**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

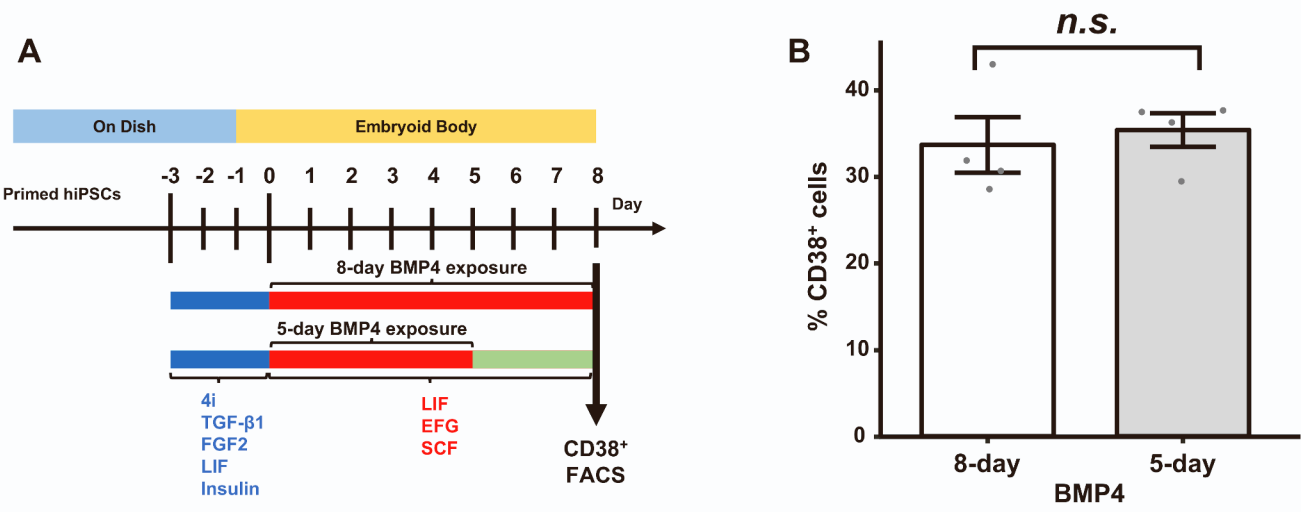


**Figure S1, Related to Figure 1.**

**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

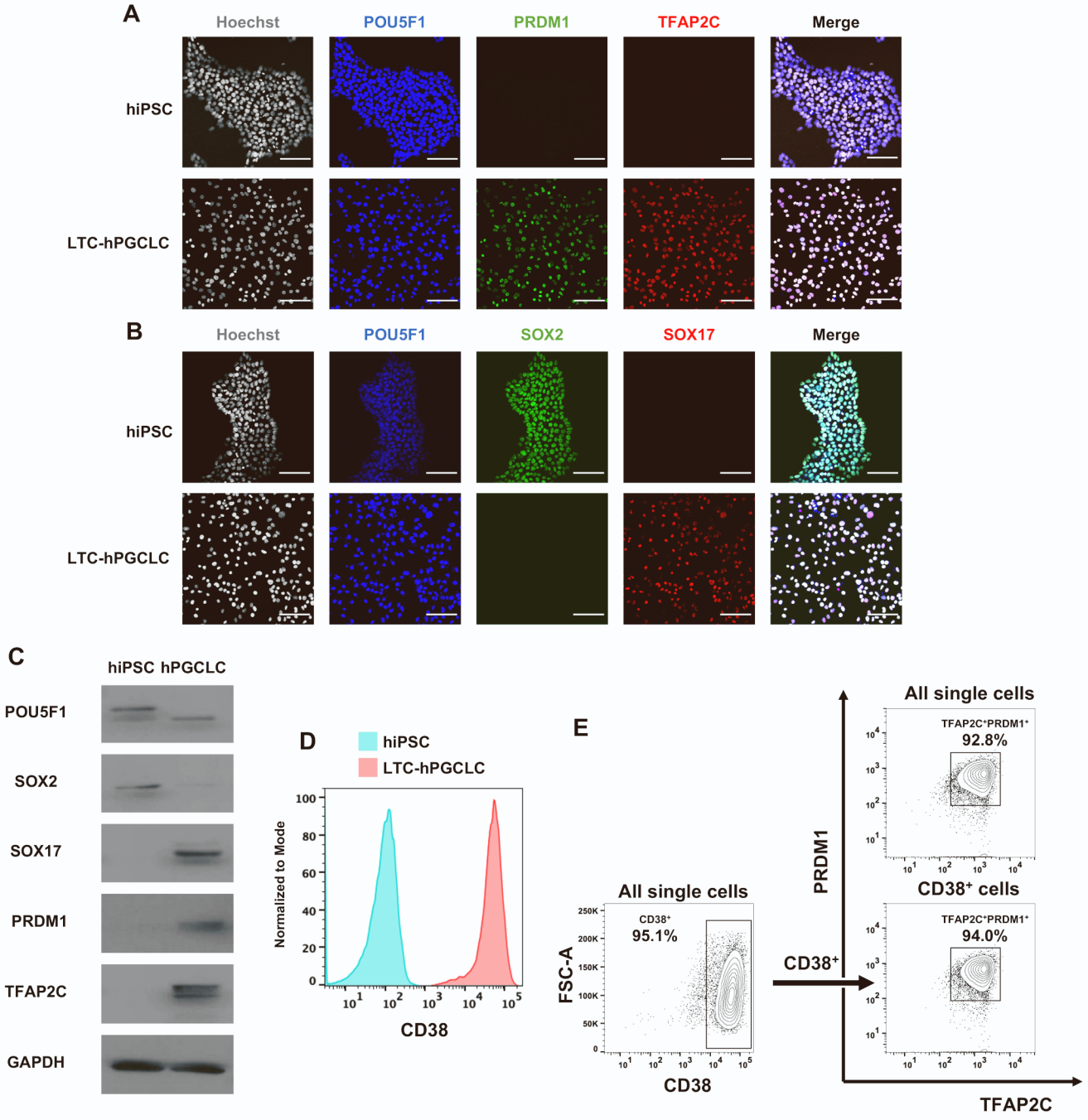


**Figure S2, Related to Figure 2.**

**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

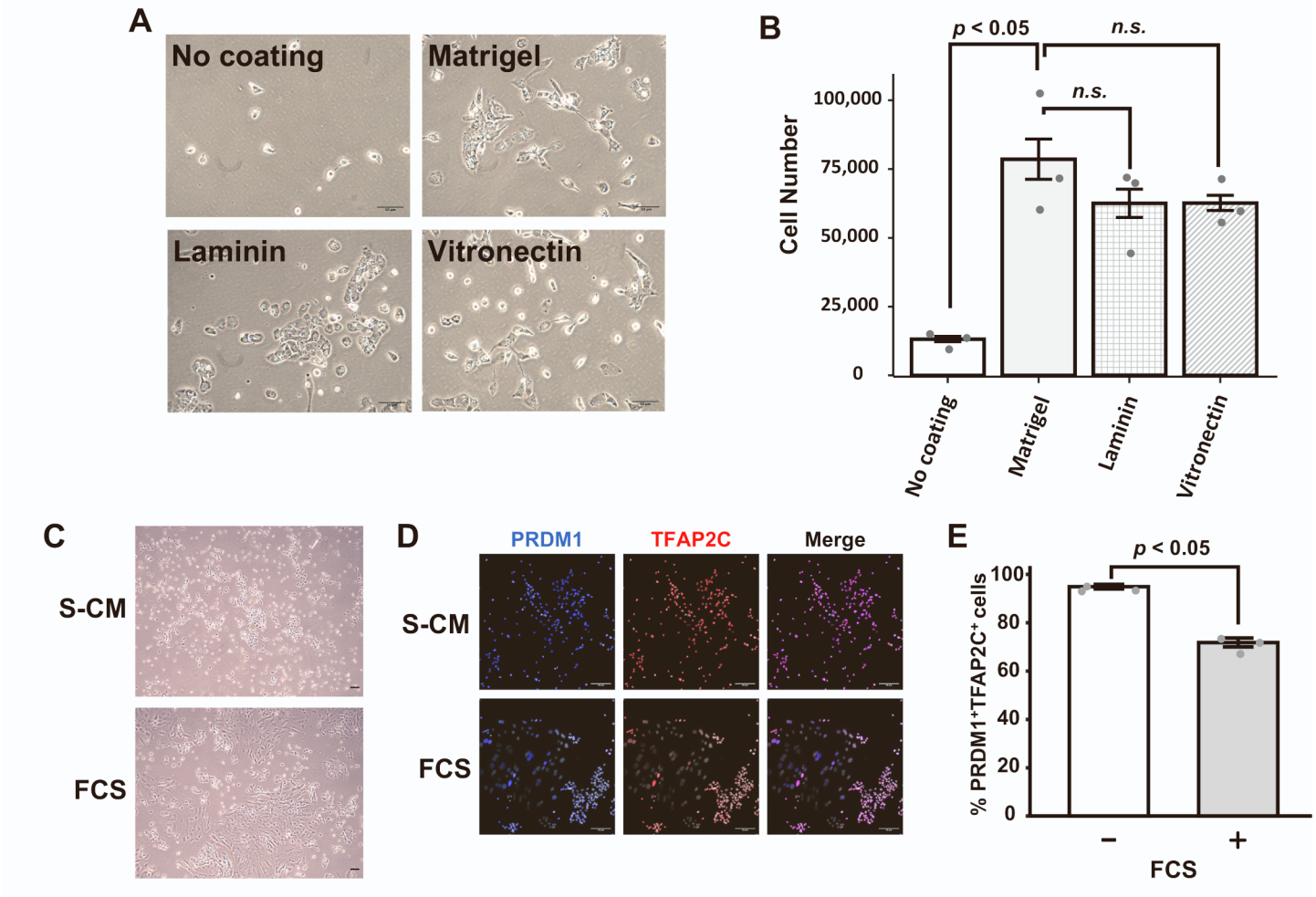


**Figure S3, Related to Figure 2**

**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

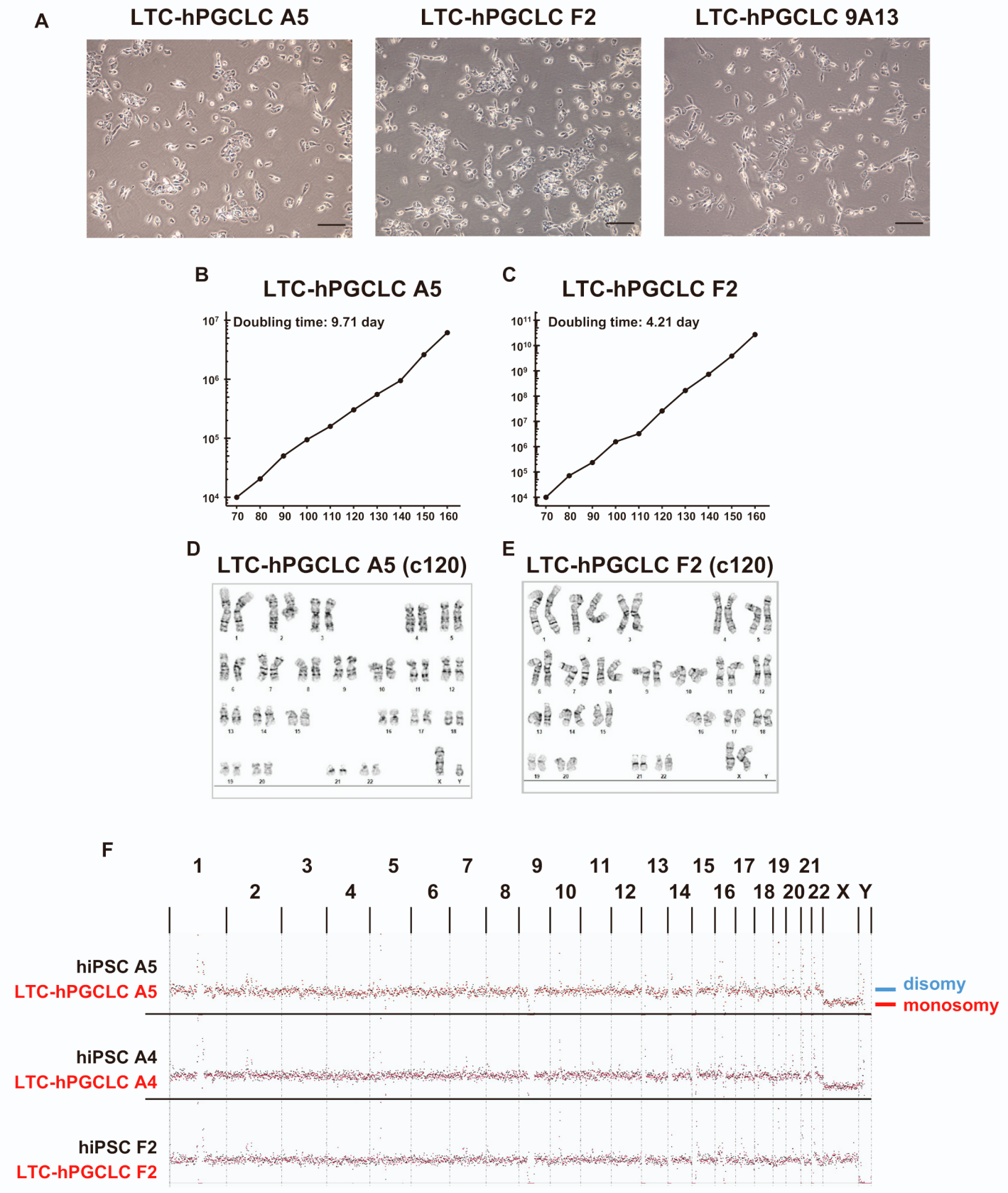


**Figure S4, Related to Figure 2**

**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

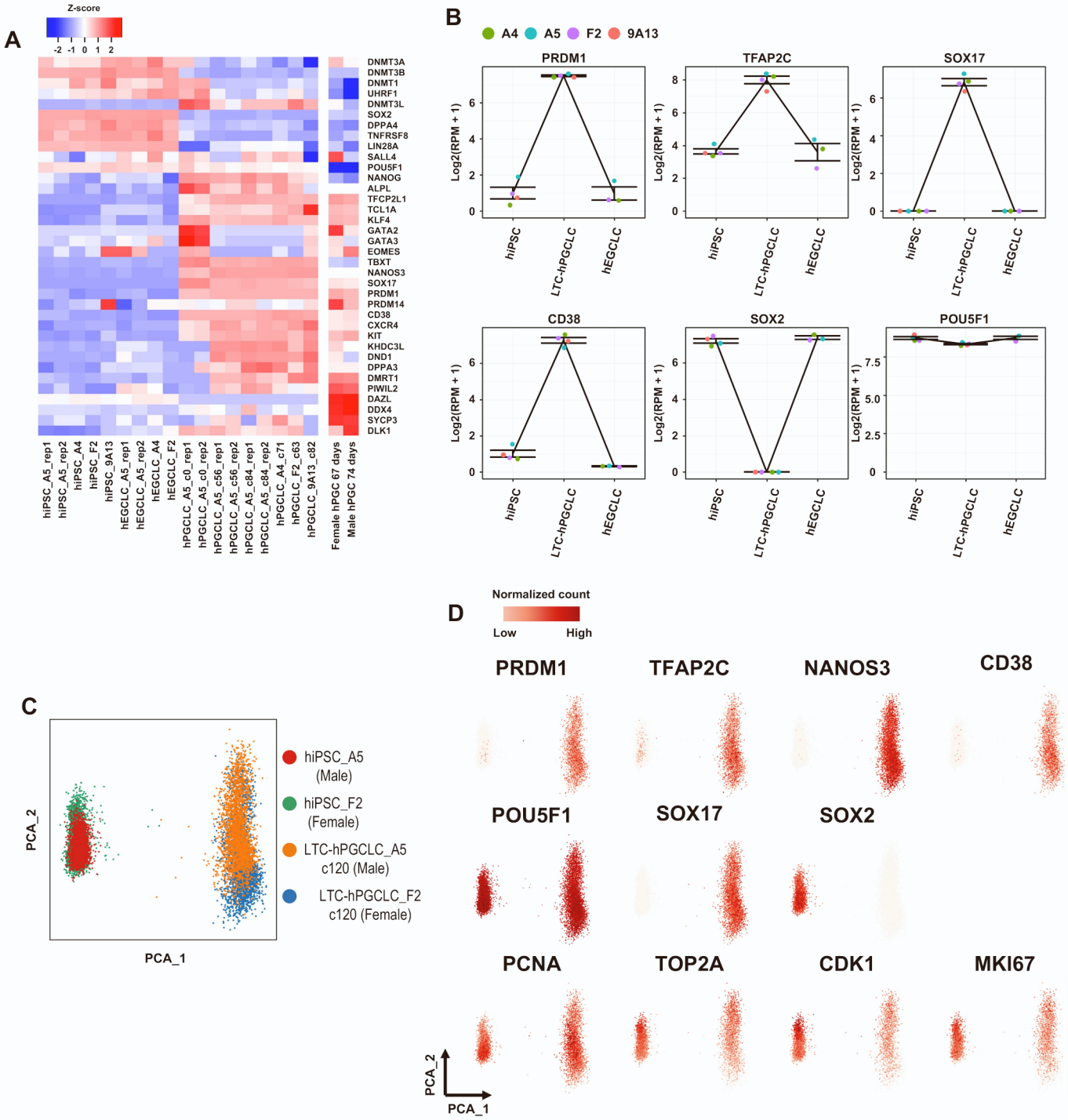
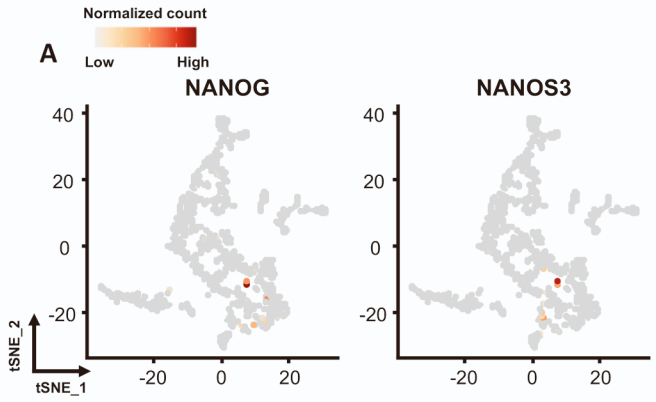


**Figure S5, Related to Figure 4**

**Transcriptomic profiles of hiPSCs, hPGCLCs, and hEGCs.**

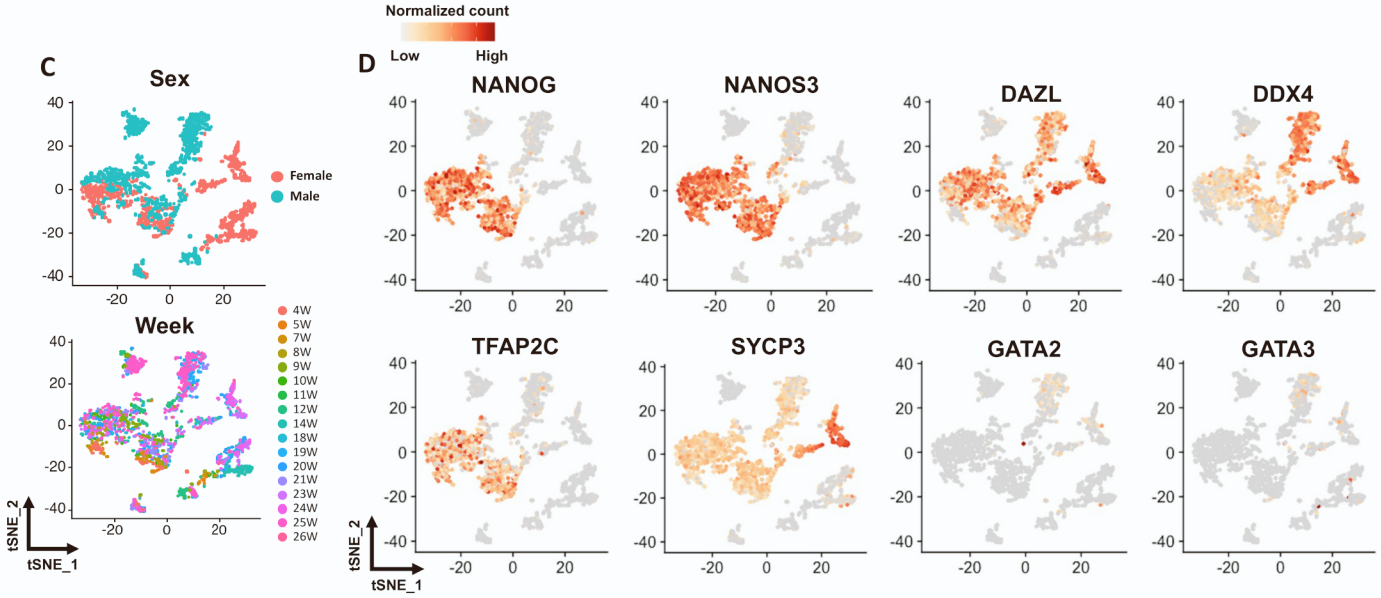
(A) Heatmap representation of marker gene expression in hiPSCs, hPGCLCs, hEGCs derived from four cell line, and human fetal PGCs. Data of human embryonic PGCs (hPGCs, gestational days 113 and 122) were from a previously published study (Gkountela *et al.,* 2015). (B) Gene expression dynamics in hiPSCs, 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).

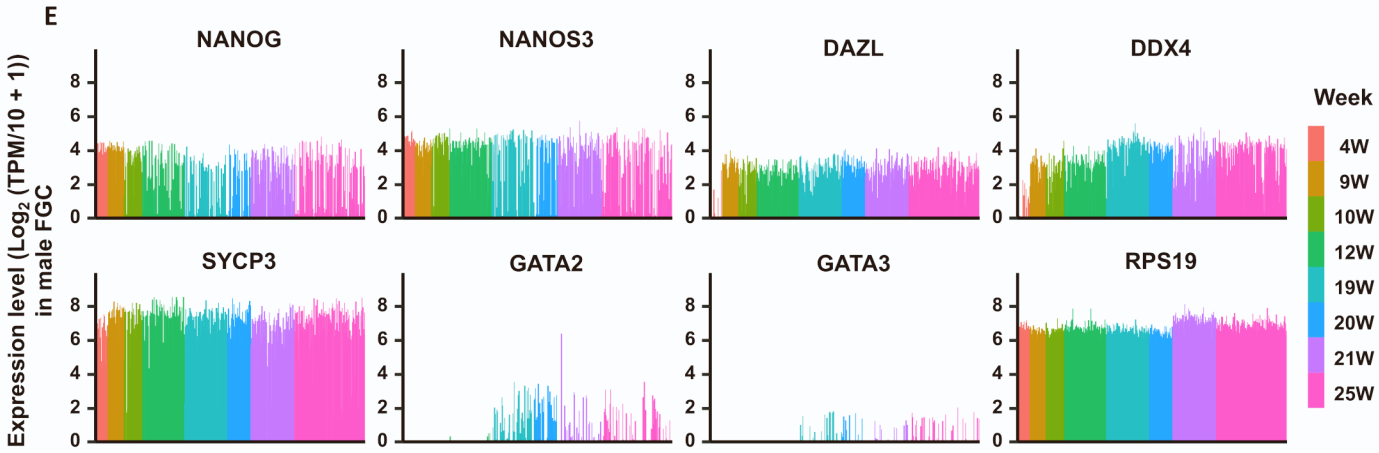
# Kobayashi et al., Figure S6




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 |



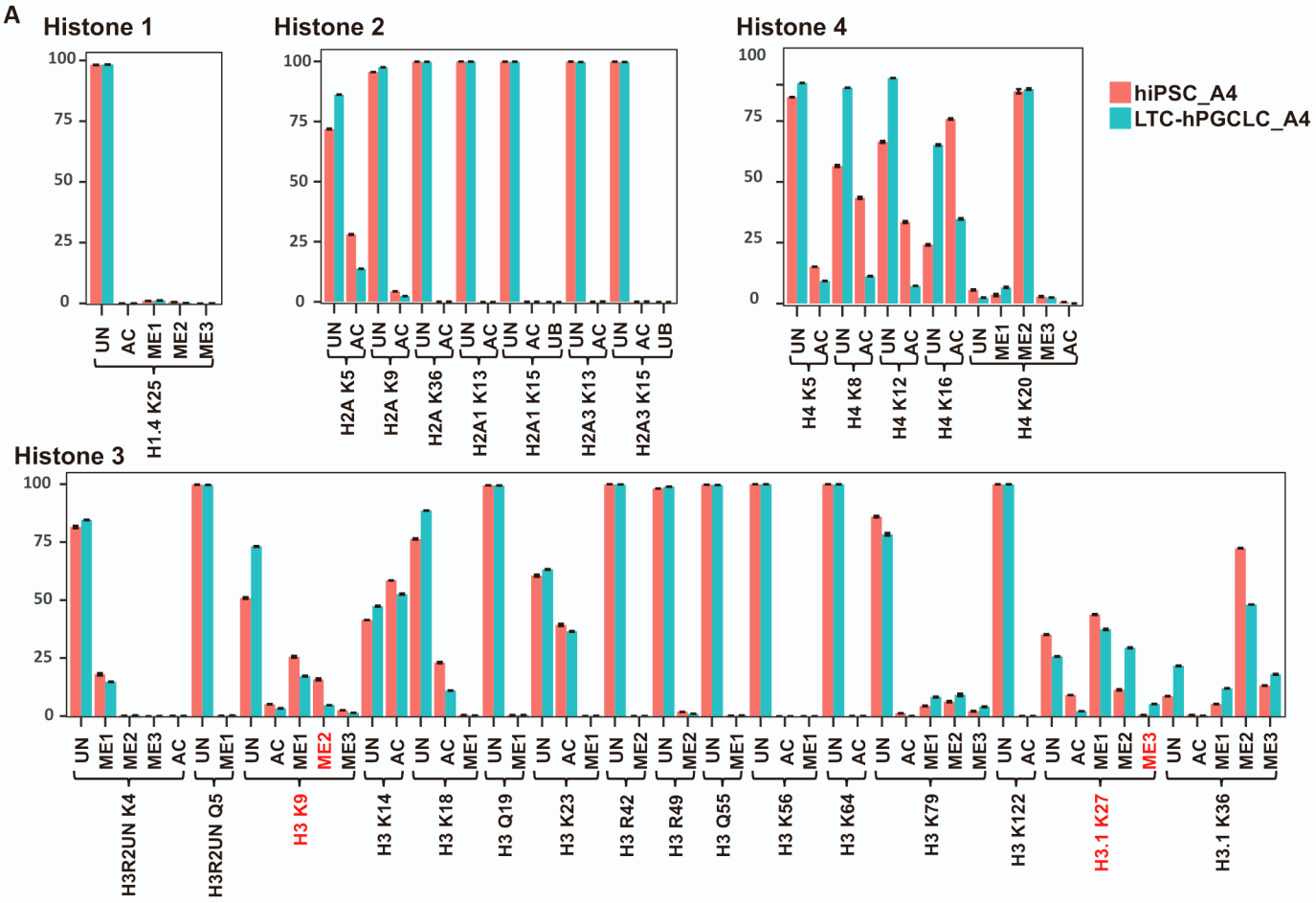



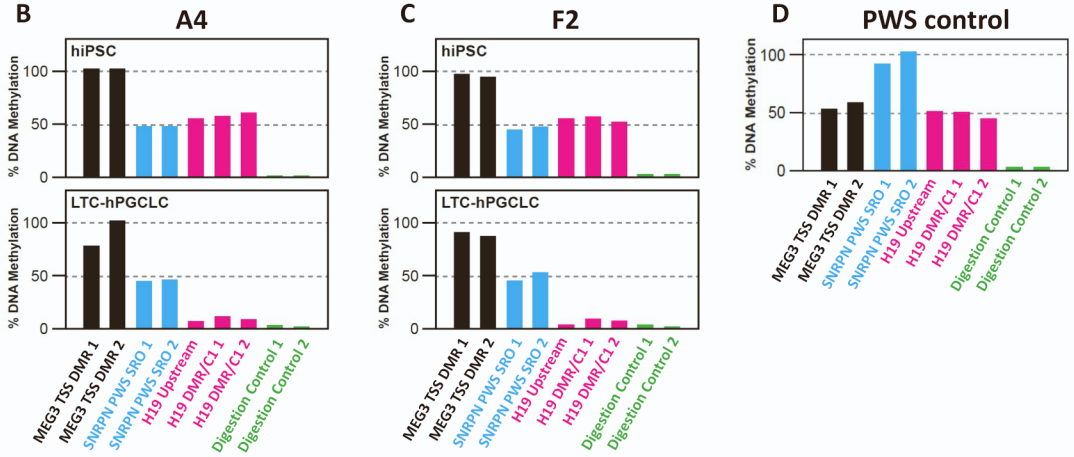
## Figure S6, Related to Figure 5

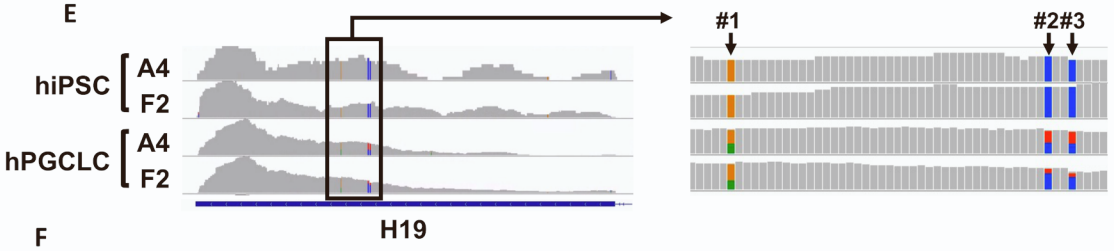
## Reanalysis of online available human fetal germ cell dataset

(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)<br>A (44%; 32/73) | C (47%; 34/72)<br>T (51%; 37/72)<br>A (1%; 1/72) | C (44%; 30/68)<br>T (54%; 37/68)<br>N (1%; 1/68) |
| | F2 | G (64%; 61/96)<br>A (36%; 35/96) | C (78%; 62/80)<br>T (23%; 18/80) | C (74%; 49/66)<br>T (24%; 16/66)<br>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 $8\times10^4$ 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.