early or late stage also remains to be determined by future studies.

Several markers associating late stages of hPGCs such as PIWIL2 or DPPA3 significantly increased their expression during the course of the long-term expansion of LTC-hPGCLCs (Figures 4C and 4D), implying a possibility that LTC-hPGCLCs alone can undergo a certain degree of maturation *in vitro*. However, the significant degree of global gDNA demethylation in LTC-hPGCLCs observed in our current study (Figure 5C; reaching to 54.2% at c84) was still far weaker than those observed for embryonic hPGCs (~30% in 59-day male PGCs; Gkountela et al., 2015) or advanced stages of hPGCLC-derived germline cells in xeno-organoid cultures (Hwang et al., 2020; Yamashiro et al., 2018). Therefore, *in vitro* maturation of LTC-hPGCLCs, if any, seems arrested at a stage mimicking hPGCs before their gonadal settlement. The LTR5 and LTR5_Hs human endogenous retroviruses showed exceptionally strong degrees of demethylation (Figures 5F and 5G) although the reason for the specific strong CpG demethylation at these endogenous retroviruses is unknown. Future studies will be required to elucidate the possible biological significance as well as epigenetic mechanism of the unique and strong CpG demethylation in (LTC)-hPGCLCs.

The global gDNA demethylation involving imprinting genes is a hallmark of epigenetic reprogramming in mammalian PGCs (Seisenberger et al., 2012; Seki et al., 2007). Some of the DMRs controlling imprinted gene expression were significantly demethylated in LTC-hPGCLCs (Figure 5E), and the strong demethylation at the IG-DMR (Figure 5E) may contribute to the increased expression of DLK1 mRNA in (LTC-)hPGCLCs (Figure 4D). Our RNA-seq data provided evidence that the *H19* imprinting gene is expressed mono-allelically in the precursor hiPSCs but bi-allelically in LTC-hPGCLCs (Figures S7E and S7F). Prevalence and mechanisms of the apparent loss of imprinting in (LTC-)hPGCLCs need to be characterized further in future studies.

The usefulness of CD38 as a surface marker for FACS enrichment of hPGCLCs was first reported by Irie et al. (2015), and our preceding study showed that the $CD38^{+}$ hPGCLCs strongly expressed representative hPGC markers such as PRDM1, TFAP2C, SOX17, and NANOS3 (Mitsunaga

et al., 2017). We also showed strong expression of CD38 in hPGCLCs generated using a different method, involving intermediate production of the incipient mesoderm-like cells (iMeLCs) and FACS-enriched using other cell surface markers, namely, EpCAM and integrin α6 (Mitsunaga et al., 2017; Sasaki et al., 2015). On the other hand, expression of several key hPGC markers such as NANOS3 or SOX17 in $CD38^-$ cells was far weaker than in $CD38^+$ cells (Mitsunaga et al., 2017). Single-cell RNA-seq profiling in our current study presents direct evidence that all individual LTC-hPGCLC cells expressing CD38 simultaneously and strongly express the hPGC markers PRDM1, TFAP2C, NANOS3, and SOX17 (Figures 4H, 4I, S5C, and S5D). Taken together, the CD38-positive hPGCLCs and LTC-hPGCLCs represent the authentic population of cells resembling human embryonic PGCs. On the other hand, there has been no evidence supporting the existence of CD38-negative subpopulations of hPGCs or hPGCLCs.

The protocol described here supports unlimited and rapid expansion of homogeneous hPGCLCs preserving hPGC-like characteristics and normal karyotype in a feeder-free condition. Even after prolonged expansion (120 days), LTC-hPGCLCs retained their capability of differentiating to DAZL-positive MLCs (Figure 7), reproducing a previous observation made with freshly isolated hPGCLCs (Hwang et al., 2020). The emergence of hPGCLC-derived $TFAP2C^+/DAZL^+$ double-positive cells in xrTestis was observed in our current study (Figure 7D) and in a preceding study (Hwang et al., 2020), using different rabbit polyclonal antibodies, whose specificity was confirmed by staining cryosections of human adult testes in the present study. Future studies will be necessary to determine whether $TFAP2C^+/DAZL^+$ double-positive germline cells also emerge in human embryonic testes. Two male and one female LTC-hPGCLC lines have been established in the current study, and their sex-specific DE-Gs clearly separated their tSNE profiles (Figures 4H, 4I, and S5E). LTC-hPGCLCs can be used for various experiments that require homogeneous populations of live cells, such as high-throughput chemical or genetic screenings, proteomics, or biochemical studies, possibly facilitating pharmacological or toxicological assessments of drugs or environmental toxicants for their germline effects. The histone proteomics experiment presented in our current study (Figure S7A) exemplifies the usefulness of access to large-yield and highly homogeneous populations of LTC-hPGCLCs offered by our method. LTC-hPGCLCs will also be valuable in studying carcinogenic mechanisms of germ cell tumors, which are believed to arise from hPGCs (Oosterhuis and Looijenga, 2019). Since frozen LTC-hPGCLCs are readily stored with high viability, a collection of LTC-hPGCLCs harboring various normal and disease-linked genetic backgrounds (e.g., sex, ethnicities, familial infertility) would be a useful resource that facilitates studies on health of human germline cells.

EXPERIMENTAL PROCEDURES

Induction of hPGCLCs

hPGCLCs were generated from hiPSCs as we previously described (Mitsunaga et al., 2017, 2019, 2021) with 5 or 8 days of exposure to BMP4 as shown in Figure S1. Day 8 EBs were dissociated to single cells, using the Embryoid Body Dissociation Kit (Militenyi Biotec, 130-096-348) and stained with an allophycocyanin (APC)-conjugated mouse anti-CD38 antibody (Abcam, Ab134399, dilution 1:20) to collect $CD38^+$ hPGCLCs by using FACSAria Fusion (BD Biosciences).

Preparation of hPGCLC maintenance medium (S-CM medium)

The hPGCLC basal medium contained 13% (v/v) KSR, 1× NEAA, 1 mM sodium pyruvate, and 1× penicillin-streptomycin in Glasgow's MEM with 2 mM glutamine (ThermoFisher, 11710035). STO-CM was prepared by maintaining 5.0×10^6 mitomycin C-treated STO cells in 12 mL of hPGCLC basal medium for 24 h, removing cells by centrifugation, and storing frozen at −20°C until use. The complete hPGCLC maintenance medium (S-CM for SCF-supplemented CM) was prepared by adding 0.1 mM β-mercaptoethanol, 50 μg/mL L-ascorbic acid, and 100 ng/mL recombinant human SCF to the CM.

hPGCLC expansion culture

FACS-enriched $CD38^+$ hPGCLCs (500–2,000 cells) were inoculated onto STO feeder layers ($8\text{-}9 \times 10^4$ cells/cm^2) in a well of six-well plates (Corning, 3506) with SCF-supplemented hPGCLC basal medium containing 10 μM Y27632 dihydrochloride. Medium was changed every other day without Y27632. During the initial phase of hPGCLC expansion, cells were dissociated using Accutase and passaged onto fresh STO feeder layers at every 5–14 days without increasing the number of wells.

When hPGCLCs expanded to ~20% confluence, cells were dissociated using Accutase and incubated in wells coated with 0.1% gelatin (STEMCELL, 07903) at 37°C for 30 min to remove STO cells, which rapidly adhere to the wells. The feeder-diminished hPGCLCs were transferred onto Matrigel-coated wells in the S-CM and passaged using Accutase. hPGCLCs were also inoculated on wells coated with laminin (Millipore, cc160) or vitronectin (Millipore, cc130).

hEGCLC derivation from hPGCLCs

The hEGCLC derivation medium consisted of the hPGCLC basal medium, 100 ng/mL recombinant human SCF (Peprotech, 300-07), and 20 ng/mL recombinant human FGF2 (R&D Systems, 4114-TC). LTC-hPGCLCs were inoculated into Matrigel-coated six well plates (1,000–2,000 cells/cm^2) and maintained in the hEGCLC derivation medium for 10 days with one-half of the medium changed every other day. Thereafter, cells were maintained in the mTeSR plus medium for at least 10 days before use.

Data availability

Bulk RNA-seq, single-cell RNA-seq, WGBS, and digital karyotyping whole genome sequencing data generated in this study are available at NCBI Gene Expression Omnibus and Sequence Read Archive: GSE174485.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.stemcr.2022.01.012.

AUTHOR CONTRIBUTIONS

Conceptualization, Mu.K. and T.S. Investigation, Mu.K., M. Kobayashi, K. Shioda, J.O., A.R.L., and T.S. Software and formal analysis, Mu.K. and P.C. Writing – original draft, Mu.K. and T.S. Writing – review and editing, Mu.K., Y.S.H., K. Sasaki, J.G., C.L., and T.S. Resources, Mu.K., Y.S.H., K.Sasaki, C.K., R.E.K., H.J., and T.S. Supervision, R.E.K., G.M.C., R.S.W., K.Sasaki, C.L., and T.S. Funding Acquisition, T.S.

CONFLICTS OF INTEREST

The authors declare no competing interests.

ACKNOWLEDGMENTS

We thank Shino Mitsunaga, Shiomi Yawata, Chie Owa, Maulik Vyas, and the MGH FACS core for technical assistance. We appreciate the Singulomics Corporation for single-cell RNA-seq and the Progressive Assessment of Therapeutics (PATh) PDX Facility at Cornell University College of Veterinary Medicine for PDX support. This work was supported by gifts from The RICBAC Foundation and The Escher Fund for Autism to T.S. The work was also supported in part by NIH grants R21ES024861, R01ES020454, R01ES023316, and R01ES031139 to T.S. and NIH fellowship F30CA247458 to A.R.L.

Received: July 22, 2021
Revised: January 13, 2022
Accepted: January 14, 2022
Published: February 10, 2022

REFERENCES

Chen, D., Gell, J.J., Tao, Y., Sosa, E., and Clark, A.T. (2017). Modeling human infertility with pluripotent stem cells. Stem Cell Res. *21*, 187–192.

Chen, D., Sun, N., Hou, L., Kim, R., Faith, J., Aslanyan, M., Tao, Y., Zheng, Y., Fu, J., Liu, W., et al. (2019). Human primordial germ cells are specified from lineage-primed progenitors. Cell Rep *29*, 4568–4582 e4565.

Gell, J.J., Liu, W., Sosa, E., Chialastri, A., Hancock, G., Tao, Y., Wamaitha, S.E., Bower, G., Dey, S.S., and Clark, A.T. (2020). An extended culture system that supports human primordial germ cell-like cell survival and initiation of DNA methylation erasure. Stem Cell Rep. *14*, 433–446.

Gkountela, S., Zhang, K.X., Shafiq, T.A., Liao, W.W., Hargan-Calvopiña, J., Chen, P.Y., and Clark, A.T. (2015). DNA demethylation dynamics in the human prenatal germline. Cell *161*, 1425–1436.

Hayashi, K., Ogushi, S., Kurimoto, K., Shimamoto, S., Ohta, H., and Saitou, M. (2012). Offspring from oocytes derived from in vitro primordial germ cell-like cells in mice. Science (New York, N.Y.) *338*, 971–975.

Hwang, Y.S., Suzuki, S., Seita, Y., Ito, J., Sakata, Y., Aso, H., Sato, K., Hermann, B.P., and Sasaki, K. (2020). Reconstitution of prospermatogonial specification in vitro from human induced pluripotent stem cells. Nat. Commun. *11*, 5656.

Irie, N., Weinberger, L., Tang, W.W., Kobayashi, T., Viukov, S., Manor, Y.S., Dietmann, S., Hanna, J.H., and Surani, M.A. (2015). SOX17 is a critical specifier of human primordial germ cell fate. Cell *160*, 253–268.

Ishikura, Y., Ohta, H., Sato, T., Murase, Y., Yabuta, Y., Kojima, Y., Yamashiro, C., Nakamura, T., Yamamoto, T., Ogawa, T., and Saitou, M. (2021). In vitro reconstitution of the whole male germ-cell development from mouse pluripotent stem cells. Cell Stem Cell *28*, 2167–2179 e2169.

Kojima, Y., Yamashiro, C., Murase, Y., Yabuta, Y., Okamoto, I., Iwatani, C., Tsuchiya, H., Nakaya, M., Tsukiyama, T., Nakamura, T., et al. (2021). GATA transcription factors, SOX17 and TFAP2C, drive the human germ-cell specification program. Life Sci. Alliance *4*, e202000974.

Li, L., Dong, J., Yan, L., Yong, J., Liu, X., Hu, Y., Fan, X., Wu, X., Guo, H., Wang, X., et al. (2017). Single-cell RNA-seq analysis maps development of human germline cells and gonadal niche interactions. Cell Stem Cell *20*, 891–892.

Mitsunaga, S., Odajima, J., Yawata, S., Shioda, K., Owa, C., Isselbacher, K.J., Hanna, J.H., and Shioda, T. (2017). Relevance of iPSC-derived human PGC-like cells at the surface of embryoid bodies to prechemotaxis migrating PGCs. Proc. Natl. Acad. Sci. U S A. *114*, E9913–E9922.

Mitsunaga, S., Shioda, K., Hanna, J.H., Isselbacher, K.J., and Shioda, T. (2021). Production and analysis of human primordial germ cell-like cells. Methods Mol. Biol. *2195*, 125–145.

Mitsunaga, S., Shioda, K., Isselbacher, K.J., Hanna, J.H., and Shioda, T. (2019). Generation of human primordial germ cell-like cells at the surface of embryoid bodies from primed-pluripotency induced pluripotent stem cells. J. Vis. Exp https://doi.org/10.3791/58297.

Miyoshi, N., Stel, J.M., Shioda, K., Qu, N., Odajima, J., Mitsunaga, S., Zhang, X., Nagano, M., Hochedlinger, K., Isselbacher, K.J., and Shioda, T. (2016). Erasure of DNA methylation, genomic imprints, and epimutations in a primordial germ-cell model derived from mouse pluripotent stem cells. Proc. Natl. Acad. Sci. U S A. *113*, 9545–9550.

Murase, Y., Yabuta, Y., Ohta, H., Yamashiro, C., Nakamura, T., Yamamoto, T., and Saitou, M. (2020). Long-term expansion with germline potential of human primordial germ cell-like cells in vitro. EMBO J. *39*, e104929.

Nicholls, P.K., Schorle, H., Naqvi, S., Hu, Y.C., Fan, Y., Carmell, M.A., Dobrinski, I., Watson, A.L., Carlson, D.F., Fahrenkrug, S.C., and Page, D.C. (2019). Mammalian germ cells are determined after

PGC colonization of the nascent gonad. Proc. Natl. Acad. Sci. U S A. *116*, 25677–25687.

Ohta, H., Kurimoto, K., Okamoto, I., Nakamura, T., Yabuta, Y., Miyauchi, H., Yamamoto, T., Okuno, Y., Hagiwara, M., Shirane, K., et al. (2017). In vitro expansion of mouse primordial germ cell-like cells recapitulates an epigenetic blank slate. EMBO J. *36*, 1888–1907.

Oosterhuis, J.W., and Looijenga, L.H.J. (2019). Human germ cell tumours from a developmental perspective. Nat. Rev. Cancer *19*, 522–537.

Saitou, M. (2021). Mammalian germ cell development: from mechanism to in vitro reconstitution. Stem Cell Rep. *16*, 669–680.

Sasaki, K., Nakamura, T., Okamoto, I., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Shiraki, N., Takakuwa, T., et al. (2016). The germ cell fate of cynomolgus monkeys is specified in the nascent amnion. Dev. Cell *39*, 169–185.

Sasaki, K., Yokobayashi, S., Nakamura, T., Okamoto, I., Yabuta, Y., Kurimoto, K., Ohta, H., Moritoki, Y., Iwatani, C., Tsuchiya, H., et al. (2015). Robust in vitro induction of human germ cell fate from pluripotent stem cells. Cell Stem Cell *17*, 178–194.

Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W., and Reik, W. (2012). The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol. Cell *48*, 849–862.

Seki, Y., Yamaji, M., Yabuta, Y., Sano, M., Shigeta, M., Matsui, Y., Saga, Y., Tachibana, M., Shinkai, Y., and Saitou, M. (2007). Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. Development *134*, 2627–2638.

Talbot, N.C., Sparks, W.O., Powell, A.M., Kahl, S., and Caperna, T.J. (2012). Quantitative and semiquantitative immunoassay of growth factors and cytokines in the conditioned medium of STO and CF-1 mouse feeder cells. In Vitro Cell Dev Biol Anim *48*, 1–11.

Tang, W.W., Dietmann, S., Irie, N., Leitch, H.G., Floros, V.I., Bradshaw, C.R., Hackett, J.A., Chinnery, P.F., and Surani, M.A. (2015). A unique gene regulatory network resets the human germline epigenome for development. Cell *161*, 1453–1467.

Tyser, R.C.V., Mahammadov, E., Nakanoh, S., Vallier, L., Scialdone, A., and Srinivas, S. (2021). Single-cell transcriptomic characterization of a gastrulating human embryo. Nature *600*, 285–289.

Yamashiro, C., Sasaki, K., Yabuta, Y., Kojima, Y., Nakamura, T., Okamoto, I., Yokobayashi, S., Murase, Y., Ishikura, Y., Shirane, K., et al. (2018). Generation of human oogonia from induced pluripotent stem cells in vitro. Science *362*, 356–360.

Stem Cell Reports, Volume *17*

Supplemental Information

Expanding homogeneous culture of human primordial germ cell-like cells maintaining germline features without serum or feeder layers

Mutsumi Kobayashi, Misato Kobayashi, Junko Odajima, Keiko Shioda, Young Sun Hwang, Kotaro Sasaki, Pranam Chatterjee, Christian Kramme, Richie E. Kohman, George M. Church, Amanda R. Loehr, Robert S. Weiss, Harald Jüppner, Joanna J. Gell, Ching C. Lau, and Toshi Shioda

Kobayashi et al., Figure S1

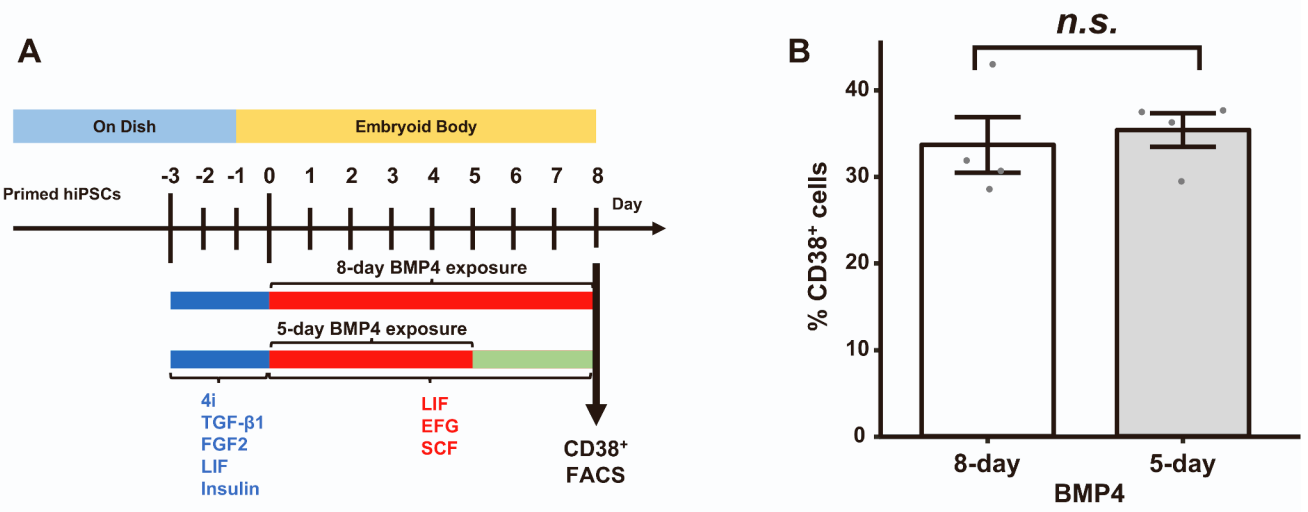


Figure S1, Related to Figure 1.

Efficient hPGCLC derivation with shortened exposure to BMP4

Efficiency of hPGCLC derivation was compared between full-length (8-day) and shortened (5-day) exposure of embryoid bodies to BMP4. (A) Diagram of the 11- day schedule of hPGCLC production from primed pluripotency hiPSCs. Timing of full-length or shortened BMP4 exposure is shown by red bars. (B) FACS analysis of $CD38^{+}$ hPGCLC population in single cell suspensions of embryoid bodies (mean ± SEM of four independent experiments). Data were pooled from eight independent experiments. n.s., not significant.

Kobayashi et al., Figure S2

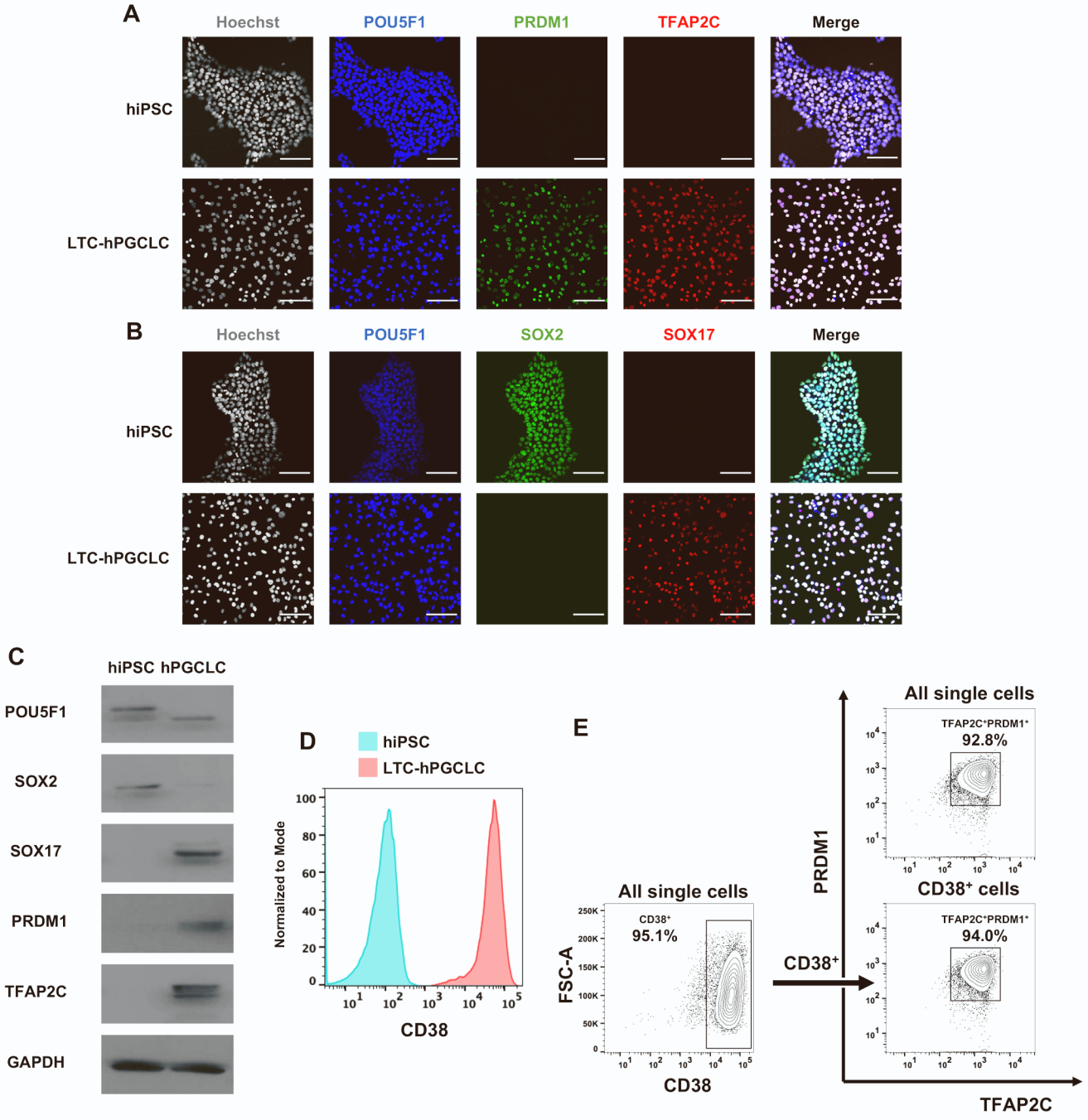


Figure S2, Related to Figure 2.

Expression of pluripotency or germline marker transcription factors in LTC-hPGCLCs.

(A, B) Multiplexed immunofluorescence images of hiPSCs and LTC-hPGCLCs. Nuclei were visualized by Hoechst staining (gray). (C) Western blotting detection of marker transcription factors in total cell lysates of hiPSCs and LTC-hPGCLCs. (D) FACS analysis of CD38 expression at cell surface. Formaldehyde-fixed hiPSCs (blue) and LTC-hPGCLCs (red) were stained with fluorescence-labeled antibodies without permeabilization. Fluorescence intensities (x-axis) and relative numbers of cells (y-axis in arbitrary units) were determined by FACS. (E) Expression of PRDM1 and TFAP2C in $CD38^+$ LTC-hPGCLCs. Permeabilized single cell suspension was subjected to immunofluorescence staining for TFAP2C and PRDM1 (*right upper*). Non-permeabilized cells were subjected to CD38 cell surface staining, and the enriched $CD38^+$ cells (*left lower*) were permeabilized and stained for PRDM1 and TFAP2C (*right lower*). Expression of the marker proteins in cell populations was determined by FACS. (F) Expression of hPGCLC markers PRDM1 and TFAP2C in hiPSCs and LTC-hPGCLCs derived from hiPSC clones A5 (153 days of culture), F2 (126 days), and 9A13 (70 days). (G) CaspaTag evaluation of apoptosis. LTC-hPGCLCs were exposed to 100 μM staurosporine (ss, *red*) or vehicle (*blue*) for 3 hours followed by the CaspaTag fluorescence assay of caspases 3/7 enzymatic activities. Fluorescence intensities (x-axis) and numbers of cells (y-axis) were determined by FACS, and proportions of apoptotic cells are shown.

Kobayashi et al., Figure S3

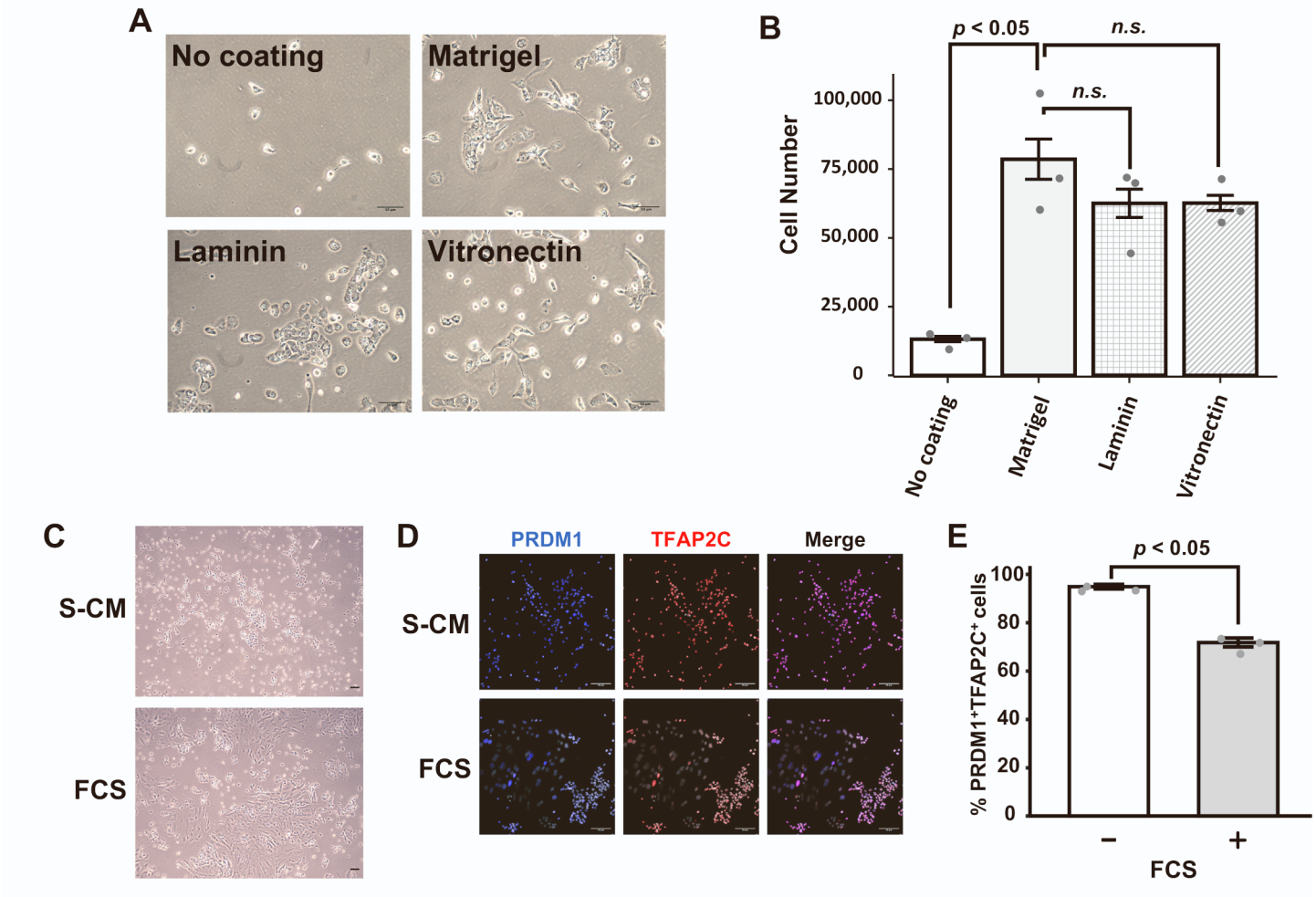


Figure S3, Related to Figure 2

Impaired PGC identify of LTC-hPGCLCs in the presence of fetal calf serum

(A) Phase contrast images of LTC-hPGCLCs cultured on wells coated with indicated extracellular matrix proteins (scale bar = 50 μm). (B) Growth of LTC-hPGCLCs expanded on extracellular matrix proteins in 7-day culture (mean ± SEM). 40,000 LTC-hPGCLCs were seeded to each wells and cultured each condition for 7 days. Data were pooled from three independent experiments. *n.s.*, not significant. (C, D) Feeder-free culture of LTC-hPGCLCs derived from A4 male hiPSCs were expanded in the S-CM long-term hPGCLC maintenance medium and subjected to an additional 7-day culture in the presence or absence of 2.5% fetal calf serum (FCS), followed by phase contrast imaging (C; scale bar = 50 μm) or immunofluorescence staining of PGC marker transcription factorsPRDM1 and TFAP2C (D; scale bar =50 μm, nuclei were visualized as gray by Hoechst staining). Note the emergence of $PRDM1^{-}$ /$TFAP2C^{-}$ double-negative cells with larger nuclei in the culture supplemented with FCS. (E) Quantification of PRDM1/TFAP2C double-positive cell population after 7-day culture in the FCS-supplemented medium (mean ± SEM). Data were pooled from three independent experiments.

Kobayashi et al., Figure S4

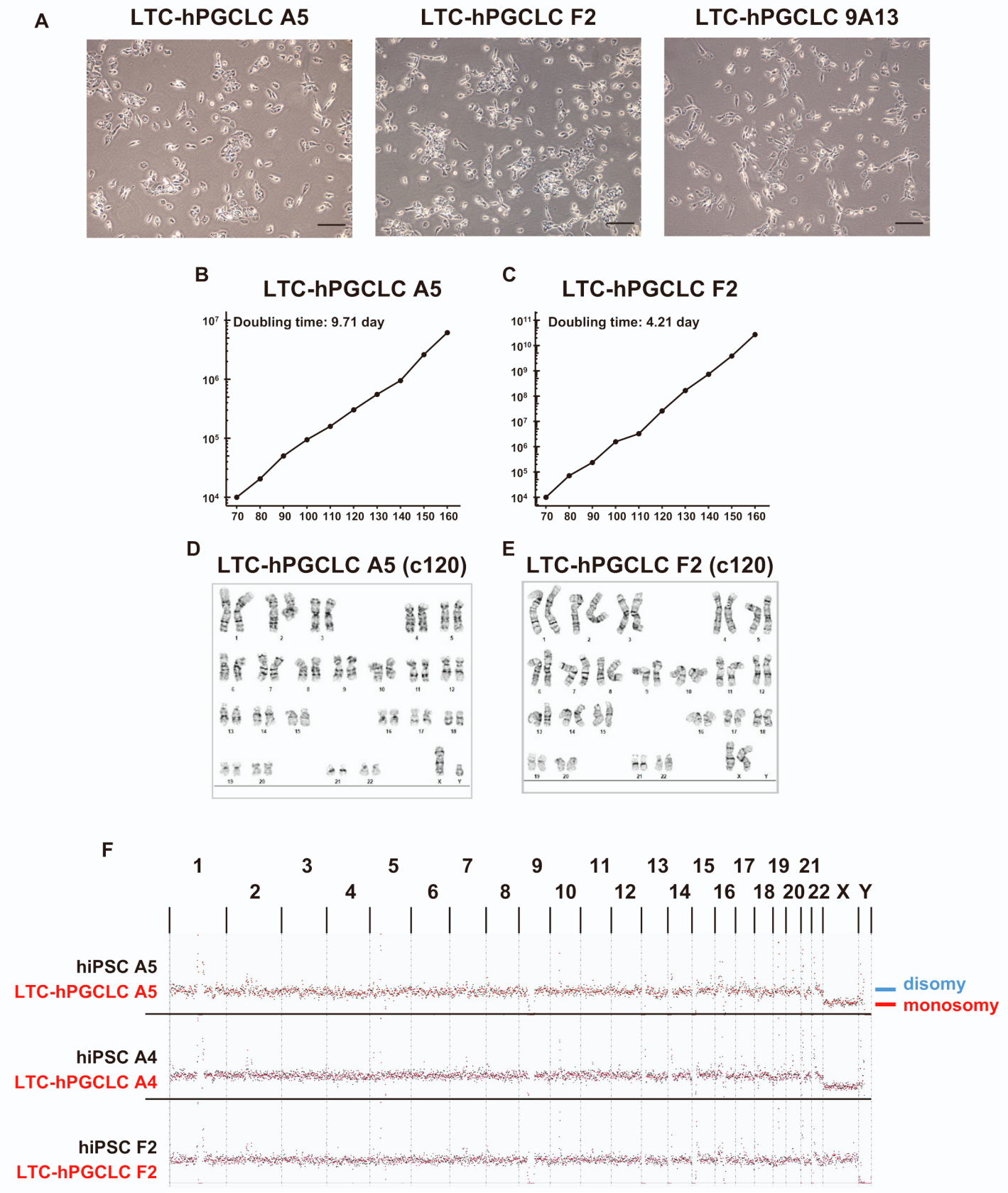


Figure S4, Related to Figure 2

Characteristics of LTC-hPGCLCs derived from hiPSC clones (A5, F2, and 9A13)

(A) Phase-contrast image of feeder-free culture at day 80. (B, C) Growth curve of feeder-free LTC-hPGCLCs supported by S-CM. LTC-hPGCLC A5 (B) and F2 (C) were pre-expanded for 70 days and then expanded additional 90 days in culture. (D, E) Normal diploid karyotype of A5 (D) and F2 (E) after 100 days of feeder-free expansion: G-banding image. (F) Digital karyotyping by whole-genome sequencing. Black and red dots represent hiPSC and LTC-hPGCLCs, respectively. Note that strong deviations from the disomy line in the autosomes observed for all cell types reflect regions of highly repetitive sequences, where digital karyotyping may not represent correct chromosomal copy numbers.

Kobayashi et al., Figure S5

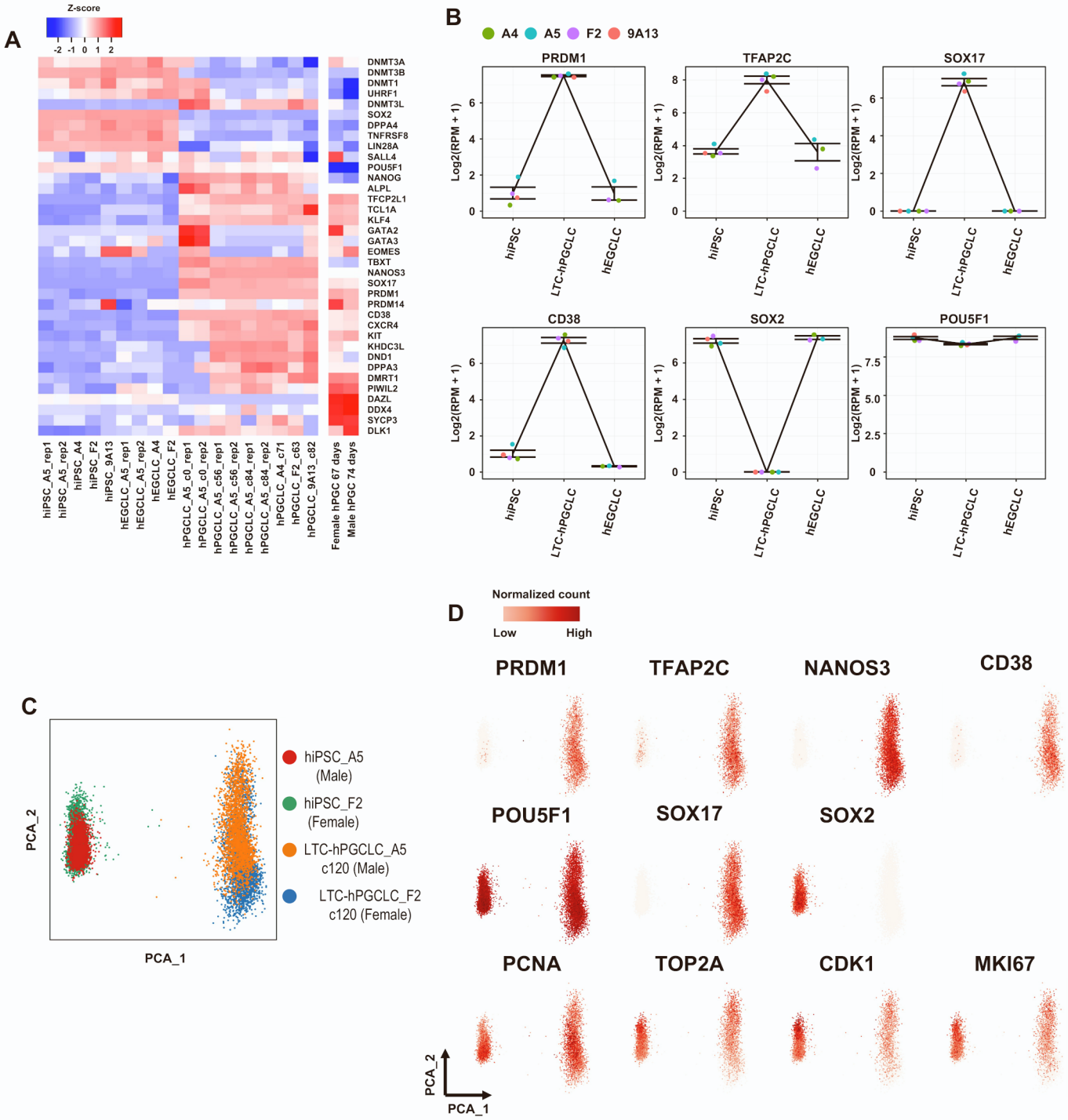
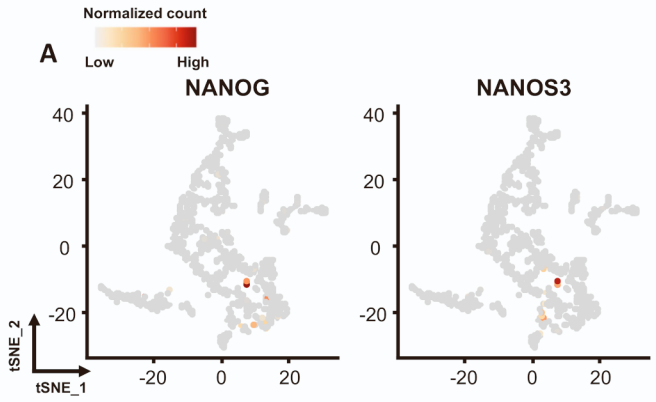


Figure S5, Related to Figure 4

Transcriptomic profiles of hiPSCs, hPGCLCs, and hEGCs.

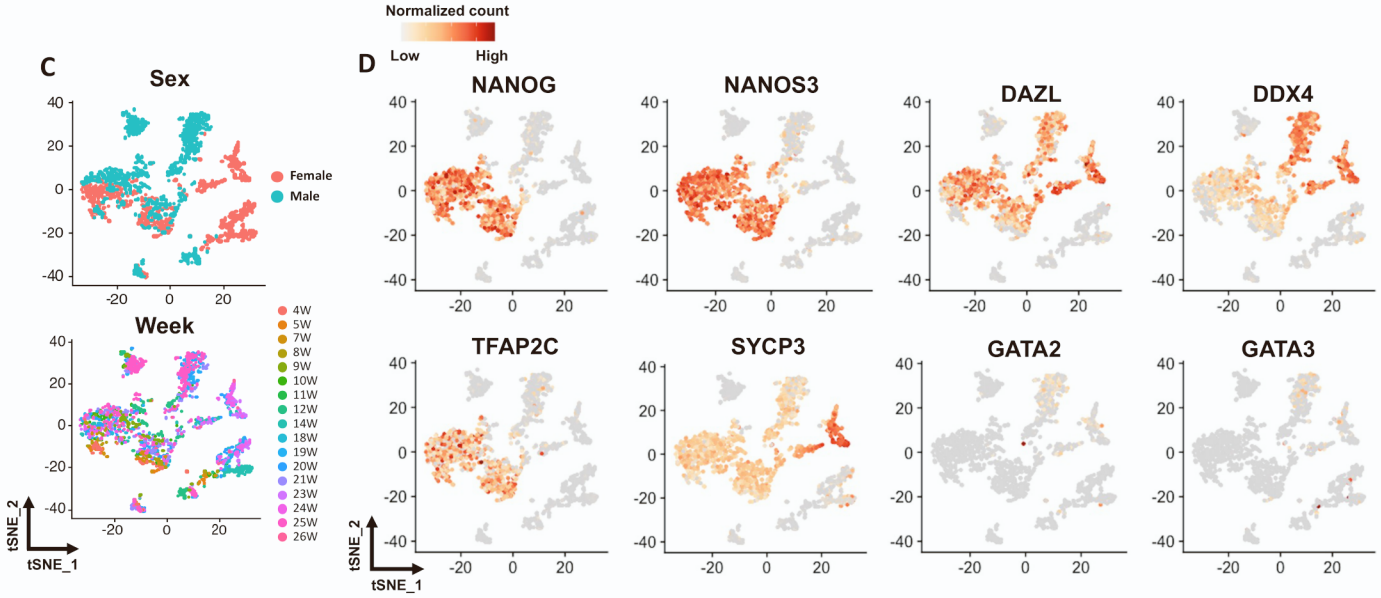
(A) Heatmap representation of marker gene expression in hiPSCs, hPGCLCs, hEGCs derived from four cell line, and human fetal PGCs. Data of human embryonic PGCs (hPGCs, gestational days 113 and 122) were from a previously published study (Gkountela *et al.,* 2015). (B) Gene expression dynamics in hiPSCs, LTC-hPGCLCs, and hEGCs in three independent iPS cell line (A4, A5, F2, and 9A13). (C, D) Single cell RNA-seq PCA clustering plot (C) and superimposed expression of pluripotency, hPGCLC, and cell cycle marker genes (D). (E) Violin plot representation of increased expression of genes located on X chromosome in female cells over male cells. Genes were filtered for significant expression (normalized counts >3 per million reads) and subjected to calculation of female/male expression rate. Whereas autosomal genes show nearly identical strength of expression between male and female cells, X chromosomal genes are more strongly expressed in females than males. Asterisk, $p < 0.05$ (*t*-test).

Kobayashi et al., Figure S6



B

	cell_1	cell_2	cell_3	cell_4	cell_5	cell_6	cell_7	cell_8
NANOG	4.43	4.54	1.03	4.32	2.22	3.64	5.27	3.48
POU5F1	5.89	5.61	4.59	4.33	6.43	6.55	6.46	0.30
NANOS3	6.37	7.09	0.69	5.29	6.84	4.13	5.16	6.65
TFAP2C	0.00	0.00	0.00	0.00	0.00	2.05	3.46	1.58
PRDM1	0.00	0.00	0.00	0.00	0.00	2.05	0.00	0.00
CD38	0.00	3.70	0.00	0.00	4.27	0.00	1.01	0.00
GATA2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GATA3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DAZL	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DDX4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SYCP3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00



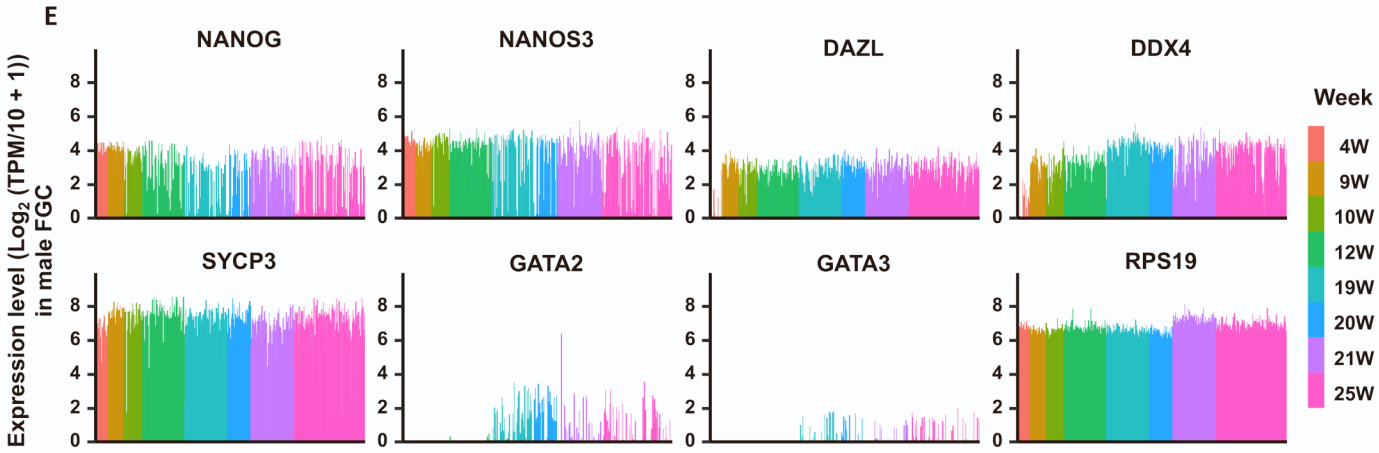
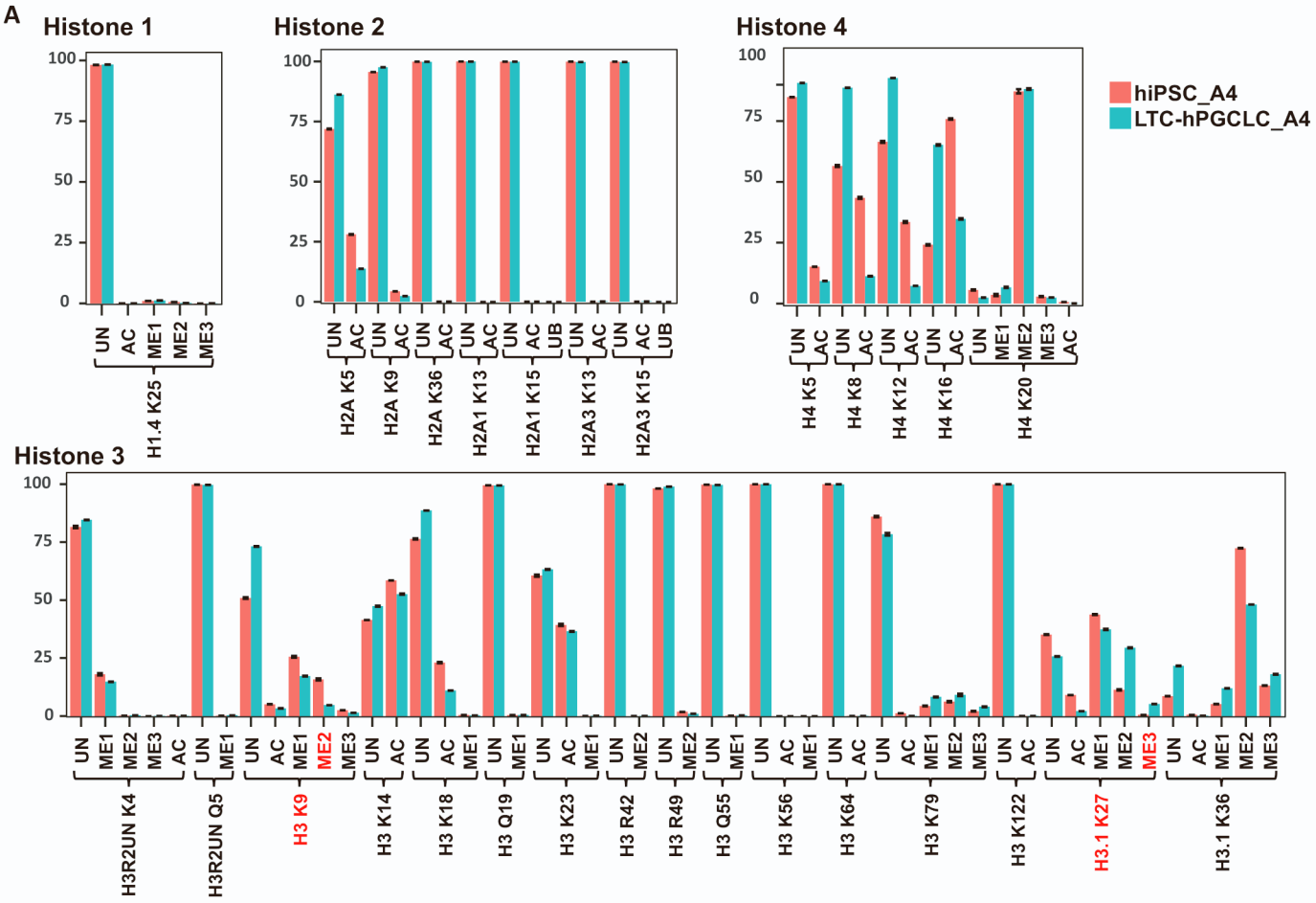


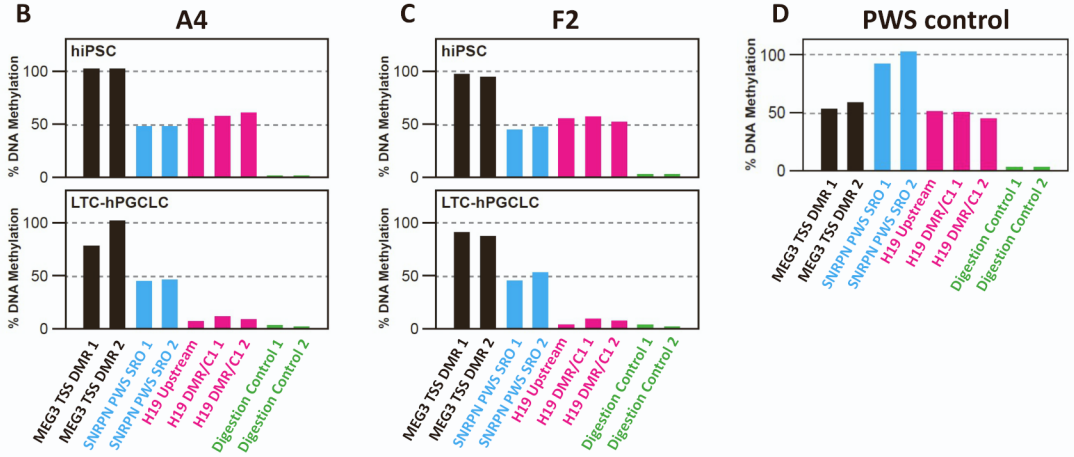
Figure S6, Related to Figure 5

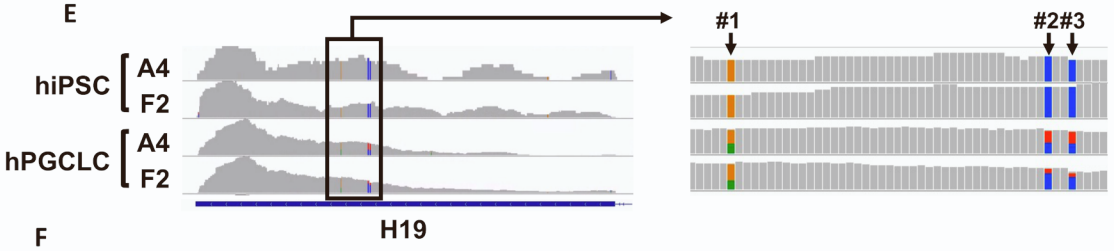
Reanalysis of online available human fetal germ cell dataset

(A, B) NANOG and NANOS3 expression in single cell RNAseq tSNE clustering plots (A) and table of pluripotency and germ cell marker gene expression in NANOG/NANOS3 double positive cells normalized by log2(TPM/10 + 1) (B) were from a previously published study (Tyser et al., 2021). (C, D) Single cell RNAseq tSNE clustering plots colored by sex (C, top), weeks of post-fertilization (C, bottom), and pluripotency and germ cell marker gene expression (D). (E) Histograms for the expression levels normalized by log2(TPM/10 + 1) of selected marker genes in male FGCs. Figure S6C, S6D and S6E were from previous published study (Li et al., 2017).

Kobayashi et al., Figure S7







F

		SNP #1	SNP #2	SNP #3
hiPSC	A4	G (100%; 6/6)	C (100%; 7/7)	C (100%; 6/6)
	F2	G (100%; 9/9)	C (100%; 12/12)	C (100%; 12/12)
hPGCLC	A4	G (56%; 41/73) A (44%; 32/73)	C (47%; 34/72) T (51%; 37/72) A (1%; 1/72)	C (44%; 30/68) T (54%; 37/68) N (1%; 1/68)
	F2	G (64%; 61/96) A (36%; 35/96)	C (78%; 62/80) T (23%; 18/80)	C (74%; 49/66) T (24%; 16/66) N (2%; 1/66)

Figure S7, Related to Figure 5

Comprehensive analysis of histone modification by histone proteomics and biallelic expression of the *H19* imprinting gene in LTC-hPGCLCs and hEGCLCs

(A) Relative amounts of modified histones in hiPSCs and day-136 LTC-hPGCLCs were determined for the four classes of histones (error bar = SD). H3K9me2 and H3K27me3 are shown in red. UN, unmodified; AC, acetylated; ME1-ME3, mono-trimethylated. UB, ubiquitilated. (B, C, D) CpG methylation quantified by Methylation-Specific Multiplex Ligation-Dependent Probe Amplification (MS-MLPA) for genomic DNA isolated from hiPSCs clone A4 (B) and F2 (C), and LTC-hPGCLCs derived from them (A4 and F2 LTC-hPGCLCs were expanded in cell culture for 71 and 63 days, respectively). The two C1 probes of H19 DMR reflect DNA methylation at the H19-IGF2 imprinting control region (ICR), and the remaining H19 probe reflects methylation at the upstream DMR. Genomic DNA of a Prader-Willi Syndrome (PWS) patient is included in the MS-MLPA analysis as positive control of aberrant hypermethylation. (E) RNA-seq tracks of hiPSCs, and LTC-hPGCLCs at *H19* loci. Right panel magnifies a part of the left panel and indicates three single nucleotide polymorphisms (SNPs) (#1-3) distinguishing the two loci. (F) Detailed RNA-seq read counts at the SNPs.

Supplemental Experimental Procedures

Cell lines

The A4 and A5 human iPSC (hiPSC) clones were generated from commercially obtained Caucasian male neonatal foreskin fibroblasts using a plasmid-based, footprint-free reprogramming method as we previously described (Mitsunaga et al., 2017). The ACS-1029 hiPSC clone purchased from American Type Culture Collection (ATCC) and mentioned as F2 hiPSCs in this study. The 9A13 hiPSC clone was previously described (Hwang et al., 2020). All hiPSC cultures were maintained in the mTeSR Plus medium (STEMCELL Technologies, 100-0276) on human ESC-qualified Matrigel (Corning, 354277) and passaged at every 3-7 days using Accutase (Innovative Cell Technologies, AT104-500) or TrypLE Select (Thermo Fisher, 12563011) in the presence of the ROCK inhibitor Y-27632 (Axon, 1683). The normal diploid karyotypes of all hiPSC clones (46XY for A4 and A5, 46XX for F2) were periodically confirmed during the study by deep sequencing digital karyotyping. The STO cell line was obtained from ATCC (CRL-1503) and maintained in DMEM supplemented with 1x penicillin/streptomycin and 10% fetal calf serum (FCS). To prepare feeder layers, STO cells were exposed to 10 μg/mL Mitomycin C (MMC; Sigma, M0503) for 4 hours at 37 °C, extensively washed, and inoculated at 8×10^4 cells/cm^2. All cell cultures were subjected to periodical PCR-based Mycoplasma testing using the MycoDect kit (Alstem, MD050) and confirmed negative.

Generation of EGFP-labelled hiPSCs

The AAVS1-Pur-CAG-EGFP donor plasmid for knock-in a CAG promoter-driven EGFP expression cassette at the AAVS1 human genome safe harbor site was a gift from Su-Chun Zhang [Addgene plasmid # 80945 (Chen *et al*., 2016)]. The plasmid was introduced into

hiPSCs with the TrueCut Cas9 protein v2 (Thermo Fisher, A36498) and a guide RNA (gRNA sequence: ggggccactagggacaggat) using the NEON electroporation system (Thermo Fisher, A36498). Cells stably expressing EGFP were enriched by FACS and cloned using the CloneR reagent (STEMCELL).

Fluorescence-activated cell sorting (FACS)

Single cell suspensions were prepared using Accutase and subjected to fixation and permeabilization as described earlier. Cells were incubated with primary antibodies: goat anti-SOX17 (R&D, AF1924; 1:1000 dilution), rabbit anti-PRDM1 (CST, 9115S; 1:100), mouse anti-TFAP2C, Alexa fluor 594 conjugated (Santa Cruz, sc-12762; 1:100), mouseanti-OCT-3/4, phycoerythrin conjugated (Santa Cruz, sc-5279; 1:100), or mouse anti-CD38, allophycocyanin conjugated (Abcam, Ab134399; 1:20). Secondary antibodies were donkey anti-goat Ig, Alexa fluor 647 conjugated (Abcam, ab150135; 1:5000 dilution) or donkey anti-rabbit Ig, Brilliant violet 510 conjugated (Biolegend, 406419; 1:100). FACS gating was determined for each experiment based on unstained control, single staining compensation control, and fluorescence minus control. FACS data were analysed using the FlowJo software.

Immunofluorescence

Cells were grown on Matrigel-coated chamber slides (Celltreat, 229168), fixed in 4% electron microscopy-grade formaldehyde in PBS(-) for one hour at 4˚C, and washed three times with PBS(-). Fixed slides were permeabilized with a buffer containing 0.1% triton-X 100 and 0.1% bovine serum albumin in PBS(-) for 10 minutes, blocked with 5% donkey serum (Abcam, ab7475) and incubated with primary antibodies: goat anti-SOX17 (R&D,

AF1924; 1:2000 dilution), rabbit anti-PRDM1 (CST, 9115S, 1:100), mouse anti-TFAP2C (Santa Cruz, sc-12762; 1:100), rabbit anti-SOX2 (Reprocell, 09-0024; 1:100), or goat anti-OCT-3/4 (Santa Cruz, sc-8629; 1:1000). Slides were then washed and stained with 1:500-diluted secondary antibodies: donkey anti-rabbit Ig, Alexa fluor 488 conjugated (Abcam, ab150061), donkey anti-goat Ig, Alexa fluor 568 (Abcam, ab175704), or donkey anti-mouse Ig, Alexa fluor 647 (Abcam, ab150111). Nuclei were stained with Hoechst33342 (Thermo Fisher).

For immunofluorescence detection of H3K27me3 and H3K9me2, LTC-hPGCLCs and hiPSCs were mixed at 1:1 ratio and fixed to obtain clear contrast between these two cell types. Primary antibodies were rabbit anti-H3K27me3 (Millipore, 07-449; 1:50) and mouse anti-H3K9me2 (Abcam, ab1220; 1:100). Secondary antibodies and nuclear staining were performed as described earlier. Stained cells were inspected on glass slides using an LSM710 confocal microscope (Zeiss), and electronic color images were processed using the Image J Fiji image analysis software.

Xenogeneic reconstituted testis (xrTestis)

The xrTestis organ culture and immunofluorescence staining were performed as described by Hwang et al. (Hwang *et al.*, 2020). LTC-hPGCLCs derived from A5 hiPSCs were expanded for 120 days and co-aggregated with E12.5 mouse embryonic testicular somatic cells depleted of mouse PGCs using anti-SSEA1 antibody microbeads and MACS (Miltenyi Biotec). For immunofluorescence, cryosections were stained with a rabbit anti-DAZL primary antibody (abcam, ab215718) and an anti-rabbit secondary antibody conjugated with AF568 (abcam, ab175693). Specific staining of DAZL expressed in human germline cells was confirmed by preliminary experiments using cryosections of normal adult human testes.

Cellular nuclei were visualized by counter-staining with DAPI.

Apoptosis assay

Apoptosis of LTC-hPGCLCs was evaluated using the CaspTag kit (Millipore, APT423) and FACS with the FITC gate as described (Mazumder et al., 2008). Apoptosis of LTC-hPGCLCs was induced by exposing cells to 100 μM staurosporine for 3 hours.

Teratoma formation assay

Transplantation of hEGCLCs into the immunocompromised NSG mice [NOD.Cg-*Prkdc*^{scid}*Il2rg*^{tm1Wjl}/SzJ, Jackson Laboratory Stock# 005557] was performed as we previously described (Loehr et al., 2021; Pierpont et al., 2017). Briefly, 1 million hEGCLCs were suspended in PBS with 10 μM ROCK inhibitor Y-27632. Immediately prior to injection, the cell suspension was mixed with an equal volumes of Matrigel (Corning, 354277) and injected subcutaneously into the flanks of the mice. Mice were monitored regularly for tumor development and sacrificed upon reaching humane endpoint criteria. Tumors were collected by dissection, fixed 4% paraformaldehyde, and subjected to the standard H&E staining.

Western blotting

Total proteins were extracted from cells using the RIPA buffer containing the cOmplete protease inhibitor cocktail (Roche, 11836153001). Histones were extracted by the acid extraction method. Briefly, cells were incubated in PBS containing 0.5% (v/v) Triton-X 100, 2 mM phenylmethylsulfonyl fluoride, and 0.02% (w/v) sodium azide (10^7 cells/mL) for 10 minutes on ice with gentle stirring. After washing with the same buffer, the pre-extracted cells

were incubated in 0.2N HCl (4 ×10^7 cells/mL) overnight on ice to solubilize histones, insoluble debris was removed by centrifugation (6,500x *g* for 10 minutes at 4 ℃), and supernatant was neutralized by adding 1/10 volume of 2 N NaOH.

Protein concentrations of the total or histone-enriched samples were determined by the Bradford method using the Protein Assay Dye Reagent Concentrate (Bio-Rad, 5000006). Equal amounts of protein were subjected to 10% SDS-PAGE and transferred onto PVDF membranes (Thermo Fisher, LC2002). After blocking in 5% skim milk for 1 h at room temperature, membranes were incubated with primary antibodies overnight at 4 °C: goat anti- OCT-3/4 (Santa Cruz, sc-8629; 1:500), rabbit anti-SOX2 (Reprocell, 09-0024; 1:1000), mouse anti-TFAP2C (Santa Cruz, sc-12762; 1:500), goat anti-SOX17 (R&D, AF1924; 1:1000), rabbit anti-PRDM1 (CST, 9115S; 1:1000), mouse anti-GAPDH (Sigma-Aldrich, G8795; 1:20000), rabbit anti-H3K27me3 (Millipore Sigma, 07-449; 1:1000), mouse anti-H3K9me2 (Abcam, ab1220; 1:1000), or rabbit anti-histone H3 (Abcam, ab1791; 1:20000). Membranes were incubated with horseradish peroxidase-conjugated secondary antibodies for 1 h at room temperature: anti-goat Ig (Santa Cruz, sc-2768; 1:3000), anti-rabbit Ig (Santa Cruz, sc2357; 1:3000), or anti-mouse Ig (Santa Cruz, sc-516102; 1:3000). Protein bands were detected by chemiluminescence using the SuperSignal West Pico Plus (Thermo Fisher, 34579) or Femto kits (Thermo Fisher, 34094).

RNAseq library preparation

Single cell suspensions of hiPSCs and hEGCLCs were prepared using Accutase, pelleted by centrifugation and snap frozen. $CD38^+$ c0-hPGCLCs and LTC-hPGCLCs were pelleted and frozen immediately after FACS isolation. cDNA was synthesised directly from pellets of 1,000 cells using the SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (Takara,

634890), purified using the AMPure XP beads (Beckman Coulter, A63881), and quantified by Tapestation using the High Sensitivity D1000 DNA screen tapes (Agilent, 5067-5584 and 5067-5585). Deep sequencing libraries were synthesized from the purified cDNA using the Nextera XT DNA Library Prep Kit (Illumina, FC-131-1096), purified using AMPure XP beads, and subjected to insert size assessment using the High Sensitivity D1000 DNA screen tapes. Multiplexed libraries were quantified by qPCR using the KAPA Illumina library quantification kit (Roche, KK4824) and sequenced using Illumina NextSeq 500 (75nt + 75nt, paired-end). Fastq files were generated by Illumina BaseSpace server.

RNAseq data analysis

Adaptor sequences and low-quality reads (Phred quality score Q < 20) were trimmed from fastq reads using the Trim Galore! tool (Babraham Institute) with its default configurations. Quality control-filtered paired end reads were then aligned to the GRC38/hg38 human genome reference sequence using the STAR aligner software (Dobin et al., 2013), and the resulting bam format alignment files were sorted, indexed, and duplication-removed using the sambamba tool package (Tarasov et al., 2015). Reads aligned to exons of the hg38 gene model were counted using the Bioconductor package *Rsubread* (Liao et al., 2019), and the counts were normalized using the negative binominal trimmed mean of M-values method implemented by the Bioconductor package *edgeR* (Robinson *et al*., 2010). Dendrogram was generated by the *ggdendro* Bioconductor package. Pearson's correlation coefficients was calculated by the *cor* function of R. Scatter plots were drawn using the *ggplot2* R package (Wickham, 2009). RNAseq data of human embryonic PGCs were described by Gkountela *et al*. (Gkountela et al., 2015) and downloaded from NCBI Gene Expression Omnibus database accession number: GEO63392). Heatmap of RNAseq data was drawn using the

gplots R package for log2 (MMT-counts data + 1) values.

Single cell RNAseq

Viable cell suspensions were prepared from cryopreserved cells and loaded into the Chromium Controller (10x Genomics, Pleasanton, CA) to capture 4,000 cells from each sample. Synthesis of cDNA from poly(A)-rich mRNA and Chromium 5' Gene Expression libraries (Chromium Next GEM Single Cell 5' Reagent Kits v2 - Dual Index) was performed following the manufacturer's instructions. Amplified cDNAs and the libraries were measured by Qubit HS DNA assay (Thermo Fisher Scientific, Wilmington, DE) and quality assessed by Bioanalyzer (Agilent Technologies, Santa Clara, CA)

Quality-controlled reads were aligned to GRCh38/hg38 using Cell Ranger v6.0.1 and filtered by cell barcode via unique molecular identifiers (UMI) employing the ScanPy Python package v.1.7.2. After cells with fewer than 5000 UMI counts were eliminated for accurate clustering, reads were normalized for each cell by total expression, multiplied by 10000, and log-transformed. tSNE and PCA calculations were performed on the resulting AnnData objects using the scanpy.tl.tsne and scanpy.tl.pca functions and visualized with matplotlib.

Whole genome bisulfite sequencing (WGBS)

Genomic DNA was purified using AllPrep DNA/RNA Micro kit (QIAGEN, 80284), and WGBS libraries were synthesized from 1.5 ng purified DNA using the Pico Methyl-Seq Library Prep Kit (Zymo Research, D5455). Unmethylated Lambda phage DNA (Thermo Fisher, SD0021) was spiked-in at 0.15% to evaluate the efficiency of bisulfite conversion. Libraries were subjected to insert size assessment by Tapestation using the High Sensitivity D5000

screening tapes (Agilent) and pooled. Multiplexed libraries were quantified using the KAPA qPCR kit and sequenced using the S4 flowcells of Illumina Novaseq6000 (150 nt + 150 nt, paired-end) to obtain fastq sequence reads.

Adapter sequence, low-quality reads (Q score < 20), and five bases from the 5'-ends were trimmed from fastq reads using the Trim Galore! tool. Quality control-filtered paired end reads were aligned to GRCh38/h38 using the Bismark with the "—non_directional" and "—score_min L,0,-0.6" command line options (Krueger and Andrews, 2011). Deduplication from the aligned BAM data were also performed by Bismark. The resulting bam format files were converted to methylation levels using the bismark_methylation_extractor tool of Bismark.

Promoter regions were defined as regions between 900-bp upstream and 400-bp downstream of transcription start sites (TSSs). High-, intermediate-, and low-CpG promoters (HCP, ICP, and LCP, respectively) were calculated as described in Borgel *et al* (Borgel *et al*., 2010). The regions of imprinted differentially methylated region (DMR) described in Court *et al* were used for analysis (Court *et al*., 2014). Repetitive element information were extracted by RepeatMasker (Smit, 2013-2015). The human endogenous retrovirus classification was from Vargiu *et al* (Vargiu *et al*., 2016). CpGs with read depth between 4 and 200 were used as methylated CpGs. Only regions in which CpG counts were more than 10 within genome-wide 1-kb bins for analysis. Heatmap of WGBS data was drawn using the *gplots* R package.

Methylation-Specific Multiplex Ligation-Dependent Probe Amplification (MS-MLPA)

MS-MLPA targeting imprinting genes was performed using the ME034-B1 probe set as previously described (Moelans et al., 2018).

Histone proteomics

Quantitative proteomic analyses of histone modifications was performed as the Mod Spec service provided by Active Motif (Carlsbad, CA, USA). This is a mass spectrometry method for determination of relative abundance of 84 different histone modifications (Kumar and Elsasser, 2019). Briefly, a pure population of 5 x 10^6 LTC-hPGCLCs were harvested using Accutase, washed once with PBS, and pelleted cells were flash frozen. Histones were acid extracted, derivatized *via* propionylation, and subjected to digested with trypsin. Newly formed N-termini were propionylated and measured using the Thermo Scientific TSQ Quantum Ultra mass spectrometer coupled with an UltiMate 3000 Dionex nano-liquid chromatography system. The outcome was quantified using Skyline to obtain relative amounts within the total pool of the tryptic peptides.

Statistics

Statistical significance was tested using Student's *t*-test (two-tails) with 0.05 as the significance level.

Data and Software availability

Bulk RNAseq, single cell RNAseq, WGBS, and digital karyotyping whole genome sequencing data generated in this study are available at NCBI Gene Expression Omnibus and Sequence Read Archive databases (https://www.ncbi.nlm.nih.gov/sra) under accession GSE174485.

References

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., Forné, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. Nature genetics *42*, 1093-1100. 10.1038/ng.708.

Chen, Y., Xiong, M., Dong, Y., Haberman, A., Cao, J., Liu, H., Zhou, W., and Zhang, S.C. (2016). Chemical Control of Grafted Human PSC-Derived Neurons in a Mouse Model of Parkinson's Disease. Cell stem cell *18*, 817-826. 10.1016/j.stem.2016.03.014.

Court, F., Tayama, C., Romanelli, V., Martin-Trujillo, A., Iglesias-Platas, I., Okamura, K., Sugahara, N., Simón, C., Moore, H., Harness, J.V., et al. (2014). Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. Genome Res *24*, 554-569. 10.1101/gr.164913.113.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21. 10.1093/bioinformatics/bts635.

Gkountela, S., Zhang, K.X., Shafiq, T.A., Liao, W.W., Hargan-Calvopiña, J., Chen, P.Y., and

Clark, A.T. (2015). DNA Demethylation Dynamics in the Human Prenatal Germline. Cell *161*, 1425-1436. 10.1016/j.cell.2015.05.012.

Hwang, Y.S., Suzuki, S., Seita, Y., Ito, J., Sakata, Y., Aso, H., Sato, K., Hermann, B.P., and Sasaki, K. (2020). Reconstitution of prospermatogonial specification in vitro from human induced pluripotent stem cells. Nat Commun *11*, 5656. 10.1038/s41467-020-19350-3.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics (Oxford, England) *27*, 1571-1572. 10.1093/bioinformatics/btr167.

Kumar, B., and Elsasser, S.J. (2019). Quantitative Multiplexed ChIP Reveals Global Alterations that Shape Promoter Bivalency in Ground State Embryonic Stem Cells. Cell Rep *28*, 3274-3284 e3275. 10.1016/j.celrep.2019.08.046.

Li, L., Dong, J., Yan, L., Yong, J., Liu, X., Hu, Y., Fan, X., Wu, X., Guo, H., Wang, X., et al. (2017). Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. Cell stem cell *20*, 858-873.e854. 10.1016/j.stem.2017.03.007.

Liao, Y., Smyth, G.K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res *47*, e47. 10.1093/nar/gkz114.

Loehr, A.R., Pierpont, T.M., Gelsleichter, E., Galang, A.M.D., Fernandez, I.R., Moore, E.S., Guo, M.Z., Miller, A.D., and Weiss, R.S. (2021). Targeting Cancer Stem Cells with Differentiation Agents as an Alternative to Genotoxic Chemotherapy for the Treatment of Malignant Testicular Germ Cell Tumors. Cancers (Basel) *13*. 10.3390/cancers13092045.

Mazumder, S., Plesca, D., and Almasan, A. (2008). Caspase-3 activation is a critical determinant of genotoxic stress-induced apoptosis. Methods Mol Biol *414*, 13-21. 10.1007/978-1-59745-339-4_2.

Mitsunaga, S., Odajima, J., Yawata, S., Shioda, K., Owa, C., Isselbacher, K.J., Hanna, J.H., and Shioda, T. (2017). Relevance of iPSC-derived human PGC-like cells at the surface of embryoid bodies to prechemotaxis migrating PGCs. Proc Natl Acad Sci U S A *114*, E9913-E9922. 10.1073/pnas.1707779114.

Moelans, C.B., Atanesyan, L., Savola, S.P., and van Diest, P.J. (2018). Methylation-Specific Multiplex Ligation-Dependent Probe Amplification (MS-MLPA). Methods Mol Biol *1708*, 537-549. 10.1007/978-1-4939-7481-8_27.

Pierpont, T.M., Lyndaker, A.M., Anderson, C.M., Jin, Q., Moore, E.S., Roden, J.L., Braxton, A., Bagepalli, L., Kataria, N., Hu, H.Z., et al. (2017). Chemotherapy-Induced Depletion of OCT4-Positive Cancer Stem Cells in a Mouse Model of Malignant Testicular Cancer. Cell

Rep *21*, 1896-1909. 10.1016/j.celrep.2017.10.078.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140. 10.1093/bioinformatics/btp616.

Smit, A., Hubley, R & Green, P. (2013-2015). RepeatMasker Open-4.0. http://www.repeatmasker.org.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics *31*, 2032-2034. 10.1093/bioinformatics/btv098.

Tyser, R.C.V., Mahammadov, E., Nakanoh, S., Vallier, L., Scialdone, A., and Srinivas, S. (2021). Single-cell transcriptomic characterization of a gastrulating human embryo. Nature *600*, 285-289. 10.1038/s41586-021-04158-y.

Vargiu, L., Rodriguez-Tome, P., Sperber, G.O., Cadeddu, M., Grandi, N., Blikstad, V., Tramontano, E., and Blomberg, J. (2016). Classification and characterization of human endogenous retroviruses; mosaic forms are common. Retrovirology *13*, 7. 10.1186/s12977-015-0232-y.

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.