Generation of human oogonia from induced pluripotent stem cells in vitro

Chika Yamashiro,[1,2] Kotaro Sasaki,[1,2] Yukihiro Yabuta,[1,2] Yoji Kojima,[1,2,3,4] Tomonori Nakamura,[1,2] Ikuhiro Okamoto,[1,2] Shihori Yokobayashi,[1,2,4] Yusuke Murase,[1,2] Yukiko Ishikura,[1,2] Kenjiro Shirane,[5,6] Hiroyuki Sasaki,[5,6] Takuya Yamamoto,[3,4,7] and Mitinori Saitou[1,2,3,4]

[1]Department of Anatomy and Cell Biology, Graduate School of Medicine, Kyoto University, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan.
[2]JST, ERATO, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan.
[3]Institute for Integrated Cell-Material Sciences, Kyoto University, Yoshida-Ushinomiya-cho, Sakyo-ku, Kyoto 606-8501, Japan.
[4]Center for iPS Cell Research and Application, Kyoto University, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan.
[5]Division of Epigenomics, Medical Institute of Bioregulation, and Epigenome Network Research Center, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582, Japan.
[6]Graduate School of Medical Sciences, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582, Japan.
[7]AMED-CREST, AMED, 1-7-1 Otemachi, Chiyoda-ku, Tokyo, 100-0004, Japan.

One Sentence Summary: Human primordial germ cell like-cells (hPGCLCs) differentiate into oogonia in xenogeneic reconstituted ovaries in vitro.

*_Correspondence should be addressed to:_
Mitinori Saitou, M.D., Ph.D.
E-mail: saitou@anat2.med.kyoto-u.ac.jp
Tel: +81-75-753-4335; Fax: +81-75-751-7286 (MS)

ABSTRACT

Human in vitro gametogenesis may transform reproductive medicine. Human pluripotent stem cells (hPSCs) have been induced into primordial germ cell-like cells (hPGCLCs); however, further differentiation to a mature germ cell has not been achieved. Here, we show that hPGCLCs differentiate progressively into oogonia-like cells during a long-term in vitro culture (~four months) in xenogeneic reconstituted ovaries with mouse embryonic ovarian somatic cells. The hPGCLC-derived oogonia display hallmarks of epigenetic reprogramming, i.e., genome-wide DNA demethylation, imprint erasure, and extinguishment of aberrant DNA methylation in hPSCs, and acquire an immediate precursory state for meiotic recombination. Furthermore, the inactive X chromosome shows a progressive demethylation and reactivation, albeit partially. These findings establish the germline competence of hPSCs and provide a critical step toward human in vitro gametogenesis.

The germ-cell lineage arises from primordial germ cells (PGCs) that go through a multi-step process to generate spermatozoa or oocytes. Methods for in vitro gametogenesis from pluripotent stem cells (PSCs) would provide a powerful tool to explore the mechanism of germ-cell development and its anomalies (*1*). Mouse PSCs have been induced into PGC-like cells (mPGCLCs), which contribute to spermatogenesis upon transplantation into testes (*2, 3*) and to oogenesis upon aggregation with embryonic ovarian somatic cells (reconstituted ovaries) followed by transplantation under ovarian bursa (*4*) or an appropriate culture (*5*). Resultant gametes from these procedures generate fertile offspring (*2-5*). Furthermore, human PSCs (hPSCs) have been induced into hPGCLCs bearing a gene-expression property of hPGCs just after their specification, opening the possibility for human in vitro gametogenesis (*6, 7*). However, further differentiation of hPGCLCs has not been successful, and whether hPGCLCs can develop as mature germ cells remains unknown.

At week 2 (Wk2) of development, hPGCs express key transcription factors (TFs) such as SOX17, TFAP2C, and BLIMP1 (also known as PRDM1). Then at Wk5 they migrate to and colonize the embryonic gonads to initiate differentiation into oogonia or gonocytes in embryonic ovaries or testes (*8, 9*) that express RNA regulators such as DAZL and DDX4 (also known as human VASA homolog: hVH) (*10-12*). Oogonia and gonocytes are very similar in morphology, gene expression, and epigenetic properties until they start sexual differentiation at Wk10 into oocytes via meiotic prophase or fetal spermatogonia (*8-12*). A characteristic event in germ-cell development is epigenetic reprogramming, which occurs by Wk10 and leads to genome-wide DNA demethylation, imprint erasure, and X reactivation (*10-12*).

We explored whether hPGCLCs can undergo further development in vitro in xenogeneic

reconstituted ovaries (xrOvaries) with mouse embryonic ovarian somatic cells (Fig. S1A). First, we induced male human induced PSCs (hiPSCs) with the *BLIMP1-tdTomato*; *TFAP2C-EGFP* alleles [585B1 BTAG (XY)] into incipient mesoderm-like cells (iMeLCs) and then into hPGCLCs (*7, 13*). BLIMP1 and TFAP2C are expressed in oogonia/gonocytes at least until Wk10 (*10-12*). We isolated BTAG-positive (BT^+AG^+) hPGCLCs at day 6 (d6) of induction by fluorescence activated cell sorting (FACS), and generated xrOvaries. In agreement with a previous report (*5*), mPGCLCs differentiated efficiently into primary oocytes and formed secondary follicles after a 21-day culture in rOvaries (Fig. S1B). At culture day 7 (ag7), the xrOvaries exhibited a round and flattened shape, and the BT^+AG^+ cells were distributed uniformly within the xrOvaries (Fig. S1C). Subsequently, the xrOvaries expanded laterally with the formation of cyst-like structures (Fig. 1A). From ag21 to ag77, the xrOvaries exhibited auto-fluorescence under fluorescence microscopy (Fig. S1D). There were about 2000 BT^+AG^+ cells per xrOvary at ag21 and ~500 at ag77 (Fig. S1E), indicating that only a fraction of the initial hPGCLCs (~5,000) survived in xrOvaries.

At ag77, the AG^+ cells in xrOvaries existed as clusters, were positive for a human mitochondrial antigen, bore faint DAPI staining, and were delineated by $FOXL2^+$ mouse granulosa cells and their basement membrane (Fig. S2A-C). The intensity of the DAPI staining in AG^+ cells appeared to decline progressively during the culture (Fig. S2D). The AG^+ cells expressed key TFs for early germ cells (TFAP2C, SOX17, POU5F1), and some were mitotically active ($Ki67^+$) or in apoptosis (cleaved $CASPASE3^+$) (Fig. S2A). Importantly, at ag77, many AG^+ cells up-regulated DAZL and DDX4 (Fig. 1B, Fig. S2E), suggesting that in xrOvaries, hPGCLCs not only survive as germ cells, but also differentiate into oogonia/gonocytes. Accordingly, at ag77, electron microscopy revealed the presence of large cells, highly similar to oogonia/gonocytes, with clear cytoplasms with sparsely located mitochondria and with round nuclei with loosely packed chromatin and prominent nucleoli (Fig. S3) (*8, 9*).

We next generated female hiPSCs bearing the AG; *DDX4/hVH-tdTomato* alleles [1390G3 AGVT (XX)] (Fig. S4) (*13*), created xrOvaries, and cultured them up to ag120. The female AG^+ hPGCLCs in xrOvaries developed similarly to the male hPGCLCs (Fig. S5), and up-regulated VT at ag77 and the AG^+VT^+ cells appeared to differentiate into AG-negative (AG^-) VT^+ cells at ag120 (Fig. 1C, S5A). We examined the expression of key genes by quantitative (q)-PCR. Both male and female hPGCLC-derived cells expressed early germ-cell genes (*BLIMP1*, *TFAP2C*, *SOX17*, *NANOS3*), core/naïve pluripotency genes (*POU5F1*, *NANOG*, *TCL1B*, *TFCP2L1*), and up-regulated genes for oogonia/gonocytes (*DAZL*, *DDX4*) from around ag35/49 onwards, and also genes for meiosis (*SYCP3*, *REC8*) (Fig. S6). The AG^-VT^+ cells at ag120 down-regulated early germ-cell and core/naïve pluripotency genes, and up-regulated *STRA8*, a gene essential for meiosis initiation (Fig. S6), suggesting their developmentally advanced character.

Consistently, the DDX4^{+} cells at ag120 expressed SYCP3, but not γH2AX, DMC1 and SYCP1 proteins, indicating that they have not yet initiated meiotic recombination (Fig. S7).

We analyzed the transcriptomes of these cell types (Fig. S8A, Table S1). Unsupervised hierarchical clustering (UHC) showed that hPGCLC-derived cells are distinct from hiPSCs/iMeLCs, and can be sub-classified according to their culture period (Fig. S8B). Principle component analysis (PCA) gave a concordant result (Fig. S8C). We identified the genes with significantly positive or negative PC1/2 loadings (453 genes) (Fig. S8D, Table S2). UHC classified them into five major clusters (Fig. 2A, S8D, Table S2): Cluster 1 represents the genes up-regulated upon hPGCLC specification or early in xrOvaries and expressed essentially continuously thereafter. Clusters 2 and 5 signify the genes transiently expressed in early hPGCLC-derived cells or down-regulated early in xrOvaries, and cluster 4 represents the genes down-regulated upon hPGCLC specification (Fig. 2A, Table S2). Cluster 3, which shows a progressive and coordinated up-regulation from ag35 onwards, signifies genes for oogonia/gonocytes and is enriched in genes with gene ontology (GO) functional terms for male meiosis, fertilization, and piRNA metabolic process (Fig. 2A, Table S2). Using the 453 genes, we compared the gene-expression properties of hPGCLC-derived cells with those of oogonia/gonocytes (*12*). hPGCLC-derived cells from ag35 onward, particularly the ag77 BT^{+}AG^{+}/ag120 AG^{+} cells, exhibited a strong similarity to Wk7/9 oogonia/gonocytes (Fig. 2B, Fig. S10).

The ag120 AG^{-}VT^{+} cells exhibited a developmentally advanced character (Fig. S6, S7, S8B, C). To clarify this point further, we examined the expression of genes that distinguish relevant human fetal germ cells [FGCs: mitotic (oogonia/gonocytes), retinoic acid (RA)-responsive (female), meiotic (female), oogenesis (female)] (*14*) in ag120 AG^{-}VT^{+} cells, which, intriguingly, revealed their similarity to RA-responsive FGCs. They down-regulate genes for early germ cells, further up-regulate *DAZL*, *DDX4*, *MAEL* and *KRBOX1*, up-regulate RA- or bone morphogenetic protein (BMP)-responsive genes (*STRA8*, *REC8*, or *ID1/2/3/4*, *MSX1/2*), yet do not sufficiently up-regulate key meiosis genes (*SYCP1*, *DMC1*, *SPO11*, *PRDM9*) (Fig. 2C, S11). Thus, hPGCLC development in xrOvaries reconstitutes human germ-cell development, albeit with protracted kinetics, leading to the generation of oogonia and RA-responsive FGCs, a state responding to the signal for the meiotic entry and in preparation for the meiotic recombination (*14*).

We determined the genome-wide DNA methylation (5-methylcytosine: 5mC) profiles of hPGCLC-derived cells in xrOvaries by whole-genome bisulfite sequencing (WGBS) (Fig. S12A, Table S3). The genome-wide 5mC levels of hiPSCs/iMeLCs were both ~80%, and they decreased progressively in hPGCLC-derived cells, reaching ~20% in ag77 cells and ~13% in ag120 cells (Fig. S12B), the levels comparable to that in oogonia/gonocytes at Wk7/10 (*11, 12*). The demethylation occurred throughout the

genome (Fig. 3, Fig. S12C, S13A, B), and the 5mC distribution profiles of the ag77/ag120 cells were very similar to those of oogonia/gonocytes at Wk7/Wk10 (Fig. 3), respectively, but not to those of the blastocysts (*15*) and naïve hESCs (*16, 17*) [and hPGCLCs reported by others (*18*)] (Fig. S12D). Thus, hPGCLC-derived cells demethylate their 5mCs in a fashion similar to oogonia/gonocytes, but not early embryonic cells and their putative in vitro counterparts.

We examined whether hPGCLC-derived cells can erase the parental imprints. Blastocysts (*15*) and somatic cells (*11*) exhibited ~50% CpG methylation in the differentially methylated regions (DMRs) of paternally and maternally imprinted genes (Fig. 4A, Table S4). In contrast, hiPSCs and primed hESCs (*12, 16, 17*) exhibited hyper-methylation in some DMRs, whereas naïve hESCs (*16, 17*) showed hypo-methylation in nearly all the DMRs (Fig. 4A), indicating that hPSCs mis-regulate the imprint states. Similar to oogonia/gonocytes, hPGCLC-derived cells progressively erased the parental imprints, including the hyper-methylated DMRs of hiPSCs (Fig. 4A). We determined the CpG sequences bearing a hyper-methylation in hiPSCs compared to three independent hESC lines (Table S5). hPGCLC-derived cells erased such hiPSC-specific methylation in a nearly complete fashion (Fig. S13C). Repeat elements are also demethylated in hPGCLC-derived cells, and the demethylation-resistant repeats in vivo, such as ERVK and SVA (*11, 12*), showed similar resistance in hPGCLC-derived cells (Fig. S13D). We defined the "escapees" that retained relatively high 5mC levels in ag77/ag120 cells and the Wk7 oogonia/gonocytes (Table S6). Essentially all the escapees in the oogonia/gonocytes (*12*) were included in those in the hPGCLC-derived cells, which were enriched around SVA, ERVK, and ERV1 (Fig. S13E).

We explored whether hPGCLC-derived cells can reactivate the inactive X chromosome (Xi). RNA fluorescence in situ hybridization (RNA FISH) revealed that the 1390G3 AGVT hiPSCs (passage 29) lack the expression of *XIST* from both alleles and show mono-allelic expression of the X-linked genes (Fig. S14A-D), indicating that they bear one active X chromosome (Xa) and one Xi without *XIST* (Xi^{XIST-}) (*19*). Accordingly, the promoters of the X-linked genes in 1390G3 AGVT hiPSCs exhibited an intermediate (~50%) 5mC level (Fig. 4B, S14E, S15), suggesting that they are unmethylated on the Xa and nearly fully methylated on the Xi (*19*). Notably, the ag120 cells, but not d6 hPGCLCs, exhibited a partial (~20%) reactivation, i.e., bi-allelic expression, of several X-linked genes, but did not re-activate *XIST* (Fig. S14A-D). Consistently, the promoters of the X-linked genes in ag120 cells were moderately demethylated (~20%) (Fig. 4B, S14E, S15). These findings demonstrate that hPGCLC-derived cells in xrOvaries undergo proper epigenetic reprogramming, i.e., genome-wide DNA demethylation, imprint erasure, and extinguishment of aberrantly acquired/persisting methylation of hiPSCs, reaching minimum 5mC levels (~13%) comparable to those reported in human germ cells. Nonetheless, the Xi^{XIST-} state in vitro is more resistant to reprogramming,

indicating a distinctive epigenetic mechanism for the Xi in hPSCs, which warrants further investigation.

We have provided evidence that primed hiPSCs (*20*) are competent to generate germ cells with their hallmark of epigenetic reprogramming. In mouse-mouse rOvaries, ovarian somatic cells, likely granulosa cells, provide timely signals/environments for mPGCLCs to mature into primary oocytes and form secondary follicles (*5*). In contrast, in xrOvaries, mouse granulosa cells create a permissive environment for hPGCLCs to gradually mature into oogonia. Although the underlying mechanism remains unclear, since mPGCLCs undergo epigenetic reprogramming upon expansion and differentiate into oocytes in response to BMP and RA (*21, 22*), hPGCLC-derived cell proliferation and signals from mouse granulosa cells (Fig. S16) would allow hPGCLC-derived cells to mature into oogonia/RA-responsive FGCs (Fig. S17). Since both male and female mPGCLCs enter into meiotic prophase under the same condition (*22*), male hPGCLCs would also enter into meiotic prophase in xrOvaries. Future studies will include exploring a strategy and the mechanism for the differentiation of hiPSC-induced oogonia into oocytes with meiotic recombination or for the differentiation of hPGCLCs into fetal spermatogonia, which promote human in vitro gametogenesis.

REFERENCES AND NOTES

1. M. Saitou, H. Miyauchi, Gametogenesis from Pluripotent Stem Cells. *Cell Stem Cell* **18**, 721-735 (2016).
2. K. Hayashi, H. Ohta, K. Kurimoto, S. Aramaki, M. Saitou, Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519-532 (2011).
3. Y. Ishikura *et al.*, In Vitro Derivation and Propagation of Spermatogonial Stem Cell Activity from Mouse Pluripotent Stem Cells. *Cell Reports* **17**, 2789-2804 (2016).
4. K. Hayashi *et al.*, Offspring from oocytes derived from in vitro primordial germ cell-like cells in mice. *Science* **338**, 971-975 (2012).
5. O. Hikabe *et al.*, Reconstitution in vitro of the entire cycle of the mouse female germ line. *Nature* **539**, 299-303 (2016).
6. N. Irie *et al.*, SOX17 is a critical specifier of human primordial germ cell fate. *Cell* **160**, 253-268 (2015).
7. K. Sasaki *et al.*, Robust In Vitro Induction of Human Germ Cell Fate from Pluripotent Stem Cells. *Cell Stem Cell* **17**, 178-194 (2015).
8. T. G. Baker, A Quantitative and Cytological Study of Germ Cells in Human Ovaries. *Proc R Soc Lond B Biol Sci* **158**, 417-433 (1963).
9. T. Fukuda, C. Hedinger, P. Groscurth, Ultrastructure of developing germ cells in the fetal human testis. *Cell Tissue Res* **161**, 55-70 (1975).
10. S. Gkountela *et al.*, DNA Demethylation Dynamics in the Human Prenatal Germline. *Cell* **161**, 1425-1436 (2015).
11. F. Guo *et al.*, The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. *Cell* **161**, 1437-1452 (2015).
12. W. W. Tang *et al.*, A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. *Cell* **161**, 1453-1467 (2015).
13. S. Yokobayashi *et al.*, Clonal variation of human induced pluripotent stem cells for induction into the germ cell fate. *Biol Reprod* **96**, 1154-1166 (2017).
14. L. Li *et al.*, Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. *Cell Stem Cell* **20**, 858-873 e854 (2017).
15. H. Okae *et al.*, Genome-wide analysis of DNA methylation dynamics during early human development. *PLoS Genet* **10**, e1004868 (2014).
16. Y. Takashima *et al.*, Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254-1269 (2014).
17. W. A. Pastor *et al.*, Naive Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory. *Cell Stem Cell* **18**, 323-329 (2016).
18. F. von Meyenn *et al.*, Comparative Principles of DNA Methylation Reprogramming during Human and Mouse In Vitro Primordial Germ Cell

Specification. *Dev Cell* **39**, 104-115 (2016).
19. S. Patel *et al.*, Human Embryonic Stem Cells Do Not Change Their X Inactivation Status during Differentiation. *Cell Reports* **18**, 54-67 (2017).
20. T. Nakamura *et al.*, A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature* **537**, 57-62 (2016).
21. H. Ohta *et al.*, In vitro expansion of mouse primordial germ cell-like cells recapitulates an epigenetic blank slate. *EMBO J* **36**, 1888-1907 (2017).
22. H. Miyauchi *et al.*, Bone morphogenetic protein and retinoic acid synergistically specify female germ-cell fate in mice. *EMBO J* **36**, 3100-3119 (2017).

ACKNOWLEDGMENTS
We thank the members of our laboratory for their helpful input on this study: K. Hayashi for his advice on rOvaries, and Y. Nagai, Y. Sakaguchi, and M. Kawasaki of the Saitou Laboratory, J. Oishi of the Sasaki Laboratory, and T. Sato and M. Kabata of the Yamamoto Laboratory for their technical assistance, and the Center for Anatomical, Pathological and Forensic Medical Research at Kyoto University for the preparation of electron microscopy. We also thank T. Mori for his encouragement and support. **Author contributions:** C.Y. performed hPGCLC induction, and xrOvary formation and analysis. K.S., Y.K. and S.Y. assisted in hPGCLC induction, and Y.I. assisted in xrOvary formation. T.N. and T.Y. contributed to the RNA-seq/WGBS, Y.M., K.S. and H.S. contributed to the WGBS, and Y.Y. contributed to the analyses of RNA-seq/WGBS data. I.O. performed the RNA FISH. Y.K. and M.S. designed the experiments and wrote the manuscript. **Funding:** This work was supported by a Grant-in-Aid for Specially Promoted Research from JSPS (17H06098) to M.S., by a JST-ERATO Grant (JPMJER1104) to M.S., and by a Grant-in-Aid for Scientific Research on Innovative Areas from JSPS to H.S. (25112010). **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The accession numbers for the RNA-seq and WBGS data generated in this study are GSE117101 (GEO) and DRA006618/DRA007077 (DDBJ), respectively.

SUPPLEMENTARY MATERIALS
Materials and Methods
Figs. S1 to S17
Tables S1 to S6
References (23-34)

Figure Legends

Fig. 1. hPGCLC differentiation in xrOvaries.

(A) Bright field (BF) images and FACS by BTAG (*BLIMP1-tdTomato*; *TFAP2C-EGFP*) of xrOvaries with 585B1 BTAG (XY) hPGCLC-derived cells from ag7 to ag77. Bars in (A, C), 500 μm.

(B) Expression of DDX4 (magenta) in AG^+ cells (yellow) in xrOvaries at ag77, with FOXL2 (cyan) and DAPI (white) staining. The boxed area in the left panel (merged image) is magnified in the right panels. Bars, 20 μm.

(C) A BF image and FACS by AGVT (*DDX4/hVH-tdTomato*) of xrOvaries at ag120 with 1390G3 AGVT (XX) hPGCLC-derived cells.

Fig. 2. Transcriptome dynamics of hPGCLC-derived cells in xrOvaries.

(A) (left) Unsupervised hierarchical clustering (UHC) and heatmap of the expression in the indicated cells of the 453 signature genes for the transitions of hPGCLC-derived cell properties in xrOvaries (Fig. S8C, D, Table S2). (right) Representative genes in each cluster and their GO enrichments (Table S2).

(B, C) Comparison of the expression of the 453 (B) or indicated genes (*14*) (C) between ag77 BT^+AG^+ and ag120 AG^+VT^+ cells (B) or ag120 AG^-VT^+ cells (C) and the indicated human germ cells (*12, 14*).

Fig. 3. Genome-wide DNA demethylation in hPGCLC-derived cells in xrOvaries.

Comparisons of the 5mC levels (genome-wide 2-kb windows) by contour representations of scatter plots, combined with histogram representations (top and right of scatter plots), between hiPSCs (top) or ag77 BT^+AG^+ cells (bottom, left)/ag120 AG^+VT^+ cells (bottom, right) and the indicated cell types.

Fig. 4. Imprint erasure and X chromosome demethylation in hPGCLC-derived cells in xrOvaries.

(A) Heatmaps showing the 5mC levels in the differentially methylated regions (DMRs) of the indicated imprinted genes in the indicated cells.

(B) Heatmaps (top) and violin plots (bottom, average: red bars) showing the 5mC levels in the CpG islands on the X chromosomes in the indicated cells.

Figure 1, Yamashiro et al.

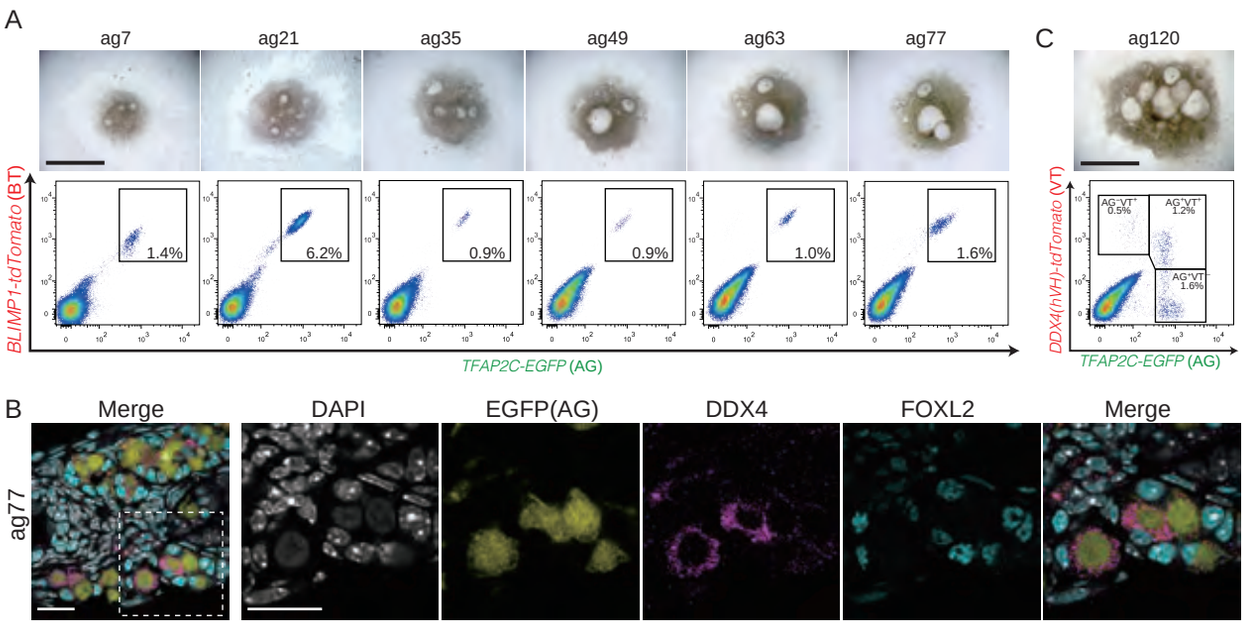


Figure 2, Yamashiro et al.

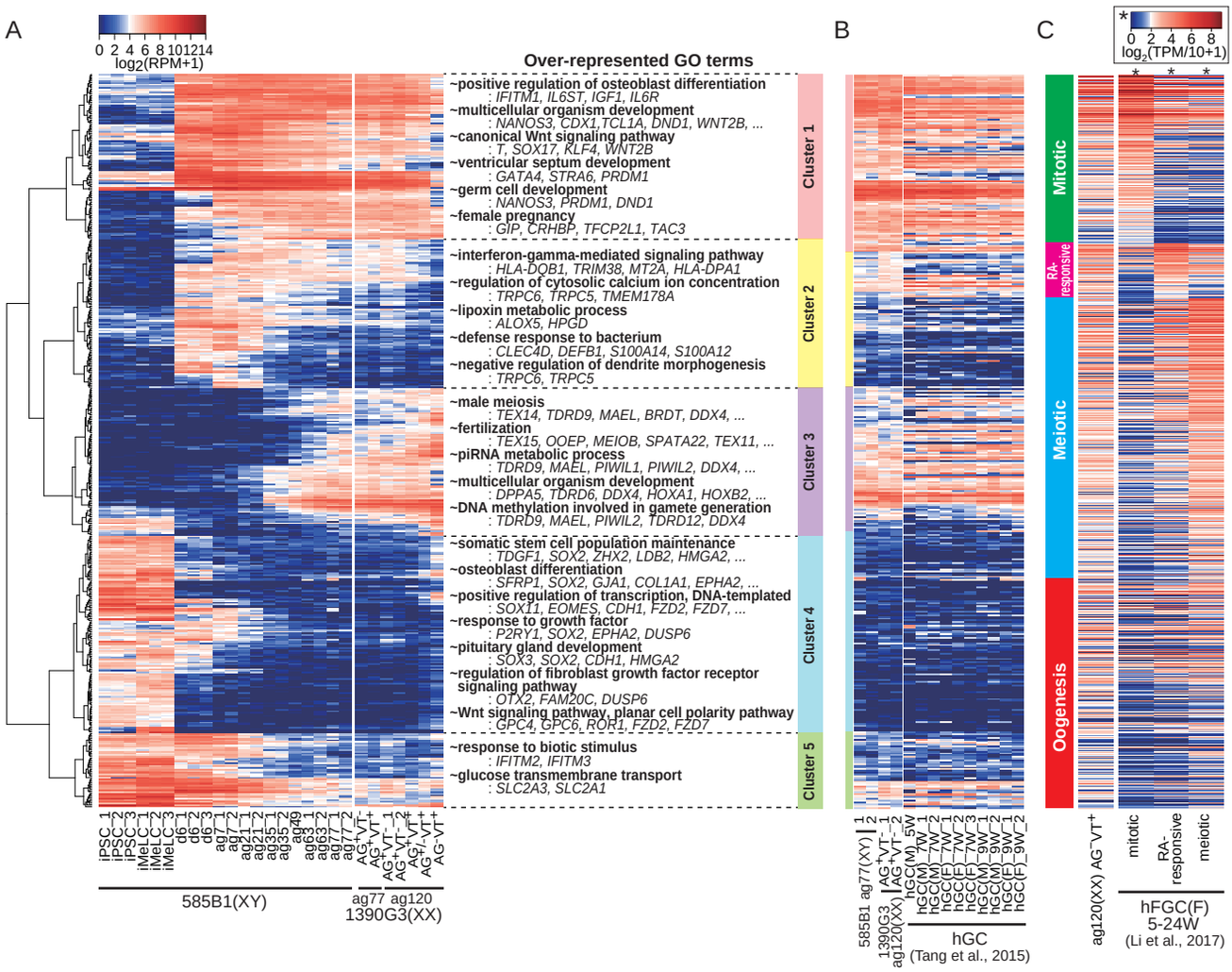


Figure 3, Yamashiro et al.

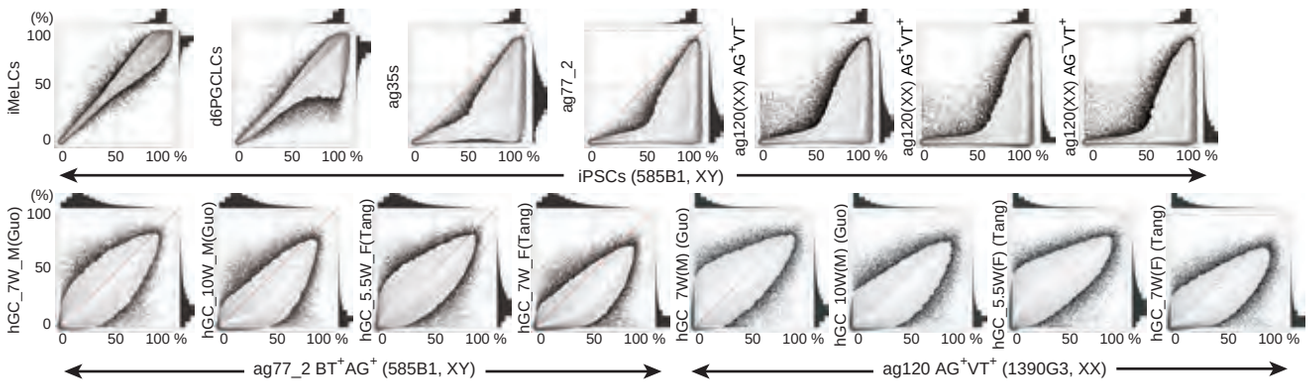
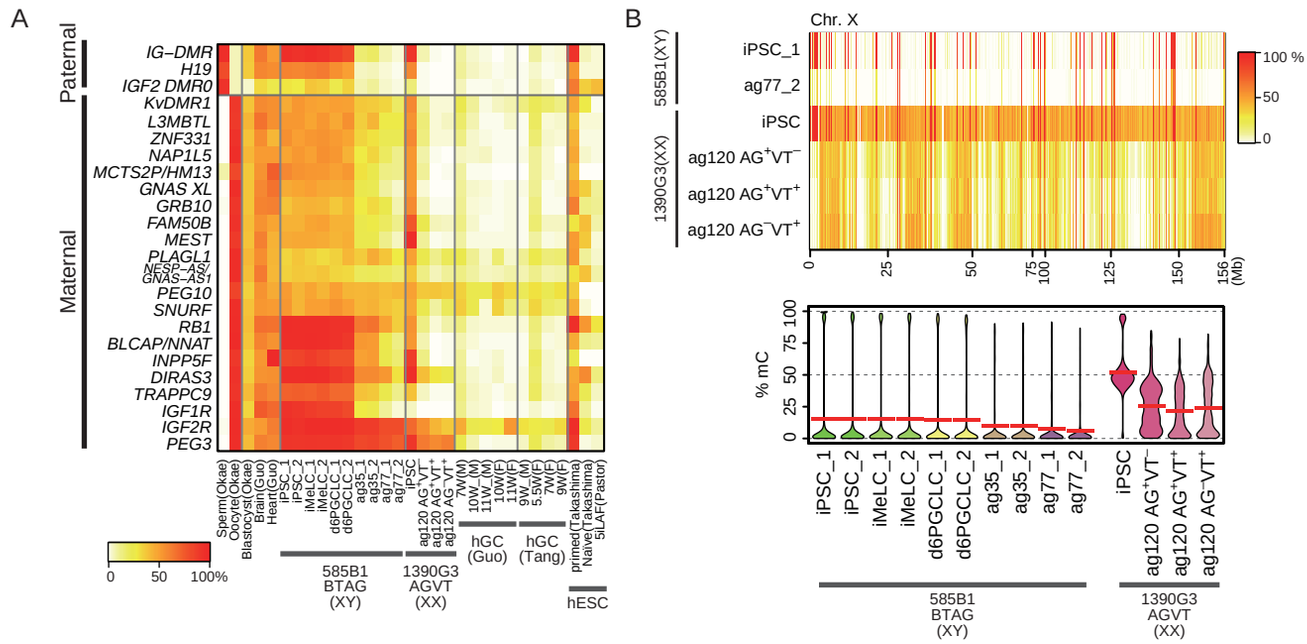


Figure 4, Yamashiro et al.



Supplementary Materials

Materials and Methods

Animals

All the animal experiments were performed under the ethical guidelines of Kyoto University. Pregnant ICR female mice were purchased from Japan SLC.

mESC culture, mPGCLC induction

The mESC culture and mPGCLC induction were performed essentially as described previously (*2, 4*). Briefly, mESCs [H18 BVSC (XX)] (*4*) were maintained under a naïve state condition in N2B27 medium with 2 chemical inhibitors (PD0325901, 0.4 μM: Stemgent, San Diego, CA; CHIR99021, 3 μM: Stemgent) and leukemia inhibitory factor (LIF) (1000 u/ml) on a dish coated with poly-L-ornithine (0.01%; Sigma) and laminin (300 ng/ml; BD Biosciences). Epiblast-like cells (EpiLCs) were induced by plating 1.0×10^5 ESCs on wells of a 12-well plate coated with human plasma fibronectin (16.7 ug/ml; Millipore) in N2B27 medium containing activin A (20 ng/ml; Peprotech), basic fibroblast growth factor (bFGF) (12 ng/ml; Life Technology) and Knock-out serum replacement (KSR) (1%; GIBCO, 10828-028). At 48 hrs after the EpiLC induction, PGCLCs were induced under a floating condition by plating 2.5×10^3 EpiLCs in wells of a low-cell-binding 96-well Lipidure-coated plate (Thermo Fisher) in 200 μl of GK15 medium [Glasgow's Minimal Essential Medium (GMEM) (GIBCO, 11710-035) supplemented with 15% KSR, 1 mM sodium pyruvate (GIBCO, 11360-070), 0.1 mM non-essential amino acids (GIBCO, 11140-050), 2 mM L-glutamin (GIBCO, 25030-081), 100 U/ml penicillin/streptomycin] with bone morphogenetic protein 4 (BMP4) (500 ng/ml; R&D Systems), LIF (1000 u/ml; Merck Millipore), stem cell factor (SCF) (100 ng/ml; R&D Systems), and epidermal growth factor (EGF) (50 ng/ml; R&D Systems).

Human iPSC culture

The experiments on the induction of hPGCLCs from hiPSCs were approved by the Institutional Review Board of Kyoto University and were performed according to the guidelines of the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan. 585B1 BTAG (XY) (*7*) and 1390G3/1390G3 AGVT (XX) hiPSCs (*13*) (see below) were cultured on a plate coated with recombinant laminin511 E8 (iMatrix-511; Nippi, 892014) and were maintained under a feeder-free condition in the AK03 medium (Ajinomoto, Tokyo) at 37 °C under an atmosphere of 5% CO_2 in air. To passage the hiPSCs, the cells were dissociated into single cells by incubating in a mixture of an equal volume of TrypLE select (GIBCO, 12563-011) and 0.5 mM EDTA/PBS (−). 10 μM of ROCK inhibitor (Y-27632) (Wako Pure Chemical Industries, 251-00514) was added and left to incubate for 24 hrs after the passage.

Generation of the AGVT-knockin reporter hiPSCs

The donor vector for generating the *TFAP2C-p2A-EGFP* (AG) allele was as described previously (*7*). To construct the donor vector for generating the *DDX4/hVH-tdTomato* (VT) allele, the homology arm of *DDX4* (the region from the 1471 base upstream to the 1291 base downstream of the stop codon) was amplified from the genomic DNA of 1390G3 iPSCs by PCR (see **Primers**), and was sub-cloned into the pCR2.1 vector using the TOPO TA cloning kit (Life Technologies; K40001SP). The *p2A-tdTdomato* fragment with the *PGK-Neo* cassette flanked by *loxP* sites was also amplified by PCR from the 585B1 BTAG hiPSCs (*7*) and inserted in-frame at the 3-prime end of the *DDX4* coding sequence of the sub-cloned vector using the GeneArt Seamless Cloning& Assembly Kit (Life Technologies, A13288). The stop codon was deleted to express the in-frame fusion protein. The *MC1-DT-A-polyA* cassette was subsequently inserted into the downstream region of the right homology arm of the VT donor vector using the restriction enzymes NotI/XbaI.

The TALEN constructs targeting the sequence close to the stop codon of *DDX4* were generated using a GoldenGate TALEN and TAL effector Kit (Addgene, #1000000016) as described previously (*7, 23*). TALEN's RVD sequences were as follows: [left TALEN] HD HD HD NI NI NG HD HD NI NN NG NI NN NI NG NN NI NG NN; [right TALEN] NN NI NI NN NN NI NG NN NG NG NG NG NN NN HD NG NG. The activity of the TALENs was validated by single-strand annealing (SSA) assay. The pGL4-SSA empty reporter plasmid, the pGL4-SSA-*HPRT1* reporter plasmid, and the TALEN pair targeting human *HPRT1* were used as a control in this assay.

The donor vectors (5 μg each) and the TALEN plasmids (2.5 μg each) were introduced into the 1390G3 hiPSCs by electroporation using a super electroporator (Nepagene, NEPA 21 type II). Single colonies were picked up after selection with neomycin, and random or targeted integration were evaluated by PCR (see **Primers**) on the genomic DNA purified by GenElute Mammalian Genomic DNA Miniprep kits (SIGMA, G1N350). The lines with the correct targeting were transfected using a plasmid expressing Cre recombinase to remove the *PGK-Neo* cassette, and verified by PCR or Southern blot analysis. The G-band analyses were performed by Nihon Gene Research Laboratories.

Induction of PGCLCs and the generation of reconstituted ovaries

hPGCLCs were induced from hiPSCs via iMeLCs as described previously (*7, 13*). Xenogeneic reconstituted ovaries (xrOvaries) were generated by aggregating d6 hPGCLCs with fetal ovarian somatic cells of E12.5 embryos according to a procedure essentially as described for the generation of isogeneic rOvaries (*4, 5*). Briefly, the d6 aggregates of iMeLCs induced from the 585B1 BTAG (XY)/1390G3 AGVT (XX) hiPSCs were dissociated into single cells, and the BT^+AG^+ or AG^+ cells (d6 hPGCLCs) were collected by FACS (see **Fluorescence-Activated Cell Sorting**). To isolate E12.5

mouse embryos, pregnant ICR females were sacrificed by cervical dislocation, and the embryos were dissected in chilled DMEM (GIBCO, 10313-021) containing 10% FBS (Hyclone), 2 mM GlutaMax (GIBCO, 35050-061), 10 mM HEPES (GIBCO, 15630-106), and 100 U/ml penicillin/streptomycin (GIBCO, 15070). Fetal ovaries were identified by their appearance, and the mesonephroi were dissected out with tungsten needles. The isolated ovaries were dissociated into single cells, and endogenous mouse PGCs were selected out by MACS (see **Magnetic-Activated Cell Sorting**). Compared to Fluorescence-activated cell sorting (FACS), MACS requires much less time and less physical stress to collect the cells for rOvary formation. The hPGCLCs (5,000 cells/well) and the fetal ovarian somatic cells (50,000 cells/well or when using frozen stocks, 75,000 cells/well) were plated on wells of a Lipidure-coated U-bottom 96-well plate (Thermo Fisher Scientific, 174925) and cultured under a floating condition in GK15+Y medium [GK15 and 10 μM ROCK inhibitor]. After a period of 2 days for the floating culture, using a glass capillary, the xrOvaries were transferred onto Transwell-COL membrane inserts (Corning, 3496) soaked in alpha-Minimum Essential Medium (GIBCO, 32571-036) containing 10% FBS, 55 μM 2-mercaptoethanol (GIBCO, 21985-023), 150 nM L-ascorbic acid (SIGMA, A4403-100MG), and 100 U/ml penicillin/streptomycin. Half the medium was changed every three days. All xrOvaries were cultured at 37 °C under an atmosphere of 5% CO_2 in air.

Histology and immunofluorescence (IF) analysis

For the histological analysis, xrOvaries were fixed with 10% buffered formalin (Nacalai Tesque, 37152-51) for 6-12 hrs at room temperature, then embedded in 1.5% agarose (Nacalai Tesque, 01158-85) and washed with distilled water for 2 hrs. The agarose blocks containing xrOvaries were dehydrated successively in 70%, 80%, 90% and 100% ethanol (Nacalai Tesque, 14713-53) and treated with chloroform (Nacalai Tesque, 038-02606) for cleaning. The blocks were then replaced with paraffin wax, and embedded in clean paraffin wax. They were sliced by a microtome (Leica, CX41) at a thickness of 5 μm, and the sections were placed on a glass slide coated with Platinum-Pro (Matsunami, 83-1922) with distilled water. The paraffin sections were re-hydrated gradually in Xylene (Nacalai Tesque, 36612-93) and serial concentrations (100%, 90%, 80%, 70%) of ethanol, and stained with hematoxylin and eosin. They were mounted in Malinol (Muto Pure Chemicals, 2009-3), and were observed by an upright microscope (Olympus, BX53) equipped with a CCD camera (Olympus, DP70).

For the IF analysis, xrOvaries were fixed with 2% paraformaldehyde (Nacalai Tesque, 26126-25) in PBS for 3-6 hrs on ice, washed three times with PBS containing 0.2% Tween-20 (PBST), and replaced with serial concentrations (10%, 30%) of sucrose (Nacalai Tesque, 30404-45) in PBS overnight at 4 °C. The samples were embedded in the OCT compound (Sakura Finetek, 4583), frozen, and cryosectioned at a thickness of 10 μm at −20 °C on a cryostat (Leica, CM1850). The sections were placed on a glass slide

coated with Platinum-Pro and dried completely. They were washed with PBS three times, then incubated in PBST containing 10% normal donkey serum (NDS) (Jackson Laboratories, 017-000-121) for 30 min, followed by incubation with the primary antibodies in 5% NDS in PBST for 2 hrs at room temperature. The sections were washed four times with PBS and incubated with the secondary antibodies and 1 μg/ml DAPI in 5% NDS in PBST for 50 min at room temperature in darkness. They were then washed four times in PBS, and mounted in VECTASHIELD mounting medium (Vector Laboratories, H-1000) with 200 ng/ml DAPI. The samples were analyzed by using a confocal microscope (Olympus, FV1000). All antibodies used in this study are listed in **Antibodies**.

Electron microscopy

The xrOvaries were fixed with the Karnovsky fixative (2% glutaraldehyde, 4% paraformaldehyde in 0.1M phosphate buffer pH7.4) for 24 hrs at 4℃. The samples were further fixed with 1% osmium tetroxide in 0.1 M phosphate buffer (pH7.2) for 90 min, and dehydrated with a graded series of ethanol. Then the samples were penetrated in propylene oxide, embedded in epoxy-resin and polymerized at 60℃ for 3 days. The samples were sliced into ultrathin sections (70 nm) with an ultramicrotome (Leica, EM UC6) and mounted on mesh grids. The ultrathin sections were stained with uranyl acetate and lead citrate, and then were observed using a transmission electron microscope (HITACHI, H-7650).

Fluorescence-activated cell sorting (FACS)

The d6 floating aggregates of iMeLCs induced from the hiPSCs were incubated in 0.25% Trypsin-EDTA (GIBCO, 15400-054) in PBS for 18 min at 37 °C, with periodical intermittent pipetting, and the dissociates were quenched by FBS, followed by pipetting to generate a single-cell suspension. After washing with PBS containing 0.8% BSA fraction V (GIBCO, 15260-037), the cell suspension was filtered through a nylon cell strainer (FALCON, 352350), and the BT^+AG^+ or AG^+ cells (d6 hPGCLCs) were sorted with a flow cytometer (BD Bioscience, AriaIII), and collected in GK15+Y medium.

For the analysis and the sorting of the BT^+AG^+ or $AG^+VT^{+/-}$ cells in the xrOvaries, the xrOvaries at various culture periods were fragmented by scratching using 30 gauge needles and were incubated in TrypLE express (GIBCO, 12604-021) for 20 min, with periodical intermittent pipetting. The resultant cell suspension was washed with PBS containing 0.1% BSA fraction V, and was filtered through a nylon cell strainer (FALCON, 352350). The BT^+AG^+ or AG^+ cells (d6 hPGCLCs) were sorted with a flow cytometer (BD Bioscience, AriaIII), and were collected in CELLOTION (Nippon Zenyaku Kogyo, CB051).

Magnetic-activated cell sorting (MACS)

Isolated mouse fetal ovaries at E12.5 were dissociated with TrypLE express for 13 min at 37 °C, and were quenched with 5 times the volume of DMEM/F12 medium (GIBCO, 11320-082) containing 0.1% BSA fraction V. The cell suspension was passed through a nylon cell strainer and centrifuged at 1,200 rpm for 3 min, and the supernatant was discarded. The cell pellet was re-suspended with MACS buffer (PBS containing 0.5% BSA fraction V and 2 mM EDTA), and then was incubated with anti-SSEA1 and anti-CD31 antibodies (Miltenyi Biotec, listed in **Antibodies,** MACS) for 20 min on ice. The cell suspension was washed with MACS buffer and centrifuged at 1,200 rpm for 3 min, and the supernatant was removed. The cell pellet was re-suspended in MACS buffer and then the cell suspension was applied to an MS column (Miltenyi Biotec, 130-042-201). The flow-through was washed with GK15+Y, and was used as fetal ovarian somatic cells for the generation of xrOvaries.

qPCR and RNA-seq analysis

The cells were lysed and the total RNAs were purified using an RNeasy Micro Kit (Qiagen, 74004) according to the manufacturer's instructions. 1 ng of total RNAs from each sample was used for the synthesis and the amplification of cDNAs. cDNA amplification, construction of cDNA libraries, and RNA sequencing using the Illumina NextSeq 500 system were performed as described previously (*3*).

qPCR on amplified cDNAs was performed using the Power SYBR Green Master Mix (Applied Biosystems, 4367659) on a real-time qPCR system (Biorad, CFX384). The expression level of each gene was evaluated relative to the averaged expression levels of *RPBP0/Rpbp0* and *PPIA/Ppia*. The primer sequences for the genes examined in this study are listed in **Primers**.

RNA FISH and DNA FISH on human iPSCs and PGCLCs

hiPSCs, hPGCLCs and hPGCLC-derived cells were transferred onto a Poly-L-lysine (Sigma) and Denhardt's-coated glass coverslip in a drop of PBS and allowed to adhere to the coverslip prior to fixation.

RNA FISH was carried out essentially as described previously (*24, 25*). For *XIST* RNA FISH, a combination of two probes covering 16 kb of *XIST* mRNA was used. For the detection of the nascent transcripts of X-linked by RNA FISH, the following bacterial artificial chromosome (BAC) probes (CHORI) were used: *TBL1X* (RP11-451G24), *HUWE1* (RP11-155O24), *ATRX* (RP11-42M11) and *XACT* (RP11-35D3). The correct chromosomal locations of the BACs were verified using DNA FISH on metaphase spreads.

For RNA FISH, coverslips carrying the cells were fixed for 10 min in 3% paraformaldehyde (PFA) (pH7.4), permeabilized on ice for 3 min in 0.5% Triton

X-100/PBS, and stored in 70% ethanol at −20℃. After dehydrating through an ethanol series, they were hybridized with the fluorescent probes at 37℃ overnight. Coverslips were counterstained with DAPI (1 μg/ml) and mounted in Vectashield (Vector Laboratories).

DNA-FISH was carried out as described previously (*24*, *25*). DNA FISH was performed following RNA FISH. Images of cells were first acquired, immediately following RNA FISH. The coverslips carrying the cells were then recovered in 2xSSC and incubated in Rnase A (100 μg/ml) in 2×SSC at 37℃ for 1 hr, before a brief rinse in 2×SSC, dehydration through an ethanol series (70%, 90%, 100%) and air drying. The DNA was then denatured in 70% formamide, 2×SSC for 30 min at 80℃ in an oven and dehydrated again through an ice-cold ethanol series. Hybridization with a fluorescent probe has been described previously. Coverslips were counterstained with DAPI (1 μg/ml), mounted and viewed under the fluorescence microscope.

Mapping reads of RNA-seq and conversion to gene-expression levels

RNA-seq read data were processed as an index of the abundance of transcripts as described previously. In brief, all reads were processed with the cutadapt v1.9.1 program to remove low quality bases and adaptor sequences. Reads of 30 bases or longer were mapped onto the human genome (GRCh38.p2/hg38) and ERCC spike-in RNA sequences with the Tophat v2.0.11/Bowtie v1.1.2 program with the "--bowtie1", "--library-type fr-secondstrand" and "--no-coverage-search" options. Mapped reads were subsequently processed with the Cufflinks v2.2.0 program with the "--compatible-hits-norm", "--no-length-correction", "--library-type fr-secondstrand", and "--max-mle-iterations 50000" options in order to estimate the abundance of the gene expressions. The reference transcript files (GFF3 file) were obtained from the NCBI ftp site and the 3' ends of transcripts were extended up to 10 kb for maximum read recovery as described previously (*26*). Samtools v1.3 (*27*) and Picard-tools v2.1.0 (http://broadinstitute.github.io/picard/) were used to analyze the mapped data. RNA-seq reads of public data were obtained from the DDBJ database ftp site and processed using the same procedure as described above except that the "--library-type fr-secondstrand" option was replaced by the "--library-type fr-unstranded" option for both Tophat and Cufflinks, and the "--no-length-correction" option in Cufflinks was not used for the abundance estimation of full-length RNA-seq data.

Data analysis of the RNA-seq

The RNA-seq data analysis was performed using the R software package (version 3.3.3, 6/3/2017). Genes showing maximum $\log_2$ (RPM+1) values > 4 [greater than ~10-20 copies per cell (*26*)] in at least one sample were used for analysis. Unsupervised Hierarchical Clustering (UHC) was performed using the hclust function with the Average method. The principal component analysis (PCA) was performed using the prcomp

function with and without scaling. Differentially expressed genes (DEGs) were defined as those bearing both $\log_2$ (RPM+1) > 4 and a more than 2 S.D. radius of scaled PC1 and PC2 values. The heatmap was generated using the heatmap.2 function in the gplot package. Gene ontology analysis was performed using DAVID (*28*).

Analysis of published RNA-seq data

To compare the gene-expression properties between hPGCLC-derived cells and human germ cells in vivo, we used the RNA-seq data on human germ cells published in Tang et al. (*12*) and Li et al. (*14*). To compare our data with those in Tang et al., genes showing maximum $\log_2$ (RPM+1) values > 4 in at least one of our samples were selected, and their $\log_2$ (RPM+1) values in d4 hPGCLCs were directly compared to their RPKM values in Tang's hPGCLCs for the regression analysis by the least square method. The RPKM values of genes expressed in human germ cells in Tang et al. were then adjusted according to this regression analysis and the adjusted relative values were used for generating a gene expression-level heatmap.

To compare our data with those in Li et al., genes showing maximum TPM values > 150 in at least one of the chosen samples (5, 8, 12, 14, 18, 20, 24 Wks; total 915 samples) were selected, and the samples with 75-percentile TPM values > 30 were sorted out as valid samples (415 samples). These samples were classified into 5 populations (mitotic/RA responsive/meiotic/oogenesis/somatic cells) by unsupervised hierarchical clustering (UHC), and a mean expression value of each gene was calculated for every population and converted to $\log_2$ (mean TPM value/10+1) for comparison.

Whole-genome bisulfite sequence analysis (WGBS)

Genomic DNAs of hiPSCs, iMeLC, and d6 hPGCLCs were purified using an All Prep DNA/RNA minikit (Qiagen) according to the manufacturer's instructions. The $BT^{+}AG^{+}$ or $AG^{+/-}VT^{+/-}$ cells isolated from xrOvaries at ag35, ag77 and ag120 (1,960 ~ 7,502 cells) were lysed directly in the cell-lysis buffer (0.1% SDS, 1 mg/ml proteinase K in DNase-free water) at 37 °C for 60 min and then at 98 °C for 15 min. Un-methylated lambda phage DNA (1/200 of the estimated gDNA mass) (Promega) was spiked into the samples, and then the samples were subjected to bisulfite conversions and library constructions for amplification-free whole-genome bisulfite sequencing analyses, using the post-bisulfite adaptor tagging (PBAT) method (*29*). Massively parallel sequencing was performed on Illumina HiSeq 2500 sequencing systems (HCS v2.0.5 and RTA v1.17.20 for the first trial; HCS v2.2.68 and RTA v1.18.66.3 for the subsequent trials) to generate 101-nucleotide single-end sequence reads. Cluster generation and sequencing were carried out in single-read mode using a TruSeq SR Cluster Kit v3-cBot-HS and TruSeq SBS Kit v3-HS (Illumina) according to the manufacturer's instructions.

Processing, mapping and conversion of the data for bisulfite sequencing

Methyl-seq read data were processed as an index of methylation levels as described previously (*30*). In brief, all reads were processed with Trim_Galore v0.4.1 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)/cutadapt v1.9.1 (http://cutadapt.readthedocs.io/en/stable/guide.html) with the "--clip_R1 4", "--trim1" and "-a AGATCGGAAGAGC" options. Qualified reads were mapped onto the human genome (GRCh38.p2/hg38) using Bismark v0.17.0/Bowtie2 v2.2.7 with the "--pbat" option. Then the bismark_methylation_extractor program in the same package was used to determine the cytosine and methyl-cytosine count at every CpG site on the genome.

Annotation of promoters, non-promoter CGIs, repetitive elements, and imprint DMRs

Promoters were defined as the regions between 900-bp upstream and 400-bp downstream of the transcription start sites (TSSs). Data for the CpG islands were obtained from Illingworth et al. (*31*) in the hg18 format, and converted into the hg38 format with the program LiftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver). Information on the repeat elements was obtained from the UCSC table browser website (https://genome.ucsc.edu/cgi-bin/hgTables). Data for the imprint loci in humans were obtained from Court et al. (*32*) in the hg19 format, and converted into the hg38 format with LiftOver.

Analysis of DNA methylation in single-copy genomic loci, promoters, and non-promoter CGIs

RStudio version 1.0.143 (R version 3.4.0) was used for the following analysis and to plot the data: All CpG sites with read depth ≥ 4 were used for the analysis. For the genome-wide analysis, the average percent methylations of the CpG sites in 2-kb non-overlapping bins were calculated. Bins with 4 or more CpGs were used for the analyses. Scatter plots were overlaid with contour plots for a clear view of the densely plotted areas. Histograms of the plots in the 5% intervals were shown at the sides of the scatter plots. Violin and beeswarm plots were drawn using the vioplot and beeswarm libraries, respectively.

Hyper-methylated regions in hiPSCs compared to hESCs in Fig. 3E were calculated as follows: First, we identified all the CpG sites with methylation levels ≥ 25% higher in hiPSCs [585B1 BTAG (XY) and 1390G3 AGVT (XX)] (*7, 13*) compared to hESCs reported previously—namely, H9 (a "primed" state) reported by Takashima et al. (*16*), H1 and H9 by Lister et al. (*33, 34*), H9 by Tang et al. (*12*), and UCLA1 by Pastor et al. (*17*). Neighboring CpG sites within 500 bp were combined, and the combined loci with more than five hyper-methylated CpGs (409 loci) were used for the analysis. "DNA demethylation escapees" were defined as described previously (*12*). Briefly, hyper-methylated regions were defined by applying the hypermr program in the MethPipe v3.4.3 package to the individual and pooled Wk7 to Wk9 hPGC data from Tang

et al. (*12*), resulting in the identification of 148,939 loci (median size = 2,124 bp). The same procedure was performed using our ag77_2 sample, resulting in the identification of 261,179 loci (median size = 1,808 bp). Overlap between the two "escapee" groups was calculated by BEDTools v2.17.0. Annotation of the "escapees" was performed using the annotatePeaks program in Homer v4.9.1 with the default settings.

Analysis of published WGBS data

All the public data used in this study are listed in Table S3. All the read data were processed using the same procedure as described above with sample-specific modifications as described in Table S3.

Accession numbers

The accession numbers of the data generated in this study: the RNA-seq data: GSE117101 (the GEO database); the WBGS data: DRA006618 and DRA007077 (the DDBJ database).

Antibodies

Immunofluorescence (*secondary antibodies)

Antibody	**Supplier**	**Product number**
Rabbit anti-OCT3/4	Abcam	ab19857
Mouse anti-SOX2	R&D Systems	MAB2018
Rabbit anti-TFAP2C	SantaCruz Biotechnology	sc-8977
Goat anti-SOX17	Neuromics	GT15094
Goat anti-NANOG	R&D Systems	AF1997
Chicken anti-GFP	Abcam	ab13970
Rat anti-GFP	Nacalai Tesque	00404-84
Mouse anti-DDX4	Abcam	ab27591
Rabbit anti-DDX4	Abcam	Ab13840
Rabbit anti-DAZL	Abcam	ab34139
Rabbit anti-Ki67	Abcam	ab15580
Rabbit anti-RFP	Abcam	ab62341
Goat anti-FOXL2	Novus	NB-100-1277
Rabbit anti-SCP3	abcam	ab15093
Rabbit anti-SCP1	Novus	NB-300-228
Rabbit anti-γ H2AX	Novus	NB-100-2280
Rabbit anti-DMC1	SantaCruz Biotechnology	sc-22868
Rabbit anti-PLZF	SantaCruz Biotechnology	sc-22839
Rabbit anti-LAMININ	Abcam	ab11575
Mouse anti-human Mitochondria, clone 133-1	Millipore	MAB1273
*AlexaFluor 488-conjugated donkey anti-rat IgG	Life Technologies	A21208
*FITC-conjugated donkey anti-chicken IgY	Abcam	ab63507
*AlexaFluor 568-conjugated	Life Technologies	A10042

donkey anti-rabbit IgG		
*AlexaFluor 568-conjugated donkey anti-mouse IgG	Life Technologies	A10037
*AlexaFluor 568-conjugated donkey anti-goat IgG	Life Technologies	A10057
*AlexaFluor 647-conjugated donkey anti-rabbit IgG	Life Technologies	A31573
*AlexaFluor 647-conjugated donkey anti-mouse IgG	Life Technologies	A31571
*AlexaFluor 647-conjugated donkey anti-goat IgG	Life Technologies	A21447

MACS

Antibody	**Supplier**	**Product number**
SSEA-1 (CD15) microbeads for human and mouse	Miltenyi Biotec	130-094-530
CD31 microbeads for mouse	Miltenyi Biotec	130-097-418

Primers

QPCR: human

Gene	**Forward primer**	**Reverse primer**
RPLP0	GAAACTCTGCATTCTCGCTTCC	ACTCGTTTGTACCCGTTGATGA
PPIA	TTGATCATTTGGTGTGTTGGGC	AAGACTGAGATGCACAAGTGGT
ERCC56.4	CCAACCCCACATTGTAACTTCG	GTCTTTACTTACGCGCTCCTCT
ERCC451.5	CAGGCAAGAGTTCAATCGCTTAG	TAGCCTTCAGTGACTGTGAGATG
ERCC1806	GATCCCGGAAGATACGCTCTAAG	CGCAGGTTGATGCTTCCAATAAA
TFAP2C	ATTAAGAGGATGCTGGGCTCTG	CACTGTACTGCACACTCACCT
PRDM1	AAACCAAAGCATCACGTTGACA	GGATGGATGGTGAGAGAAGCAA
SOX17	TTCGTGTGCAAGCCTGAGAT	TAATATACCGCGGAGCTGGC
NANOS3	TGGCAAGGGAAGAGCTGAAATC	TTATTGAGGGCTGACTGGATGC
DDX4	ACTGATACAAATGGTGTTAACTGGGA	AAACATGTCTAAGCCCCCTAAAGAA
DAZL	TTTTTGTCTTTGTTGGAGTGAAGCA	ACAGTATCAGCAATAGGCAGAAGCA
DPPA3	AAGCCCAAAGTCAGTGAGATGA	GCTATAGCCCAACTACCTAATGC
POU5F1	CTGTCTCCGTCACCACTCTG	AAACCCTGGCACAAACTCCA
NANOG	AGAGGTCTCGTATTTGCTGCAT	AAACACTCGGTGAAATCAGGGT
SOX2	TGAATCAGTCTGCCGAGAATCC	TCTCAAACTGTGCATAATGGAGT
TCL1B	CAAATCCCCTTCATACCCACCA	TTCTAACCCAAGCACAGATCCC
TFCP2L1	AGCTCAAAGTTGTCCTACTGCC	TTCTAACCCAAGCACAGATCCC
KLF2	ACTAGAGGATCGAGGCTTGTGA	TGCCCACCTGTCTCTCTATGTA
KLF4	AGCCTAAATGATGGTGCTTGGT	CCTTGTCAAAGTATGCAGCAGT
PRDM14	TATCATACTGTGCACTTGGCAGAA	AGCAACTGGGACTACAGGTTTGT
T	AGCCAAAGACAATCAGCAGAAA	CACAAAAGGAGGGGCTTCACTA
MIXL1	TGCTTTCAAAACACTCGAGGAC	GAGTGATCGAAGTAACAGGTGC
EOMES	AAGGGGAGAGTTTCATCATCCC	GGCGCAAGAAGAGGATGAAATAG
SYCP3	CTTCCATGAAACAGCAGCAGCA	GGTTCAAGTTCTTTCTTCAAAGAGTCA
STRA8	GCCAGCTGCAACCCAGAAAACC	ACGGGAAAGGATGCTTCCTGCTT
REC8	CCCTCTCCTCGCCTCTTGACCA	GAATCTGGGCCCCGGCTGGAT
DNMT1	TTCTGGCACCAGGAATCCCCAA	TGTCAGCCAAGGCCACAAACAC
UHRF1	GAGTCCCCTTGAGGCCATTTCT	AAAGAGGAAACATCTCGGGCCT
DNMT3A	TGGGATTCATCCAGACTCATGC	AAAGTGAGAAACTGGGCCTGAA

DNMT3B	TAACTGGAGCCACGACGTAAC	GCATCCGTCATCTTTCAGCCTA
DNMT3L	AGCCATAAGGAGCAGGCACT	GGGGAGAAAGCAGTTCTTCACCA
EHMT1	CGCACTGGCATCACCTTCTGAG	AGGAGAACAGAGCAACCCCGTG

QPCR: mouse

Gene	Forward primer	Reverse primer
Rplp0	CAAAGCTGAAGCAAAGGAAGAG	AATTAAGCAGGCTGACTTGGTTG
Ppia	TTACCCATCAAACCATTCCTTCTG	AACCCAAAGAACTTCAGTGAGAGC
Bmp2	CACCCTCCCTGCAGCAAGAACA	AGATCAGCCCCCTGGAAGGGAT
Bmp4	AACCATGCCATTGTGCAGACCC	GTTCAGTGGGGACACAACAGGC
Bmp5	TTTCTCCAGCCCACGCTGACTT	ATTGCTACAAGCCTCACCACGG
Bmp7	AGCTTCTACTCTGCCCATTCGATGT	AAAAGCCCCAGATCTGCAAACACA
Rspo1	ACAAAGGGCAACAGCAGCCAC	ATCTGGAGACCGGTCACTGTGC
Wnt4	ATCTGGAGACCGGTCACTGTGC	CCAGCTGTGAGACTTGATTGTCGG

Cloning of the homology arms

Gene	Forward primer	Reverse primer
DDX4_C-terminal	CGAATTGGGCGTACTGGTCGTTGTGGGAATA	TGTCTCCCCAGTATTTACACCTCACTGGGTC

Genotyping

Gene	Forward primer	Reverse primer
DDX4_C_5'	GGAAAGTGCCCAGTTCTTGTTG	CCATTGAGTACAGAGCTCCCATC
DDX4_C_3'	AGGCTGAACACCGCCTTTAT	ACATGTCACTTTGGATTGTTTAGGT
TFAP2C_C_5'	CCCTCCACAGTCCCGATGGCTCAACAAAGAA	CTAAGTGTGTGGGTCTGTTGGCCCAGGGAAT
TFAP2C_C_3'	GCTCGCCCCAGTCTTGGAGACGAACATACA	GGCTTGACTGTGGCTTGGTTAATGCCCTGA
pCR2.1 vector	CCGGATGAATGTCAGCTACTGGGCTATCTG	TTCGGGGCGAAAACTCTCAAGGATCTTACC

Oligonucleotides

The oligonucleotides used for the SC3-seq are listed below.

Oligo	Grade	Sequence
V1(dT)24	HPLC	ATATGGATCCGGCGCGCCGTCGACTTTTTTTTTTTTTTTTTTTTTT
V3(dT)24	HPLC	ATATCTCGAGGGCGCGCCGGATCCTTTTTTTTTTTTTTTTTTTTTT
N-V3(dT)24	HPLC	(NH2)-ATATCTCGAGGGCGCGCCGGATCCTTTTTTTTTTTTTTTTTTTTTT
tRd2SPV1(dT)20	HPLC or Column	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCATATGGATCCGGCGCGCCGTCGACTTTTTTTTTTTTTTTTTT
tRd1SPTs	HPLC	TCTTTCCCTACACGACGCTCTTCCGATC~T
tRd1SPTas	HPLC	GATCGGAAGAGCGTCGTGTAGGGAAAGA
S502t	HPLC	AATGATACGGCGACCACCGAGATCTACACCTCTCTATACACTCTTTCCCTACACGACGCTCT
S503t	HPLC	AATGATACGGCGACCACCGAGATCTACACTATCCTCTACACTCTTTCCCTACACGACGCTCT
S505t	HPLC	AATGATACGGCGACCACCGAGATCTACACGTAAGGAGACACTCTTTCCCTACACGACGCTCT

S506t	HPLC	AATGATACGGCGACCACCGAGATCTACACACTGCATAACACTCTT TCCCTACACGACGCTCT
S507t	HPLC	AATGATACGGCGACCACCGAGATCTACACAAGGAGTAACACTCT TTCCCTACACGACGCTCT
S508t	HPLC	AATGATACGGCGACCACCGAGATCTACACCTAAGCCTACACTCTT TCCCTACACGACGCTCT
S510t	HPLC	AATGATACGGCGACCACCGAGATCTACACCGTCTAATACACTCTT TCCCTACACGACGCTCT
S511t	HPLC	AATGATACGGCGACCACCGAGATCTACACTCTCTCCGACACTCTT TCCCTACACGACGCTCT
S513t	HPLC	AATGATACGGCGACCACCGAGATCTACACTCGACTAGACACTCTT TCCCTACACGACGCTCT
S515t	HPLC	AATGATACGGCGACCACCGAGATCTACACTTCTAGCTACACTCTT TCCCTACACGACGCTCT
S516t	HPLC	AATGATACGGCGACCACCGAGATCTACACCCTAGAGTACACTCTT TCCCTACACGACGCTCT
S517t	HPLC	AATGATACGGCGACCACCGAGATCTACACGCGTAAGAACACTCT TTCCCTACACGACGCTCT
S518t	HPLC	AATGATACGGCGACCACCGAGATCTACACCTATTAAGACACTCTT TCCCTACACGACGCTCT
S520t	HPLC	AATGATACGGCGACCACCGAGATCTACACAAGGCTATACACTCTT TCCCTACACGACGCTCT
S521t	HPLC	AATGATACGGCGACCACCGAGATCTACACGAGCCTTAACACTCTT TCCCTACACGACGCTCT
S522t	HPLC	AATGATACGGCGACCACCGAGATCTACACTTATGCGAACACTCTT TCCCTACACGACGCTCT
N701t	HPLC	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTGACTGGAGTT CAGACGTGT
N702t	HPLC	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTGACTGGAGTT CAGACGTGT
N703t	HPLC	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTGACTGGAGTT CAGACGTGT
N704t	HPLC	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTGACTGGAGT TCAGACGTGT
N705t	HPLC	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTGACTGGAGTT CAGACGTGT
N706t	HPLC	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTGACTGGAGTT CAGACGTGT
N707t	HPLC	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTGACTGGAGT TCAGACGTGT
N710t	HPLC	CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTGACTGGAGTT CAGACGTGT
N711t	HPLC	CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTGACTGGAGTT CAGACGTGT
N712t	HPLC	CAAGCAGAAGACGGCATACGAGATTCCTCTACGTGACTGGAGTT CAGACGTGT
N714t	HPLC	CAAGCAGAAGACGGCATACGAGATTCATGAGCGTGACTGGAGTT CAGACGTGT
N715t	HPLC	CAAGCAGAAGACGGCATACGAGATCCTGAGATGTGACTGGAGTT CAGACGTGT
N716t	HPLC	CAAGCAGAAGACGGCATACGAGATTAGCGAGTGTGACTGGAGTT CAGACGTGT
N718t	HPLC	CAAGCAGAAGACGGCATACGAGATGTAGCTCCGTGACTGGAGTT CAGACGTGT
N719t	HPLC	CAAGCAGAAGACGGCATACGAGATTACTACGCGTGACTGGAGTT CAGACGTGT
N720t	HPLC	CAAGCAGAAGACGGCATACGAGATAGGCTCCGGTGACTGGAGTT CAGACGTGT

N721t	HPLC	CAAGCAGAAGACGGCATACGAGATGCAGCGTAGTGACTGGAGTT CAGACGTGT
N722t	HPLC	CAAGCAGAAGACGGCATACGAGATCTGCGCATGTGACTGGAGTT CAGACGTGT
N723t	HPLC	CAAGCAGAAGACGGCATACGAGATGAGCGCTAGTGACTGGAGTT CAGACGTGT
N724t	HPLC	CAAGCAGAAGACGGCATACGAGATCGCTCAGTGTGACTGGAGTT CAGACGTGT
N726t	HPLC	CAAGCAGAAGACGGCATACGAGATGTCTTAGGGTGACTGGAGTT CAGACGTGT
N727t	HPLC	CAAGCAGAAGACGGCATACGAGATACTGATCGGTGACTGGAGTT CAGACGTGT
N728t	HPLC	CAAGCAGAAGACGGCATACGAGATTAGCTGCAGTGACTGGAGTT CAGACGTGT
N729t	HPLC	CAAGCAGAAGACGGCATACGAGATGACGTCGAGTGACTGGAGT TCAGACGTGT

Supplementary Figure Legends

Fig. S1. Generation and culture of xrOvaries.

(A) Scheme for the generation and culture of xrOvaries between d6 hPGCLCs (~5,000 cells/well) and mouse ovarian somatic cells at embryonic day (E) 12.5 (~50,000 cells/well or when using frozen stocks, 75,000 cells/well).

(B) Bright field (BF) and fluorescence (green: *Blimp1-mVenus*; cyan: *Stella-ECFP*) images of a rOvary. Bar, 500 μm.

(C) BF and fluorescence [green: *TFAP2C-EGFP* (AG); red: *BLIMP1-tdTomato* (BT)] images of xrOvaries with 585B1 BTAG (XY) hPGCLC-derived cells at ag7. Bar, 500 μm.

(D) Compensatory removal of cells in xrOvaries exhibiting autofluorescence (APC-A channel) by FACS by forward scatter (FSC-A) and APC-A (middle). The cells negative for autofluorescence of APC-A were analyzed for BTAG (bottom) (Fig. 1A, Fig. S1C) or AGVT expression (Fig. 1C, Fig. S5A).

(E) Numbers of the BT^+AG^+ cells (derived from 585B1 BTAG hiPSCs) per xrOvary from ag7 to ag77. The averages are shown as bars.

Fig. S2. The properties of hPGCLC (585B1 BTAG)-derived cells in xrOvaries.

(A) Immunofluorescence (IF) analysis of the expression of indicated key proteins (human mitochondrial antigen, FOXL2, LAMININ, TFAP2C, SOX17, POU5F1, SOX2, Ki-67, cleaved CASPASE3, DAZL and DDX4: magenta or cyan) in AG^+ cells (yellow) in xrOvaries at ag77, with DAPI (white) and merges. Bars, 10 μm.

(B) IF analysis of the expression of a human mitochondrial antigen (cyan) in a mouse embryonic ovary at E12.5 (top) and in an xrOvary at ag15 (bottom), which were co-stained with FOXL2 (magenta) and DAPI (white) (top) and with GFP (AG) (yellow), DDX4 (magenta) and DAPI (white) (bottom), respectively. Note that there was no immunoreactivity for the human mitochondrial antigen in mouse ovarian somatic cells (top) and in developing mouse oocytes remaining to be sorted by MACS (a $DDX4^+$ cell) (bottom). Bars, 20 μm.

(C) Hematoxylin-eosin (HE) staining of the sections of xrOvaries at ag7, ag35, and ag77. The putative BT^+AG^+ cells with a distinct morphology (large round cells with clear cytoplasm) are indicated by arrowheads. Bar, 20 μm.

(D) Nuclear architecture shown by DAPI (white) staining of the AG^+ cells (yellow) and surrounding mouse ovarian somatic cells in xrOvaries at ag7, ag35, and ag77. Merged images are shown on the right. Bars, 10 μm.

(E) IF analysis of the expression of DDX4 (magenta) in AG^+ cells (yellow) in xrOvaries at ag7, ag35, and ag77, with FOXL2 (cyan) and DAPI (white) staining. The boxed area in the left panel (merged image) is magnified in the right panels. The images at ag77 is also shown in Fig. 1B. Bar, 20 μm.

Fig. S3. Ultrastructure of hPGCLC (585B1 BTAG)-derived cells in xrOvaries.

(A) A portion of an xrOvary at ag77 examined by electron microscopy. Note that large, round cells with low electron density (most likely hPGCLC-derived cells, arrows) are surrounded by small, squamous cells with high electron density (most likely mouse granulosa cells, arrowheads). The magnified areas in (B) and (C) are boxed. Bar, 10 μm.
(B) A magnified view of the boxed area in (A), showing apparent direct contacts (arrowheads) between a hPGCLC-derived cell and a mouse granulosa cell. Bar, 500 nm.
(C) A magnified view of the boxed cell (most likely a hPGCLC-derived cell) in (A). The cell bears a clear cytoplasm with sparsely distributed mitochondria with villiform cristae and an ovoid nucleus with loosely packed chromatin and a prominent granular nucleolus. These properties are highly similar to those of human oogonia/gonocytes (*8, 9*). Bar, 2 μm.

Fig. S4. Generation of 1390G3 AGVT (XX) hiPSCs.
(A) Single-strand annealing (SSA) activities of TALEN pairs targeting the human *DDX4* (human VASA homolog) locus compared with those of a control TALEN pair targeting the human *HPRT1* locus as assessed by a Dual-Glo luciferase assay system.
(B) Schematic illustration of the human *DDX4* (human VASA homolog) locus, and the constructs for knocking in *2A-tdTomato* into the locus. Black boxes indicate the exons.
(C) Screening by PCR of the homologous recombinants for *DDX4* (human VASA homolog)-*2A-tdTomato* (VT) (left) and *TFAP2C-2A-EGFP* (AG) (*7*) (middle), and of random integration of the targeting vector (right). Targeted (black arrowheads): bands for the targeted allele; targeted and recombined (black arrows): bands for the targeted allele with the selection cassettes (*loxP-pgk-neo-polyA-loxP* or *loxP-pgk-puro-polyA-loxP*) excised by the Cre recombinase; non-targeted (white arrowheads): bands for the non-targeted, wild-type alleles.
(D) Sequences of the alleles that did not show homologous recombination in several targeted clones are shown. Stop codons are marked by red squares. Red bars indicate deleted sequences. The 1390G3_5-2-6 clone bearing the heterozygous AG and the VT alleles without mutations is selected for experiments.
(E) Southern blot analysis of the homologous recombination of the AGVT alleles in 1390G3-derived clones. The sites for the restriction enzyme digestions and the positions of the probes are shown. Arrows indicate the expected targeted bands.
(F) A representative result for the G-band analysis of the 1390G3 AGVT_5-2-6 clone bearing a normal karyotype (46, XX).

Fig. S5. xrOvaries with 1390G3 AGVT.
(A) Bright field (BF) images (top) and FACS by AGVT of d6 hPGCLCs and xrOvaries at ag56, 77, 98 and 120 with 1390G3 AGVT (XX) hiPSC-derived cells. The percentages of the cells in the indicated gates are shown. The image and FACS at ag120 are also shown in Fig. 1C. Bars, 500 μm.
(B) Numbers of the hPGCLC-derived cells (sum of the AG^+VT^-, AG^+VT^+, AG^-VT^+

cells) per xrOvary at ag56, 77, 98 and 120. Where applicable, the averages are shown as bars.
(C, D) Expression of key proteins (magenta) in AG^+ cells (yellow) in xrOvaries at ag77 (C) and ag120 (D), with DAPI (white) and merges. The bottom panel in (D) highlights the expression of DDX4 (magenta) in AG^-VT^+ (cyan) cells. Note also the presence of an AG^+VT^- cell. Bars, 20 μm.

Fig. S6. Expression of key genes during hPGCLC induction and hPGCLC development in xrOvaries.
Dynamics of the expression of key genes during hPGCLC induction and development [hiPSCs, iMeLCs, the 585B1 BTAG (XY) d6 hPGCLCs, the hPGCLC-derived cells (ag7, 21, 35, 49, 63 and 77 BT^+AG^+ cells), and the 1390G3 AGVT (XX) hPGCLC-derived cells (ag77 AG^+VT^-, AG^+VT^+, ag120 AG^+VT^-, AG^+VT^+, $AG^{+/-}VT^+$, AG^-VT^+ cells)] as measured by qPCR. For each gene examined, the Δ*Ct* from the average *Ct* values of the two independent housekeeping genes *RPLP0* and *PPIA* (set as 0) were calculated and plotted for 2 independent experiments. Mean values are connected by a line. *Not detected.

Fig. S7. Expression of key meiosis markers in xrOvaries.
(A) IF analysis of the expression of the indicated meiosis markers (green) in an adult mouse testis, with DAPI (white) staining, as positive controls for the antibody activities. Bars, 20 μm.
(B, C) IF analysis of the expression of the indicated meiosis markers (cyan) in the 1390G3 AGVT (XX) hPGCLC-derived AG^+ cells (yellow, B) or $DDX4^+$ cells (magenta, C) in xrOvaries at ag120, with DAPI (white) staining. Note that the $DDX4^+$ cells become $SCP3^+$, but $DMC1^-$, $\gamma H2X^-$ and $SCP1^-$. Bars, 20 μm.
(D) IF analysis of the expression of DMC1 or PLZF (magenta) in the 585B1 BTAG (XY) hPGCLC-derived AG^+ cells (yellow) in xrOvaries at ag77, with DAPI (white) staining. A merged image is shown in the left panels and the boxed area is magnified in the right panels. Bars, 20 μm.

Fig. S8. Transcriptome analysis during hPGCLC induction and hPGCLC development in xrOvaries.
(A) Statistics of RNA-seq analyses for the key cell types during hPGCLC induction and hPGCLC-derived cell differentiation in xrOvaries. The color coding is as indicated.
(B) Unsupervised hierarchical clustering (UHC) of the transcriptomes of hiPSCs, iMeLCs, d6 hPGCLCs and hPGCLC-derived cells in xrOvaries. The biological replicate numbers are shown after the cell types in (A and B).
(C) Principal component analysis (PCA) of the transcriptomes of hiPSCs, iMeLCs, d6 hPGCLCs and hPGCLC-derived cells in xrOvaries.
(D) Scatter plot of the normalized loading scores of PCA in (C). Dots colored according

to the clusters in (Fig. 2A) (453 genes, Table S2) indicate genes that contributed highly to the PC1/2 axes: i.e., those with a more than 2 s.d. radius of PC1/2. Key genes are annotated.

Fig. S9. Differentially expressed genes in hPGCLC-derived cells in xrOvaries.
(left) Scatter-plot representations of differentially expressed genes (DEGs) between the indicated samples. The genes up/down-regulated (> 2-fold in the top to fourth panels, > 4-fold in the bottom panel) in each cell-fate transition are plotted in red and blue, respectively. (right) The GO analyses of the DEGs. Representative genes in each GO category are indicated.

Fig. S10. Comparison of gene-expression properties between 585B1 BTAG (XY) hPGCLC-derived cells and human germ cells in vivo.
(A) Scatter-plot comparisons of the transcriptome data between d4 hPGCLCs of this study [$\log_2$(RPM+1)] and PGCLCs by Tang et al. [$\log_2$FPKM] (*12*). The regression line between the two samples was defined as $y = 0.992x + 1.065$ by the least square method, and the RPKM values of genes expressed in human germ cells in Tang et al. were adjusted according to this regression analysis.
(B) Comparison of the expression of the indicated genes between ag77 BT^+AG^+ cells and human germ cells at the indicated week of development (*12*). The adjusted relative values based on the calculation as in (A) were used for generating a gene expression-level heatmap. The color coding is as indicated.

Fig. S11. The gene expression property of the 1390G3 AGVT (XX) ag120 $AG^{+/-}VT^+$ cells.
(A) The expression of the genes that distinguish the characteristics of human fetal germ cells (*14*) in the indicated cell types. The color coding is as indicated. See also **Fig. 2C**.
(B) The expression of the selected key genes in the indicated cell types. The color coding is as indicated in (A).

Fig. S12. Genome-wide DNA methylation levels in hPGCLC-derived cells in xrOvaries and other relevant cell types.
(A, B) Total read counts, the coverage of the CpG sequences (A) and violin plots of the 5mC-level distributions (B, averages: red bars) in the indicated cell types (see also Table S3).
(C, D) Comparisons of the 5mC levels (genome-wide 2-kb windows) by contour representations of scatter plots, combined with histogram representations (top and right of scatter plots), between the indicated biological replicates (C, top, left), between iMeLCs and the indicated cell types (C, bottom, left), between male (585B1 BTAG) and female (1390G3 AGVT) hiPSCs on the autosomes and the X chromosomes (C, top and bottom, the fourth column from the rightmost), between female (1390G3 AGVT) hiPSCs

and the indicated cell types on the autosomes and the X chromosomes (C, top and bottom, right) and between the 585B1 BTAG (XY) ag77 BT^+AG^+ cells and the indicated cell types (D).

Fig. S13. DNA methylation levels in key elements in hPGCLC-derived cells in xrOvaries and other relevant cell types.
(A) Heatmaps showing the 5mC levels of the indicated genomic elements (200 randomly chosen regions) on the autosomes and the X chromosomes in the indicated cells. HCP, ICP, LCP: high, intermediate, and low CpG promoters, respectively (*31*). CGI: CpG islands. The color coding is as indicated.
(B) Violin plots of the 5mC levels (averages: red bars) of the promoters of the genes in clusters 1 to 5 in Fig. 2A in the indicated cells. The genes are classified with those bearing HCP, ICP and LCP, respectively.
(C, D) Violin plots of the 5mC levels (% 5mC; averages: red bars) in regions bearing hyper-methylation in hiPSCs compared to hESCs (C) (Table S5) or the 5mC (%) and expression levels [$\log_2$(RPM+1)] of indicated repeat elements (D left and right; averages: red bars) in the indicated cells. The asterisk in (C) indicates a p value < 0.01 by t-test.
(E) (left) Venn diagram showing the overlap of the DNA-demethylation "escapees" (Table S6) between ag77 BT^+AG^+/ag120 AG^+VT^+ cells and germ cells at week 7 of development (*12*). (right) Enrichment of the escapees in the indicated genomic elements.

Fig. S14. X chromosome activity in the 1390G3 AGVT (XX) hiPSCs and hPGCLC-derived cells.
(A-D) RNA fluorescence in situ hybridization (RNA-FISH) analyses for the expression of the X-linked genes [(A) *TBL1X* (red), (B) *POLA1* (red), (C) *ATRX* (red), (D) *XACT* (red) and (A-D) *XIST* (green)] in 1390G3 AGVT (XX) hiPSCs, d6 hPGCLCs, and ag120 AG^+VT^-/AG^+VT^+/AG^-VT^+ cells in xrOvaries. Representative RNA-FISH images of the cells for bi-allelic or mono-allelic expression of the indicated genes (left), their percentages (right), and the hybridization patterns and the approximate positions of the X-linked genes on the X chromosome (bottom) are shown.
(E) 5mC levels of the promoters of the indicated genes in the indicated cell types.

Fig. S15. DNA methylation dynamics on the autosomes and the X chromosomes during hPGCLC specification and development in xrOvaries.
Violin plots showing the 5mC-level dynamics in the promoters (HCP, ICP, LCP) (*31*), non-promoter CGIs, and genome-wide unique regions (genome-wide 2 Kb bins) on the autosomes (left) and the X chromosomes (right) during 585B1 BTAG (XY) and 1390G3 AGVT (XX) hPGCLC specification and development in xrOvaries. The average 5mC level in each cell type is indicated by a red bar.

Fig. S16. Expression of key signaling molecules in hPGCLC-derived cells and

mouse ovarian somatic cells in xrOvaries.
(A) Expression of the genes for key signaling molecules (*Bmp2/4/5/7* and *Rspo1/Wnt4*) in mouse ovarian somatic cells at E12.5 and cultured as xrOvaries (ag56, 77, 130) as measured by qPCR. The mouse somatic cells in xrOvaries were collected by FACS as $AG^{-}VT^{-}$ cells (ag56, 77) or $BT^{-}AG^{-}$ (ag130) cells. Note that *Bmp2/4* and *Rspo1/Wnt4* were expressed continuously in mouse ovarian somatic cells in xrOvaries. For each gene examined, the ΔCt from the average Ct values of the two independent housekeeping genes *Rplp0* and *Ppia* (set as 0) was calculated and plotted for 3 independent experiments. Mean values are connected by a line. *Not detected.
(B) Expression of the genes for key signaling receptors and their immediate targets during 585B1 BTAG (XY) and 1390G3 AGVT (XX) hPGCLC specification and development in xrOvaries measured by RNA-seq. The expression values (the mean values where applicable) are connected by a line.

Fig. S17. A model for the differentiation of oogonia from hPGCLCs in xrOvaries.
(top) Differentiation of primary oocytes from mPGCLCs and formation of secondary follicles in syngeneic rOvaries.
(bottom) Differentiation of oogonia from hPGCLCs and their epigenetic reprogramming in xrOvaries.

Supplementary Tables

Table S1. RNA-seq data generated in this study.

Table S2. List of differentially expressed genes (DEGs) identified in Fig. 2B/2C and their gene ontology enrichments.

Table S3. Mapping statistics and accession numbers for the WGBS data of this study and the relevant published studies.

Table S4. Lists and chromosomal coordinates of the imprint loci used in this study.

Table S5. Lists and chromosomal coordinates of loci bearing hyper-methylation in hiPSCs compared to hESCs.

Table S6. Lists and chromosomal coordinates of DNA-demethylation "escapees".

References

23. T. Sakuma *et al.*, Efficient TALEN construction and evaluation methods for human cell and animal applications. *Genes Cells* **18**, 315-326 (2013).
24. J. Chaumeil, I. Okamoto, E. Heard, X-chromosome inactivation in mouse embryonic stem cells: analysis of histone modifications and transcriptional activity using immunofluorescence and FISH. *Methods Enzymol* **376**, 405-419 (2004).
25. J. Chaumeil, S. Augui, J. C. Chow, E. Heard, Combined immunofluorescence, RNA fluorescent in situ hybridization, and DNA fluorescent in situ hybridization to study chromatin changes, transcriptional activity, nuclear organization, and X-chromosome inactivation. *Methods Mol Biol* **463**, 297-308 (2008).
26. T. Nakamura *et al.*, SC3-seq: a method for highly parallel and quantitative measurement of single-cell gene expression. *Nucleic Acids Res* **43**, e60 (2015).
27. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
28. W. Huang da, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57 (2009).
29. F. Miura, Y. Enomoto, R. Dairiki, T. Ito, Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res* **40**, e136 (2012).
30. K. Shirane *et al.*, Global Landscape and Regulatory Principles of DNA Methylation Reprogramming for Germ Cell Specification by Mouse Pluripotent Stem Cells. *Dev Cell* **39**, 87-103 (2016).
31. R. S. Illingworth *et al.*, Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet* **6**, e1001134 (2010).
32. F. Court *et al.*, 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 (2014).
33. R. Lister *et al.*, Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315-322 (2009).
34. R. Lister *et al.*, Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68-73 (2011).