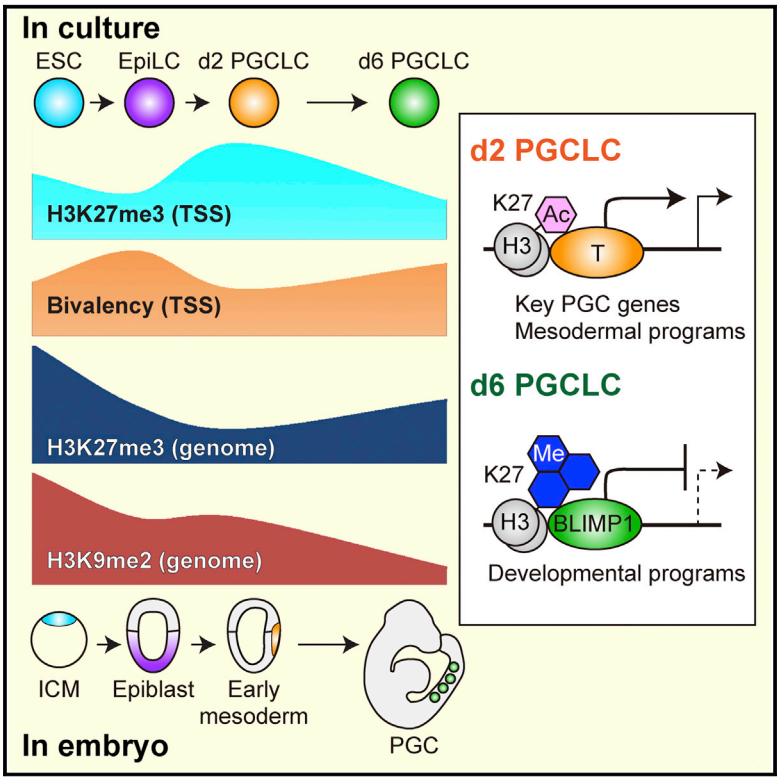
Article

Cell Stem Cell

Quantitative Dynamics of Chromatin Remodeling during Germ Cell Specification from Mouse Embryonic Stem Cells

Graphical Abstract



Highlights

- Chromatin dynamics were quantitatively assessed during PCG specification
- EpiLCs represent a primed state with abundant bivalency with low H3K27me3
- PGCLCs deplete H3K9me2 throughout the genome including from lamina-associated domains
- BLIMP1 serves as a potential nucleator for H3K27me3 accumulation and spread

Authors

Kazuki Kurimoto, Yukihiro Yabuta, ..., Katsuhiko Shirahige, Mitinori Saitou

Correspondence

kurimoto@anat2.med.kyoto-u.ac.jp (K.K.),
saitou@anat2.med.kyoto-u.ac.jp (M.S.)

In Brief

Kurimoto et al. perform careful analyses of chromatin remodeling during mouse germ cell specification from embryonic stem cells. Widespread epigenetic reprogramming events included re-organization of H3K27me3 and bivalent signatures as well as progressive deletion of H3K9me2 throughout the genome, creating a unique foundation for the epigenome of the next generation.

Accession Numbers

GSE60204
GSE60018

Kurimoto et al., 2015, Cell Stem Cell *16*, 517–532
May 7, 2015
http://dx.doi.org/10.1016/j.stem.2015.03.002

CellPress

Cell Stem Cell
Article

Quantitative Dynamics of Chromatin Remodeling during Germ Cell Specification from Mouse Embryonic Stem Cells

Kazuki Kurimoto,[1,2,11,*] Yukihiro Yabuta,[1,2,11] Katsuhiko Hayashi,[1,3,4] Hiroshi Ohta,[1,2] Hiroshi Kiyonari,[5] Tadahiro Mitani,[1] Yoshinobu Moritoki,[1,6] Kenjiro Kohri,[6] Hiroshi Kimura,[7] Takuya Yamamoto,[8,10] Yuki Katou,[9] Katsuhiko Shirahige,[9] and Mitinori Saitou[1,2,8,10,*]

[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]Department of Developmental Stem Cell Biology, Faculty of Medical Sciences, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582, Japan
[4]JST, PRESTO, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582, Japan
[5]Laboratories of Animal Resource Development and Genetic Engineering, RIKEN Center for Life Science Technologies, 2-2-3 Minatojima Minami, Chuou-ku, Kobe 650-0047, Japan
[6]Department of Nephro-Urology, Graduate School of Medical Sciences, Nagoya City University, Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya 467-8601, Japan
[7]Department of Biological Sciences, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama 226-8501, Japan
[8]Center for iPS Cell Research and Application, Kyoto University, 53 Kawahara-cho, Shogoin Yoshida, Sakyo-ku, Kyoto 606-8507, Japan
[9]Laboratory of Genome Structure and Function, Center for Epigenetic Disease, Institute of Molecular and Cellular Biosciences, The University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-0032, Japan
[10]Institute for Integrated Cell-Material Sciences, Kyoto University, Yoshida-Ushinomiya-cho, Sakyo-ku, Kyoto 606-8501, Japan
[11]Co-first author
*Correspondence: kurimoto@anat2.med.kyoto-u.ac.jp (K.K.), saitou@anat2.med.kyoto-u.ac.jp (M.S.)
http://dx.doi.org/10.1016/j.stem.2015.03.002

SUMMARY

Germ cell specification is accompanied by epigenetic remodeling, the scale and specificity of which are unclear. Here, we quantitatively delineate chromatin dynamics during induction of mouse embryonic stem cells (ESCs) to epiblast-like cells (EpiLCs) and from there into primordial germ cell-like cells (PGCLCs), revealing large-scale reorganization of chromatin signatures including H3K27me3 and H3K9me2 patterns. EpiLCs contain abundant bivalent gene promoters characterized by low H3K27me3, indicating a state primed for differentiation. PGCLCs initially lose H3K4me3 from many bivalent genes but subsequently regain this mark with concomitant upregulation of H3K27me3, particularly at developmental regulatory genes. PGCLCs progressively lose H3K9me2, including at lamina-associated perinuclear heterochromatin, resulting in changes in nuclear architecture. T recruits H3K27ac to activate BLIMP1 and early mesodermal programs during PGCLC specification, which is followed by BLIMP1-mediated repression of a broad range of targets, possibly through recruitment and spreading of H3K27me3. These findings provide a foundation for reconstructing regulatory networks of the germline epigenome.

INTRODUCTION

Epigenetic reprogramming refers to genome-wide chromatin-state remodeling leading to fundamental changes of cellular phenotypes, and it has critical implications for broad areas of medical science. The key examples include physiologic reprogramming during germ cell development (Kafri et al., 1992; Monk et al., 1987) and artificial reprogramming in induced pluripotency (Takahashi and Yamanaka, 2006) or somatic cell nuclear transfer into oocytes (Gurdon, 1962). While the mechanism for the latter examples is inherently difficult to analyze due to the low efficiency and high stochasticity, the mechanism for the former should be amenable to precise genetic/biochemical analysis as a deterministic activity imposed by a program for germ cell specification (Saitou et al., 2012). Furthermore, and importantly, impaired epigenetic reprogramming in germ cells leads to a variety of critical defects, including loss of germ cells, impaired development of embryos, and impaired development/physiology of offspring. Thus, it is essential in reproductive biology/medicine to understand epigenetic regulation and its consequences during germ cell development.

The germ cell lineage in mice arises as primordial germ cells (PGCs) at around embryonic day (E) 7.25 (Ginsburg et al., 1990; Saitou et al., 2002). PGCs repress a somatic mesodermal program, regain a transcriptional network for pluripotency, and, during their migration, initiate epigenetic reprogramming that includes reduction of histone H3 lysine 9 di-methylation (H3K9me2), elevation of H3K27 tri-methylation (H3K27me3), and demethylation of 5-methylcytosine (5mC) (Seisenberger

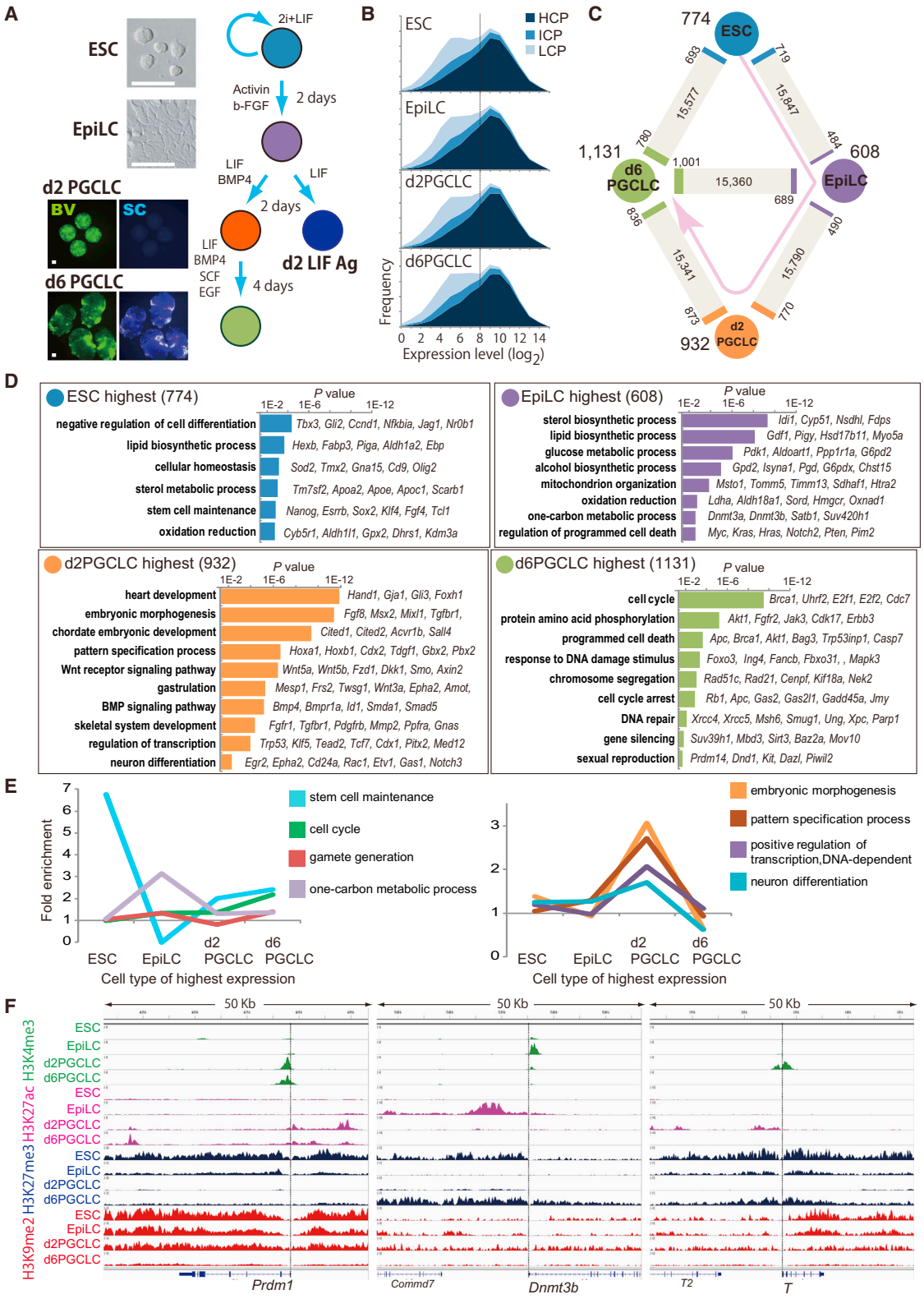


Figure 1. Transcriptional Foundation for In Vitro PGC Specification

(A) Scheme for in vitro PGC specification. BV, *Blimp1*-mVenus; SC, *Stella*-ECFP. Bar, 50 μm.

(B) The $\log_2$ expression level-frequency plots for the expression of genes with high, intermediate, and low CpG promoters (HCP, ICP, and LCP, respectively). Dotted lines indicate a $\log_2$ expression level of 8, which corresponds to the expression of ~20 copies of genes per cell (Kurimoto et al., 2006).

(legend continued on next page)

et al., 2012; Seki et al., 2005). Consequently, by E13.5 in embryonic gonads, PGCs acquire a unique epigenetic state with an extremely low level of genome-wide 5mC (Seisenberger et al., 2012). Thus, in reprogramming PGCs, DNA methylation and transcription are largely uncoupled and a mechanism independent of DNA methylation, most likely, that by histone modifications promotes transcriptional control (Seisenberger et al., 2012). Subsequently, male and female germ cells establish unique epigenomes for the spermatogenic and oogenic pathways, respectively. Although recent advances in sequencing technologies have uncovered consequences of epigenetic reprogramming—particularly the genome-wide distribution of 5mC—during germ cell development (Lee et al., 2014), there is a lack of information as to how PGCs establish their epigenome—particularly the genome-wide histone modification states—at the outset and what the overall mechanism of epigenetic reprogramming by germ cells might be. This has mainly been because PGCs are specified in small numbers (~40) and are intractable to quantitative analysis, and because there is no appropriate in vitro system that recapitulates germ cell development.

Recent studies have shown that embryonic stem cells (ESCs)/induced pluripotent stem cells (iPSCs) are induced into epiblast-like cells (EpiLCs), and EpiLCs are in turn induced into PGC-like cells (PGCLCs) with capacity for both spermatogenesis and oogenesis (Hayashi et al., 2011, 2012). The global gene-expression dynamics during PGCLC induction are highly similar to those for PGC specification and development until around E9.5, and PGCLCs appear to acquire epigenetic properties similar to those of PGCs at around E9.5 (Hayashi et al., 2011), demonstrating that the in vitro PGC specification is a proper reconstitution of the in vivo pathway and provides an appropriate model for analysis of the mechanisms underlying PGC specification and development. Here, using this system, we investigate the chromatin-state reprogramming, and in particular, the histone modification reprogramming, for PGC specification and development.

RESULTS

Global Transcription Regulation

We first performed a systematic evaluation of transcriptomes of four key cell types (ESCs, EpiLCs, day 2 [d2] PGCLCs, and d6 PGCLCs) during in vitro PGC specification (Figure 1A) (Hayashi et al., 2011; Nakaki et al., 2013). We classified the 17,050 genes into those bearing high (10,848), intermediate (2,438), and low (3,764) CpG density promoters (HCP, ICP, and LCP, respectively) (Weber et al., 2007). The HCP genes include housekeeping genes and developmental regulators, whereas the LCP genes include highly tissue-specific genes. All four cell types predominantly expressed the HCP genes (Figure 1B), indicating that the HCP genes regulate key properties of the PGC specification pathway.

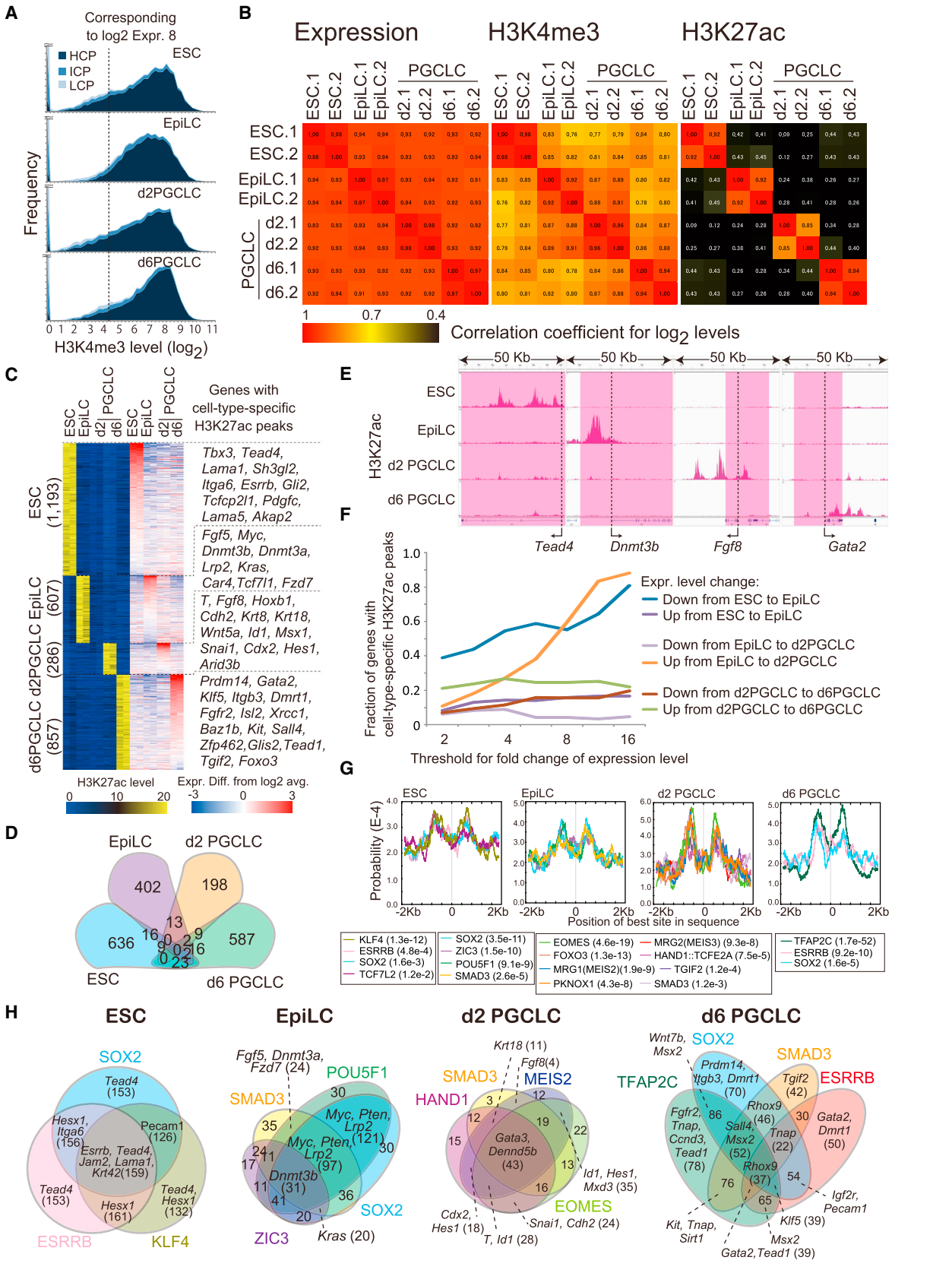
We explored genes that were upregulated or downregulated (>2-fold) between the successive stages (~2.8%–5.9%) and identified those expressed at the highest level in ESCs (774, the ESC genes), EpiLCs (608, the EpiLC genes), d2 PGCLCs (932, the d2 PGCLC genes), or d6 PGCLCs (1,131, the d6 PGCLC genes) (Figure 1C). The Gene Ontology (GO) function category analysis revealed that genes for "stem cell maintenance" are enriched in the ESC genes but are depleted in the EpiLC genes, and then again become prevalent in the d2 and d6 PGCLC genes (Figures 1D and 1E and Figure S1A). In contrast, genes for "one-carbon metabolic process," including *Dnmt3a*, *Dnmt3b*, *Satb1*, and *Suv420h1*, were prevalent among the EpiLC genes (Figures 1D and 1E and Figure S1A). Notably, the EpiLC genes did not show overt enrichment of genes for developmental regulators (Figures 1D and 1E), despite the fact that EpiLCs were induced from ESCs by withdrawal of an inhibitor of the mitogen-activated protein kinase (MAPK) pathway (PD0325901) and stimulation by Activin A and basic fibroblast growth factor (bFGF). Instead, the genes for developmental regulators such as those for "embryonic morphogenesis" and "pattern specification process" were highly enriched in the d2 PGCLC genes (Figures 1D and 1E and Figure S1A).

The ESC and d6 PGCLC genes were distinguished mainly by their unique metabolic, signaling, and cell cycle properties (Figure 1D). For example, the ESC genes were enriched with genes for "lipid biosynthetic process," "sterol metabolic process," and "oxidation reduction," whereas the d6 PGCLC genes were characterized by genes playing roles in the "cell cycle," "protein amino acid phosphorylation," and "response to DNA damage stimulus." The d6 PGCLC genes were also characterized by those for "programmed cell death," including both positive (*Apc* and *Brca1*) and negative (*Akt1* and *Bag3*) regulators of apoptosis, consistent with the fact that migrating PGCs are balanced in transducing apoptotic signals and readily eliminated by apoptosis when mis-located or depleted from survival signals (Runyan et al., 2006). The EpiLC genes were also characterized by unique metabolic properties, with genes for the "sterol biosynthetic process," "lipid biosynthetic process," and "glucose metabolic process" (Figure 1D). These analyses provide a clear delineation of key transcriptional properties of the four cell types during in vitro PGC specification.

Quantification of Chromatin-State Dynamics and BLIMP1 Binding

To explore chromatin-state dynamics during in vitro PGC specification, we developed a protocol for the representative amplification of chromatin-immunoprecipitated (ChIP-ed) DNA from a relatively small number of samples (Experimental Procedures), as the number of PGCLCs obtained per experiment is

(C) Gene expression changes. The numbers of genes upregulated (>2 fold, marked by corresponding colors) and unchanged (gray) in each cell type compared to a neighboring cell type are indicated. The numbers of genes expressed at the highest level in each cell type are indicated in bold next to each cell type.
(D) GO analysis of genes expressed at the highest levels in the four cell types.
(E) Enrichment of genes for the indicated GO terms.
(F) ChIP-seq track transitions for H3K4me3, H3K27ac, H3K27me3, and H3K9me2 in the *Prdm1*, *Dnmt3b*, and *T* loci (50 kb).
See also Figure S1.



(legend on next page)

on the order of $\sim 10^5$. We measured the distribution of H3 (nucleosome), H3K4me3 (promoters), H3K27ac (acetylation) (active enhancers), H3K27me3 (repression by polycomb complex 2 [PRC2]], and H3K9me2 (repression by G9A/GLP). For the detection of BLIMP1-binding sites, we generated a knockin mouse strain in which the EGFP sequence is fused in-frame to the first ATG of *Blimp1*, derived ESCs homozygously bearing the *EGFP-Blimp1* knockin alleles, and induced PGCLCs for the ChIP-sequencing (ChIP-seq) analysis using an anti-GFP antibody (Figures S1B–S1F). See Figures S2 and S3A–S3C and Supplemental Experimental Procedures for the normalization of the ChIP-seq data.

Figure 1F shows ChIP-seq track transitions for the histone modifications around *Prdm1*, *Dnmt3b*, and *T* loci. Consistent with their expression, while *Prdm1* showed a specific H3K4me3 peak around its promoter in d2 and d6 PGCLCs, *Dnmt3b* showed such a peak in EpiLCs and *T* showed it in d2 PGCLCs (Figure 1F). Consistently, the three loci exhibited several distinct H3K27ac peaks in relevant cell types, which would represent their enhancers (Figure 1F). Indeed, one prominent peak around 10 kb upstream from a transcription start site (TSS) of *Prdm1* corresponds to a T-binding site for *Prdm1* upregulation in PGCLCs (Aramaki et al., 2013). On the other hand, while the *Prdm1* and *T* loci were deposited with a "lawn" of high-level H3K27me3 in ESCs and to a lesser extent in EpiLCs, they exhibited no H3K27me3 in d2 PGCLCs: the *T* locus regained H3K27me3 in d6 PGCLCs. As to the *Dnmt3b* locus, the upstream region exhibited high-level H3K27me3 in ESCs and lost it in EpiLCs, and the whole locus regained it in d6 PGCLCs (Figure 1F). Regarding H3K9me2, the three loci exhibited similar modification patterns in ESCs, EpiLCs, and d2 PGCLCs, albeit at different levels, and all three loci exhibited a significant reduction of this modification in d6 PGCLCs (Figure 1F). Thus, the histone modifications of the three key genes provide not only a precise reflection/prediction of their active/repressed states, but also novel information as to their regulatory mechanism during in vitro PGC specification.

Regulation of Promoter and Enhancer Usage

We analyzed the transitions of H3K4me3 and H3K27ac during in vitro PGC specification. In all cell types, a vast majority of the genes with high H3K4me3 were HCP genes (Figure 2A) and the $\log_2$ H3K4me3 levels around the TSSs were positively correlated with the $\log_2$ gene-expression levels (Figure S4A). Consistent with the notion that gene expression involves complex and dynamic enhancer usage (Calo and Wysocka, 2013), the global changes in H3K27ac were more pronounced than those in gene expression and in H3K4me3 (Figure 2B). We identified H3K27ac peaks (<15 kb from the TSSs and gene bodies) specific to each cell type: many of the genes with such peaks show specific and high-level expression in each cell type (Figure 2C). Accordingly, a vast majority of cell-type-specific peaks were associated with different genes in each cell type (Figure 2D). For example, *Tead4*, *Dnmt3b*, *Fgf8*, and *Gata2* exhibited specific H3K27ac peaks and the highest expression in ESCs, EpiLCs, d2 PGCLCs, and d6 PGCLCs, respectively (Figures 2C and 2E).

The cell-type-specific H3K27ac peaks were associated with ~6%–39% of the genes differentially expressed between adjacent stages, and such peaks were, with increasing threshold levels, prevalent in genes downregulated in the ESC-to-EpiLC transition and upregulated in the EpiLC-to-d2 PGCLC transition (Figure 2F), suggesting a critical role of specific enhancer usage for the maintenance of naive pluripotency and the induction of a mesoderm program/specification of germ cell fate.

We explored whether consensus sequences for any transcription factor (TF) binding were enriched in H3K27ac peaks specific to each cell type, and we identified a high prevalence of such sequences for unique, but partially overlapping, TFs for each cell type (Figure 2G): for example, while SOX2 exhibited enrichment in three cell types (ESCs, EpiLCs, and d6 PGCLCs), KLF4, ZIC3, and TFAP2C showed specific enrichment in ESCs, EpiLCs, and d6 PGCLCs, respectively. Reflecting highly specific enhancer usage for the induction of a mesoderm program/specification of germ cell fate in d2 PGCLCs (Figure 2F), the binding sites for TFs involved in mesoderm formation were enriched in d2 PGCLCs (Figure 2G). Importantly, a majority of the H3K27ac peaks bear multiple TF-binding motifs (Figure 2H), suggesting that genes specific to each cell type are regulated by a combination of distinct yet partially overlapping sets of TFs.

We looked into the relationship between the binding sites of OCT4, a key pluripotent factor, in ESCs and EpiLCs (Buecker et al., 2014) and H3K27ac peaks. Notably, OCT4-binding sites specific to ESCs exhibited substantial enrichment of H3K27ac in ESCs and d6 PGCLCs, whereas those present in both ESCs and EpiLCs showed enrichment of H3K27ac in all cell types (Figure S3D). These findings indicate that OCT4 distinctively regulates a specific set of common genes in ESCs and d6 PGCLCs, while it controls many other targets similarly in all the cell types, highlighting a similarity between ESCs and d6 PGCLCs in the regulatory network for pluripotency. Thus, the potential enhancers we have identified provide a foundation for delineating regulatory networks orchestrating PGC specification.

Figure 2. Promoter and Enhancer Usage for In Vitro PGC Specification

(A) The $\log_2$ H3K4me3 level-frequency plots for the HCP, ICP, and LCP genes. Dotted lines correspond to a $\log_2$ expression level of 8.

(B) Heat map representation of the changes in gene expression, H3K4me3, and H3K27ac levels (correlation coefficients). The values for two biological replicates (1, 2) for each cell type are shown.

(C) Heat map representation of cell-type-specific H3K27ac peaks and expression of associated genes.

(D) Venn diagram showing that genes (numbers indicated) with cell-type-specific H3K27ac peaks exhibit little overlap.

(E) ChIP-seq tracks for cell-type-specific H3K27ac peaks in the *Tead4*, *Dnmt3b*, *Fgf8*, and *Gata2* loci (50 kb).

(F) Fraction of genes (color codes indicated) with cell-type-specific H3K27ac peaks plotted against expression level changes.

(G) Enrichment of transcription factor (TF)-binding sites in cell-type-specific H3K27ac peaks in the four cell types.

(H) Venn diagrams showing the co-existence of multiple TF-binding sites in cell-type-specific H3K27ac peaks in the four cell types. Representative genes and the gene numbers in respective categories are indicated.

See also Figures S3 and S6 and Table S5.

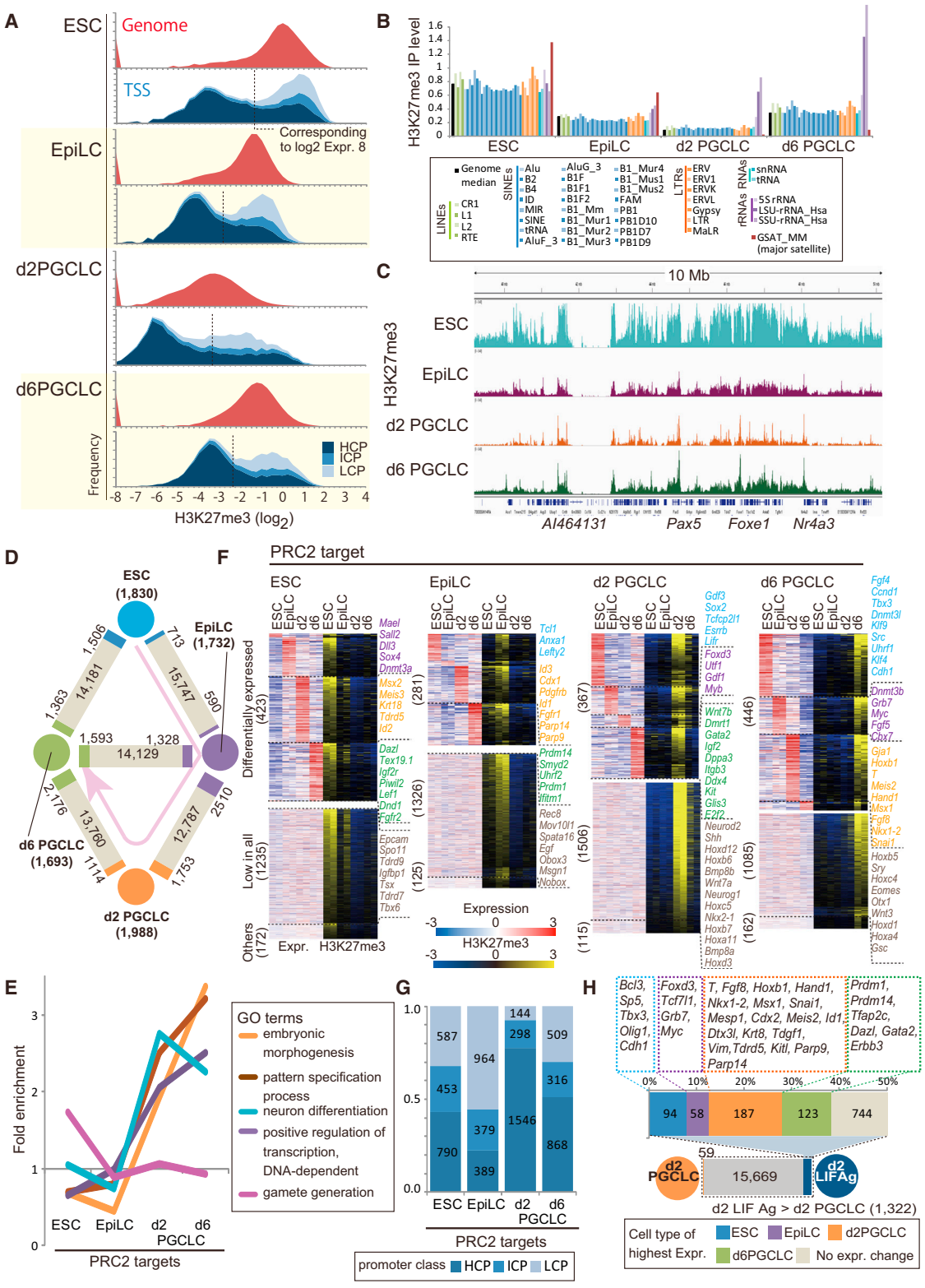


Figure 3. Dynamics of H3K27me3 for In Vitro PGC Specification

(A) The $\log_2$ H3K27me3 IP level-frequency plots for the genome (single-copy regions, 2 kb sliding windows with 1 kb overlaps, red) and the HCP, ICP, and LCP genes (color codes indicated). The $\log_2$ H3K27me3 IP levels corresponding to a $\log_2$ expression level of 8 are indicated by dotted lines.

(B) The H3K27me3 IP levels in repeat elements in the four cell types.

(C) ChIP-seq track transition for the H3K27me3 IP levels in the 10 Mb region around *Pax5*.

(legend continued on next page)

Regulation of Polycomb Repression

We next examined the transitions of H3K27me3 during in vitro PGC specification. Remarkably, the genome-wide IP levels of H3K27me3 (for definition, see Supplemental Experimental Procedures) changed in a dynamic fashion during in vitro PGC specification (Figures 3A–3C and Figures S4A and S4B), a finding demonstrating observations by immunofluorescence and western blot analysis (Seki et al., 2005): the H3K27me3 IP level was highest throughout the genome and around the TSSs in ESCs and decreased significantly in EpiLCs. The genome-wide H3K27me3 reached the lowest IP level in d2 PGCLCs, which, however, exhibited a high level of H3K27me3 specifically around a subset of TSSs (Figures 3A and 3C and Figures S4A and S4B). The H3K27me3 IP level recovered significantly throughout the genome, around the TSSs, and in the gene bodies in d6 PGCLCs (Figures 3A and 3C and Figure S4B). A vast majority of the TSSs of the LCP genes were marked by high-level H3K27me3 in all cell types (Figure 3A).

The H3K27me3 IP levels around the TSSs exhibited bimodal distributions with negative correlation to the gene expression levels in all cell types (Figures 3A and 3C and Figure S4A), and the $\log_2$ expression levels of 8 (~20 copies/cell, a lower limit of appreciable expression) (Figure 1B) (Kurimoto et al., 2006) corresponded to the saddles of the bimodal distributions (Figure 3A). We therefore defined the "enrichment levels" of H3K27me3 around the TSSs as the fold difference from the IP levels corresponding to a $\log_2$ expression level of 8, so that the enrichment levels were log-proportionally correlated with the expression levels.

Accordingly, we identified the TSSs with significant changes of the H3K27me3 enrichment levels (>2 fold) between the successive stages and, consequently, the TSSs with the highest H3K27me3 enrichment levels in ESCs (1,830, the ESC PRC2 targets), EpiLCs (1,732, the EpiLC PRC2 targets), d2 PGCLCs (1,988, the d2 PGCLC PRC2 targets), or d6 PGCLCs (1,693, the d6 PGCLC PRC2 targets) (Figure 3D). Accordingly, the ESC PRC2 targets were enriched in genes for "gamete generation" and "spermatogenesis," whereas the d2 and d6 PGCLC PRC2 targets were enriched in genes for "neuron differentiation," "embryonic morphogenesis," and "pattern specification process" (Figure 3E and Figure S4C). Thus, key developmental regulators are more strongly repressed by PRC2 in PGCLCs than in ESCs and EpiLCs. The fact that the d2 PGCLC genes were enriched with distinct developmental regulators (Figures 1D and 1E) indicates that a subset of developmental regulators are specifically activated, whereas the remaining majority of developmental regulators are repressed by PRC2 in d2 PGCLCs. Indeed, many developmental regulators did not show significant expression during in vitro PGC specification and exhibited particularly high H3K27me3 enrichment levels in d2 and d6 PGCLCs (Figure 3F). The four PRC2 target classes show expression in distinct cell types and there is no apparent link between the timings of the genes acquiring robust PRC2-based repressive states and the activation of their expression (Figure 3F).

The ESC and d6 PGCLC PRC2 targets consisted of ~40% to ~50% HCPs, ~20% ICPs, and ~30% LCPs, whereas the EpiLC PRC2 targets showed depletion of HCPs (~20%) and enrichment with LCPs (~55%), and the d2 PGCLC PRC2 targets were highly enriched with HCPs (~80%) (Figure 3G). This suggests that during the ESC-to-EpiLC transition, many developmental regulators with HCPs reduce their H3K27me3 levels for differentiation toward three germ layers, whereas during the EpiLC-to-d2 PGCLC transition, only specific developmental regulators, in response to BMP4, are activated, while those for unrelated lineages are strongly repressed by PRC2.

To examine the impact of BMP4 signaling on H3K27me3 acquisition, we compared the H3K27me3 states between d2 PGCLCs and the EpiLC aggregates stimulated only by LIF (leukemia inhibitory factor) for two days (d2 LIF Ag). Many developmental regulators and PGC genes upregulated in d2 PGCLCs acquired a high level of H3K27me3 in d2 LIF Ag, indicating that BMP4 signaling prevents the acquisition of or depletes H3K27me3 from these genes (Figure 3H and Figures S4D and S4E).

Transitions in Bivalency

Bivalent promoters with both activating H3K4me3 and repressing H3K27me3 are prevalent in developmental regulators and may represent a poised state for a timely activation of associated genes (Voigt et al., 2013). The bivalency may be key for the germ cell lineage to activate developmental genes at an appropriate time point after fertilization (Sachs et al., 2013). However, how the germ cell lineage may acquire the bivalency has been unknown. We therefore looked at the transitions in bivalency during in vitro PGC specification.

We defined the bivalent genes as those that show no/low expression ($\log_2$ expression levels <8) and bear substantial levels of both H3K27me3 and H3K4me3 (Figures 4A and 4B and Figure S5A). Throughout in vitro PGC specification, a large number of genes were classified as bivalent (1,798, 2,863, 1,559, and 2,435 genes for ESCs, EpiLCs, d2 PGCLCs, and d6 PGCLCs, respectively), and such genes were most significantly enriched with those for "neuron differentiation" (Figures 4C and 4D and Figure S5B). We verified by sequential ChIP-qPCR that the 12 gene loci identified as bivalent in EpiLCs

(D) Changes in the H3K27me3 enrichment level around TSSs. The numbers of TSSs upregulated (>2-fold, marked by corresponding colors) and unchanged (gray) in each cell type compared to a neighboring cell type are indicated. The numbers of TSSs with the highest H3K27me3 enrichment levels in each cell type are indicated in parentheses below each cell type.

(E) Enrichment of the indicated GO terms in the PRC2 targets of the four cell types.

(F) Heat map representation of the PRC2 targets of the four cell types and their expression.

(G) Classification of the PRC2 targets of the four cell types by promoter classes.

(H) Comparison of the H3K27me3 IP levels around TSSs between d2 PGCLCs and d2 LIF Ag. The numbers of TSSs upregulated (>2-fold, marked by corresponding colors) and unchanged (gray) in each cell type are indicated. Genes that show higher H3K27me3 IP levels in d2 LIF Ag are classified by the gene expression classes (d2 and d6 PGCLC genes are enriched).

See also Figures S2, S4, and S6.

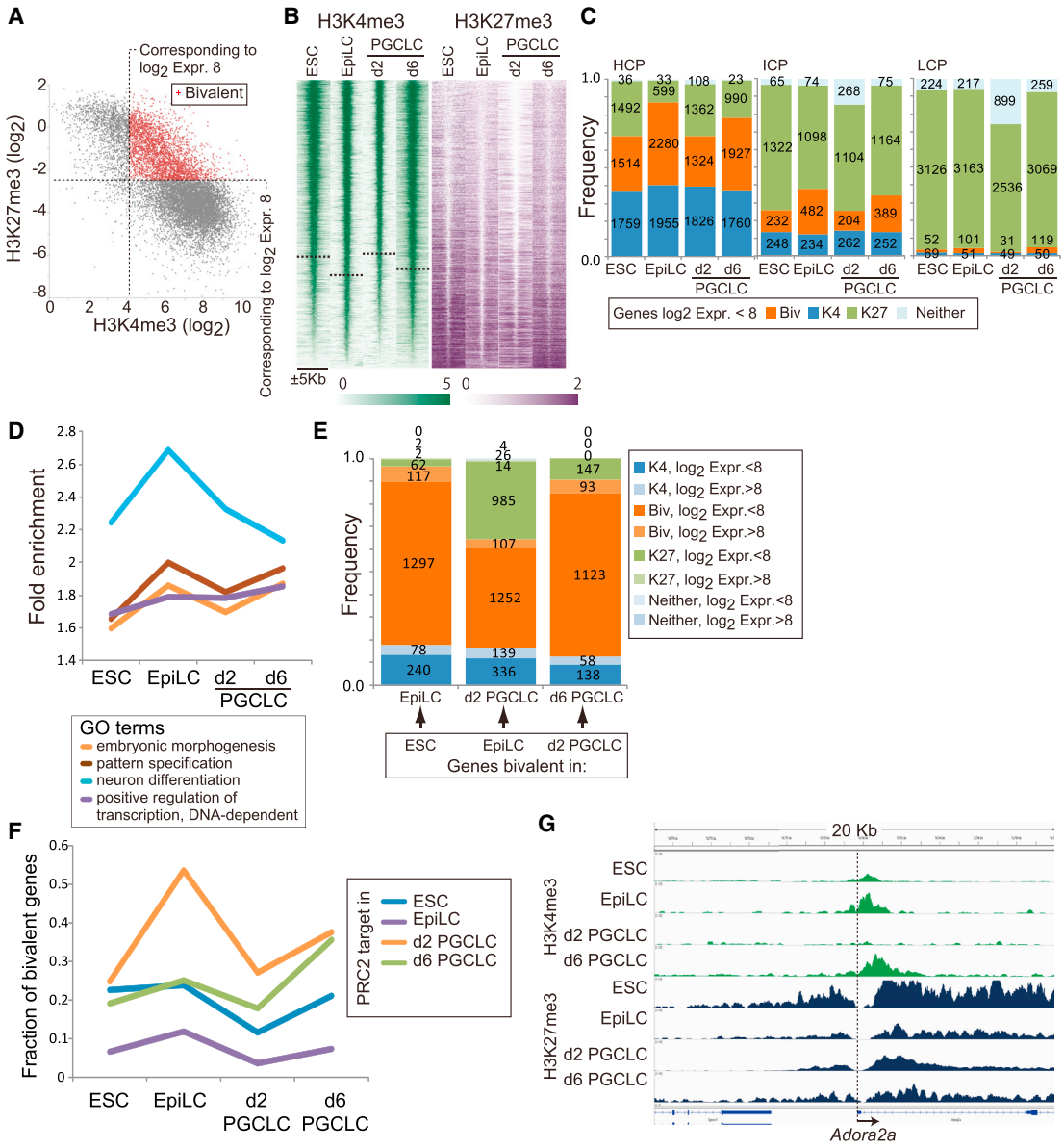


Figure 4. Dynamics of Bivalency for In Vitro PGC Specification

(A) Definition of the bivalent state. The scatter plots show the $\log_2$ levels of the H3K4me3 peaks around the TSSs and those of the neighboring (<1 kb from the centers of the H3K4me3 peaks) H3K27me3 (data for d6 PGCLCs). The modification levels corresponding to a $\log_2$ expression level of 8 are indicated by dotted lines (see Supplemental Experimental Procedures). Bivalent genes (indicated by red dots) are defined as those bearing both H3K27me3 and H3K4me3 at levels higher than those indicated by the dotted lines.

(B) Heat map representation of the H3K4me3 levels ±5 kb from the TSSs and the corresponding H3K27me3 levels. Dotted lines correspond to the genes with a $\log_2$ expression level of 8.

(C) Classification of the bivalent genes (orange, with their numbers) by promoter classes. Only those genes that were expressed at levels lower than a $\log_2$ expression level of 8 were evaluated (see also Figure S5A). The genes that bear H3K4me3 only (blue), H3K27me3 only (green), or neither modification (pale) are also shown, with their numbers.

(D) GO term enrichment in the bivalent genes in the four cell types.

(E) The change of the H3K4me3 and H3K27me3 states of the bivalent genes in the stage transitions. The numbers of the genes whose chromatin states are changed from bivalency to the indicated states are shown.

(F) Fraction of the bivalent genes among the PRC2 targets in the four cell types.

(G) ChIP-seq track transition for H3K4me3 and H3K27me3 in the 20 kb region around *Adora2a*, which upregulates H3K4me3 and regains bivalency in d6 PGCLCs.

See also Figures S5 and S6.

indeed bear both H3K4me3 and H3K27me3 on the same loci (Figure S5C).

A majority (1,297/1,798, ~72%) of the bivalent genes in ESCs remained bivalent during the ESC-to-EpiLC transition and a large number (1,566) of genes were newly recruited as bivalent in this transition (Figure 4E and Figure S5D); consequently, EpiLCs bore the largest number of bivalent genes (Figure 4C). During the EpiLC-to-d2 PGCLC transition, however, only ~44% (1,252/2,863) of the bivalent genes in EpiLCs remained bivalent (307 genes became newly bivalent in d2 PGCLCs) and, remarkably, as many as ~34% (985/2,863), which were highly enriched with genes for "neuron differentiation" ($p = 2.3 \times 10^{-24}$), only became H3K27me3 only in d2 PGCLCs (Figure 4E and Figure S5D). Accordingly, among the d2 PGCLC PRC2 targets, more than 50% were bivalent in EpiLCs, but less than 30% were bivalent in d2 PGCLCs (Figure 4F), indicating that the d2 PGCLC PRC2 targets mainly reflect the bivalent genes in EpiLCs and that, compared to EpiLCs, d2 PGCLCs lose their potential for "neuron differentiation" and are restricted to an early mesodermal/PGC state.

In contrast, during the d2-to-d6 PGCLC transition, a majority (1,123/1,559, ~72%) of the bivalent genes in d2 PGCLCs remained bivalent and a larger number (1,312) of genes were newly recruited as bivalent (Figure 4E and Figure S5D). As a result, ~40% of the d2 and d6 PGCLC PRC2 targets were bivalent in d6 PGCLCs (Figure 4F), indicating that d6 PGCLCs re-acquire a potentially developmentally more permissive state; i.e., d6 PGCLCs gain H3K4me3 around the TSSs of the genes repressed by H3K27me3 (Figure 4G). Additionally, consistent with the depletion of HCPs in the EpiLC PRC2 targets, only ~10% of the EpiLC PRC2 targets were bivalent throughout in vitro PGC specification (Figures 3G and 4F). Thus, developmental regulators with HCPs are in more permissive states with relatively low H3K27me3 and high H3K4me3 in EpiLCs, acquire highly repressive states in d2 PGCLCs, and regain potentially more permissive states in d6 PGCLCs, although it is important to note that d2 and d6 PGCLCs exhibit higher H3K27me3 enrichment levels in developmental regulators.

We compared the bivalent genes in d6 PGCLCs with those in E11.5 PGCs (Sachs et al., 2013). A majority of the bivalent genes in d6 PGCLCs were in common with those in E11.5 PGCs (1,423/2,435) (Figure S5E). While the bivalent genes specific to d6 PGCLCs (1,012/2,435) did not show enrichment in GO terms for major developmental processes such as "neuron differentiation," "pattern specification," and "embryonic morphogenesis," those specific to E11.5 PGCs showed enrichment of all such terms (Figure S5E). Thus, d6 PGCLCs are similar in bivalency to E11.5 PGCs and may acquire more bivalency when cultured further under an appropriate condition.

Genome-wide Reduction of H3K9me2

We next evaluated the transitions of H3K9me2 during in vitro PGC specification. Strikingly, the H3K9me2 level was drastically reduced throughout the genome and around the TSSs in d6 PGCLCs (Figures 5A–5D), reinforcing the observations made by immunofluorescence and western blot analysis (Seki et al., 2005). Throughout the in vitro PGC specification, the TSSs associated with high-level H3K9me2 (IP level > 1) were predominantly those of the LCP genes and were enriched with genes for "inflammatory response" and "defense response" (Figure 5E), whereas a majority of the HCP and ICP genes were essentially devoid of this modification (Figure 5B). The H3K9me2 levels around the TSSs were highly anti-proportional to the expression levels of the associated genes (Figure 5D and Figure S4A).

Figure 5F illustrates ChIP-seq tracks for the H3K27me3 and H3K9me2 transitions in a 50 Mb region on chromosome 11: the H3K9me2 level was high in ESCs, EpiLCs, and d2 PGCLCs and significantly downregulated in d6 PGCLCs. The H3K9me2 levels in d2 PGCLCs were distributed in a narrower range compared to those of the other cell types (Figure 5B and Figure S4A), which may signify a unique property of d2 PGCLCs (i.e., transient acquisition of a somatic program) during in vitro PGC specification. In ESCs, EpiLCs, and d2 PGCLCs, H3K9me2 was enriched in lamina-associated domains (LADs), the genome regions that are associated with the nuclear lamina (NL) and are long interspersed nuclear element (LINE)-rich and gene-poor, and typically harbor low or no gene-expression levels (Figure 5F) (Guelen et al., 2008; Peric-Hupkes et al., 2010). Remarkably, d6 PGCLCs showed a reduction of H3K9me2 in LADs as well (Figures 5F and 5G).

We found that the expression of *Lamin A* is progressively decreased, while expression of *Lamin B1* and *B2* remain relatively constant during in vitro PGC specification (Figure 5H). The immunofluorescence analysis, however, revealed that, compared to EpiLCs, d6 PGCLCs showed a reduced level of LAMIN B1 and appeared to have a diminished level of DAPI-dense heterochromatin around the nuclear membrane (Figure 5I). However, similar to ESCs and EpiLCs, d6 PGCLCs exhibited a relative depletion of H3K27me3 in LADs (Figure 5G). d6 PGCLCs may thus bear a chromosome structure relatively free from constraints imposed by LAD-NL interactions, although they may still retain some features distinguishing LADs from other parts of the genome.

Repression of "Germline Genes" by H3K27me3 and H3K9me2

Upon colonization of embryonic gonads, PGCs upregulate a set of genes, often referred to as "germline genes," that are crucial for their progression into meiosis in females and for transposon repression in males. It has been proposed/reported that these genes are activated by the genome-wide DNA demethylation in PGCs (Hackett et al., 2012; Maatouk et al., 2006) and lack histone modifications in somatic cells (Hackett et al., 2012). It is unknown how they are regulated by histone modifications in early PGCs. We therefore explored their histone modifications during in vitro PGC specification, when a majority of them were in a repressed state (Figure 6A).

The key germline genes, *Ddx4* and *Dazl*, lacked H3K4me3 and H3K27ac throughout the in vitro PGC specification, despite the fact that they showed low-level expression in d6 PGCLCs (Figures 6A and 6B). Remarkably, a majority of the germline genes harbored a high level of H3K27me3 throughout the in vitro PGC specification (Figures 6B and 6C). Because they essentially lacked H3K4me3, less than 10% of them showed bivalency (Figure S5F). Interestingly, they were marked by an exceptionally high level of H3K9me2 in ESCs and EpiLCs, and

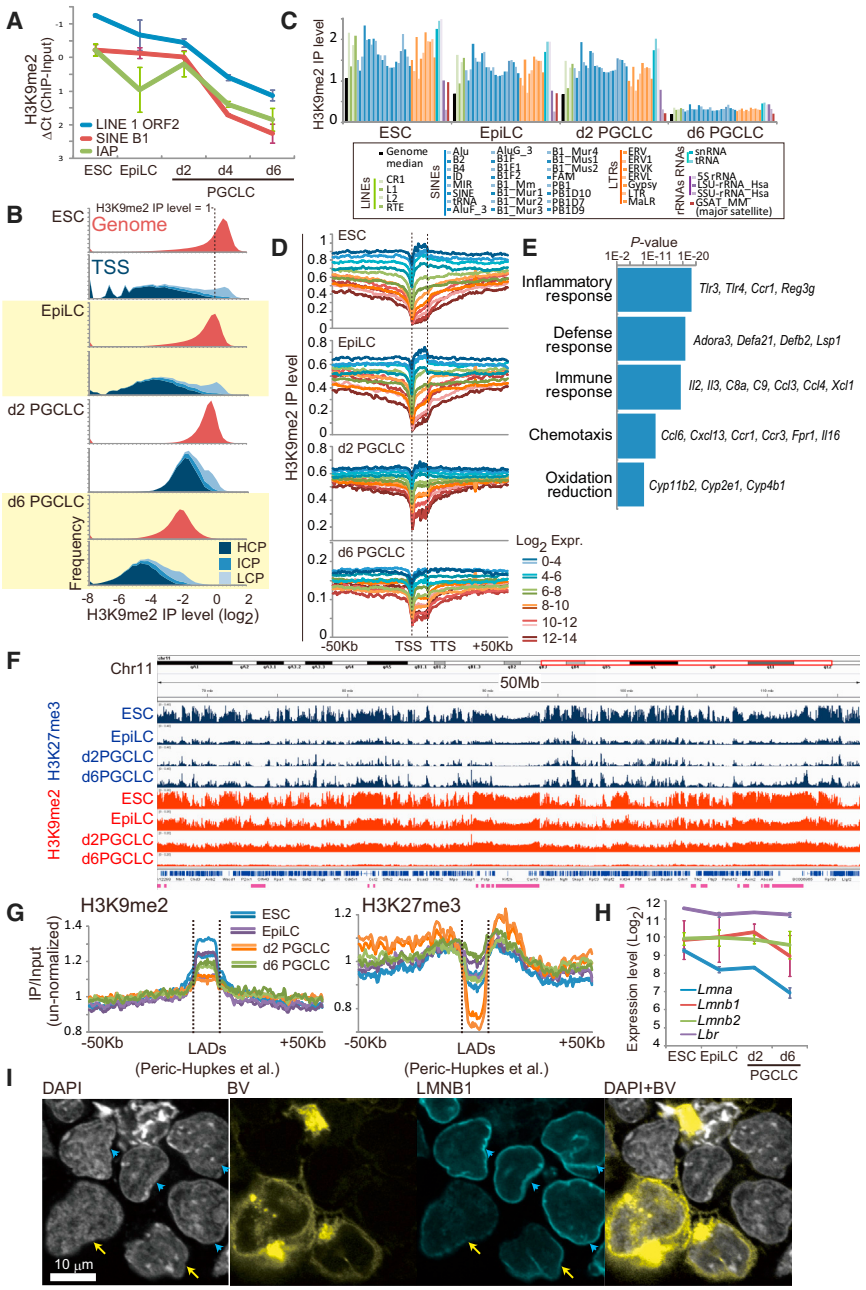


Figure 5. Dynamics of H3K9me2 for In Vitro PGC Specification

(A) ChIP-qPCR (with SDs, two biological replicates) for H3K9me2 in the indicated repeat elements.

(B) The $\log_2$ H3K9me2 IP level-frequency plots for the genome and the HCP, ICP, and LCP genes.

(C) The H3K9me2 IP levels in repeat elements in the four cell types.

(D) The relationship between the $\log_2$ expression levels of genes (color codes indicated) and the H3K9me2 levels around them (±50 kb from their TSSs and TTSs [transcription termination sites]) in the four cell types.

(E) GO term enrichment in genes with high H3K9me2 levels (IP level > 1) in ESCs.

(F) ChIP-seq track transitions for H3K27me3 (top, blue) and H3K9me2 (bottom, red) IP levels in a 50 Mb region on chromosome 11. The lamina-associated domains (LADs) (Peric-Hukpes et al., 2010) are indicated by red bold lines at the bottom.

(G) H3K9me2 and H3K27me3 distributions around LADs (±50 kb) in the four cell types (color codes indicated).

(H) Expression of *Lmna*, *Lmnb1/b2*, and *Lbr* (SDs, two biological replicates).

(I) Immunofluorescence analysis of heterochromatin structure identified by DAPI staining (white) and LAMIN B1 expression (blue) in d6 PGCLCs (BV-positive cells, yellow, indicated by yellow arrows) and EpiLCs (BV-negative cells, indicated by blue arrows). Bar, 10 μm.

See also Figure S2.

they continued to show such a trend in d6 PGCLCs with a very low level of genome-wide H3K9me2 (Figures 6B and 6D).

We analyzed the published data on expression of H3K27me3 and its distribution in the germline genes in PGCs at E11.5 (Kagiwada et al., 2013; Sachs et al., 2013). Many germline genes were expressed in PGCs at E11.5 (Figure 6E), and in general, there was a negative correlation between the expression and the H3K27me3 levels of the germline genes, although some germline genes expressed at this stage still exhibited relatively high levels of H3K27me3 (Figures 6F and 6G). Thus, we propose that, unlike in somatic cells, the germline genes are repressed by both H3K27me3 and H3K9me2 during PGC/PGCLC specification and development.

EpiLCs versus EpiSCs

Having established the chromatin-state dynamics during in vitro PGC specification, we compared the chromatin states between EpiLCs and epiblast stem cells (EpiSCs) (Factor et al., 2014), which bear robust and little, if any, competence for the germ cell fate, respectively (Hayashi et al., 2011).

Genes bearing higher H3K27ac peaks in EpiSCs than in EpiLCs were enriched in the GO term for "regulation of transcription, DNA-dependent," including characteristic TFs (Figure S6A), indicating that EpiSCs bear distinct transcriptional circuitry with unique epigenetic foundations. EpiSCs, but not EpiLCs, expressed genes associated with pattern specification/tissue morphogenesis such as *Hoxa1*, *Hoxb1-b3*, *Snai1*, *Id3*, and *Fgf8*, albeit at low levels, showing substantial H3K4me3 and H3K27ac peaks around such genes (Figure S6B). In good agreement with these findings, genes bearing higher H3K27ac peaks in EpiSCs than in EpiLCs included higher numbers of genes for "neuron differentiation," "embryonic development," and "pattern specification" than genes bearing higher H3K27ac peaks in EpiLCs than in EpiSCs (Figure S6C). Moreover, we found that the bivalent genes specific to EpiSCs, but not EpiLCs, were enriched with those for "pattern specification" and "embryonic morphogenesis" (Figure S6D). Thus,

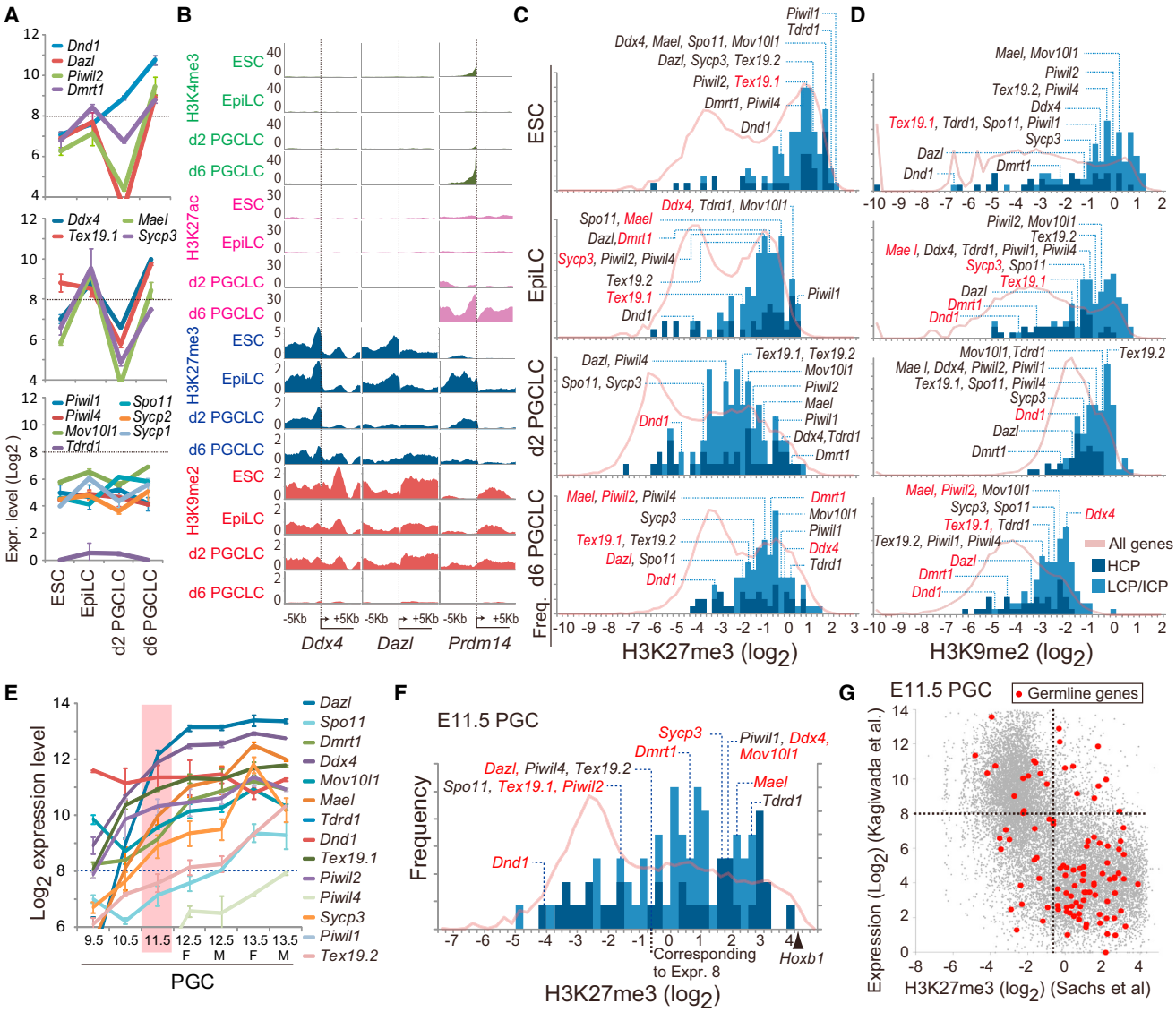


Figure 6. Regulation of Germline Genes for In Vitro PGC Specification

(A) Expression of the indicated germline genes (SDs, two biological replicates).

(B) ChIP-seq track transitions for H3K4me3, H3K27ac, H3K27me3, and H3K9me2 ±5 kb from the TSSs (dotted lines) of *Ddx4*, *Dazl*, and *Prdm14*.

(C and D) The $\log_2$ H3K27me3 (C) or H3K9me2 (D) IP level-frequency plots for germline genes (102 genes). Representative germline genes are indicated, and those expressed at more than a $\log_2$ expression level of 8 are colored in red. Pale red lines represent the plots for all genes as a reference.

(E) Expression of the indicated germline genes in PGCs from E9.5 to E13.5 (F, female; M, male) (SDs, two biological replicates) (Kagiwada et al., 2013).

(F) The $\log_2$ H3K27me3 level-frequency plots for germline and all genes in PGCs at E11.5 (Sachs et al., 2013) as in (C). The H3K27me3 level for *Hoxb1* is indicated by a triangle.

(G) The $\log_2$ H3K27me3 level-$\log_2$ expression level plots for germline (red dots) and all (gray dots) genes in PGCs at E11.5 (Kagiwada et al., 2013; Sachs et al., 2013). Dotted lines indicate a $\log_2$ expression level of 8 and the $\log_2$ H3K27me3 level corresponding to a $\log_2$ expression level of 8.

See also Figure S6.

EpiSCs bear an epigenetic property representing a more primed state for somatic differentiation than EpiLCs. Interestingly, the germline genes exhibited much higher H3K27me3 in EpiLCs than in EpiSCs, indicating that they are properly repressed in EpiLCs for appropriate later expression (Figure S6E). Thus, EpiLCs and EpiSCs bear different epigenetic properties reflecting their transcriptional and functional differences.

Distribution of BLIMP1 and T

We determined the binding sites of BLIMP1 and T, two key transcriptional regulators for PGC specification (Aramaki et al., 2013; Ohinata et al., 2005), and explored their relationship with the chromatin states during in vitro PGC specification. Figures 7A and 7B provide examples of the BLIMP1-binding peaks in the *Klf9* and *Hoxa* loci and the T-binding peak in the *Prdm1* locus,

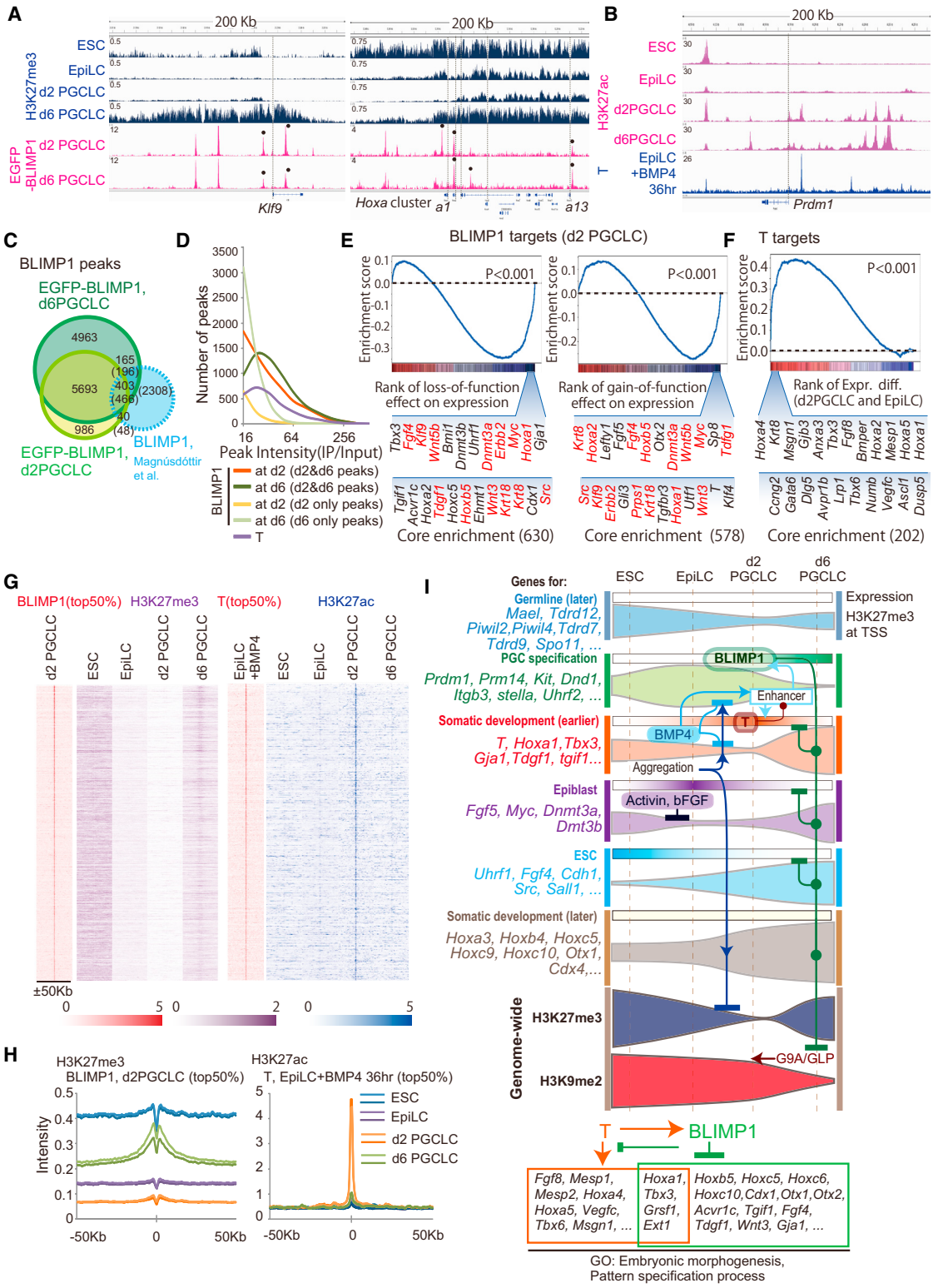


Figure 7. Distribution of BLIMP1 and T for In Vitro PGC Specification

(A) ChIP-seq track transitions for H3K27me3 binding (top, blue) and BLIMP1 binding (bottom, red) in the 200 kb regions around *Klf9* (left) and the *Hoxa* cluster (right).

(B) ChIP-seq track transitions for H3K27ac-binding (top, red) and T-binding (bottom, blue) in the 200 kb region around *Prdm1*. The genome-wide binding profiles of T were determined in floating aggregates of EpiLCs stimulated by BMP4 and LIF for 36 hr (Aramaki et al., 2013). In (A) and (B), the dotted lines indicate the TSSs of the genes bound by the TFs.

(legend continued on next page)

respectively (see also Figure S7A). We identified 7,122 and 11,224 EGFP-BLIMP1 binding peaks in d2 and d6 PGCLCs, respectively, which were highly overlapped with each other (Figures 7C and 7D), but were only partially overlapped with those identified in P19 cells overexpressing *Blimp1* (Figure 7C) (Magnúsdóttir et al., 2013), indicating that the cellular context plays a key role in the distribution and function of TFs (Nakaki et al., 2013). Both BLIMP1- and T-binding sites were enriched around TSSs (Figure S7B) and were distributed in different promoter classes in a proportional manner (Figure S7C). The motifs for BLIMP1 and T binding showed partial, but not perfect, overlaps with previously identified motifs, extending their potential targets (Figures S7D–S7G).

We explored the relationship between the genes bound by BLIMP1 and (1) the genes upregulated or downregulated in *Blimp1*-knockout cells (Kurimoto et al., 2008) and (2) the genes upregulated or downregulated in EpiLCs overexpressing *Blimp1* (Nakaki et al., 2013) by gene set enrichment analysis (GSEA). Both the genes upregulated in *Blimp1*-knockout cells and those downregulated in EpiLCs overexpressing *Blimp1* were enriched with those bound by BLIMP1, revealing 630 and 578 targets for BLIMP1 repression, respectively (Figure 7E and Figure S7H). However, it should be noted that BLIMP1 binds to a much higher number of genes than it appears to regulate at an appreciable level (Figure 7E). We also found by GSEA that T preferentially targets at least 202 genes upregulated in d2 PGCLCs (Figure 7F and Figure S7H).

The targets of BLIMP1 in d2 PGCLCs showed only a small overlap with those of T (Figure S7I): while both BLIMP1 and T targets were enriched in genes for embryonic development/pattern specification processes, targets of T as an activator were skewed to genes involved in early mesoderm formation and those of BLIMP1 as a repressor were for more various biological processes (Figures S7J and S7K). Remarkably, the BLIMP1-binding sites in d2 PGCLCs, which were mostly preserved in d6 PGCLCs, exhibited an enrichment of H3K27me3 in d6, but not d2, PGCLCs (Figures 7G and 7H), suggesting that BLIMP1 functions as a nucleator for PRC2 recruitment and spreading for a stable repression of gene expression. In contrast, the T-binding sites were enriched with H3K27ac (Figures 7G and 7H), consistent with the notion that T activates transcription by recruiting histone acetyltransferase in a sequence-dependent fashion.

DISCUSSION

We have established the concept of epigenetic reprogramming upon PGC specification by a comprehensive analysis of transcription and chromatin-state dynamics during in vitro PGC specification (Figure 7I). We showed that the H3K27me3 IP levels, including those around the TSSs, exhibited dynamic global changes with negative correlations with gene expression levels in all cell types (Figure 3A and Figure S4A), suggesting that the absolute levels of H3K27me3 around the TSSs in a certain cell type may not necessarily be indicative of the expression levels of the associated genes. This is consistent with the observations that ESCs lacking PRC2 activity appear to proliferate in a normal fashion, although their differentiation potentials are severely perturbed (Margueron and Reinberg, 2011). The significance of the genome-wide deposition of high H3K27me3 in ESCs therefore remains unclear. The high H3K27me3 levels may reflect gene expression circuitry for (potentially) pluripotent cells, since d6 PGCLCs, which regain the relevant genes (Figures 1D and 1E and Figure S1A), also exhibit high H3K27me3 throughout the genome (Figure 3A and Figure S4A). Accordingly, there appears to be a tendency for the PRC2 components to be expressed at higher levels in both ESCs and d6 PGCLCs (Figure S4F). BLIMP1 as a potential nucleator for H3K27me3 deposition and spreading (Figures 7A, 7G, and 7H and Figure S7A) may also contribute to the high H3K27me3 level in d6 PGCLCs.

Coupled with the dynamic regulation of H3K27me3, genes bearing bivalent promoters exhibit dynamic changes during in vitro PGC specification (Figure 4). Notably, EpiLCs bear the highest number of bivalent genes with the highest enrichment for developmental regulators, and at the same time, they show the narrowest range of the H3K27me3 IP level among the four cell types (Figure 3A and Figure 4), suggesting that EpiLCs acquire an epigenome in which a wide range of developmental regulators are repressed by relatively low levels of H3K27me3 and bear low levels of H3K4me3, which may represent a bona fide primed state immediately ready for differentiation into cells of the three germ layers. In good agreement with this idea, EpiLCs do not overtly upregulate genes for developmental regulators (Figures 1D and 1E), and EpiLCs, but not ESCs, respond quickly to BMP4 to deplete or prevent the acquisition of H3K27me3 from genes for early mesoderm induction/PGC

(C) Venn diagrams showing the overlap of BLIMP1 peaks in d2 and d6 PGCLCs and those reported previously (Magnúsdóttir et al., 2013).
(D) The number of BLIMP1 peaks plotted against the intensity of BLIMP1 peaks in d2 and d6 PGCLCs (color codes as indicated).
(E) Gene set enrichment analysis (GSEA) of BLIMP-bound genes and gene expression difference caused by loss of function (Kurimoto et al., 2008) (left; upregulated genes on the right) and gain of function (Nakaki et al., 2013) (right; downregulated genes on the right) of *Blimp1*. Genes marked in red are core enrichment genes determined by both loss- and gain-of-function effects of *Blimp1*.
(F) GSEA of T-bound genes and gene expression difference between EpiLCs and d2 PGCLCs (genes upregulated in d2 PGCLC on the left). In (E) and (F), all genes were ordered according to the rank of the difference of expression levels.
(G) (Left) Heat map representation of BLIMP1 peaks in d2 PGCLCs (top 50%), which show enrichment of H3K27me3 (±50 kb around the TSSs) in d6, but not d2, PGCLCs. (Right) Heat map representation of T peaks in EpiLC aggregates stimulated by BMP4 (Aramaki et al., 2013), which show enrichment in H3K27ac (±50 kb around the TSSs) in d2, but not d6, PGCLCs.
(H) (Left) Averaged intensity plots for H3K27me3 levels in the four cell types around ±50 kb of the top 50% BLIMP1-binding peaks in d2 PGCLCs. (Right) Averaged intensity plot for H3K27ac levels in the four cell types around ±50 kb of the top 50% T-binding peaks in EpiLC aggregates stimulated by BMP4 for 36 hr. Color codes for the cell types are indicated on the right.
(I) (Top) A model for transcription and chromatin-state dynamics during in vitro PGC specification. (Bottom) Differential regulation of genes for "embryonic morphogenesis" and "pattern specification" by BLIMP1 (repression) and T (activation).
See also Figures S1 and S7and Table S6.

specification to be induced as PGCLCs (Figure 3H) (Hayashi et al., 2011): d2 PGCLCs instead deplete H3K4me3 from, and deposit high H3K27me3 onto, a large number of developmental regulators for unrelated lineages for their repression (Figures 3 and 4). Additionally, compared to ESCs, EpiLCs acquire a distinct enhancer pattern (Figure 2) (Buecker et al., 2014), which would also contribute to their developmental potential. It will be important to explore whether pre-gastrulating epiblasts bear an epigenomic equivalent to EpiLCs.

d6 PGCLCs upregulate the global H3K27me3 levels (Figure 3) and newly recruit a large number of genes, many of which are marked only by H3K27me3 in d2 PGCLCs, as bivalent genes (Figure 4 and Figure S5D). This indicates that d6 PGCLCs introduce H3K4me3 onto many genes at levels that maintain their H3K27me3-based repression. Accordingly, d6 PGCLCs upregulated a gene encoding an H3K4 methyltransferase (MTase) (*Mll3*) (Figure S4G). The mechanism by which d6 PGCLCs restore the bivalency and the functional significance of bivalency in PGCs/PGCLCs warrants further investigation.

Our results demonstrated a striking reduction of the H3K9me2 levels in d6 PGCLCs (Figure 5). Repression of the H3K9 MTases G9A and GLP in PGCs/PGCLCs may contribute to the creation of this unique state (Seki et al., 2007), a possibility that requires experimental validation. What, then, would be the potential consequences of the genome-wide reduction of H3K9me2 in PGCs/PGCLCs? We showed that d6 PGCLCs bear altered nuclear architecture with diminution of DAPI-dense heterochromatin and reduced concentration of LAMIN B1 around the nuclear periphery (Figure 5I). This unique nuclear architecture of d6 PGCLCs is reminiscent of that of PGCs (Kagiwada et al., 2013) and may be a key of epigenetic reprogramming in PGCs/PGCLCs. Considering the reported link between the loss of H3K9me2 and DNA demethylation (Tachibana et al., 2008), another potential result would be that PGCs undergo genome-wide DNA demethylation (Seisenberger et al., 2012). Indeed, compared to EpiLCs, PGCLCs show substantial genome-wide DNA demethylation (K. Shirane, K.K., M.S., and H. Sasaki, data not shown). Further studies will be necessary to explore these possibilities.

A recent study showed that the germline genes lack any repressive histone modifications and are repressed solely by DNA methylation in somatic cells (Hackett et al., 2012). We showed that the germline genes are enriched in ESC PRC2 targets and a majority of them bear high H3K27me3 and H3K9me2 levels throughout in vitro PGC specification (Figure 6 and Figure S4C). These findings suggest that the germline genes are repressed by both DNA methylation and histone modifications in PGCs. In good agreement with this idea, many of the germline genes expressed in PGCs at E11.5 are depleted of H3K27me3 from their promoters at the same stage (Figures 6E–6G) (Sachs et al., 2013), and the germline genes are upregulated prematurely in PGCs with impaired PRC1 activity (Yokobayashi et al., 2013). The comprehensive profiling of transcription and chromatin-state transitions for the in vitro PGC specification pathway provided herein should serve as a robust foundation not only for reconstituting the regulatory networks underlying epigenetic reprogramming in PGCs, but also for developing culture conditions for the induction of PGCLCs into a more mature phenotype.

EXPERIMENTAL PROCEDURES

The experimental procedures for analysis of the gene expression for in vitro PGC specification, generation of *EGFP-Blimp1* knockin mice and ESCs, induction from ESCs of EpiLCs and PGCLCs, normalization of the ChIP-seq data, ChIP-seq data analysis, sequential ChIP-qPCR, comparison of the ChIP-seq data for EpiLCs with those for EpiSCs, GSEA, immunofluorescence analysis of spread cells, immunofluorescence analysis of cryosections, histology, and western blot analysis are available in the Supplemental Information.

ChIP

BVSC and *EGFP-Blimp1* knockin cells were used for ChIP of histone H3/histone H3 modifications (H3K4me3, H3K9me2, H3K27ac, and H3K27me3) and EGFP-BLIMP1 respectively. Mouse monoclonal antibodies were used for ChIP-seq/ChIP-qPCR for histone H3 modifications (Hayashi-Takanaka et al., 2011), and rabbit polyclonal antibodies were used for histone H3 and EGFP-BLIMP1 (Table S2 and Table S3).

Prior to IP, the mouse antibodies and the rabbit antibodies were reacted with 20 μl of M280 Dynabeads Sheep anti-mouse IgG and 10 μl of M280 Dynabeads Protein G (Life Technologies), respectively.

ChIP was performed as described previously (Luo et al., 1998) with modification. Cells were suspended in 1 ml of PBS and fixed by the addition of 27 μl of 36.5% formalin (Sigma) for 10 min at room temperature. Fixation was terminated by the addition of 110 μl of 1.5M glycine. The fixed cells were washed twice in PBS containing 0.1% BSA and were lysed in 400 μl SDS-lysis buffer (50 mM Tris-HCl [pH 8.0] containing 1% SDS, 10 mM EDTA, and 1 mM PMSF) on ice for 10 min. Chromatins were then solubilized using a Bioruptor UCD250 with ice-chilled water, with 10 cycles of 30 s sonication at high power and 60 s intervals. The sonication products were centrifuged at 16,000 × *g* for 5 min. The supernatants were diluted in 1 ml ChIP dilution buffer (16.7 mM Tris-HCl [pH 8.0], 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 167 mM NaCl, 1 mM PMSF, 2 mg/ml tRNA [Roche]) and were divided into 0.1 ml as input DNA and 1.3 ml as ChIP samples.

The ChIP samples were mixed with the Dynabeads-antibody complexes and rotated at 4°C overnight. The ChIP-ed Dynabeads were recovered using a DynaMag-2 magnet (Life Technologies) and washed with 200 μl of low salt buffer (20 mM Tris-Cl [pH 8.0] containing 0.1% SDS, 1% Triton X-100, 2 mM EDTA, and 150 mM NaCl), a high salt buffer (20 mM Tris-Cl [pH 8.0] containing 0.1% SDS, 1% Triton X-100, 2 mM EDTA, and 500 mM NaCl), RIPA buffer (10 mM Tris-Cl [pH 8.0] containing 1% sodium deoxycholate, 1% NP40, 250 mM LiCl, and 1 mM EDTA), and TE buffer (10 mM Tris-Cl [pH 8.0] containing 1 mM EDTA). The ChIP-ed Dynabeads were then incubated in new tubes containing 50 μl of elution buffer (0.1 M $NaHCO_3$, 1% SDS, 10 mM DTT, and 60 ng/μl tRNA) for 15 min at room temperature. The supernatants were collected using DynaMag-2, and they were reverse-crosslinked by the addition of 8 μl of 2.5 M NaCl and being incubated at 65°C overnight. Then, 2 μl of 0.5 M EDTA, 4 μl of 1M Tris-Cl (pH 6.5), and 0.2 μl of 20 mg/ml proteinase K solution were added to the reaction mixture, which was further incubated at 45°C for 1 hr. The ChIP-ed DNAs were then purified with Qiaquick PCR purification columns, using the buffer EB supplemented with 4 ng/μl tRNA (QIAGEN).

Library Preparation for Sequencing

The ChIP-ed and input DNAs were sheared to an average size of about 150 bp by ultra-sonication (Covaris). The ChIP-ed DNAs obtained from 1 × 10^6 cells were then subjected to the procedures for SOLiD Fragment Library Preparation (Life Technologies) according to the manufacturer's instructions, except for the ligation reaction, which was performed overnight.

For histone H3 and its modifications, ChIP-ed DNAs from 1 × 10^5 cells were subjected to the procedures for SOLiD Fragment Library Preparation with a modification, in which the size fractionation step using AMPure XP beads (Agencourt) was omitted and a ligation reaction was performed overnight.

We developed a new method to prepare and precisely amplify libraries of ChIP-ed DNA from a small number of cells, and we applied this method to H3K4me3 from 1 × 10^5 cells and TFs from 3 × 10^5 ~1 × 10^6 cells. We prepared the modified adaptor for the SOLiD library by annealing 50 μM each of P1-T Adaptor/F and Barcode-Internal+12-mer/R (Table S1) in 1 × Linker buffer (50 mM Tris-HCl [pH 8.0] containing 100 mM NaCl and 1 mM EDTA) with a

step-wise cooling procedure (95°C, 70°C, 50°C, 40°C, and 25°C for 15 min each). We end-repaired the sheared DNAs by using an End-It DNA End-Repair Kit (Epicenter) on a 120 μl scale according to the manufacturer's instructions, and we purified them by using a MinElute PCR purification kit (QIAGEN) with 540 μl of binding buffer PB3 (5.6 M Gu HCl, 33.3% isopropanol) supplemented with 10 ng/μl tRNA and 10 μl of elution buffer EB (QIAGEN) with 4 ng/μl tRNA. The end-repaired DNAs were incubated in 20 μl of dA tailing mixture (0.2 mM dATP [GE Healthcare], 1 × High Yield buffer [Greiner], 4 ng/μl tRNA, 2.5 units of Taq DNA polymerase [Greiner]) for 10 min at 70°C, and we purified them by using a MinElute PCR purification kit with the buffers PB (QIAGEN) and EB (11 μl) supplemented with tRNA. The dA-tailed DNAs were incubated in 15 μl ligation mixture (1 × DNA Ligase Reaction Buffer [Life Technologies], 0.1 μM of the modified adaptor for the SOLiD library, and 2.5 units of ExpressLink T4 DNA Ligase [Life Technologies]) at 16°C for more than 22 hr, and we purified them using a MinElute PCR purification kit with the buffers PB and EB (20 μl) supplemented with tRNA. We then amplified the adaptor-ligated DNAs by using the Platinum Taq DNA Polymerase on a 50 μl reaction scale (Life Technologies) with 1 μM each of Library PCR primer 1 (Life Technologies) and Library PCR primer Barcode001+Internal adaptor (Table S1). We purified the PCR products by using 60 μl AMPure XP beads supplemented with 3.33 μl of 25% PEG8000 and 13.7 μl of 2.5 M NaCl, and they were eluted in 30 μl of buffer EB supplemented with tRNA. We then subjected the PCR products to another PCR by using 1 μM each of Library PCR Primer 1 and P2-Barcode-*N*-Internal primer (Table S1), and we purified them twice with AMPure XP beads supplemented with PEG8000 and NaCl. The barcode-tagged library DNAs were sequenced on a SOLiD 5500xl platform (Life Technologies) to generate single-end 50 bp reads. The experimental outlines, including PCR cycle numbers, and mapping statistics are summarized in Table S3.

ACCESSION NUMBERS

The ChIP-seq and normalized microarray data used in this study were deposited in the NCBI database (GEO accession numbers: GSE60204 and GSE60018).

SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes seven figures, six tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2015.03.002.

AUTHOR CONTRIBUTIONS

K.K. and Y.Y. conducted the overall experiments and analyzed the data. K.H., H.O., Y.M., and K.K. contributed to the PGCLC induction. H.K. contributed to the generation of *EGFP-BLIMP1* knockin mice. T.M., H.K., T.Y., Y.K., and K.S. contributed to the ChIP-seq experiments. K.K. and M.S. conceived the project, designed the experiments, and wrote the manuscript.

ACKNOWLEDGMENTS

We thank H. Sasaki, K. Shirane, T. Abe, and M. Yamaji for discussion and R. Kabata, S. Ohsako, N. Konishi, S. Watanabe, M. Kabata, and T. Satoh for their assistance. We used the supercomputer of ACCMS, Kyoto University. The authors were supported in part by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, by JST-CREST/ERATO, by the Takeda Science Foundation, and by the Tanabe Medical Frontier Conference.

Received: August 4, 2014
Revised: January 14, 2015
Accepted: February 27, 2015
Published: March 19, 2015

REFERENCES

Aramaki, S., Hayashi, K., Kurimoto, K., Ohta, H., Yabuta, Y., Iwanari, H., Mochizuki, Y., Hamakubo, T., Kato, Y., Shirahige, K., and Saitou, M. (2013). A mesodermal factor, T, specifies mouse germ cell fate by directly activating germline determinants. Dev. Cell *27*, 516–529.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of enhancer patterns in transition from naive to primed pluripotency. Cell Stem Cell *14*, 838–853.

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? Mol. Cell *49*, 825–837.

Factor, D.C., Corradin, O., Zentner, G.E., Saiakhova, A., Song, L., Chenoweth, J.G., McKay, R.D., Crawford, G.E., Scacheri, P.C., and Tesar, P.J. (2014). Epigenomic comparison reveals activation of "seed" enhancers during transition from naive to primed pluripotency. Cell Stem Cell *14*, 854–863.

Ginsburg, M., Snow, M.H., and McLaren, A. (1990). Primordial germ cells in the mouse embryo during gastrulation. Development *110*, 521–528.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., and van Steensel, B. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948–951.

Gurdon, J.B. (1962). Adult frogs derived from the nuclei of single somatic cells. Dev. Biol. *4*, 256–273.

Hackett, J.A., Reddington, J.P., Nestor, C.E., Dunican, D.S., Branco, M.R., Reichmann, J., Reik, W., Surani, M.A., Adams, I.R., and Meehan, R.R. (2012). Promoter DNA methylation couples genome-defence mechanisms to epigenetic reprogramming in the mouse germline. Development *139*, 3623–3632.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell *146*, 519–532.

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

Hayashi-Takanaka, Y., Yamagata, K., Wakayama, T., Stasevich, T.J., Kainuma, T., Tsurimoto, T., Tachibana, M., Shinkai, Y., Kurumizaka, H., Nozaki, N., and Kimura, H. (2011). Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. Nucleic Acids Res. *39*, 6475–6488.

Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H., and Razin, A. (1992). Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. Genes Dev. *6*, 705–714.

Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M., and Saitou, M. (2013). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. EMBO J. *32*, 340–353.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Ono, Y., Uno, K.D., Yamada, R.G., Ueda, H.R., and Saitou, M. (2006). An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. Nucleic Acids Res. *34*, e42.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K., and Saitou, M. (2008). Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. Genes Dev. *22*, 1617–1635.

Lee, H.J., Hore, T.A., and Reik, W. (2014). Reprogramming the methylome: erasing memory and creating diversity. Cell Stem Cell *14*, 710–719.

Luo, R.X., Postigo, A.A., and Dean, D.C. (1998). Rb interacts with histone deacetylase to repress transcription. Cell *92*, 463–473.

Maatouk, D.M., Kellam, L.D., Mann, M.R., Lei, H., Li, E., Bartolomei, M.S., and Resnick, J.L. (2006). DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. Development *133*, 3411–3418.

Magnúsdóttir, E., Dietmann, S., Murakami, K., Günesdogan, U., Tang, F., Bao, S., Diamanti, E., Lao, K., Gottgens, B., and Azim Surani, M. (2013). A tripartite transcription factor network regulates primordial germ cell specification in mice. Nat. Cell Biol. *15*, 905–915.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343–349.

Monk, M., Boubelik, M., and Lehnert, S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. Development *99*, 371–382.

Nakaki, F., Hayashi, K., Ohta, H., Kurimoto, K., Yabuta, Y., and Saitou, M. (2013). Induction of mouse germ-cell fate by transcription factors in vitro. Nature *501*, 222–226.

Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S.C., Obukhanych, T., Nussenzweig, M., Tarakhovsky, A., et al. (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. Nature *436*, 207–213.

Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S.W., Solovei, I., Brugman, W., Gräf, S., Flicek, P., Kerkhoven, R.M., van Lohuizen, M., et al. (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. Mol. Cell *38*, 603–613.

Runyan, C., Schaible, K., Molyneaux, K., Wang, Z., Levin, L., and Wylie, C. (2006). Steel factor controls midline cell death of primordial germ cells and is essential for their normal proliferation and migration. Development *133*, 4861–4869.

Sachs, M., Onodera, C., Blaschke, K., Ebata, K.T., Song, J.S., and Ramalho-Santos, M. (2013). Bivalent chromatin marks developmental regulatory genes in the mouse embryonic germline in vivo. Cell Rep. *3*, 1777–1784.

Saitou, M., Barton, S.C., and Surani, M.A. (2002). A molecular programme for the specification of germ cell fate in mice. Nature *418*, 293–300.

Saitou, M., Kagiwada, S., and Kurimoto, K. (2012). Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. Development *139*, 15–31.

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

Seki, Y., Hayashi, K., Itoh, K., Mizugaki, M., Saitou, M., and Matsui, Y. (2005). Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. Dev. Biol. *278*, 440–458.

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

Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H., and Shinkai, Y. (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. EMBO J. *27*, 2681–2690.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Voigt, P., Tee, W.W., and Reinberg, D. (2013). A double take on bivalent promoters. Genes Dev. *27*, 1318–1338.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat. Genet. *39*, 457–466.

Yokobayashi, S., Liang, C.Y., Kohler, H., Nestorov, P., Liu, Z., Vidal, M., van Lohuizen, M., Roloff, T.C., and Peters, A.H. (2013). PRC1 coordinates timing of sexual differentiation of female primordial germ cells. Nature *495*, 236–240.