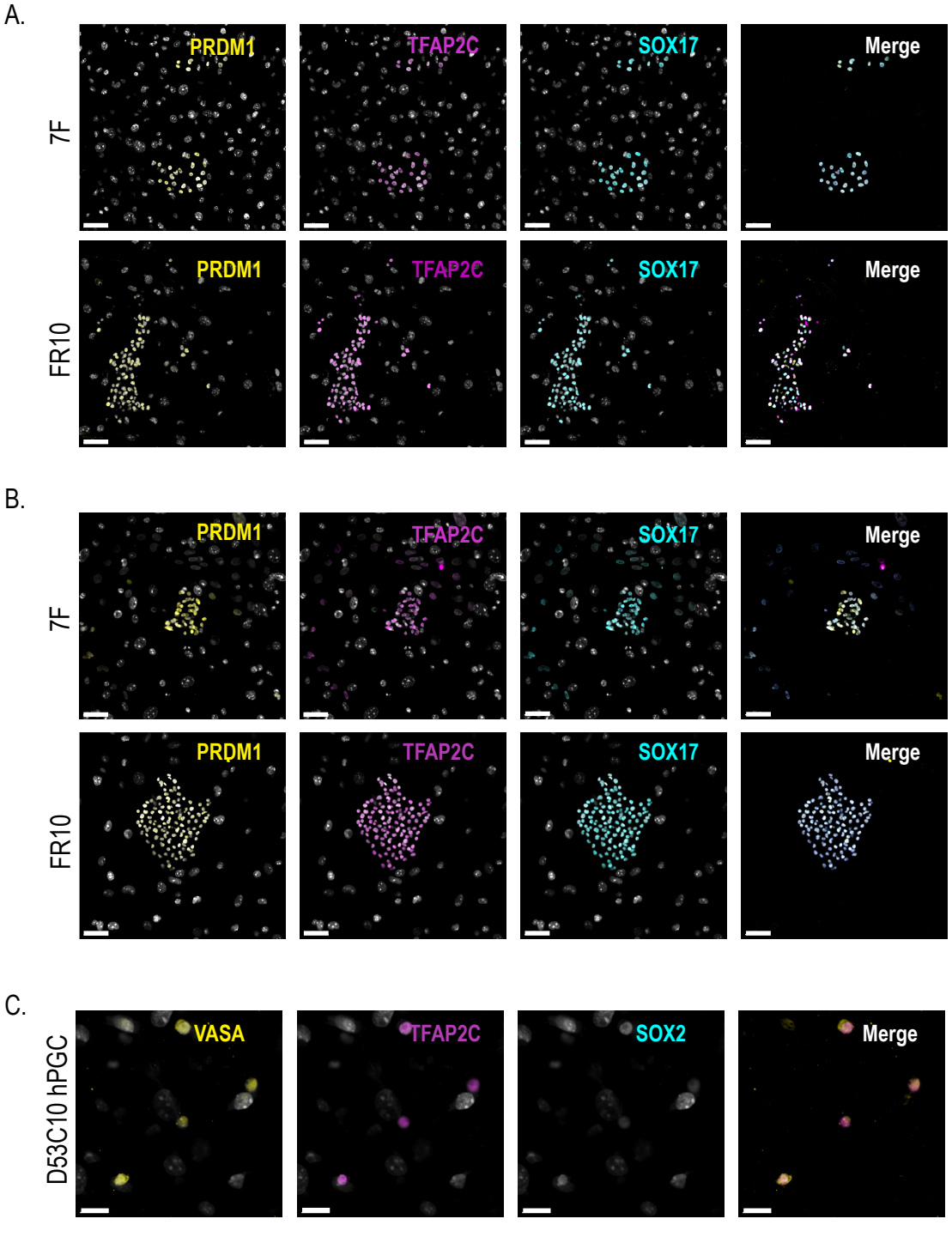
Stem Cell Reports, Volume *14*

Supplemental Information

An Extended Culture System that Supports Human Primordial Germ Cell-like Cell Survival and Initiation of DNA Methylation Erasure

Joanna J. Gell, Wanlu Liu, Enrique Sosa, Alex Chialastri, Grace Hancock, Yu Tao, Sissy E. Wamaitha, Grace Bower, Siddharth S. Dey, and Amander T. Clark

Supplement 1.



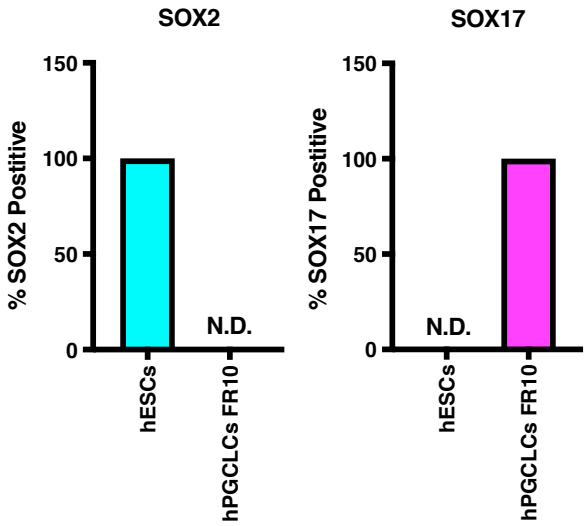
Supplemental Figure 1: Germline identify in male and female hPGCLCs and hPGCs in extended culture. Related to Figure 1 and Table 1

A. IF images of UCLA1 D4C10 hPGCLCs in 7F media (top) and FR10 media (bottom). Germline identity was evaluated using PRDM1 (yellow), TFAP2C (magenta) and SOX17 (yellow). Scale bars 50 µm.

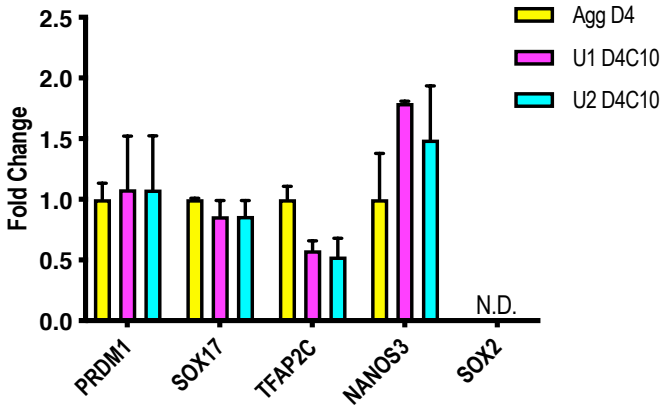
B IF images of UCLA6 D4C10 hPGCLCs in 7F medial (top) and FR10 media (bottom). Germline identity was evaluated using PRDM1 (yellow), TFAP2C (magenta) and SOX17 (yellow). Scale bars 50 µm.

C. IF images of embryonic day 53 (D53) hPGCs cultured for ten days in FR10 media. Germline identity was evaluated using VASA (yellow) and TFAP2C (magenta), additionally demonstrating the lack of SOX2 (cyan). Scale bars 20 µm.

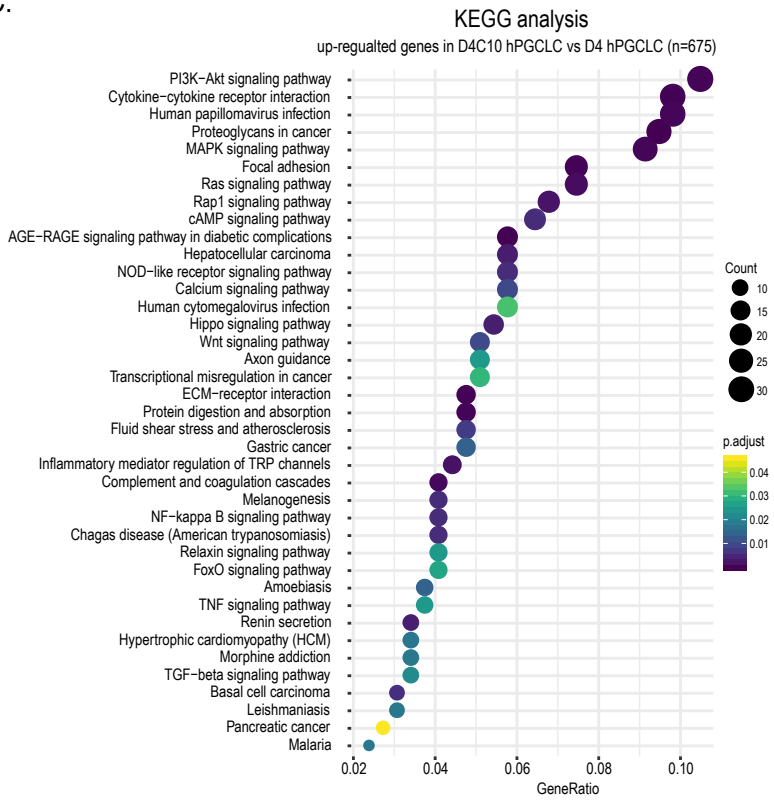
A.


SOX2
% SOX2 Postitive
150
100
50
0
N.D.
hESCs
hPGCLCs FR10
SOX17
% SOX17 Postitive
150
100
50
0
N.D.
hESCs
hPGCLCs FR10

B.


Agg D4
U1 D4C10
U2 D4C10
Fold Change
2.5
2.0
1.5
1.0
0.5
0.0
N.D.
PRDM1
SOX17
TFAP2C
NANOS3
SOX2

C.


KEGG analysis
up-regualted genes in D4C10 hPGCLC vs D4 hPGCLC (n=675)
PI3K−Akt signaling pathway
Cytokine−cytokine receptor interaction
Human papillomavirus infection
Proteoglycans in cancer
MAPK signaling pathway
Focal adhesion
Ras signaling pathway
Rap1 signaling pathway
cAMP signaling pathway
AGE−RAGE signaling pathway in diabetic complications
Hepatocellular carcinoma
NOD−like receptor signaling pathway
Calcium signaling pathway
Human cytomegalovirus infection
Hippo signaling pathway
Wnt signaling pathway
Axon guidance
Transcriptional misregulation in cancer
ECM−receptor interaction
Protein digestion and absorption
Fluid shear stress and atherosclerosis
Gastric cancer
Inflammatory mediator regulation of TRP channels
Complement and coagulation cascades
Melanogenesis
NF−kappa B signaling pathway
Chagas disease (American trypanosomiasis)
Relaxin signaling pathway
FoxO signaling pathway
Amoebiasis
TNF signaling pathway
Renin secretion
Hypertrophic cardiomyopathy (HCM)
Morphine addiction
TGF−beta signaling pathway
Basal cell carcinoma
Leishmaniasis
Pancreatic cancer
Malaria
0.02
0.04
0.06
0.08
0.10
GeneRatio
Count
10
15
20
25
30
p.adjust
0.04
0.03
0.02
0.01

Supplemental Figure 2: hPGCLCs in extended culture have a transcriptional identity resembling early PGCs. Related to Figure 2

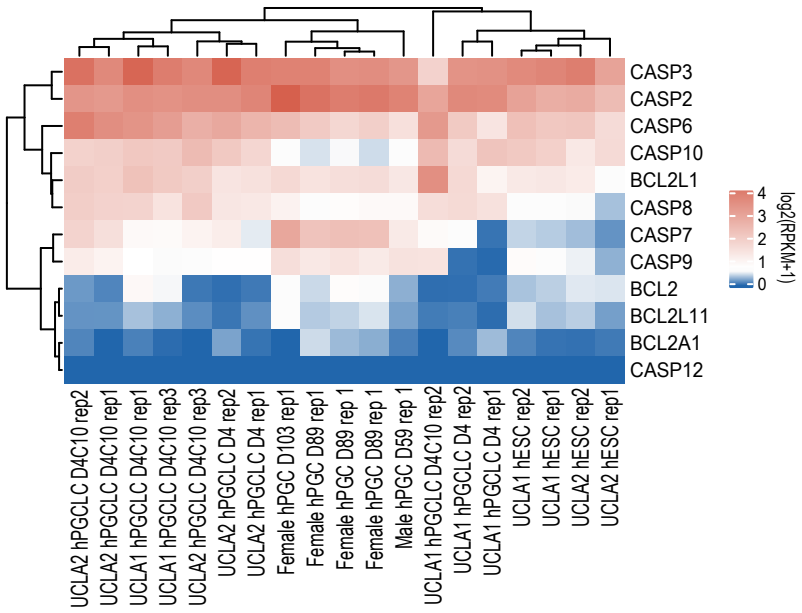
A. Bar graphs representing percentage of SOX2 positive UCLA2 hESCs (cyan) and D4C10 hPGCLCs (magenta) in FR10 media (left) and SOX17 positive hESCs and D4C10 hPGCLCs (right) in FR10 media. N.D. = not detected.

B. Gene expression measured by RT-PCR of FACs isolated hPGCLCs, aggregates at day 4 (yellow), UCLA1 D4C10 (magenta) and UCLA2 D4C10 (cyan). N=2 reps. N.D. = not detected.

C. Dot plot depicting Kyoto Encyclopedia of Genes and Genomes (KEGG) terms identified for the N=675 up-regulated genes in D4C10 hPGCLCs versus D4 hPGCLCs.

Supplement 3.

A.

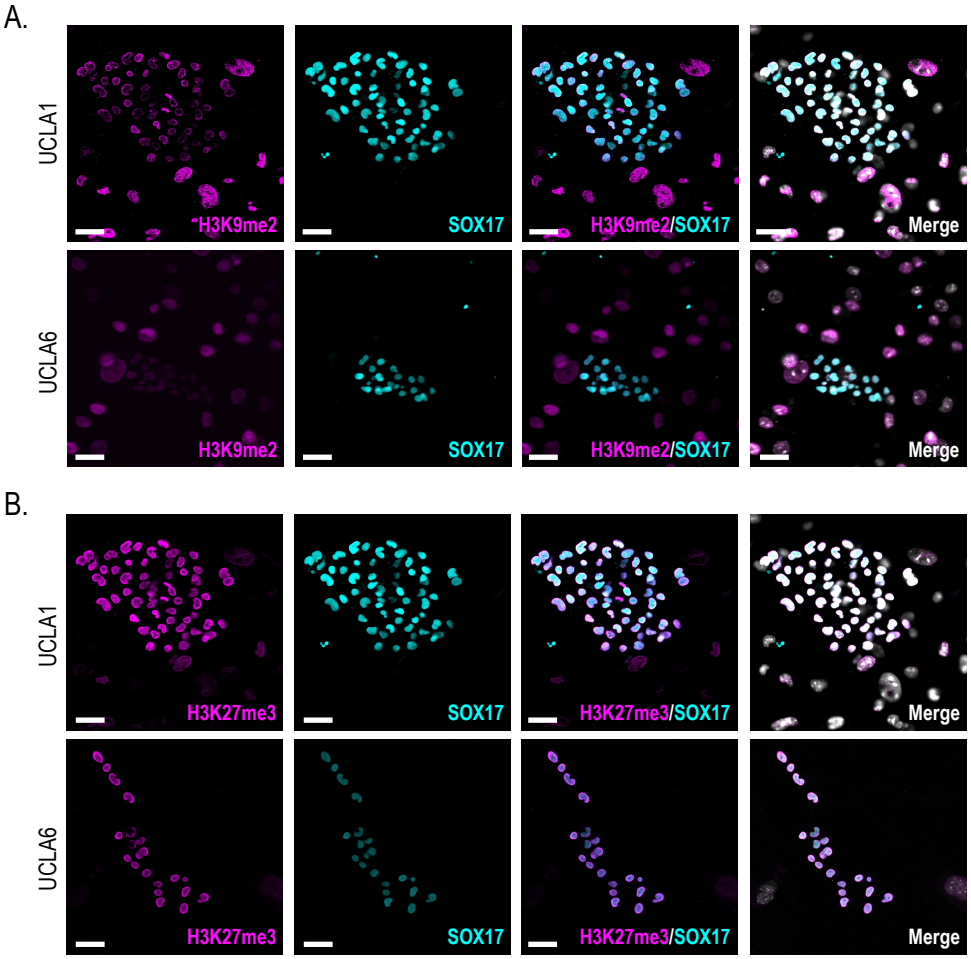


Supplemental Figure 3: Apoptotic genes are not altered in D10C4 hPGCLCs cultured in FR10 media.

Related to Figure 2

A. Heatmap of select apoptotic genes in UCLA1 and UCLA2 primed hESCs (N=2 reps), gonadal hPGCs (N=5 reps), UCLA1 & UCLA2 D4 hPGCLCs (N=2 reps), UCLA1 & UCLA2 D4C10 hPGCLCs (N=3 reps).

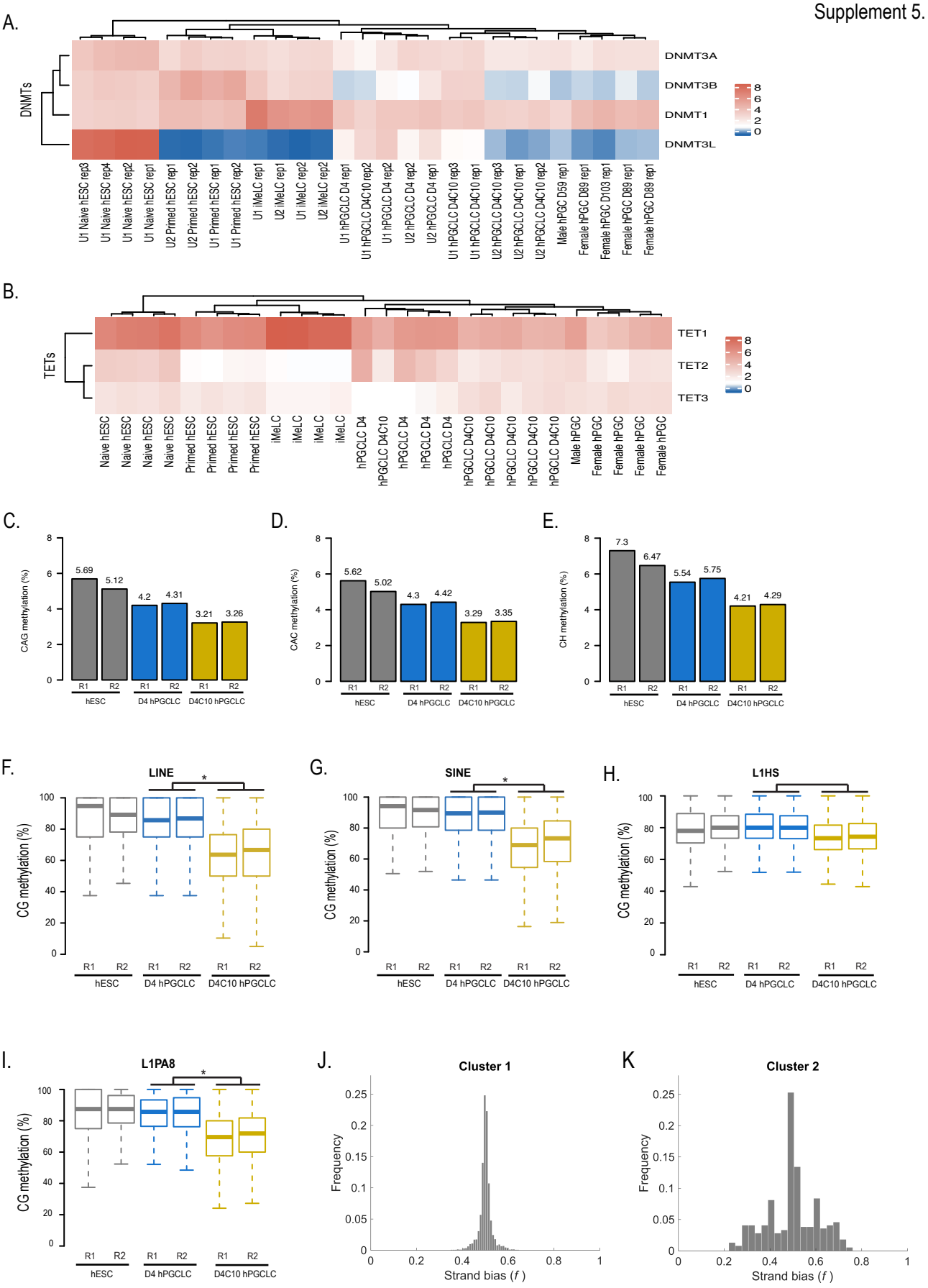
Supplement 4.



Supplemental Figure 4: Epigenetic reprogramming in UCLA1 and UCLA6 hPGCLCs, mirror that of UCLA2 hPGCLCs. Related to Figure 4

A. IF images of H3K9me2 (magenta) pattern in SOX17 (cyan) positive UCLA1 (top) and UCLA6 (bottom) D4C10 hPGCLCs cultured in FR10 media. Scale bars 50 μm.

B. IF images of H3K27me3 (magenta) pattern in SOX17 (cyan) positive UCLA1 (top) and UCLA6 (bottom) D4C10 hPGCLCs cultured in FR10 media. Scale bars 50 μm.


A.
DNMTs
DNMT3A
DNMT3B
DNMT1
DNMT3L
8
6
4
2
0
U1 Naive hESC rep3
U1 Naive hESC rep4
U1 Naive hESC rep2
U1 Naive hESC rep1
U2 Primed hESC rep1
U2 Primed hESC rep2
U1 Primed hESC rep1
U1 Primed hESC rep2
U1 iMeLC rep1
U2 iMeLC rep1
U1 iMeLC rep2
U2 iMeLC rep2
U1 hPGCLC D4 rep1
U1 hPGCLC D4C10 rep2
U1 hPGCLC D4 rep2
U2 hPGCLC D4 rep2
U2 hPGCLC D4 rep1
U1 hPGCLC D4C10 rep3
U1 hPGCLC D4C10 rep1
U2 hPGCLC D4C10 rep3
U2 hPGCLC D4C10 rep1
U2 hPGCLC D4C10 rep2
Male hPGC D59 rep1
Female hPGC D89 rep1
Female hPGC D103 rep1
Female hPGC D89 rep1
Female hPGC D89 rep1
B.
TETs
TET1
TET2
TET3
8
6
4
2
0
Naive hESC
Naive hESC
Naive hESC
Naive hESC
Primed hESC
Primed hESC
Primed hESC
Primed hESC
iMeLC
iMeLC
iMeLC
iMeLC
hPGCLC D4
hPGCLC D4C10
hPGCLC D4
hPGCLC D4
hPGCLC D4
hPGCLC D4C10
hPGCLC D4C10
hPGCLC D4C10
hPGCLC D4C10
hPGCLC D4C10
Male hPGC
Female hPGC
Female hPGC
Female hPGC
Female hPGC
C.
CAG methylation (%)
5.69
5.12
4.2
4.31
3.21
3.26
R1
R2
hESC
D4 hPGCLC
D4C10 hPGCLC
D.
CAC methylation (%)
5.62
5.02
4.3
4.42
3.29
3.35
E.
CH methylation (%)
7.3
6.47
5.54
5.75
4.21
4.29
F.
LINE
CG methylation (%)
G.
SINE
H.
L1HS
I.
L1PA8
J.
Cluster 1
Frequency
Strand bias (f)
K
Cluster 2

Supplemental Figure 5: Epigenetic modifiers and DNA demethylation pattern bears resemblance to gonadal hPGCs. Related to Figure 4 and Figure 5.

A. Heatmap of gene expression of DNA methyltransferases (DNMT) in UCLA1 (U1) naïve hESCs (N=4 replicates), U1 & U2 primed hESCs (N=2 reps), U1 & U2 iMeLCs (N=2 reps), gonadal hPGCs (N=5 reps), U1 & U2 D4 hPGCLCs (N=2 reps), U1 & U2 D4C10 hPGCLCs (N=3 reps).

B. Heatmap of gene expression of Ten eleven translocation (TET), mirroring the cells utilized in the above heatmap (Figure S5A).

C-E. Boxplot of non-CG methylation levels across the genome in UCLA2 hESCs (grey), D4 hPGCLCs (blue) and D4C10 hPGCLCs (Yellow) (N=2 replicates for each conditions). C. CAG methylation; D. CAC methylation; E. CH methylation.

F-I. Boxplot of CG methylation levels over major transposons elements. F. Long interspersed nuclear elements (LINE); G. short interspersed nuclear elements (SINE); H. human endogenous retrovirus K (HERV-K) subgroup, LINE-1 elements (L1) Homo sapiens (L1HS); I. HERV-K subgroup L1PA8, in hESCs (grey), D4 hPGCLCs (blue), D4C10 hPGCLCs (yellow) were performed in replicates of N=2, (R1 and R2). *= p-value $<2.2e^{-16}$

J. Histogram showing strand bias for all autosomal chromosomes cluster 2 cells in Figure 5F.

K. Histogram showing strand bias for all autosomal chromosomes cluster 1 cells in Figure 5F.

Supplemental Tables:

Table S1. Summary information for RNA-seq data used in this study. Relating to Figure 2

Type	seqType	Kit	Sample	Adaptor	Adaptor sequence	Raw reads	Left mapped	Right mapped	Aligned pairs	Map %
RNAseq	PE50	Ovation RNA amp system V2 and Encore Rapid	hPGCLC D4C10 UCLA1 rep1	BC9	ACACGA	21740348	19714147	19260599	18273284	84.05%
RNAseq	PE50		hPGCLC D4C10 UCLA1 rep2	BC10	CACACA	32395255	28748441	28202012	26094632	80.55%
RNAseq	PE50		hPGCLC D4C10 UCLA1 rep3	BC13	ACAAAC	46668321	41644372	41042148	38765254	83.07%
RNAseq	PE50		hPGCLC D4C10 UCLA2 rep1	BC14	CACCTC	35299597	31306797	30830937	28703048	81.31%
RNAseq	PE50		hPGCLC D4C10 UCLA2 rep2	BC15	GTGGCC	28014273	25561170	25267928	24051776	85.86%
RNAseq	PE50		hPGCLC D4C10 UCLA2 rep3	BC16	TGTTGC	19342190	17278008	17017022	16028156	82.87%

Table S2. Summary information for WGBS data used in this study. Relating to Figure 5

Type	seqType	Kit	Sample	Adaptor	Adaptor sequence	Total Reads	Aligned Reads	Aligned %
WGBS	SE100	NugenBisulfite	hESC UCLA2 rep1	1	AACCAG	466856712	320670761	68.69%
WGBS	SE100		hESC UCLA2 rep2	2	TGGTGA	501039107	368950817	73.64%
WGBS	SE100		hPGCLC UCLA2 D4 rep1	3	AGTGAG	390547290	290454998	74.37%
WGBS	SE100		hPGCLC UCLA2 D4 rep2	4	GCACTA	473938105	347974868	73.42%
WGBS	SE100		hPGCLC UCLA2 D4C10 rep1	5	ACCTCA	443735742	327957432	73.91%
WGBS	SE100		hPGCLC UCLA2 D4C10 rep2	6	GTGCTT	473238502	360474596	76.17%

Supplemental Experimental Procedures

Cell lines and Cell Culture

Primed hESC lines were cultured on mitomycin C-inactivated mouse embryonic fibroblasts (MEFs) in hESC media, which is composed of 20% knockout serum replacement (KSR) (GIBCO, 10828-028), 100μM L-Glutamine (GIBCO, 25030-081), 1x Non-Essential Amino Acids (NEAA) (GIBCO, 11140-050), 55μM 2-Mercaptoethanol (GIBCO, 21985-023), 10ng/mL recombinant human FGF basic (R&Dsystems, 233-FB), 1x Penicillin-Streptomycin (GIBCO, 15140-122), and 50ng/mL primocin (InvivoGen, ant-pm-2) in DMEM/F12 media (GIBCO, 11330-032). All hESC lines were split every 7 days with Collagenase type IV (GIBCO, 17104-019). All hESC lines used in this study are registered with the National Institute of Health Human Embryonic Stem Cell Registry and are available for research use with NIH funds. Specifically, the following hESC lines were used in this study: UCLA1(46XX), UCLA2 (46XY), UCLA2-GFP (46XY), UCLA6 (46XY). The derivation and basic characterization (karyotype and teratoma analysis) of UCLA1, UCLA2 and UCLA6 were previously reported (Diaz Perez et al., 2012). Experiments were performed between passage 15-25, two passages were performed between thaw and use in experiments. SIM 6-thioguanine resistant ouabain (STO) line was obtained from America Type Culture Collection (ATCC) (ATCC-CRL-1503). STO cells were mitomycin C (MMC)-inactivated with 10ug/mL of MMC. Passage 1-3 were used for extended culture experiments. All cell lines used in these experiments were Mycoplasma negative. Mycoplasma testing was performed every 6-9 weeks, using MycoAlert Detection Kit (Lonza, LT07-418).

Induction of hPGCLCs though iMeLCs from primed hESCs

hPGCLCs were induced from primed hESCs as described in Sasaki et al., 2015 with some modifications (Sasaki et al., 2015). Day 7 hESCs were dissociated into single cells with 0.05% Trypsin-EDTA (Gibco, 25300-054) and plated onto Human Plasma Fibronectin (Invitrogen,33016-015)-coated 12-well-plate at the density of 200,000 cells/well in 2mL/well of iMeLC media, which is composed of 15% KSR, 1x NEAA, 0.1mM 2-Mercaptoethanol, 1x PSG (Gibco, 10378-016), 1mM sodium pyruvate (Gibco,11360-070),

50ng/mL Activin A (Peprotech, AF-120-14E), 3μM CHIR99021 (Stemgent, 04-0004), 10μM of ROCKi (Y27632, Stemgent, 04-0012-10), and 50ng/mL primocin in Glasgow's MEM (GMEM) (Gibco, 11710-035). iMeLCs were dissociated into single cells with 0.05% Trypsin-EDTA after 24 hours of incubation unless otherwise mentioned and plated into ultra-low cell attachment U-bottom 96-well plates (Corning, 7007) at the density of 3,000 cells/well in 200μl/well of PGCLC media, which is composed of 15% KSR, 1x NEAA, 0.1mM 2-Mercaptoethanol, 1x PSG (Gibco, 10378-016), 1mM sodium pyruvate (Gibco, 11360-070), 10ng/mL human LIF (Millipore, LIF1005), 200ng/mL human BMP4 (R&D systems, 314-BP), 50ng/mL human EGF (R&D systems, 236-EG) 10μM of ROCKi (Y27632, Stemgent, 04-0012-10), and 50ng/mL primocin in Glasgow's MEM (GMEM) (Gibco, 11710-035).

Immunofluorescence and microscopy

Preparation for immunostaining of paraffin slides and chamber slides was performed as previously described (Gell et al., 2018). The primary antibodies used for immunofluorescence in this study include: rabbit-anti-PRDM1 (Cell Signaling, 9115, 1:100), mouse-anti-TFAP2C (Santa Cruz Biotechnology, sc-12762, 1:200) goat-anti-SOX17 (Neuromics, GT15094, 1:100), rabbit-anti-SOX2 (Abcam, 97959, 1:200), mouse-anti-Ki67 (BDpharmgen, 556003, 1:100), mouse-anti-UHRF1 (Santa Cruz, sc-373750, 1:100), mouse-anti-H3K9me2 (Abcam, ab1220, 1:100), rabbit-anti-H3K27me3 (Millipore, 07-449, 1:50), goat-anti-VASA (R&D, AF2030, 1:100). The secondary antibodies used in this study are donkey anti-rabbit-488 (Jackson ImmunoResearch Laboratories, 711-545-152), donkey anti-rabbit-594 (Jackson ImmunoResearch Laboratories, 711-585-152) donkey anti-goat-647 (Life Technologies, A21447), donkey-anti-mouse-594 (Life technologies, A21447), and donkey-anti-mouse-488 (Jackson ImmunoResearch Laboratories, 715-545-150). DAPI is counterstained to indicate nuclei. EdU staining was performed using Click-iT™ EdU Cell Proliferation Kit for Imaging, Alexa Fluor™ 488 dye

(ThermoFisher, C10337). All slides were imaged with an LSM 780 confocal microscope (Zeiss) using ZEN 2011 software. Images were processed using image analysis software Imaris 9.3.1 (Bitplane).

Fluorescence activated cell sorting

Day 4 aggregates were dissociated with 0.05% Trypsin-EDTA for 10 minutes at 37^{o}C. The dissociated cells were stained with conjugated antibodies, washed with FACS buffer (1% BSA in PBS) and resuspended in FACS buffer accompanying with 7-AAD (BD Pharmingen, 559925). D4C10 hPGCLCs were isolated from the STOs by harvesting cells with 0.05% Trypsin-EDTA for 3 min at 37^{o}C. The day 53 Human embryonic testes were obtained from the University of Washington Birth defects laboratory. The testis was dissociated to a single cell suspension according to the methods used in Chen et al., 2017. The conjugated antibodies used in this study are: ITGA6 conjugated with BV421 (BioLegend, 313624), EPCAM conjugated with 488 (BioLegend, 324210), PE-anti-Human TRA-1-85 (R&D systems, FAB3195P, 1:70), APC-anti-mouse CD-29 (BioLegend, 102216, 1:70), CD117 (cKIT) conjugated to APC (BD Pharmingen, 550412) and TNAP conjugated with PE (BD Pharmingen 561433). Gating for each experiment was based off of unstained negative controls, single staining compensation controls and fluorescence minus one (FMO) controls. hPGCLCs were either sorted into respective culture media or appropriate reagent for RNA or DNA preparation. FLOWJO was utilized for analysis of FACs data.

Real time quantitative PCR

hPGCLC were sorted into 350µL of RLT buffer (QIAGEN) and RNA was extracted using RNeasy micro kit (QIAGEN, 74004). cDNA was synthesized using SuperScript®II Reverse Transcriptase (Invitrogen, 18064-014). Real time quantitative PCR was performed using TaqMan® Universal PCR Master Mix (Applied Biosystems, 4304437). The expression level of genes-of-interest were normalized to the housekeeping gene, GAPDH. The TaqMan probes used include: GAPDH (Applied Biosystems, Hs99999905_m1),

NANOS3 (Applied Biosystems, Hs00928455_s1), PRDM1 (Applied Biosystems, Hs01068508_m1), TFAP2C (Applied Biosystems, Hs00231476_m1), SOX17 (Applied Biosystems, Hs00751752), SOX2 (Applied Biosystems, Hs01053049_s1). Two independent experiments were performed. R-studio was used to analyze data, and PRISM GraphPad was used to generate graphs.