Mapping developmental paths of monkey primordial germ-like cells differentiation from pluripotent stem cells by single cell ribonucleic acid sequencing analysis†

Authors: Zhang, Puyao, Xue, Sengren, Guo, Rongrong, Liu, Jian, Bai, Bing, et al.

OXFORD

Mapping developmental paths of monkey primordial germ-like cells differentiation from pluripotent stem cells by single cell ribonucleic acid sequencing analysis[†]

Puyao Zhang[1], Sengren Xue[3], Rongrong Guo[2], Jian Liu[2], Bing Bai[2], Dexuan Li[2], Ahjol Hyraht[2], Nianqin Sun[2], Honglian Shao[2], Yong Fan[4], Weizhi Ji[2], Shihua Yang[3,*], Yang Yu[1,*] and Tao Tan[2,*]

[1]Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Beijing Key Laboratory of Reproductive Endocrinology and Assisted Reproductive Technology and Key Laboratory of Assisted Reproduction, Ministry of Education, Peking University Third Hospital, Beijing, China
[2]State Key Laboratory of Primate Biomedical Research, Institute of Primate Translational Medicine, Kunming University of Science and Technology, Kunming, China
[3]College of Veterinary Medicine, South China Agricultural University, Guangzhou, China
[4]Department of Gynecology and Obstetrics, Key Laboratory for Major Obstetric Diseases of Guangdong Province, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

*Correspondence: Tao Tan, Tel: +86 871 65952802; E-mail: tant@lpbr.cn; Yang Yu, Tel: +86 010 82266699; Email: yuyang5012@hotmail.com; Shihua Yang, Tel: +86 020 85284899; E-mail: yangsh@scau.edu.cn

Authors Puyao Zhang, Sengren Xue, Rongrong Guo, Jian Liu, and Bing Bai contributed equally to this work.

Abstract

The induction of primordial germ-like cells (PGCLCs) from pluripotent stem cells (PSCs) provides a powerful system to study the cellular and molecular mechanisms underlying germline specification, which are difficult to study in vivo. The studies reveal the existence of a species-specific mechanism underlying PGCLCs between humans and mice, highlighting the necessity to study regulatory networks in more species, especially in primates. Harnessing the power of single-cell RNA sequencing (scRNA-seq) analysis, the detailed trajectory of human PGCLCs specification in vitro has been achieved. However, the study of nonhuman primates is still needed. Here, we applied an embryoid body (EB) differentiation system to induce PGCLCs specification from cynomolgus monkey male and female PSCs, and then performed high throughput scRNA-seq analysis of approximately 40 000 PSCs and cells within EBs. We found that EBs provided a niche for PGCLCs differentiation by secreting growth factors critical for PGCLC specification, such as bone morphogenetic protein 2 (BMP2), BMP4, and Wnt Family Member 3. Moreover, the developmental trajectory of PGCLCs was reconstituted, and gene expression dynamics were revealed. Our study outlines the roadmap of PGCLC specification from PSCs and provides insights that will improve the differentiation efficiency of PGCLCs from PSCs.
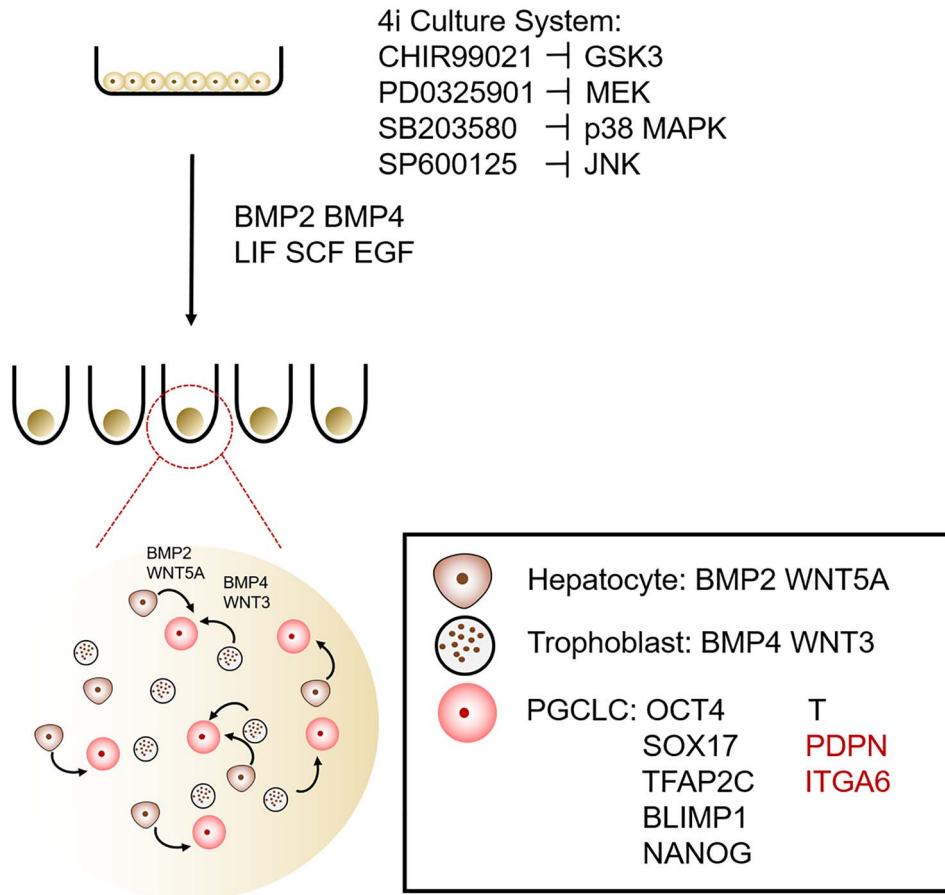
Summary Sentence
This study outlines the roadmap of monkey primordial germ-like cells (PGCLCs) specification from pluripotent stem cells by single cell ribonucleic acid sequencing analysis and provides insights for improving the differentiation efficiency of PGCLCs.

Graphical Abstract

Dr. Tao Tan is Professor State Key Laboratory of Primate Biomedical Research, Institute of Primate Translational Medicine, Kunming University of Science and Technology. Dr. Tan received his PhD from Kunming Institute of Zoology (KIZ) in 2010. Dr. Tan works on early embryonic development of non-human primates. Dr. Tan has so far published >30 peer-reviewed research articles with >2,000 citations (as of April 2022). Dr. Tan has received numerous academic awards, including the 2020 Chinese Society for Stem Cell Research (CSSCR) Young Investigator Award, the 2022 Chinese Society for Cell Biology (CSCB) Young Investigator Award. For more information, please visit http://www.lpbr.cn/2.

mental and molecular programs that include chromatin modifications, global erasure of DNA methylation, and re-establishment of sex-specific DNA methylation, eventually forming either spermatozoa or oocytes. Then, the genetic and epigenetic information of the parents is transmitted to the next generations [1, 2]. Thus, a precise understanding of the mechanisms underlying PGCs' specification, and the generation of functional germ cells are key goals of reproductive biology and regenerative medicine.

To date, most of the knowledge regarding mammalian PGC specifications has been obtained from mouse studies. These studies have provided fundamental knowledge about PGC development, such as that bone morphogenetic protein 4 (BMP4) from extra-embryonic ectoderm induces *BLIMP1* (also known as *PRDM1*) expression and that PRDM14 in the epiblast initiates the PGC program [3, 4]. However, other evidence has indicated the existence of species-specific divergences in PGC specification between humans and mice; for example, murine PGCs express SOX2, whereas human PGCs express SOX15 and SOX17 [5]. Thus, studies on PGC specification in different species are needed, especially in primates.

To investigate the mechanisms underlying human PGC specification, several approaches have been developed to induce primordial germ-like cells (PGCLCs) from human pluripotent stem cells (PSCs); these approaches involve two steps: first is a preinduction of PSCs into an intermediate state, such as incipient mesoderm-like cells (iMeLCs), and then

Introduction

Primordial germ cells (PGCs) are the precursors of gametes (sperm and oocyte). PGCs undergo multistep develop-

these intermediates are further differentiated into PGCLCs in embryoid body (EB) [6–9]. These pioneering studies identified a spectrum of key transcription factors (TFs) that are involved in human PGCLCs specification, such as SOX17, TFAP2C, BLIMP1, T, MIXL1, and EOMES. Despite these advances, there remain several key unsolved questions, including whether an EB provides a niche that is comparable to that of the in vivo situation to induce PGCLC differentiation and whether the population of PGCLCs is homogeneous or heterogeneous.

To answer these questions, we combined high through-put single-cell RNA sequencing (scRNA-seq) analysis with a modified two-step PGCLCs induction method to study the niche within EBs and the developmental trajectory during PGCLC specification. We showed that cells within EBs formed a niche that facilitated the specification of PGCLCs by secreting growth factors that are indispensable for PGC specification in vivo. We found that in addition to PGCLCs, the mesoderm (mesenchymal cell) and endoderm lineages are the other two major alternative cell types within EBs. We also confirmed that the population of PGCLCs is heterogeneous and was composed of PGCLCs at different developmental stages. Taken together, we have delineated the roadmap of PGCLC specification in vitro and have offered insight into optimizing the differentiation system for functional PGCLCs from PSCs.

Materials and methods

Cell culture

Cynomolgus monkey female (XX) and male (XY) embryonic stem cells (ESCs) were cultured on mitomycin C-inactivated mouse embryonic fibroblasts (MEFs) as previously described [10]. Briefly, ESC colonies were cultured in Nutrient Mixture F-12 (DMEM/F12, Invitrogen) medium containing 20% KOSR (Invitrogen) supplemented with 4 ng/mL bFGF (Chemicon), 1% nonessential amino acids (Invitrogen), 0.1 mM β-mercaptoethanol (Invitrogen), and 1% penicillin–streptomycin-L-glutamate (Invitrogen). For 4i-state conversion, ESC colonies were dissociated by collagenase IV and reseeded on MEFs with a 4i conversion medium as previously reported [11]. The 4i conversion medium comprised ESC culture medium supplemented with 10 ng/mL human LIF (PeproTech), CHIR99021(3 μM) (Tocris), PD0325901 (0.5 μM) (R&D Systems), SP600125 (10 μM) (Selleck), and SB203580 (10 μM) (Selleck). The medium was changed daily. All chemicals were from Sigma-Aldrich unless otherwise stated.

Monkey PGCLCs differentiation

To induce PGCLC differentiation, 4i-state cells were dissociated with TrypLE Express (GIBCO), then the cells were seeded into ultra-low cell attachment U-bottom 96-well plates (Corning) at a density of 3000 cells/well in 200 μL PGCLC differentiation medium containing Glasgow's MEM (GMEM, GIBCO), 15% KOSR, 0.1 mM nonessential amino acids, 0.1 mM β-mercaptoethanol, 1% penicillin–streptomycin-L-glutamate, 1 mM sodium pyruvate, 250 ng/mL BMP4 (R&D Systems), 250 ng/mL BMP2 (PeproTech), 1 μg/mL human LIF (PeproTech), 100 ng/mL SCF (PeproTech), 50 ng/mL EGF (PeproTech), and 10 μM ROCK inhibitor, Y-27632 (Selleck). The medium was changed daily.

Immunofluorescence

EBs were harvested on day 2 and day 4 and fixed in 4% paraformaldehyde for 30 min at 25°C. Then the EBs were rinsed thrice with phosphate-buffered saline (PBS) and dehydrated with sucrose solutions of increasing concentrations of 15%, 20%, and 30% (wt/v) each for 2 h. Next, EBs were embedded in optimal cutting temperature compound (Tissue-Tek OCT, Sakura Finetek), frozen, and stored at −80°C. Before immunofluorescence staining, samples were prepared as cryosections with an 8-μm thickness on pretreated glass slides (CITOGLAS) and air-dried for 1 h. Subsequently, the slides were washed with PBS thrice for 5 min each. Then, the slides were permeabilized in PBST (PBS with 0.2% Triton X-100) for 30 min, washed with PBS thrice for 5 min each, and blocked in 3% (w/v) bovine serum albumin in PBS for 1 h. Slides were then incubated with primary antibodies overnight at 4°C. Subsequently, the slides were washed in PBST thrice for 10 min each. Then, Alexa Fluorophore (488, 594, and/or 647)-conjugated secondary antibodies were incubated with the slides in the dark at 25°C for at least 1 h. Images were taken using a Leica TCS SP8 confocal microscope (Leica). After wash, fluorescence-conjugated secondary antibodies, and 4′,6-diamidino-2-phenylindole (DAPI) were incubated with the slides in dark at 25°C for 2 h. Images were taken using Leica TCS SP8 confocal microscope (Leica). In vitro cultured cynomolgus monkey embryos [12] and 4i condition or MEF cells were included as positive or negative controls, respectively. For antibodies used, see Table S1.

Quantification of the proportion of PGCLCs within EB during differentiation

Cryosections from day 2 (three EBs) and day 4 EBs (three EBs) were stained with BLIMP1 TFAP2C, OCT4, and T antibodies. Further, 500 cells within day 2 or 4 cryosections were randomly selected and the proportion of BLIMP1, TFAP2C and double-positive cells or OCT4, T, and double-positive cells were calculated. The experiments were repeated thrice.

Quantification of the Ki67-positive PGCLCs and 5hmc levels of PGCLCs

For quantification of the Ki67 positive PGCLC cells, cells cultured in 4i condition and cryosections from day 2 (three EBs) and day 4 EBs (three EBs) were stained with BLIMP1, TFAP2C, and Ki67 antibodies. Further, 500 cells within 4i condition day 2, and day 4 cryosections were randomly selected and the proportion of SOX17, TFAP2C, Ki67, and triple-positive cells were calculated. Notably, the proportion of Ki67 of cells in the 4i condition was 100% although SOX17 and TFAP2C were not expressed. The experiments were repeated thrice.

For quantification of the 5hmc levels of BLIMP1/SOX17 double-positive PGCLCs, the mean fluorescence intensity (MFI) values were collected using the mask channel option on 5hmC and DAPI channels. Relative intensity values for individual nuclei were calculated by dividing the MFI of 5hmC by the MFI of DAPI.

Quantitative real-time polymerase chain reaction

Total RNA was extracted using TRIzol reagent (Invitrogen). Complementary DNA (cDNA) synthesis was performed using 1 μg of total RNA with HiScript II 1st Strand cDNA Synthesis Kit (Vazyme). Quantitative polymerase chain reaction (qPCR)

Table 1. Primers used for qPCR

Genes	Sense (5′ → 3′)	Anti-sense (5′ → 3′)
ß-ACTIN	CAATAGACAGGAGCGCTATCCG	GCAGGGTTTAAAGACGATGTCAC
PPIA	CTTTACATCCTGAGCTGGGAAGC	TCACATCCCAGCATTATCTCCTT
SOX15	CTCTCCCACTCCATACAATCCTC	GTAGGGGATGCTGCTGGATTAAA
SOX17	GAACAGTATCTGCACTTCGTGTG	AGTAATATACTGCGGAGCTGGC
TFAP2C	CACCATGAGCTTCCAGTTCCTA	GATATTCGGCGACTCCAGTGTG
BRACHYURY (*T*)	CTTGATGCAAAGGAAAGAAGCGA	TGAGGATTTGTAGGTGGACACAG
BLIMP1/PRDM1	CAAAAGAAACATGACCGGCTACA	CTTGTGGCAGACCTTCAGATTGG
PRDM14	GAACTGGGAGCGGAGAGTAAGG	CACACCTTGTCCTGAGTTGCG
OCT4	AGCAGTGACTATGCACAACGA	ACAGTGCAGTGAAGTGAGGG
DDX4/VASA	CAAACAGCGGAGAGAACTTGAA	TCCATTTTCTCCAGAAGAATACCT
cKIT/KIT	ACAATGGCACGGTTGAATGTAAG	CAATCAGCAAAGGAGTGAACAGG

analysis was performed according to the manufacturer's instructions and monitored with SYBR green PCR master mix [FastStart Universal SYBR Green Master (Rox); Roche] using the Applied Biosystems QuantStudio 7 Real-Time PCR Detection System (Thermo Fisher). *PPIA/ß-ACTIN* was used as a control for normalization. Fold-change in mRNA expression level was normalized to the cycle threshold using naïve cells. The primers used are listed in Table 1.

The scRNA-seq library preparation and sequencing

The scRNA-seq libraries were generated using the Chromium Single Cell 3′Reagent Kit (10× Genomics). Briefly, a single-cell suspension (500 cells/μL in PBS) was mixed with real-time quantitative (RT-PCR) master mix and loaded together with Single Cell 3′Gel Beads and Partitioning Oil into a Single Cell 3′Chip (10× Genomics) according to the manufacturer's instructions. RNA transcripts from single cells were uniquely barcoded and reverse transcribed within droplets. cDNA molecules were preamplified and pooled followed by library construction. All libraries were quantified on a 2100 Bioanalyzer (Agilent) with RT-PCR and then subjected to 150-bp paired-end sequencing on an Illumina Xten platform (sequenced by Novogene).

The scRNA-seq data processing

Sequencing data were first processed using the 10× Genomics Cell Ranger pipeline (https://www.10xgenomics.com) and sequences were aligned using STAR [13] against the *Macaca fascicularis* reference genome (ensemble release-91). Subsequently, unique molecular identifier (UMI) matrices from each sample were merged in R and imported to Seurat [14]. Cells with less than 700 detected genes and more than 4000 genes, or less than 700 UMIs and more than 30 000 UMIs were excluded from further analysis, as well as cells with more than 10% of UMIs mapped to mitochondrial genes. Next, the raw UMI matrix was normalized by LogNormalize and scaled by ScaleData. Highly variable genes were identified from these cells using FindVariableGenes with the following parameters: x.low.cutoff = 0.0125, x.high.cutoff = 3, and y.cutoff = 0.5.

Principal component analysis was then performed, and the top 10 principal components were used for *t*-distributed stochastic neighbor embedding (*t*-SNE) and unsupervised clustering. Clusters with fewer than 200 cells were omitted from further analysis. Differentially expressed genes of each cluster were produced by FindAllMarkers. Cell type identifications were based on known genetic markers. Gene Ontology (GO) analysis was performed with ClusterProfiler [15], and GO terms with *P*-value less than 0.05 were considered statistically significant.

Single-cell-pseudotime analysis

For pseudotime analysis, the UMI matrix calculated in Seurat was passed into Monocle2 [16], a computational method that performs lineage trajectory reconstruction based on scRNA-seq data, and highly variable genes identified by Seurat were used as ordering genes for trajectory calculation. After trajectory reconstruction, we used BEAM to identify genes with branch-dependent expression and used plot_multiple_branches_heatmap to visualize these genes along the branches.

Statistical analysis

The results were presented as mean ± standard error of the mean. Significance was assessed using unpaired *t*-tests (two-tailed); *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$; ns, nonsignificant. Statistical analyses are described in detail in the figure legends for each panel.

Data availability

All sequencing data were deposited at the NCBI GEO under accession number GSE139432.

Results
PGCLC differentiation from male and female monkey ESCs

We detected expression of OTX2 (Figure S1A), a marker of human formative PSCs [17, 18] in monkey PSCs cultured in a previously reported 4i-based medium and cells showed higher correlation coefficients to human formative PSCs [17, 18] than naïve PSCs [19] (Figure S1B), indicating the capability of PGCLC specification, male and female monkey ESCs were cultured in 4i conversion medium for at least two passages as previously reported before PGCLC induction (9). The cells cultured in the 4i conversion state formed dome-shaped colonies (Figure 1A and Figure S1C) and expressed pluripotent markers, such as OCT4, SOX2, and NANOG. These cells also expressed primate-specific PSC surface markers, such as TRA-1-81 (Figure 1B and Figure S1D). Then, 3000 4i-state ESCs were used to form EBs in low-attachment 96-well plates for 4 days. Cell aggregates with clear boundaries could be observed on day 2, and EBs had gradually grown up on day 4 with increased diameters (Figure S1E). RT-qPCR analysis confirmed PGCLC differentiation. As shown in Figure 1D, after 2 days of induction, EBs began to express conserved and primate-specific PGC markers, such as PRDM1 (also known as BLIMP1), TFAP2C, KIT, DDX4 (also known as VASA), and SOX17. On day 4, the expression levels of these PGC markers peaked, although levels of SOX17 and
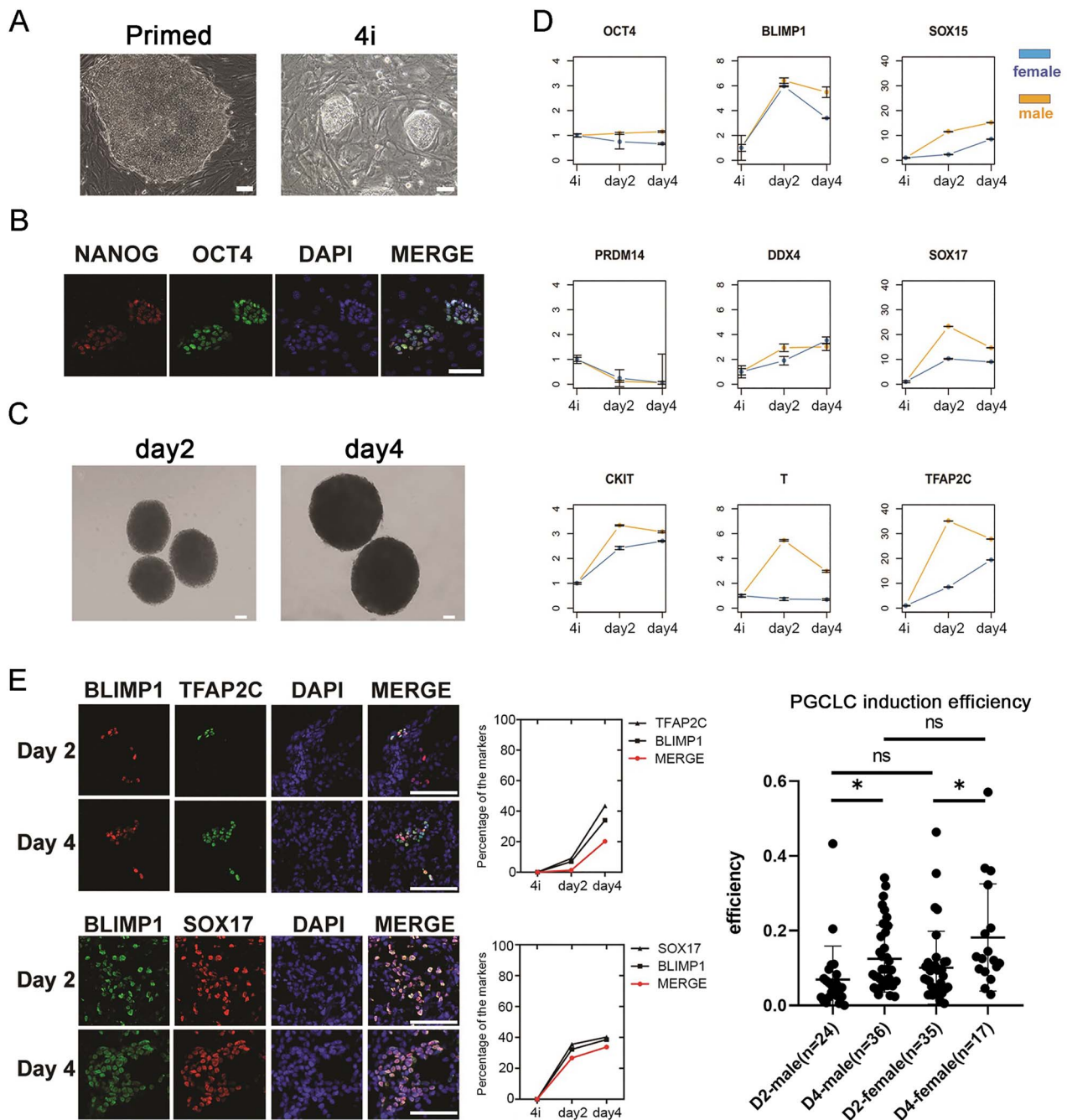
Figure 1. PGCLC differentiation of naïve-state monkey ESCs. (A) Representative bright-field images showing colony morphologies of female primed and 4i-state ESCs. Scale bar = 50 μm. **(B)** Representative immunofluorescence images of female 4i-state ESCs stained with anti-OCT4 and anti-NANOG antibodies. Scale bar = 50 μm. All experiments were performed in triplicates, unless otherwise stated. **(C)** Development of female EBs on days 2 and 4. Scale bar = 50 μm. All experiments were performed in triplicates, unless otherwise stated. **(D)** Dynamics of the expression of key genes during PGCLC induction by real-time PCR. Expression is normalized to ß-ACTIN or PPIA. Fold change was calculated relative to the expression levels of each gene in 4i-state cells, which was given a value of 1.0. All experiments were performed in triplicates, unless otherwise stated. **(E)** Immunostaining of BLIMP1/TFAP2C and BLIMP1/SOX17 expression in days 2 and 4 female EBs (left panel). Percentage of BLIMP1/TFAP2C and BLIMP1/SOX17 cells in day 2 and day 4 female EBs (middle panel). Scale bar = 50 μm. Percentage of SOX17/BLIMP1/TFAP2C triple-positive PGCLCs within days 2 and 4 EBs (right panel, shown is mean and SEM). ns, not significant; *$P < 0.05$. n, independent experiments, at least four cell lines were used.

PRDM1 were slightly decreased, whereas OCT4 expression was relatively stable throughout differentiation (Figure 1D). Immunofluorescence staining of PGC markers showed that the ratios of BLIMP1/TFAP2C and SOX17/BLIMP1 double-positive cells increased during the 4-day differentiation [male: 1.29% ($n = 620$) to 20.21% ($n = 876$) and 26.71% ($n = 307$) to 33.80% ($n = 639$), female: 6.57% ($n = 624$) to 15.44% ($n = 1036$) and 25.59% ($n = 758$) to 31.38% ($n = 835$), respectively] (Figure 1E and Figure S2A). Furthermore, the positive ratios of other PGC makers, such as OCT4 (POU5F1) and T also increased during the differentiation of male and female cells (Figure S2B). In addition, we also found a significant increase in 5-hydroxymethylcytosine (5hmC) in monkey PGCLCs (mPGCLCs) during differentiation, which is

consistent with a study in humans (Figure S2C). The PGCLCs induction in our modified approach was reproducible as the percentage of SOX17/BLIMP1/TFAP2C triple-positive PGCLCs increased from day 2 to day 4 within male and female EBs (average 6.9% in day 2 male EBs, 12.49% in day 4 EBs, 10.06% in day 2 female EBs, and 18.17% in day 4 female EBs) (Figure 1E). Taken together, these data indicated that PGCLCs were induced from monkey PSCs after induction in a 4i medium.

Dissecting the cell types within EBs during PGCLC differentiation

As cell types within EBs during PGCLC specification from PSCs are elusive, we next examined the cell types within EBs. The 4i-state cells and EBs (96 EBs from day 2 and 96 EBs from day 4) were dissociated into single cells and subjected to high throughput droplet-based scRNA-seq analysis as in previous reports [20, 21]. After quality filtration, we obtained transcriptomic profiles of 38 193 cells from six samples (female and male 4i-state cells, and day 2 and day 4 EBs (Table S2). The average number of cells detected was 6478 (range from 4200 to 9005), and genes detected were 2442 (range from 746 to 3997), with a mean coverage of 10 472 reads (UMI) per cell. Also, the average percentage of reads mapping to mitochondrial genes (percent. mito) was 2.3% (range from 0% to 9.9%), suggesting our cells are with high cell viability (Figure S3A).

Then, the *t*-SNE analysis was performed and eight clusters of cells were identified (Figure 2A). Based on the expression of marker genes, the identity and percentage of each cluster were investigated (Figure 2B and Figure S3B). Cluster 1 was a hepatocyte with an expression of AFP and F2, two well-known markers of hepatocytes [22–24]. Cluster 2 was trophoblast, as cells in cluster 2 expressed trophoblast markers, such as KRT7 and WFDC2 [25]. Cluster 3 was identified as 4i cells according to the expression of the pluripotent markers such as SOX2 and OCT4 (POU5F1) [26]. Cluster 4 was differentiating (Differ.) cells within the 4i population, as they expressed the differentiated cell markers, such as STMN2 and FOXA2 [27, 28]. Interestingly, we identified two types of PGCLCs (cluster 5 and cluster 6): PGC1 (cluster 5) and PGC2 (cluster 6). Both of these cell types expressed PGC markers, such as TFAP2C and POU5F1 (OCT4) [6, 8], whereas PGC2 cells expressed genes involved in epithelial-to-mesenchymal transition (EMT), such as MMP9, SNAI2 (also known as SLUG), and TGFβ1 [29–31]. Cluster 7 was identified as neural progenitor cells (NPCs), as they expressed SOX2 and PCSK1N, the two well-known NPCs makers [32]. Finally, the last identified cell type (cluster 8) was mesenchymal cells (cluster 8), which expressed COL3A1 and COL1A1 [33]. In addition, the levels of previously reported G1/S- and G2/M-specific genes [34, 35] were analyzed in cells within EBs. Interestingly, PGCLCs were relatively more quiescent than the 4i-state cells, and there was no difference between PGC1 and PGC2 cells (Figure S3C).

To verify cell identities, the differentially expressed genes of each cluster of cells were further identified and GO analysis was performed (Figure 2C and Table S3). Genes involved in cellular homeostasis and blood coagulation were enriched in hepatocytes (cluster 1). Trophoblast cells (cluster 2) were enriched for specific genes involved in tissue morphogenesis, etc. The 4i-state cells (cluster 3) specific genes were enriched in GO terms such as RNA processing process. Differ. cells (cluster 4) expressed genes related to tissue morphogenesis, epithelium development, and embryonic morphogenesis. PGCLCs (clusters 5 and 6) expressed genes involved in cellular modified amino acid metabolic process, mitochondrial membrane organization, cell migration, etc. Genes belonging to the nervous system, head, and tissue development were enriched in NPCs (cluster 7). Finally, mesenchymal cells (cluster 8) expressed genes enriched in GO terms for tissue development, stem cell differentiation, and embryo development. Overall, our findings revealed the cellular heterogeneity within EBs and suggested the trophoblast, endoderm (hepatocyte, cluster 1), mesoderm (mesenchymal cells, cluster 8), and PGCLC lineage were the major cell types in the differentiation system.
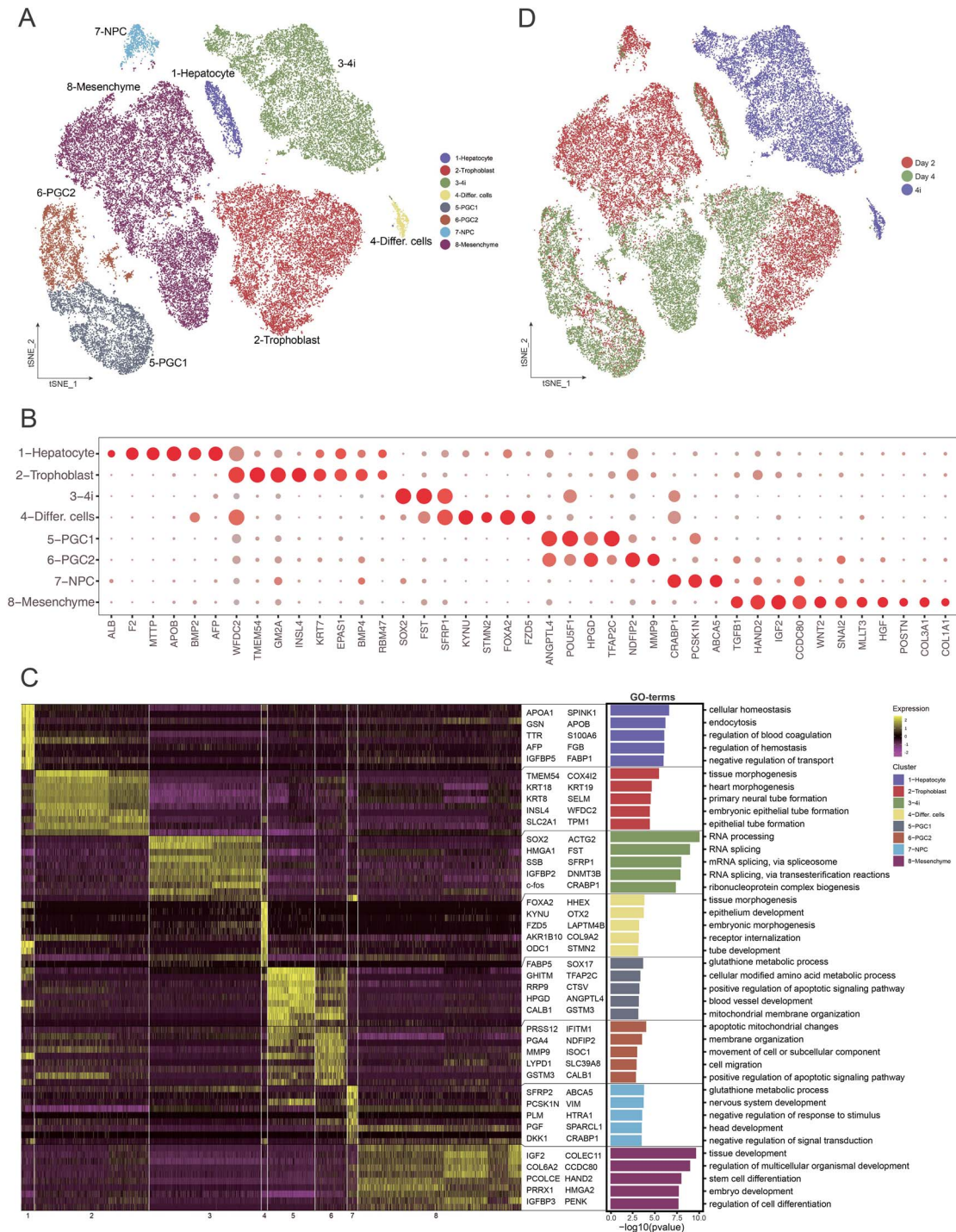
EBs as a niche for monkey PGCLC specification

As discussed earlier, we identified eight clusters of cells during PGCLC differentiation. To reconstruct the developmental trajectory of all cell types, pseudotime trajectory analysis was applied following a previously described method [36] (Figure 3A). We found that all cells formed a Y-like trajectory: 4i-state PSCs occupied the one branch of the pseudotime trajectory, and endoderm (hepatocyte) and mesoderm (mesenchymal cell) lineages were the right and left branches, respectively (Figure 3B). We also found that PGC1 cells were located with 4i-state PSCs in pseudotime trajectory, whereas part of the PGC2 cells was located in the mesoderm lineage branch (Figure 3B), suggesting PGC1 and PGC2 might be early- and late-stage PGCs, respectively. Expression patterns of early- and late-stage PGC marker genes [37] also confirmed this speculation (Figure 3C). As a small portion of 4i cells (named 4i-overlap) were located with PGC6, we investigated the differentially expressed genes between PGC6 and 4i-overlap cells. As small numbers of 4i cells (named 4i-overlap) were located with PGC2, we investigated the differentially expressed genes between PGC6 and 4i-overlap cells. We observed that 80 genes were upregulated in 4i-overlap cells, such as EEF1A1, HMGA1, and EEF1G, suggesting rapid proliferation of these cells (Figure S3D). Collectively, these findings imply that although small numbers of 4i cells overlap with PGC2 in pseudotime trajectory, they display distinct transcriptome profiles.

We next determined the gene expression dynamics along the site pseudotime trajectory. As shown in Figure 3D, five clusters of gene expression patterns (here, "here" referred to as cluster 1–5) from prebranch (4i-state PSCs and PGC1) to the endoderm (hepatocyte) or mesoderm (mesenchymal cell and part of PCG2 cells) lineages were observed. Genes enriched in clusters 1 and 2 were highly expressed in the endoderm lineage, and genes belonging to clusters 3 and 5 were specifically expressed in prebranch cells (4i-state PSCs and PGC1) and the mesoderm lineage. Finally, cluster 4 represented genes enriched in the mesoderm lineage. We then performed GO analysis to reveal the functional enrichment of these clusters. For example, mRNA catabolic process (clusters 3 and 5), and ossification were enriched in prebranch, and genes belonging to lipid catabolic process, cell cycle regulation, and response to nutrient levels were enriched in other branches (clusters 1, 2, and 4) (Figure S3E). Collectively, these findings revealed the unique developmental trajectory of cells within EBs.

As extra-embryonic mesenchyme (EXMC), amnion, and cytotrophoblasts cells formed a niche to induce PGC

Figure 2. Single cell transcriptomic profiles and cell type identification. (A) *t*-SNE plots of male and female 4i-state cells, days 2 and 4 EB cells. Cells are colored by identified cell types. (B) A bubble plot showing expression of cluster-specific genes. The color of the dot represents the average expression and the size represents the percentage of cells expressing that gene. (C) Heatmaps showing the differential gene expression patterns of each cluster. Representative genes and key gene ontology enrichments are shown. (D) *t*-SNE plots of male and female 4i-state cells, day 2 and day 4 EB cells. Cells are colored by differentiation time points (days 2 and 4).

specification from amnion in vivo through secretion of BMP4, BMP2, and WNT3A [28], we wondered whether EBs also provided a niche for PGCLC specification as in vivo. We determined gene expression levels of the classical signaling pathways (NOTCH, WNT, TGF, and BMP) within EBs (Figure 3E). We found that WNT2, TGFβ1, and TGFβ2 were highly expressed in mesenchymal cells, hepatocytes expressed BMP2 and WNT5A, and trophoblast cells expressed BMP4

and WNT3. Moreover, the expression of downstream genes of the WNT and TGFβ superfamilies were upregulated in PGC1 and PGC2 cells, for example, RUVBL1, CTBP2, CSNK2A1, and ID1–4 (Figure 3E). The expression of BMP2, BMP4, WNT3A, and WNT5A in day 4 EBs was validated by immunostaining (Figure S3F).

Next, the expression of hepatocyte (HNF4A), trophoblast (CK7), and PGCLC (SOX17) markers within EB was
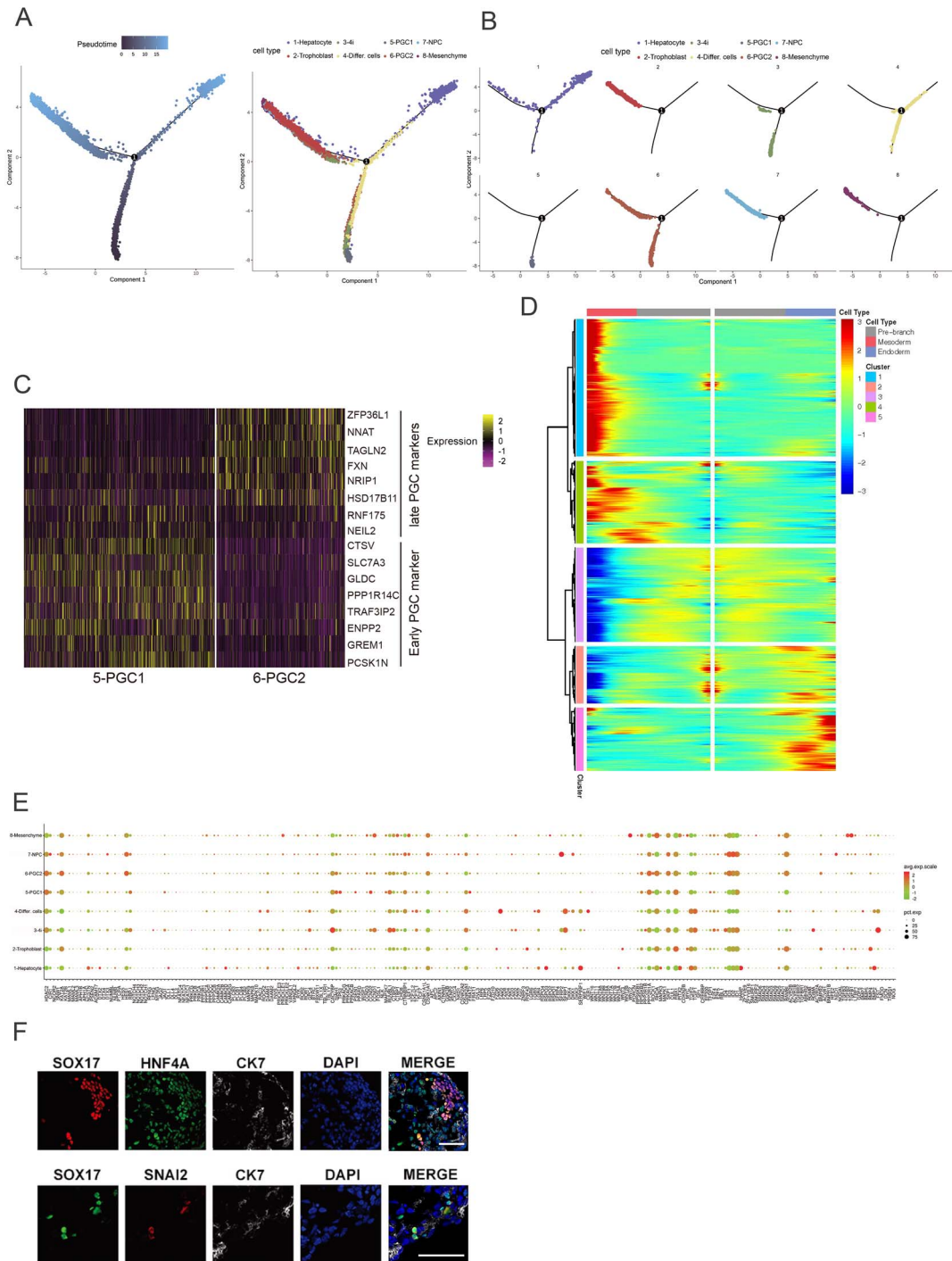
Figure 3. Reconstruction of the niche for PGCLC induction. (A) Pseudotime analysis of the differentiation trajectory and cell types was plotted in the pseudotime trajectory (B). (C) Expression of early- and late-stage PGC markers in PGC1 and PGC2. The color key from purple to yellow indicates low to high expression, respectively. (D) Heatmap showing the gene expression dynamics during PGCLC differentiation. Genes are clustered according to expression patterns, and cell types are colored. (E) Dot plots showing expression of members of the WNT, TGFβ, and FGF signaling pathways in identified cell types. (F) Representative immunofluorescence images of EB cryosections stained with anti-SOX17, anti-HNF4A, anti-CK7, and anti-SNAI2 antibodies. Arrow indicates SOX17- and HNF4A-positive cells were enclosed by trophoblast cells (CK7 positive) and arrowhead indicates SNAI2 and SOX17 double-positive cells. Scale bar= 50 μm. All experiments were performed in triplicates unless otherwise stated.

detected by immunofluorescence staining. We observed that on day 4 the majority of SOX17- and HNF4A-positive cells were located on the outer surface of EBs and were enclosed by trophoblast cells (CK7-positive) (Figure 3F). Moreover, SNAI2 expression was also confirmed in part of the SOX17-positive population (Figure 3F),

consistent with our transcriptomics data that PGC2 had migratory characteristics and expressed EMT genes. Taken together, these findings indicated that cells within EBs formed an autocrine niche to induce PGCLC differentiation by secreting growth factors important for PGCLC specification.
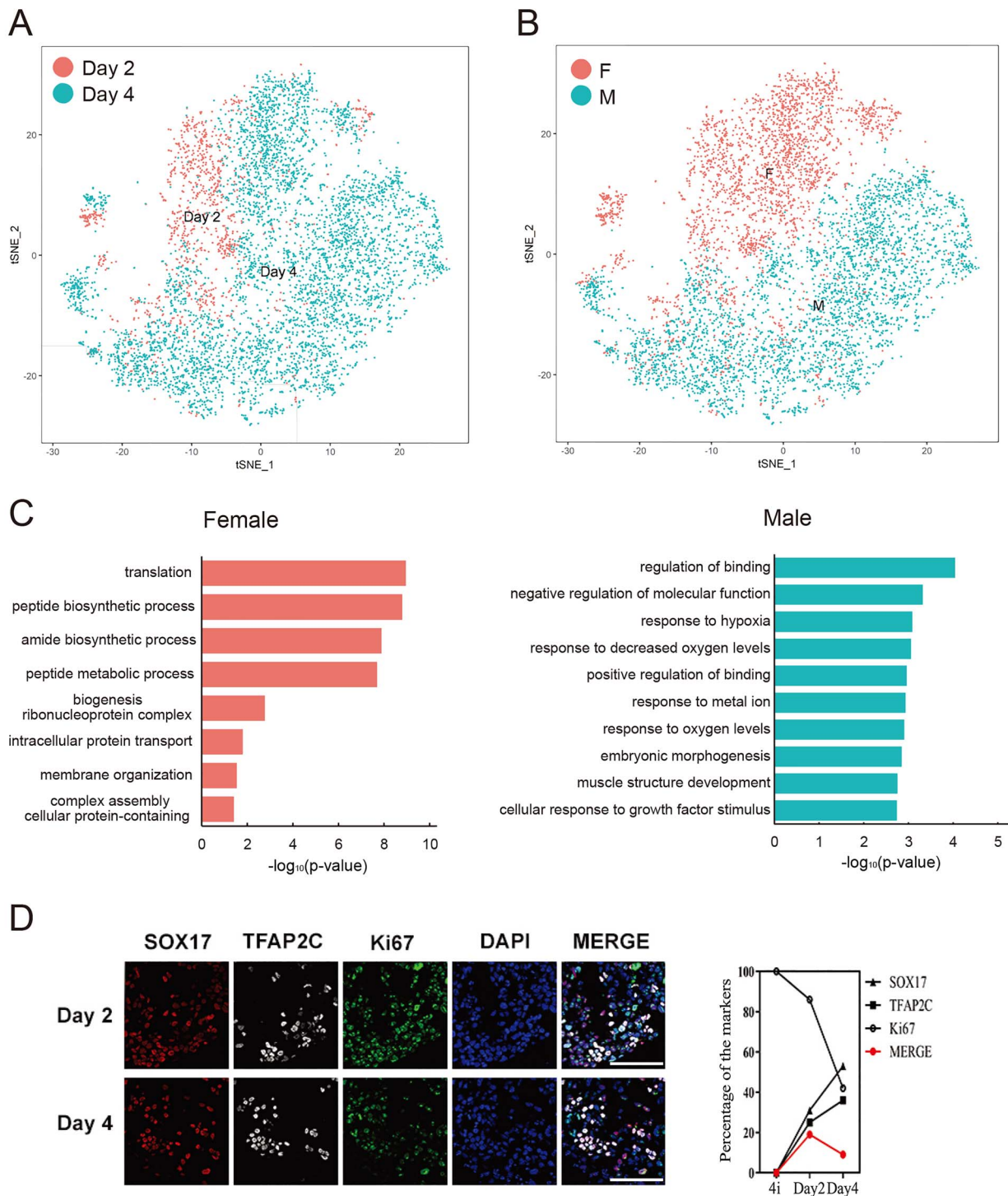
Figure 4. Transcriptional differences between male and female PGCLCs. (A) *t*-SNE plots of PGCLCs. Cells are colored by differentiation time points (days 2 and 4). (B) *t*-SNE plots of PGCLCs. Cells are colored by cell types (M, male or F, female). (C) GO term enrichment analysis of genes enriched in male and female PGCLCs. (D) Quantification of cells (n = 1000 cells) with expression of SOX17, TFAP2C, and Ki67, respectively (black line) or both (red line) in 4i-state and in EB cryosections. Scale bar= 50 μm. All experiments were performed in triplicates unless otherwise stated.

Transcriptional landscape of monkey male and female PGCLCs specification in vitro

We next determined the transcriptomic landscape of PGCLCs during in vitro differentiation. As the majority of PGCLCs emerged on day 4 (Figure 2D and Figure 4A), we wondered whether there were differences in the timing of PGCLC specification between male and female cells. As shown in Figure 4B, besides a portion of female germ cells emerging on day 2, the majority of male PGCLC differentiation occurred on day 4, suggesting that the differentiation rate of male cells might be slower than that of female cells. We then sought to determine the developmental stage (early or late) of PGCLCs.

We found that the majority of female PGCLCs were early-stage (PGC1) while male PGCLCs were both early- and late-stage (PGC1 and PGC2) (Figure S4A).

Next, to dissect the differences in gene expression patterns between male and female PGCLCs, differentially expressed genes were identified in male and female PGCLCs. GO analysis indicated that genes enriched in female PGCLCs were related to translation and peptide synthesis, while those in male PGCLCs were enriched in the GO terms regulation of binding and response to hypoxia (Figure 4C and Table S4), implying that male PGCLCs might be migrating, as hypoxia promotes the migration of murine germline stem cells [38]. As the ratio of PGCLCs decreased beyond day 4 of differentiation in both humans [6], we next determined whether the proliferation of PGCLCs was impaired during prolonged differentiation. As shown in Figure 4D, the percentage of Ki67-positive PGCLCs was decreased, suggesting mitotic arrest might be a reason for the loss of PGCLCs during differentiation.

Developmental trajectory of PGCLCs

We showed that the population and differentiation of PGCLCs were highly asynchronous and heterogeneous. To reconstitute the PGCLC developmental trajectory, we dissected subpopulations of PGCLCs using *t*-SNE analysis. Based on the *t*-SNE analysis, PGCLCs could be clearly grouped into six clusters: IGF2-expressing PGCs (cluster 1), HPGD-expressing PGCs (cluster 2), PDPN-and IGTA6 expressing PGCs (cluster 3), YWHAZ-expressing PGCs (cluster 4), CD24-expressing PGCs (cluster 5), and KRT18-expressing PGCs (cluster 6) (Figure 5A and B). Also, we verified that PDPN and IGTA6 were expressed in PGCLCs by immunostaining on day 4 EBs (Figure S4B). Pseudotime analysis was then used to investigate the cell fate transitions of PGCLCs. Further, PGCLCs formed a "U-like"structure, with cluster 5 (CD24-expressing PGCs) and cluster 6 (KRT18-expressing PGCs) occupying the beginning heads and other PGCLCs at the tail and sub-branch at the turn (Figure 5C). Thus, pseudotime analysis delineated the stepwise developmental path of PGCLC differentiation in vitro.

We next examined gene expression dynamics of all branches within the pseudotime trajectory (states 1–3; clusters 1–7). We found three types of gene expression patterns: genes enriched in prebranch (state 2; clusters 5 and 6) and other states (states 1 and 3; clusters 1–4, and 7) PGCLCs, respectively (Figure 5D). According to GO analysis, translation- and peptide biosynthetic-related genes were enriched in prebranches. Response to biotic stimulus, microtubule-based processes, and ribosome biogenesis-related genes were enriched in other branches (states 1 and 3) (Figure S4C).

To evaluate whether our monkey PGCLC induction system could be used to study primates (humans and monkey) conserved or monkey-specific mechanisms underlying PGC specification, we compared marker gene expression patterns between in vivo human PGCs (hPGCs) [35] and in vitro mPGCLCs. We found that expression patterns of a portion of PGC markers were conserved between hPGCs and mPGCLCs, such as TFAP2C and POU5F1, while the expression patterns of another portion of genes were species-specific, such as NANOS3 (Figure 5E).

Discussion

A strategy involving the first preinduction of human PSCs is applied to induce human PGCLCs specification [7, 9], whereas a previous study in cynomolgus monkey indicates that this preinduction is dispensable [39]. In this study, we observed that cynomolgus monkey PSCs cultured in 4i condition [11], with germline competence, of which transcriptomic profile is similar to human formative PSCs. This finding may suggest that there are species differences in germline competence between monkey and human PSCs, which is still worth pursuing.

In this study, we performed an scRNA-seq analysis of in vitro monkey PGCs, including 4i-state ESCs, and day 2 and day 4 EBs. By identifying cell types within EBs, we reconstituted the niche for PGCLC differentiation. In addition to PGCLCs, the endoderm (hepatocytes), mesoderm (mesenchymal cells), and trophoblastic lineages were induced during differentiation. Our scRNA-seq analysis revealed that hepatocytes secreted BMP2 and WNT5A, while trophoblast cells expressed BMP4 and WNT3 (Figure 3F). Interestingly, during PGC specification of the cynomolgus monkey (*M. fascicularis*) from the amnion in vivo, the amnion expressed BMP4 and WNT3A; cytotrophoblasts expressed WNT3A, and EXMC expressed BMP4 and BMP2, providing key signals for PGC specification [37]. Thus, we clearly identified the cellular composition of EBs during differentiation, which indicated that cells within EBs secreted cytokines that are potentially indispensable for PGC specification, similar to the in vivo situation [37]. Interestingly, we found mesenchymal cells expressed TGFβ1, which inhibited the proliferation of murine PGCs in vitro [40]. As PGCLC proliferation decreased during prolonged culture, it is worth testing whether inhibiting TGFβ signaling will promote PGCLC proliferation.

We found that in addition to germ cell fate, the mesoderm and endoderm lineages were induced in the current differentiation system. A study in human PSCs revealed that predifferentiation through premesendoderm intermediate activates mesendoderm enhancers and endows human PSCs with competence for PGCLCs differentiation [41] and further activation induces mesoderm and endoderm lineage specification, whereas downregulation of mesendodermal marker gene OTX2 promotes the PGC's fate. Moreover, the synergist interaction of SOX17 and TFAP2C guarantees the PGCLCs' differentiation [41]. Taken together, it is worthy to test whether the same mechanism is used to induce mPGCLCs and the inhibition of the mesoderm and endoderm cell fates using small molecules could further improve the frequency of PGCLC differentiation from primate PSCs.

First, differentiating through an iMELC intermediate promotes human PGCLCs' induction from primed human PSCs [7]. Mesoderm fate (mesenchyme) was also observed in our mPGCLCs' induction system. As iMELCs express nascent mesoderm or primitive streak marker genes (T, EOMES, SP5, and MIXL1), and mesenchyme (cluster 8) expressed mature mesoderm marker genes, such as SNAI 2 and MLLT3, we speculate that mesenchyme might directly differentiate from 4i-state PSCs instead of an intermediate in PGCLCs specification.

Using scRNA-seq analysis, we clearly indicated that BLIMP1 and TFAP2C double-positive PGCLCs are a heterogeneous population, comprising cells of different developmental stages. Recently, oogonia-like cells were
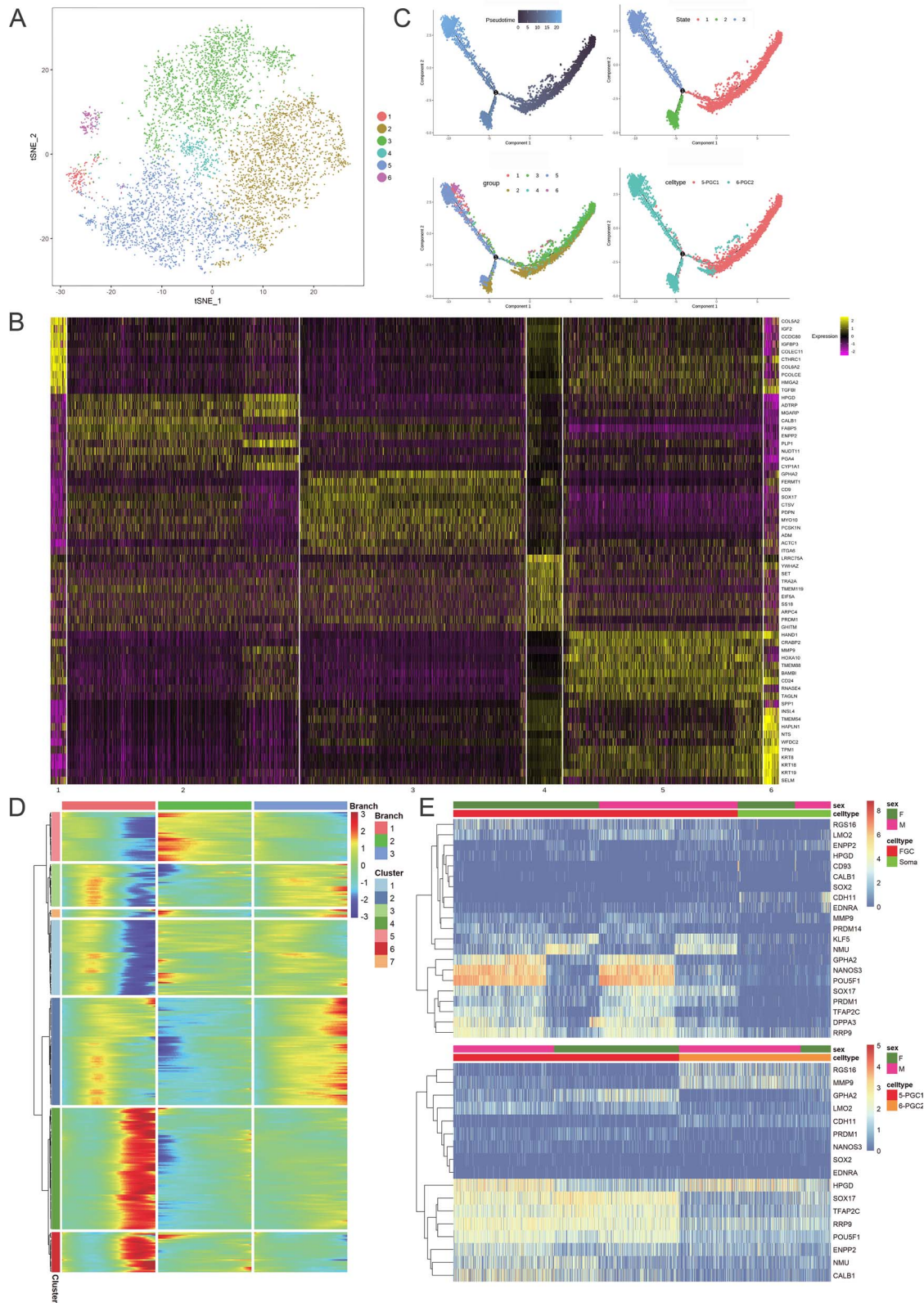
Figure 5. Developmental trajectories of PGCLCs. (A) *t*-SNE plots of PGCLCs. Cells are colored by cell types. (B) Heatmaps showing the differential gene expression patterns of each cluster. Top 10 differentially expressed genes of each cluster are shown. (C) Pseudotime analysis of developmental trajectories of PGCLCs. Cells colored by identified cell clusters (1–6) and cell types (PGC1 and PGC2). (D) Heatmap showing the gene expression dynamics during PGCLC differentiation. Genes are clustered according to expression patterns (clusters) and cell types are colored (branches). (E) Unsupervised hierarchical clustering (UHC) and a heatmap of the levels of selected marker genes in in vivo human PGCs (FGC) and in vitro monkey PGCLCs (PGC1 and PGC2). Color bars indicate cell types (PGC or soma cells) and sex (male or female).

induced from human PSCs. In this study, BLIMP1 and TFAP2C double-positive PGCLCs were induced from PSCs and were further cultured with murine embryonic ovarian somatic cells for a long period, approximately 4 months [42]. Based on these observations, we speculate that the heterogeneity of PGCLCs could be a potential reason for the long time required to induce oogonia-like cells from PGCLCs. In this study, we identified putative surface markers (PDPN and ITGA6) of distinct clusters of PGCLCs. In the future, using these markers to sort specific PGCLC types will allow us to study the time frame and efficiency of functional oogonia specification of distinct PGCLC types from PSCs, which warrants further investigation.

In conclusion, our study clarified the niche for PGCLC specification and identified putative signaling pathways that promote and inhibit PGCLC differentiation. More importantly, the developmental trajectory of PGCLCs was delineated. Finally, our work will offer new insights and clues to further optimize and improve the efficiency of germ cell differentiation from PSCs.

Supplementary material

Supplementary material is available at *BIOLRE* online.

Conflict of interest

The authors have declared that no conflict of interest exists.

Author contributions

W.J., S.Y., Y.Y., and T.T. designed the study; P.Z., S.X., R.G., N.S., H.S., and Y.F. performed the experiments; J.L., B.B., and T.T. analyzed the data; and T.T. wrote the article.

References

1. Sasaki H, Matsui Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat Rev Genet* 2008; **9**:129–140.
2. Leitch HG, Tang WW, Surani MA. Primordial germ-cell development and epigenetic reprogramming in mammals. *Curr Top Dev Biol* 2013; **104**:149–187.
3. Gunesdogan U, Magnusdottir E, Surani MA. Primordial germ cell specification: a context-dependent cellular differentiation event [corrected]. *Philos Trans R Soc Lond B Biol Sci* 2014; **369**:20140314.
4. Magnusdottir E, Surani MA. How to make a primordial germ cell. *Development* 2014; **141**:245–252.
5. Tang WW, Kobayashi T, Irie N, Dietmann S, Surani MA. Specification and epigenetic programming of the human germ line. *Nat Rev Genet* 2016; **17**:585–600.
6. Irie N, Weinberger L, Tang WW, Kobayashi T, Viukov S, Manor YS, Dietmann S, Hanna JH, Surani MA. SOX17 is a critical specifier of human primordial germ cell fate. *Cell* 2015; **160**:253–268.
7. Sasaki K, Yokobayashi S, Nakamura T, Okamoto I, Yabuta Y, Kurimoto K, Ohta H, Moritoki Y, Iwatani C, Tsuchiya H, Nakamura S, Sekiguchi K *et al*. Robust in vitro induction of human germ cell fate from pluripotent stem cells. *Cell Stem Cell* 2015; **17**:178–194.
8. Kojima Y, Sasaki K, Yokobayashi S, Sakai Y, Nakamura T, Yabuta Y, Nakaki F, Nagaoka S, Woltjen K, Hotta A, Yamamoto T, Saitou M. Evolutionarily distinctive transcriptional and Signaling pro-

grams drive human germ cell lineage specification from pluripotent stem cells. *Cell Stem Cell* 2017; **21**:517–532.e5.
9. Chen D, Sun N, Hou L, Kim R, Faith J, Aslanyan M, Tao Y, Zheng Y, Fu J, Liu W, Kellis M, Clark A. Human primordial germ cells are specified from lineage-primed progenitors. *Cell Rep* 2019; **29**:4568–4582.e5.
10. Sun Z, Wei Q, Zhang Y, He X, Ji W, Su B. MicroRNA profiling of rhesus macaque embryonic stem cells. *BMC Genomics* 2011; **12**:276.
11. Fang R, Liu K, Zhao Y, Li H, Zhu D, Du Y, Xiang C, Li X, Liu H, Miao Z, Zhang X, Shi Y *et al*. Generation of naive induced pluripotent stem cells from rhesus monkey fibroblasts. *Cell Stem Cell* 2014; **15**:488–497.
12. Niu Y, Sun N, Li C, Lei Y, Huang Z, Wu J, Si C, Dai X, Liu C, Wei J, Liu L, Feng S *et al*. Dissecting primate early post-implantation development using long-term in vitro embryo culture. *Science* 2019; **366**:eaaw5754.
13. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; **29**:15–21.
14. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018; **36**:411–420.
15. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012; **16**:284–287.
16. Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C. Single-cell mRNA quantification and differential analysis with census. *Nat Methods* 2017; **14**:309–315.
17. Kinoshita M, Barber M, Mansfield W, Cui Y, Spindlow D, Stirparo GG, Dietmann S, Nichols J, Smith A. Capture of mouse and human stem cells with features of formative pluripotency. *Cell Stem Cell* 2021; **28**:453–471.e8.
18. Yu L, Wei Y, Sun HX, Mahdi AK, Pinzon Arteaga CA, Sakurai M, Schmitz DA, Zheng C, Ballard ED, Li J, Tanaka N, Kohara A *et al*. Derivation of intermediate pluripotent stem cells amenable to primordial germ cell specification. *Cell Stem Cell* 2021; **28**:550–567.e12.
19. Messmer T, von Meyenn F, Savino A, Santos F, Mohammed H, Lun ATL, Marioni JC, Reik W. Transcriptional heterogeneity in naive and primed human pluripotent stem cells at single-cell resolution. *Cell Rep* 2019; **26**:815–824.e4.
20. Krentz NAJ, Lee MYY, Xu EE, Sproul SLJ, Maslova A, Sasaki S, Lynn FC. Single-cell transcriptome profiling of mouse and hESC-derived pancreatic progenitors. *Stem Cell Reports* 2018; **11**:1551–1564.
21. Biddy BA, Kong W, Kamimoto K, Guo C, Waye SE, Sun T, Morris SA. Single-cell mapping of lineage and identity in direct reprogramming. *Nature* 2018; **564**:219–224.
22. Mann KG, Elion J, Butkowski RJ, Downing M, Nesheim ME. Prothrombin. *Methods Enzymol* 1981; **80**:286–302.
23. Spear BT. Alpha-fetoprotein gene regulation: lessons from transgenic mice. *Semin Cancer Biol* 1999; **9**:109–116.
24. Hay DC, Zhao D, Fletcher J, Hewitt ZA, McLean D, Urruticoechea-Uriguen A, Black JR, Elcombe C, Ross JA, Wolf R, Cui W. Efficient differentiation of hepatocytes from human embryonic stem cells exhibiting markers recapitulating liver development in vivo. *Stem Cells* 2008; **26**:894–902.
25. Yabe S, Alexenko AP, Amita M, Yang Y, Schust DJ, Sadovsky Y, Ezashi T, Roberts RM. Comparison of syncytiotrophoblast generated from human embryonic stem cells and from term placentas. *Proc Natl Acad Sci U S A* 2016; **113**:E2598–E2607.
26. Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P. Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem* 2005; **280**:24731–24737.
27. Zeisel A, Munoz-Manchado AB, Codeluppi S, Lonnerberg P, La Manno G, Jureus A, Marques S, Munguba H, He L, Betsholtz C, Rolny C, Castelo-Branco G *et al*. Brain structure. Cell types in the

mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 2015; **347**:1138–1142.

28. Fuzik J, Zeisel A, Mate Z, Calvigioni D, Yanagawa Y, Szabo G, Linnarsson S, Harkany T. Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nat Biotechnol* 2016; **34**:175–183.

29. Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. *Cell* 2009; **139**:871–890.

30. Xu J, Lamouille S, Derynck R. TGF-beta-induced epithelial to mesenchymal transition. *Cell Res* 2009; **19**:156–172.

31. Tan TK, Zheng G, Hsu TT, Wang Y, Lee VW, Tian X, Wang Y, Cao Q, Wang Y, Harris DC. Macrophage matrix metalloproteinase-9 mediates epithelial-mesenchymal transition in vitro in murine renal tubular cells. *Am J Pathol* 2010; **176**:1256–1270.

32. Zhong S, Zhang S, Fan X, Wu Q, Yan L, Dong J, Zhang H, Li L, Sun L, Pan N, Xu X, Tang F *et al.* A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* 2018; **555**:524–528.

33. Schelker M, Feau S, Du J, Ranu N, Klipp E, MacBeath G, Schoeberl B, Raue A. Estimation of immune cell content in tumour tissue using single-cell RNA-seq data. *Nat Commun* 2017; **8**:2032.

34. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, Fallahi-Sichani M, Dutton-Regester K *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 2016; **352**:189–196.

35. Li L, Dong J, Yan L, Yong J, Liu X, Hu Y, Fan X, Wu X, Guo H, Wang X, Zhu X, Li R *et al.* Single-cell RNA-Seq analysis maps development of human germline cells and gonadal niche interactions. *Cell Stem Cell* 2017; **20**:858–873.e4.

36. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 2014; **32**: 381–386.

37. Sasaki K, Nakamura T, Okamoto I, Yabuta Y, Iwatani C, Tsuchiya H, Seita Y, Nakamura S, Shiraki N, Takakuwa T, Yamamoto T, Saitou M. The germ cell fate of Cynomolgus monkeys is specified in the nascent amnion. *Dev Cell* 2016; **39**:169–185.

38. Kuo YC, Au HK, Hsu JL, Wang HF, Lee CJ, Peng SW, Lai SC, Wu YC, Ho HN, Huang YH. IGF-1R promotes symmetric self-renewal and migration of alkaline phosphatase(+) germ stem cells through HIF-2alpha-OCT4/CXCR4 loop under hypoxia. *Stem Cell Reports* 2018; **10**:524–537.

39. Sakai Y, Nakamura T, Okamoto I, Gyobu-Motani S, Ohta H, Yabuta Y, Tsukiyama T, Iwatani C, Tsuchiya H, Ema M, Morizane A, Takahashi J *et al.* Induction of the germ cell fate from pluripotent stem cells in cynomolgus monkeys†. *Biol Reprod* 2020; **102**: 620–638.

40. Godin I, Wylie CC. TGF beta 1 inhibits proliferation and has a chemotropic effect on mouse primordial germ cells in culture. *Development* 1991; **113**:1451–1457.

41. Tang WWC, Castillo-Venzor A, Gruhn WH, Kobayashi T, Penfold CA, Morgan MD, Sun D, Irie N, Surani MA. Sequential enhancer state remodelling defines human germline competence and specification. *Nat Cell Biol* 2022; **24**:448–460.

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