The embryonic ontogeny of the gonadal somatic cells in mice and monkeys

AUTHOR(S):

Sasaki, Kotaro; Oguchi, Akiko; Cheng, Keren; Murakawa, Yasuhiro; Okamoto, Ikuhiro; Ohta, Hiroshi; Yabuta, Yukihiro; ... Yamamoto, Takuya; Seita, Yasunari; Saitou, Mitinori

CITATION:

Sasaki, Kotaro ...[et al]. The embryonic ontogeny of the gonadal somatic cells in mice and monkeys. Cell Reports 2021, 35(5): 109075.

ISSUE DATE:

2021-05

URL:

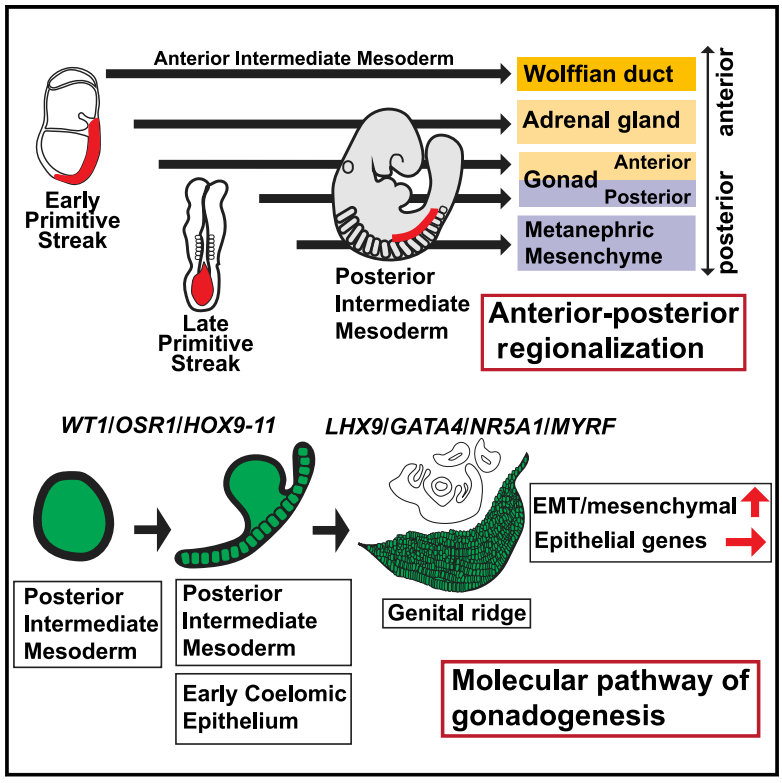
http://hdl.handle.net/2433/275697

Resource

Cell Reports

The embryonic ontogeny of the gonadal somatic cells in mice and monkeys

Graphical abstract



Highlights

- Ontogeny of the gonads in mice and monkeys defined by single-cell transcriptomics
- Temporally distinct contribution of T^+ descendants to the gonads and related organs
- The gonads originate from the T^+ primitive streak through posterior intermediate mesoderm
- The gonads acquire an epithelial-mesenchymal hybrid state upon specification

Authors

Kotaro Sasaki, Akiko Oguchi, Keren Cheng, ..., Takuya Yamamoto, Yasunari Seita, Mitinori Saitou

Correspondence

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

In brief

Sasaki et al. map the ontogenic ancestries of the gonads in mice and monkeys. Remarkably, the gonads originate from a T^+ primitive streak through $WT1^+$ posterior intermediate mesoderm, and this process involves sustained expression of epithelial genes and upregulation of mesenchymal genes, thereby conferring an epithelial-mesenchymal hybrid state.

Sasaki et al., 2021, Cell Reports *35*, 109075
May 4, 2021
https://doi.org/10.1016/j.celrep.2021.109075

CellPress

Cell Reports
CellPress OPEN ACCESS

Resource

The embryonic ontogeny of the gonadal somatic cells in mice and monkeys

Kotaro Sasaki,[1,2,11,*] Akiko Oguchi,[3,4] Keren Cheng,[1,2] Yasuhiro Murakawa,[3,4] Ikuhiro Okamoto,[4,5] Hiroshi Ohta,[4,5] Yukihiro Yabuta,[4,5] Chizuru Iwatani,[6] Hideaki Tsuchiya,[6] Takuya Yamamoto,[4,7,8,9] Yasunari Seita,[1,2,10] and Mitinori Saitou[4,5,7,*]

[1]Institute for Regenerative Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]Department of Biomedical Sciences, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[3]RIKEN Center for Integrative Medical Sciences, Yokohama, Kanagawa 230-0045, Japan
[4]Institute for the Advanced Study of Human Biology (ASHBi), Kyoto University, Kyoto 606-8501, Japan
[5]Department of Anatomy and Cell Biology, Graduate School of Medicine, Kyoto University, Kyoto 606-8501, Japan
[6]Research Center for Animal Life Science, Shiga University of Medical Science, Otsu, Shiga 520-2192, Japan
[7]Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto 606-8507, Japan
[8]AMED-CREST, AMED, Tokyo 100-0004, Japan
[9]Medical-risk Avoidance based on iPS Cells Team, RIKEN Center for Advanced Intelligence Project (AIP), Kyoto 606-8507, Japan
[10]Bell Research Center for Reproductive Health and Cancer, Nagoya 460-0003, Japan
[11]Lead contact
*Correspondence: ksasaki@upenn.edu (K.S.), saitou@anat2.med.kyoto-u.ac.jp (M.S.)
https://doi.org/10.1016/j.celrep.2021.109075

SUMMARY

In the early fetal stage, the gonads are bipotent and only later become the ovary or testis, depending on the genetic sex. Despite many studies examining how sex determination occurs from biopotential gonads, the spatial and temporal organization of bipotential gonads and their progenitors is poorly understood. Here, using lineage tracing in mice, we find that the gonads originate from a T^+ primitive streak through $WT1^+$ posterior intermediate mesoderm and appear to share origins anteriorly with the adrenal glands and posteriorly with the metanephric mesenchyme. Comparative single-cell transcriptomic analyses in mouse and cynomolgus monkey embryos reveal the convergence of the lineage trajectory and genetic programs accompanying the specification of biopotential gonadal progenitor cells. This process involves sustained expression of epithelial genes and upregulation of mesenchymal genes, thereby conferring an epithelial-mesenchymal hybrid state. Our study provides key resources for understanding early gonadogenesis in mice and primates.

INTRODUCTION

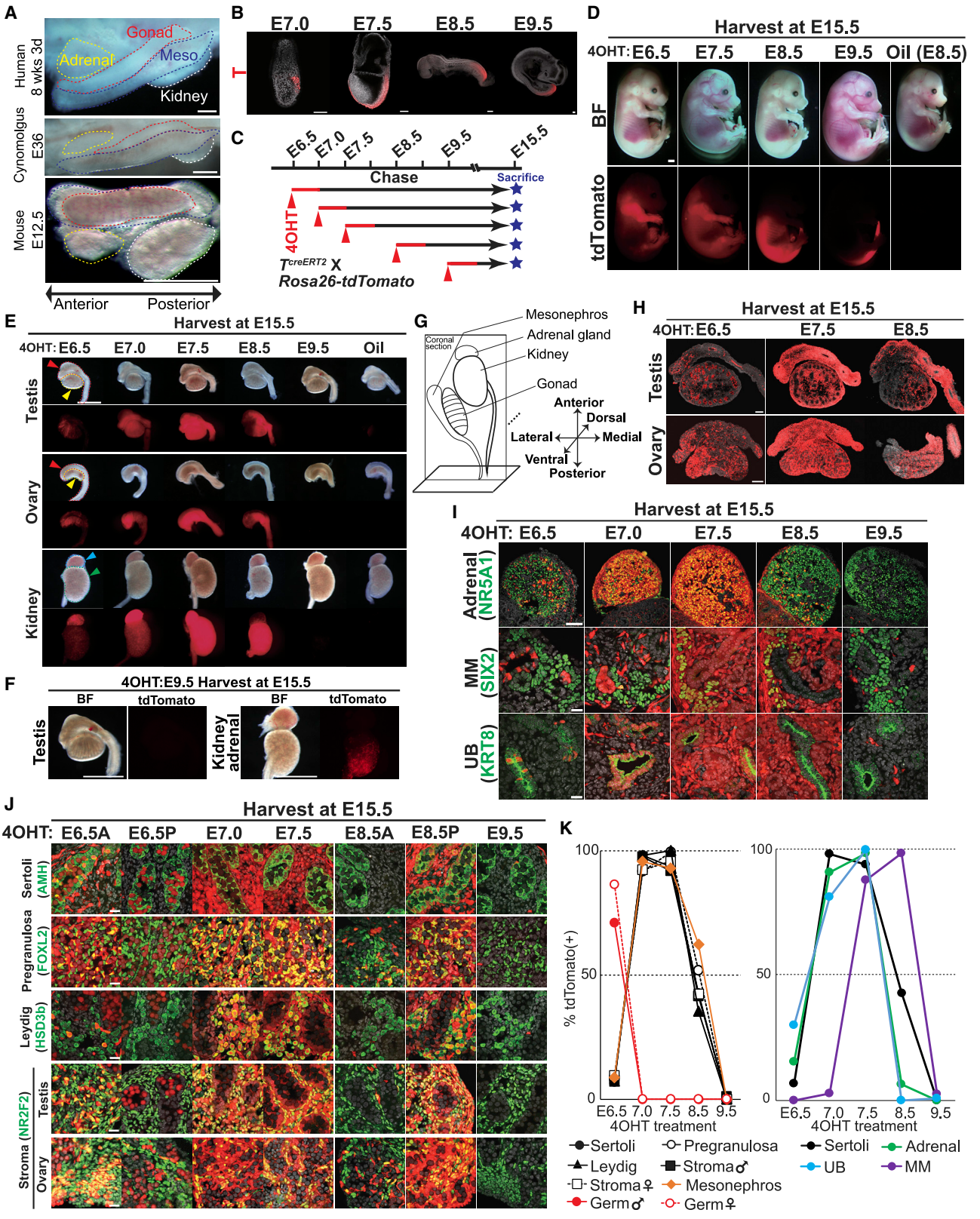
The formation of sexual dimorphism involves critical and pervasive events in sexually reproducing animals, thereby ensuring reproductive success. The gonads play a central role in such processes, because their differentiation into either ovaries or testes subsequently dictates the overall sex characteristics of the individual by regulating development of the external and internal sex organs and the brain via sex hormones (Capel, 2017; Yang et al., 2018). The embryonic gonads initially form as sexually bipotent structures known as genital ridges (GRs), which consist of GR progenitor cells (GPCs). Subsequently, the Y-linked sex-determining gene SRY, expressed in XY gonads, endows testicular characteristics, whereas WNT4/RSPO1/β-catenin signaling in XX gonads induces ovarian development.

Although most previous studies have focused on the cellular, genetic, and epigenetic cues that direct the formation of gonadal sexual dimorphism from GRs (Capel, 2017), the ontogeny of bipotential gonads and the accompanying molecular events driving their development remain poorly understood, particularly in humans and non-human primates. The formation of the GR starts around embryonic day (E) 10.5 through the proliferation of the coelomic epithelium (CE) at the ventromedial aspect of mesonephros, a process that appears to be accompanied by the dissolution of basement membranes and the epithelial-to-mesenchymal transition (EMT) (Karl and Capel, 1998; Kusaka et al., 2010; Yang et al., 2018). Prior studies have identified key genetic determinants of nascent GR formation, including *Gata4*, *Nr5a1*, *Lhx9*, *Emx2*, *Six1/4*, *Cbx2*, and *Tcf21* (Birk et al., 2000; Cui et al., 2004; Fujimoto et al., 2013; Hu et al., 2013; Kusaka et al., 2010; Luo et al., 1994; Miyamoto et al., 1997). Mutations in these genes result in gonadal depletion/hypoplasia. Among them, GATA4 marks most nascent GPCs that emerge at the CE ~E10.0, and activation of LHX9 and NR5A1 are required to fully specify GRs (Hu et al., 2013). However, the developmental origin of $GATA4^+$ GPCs and the cellular trajectories leading to GPCs from their progenitors remain unknown.

Analyses of mutant mice have revealed that genetic ablation of *Wt1*, *Osr1*, or *T* leads to loss of gonads, adrenal glands, mesonephros, and metanephros (i.e., the kidneys), suggesting that these genes are expressed in the ontogenic ancestries of GPCs and that these organs share a developmental origin at

CellPress OPEN ACCESS

Cell Reports
Resource



(legend on next page)

the early mesodermal stages (Chesley, 1935; Kreidberg et al., 1993; Wang et al., 2005). Although these genes are broadly expressed in various mesodermal derivatives and the primitive streak (PS), when and where these genes are activated in GPC progenitors remain unknown. To fill these knowledge gaps, we used lineage tracing of *Wt1*, *Osr1*, or *T*, combined with immunohistochemistry and single-cell RNA sequencing (scRNA-seq), to spatially and temporally map the ontogenic ancestries of GPCs and related organs at high resolution, thus providing resources and knowledge for understanding mammalian gonadogenesis.

RESULTS

Descendants of T^+ cells contribute to the gonads and related organs in a temporally distinct manner

In vertebrate embryos, the body axis progressively extends during development in an anterior to posterior direction, such that new tissues are continuously deposited at the posterior end and are fueled by proliferating T^+ progenitors residing in the PS, from which the anterior mesodermal organs exit early and posterior organs exit late (Deschamps and Duboule, 2017; Takasato and Little, 2015).

In the early development of human, cynomolgus monkey, and mouse embryos, a spatially distinct distribution along the anterior-posterior (AP) axis was observed for the adrenal glands, gonads, mesonephros, and kidneys. The gonads and overlying mesonephros were the most elongated organs along the AP axis, with the anterior end overlapping with the adrenal glands and the posterior end overlapping with the metanephric mesenchyme (MM) (Figure 1A). We hypothesized that the AP regionalization of the gonads and related organs would reflect their timing of exit from the PS. Whole-mount immunofluorescence (IF) revealed T^+ PS cells along the midline of posterior embryos during E7.0–E7.5; after E8.5, more restricted localization was observed at the posterior end, previously termed the posterior growth zone (Figure 1B) (Deschamps and Duboule, 2017). To determine the contribution of T^+ PS cells to the gonads and related organs at specific developmental time points, we traced the lineage of the descendants of T^+ PS cells at the associated developmental stages (Figures 1C and 1D; STAR Methods). Interestingly, 4-hydroxytamoxifen (4OHT) injection at E6.5, which was expected to label T^+ PS cells ~E6.5–E7.0, showed that the contribution of the descendant cells was restricted to the anterior half of the gonads (testis and ovary)/mesonephros, as well as the entire adrenal glands at E15.5, with minimal contribution to the kidneys (Figure 1E). In contrast, 4OHT injection at E8.5 labeled only the posterior half of the gonads and mesonephros, as well as the entire kidney, without significantly labeling adrenal glands (Figures 1E, 1G, and 1H). These data confirmed, at the tissue level, that the early PS contributes to the anterior organs (adrenal glands and the anterior half of the gonads/mesonephros), whereas the late PS contributes to the posterior organs (kidneys and the posterior half of the gonads/mesonephros). When 4OHT was injected at E9.5, the contribution of T^+ descendent cells was restricted to the posterior half of the kidneys, albeit weakly, whereas essentially no contribution was seen in other tissues (Figure 1F). These data suggest that the posterior kidney is derived from T^+ descendant cells that remain in the PS until a late stage of development, when the T^+ descendants that contribute to other organs have already exited the PS and initiated mesodermal differentiation.

We next evaluated the contributions of labeled T^+ descendants at the cellular level (Figures 1G–1K). The contribution of T^+ descendants to all major somatic cell types in the testes and ovaries exhibited kinetics similar to that observed at the tissue level (Figures 1H–1K). In the kidneys, the $KRT8^+$ ureteric bud

Figure 1. Descendants of the T^+ primitive streak contribute to the gonads and related organs in a temporally distinct manner

(A) Bright-field images of the gonads (red), adrenal glands (yellow), mesonephros (blue), and kidneys (white) in embryos from humans (8 weeks, 5 days), cynomolgus monkeys (E36), and mice (E12.5) are shown with topological orientation maintained. Scale bar, 0.5 mm.

(B) Whole-mount IF images for T in embryos at the indicated stages (embryos n = 7, 8, 5, and 4). E7.0–E8.5 embryos are oriented from anterior (left) to posterior (right). Scale bar, 100 μm.

(C) Schematic for the lineage tracing of T^+ descendants.

(D) Bright-field and fluorescence images (for tdTomato, red) of embryos at E15.5 after lineage tracing for T through 4OHT injection at the indicated stages (embryos n = 9, 10, 11, and 6 for E6.5, E7.5, E8.5, and E9.5, respectively). As a control, oil was injected in the dams during pregnancy at E8.5 (embryos n = 3). Scale bar, 1 mm.

(E) Bright-field and fluorescence images (for tdTomato, red) of testes/ovaries (with attached mesonephros), kidneys, and adrenal glands at E15.5 after lineage tracing for T, as in (C). Testes/ovaries/mesonephros are oriented from anterior (left) to posterior (right) (embryos n = 2, 2, 3, 4, 2, 2 [male]; 2, 1, 2, 2, 1, 1 [female]). Arrowheads indicate mesonephros (red), testes/ovaries (yellow), adrenal glands (blue), and kidneys (green). Scale bar, 1 mm.

(F) Bright-field and fluorescence images of a testis/mesonephros (left) and a kidney/adrenal gland (right) at E15.5 after lineage tracing for T through 4OHT injection at E9.5. Images were taken with a longer exposure time than those in (E). Scale bar, 1 mm.

(G) Diagram of the coronal sections taken for histologic studies as in (H), (I), and (J).

(H) Representative coronal sections of testes/ovaries/mesonephros at E15.5 after lineage tracing for T by 4OHT injection at E6.5, E7.5, and E8.5 (embryos n = 2, 3, and 4 [male] and 2, 2, and 2 [female], respectively). Fluorescence images for tdTomato (red) merged with DAPI (white). Scale bar, 100 μm.

(I) Representative coronal sections of adrenal glands and kidneys at E15.5 after lineage tracing for T, as in (C). Immunofluorescence (IF) staining for NR5A1, SIX2, and KRT8 (green) at the indicated stages (embryos n = 4, 3, 2, 5, 6, 6) marks adrenal gland cells, metanephric mesenchyme (MM), and the ureteric bud (UB) of the kidneys, respectively. Merged images with tdTomato (red) and DAPI are shown. Scale bar, 100 μm (top row) or 20 μm (middle and bottom rows).

(J) IF images of coronal sections of ovaries and testes at E15.5 after lineage tracing for T, as in (C). The 4OHT was injected at the indicated stages (embryos n = 2, 2, 3, 4, 2 [male]; 2, 1, 2, 2, 1 [female]). For E6.5 and E8.5 embryos, both anterior (E6.5A and E8.5A) and posterior (E6.5P and E8.5P) portions of the testes/ovaries are shown. Markers of Sertoli cells (AMH), pregranulosa (FOXL2), Leydig cells (HSD3b), and stroma (NR2F2) (green), merged with tdTomato (red) and DAPI (white), are shown. Scale bar, 20 μm.

(K) Percentage of tdTomato-positive cells in the indicated cell types at E15.5. 4OHT was injected at the indicated time point. Markers of germ cells (DDX4) and mesonephros (WT1) were used, together with the markers used in (I) and (J).

See also Figure S1.

(UB) showed labeling patterns distinct from those of the MM. For example, when 4OHT was injected at E8.5, only the MM was labeled, whereas no labeling was observed in the UB in E15.5 kidneys; this suggested that the UB progenitor cells had already exited the PS and lost T expression at E8.5, whereas the MM progenitor cells remained within the T^+ PS (Figures 1H and 1J) (Taguchi et al., 2014). The gonads exhibited an intermediate contribution pattern between those of UB/adrenal glands (no/very low contribution) and those of MM (~100% contribution): ~50% of the gonadal somatic cells and mesonephros were derived from T^+ descendants labeled by E8.5 4OHT injection, thus suggesting that their origin spans both early and late PS (Figures 1H–1J). Altogether, our data provide overall support for our hypothesis that the temporal sequence of exit from the PS translates to position along the AP axis.

Transition of T^+ PS cells into $OSR1^+$ intermediate mesoderm/lateral plate mesoderm progenitors

The gonads, mesonephros, adrenal glands, and kidneys are composed of descendant cells that express OSR1, a broad marker of the intermediate mesoderm (IM), lateral plate mesoderm (LPM), and their derivatives (James et al., 2006; Mugford et al., 2008; Taguchi et al., 2014; Wang et al., 2005). *Osr1* promoter activity started after E8.0 in bilateral streak-like structures along the AP axis of embryos and persisted until E9.5 (Figure S1A). By E10.5, the *Osr1* activity was mostly lost except at a few regions (Figure S1A). IF studies indicated that T^+ PS cells gradually lost T and acquired OSR1-EGFP as they moved ventrally and anteriorly from the PS/posterior growth zone, thus forming the putative progenitors of the IM/LPM (Figures S1B–S1D).

A lineage tracing study revealed the broad distribution of $OSR1^+$ descendants in the trunk of embryos, in agreement with their contributions to IM and LPM derivatives (Figures S1E and S1F). Further inspection of $OSR1^+$ descendants at the tissue level showed contribution dynamics similar to those of T^+ descendants, albeit with delayed kinetics (Figure S1G). These results suggested that T^+ descendants sequentially activate *Osr1* across the developmental time course as they exit the PS and contribute to generation of the gonads and related organs.

The gonads originate from $WT1^+$ posterior intermediate mesoderm

Previous studies have shown that the gonads are derived from $WT1^+$ CE present at the ventromedial aspect of the mesonephros (Bandiera et al., 2013; Hu et al., 2013). However, the origin of such CE remains unclear. Whole-mount IF revealed that WT1 expression was first detected at E9.0 (~13 somites [SMs]) as a short streak of cells present at the ventral aspect of the posterior SM (Figure 2A). Strong expression was also observed in the proepicardium, as reported previously (Zhou et al., 2008). Central cross sections of the $WT1^+$ structures revealed that $WT1^+$ cells co-expressing SALL1 and SALL4 were localized at the IM (Figures S2A–S2D). WT^+ IM was encompassed by the $OSR1\text{-}EGFP^+$ regions but excluded from both the $FOXF1^+$ LPM and a posteriorly extending $LHX1^+$ Wolffian duct (WD), a derivative of the anterior IM (AIM) that is known to be specified ~E8.5 at the pronephric region (Figures S2A and S2B) (Taguchi et al., 2014). These data suggest that $WT1^+$ IM represents the posterior IM (PIM), which is specified as a lineage distinct from the AIM or LPM as early as E9.0. Concomitantly with body axis elongation, $WT1^+$ regions extended along the AP axis, and WT1 expression increased in the anterior direction (Figures 2A and 2B). Cross sections revealed that $WT1^+$ cells initially localized at the IM and expanded ventrolaterally as a continuous mass while increasing cell number, thus forming the early CE (ECE), which by E10.5 expanded to encompass the entire urogenital ridge consisting of CE containing $NR5A1^+WT1^+$ GPCs and underlying $NR5A1^-WT1^+$ mesonephros (Figures 2C and S2D–S2F).

We next explored the spatial and temporal dynamics of GPC specification in $WT1^+$ cells. Whole-mount IF at different stages revealed that $NR5A1^+GATA4^+$ GPCs emerged at E10.0 at the anterior end of $WT1^+$ CE, whereas patchy $NR5A1^+$ cells emerged within the diffusely $GATA4^+$ CE, thus suggesting that GATA4 may be upregulated earlier than NR5A1 in the CE (Figures 2D and 2E) (Hu et al., 2013). Thereafter, GPCs extended in the posterior direction while increasing their thickness and extending to two-thirds the length of the entire $WT1^+$ CE/mesonephros by E11.0 (Figure 2D).

Comparison of cross sections from anterior and posterior regions of embryos at E9.0 or E10.5 revealed that anterior regions exhibited more developmentally mature features (i.e., the emergence of $WT1^+$ ECE at E9.0 or the thickening of $NR5A1^+$ GPCs at E10.5) (Figures S2D–S2F). Overall, these results support IM development progressing in the AP direction, which is consistent with prior literature (Brambell, 1927; Hu et al., 2013).

Anterior and posterior gonads are formed by temporally distinct $WT1^+$ descendants

To corroborate our IF analysis, we performed lineage tracing of $WT1^+$ descendants (Figure 3A). To ensure specific labeling of $WT1^+$ PIM at E9.0, not their descendants (i.e., CE), we injected the 4OHT 24 h earlier than the time at which $WT1^+$ PIM first emerges.

4OHT injection at E8.0 did not result in overt labeling of tissues at E15.5. However, when 4OHT was injected at E8.5 and E9.5, strong and specific labeling was observed at the region representing the gonads/mesonephros/adrenal glands at E15.5, indicating that in contrast to $OSR1^+$ descendants, $WT1^+$ descendants have limited developmental potential (Figure 3B). Indeed, with the exception of the epicardium, most other internal organs of either mesodermal or endodermal origins were not labeled (Figures 3B, 3C, and S3A). At the tissue level, $WT1^+$ descendants labeled by E8.5 4OHT injection contributed to the anterior gonads/mesonephros, as well as the adrenal glands (Figure 3C). The contribution of the $WT1^+$ descendants to the posterior gonads/mesonephros or the kidneys occurred when 4OHT was injected at E9.5 or E10.5 onward, respectively (Figure 3C). These results are in overall agreement with those obtained from T- and OSR1-lineage tracing, and suggest that early and late $WT1^+$ descendants contribute to the anterior and posterior PIM derivatives, respectively.

Histologic quantification of the lineage-traced cellular constituents in E15.5 embryos revealed that most gonadal somatic cells, including surface (coelomic) epithelium, were $WT1^+$ descendants labeled by 4OHT injection during E8.5–E9.5 (Figures 3D–3F and

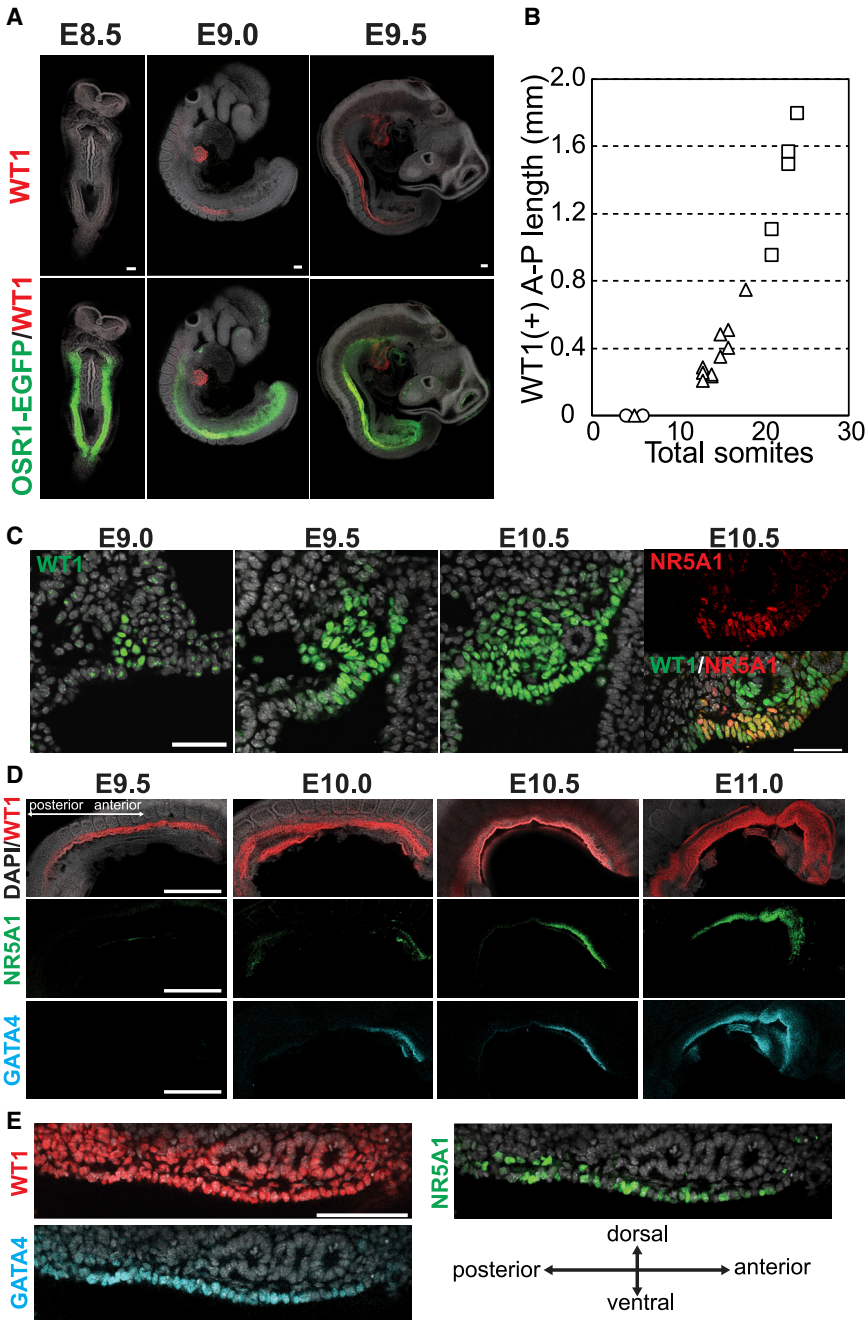


Figure 2. The gonads originate from WT1$^+$ posterior intermediate mesoderm

(A) Whole-mount IF images of WT1 (red) and OSR1-EGFP (green) in embryos at E8.5, E9.0, and E9.5, merged with DAPI (white) (embryos n = 2, 11, and 5, respectively). Scale bar, 100 μm.

(B) Lengths of WT1$^+$ regions along the AP axis for embryos at E8.5 (circle), E9.0 (triangle), and E9.5 (square) and their somite numbers, as in (A).

(C) IF on the transverse sections of embryos at the indicated stages for WT1 (green) merged with DAPI (white) (left) or DAPI and NR5A1 (red) (right) (embryos n = 3, 2, and 2 for E9.0, E9.5, and E10.5, respectively). The sections were obtained from the centers of WT1$^+$ regions along the AP axis, and the magnified images of the coelomic angles are shown, oriented dorsal (up) to ventral (down) and medial (left) to lateral (right). Scale bar, 50 μm.

(D) Whole-mount IF images for WT1 (red) merged with DAPI (white), NR5A1 (green), and GATA4 (cyan) in embryos at the indicated stages (embryos n = 3, 7, 5, 2). Embryos are oriented posterior (left) to anterior (right). Scale bar, 500 μm.

(E) Magnified images of an E10.0 embryo for NR5A1 (green), WT1 (red), and GATA4 (cyan) merged with DAPI, as in (D) with a slightly different plane on the z axis. Coelomic epithelium with underlying mesonephros. Scale bar, 100 μm.

See also Figures S1 and S2.

Cells with gonadal ancestry do not express SIX2

Previous studies have suggested that at E9.5, SIX2 is expressed in a portion of the IM, which has limited nephrogenic potential according to *ex vivo* culture assays (Taguchi et al., 2014). We observed scattered SIX2$^+$ cells within OSR1-EGFP$^+$WT1$^+$ PIM at E9.5 (Figure S3E). To determine whether SIX2$^+$ PIM cells are already committed to a nephrogenic fate or have the potential to differentiate into gonadal/adrenal gland lineages, we trace the fate of SIX2$^+$ descendants. Remarkably, the descendants of SIX2$^+$ cells labeled by 4OHT injection at E9.0 or E9.5 did not contribute to the gonads or adrenal glands. Instead, these cells contributed to a small portion of the mesonephros, thus suggesting that commitment to the mesonephros occurs in a portion of the SIX2$^+$ PIM (Figure S3F). The kidneys were populated by SIX2$^+$ descendants labeled by 4OHT injection at E9.5 or E10.5 (Figure S3F) (Self et al., 2006).

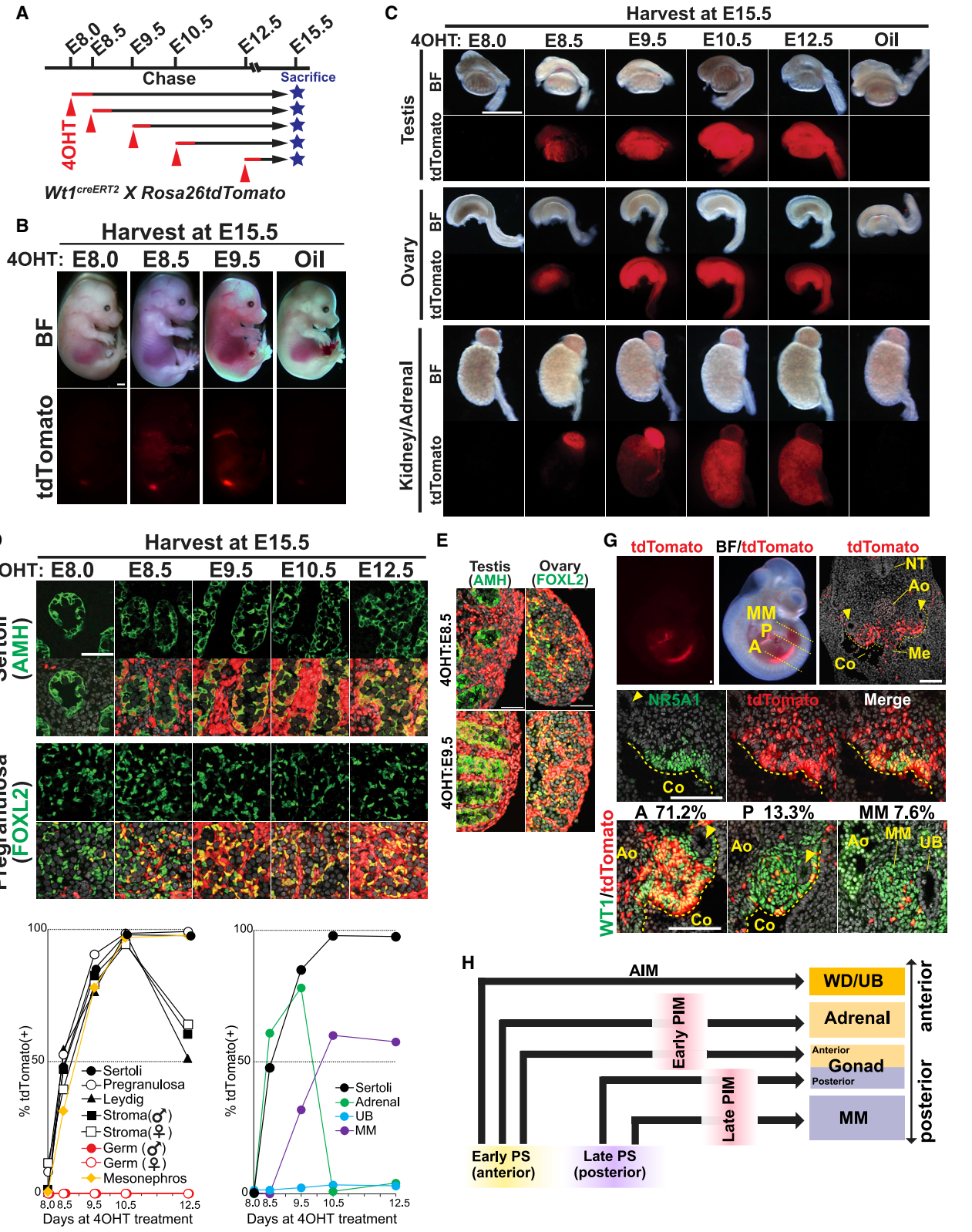
S3B). IF on embryos at E10.5 (4OHT injected at E8.5) showed that WT1$^+$ descendants labeled by this 4OHT treatment specifically contributed to the mesonephric stroma; overlying CE, including NR5A1$^+$ GPCs; and to a lesser extent, the mesenteric stroma (Figure 3G). Altogether, these observations support the notion that WT1$^+$ PIM contributes to gonadal lineages via CE. The contribution of WT1$^+$ descendants to mesonephric cells and gonadal cells showed similar dynamics (Figures 3F, S3C, and S3D). In the kidneys, the contribution of WT1$^+$ descendants to the MM started at E9.5 and peaked at E10.5, whereas UBs were not populated by WT1$^+$ descendants at any developmental time point.

Transcriptomic changes accompanying GPC specification from the PIM

We next determined the transcriptomic changes accompanying the specification and early development of GPCs (Figures S4A–S4C; STAR Methods). Because our interest was in GPC specification, not sex determination, we focused on analysis of female cells.

CellPress
OPEN ACCESS

Cell Reports
Resource



(legend on next page)

Unsupervised hierarchical clustering (UHC) revealed five distinct cell clusters: *Eya1*$^+$/*Osr1*$^+$/*Wt1*$^+$/*Sall1*$^+$ mouse PIM (mPIM)/ECE (Figures 2C and S2A–S2C) (James et al., 2006; Sajithlal et al., 2005; Taguchi et al., 2014), *Gata4*$^+$/*Nr5a1*$^+$/*Lhx9*$^+$ mouse GPCs (mGPCs) (Birk et al., 2000; Hu et al., 2013; Luo et al., 1994), *Maf*$^+$/*Tcf21*$^+$/*Col6a2*$^+$ mouse gonadal stroma (mST) (Stévant et al., 2018, 2019), *Rspo1*$^+$/*Wnt4*$^+$/*Runx1*$^+$ mouse pregranulosa early (mpGrE) and mouse pregranulosa late (mpGrL) (Niu and Spradling, 2020; Stévant et al., 2019) (Figures 4A–4C). Principal-component analyses (PCAs) revealed continuous lineage trajectories of these distinct clusters. The PIM/ECE first transitioned into mGPCs by increasing the principal component (PC)3 scores (Figures 4B and S4D). During this transition, mPIM/ECE genes (e.g., *Osr1*, *Eya1*, and *Sall1*) were downregulated, and genes involved in early gonadogenesis (e.g., *Nr5a1*, *Gata4*, *Lhx9*, and *Zfpm2*) were upregulated (Figures 4C and S4H). After specification of mGPCs, the lineage bifurcated and became either mST or pregranulosa lineages (mpGrE and mpGrL) (Figures 4B and S4D). The transcriptomes of mGPCs were closer to those of mST than mpGrE (Figure S4F), and mGPCs had already started to show upregulation of many stromal markers (e.g., *Maf*, *Tcf21*, and *Col6a2*), thus suggesting that GPC specification may accompany EMT-like changes (Figures 4A, 4C, and S4H). For mST lineage progression, we identified 279 genes with significantly positive scores for PC1 loading (SD > 2) (mST genes) that were enriched in Gene Ontology (GO) terms such as "proteinaceous extracellular matrix" and "male gonad development" (Figure 4D). For pregranulosa lineages that progressed from the PIM/ECE toward mpGrL via mGPCs, we found 391 genes with significantly positive scores for PC1 loading (pregranulosa genes) that were enriched in GO terms such as "lipid metabolic process" or "epithelial differentiation," in agreement with previous studies (Figure 4E) (Niu and Spradling, 2020; Stévant et al., 2019).

The 726 differentially expressed genes (DEGs) upregulated in mPIM/ECE included *Hox* genes and genes involved in WNT signaling (Figure S4G; Table S1). The 366 DEGs upregulated in mGPCs were enriched in GO terms such as "cell division," thereby suggesting that mGPCs are highly proliferative (Figures S4G and S4H; Table S1). Overall, these data revealed the transcriptional properties of mGPCs and delineated the lineage progression of mGPCs and their derivatives from mPIM/ECE.

The ontogeny of GPCs in primates

We next explored the spatial and temporal dynamics of GPC specification in primates. To this end, we first used human induced pluripotent stem cell (hiPSC)-based induction of MM (Taguchi et al., 2014), which appears to share early mesodermal ancestry with GPCs (Figures 1E, 1K, 3C, 3F, and S1G). *OSR1-EGFP* knockin hiPSCs showed gradual downregulation of *POU5F1*, a marker of pluripotency, and sequential upregulation of *T* (day 3), *OSR1* (day 5), and *WT1* (day 7) (Figures 5A and 5B), thus suggesting that the specification of WT1$^+$ PIM may be conserved between mice and humans.

We next set out to determine the spatial and temporal dynamics of WT1$^+$ PIM specification *in vivo* using cynomolgus monkey embryos. Serial sections of a E23 embryo (Carnegie stage [CS] 10) revealed a distinct cluster of cells, which were diffusely LHX1$^+$ and focally PAX2$^+$ (Figures S5A and S5B) but WT1$^-$ (Figure 5C) and therefore represented cynomolgus monkey AIM (cAIM). Notably, the cells were FOXF1$^-$, in contrast to the LPM, thus suggesting the segregation of the AIM from the LPM both morphologically and immunophenotypically at this stage (Figure S5B).

WT1$^+$ PIM first emerged at E24 (n = 2, CS11) as an anatomically distinct structure between the PM and the LPM (Figure 5C). The PIM was distinguishable from the AIM by no/very weak expression of AIM markers (Figure S5C). Similar to the AIM, the PIM was clearly demarcated from the FOXF1$^+$ LPM (Figure S5C). Serial transverse sections revealed that the AIM was localized anteriorly and the PIM was localized posteriorly with some overlap at the center, where the AIM was lateral to the PIM (Figure 5D). At E26 (n = 2, CS11, CS12), the AIM and PIM ran in parallel along the AP axis, probably as a result of the posterior extension of the AIM (Figure S5D) (Grote et al., 2006). At this stage, the AIM acquired the epithelial marker CDH1, thereby

Figure 3. Lineage tracing of WT1$^+$ descendants

(A) Scheme for lineage tracing of WT1$^+$ descendants.

(B) Bright-field and fluorescence images for tdTomato in embryos at E15.5 after lineage tracing of WT1$^+$ descendants with 4OHT injection at the indicated stages (embryos n = 10, 5, 4, 5), as illustrated in (A). Scale bar, 1 mm.

(C) Bright-field and fluorescence images of testes/ovaries (with attached mesonephros) or kidneys/adrenal glands at E15.5, as illustrated in (A) (embryos n = 5, 2, 1, 5, 2 [male]; 5, 3, 3, 3, 2 [female]). As a control, oil was injected at E8.5 (embryos n = 3 [male]; 2 [female]). Scale bar, 1 mm.

(D) IF images of sections for ovaries and testes at E15.5 after lineage tracing for WT1$^+$ descendants, as illustrated in (A). 4OHT was injected at the indicated stages (embryos n = 5, 2, 1, 5, 2 [male]; 5, 3, 3, 3, 2 [female]). Markers of Sertoli cells (AMH) or pregranulosa cells (FOXL2) (green), merged with tdTomato (red) and DAPI (white), are shown. The anterior portions of gonads are shown for embryos with 4OHT injection at E8.5. Scale bar, 50 μm.

(E) IF images of ovaries and testes at E15.5, as shown in (D), highlighting the diffuse tdTomato expression in FOXL2$^-$ surface (coelomic) epithelium.

(F) Percentage of tdTomato-positive cells in the indicated cell types at E15.5; 4OHT was injected at the indicated time point. To identify cell types, the markers in Figure 1J were used.

(G) Images of E10.5 embryos; 4OHT was injected at E8.5 to trace the lineage of WT1$^+$ descendants (embryos n = 2). Fluorescence images for tdTomato (top left), merged with a bright-field image (top center), and the transverse sections at the indicated levels (A, anterior; P, posterior; MM, metanephric mesenchyme) are shown. Low-magnification images of coelomic angles at level A for tdTomato (red) merged with DAPI (white) are shown at the top right, and high-magnification images at level A for NR5A1 (green) and tdTomato (red) merged with DAPI are shown in the second row. IF images for WT1 (green) and tdTomato (red) merged with DAPI at levels A, P, and MM are shown at the bottom, with the percentage of tdTomato$^+$ cells among WT1$^+$ cells indicated. Arrowheads and dotted lines indicate wolffian duct and coelomic surface, respectively. Ao, aorta; Co, coelom; Me, dorsal mesentery; MM, metanephric mesenchyme; NT, neural tube; UB, ureteric bud. Scale bar, 100 μm.

(H) Model for AP regionalization of PIM derivatives.

See also Figure S3.

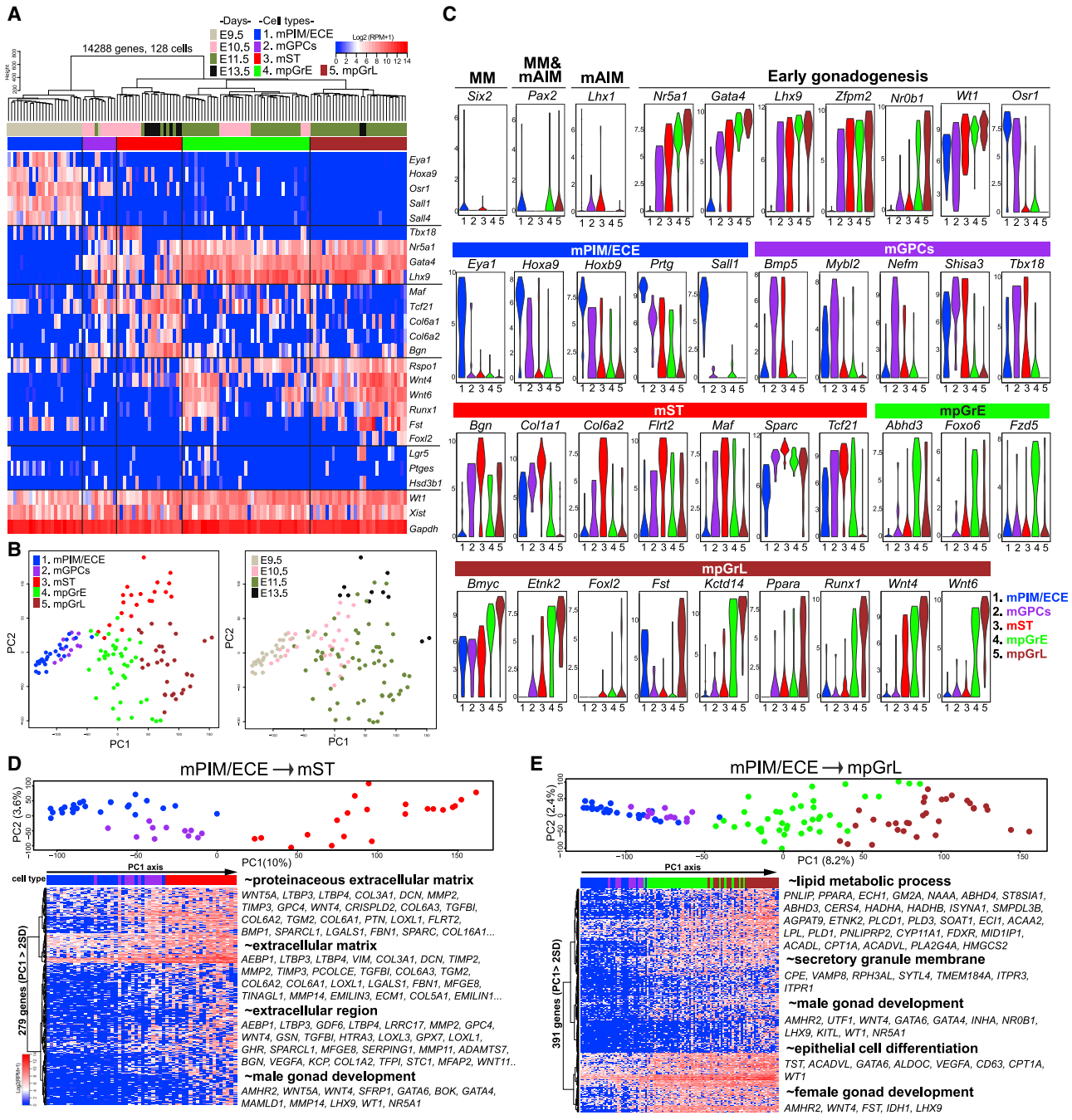


Figure 4. Transcriptomic changes accompanying mGPC specification from mPIM/ECE

(A) UHC (log2[RPM{reads per million-mapped reads}+1] ≥ 6 in at least one sample among 128 cells, 14,288 genes) and a heatmap of the levels of selected marker genes. Color bars under the dendrogram indicate embryonic days (top) and cell types (bottom). mPIM/ECE, mouse posterior intermediate mesoderm/early coelomic epithelium; mGPCs, mouse GR progenitor cells; mST, mouse gonadal stroma; mpGrE, mouse pregranulosa early; mpGrL, mouse pregranulosa late.

(B) PCA of all cells, as in (A). The cells (color coded as in A: left, cell types; right, embryonic days) are plotted in two-dimensional space defined by PC1 and PC2.

(C) Violin plots showing the expression of known markers of the indicated cell types/process (top) or genes upregulated in the indicated cell types, compared with others, as revealed by DEG analysis in Figure S4G.

(D and E) PCA showing the lineage trajectory for the stromal lineage (D) or pregranulosa lineage (E) (top), and heatmap of the expression of genes defined as highly contributing to the PC1 axis for each PCA plot (PC1 > 2 SD) (bottom). Cells in the heatmap were ordered by scores for PC1 loading. Representative genes and key GO enrichments are shown (bottom right).

See also Figure S4.

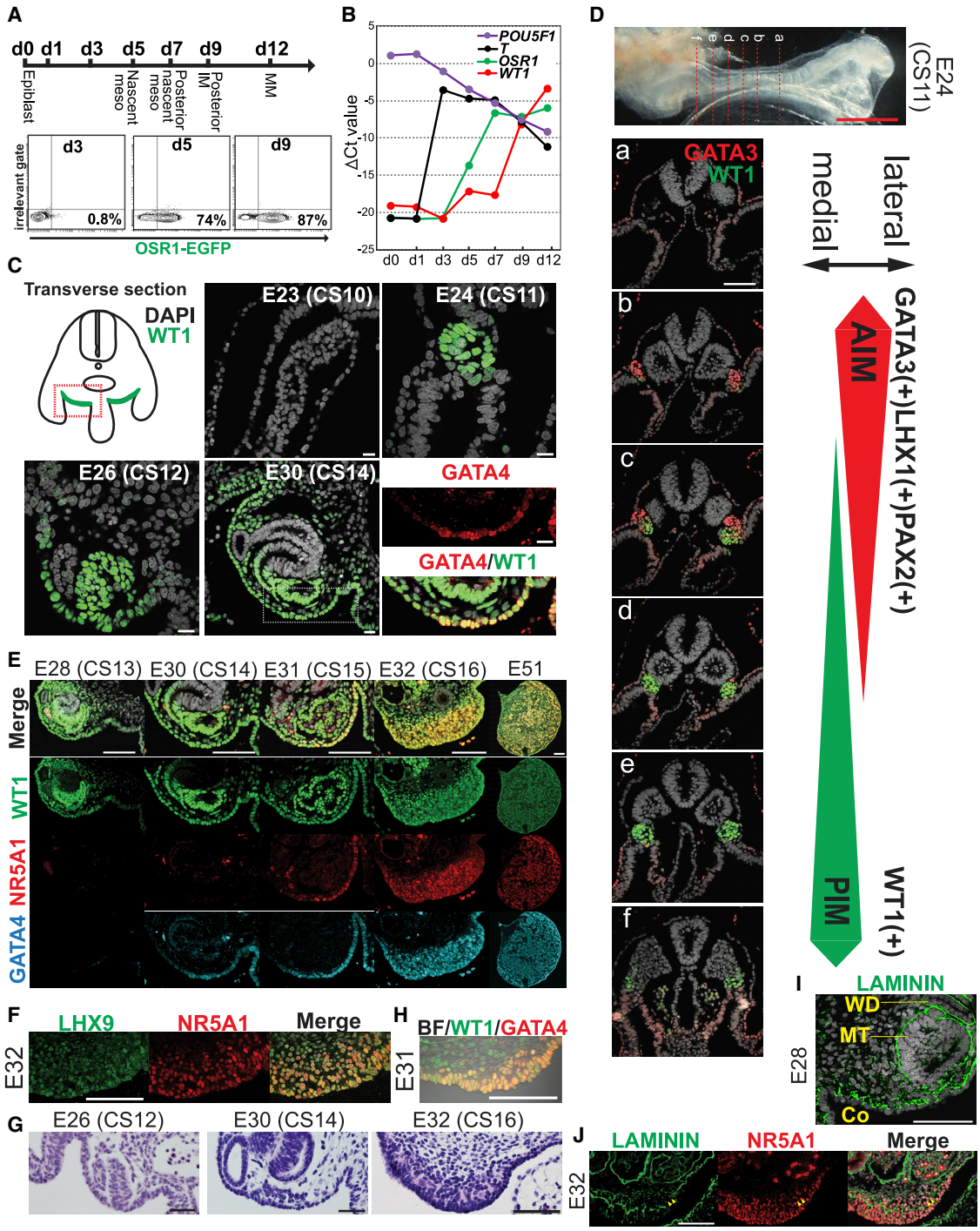


Figure 5. The ontogeny of GPCs in primates

(A) Induction of MM from hiPSCs by 3D spheres. Induction schemes (top) and the percentage of OSR1-EGFP$^+$ cells as detected by flow cytometry on the indicated days. Representative plots of two biological replicates are shown.

(B) Quantitative PCR (qPCR) analysis of the expression of key genes in cDNA generated from bulk spheres at the indicated time point during MM induction, as in (A). ΔCt values were calculated by subtraction of the raw Ct values of each gene (mean value of two biological replicates) from the averaged Ct values of the housekeeping genes *ARBP* and *PPIA*.

(C) Schematic of IF of cynomolgus monkey embryo sections (upper left) and transverse sections obtained from cynomolgus monkey embryos at the indicated stages (embryonic days and Carnegie stages, embryos n = 1 [E23], 2 [E24], 2 [E26], 2 [E30]). WT1 (green) merged with DAPI (white) is shown. For the E30 section, magnified IF images of the dotted regions further merged with GATA4 (red) are shown. Scale bar, 20 μm.

(legend continued on next page)

suggesting that differentiation into the WD had occurred (Figure S5E) (Grote et al., 2006). The anterior regions (CS11) or midsections (CS12) of E26 embryos showed the PIM extending ventrolaterally and forming ECE, similar to the findings in mice at E9.5 (Figures 5C, S5D, and S5F). Likewise, within the same embryos at E26, differences in PIM maturation were observed; more anterior sections showed increased lateral ECE extension (Figures S5D and S5F). The posterior ends of embryos at E26 showed a T^+ caudal eminence (Figure S5F) corresponding to the posterior growth zone in mice (Figure 1B) (Müller and O'Rahilly, 2004), thus suggesting ongoing mesodermal specification.

By E28, $WT1^+$ cells formed mesospheric stroma, mesonephric tubules connected to WDs, and CE entirely covering the urogenital ridge (Figure 5E). CE with early gonadal features first emerged as $WT1^+GATA4^+NR5A1^-$ cells at E30 (CS14) (Figures 5C and 5D). At E31 (CS15), more mature $WT1^+GATA4^+NR5A^+$ cells emerged within the $WT1^+GATA4^+$ CE at anterior cross sections (approximately two SMs beneath the forelimb bud) (Figure S5G). The formation of histologically recognizable GRs started focally in one embryo at E31 (CS15) (Figure 5H) and more diffusely in all embryos at E32 (CS16) (Figures 5E–5G and 5J), as revealed by nuclear overlapping and fragmentation of basement membranes suggestive of EMT-like changes (Figures 5G–5J). Interestingly, we observed expression of KRT19 (an intermediate filament and epithelial marker) throughout GPC development; KRT19 was expressed somewhat weakly in the PIM (E24) and strongly in CE/GPCs during E30–E32 (Figures S5H and S5I). Further IF profiling of GPCs at E31/E32 revealed that GPCs also expressed KRT18 (an epithelial marker) and VIM (a mesenchymal marker). CDH1, a canonical epithelial marker, was not expressed in GPCs, CE, or PIM (Figures S5H and S5I). Altogether, our results indicated that GPCs acquired an epithelial-mesenchymal hybrid state through EMT-like morphogenetic changes.

Lineage trajectory of cGPCs

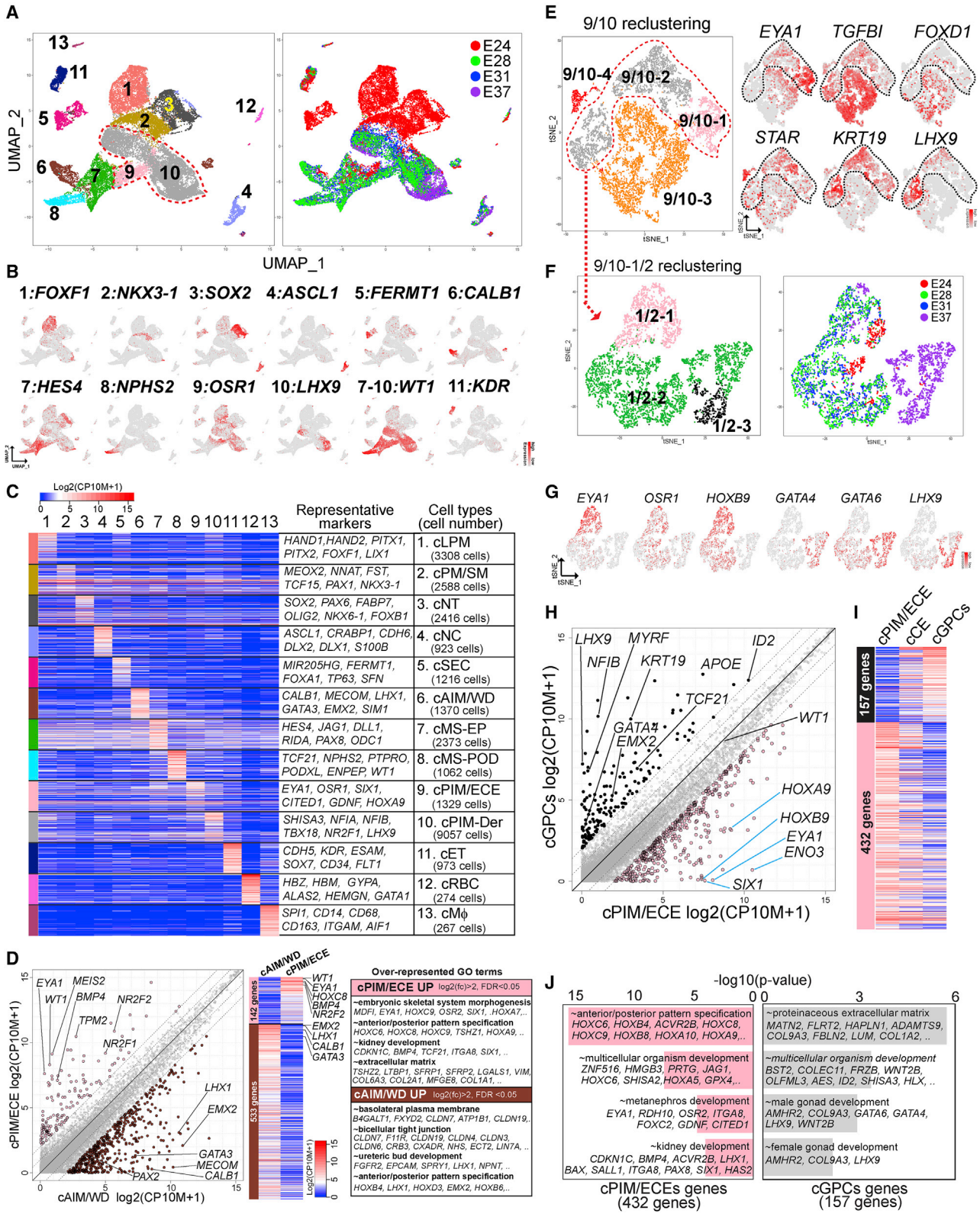
We next sought to understand the lineage trajectory of cynomolgus monkey GPCs (cGPCs) and the accompanying transcriptional changes. We focused on GPC specification and therefore used only female embryos. Female-specific gene expression, as marked by FOXL2 (Georges et al., 2013), started ~E40 (CS19) (Figure S6A); consequently, we used embryos at E28 (CS13), E31 (CS14), and E37 (CS18) for the collection of cGPCs and/or cynomolgus monkey CE (cCE) and E24 embryos for the collection of cynomolgus monkey PIM (cPIM)/ECE (Figures S6A–S6D; STAR Methods).

Of the ~27,585 cells for which transcriptomes were available, 27,156 cells remained for downstream analysis after removal of low-quality cells (Figures S6D and S6E). By profiling the expression of known marker genes in a uniform manifold approximation and projection (UMAP) plot, we identified 13 clusters (Figures 6A–6C; Table S2). Clusters 6–10 were tight and continuous, thus suggesting a close ontological relationship (Figure 6A). All clusters except cluster 6 were strongly positive for WT1 and consequently appeared to be derivatives/close kin of $WT1^+$ PIM (Figure 6B). Accordingly, cluster 9 expressed *OSR1*, *EYA1*, and *HOXA9*, markers of mPIM/ECE (Figures 4A and S4H), and consisted predominantly of cells from E24/E28; consequently it was annotated as cPIM/ECE (Figures 6A–6C). To the left of cPIM/ECE, three clusters were identified (Figure 6A). Cluster 6 expressed *CALB1*, *LHX1*, and *GATA3*, markers of the AIM (Figures S2B and S5A–S5E); consequently, it was annotated as cAIM/WD. Cluster 7 expressed *HES4*, *JAG1*, and *DLL1*, markers of renal vesicle/S-shaped bodies, and cluster 8 expressed *NPHS2* and *PODXL*, markers of podocytes; therefore, these clusters were termed cynomolgus monkey mesonephric early progenitors (cMS-EPs) and cynomolgus monkey mesonephric podocytes (cMS-PODs), respectively (Figures 6B and 6C). To the right of cluster 9, cluster 10 contained cells expressing *LHX9*, *TBX18*, and *NFIA*, indicative of CE/gonadal differentiation, and were tentatively annotated as cPIM derivatives (cPIM-Der) for further in-depth analyses, as described later (Figures 6A–6C).

To delineate the lineage progression of cGPCs, we focused our analyses on clusters 9 and 10. Re-clustering of these two clusters revealed $KRT19^+$ and $KRT19^-$ populations (Figure 6E; Table S2). Among the $KRT19^-$ clusters, cluster 9/10-3 expressed the nephrogenic stromal markers *TGFBI* and *FOXD1* (Combes et al., 2019; Li et al., 2014) and therefore likely represented mesonephric stroma; this cluster was named cynomolgus monkey mesonephric stroma (cMS-ST). Another $KRT19^-$ cluster, 9/10-4, strongly expressed *STAR*, a marker of the adrenal gland lineage, and consequently was named cynomolgus monkey fetal adrenal gland (cAD). Given our IF results showing that GPCs and their ancestries were $KRT19^+$, we focused on $KRT19^+$ clusters (9/10-1 and 9/10-2) for further analysis. Re-clustering analysis of the $KRT19^+$ population revealed a continuous lineage trajectory (Figure 6F). One end of the trajectory was cluster 1/2-1, which expressed *OSR1*, *EYA1*, and *HOXB9*, and consequently was (re)annotated as cPIM/ECE. On the opposite end of the trajectory was cluster 1/2-3, which expressed *GATA4*, *GATA6*, and *LHX9*, suggestive of a GPC

(D) (Top) Bright-field image of a cynomolgus monkey embryo at E24 (CS11) (left, posterior; right, anterior) indicating the approximate planes (a–f) where the sections were obtained (embryos n = 2). Scale bar, 1 mm. (Bottom) Transverse sections of the embryos at the indicated planes, stained for WT1 (green) and GATA3 (red) and with DAPI (white) (left), and schematic showing the regions with the AIM and/or PIM present (right). Scale bar, 100 μm.
(E) IF images of the coelomic angles (E28–E32) or the ovaries (E51) at the indicated stages for WT1 (green), NR5A1 (red), and GATA4 (cyan) merged with DAPI (white) (embryos n = 2, 2, 2, 2, 1). Scale bar, 100 μm.
(F) IF of E32 GR stained for LHX9 (green) and NR5A1 (red) merged with DAPI (white) (embryos n = 1). Scale bar, 100 μm.
(G) H&E-stained sections showing the coelomic angles of cynomolgus monkey embryos at the indicated stages (n = 1 for each stage). Scale bar, 50 μm.
(H) IF image of nascent GR at E31 for WT1 (green) and GATA4 (red) merged with bright-field images, highlighting the crowding and pseudostratification of $WT1^+GATA4^+$ GPCs (embryos n = 2). Scale bar, 50 μm.
(I) IF image of the coelomic angle at E28, with staining for laminin (green) and DAPI (white) (embryos n = 2). Laminin is thin and somewhat irregularly outlines the basement membranes of the coelomic epithelium. Co, coelom; MT, mesonephric tubule; WD, Wolffian duct. Scale bar, 100 μm.
(J) IF image of GR at E32 for laminin (green) and NR5A1 (red) merged with DAPI (white) (embryos n = 2). Arrowheads indicate the regions with basement membrane disruption. Scale bar, 100 μm. See also Figure S5.

Cell Reports
Resource
CellPress
OPEN ACCESS



(legend on next page)

identity, and consequently was annotated as cGPCs (Figure 6G). The presence of cGPCs bearing these marker genes in E36 ovaries was confirmed by qPCR and single-cell mRNA 3-prime end sequencing (SC3-seq) (Nakamura et al., 2015) (Figure S6F). Between cPIM/ECE and cGPCs was cluster 1/2-2, which expressed *KRT19* and gradually lost cPIM/ECE markers and gained cGPC markers along the trajectory (Figures 6F–6I). Because *KRT19* is specifically expressed in CE or GPCs, but not in mesonephric tubules or stroma, particularly at and after E28 (Figures S5H and S5I), we annotated these cells as cCE (Figures 6I, S7A, and S8E). These analyses revealed the identities and developmental connection of various lineages, particularly the progression of cGPCs from cPIM/ECE, as supported by pseudotime analysis and UHC (Figures S7B–S7D and S8A).

Distinct HOX genes dictate the regionalization of IM along the AP axis in cynomolgus monkey embryos

To identify the genetic determinants of IM regionalization along the AP axis in cynomolgus monkey embryos, we performed pairwise DEG analysis on cAIM/WD and cPIM/ECE. This analysis identified unique gene sets between these cell types, along with known and novel markers (Figure 6D; Table S3). Interestingly, several HOX genes were differentially expressed between cPIM/ECE and cAIM/WD (Figure 6D). *HOX* genes confer positional information, the HOX code, to emerging embryonic axial tissues through a timed activation sequence of clustered *HOX* genes from the 3′ to 5′ direction when mesodermal cells migrate from the posterior growth zone (Deschamps and Duboule, 2017; Neijts et al., 2014). Accordingly, the anterior/early progenitor population relays an early number of (anterior) *HOX* genes, whereas the posterior/late progenitor population transmits progressively later numbers of (posterior) *HOX* genes, along with earlier HOX genes, to their descendants. Previous studies have suggested that the MM has posterior HOX codes, such as *HOX11* genes (Lindström et al., 2018; Patterson et al., 2001; Taguchi et al., 2014). Therefore, we next explored the HOX codes of the cPIM/ECE-cGPC lineages to gauge the timing of the exit from the PS in ancestral cells compared with cAIM/WD or kidney cells. We found that cGPCs retained *HOX* gene expression, albeit with generally lower levels than cPIM/ECE (Figures 6H, 6J, and S8B). *HOX1–HOX9* were generally expressed in both cAIM/WD and cPIM/ECE-cGPCs. However, *HOX10*-expressing cells were less common in cAIM/WD than in cPIM/ECE-cGPCs, and *HOX11* genes were essentially not expressed in cAIM/WD, thus suggesting that *HOX9–HOX10* is the posterior boundary of the AIM/WD (Figures S8B and S8C). In contrast, some cPIM/ECE-cGPCs expressed *HOX11* (Figures S8B and S8C), but none of them expressed *HOX12* (data not shown), thus suggesting that *HOX11* is the posterior boundary of cPIM/ECE-cGPCs, which appears to overlap with the anterior boundary of the kidneys. Importantly, $HOX11^{+}$ GPCs were a minority among cGPCs (~10%), which nonetheless expressed cGPC marker genes similarly to $HOX11^{-}$ cGPCs (Figure S8D). These results also support the findings of a prior mouse lineage tracing study for T revealing that the posterior gonads and the kidneys are both derived from the posterior PS (Figures 1E, 1K, and 3H).

cGPC programs are highly conserved between mice and monkeys

We next explored the genetic pathways accompanying cGPC specification from cPIM/ECE. Pairwise/multigroup DEG and pseudotime analyses revealed genes uniquely regulated between stages (Figures 6H, 6I, S7A, S7C, and S7D; Table S3). A heatmap of these DEGs showed a gradual transition from cPIM/ECE to cGPCs via cCE (Figures 6I and S7A). Genes upregulated in cGPCs—cGPC genes (Figures 6I, 6J, and S7A) or group 3 genes (Figure S7D)—were enriched in GO terms such as "proteinaceous extracellular matrix" or "female gonad

Figure 6. Lineage trajectory of cGPCs

(A) UMAP plot of all cells based on computationally aggregated scRNA-seq data obtained from four samples (E24, E28, E31, and E37). Cells are colored according to 13 clusters identified by Seurat clustering analysis (left) or by embryonic days (right).

(B) UMAP feature plots of marker genes characteristic of the 13 clusters defined in (A).

(C) Heatmap showing the averaged expression pattern of the top 100 DEGs (by fold-change value) identified from a multigroup comparison (false discovery rate [FDR] < 0.01, fold change > 1 compared with other cell types) among cell types, representative DEGs (markers), and identities of cell types annotated according to analyses in (A)–(C). cLPM, cynomolgus monkey lateral plate mesoderm; cPM/SM, cynomolgus monkey paraxial mesoderm (PM)/somitic mesoderm (SM); cNT, cynomolgus monkey neural tube; cNC, cynomolgus monkey neural crest; cSEC, cynomolgus monkey surface ectoderm; cAIM/WD, cynomolgus monkey anterior intermediate mesoderm/WD; cMS-EP, cynomolgus monkey mesonephric epithelial progenitor; cMS-POD, cynomolgus monkey mesonephric podocyte; cPIM/ECE, cynomolgus monkey posterior intermediate mesoderm/early coelomic epithelium; cPIM-Der, cynomolgus monkey posterior intermediate mesoderm derivative; cET, cynomolgus monkey endothelium; cRBC, cynomolgus monkey red blood cell; cMΦ, cynomolgus monkey macrophage.

(D) Scatterplot comparison (left) and heatmap (middle) of the averaged values of DEGs between cAIM/WD and cPIM/ECE. Brown, 533 genes higher in cAIM/WD; pink, 142 genes higher in cPIM/ECE (more than 4-fold differences [flanking diagonal lines], FDR < 0.05). Key genes are annotated. Representative genes and their GO enrichments for DEGs are shown at the right.

(E) Re-clustering of clusters 9 and 10, defined in (A), projected on the T-distributed Stochastic Neighbor Embedding (tSNE) plot (left) and expression of key marker genes (right).

(F and G) Re-clustering of clusters 9/10-1 and 9/10-2, defined in (E) and projected on the tSNE plot. Cells are colored according to three clusters (1/2-1, 1/2-2, and 1/2-3) identified by Seurat clustering analysis (F, left) or embryonic days (F, right); clusters 1/2-1, 1/2-2, and 1/2-3 are annotated as cPIM/ECE, cCE, and cGPCs, respectively, on the basis of expression of marker genes projected on the tSNE plot (G).

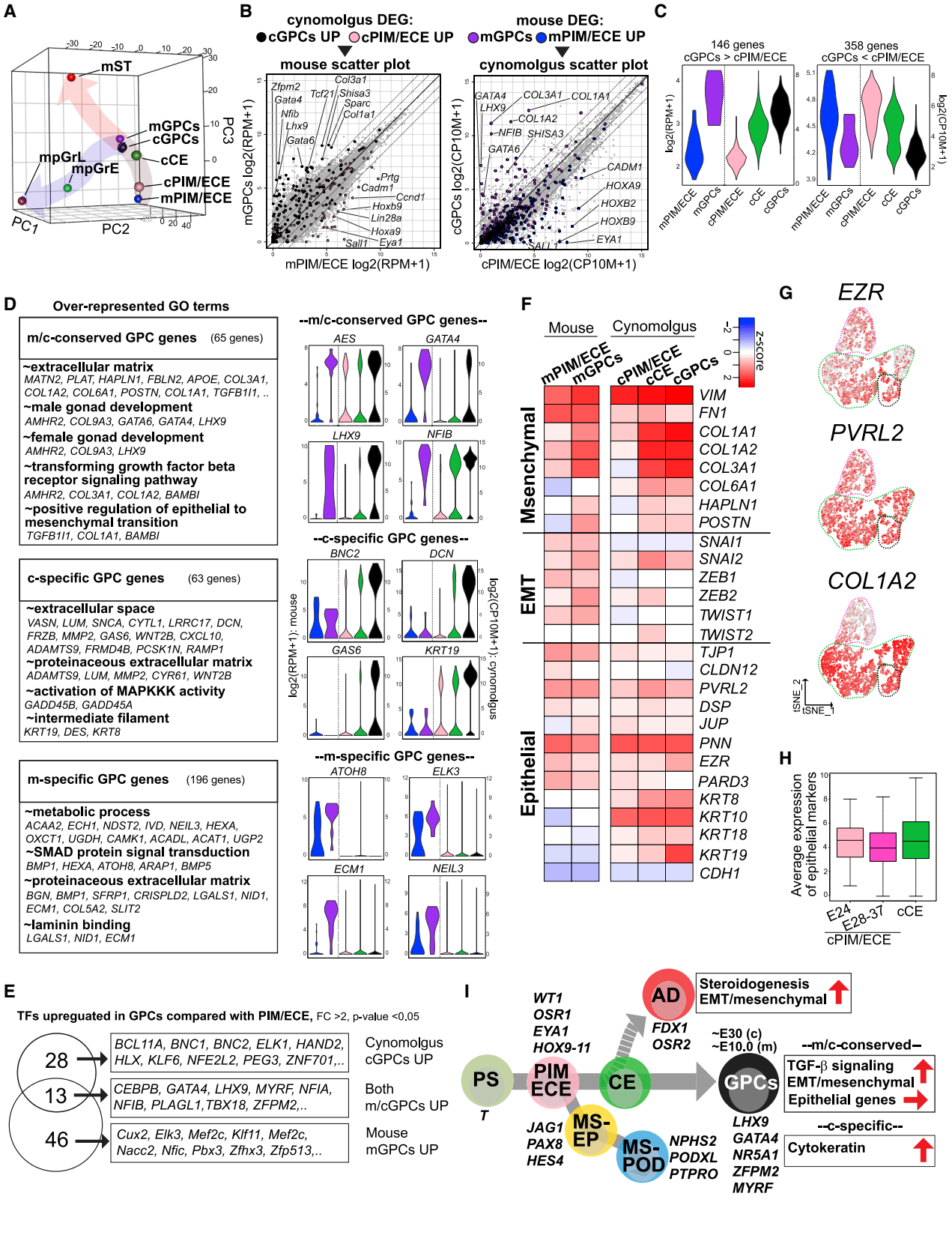
(H) Scatterplot showing the DEGs between cPIM/ECE and cGPCs. Genes upregulated by more than 4-fold (FDR < 0.05) in the indicated cell clusters are plotted with the colors in (A) and (F). Key genes are shown.

(I) Heatmap of the averaged expression of DEGs defined in (H) (432 and 157 genes upregulated in cPIM/ECE and cGPCs, respectively) in the indicated cell types defined in (F).

(J) GO analysis of DEGs (>4-fold differences, FDR < 0.05), as defined in (H).

See also Figures S6–S8.

Cell Reports
Resource
CellPress
OPEN ACCESS



(legend on next page)

development,'' thereby suggesting that gene programs similar to mGPCs are operative in cGPCs (Figures 4D, 6J, S4H, S7A, and S7D).

Plotting the average expression values of mouse and cynomolgus monkey cell types on a PCA space revealed the separation of the two species according to PC1 values. However, when these cell types were viewed through the PC2/PC3 plane, trajectories from the PIM/ECE to GPCs appeared highly conserved between species, with cCE situated between c/mPIM/ECE and c/mGPCs (Figure 7A). Projection of the DEGs between GPCs and PIM/ECE from one species to the other also revealed high conservation (Figures 7B and 7C). For example, many known and previously poorly characterized GPC markers were upregulated in both mGPCs and cGPCs, including *LHX9*, *GATA4*, *TCF21*, *SHISA3*, *NFIA*, *ZFPM2*, and *COL1A1*, which were enriched in GO terms such as ''female gonad development'' or ''extracellular matrix'' (Figures 7B and 7D; Table S4). Many mesenchymal genes that were upregulated as GPCs progressed from c/mPIM/ECE, and this was accompanied by the upregulation of key EMT transcription factor genes, such as *SNAI2*, *ZEB2*, or *TWIST2*, supporting an EMT-like change (Dong et al., 2018; Skrypek et al., 2017). Many epithelial markers were also expressed in the PIM/ECE and were generally maintained during the course of GPC progression from the PIM/ECE (Figures 7F–7H). Moreover, CDH1 (E-cadherin), a prototypical marker of epithelial cells, remained unexpressed throughout GPC progression (Figure 7F). Altogether, these observations suggest that GPC specification in both mice and primates involves the acquisition of an epithelial-mesenchymal hybrid state that differs from canonical EMT, in which the acquisition of mesenchymal features is accompanied by a loss of epithelial features.

We also observed species-specific differences in the GPC program. For example, some genes involved in fatty-acid oxidization or carbohydrate metabolism were mGPC specific, a finding that might have evolutionary importance (Figure 7D). In contrast, cGPC-specific genes were enriched in the GO term ''intermediate filament,'' thus reflecting the expression of *KRT8* or *KRT19* in cGPCs (Figure 7D). This analysis also revealed a set of transcription factors with conserved or divergent expression patterns between species (Figure 7E). Among them, *MYRF* has been identified as a candidate gene for some disorders of sex development (DSD) cases, but its function in murine gonadogenesis has not been addressed (Hamanaka et al., 2019). Collectively, these findings suggest that although the overall GPC program is highly conserved, some species-specific gene expression exists.

Early adrenal gland fate determination accompanies a gene expression program

The adrenal glands are closely related to GPCs (Figures 1E, 3C, 3H, and S1G). DEGs upregulated in cAD included genes enriched in GO terms such as ''cholesterol metabolic process'' and ''adrenal gland development'' (Figures S8E and S8F; Table S3) (del Valle et al., 2017), or Jensen Diseases terms such as ''congenital adrenal gland insufficiency,'' thereby suggesting functional importance in human adrenal gland development (Figure S8G). Pairwise DEGs between cAD and cPIM/ECE revealed 95 gene, that were upregulated in cGPCs, but not cAD (compared with cPIM/ECE), including *LHX9*, *GATA4*, and *WT1* (Figures S8H and S8I; Table S3), thus suggesting that loss of these key transcription factors might influence the segregation of the adrenal glands from the gonadal lineage (Bandiera et al., 2013). In addition, the gene expression of cAD was most highly correlated with cCE rather than cPIM/ECE or cGPCs (Figure S8J), thus raising the possibility that cAD might be segregated from gonadal lineages before cGPCs are specified.

DISCUSSION

In this study, we spatially and temporally mapped the developmental connections of various IM lineages. We found that anteriorly located IM derivatives (UB, adrenal glands, and anterior gonads/mesonephros) exit the PS (marked by loss of *T*) and activate IM markers earlier than posteriorly located IM derivatives (MM and posterior gonads/mesonephros) (Figure 3H). The observation that gonads/mesonephros are derived from both early and late PS derivatives is reflected by their long, narrow structure along the AP axis (Harikae et al., 2013). The

Figure 7. cGPC programs are highly conserved between mice and monkeys

(A) PCA of averaged transcriptome values for cell types defined in Figures 4A and 6F. Genes annotated in both mice and cynomolgus monkeys were used for analyses.

(B) Scatterplots comparing the averaged gene expression levels between mPIM/ECE and mGPCs (left) or cPIM/ECE and cGPCs (right). Genes upregulated in cGPCs (black) and cPIM/ECE (pink), as defined in Figure 6H (4-fold differences, FDR < 0.05), are projected on the scatterplot at the left. Genes upregulated in mGPCs (purple) and mPIM/ECE (blue), as defined in Figure S4H (4-fold differences, FDR < 0.01), are projected on the scatterplot at the right. Key genes are shown.

(C) Violin plots of the expression values of 146 genes (cGPCs > cPIM/ECE, 4-fold differences, FDR < 0.05) (left) or 358 genes (cGPCs < cPIM/ECE, 4-fold differences, FDR < 0.05) (right) in the indicated mouse and cynomolgus monkey cell types. Among the DEGs defined in Figure 6H, genes annotated for both mice and cynomolgus monkeys were used (Table S6).

(D) Representative genes upregulated in GPCs compared with PIM/ECE in both mice and cynomolgus monkeys (top, m/c-conserved GPC genes, 65 genes), cynomolgus monkeys only (middle, c-specific GPC genes, 63 genes), or mice only (bottom, m-specific GPC genes, 196 genes), with their over-represented GO terms (left). Violin plots for the expression of representative genes in mouse and cynomolgus monkey cell types, colored as in Figures 4A and 6F.

(E) Overlap between mouse and cynomolgus monkey transcription factors upregulated in GPCs compared with the PIM/ECE (>2-fold differences, p < 0.05).

(F) Heatmap showing the averaged expression of mesenchymal and epithelial marker genes or EMT-related transcription factors in the indicated cell types in mice and cynomolgus monkeys.

(G) Expression of key epithelial (*EZR* and *PVRL2*) or mesenchymal markers (*COL1A2*) projected on the tSNE plot shown in Figure 6F.

(H) Boxplot showing the averaged expression of epithelial markers as in (F) in cPIM/ECE (derived from E24 or E28–E37 embryos) and cCE (from all embryos).

(I) Model of GPC specification.

See also Figures S6–S8.

morphology of the gonads is a conserved feature across multiple taxa and might confer some advantages in reproductive fitness, e.g., ensuring receipt of migrating primordial germ cells (PGCs), which are widely scattered throughout the hindgut.

Using lineage tracing, IF, and transcriptomics, we determined that gonads originate from $WT1^+$ PIM (Figure 7I). The anterior end of the $WT1^+$ region at E9.0 and most $WT1^+$ regions at E9.5 showed $WT1^+$ ECE as a continuous structure abutting the PIM (Figures 2D and S2D). We believe that the ECE is an early PIM descendant rather than a tissue of different origin, partly because it is (1) seen only in the anterior (more mature) region, (2) anatomically contiguous with the PIM, and (3) $WT1^+FOXF1^-$ (Figures 2D, S2A, and S2D). Importantly, $WT1^+$ cells randomly isolated from E9.5 embryos on the basis of $WT1^+/OSR1^+$ status revealed a highly homogeneous cell cluster in scRNA-seq, without readily identifiable subclusters (Figures 4A, 4B, S4D, and S4E). Moreover, PCA showed that this cluster seamlessly transitioned into the GPC cluster (Figures 4B and S4D). These observations suggest that the PIM and ECE are highly similar or identical transcriptionally, despite some morphologic distinctions.

By E10.0–E10.5, the PIM/ECE further expands and forms the mesonephros and overlying CE containing GPCs (Figures 2C–2E). The timing and mechanisms of segregation between the mesonephros and the GPC remain unknown. We demonstrated that a portion of the PIM expresses SIX2 and partially contributes to mesonephros, but not the gonads, according to lineage tracing analysis (Figures S3E and S3F). Although lineage segregation likely starts within the PIM, our comprehensive analysis provides only a snapshot within the continuous progression of GPC specification. Fortunately, the emergence of the PIM to GPC specification occurs within ~1 day in mice; therefore, time-lapse imaging of slice cultures might enable visualization of the lineage segregation of GPCs and mesonephros within the PIM at high resolution in the future.

Recent studies have highlighted the importance of fetal gonadal somatic cells for *in vitro* gametogenesis, because in both mice and humans, an oogonia/spermatogonia-like state can be induced from pluripotent stem cells if germ-cell-like cells are cultured in the presence of fetal gonadal somatic cells (Hikabe et al., 2016; Hwang et al., 2020; Ishikura et al., 2016; Morohaku et al., 2016; Yamashiro et al., 2018). The induction of gonadal somatic cells from hiPSCs would enable the induction of human gametes in a scalable and clinically applicable manner. Accordingly, the spatial and temporal dynamics of the GPC specification process illustrated at high resolution in this study should provide critical reference information for the directional induction of gonadal lineages *in vitro*. Moreover, our comparative transcriptomics should provide resources for identifying the mechanisms of early gonadogenesis and diagnosis and treatment of DSD.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.109075.

ACKNOWLEDGMENTS

We thank L. King for carefully reviewing the manuscript and providing insightful comments. We thank C. Lengner, M. Takasato, and members of the Saitou and Sasaki labs for discussions. We are grateful to the Center for Anatomical, Pathological and Forensic Medical Researches, Kyoto University, for the preparation of paraffin sections. We thank Dr. Kenji Osafune at Kyoto University for providing us with 3D45 OSR1-EGFP hiPSCs. This work was supported in part by a JST-ERATO grant (JPMJER1104), Grant-in-Aid for Specially Promoted Research from JSPS (17H06098) and Pythias Fund (to M.S.), and the Open Philanthropy fund from the Silicon Valley Community Foundation (2019-197906 to K.S.).

AUTHOR CONTRIBUTIONS

K.S. conceived the project. K.S. and M.S. supervised the projects. K.S. designed and conducted the overall experiments and analyzed the data. K.S. wrote the manuscript. K.S., K.C., A.O., Y.M., T.Y., and Y.Y. contributed to the processing and analyses of scRNA-seq data. K.S., C.I., H.T., I.O., and Y.S. contributed to the isolation of cynomolgus monkey embryos. K.S. and H.O. contributed to mouse breeding.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 6, 2020
Revised: January 21, 2021
Accepted: April 12, 2021
Published: May 4, 2021

REFERENCES

Bandiera, R., Vidal, V.P.I., Motamedi, F.J., Clarkson, M., Sahut-Barnola, I., von Gise, A., Pu, W.T., Hohenstein, P., Martinez, A., and Schedl, A. (2013). WT1 maintains adrenal-gonadal primordium identity and marks a population of AGP-like progenitors within the adrenal gland. Dev. Cell *27*, 5–18.

Birk, O.S., Casiano, D.E., Wassif, C.A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J.A., Parker, K.L., et al. (2000). The LIM homeobox gene Lhx9 is essential for mouse gonad formation. Nature *403*, 909–913.

Brambell, F.W.R. (1927). The development and morphology of the gonads of the mouse.—Part I. The morphogenesis of the indifferent gonad and of the ovary. Proc. R. Soc. Lond., B *101*, 391–409.

Capel, B. (2017). Vertebrate sex determination: evolutionary plasticity of a fundamental switch. Nat. Rev. Genet. *18*, 675–689.

Cazzulino, A.S., Martinez, R., Tomm, N.K., and Denny, C.A. (2016). Improved specificity of hippocampal memory trace labeling. Hippocampus *26*, 752–762.

Chesley, P. (1935). Development of the short-tailed mutant in the house mouse. J. Exp. Zool. *70*, 429–459.

Combes, A.N., Phipson, B., Lawlor, K.T., Dorison, A., Patrick, R., Zappia, L., Harvey, R.P., Oshlack, A., and Little, M.H. (2019). Single cell analysis of the developing mouse kidney provides deeper insight into marker gene expression and ligand-receptor crosstalk. Development *146*, dev178673.

Cui, S., Ross, A., Stallings, N., Parker, K.L., Capel, B., and Quaggin, S.E. (2004). Disrupted gonadogenesis and male-to-female sex reversal in Pod1 knockout mice. Development *131*, 4095–4105.

del Valle, I., Buonocore, F., Duncan, A.J., Lin, L., Barenco, M., Parnaik, R., Shah, S., Hubank, M., Gerrelli, D., and Achermann, J.C. (2017). A genomic atlas of human adrenal and gonad development. Wellcome Open Research *2*, 25.

Deschamps, J., and Duboule, D. (2017). Embryonic timing, axial stem cells, chromatin dynamics, and the Hox clock. Genes Dev. *31*, 1406–1416.

Dong, J., Hu, Y., Fan, X., Wu, X., Mao, Y., Hu, B., Guo, H., Wen, L., and Tang, F. (2018). Single-cell RNA-seq analysis unveils a prevalent epithelial/mesenchymal hybrid state during mouse organogenesis. Genome Biol. *19*, 31.

Fujimoto, Y., Tanaka, S.S., Yamaguchi, Y.L., Kobayashi, H., Kuroki, S., Tachibana, M., Shinomura, M., Kanai, Y., Morohashi, K.-I., Kawakami, K., and Nishinakamura, R. (2013). Homeoproteins Six1 and Six4 regulate male sex determination and mouse gonadal development. Dev. Cell *26*, 416–430.

Georges, A., Auguste, A., Bessière, L., Vanet, A., Todeschini, A.L., and Veitia, R.A. (2013). FOXL2: a central transcription factor of the ovary. J. Mol. Endocrinol. *52*, R17–R33.

Grote, D., Souabni, A., Busslinger, M., and Bouchard, M. (2006). Pax 2/8-regulated Gata 3 expression is necessary for morphogenesis and guidance of the nephric duct in the developing kidney. Development *133*, 53–61.

Haghverdi, L., Büttner, M., Wolf, F.A., Buettner, F., and Theis, F.J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. Nat. Methods *13*, 845–848.

Hamanaka, K., Takata, A., Uchiyama, Y., Miyatake, S., Miyake, N., Mitsuhashi, S., Iwama, K., Fujita, A., Imagawa, E., Alkanaq, A.N., et al. (2019). MYRF haploinsufficiency causes 46,XY and 46,XX disorders of sex development: bioinformatics consideration. Hum. Mol. Genet. *28*, 2319–2329.

Harikae, K., Miura, K., and Kanai, Y. (2013). Early gonadogenesis in mammals: significance of long and narrow gonadal structure. Dev. Dyn. *242*, 330–338.

Hikabe, O., Hamazaki, N., Nagamatsu, G., Obata, Y., Hirao, Y., Hamada, N., Shimamoto, S., Imamura, T., Nakashima, K., Saitou, M., and Hayashi, K. (2016). Reconstitution *in vitro* of the entire cycle of the mouse female germ line. Nature *539*, 299–303.

Hu, Y.-C., Okumura, L.M., and Page, D.C. (2013). Gata4 is required for formation of the genital ridge in mice. PLoS Genet. *9*, e1003629.

Hwang, Y.S., Suzuki, S., Seita, Y., Ito, J., Sakata, Y., Aso, H., Sato, K., Hermann, B.P., and Sasaki, K. (2020). Reconstitution of prospermatogonial specification *in vitro* from human induced pluripotent stem cells. Nat Commun. *11*, 5656.

Imuta, Y., Kiyonari, H., Jang, C.-W., Behringer, R.R., and Sasaki, H. (2013). Generation of knock-in mice that express nuclear enhanced green fluorescent protein and tamoxifen-inducible Cre recombinase in the notochord from Foxa2 and T loci. Genesis *51*, 210–218.

Ishikura, Y., Yabuta, Y., Ohta, H., Hayashi, K., Nakamura, T., Okamoto, I., Yamamoto, T., Kurimoto, K., Shirane, K., Sasaki, H., and Saitou, M. (2016). *In Vitro* Derivation and Propagation of Spermatogonial Stem Cell Activity from Mouse Pluripotent Stem Cells. Cell Rep. *17*, 2789–2804.

James, R.G., Kamei, C.N., Wang, Q., Jiang, R., and Schultheiss, T.M. (2006). Odd-skipped related 1 is required for development of the metanephric kidney and regulates formation and differentiation of kidney precursor cells. Development *133*, 2995–3004.

Karl, J., and Capel, B. (1998). Sertoli cells of the mouse testis originate from the coelomic epithelium. Dev. Biol. *203*, 323–333.

Kobayashi, A., Valerius, M.T., Mugford, J.W., Carroll, T.J., Self, M., Oliver, G., and McMahon, A.P. (2008). Six2 defines and regulates a multipotent self-renewing nephron progenitor population throughout mammalian kidney development. Cell Stem Cell *3*, 169–181.

Kreidberg, J.A., Sariola, H., Loring, J.M., Maeda, M., Pelletier, J., Housman, D., and Jaenisch, R. (1993). WT-1 is required for early kidney development. Cell *74*, 679–691.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44* (*W1*), W90–W97.

Kusaka, M., Katoh-Fukui, Y., Ogawa, H., Miyabayashi, K., Baba, T., Shima, Y., Sugiyama, N., Sugimoto, Y., Okuno, Y., Kodama, R., et al. (2010). Abnormal epithelial cell polarity and ectopic epidermal growth factor receptor (EGFR) expression induced in Emx2 KO embryonic gonads. Endocrinology *151*, 5893–5904.

Li, W., Hartwig, S., and Rosenblum, N.D. (2014). Developmental origins and functions of stromal cells in the normal and diseased mammalian kidney. Dev. Dyn. *243*, 853–863.

Lindström, N.O., McMahon, J.A., Guo, J., Tran, T., Guo, Q., Rutledge, E., Parvez, R.K., Saribekyan, G., Schuler, R.E., Liao, C., et al. (2018). Conserved and divergent features of human and mouse kidney organogenesis. J. Am. Soc. Nephrol. *29*, 785–805.

Lun, A.T.L., McCarthy, D.J., and Marioni, J.C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Res. *5*, 2122.

Luo, X., Ikeda, Y., and Parker, K.L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. Cell *77*, 481–490.

Mae, S.I., Shono, A., Shiota, F., Yasuno, T., Kajiwara, M., Gotoda-Nishimura, N., Arai, S., Sato-Otubo, A., Toyoda, T., Takahashi, K., et al. (2013). Monitoring and robust induction of nephrogenic intermediate mesoderm from human pluripotent stem cells. Nat. Commun. *4*, 1367.

Miyamoto, N., Yoshida, M., Kuratani, S., Matsuo, I., and Aizawa, S. (1997). Defects of urogenital development in mice lacking Emx2. Development *124*, 1653–1664.

Morohaku, K., Tanimoto, R., Sasaki, K., Kawahara-Miki, R., Kono, T., Hayashi, K., Hirao, Y., and Obata, Y. (2016). Complete *in vitro* generation of fertile oocytes from mouse primordial germ cells. Proc. Natl. Acad. Sci. USA *113*, 9021–9026.

Mugford, J.W., Sipilä, P., McMahon, J.A., and McMahon, A.P. (2008). Osr1 expression demarcates a multi-potent population of intermediate mesoderm that undergoes progressive restriction to an Osr1-dependent nephron progenitor compartment within the mammalian kidney. Dev. Biol. *324*, 88–98.

Müller, F., and O'Rahilly, R. (2004). The primitive streak, the caudal eminence and related structures in staged human embryos. Cells Tissues Organs *177*, 2–20.

Nakamura, T., Yabuta, Y., Okamoto, I., Aramaki, S., Yokobayashi, S., Kurimoto, K., Sekiguchi, K., Nakagawa, M., Yamamoto, T., and Saitou, M.

(2015). SC3-seq: a method for highly parallel and quantitative measurement of single-cell gene expression. Nucleic Acids Res. *43*, e60.

Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T., and Saitou, M. (2016). A developmental coordinate of pluripotency among mice, monkeys and humans. Nature *537*, 57–62.

Neijts, R., Simmini, S., Giuliani, F., van Rooijen, C., and Deschamps, J. (2014). Region-specific regulation of posterior axial elongation during vertebrate embryogenesis. Dev. Dyn. *243*, 88–98.

Niu, W., and Spradling, A.C. (2020). Two distinct pathways of pregranulosa cell differentiation support follicle formation in the mouse ovary. Proc. Natl. Acad. Sci. USA *117*, 20015–20026.

Ohinata, Y., Sano, M., Shigeta, M., Yamanaka, K., and Saitou, M. (2008). A comprehensive, non-invasive visualization of primordial germ cell development in mice by the Prdm1-mVenus and Dppa3-ECFP double transgenic reporter. Reproduction *136*, 503–514.

Patterson, L.T., Pembaur, M., and Potter, S.S. (2001). Hoxa11 and Hoxd11 regulate branching morphogenesis of the ureteric bud in the developing kidney. Development *128*, 2153–2161.

Robinson, S.P., Langan-Fahey, S.M., Johnson, D.A., and Jordan, V.C. (1991). Metabolites, pharmacodynamics, and pharmacokinetics of tamoxifen in rats and mice compared to the breast cancer patient. Drug Metab. Dispos. *19*, 36–43.

Sajithlal, G., Zou, D., Silvius, D., and Xu, P.X. (2005). Eya 1 acts as a critical regulator for specifying the metanephric mesenchyme. Dev. Biol. *284*, 323–336.

Self, M., Lagutin, O.V., Bowling, B., Hendrix, J., Cai, Y., Dressler, G.R., and Oliver, G. (2006). Six2 is required for suppression of nephrogenesis and progenitor renewal in the developing kidney. EMBO J. *25*, 5214–5228.

Skrypek, N., Goossens, S., De Smedt, E., Vandamme, N., and Berx, G. (2017). Epithelial-to-Mesenchymal Transition: Epigenetic Reprogramming Driving Cellular Plasticity. Trends Genet. *33*, 943–959.

Stévant, I., Neirijnck, Y., Borel, C., Escoffier, J., Smith, L.B., Antonarakis, S.E., Dermitzakis, E.T., and Nef, S. (2018). Deciphering Cell Lineage Specification during Male Sex Determination with Single-Cell RNA Sequencing. Cell Rep. *22*, 1589–1599.

Stévant, I., Kühne, F., Greenfield, A., Chaboissier, M.C., Dermitzakis, E.T., and Nef, S. (2019). Dissecting Cell Lineage Specification and Sex Fate Determination in Gonadal Somatic Cells Using Single-Cell Transcriptomics. Cell Rep. *26*, 3272–3283.e3.

Taguchi, A., Kaku, Y., Ohmori, T., Sharmin, S., Ogawa, M., Sasaki, H., and Nishinakamura, R. (2014). Redefining the in vivo origin of metanephric nephron progenitors enables generation of complex kidney structures from pluripotent stem cells. Cell Stem Cell *14*, 53–67.

Takasato, M., and Little, M.H. (2015). The origin of the mammalian kidney: implications for recreating the kidney *in vitro*. Development *142*, 1937–1947.

Wang, Q., Lan, Y., Cho, E.S., Maltby, K.M., and Jiang, R. (2005). Odd-skipped related 1 (Odd 1) is an essential regulator of heart and urogenital development. Dev. Biol. *288*, 582–594.

Wilson, J.F., and Erlandsson, R. (1998). Sexing of human and other primate DNA. Biol. Chem. *379*, 1287–1288.

Yamasaki, J., Iwatani, C., Tsuchiya, H., Okahara, J., Sankai, T., and Torii, R. (2011). Vitrification and transfer of cynomolgus monkey (Macaca fascicularis) embryos fertilized by intracytoplasmic sperm injection. Theriogenology *76*, 33–38.

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

Yang, Y., Workman, S., and Wilson, M. (2018). The molecular pathways underlying early gonadal development. J. Mol. Endocrinol. *62*, R47–R64.

Zhou, B., Ma, Q., Rajagopal, S., Wu, S.M., Domian, I., Rivera-Feliciano, J., Jiang, D., von Gise, A., Ikeda, S., Chien, K.R., and Pu, W.T. (2008). Epicardial progenitors contribute to the cardiomyocyte lineage in the developing heart. Nature *454*, 109–113.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
AlexaFluor 488 conjugated donkey anti-rabbit IgG	Life Technologies	A21206; RRID:AB_2535792
AlexaFluor 488 conjugated donkey anti-rat IgG	Life Technologies	A21208; RRID:AB_2535794
AlexaFluor 488 conjugated donkey anti-mouse IgG	Life Technologies	A32766; RRID:AB_2762823
AlexaFluor 568 conjugated donkey anti-mouse IgG	Life Technologies	A10037; RRID:AB_2534013
AlexaFluor 568 conjugated donkey anti-rabbit IgG	Life Technologies	A10042; RRID:AB_2534017
AlexaFluor 647 conjugated donkey anti-goat IgG	Life Technologies	A32849; RRID:AB_2762840
AlexaFluor 647 conjugated donkey anti-rabbit IgG	Life Technologies	A31573; RRID:AB_2536183
Goat anti-FOXF1	R&D Systems	AF4798; RRID:AB_2105588
Goat anti-FOXL2	Novus Biologicals	NB100-1277; RRID:AB_2106188
Goat anti-GATA4	Santa Cruz Biotechnology	sc-1237; RRID:AB_2108747
Goat anti-LHX9	Santa Cruz Biotechnology	sc-19348; RRID:AB_2135970
Goat anti-AMH	Santa Cruz Biotechnology	sc-6886; RRID:AB_649207
Mouse anti-COUP-TFII/NR2F2	R&D Systems	PP-H7147-00; RRID:AB_2155627
Mouse anti-DDX4	Abcam	ab27591; RRID:AB_11139638
Mouse anti-GATA3	Biocare Medical	CM405A; RRID:AB_10895444
Mouse anti-NR5A1	Perseus Proteomics	PP-N1665; RRID:AB_1962633
Mouse anti-LHX1	DSHB	4F2; RRID:AB_531784
Mouse anti-KRT5/6	DAKO	M723729-2
Mouse anti-Pan Cytokeratin (AE1/AE3)	Biocare Medical	ACR011A
Mouse anti-SALL1	R&D Systems	PP-K9814-00; RRID:AB_2183228
Mouse anti-SALL4	Biocare Medical	CM384A; RRID:AB_10582500
Mouse anti-VIM	Thermofisher Scientific	14-9897-82; RRID:AB_10597910
Rabbit anti-CDH1	Cell Signaling	3195S; RRID:AB_2291471
Rabbit anti-HSD3B1	Cosmo Bio	KAL-KO607; RRID:AB_2722746
Rabbit anti-LAMININ	Abcam	ab11575; RRID:AB_298179
Rabbit anti-SIX2	proteintech	11562-1-AP; RRID:AB_2189084
Rabbit anti-mouse T	Abcam	ab209665; RRID:AB_2750925
Rabbit anti-human T	Santa Cruz Biotechnology	sc-20109; RRID:AB_2255702
Rabbit anti-WT1	Abcam	ab89901; RRID:AB_2043201
Rabbit anti-KRT18	Abcam	ab133263; RRID:AB_11155892
Rabbit anti-KRT19	Abcam	ab52625; RRID:AB_2281020
Rabbit anti-DDX4	Abcam	ab13840; RRID:AB_443012
Rabbit anti-PAX2	Biolegend	901002; RRID:AB_2734656
Rat anti-KRT8	DSHB	TROMA-1
Rat anti-GFP	Nacalai Tesque	04404-26
Biological samples		
Cynomolgus fetal samples	Shiga University of Medical Science	N/A
Human embryo sample	Daigo Watanabe Clinic	N/A
Chemicals, peptides, and recombinant proteins		
iMatrix-511 (recombinant laminin-511 E8)	Nippi	892 012
Y-27632	Wako Pure Chemical Industries	259-00613
CHIR99021	Biovision	1677-5

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Retinoic acid	Sigma	R2625
Recombinant human FGF9	R&D Systems	273-F9
Recombinant human/murine/rat Activin-A	Peprotech	120-14
Recombinant human BMP4	HumanZyme	HZ-1078
Critical commercial assays		
Chromium Single Cell 3′ Library & Gel Bead Kit v2	10X genomics	120267
Deposited data		
Cynomolgus embryo 10X Genomics scRNA-seq data	This paper	GEO: GSE160043
Cynomolgus embryo SC3-seq data	This paper	GEO: GSE160049
Mouse embryo SC3-seq data	This paper	GEO: GSE160049
Experimental models: Cell lines		
3D45 (OSR1-EGFP knock-in hiPSCs	(Mae et al., 2013)	N/A
Experimental Models: Organisms/Strains		
Osr1$^{Egfp/+}$ mice	Dr. Ryuichi Nishinakamura (Kumamoto University)	N/A
T$^{nEGFP\text{-}CreERT2/+}$ mice	Dr. Hiroshi Sasaki (Osaka University)	Acc. No. CB0604K
Dppa3-Ecfp transgenic mice	Ohinata et al., 2008	N/A
Osr1$^{tm1(EGFP/cre/ERT2)Amc}$ mice	Jackson Laboratory	009061
Six2$^{tm3(EGFP/cre/ERT2)Amc}$ mice	Jackson Laboratory	009600
Wt1$^{CreERT2/+}$ mice	Jackson Laboratory	010912
Gt(ROSA)26Sortm9(CAG-tdTomato)Hze mice	Jackson Laboratory	007909
Oligonucleotides		
See Table S5 for primers used in this study	This paper	Table S5
Software and algorithms		
FACSDiva software	BD Biosciences	N/A
ImageJ / Fiji	National Institutes of Health (NIH)	N/A
Cell Ranger v3.1.0	10X Genomics, https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest	N/A
R (v3.6.1)	https://www.R-project.org	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to, and will be fulfilled by, the lead contact, Kotaro Sasaki (ksasaki@upenn.edu).

Materials availability

cDNAs generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

Data and code availability

Accession numbers generated in this study are GSE160049 (SC3-seq data for mouse embryos and cynomolgus monkey GRs at E36) and GSE160043 (10x Chromium scRNA-seq data for cynomolgus monkey embryos). The codes used for pseudotime analysis are available at GitHub repository: https://doi.org/10.5281/zenodo.4632324.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

Animal procedures were conducted in compliance with the ethical guidelines of Kyoto University and Shiga University of Medical Science. *Osr1*$^{Egfp/+}$ mice were obtained from Dr. Ryuichi Nishinakamura (Kumamoto University) (Taguchi et al., 2014).

T$^{nEGFP\text{-}CreERT2/+}$ (referred to as *T*$^{creERT2/+}$) mice were obtained from Dr. Hiroshi Sasaki (Osaka University, Acc. No. CB0604K) (Imuta et al., 2013). *Osr1*$^{tm1(EGFP/cre/ERT2)Amc}$ (referred to as *Osr1*$^{creERT2/+}$) mice (Mugford et al., 2008), *Six2*$^{tm3(EGFP/cre/ERT2)Amc}$ (referred to as *Six2*$^{creERT2/+}$) mice (Kobayashi et al., 2008), *Wt1*$^{CreERT2/+}$ mice (Zhou et al., 2008), and *Gt(ROSA)26Sortm9(CAG-tdTomato)Hze* (referred to as *Rosa26-LTL-tdTomato*) mice were purchased from the Jackson Laboratory. *Dppa3-Ecfp* transgenic mice were generated as described previously (Ohinata et al., 2008). All mice were maintained on a largely C57BL/6 background. For lineage tracing experiments, homozygous *Rosa26-LTL-tdTomato* mice (female, aged 6-16 weeks) were crossed with heterozygous Cre-driver mice (male, aged 7-18 weeks). To obtain *Osr1*$^{Egfp/+}$ or *Dppa3-Ecfp* transgenic embryos, C57BL6 mice (female, aged 6-16 weeks) were crossed with heterozygous *Osr1*$^{Egfp/+}$ or *Dppa3-Ecfp* transgenic mice (male, aged 7-18 weeks), respectively.

Mice were housed under a 12:12 h light:dark cycle at 22–26 °C, with a humidity of 40%–60%.

Collection of human embryo samples

Fetal urogenital organs including the ovaries, mesonephros, adrenal glands and kidneys at 8 wks and 3 days of gestation were obtained from a donor undergoing elective abortion at the Daigo Watanabe Clinic. All experimental procedures were approved by the IRB at Kyoto University (G1047). Informed consent was obtained from all human participants. The sex of the fetus was female, which was determined by sex-specific PCR on genomic DNA isolated from trunk tissues by using primers for the ZFX/ZFY loci (Table S5) (Wilson and Erlandsson, 1998). Fetal urogenital organs were dissected *en bloc* in RPMI-1640 (Roche), and images were taken under a dissection stereomicroscope while maintaining the topological orientation of the organs.

Collection of cynomolgus monkey embryo and fetal samples

Procedures using cynomolgus monkeys were approved by the Animal Care and Use Committee of the Shiga University of Medical Science. The series of assisted reproductive technologies using cynomolgus monkeys, including oocyte collection, intra-cytoplasmic sperm injection, pre-implantation embryo culture and transfer of pre-implantation embryos into foster mothers, were as reported previously (Yamasaki et al., 2011). The light cycle was 12 h of artificial light from 8 a.m. to 8 p.m. Water was available *ad libitum*. Temperature and humidity in the animal rooms were maintained at 23-27 °C and 45%–55%, respectively.

Implanted embryos were scanned with transabdominal ultrasound monitoring and recovered by cesarean section under full anesthesia. A total of 28 embryos were harvested ranging in embryonic age from E23 to E51 (embryos n = 8 [male], 18 [female], 2 [unknown]). The foster mothers were maintained after surgery. The sex of each sampled fetus was determined by sex-specific PCR with primers targeting the ZFX/ZFY loci (Table S5) (Wilson and Erlandsson, 1998). Embryos were dissected in RPMI-1640 (Roche).

Culture of hiPSCs

OSR1-EGFP hiPSCs (clone 3D45, gift from Dr. Kenji Osafune, Kyoto University) were cultured on plates coated with recombinant laminin-511 E8 (iMatrix-511, Nippi, Japan) and were maintained under feeder-free conditions in StemFit AK03N medium (Ajinomoto, Japan) at 37°C under an atmosphere of 5% CO_2 in air (Mae et al., 2013). Before passaging or induction of differentiation, hiPSC cultures were treated with a 1:1 mixture of TrypLE Select (Life Technologies) and 0.5 mM EDTA/PBS for 12 min at 37°C to dissociate them into single cells. For 1 day after passaging hiPSCs, 10 μM ROCK inhibitor (Y-27632; Wako Pure Chemical Industries) was added in the culture medium.

METHOD DETAILS

Immunofluorescence (IF) analysis on frozen sections and whole mount IF analysis

For IF analysis of mouse embryos and some cynomolgus monkey embryos (for laminin), the trunks of embryos posterior to the level of the heart or isolated organs were fixed with 4% paraformaldehyde in PBS for 2 hr on ice, washed three times with PBS containing 0.2% Tween-20 (PBST), and then successively immersed in 10% and 30% sucrose in PBS overnight at 4°C. The fixed tissues were embedded in OCT compound (Sakura), frozen and sectioned to 10 μm thickness. The trunks of the embryos and organs were embedded perpendicularly or en face unless otherwise stated. Sections were placed on glass slides (Platinum Pro, Matsunami) that were then air-dried and stored at −80°C until use. Air-dried sections were washed three times with PBS, then incubated with blocking solution for 1 hr. The sections were incubated with primary antibodies in blocking solution for 2 hr at room temperature, then washed six times with PBS (20 min each). They were subsequently incubated with secondary antibodies and 1 μg/ml DAPI in blocking solution for 50 min at room temperature. After being washed six times with PBS (20 min each), the sections were mounted in Vectashield mounting medium (Vector Laboratories) for confocal laser scanning microscopy analysis (Olympus FV1000 or Zeiss LSM 780). For whole mount IF, mouse embryos were fixed in 4% PFA in PBS for 1 hr on ice, then washed three times with PBST. The embryos were then treated with blocking solution (1% donkey serum, 0.2% BSA, 0.2% Tween 20 and 1 × PBS) overnight at 4°C. Embryos were subsequently incubated with primary antibodies for 3 days at 4°C, then washed eight times (45 min each) with PBST. The embryos were then treated with secondary antibodies for 2 days at 4 × , then washed eight times (45 min each) with PBST before being mounted in Vectashield mounting medium for confocal laser scanning microscopic analysis.

To quantify the percentage of cells labeled by tdTomato for each cell type stained with specific markers, we captured images from randomly selected regions (two to five images per section) from two to four sections. The numbers of tdTomato^{+} cells among each cell type labeled by specific markers were counted in ImageJ.

IF analyses on paraffin sections

IF analyses for cynomolgus monkey embryos were performed on paraffin sections, except IF for Laminin. Embryos at E23 and E26 were fixed *in toto* in 10% buffered formalin (Nacalai Tesque). For embryos at E28 or older, anterior tissues (above the level of the heart) were removed, and the abdominal ectoderm was incised vertically before fixation to increase the perfusion by formalin. Samples were incubated overnight at room temperature with gentle rocking. After dehydration, tissues were embedded in paraffin and serially sectioned at 4 mm thickness with a microtome and placed on glass slides (Platinum Pro). Embryos were carefully oriented perpendicularly to the surface of the mold, and transverse sections were obtained. Paraffin sections were then de-paraffinized with xylene. Antigens were retrieved by treatment of sections with HistoVT one (Nacalai Tesque) for 35 min at 90°C and then for 15 min at room temperature. The staining and incubation procedure for paraffin sections was similar to that for frozen sections, with minor modifications: the primary antibody incubation was performed overnight at 4°C, and slides were washed with PBS six times after each incubation (20 min each). Slides were mounted in Vectashield mounting medium for confocal microscopy analysis.

Mouse mating, 4OHT administration and lineage tracing

The *Osr1*$^{Egfp/+}$ mice were mated with wild-type C57BL6 mice to obtain *Osr1*$^{Egfp/+}$ embryos. For lineage tracing, *T*$^{CreERT2/+}$, *Osr1*$^{CreERT2/+}$, *Wt1*$^{creERT2/+}$ and *Six2*$^{creERT2/+}$ male mice were mated with *Rosa26-LTL-tdTomato* females. Times starting at noon on the day on which a vaginal plug was found were designated as days post-coitum (dpc), e.g., E0.5. The 4-OHT (H6278, Sigma) was dissolved in sunflower oil (S5007, Sigma) at 5 mg/ml by sonication for 15 min and was stored at 4°C until use (within less than 1 month) with protection from light. For induction of Cre recombination, dams were injected intraperitoneally with 1.5 mg 4-OHT per 25 g body weight at the designated stages. This low 4OHT dose primarily restricted Cre activity to within 12 hr after treatment (Cazzulino et al., 2016; Robinson et al., 1991). The mice were sacrificed at E15.5 (E10.5 in some experiments), and embryos were isolated in PBS. For the control experiments, sunflower oil (oil) was injected at E8.5 or E9.5. All embryos were genotyped for the presence of Cre driver alleles by PCR. The sex of each sampled fetus was determined by the presence of testes or ovaries. The embryos were dissected in chilled STO medium (Dulbecco's modified Eagle's medium containing 10% fetal bovine serum, 2 mM Glutamax [35050–061, GIBCO] and 100 U/ml of penicillin/streptomycin [15140–133, GIBCO]). In some experiments, fetal gonads (ovaries and testes) with attached mesonephros or kidneys with attached adrenal glands were carefully isolated with forceps and washed with PBS. These organs were inspected for tdTomato fluorescence under a fluorescence dissection microscope (M205C, Leica), and images were captured with the same exposure time throughout an experiment unless otherwise stated.

Induction of metanephric mesenchyme by three-dimensional culture

MM induction was performed according to a previous report with modifications (Taguchi et al., 2014). First, 3D45 hiPSCs were aggregated by seeding 1×10^4 cells/well in 96-well low cell binding V-bottom plates (Thermo, 91100574) in the presence of 10 μM Y-27632 and 0.5 ng/ml human BMP4 (R&D Systems). After 24 hr (on day[d] 1), the medium was changed to BC10 medium (1 ng/ml BMP4 and 10 μM CHIR99032). Subsequently, from d3 onward, half the culture medium volume was changed every other day (BC10Y: BC10 plus 10 μM Y-27632). On d7, the medium was switched to ABC3R medium (10 ng/ml Activin-A, 3 ng/ml BMP4, 3 μM CHIR99032, 0.1 μM retinoic acid and 10 μM Y-27632). On d9, the medium was switched to C1F medium (1 μM CHIR99032 and 5 ng/ml FGF9). The samples at d0 (hiPSCs), d1, d3, d5, d7, d9 and d12 were harvested and dissociated with 0.1% trypsin/EDTA treatment for 15 min at 37°C with periodic pipetting. After the reaction was quenched by the addition of a half volume of FBS, cells were resuspended in FACS buffer (0.1% BSA in PBS) and strained through a 70 μm nylon cell strainer (BD Biosciences) to remove cell clumps; the cells were then used for flow cytometric analysis. For qPCR analysis, cells were pelleted by centrifugation (500 g for 10 min) and subsequently lysed for RNA isolation and cDNA preparation.

10x Genomics single-cell RNA-seq library preparation

Fragments of cynomolgus monkey embryos at E24 (CS11), E28 (CS13), E31 (CS14) and E37 (CS18) were used for scRNA-seq with a Chromium Single Cell 3' Reagent Kit (v2 chemistry). In contrast to mouse embryos, genetic tools to trace or mark specific lineages in cynomolgus monkey embryos are not available. Therefore, we used anatomic landmarks to manually isolate the regions in which GPCs and their ancestries were enriched, and removed as much irrelevant tissue as possible. To ensure collection of sufficient numbers of GPC lineages among contaminated lineages for transcriptome analyses, we used scRNA-seq on the 10x Genomics platform, which has higher throughput than SC3-seq.

For preparation of E24 (CS11) embryos, the yolk sac and amnion were first carefully trimmed with forceps and ophthalmic scissors, and then the heart (and anterior tissues above the heart) and the posterior growth zone including the PS were removed.

At E31, MM condensation around LHX1$^+$ UB, an initial morphologic sign of metanephric kidney formation, was observed at the regions medial to the hind limb buds (Figure S6B). Anteriorly, a cluster of NR5A1$^+$WT1$^-$GATA4$^-$ adrenal glands cells was observed at the mediodorsal aspect of the nascent GPCs, dorsal to the liver (Figure S6B). Therefore, the heart (and anterior tissues above the heart), and hindlimb (and tissues beneath the hindlimb) were first removed, and then the urogenital ridge containing the mesonephros, CE and a portion of the proximal mesentery were isolated for E28 (CS13) and E31 (CS14) embryos.

For E37 embryos (CS18), similarly to E28/31 embryos, after the anterior and posterior tissues were removed, the urogenital ridges were isolated en bloc. Subsequently, most mesonephros and mesentery were trimmed, and tissues consisting of predominantly GR with an attached portion of mesonephros, putative adrenal glands and proximal mesentery were used for downstream sample processing.

Isolated embryonic fragments were washed twice with PBS and then minced with scissors in 500 μL of 0.1% trypsin/EDTA solution, then incubated for 9 min at 37°C with gentle pipetting every 3 min. After quenching of the reaction by addition of 500 μL of STO medium, cell suspensions were strained through a 70 μm nylon cell strainer and centrifuged for 220 g for 5 min. Cell pellets were resuspended in 0.1% BSA in PBS and counted. All samples were stained with trypan blue and confirmed to be > 80% viable. Cells were loaded into Chromium microfluidic chips and used to generate single cell gelbead emulsions with a Chromium controller (10x Genomics) according to the manufacturer's protocol. Gelbead emulsion-RT was performed with a T100 Touch Thermal Cycler (Bio-Rad). All subsequent cDNA amplification and library construction steps were performed according to the manufacturer's protocol. Libraries were sequenced with a NextSeq 500/500 high output kit v2 (150 cycles) (FC-404–2002) on an Illumina NextSeq 550 sequencer.

Mapping reads of 10x Chromium scRNA-seq and data analysis

Raw data were demultiplexed with the mkfastq command in Cell Ranger (v2.1.0) to generate Fastq files. Trimmed sequence files were mapped to the reference genome for cynomolgus monkeys (MacFas5.0). Read counts were obtained from outputs from Cell Ranger.

Secondary data analyses were performed in R software version (v.3.6.1) with the Seurat (v.3.1.5), ggplot2 (v.3.3.2), gplots (v.3.0.3), qvalue (v.2.18.0), maptools (v.0.9-9), genefilter (v.1.68.0), rgl (v.0.100.54), dplyr (v.0.8.3) and matrix (v.1.2-18) packages and Excel (Microsoft). UMI count tables were first loaded into R by using the Read10x function, and Seurat objects were built from each sample. Cells with more than 5500 genes or fewer than 1000 genes were filtered out. Mitochondrial content was not used for normalization since the majority of mitochondrial genes starting with "MT-" were not annotated in MacFas5.0, a reference genome for cynomolgus monkeys. Samples were combined, and UMI counts were normalized with the NormalizeData function (method = "RC"; scale factor = 10 M) and then converted to $\log_2$ (CP10M+1) values. Combined data were then analyzed with UMAP and tSNE clustering. Clusters were annotated on the basis of previously characterized marker gene expression by using the FeaturePlot function and the gene expression matrix file, and cluster annotation was generated for downstream analyses.

For identification of DEGs (marker genes) among groups, *p-value*s for expressed gene ($\log_2$[CP10M+1] > 0) were first calculated with the Kruskal-Wallis test, and FDR values < 0.05 for these genes were obtained post hoc with Storey's method. The DEGs were then defined as the genes exhibiting higher expression in one cluster than the remaining clusters with FDR < 0.05. In some analyses, only the top 100 or 200 DEGs ranked by fold change (compared with other cell types) are shown.

For identification of DEGs in pairwise comparisons, Welch's t test and the Benjamini-Hochberg method were applied to calculate the p value and FDR, respectively. Among genes with FDR < 0.05, those with a more than 4-fold difference (2-fold difference in some analyses) were defined as DEGs and were plotted over the scatterplot of averaged transcriptome values for cell clusters.

For pseudotime analysis, raw counts from single cell RNA-seq were normalized by log2 transformation. Highly variable genes were identified by getTopHVGs() in R packages scran (Lun et al., 2016). Then, only variable genes with FDR < 0.01 were used to calculate diffusion components (DC), and diffusion pseudotime (DPT) (Haghverdi et al., 2016).

For the GO analyses using the DAVID web tool, because the annotation of *M. fascicularis* genes was relatively incomplete, the human annotation corresponding to that of cynomolgus monkeys was used. The Enrichr web tool was used for the analysis of Jensen DISEASES enrichment scores (Kuleshov et al., 2016).

Single-cell cDNA preparation and primary transcriptome analysis by SC3-seq

To isolate PIM/ECE cells, we dissected the posterior trunks (below the heart and above the hind limb bud) from E9.5 *Osr1*$^{Egfp/+}$ mouse embryos. We pooled seven embryos (including both male and female embryos) and FACS-sorted EGFP$^+$ cells. To isolate GPCs and their derivatives, we isolated GRs from *Dppa3*$^{Ecfp/+}$ mouse embryos (E10.5, one female embryo; E11.5, seven embryos [containing both male and female embryos]; E13.5, one female embryo) by using the accumulation of ECFP signals (representing PGCs) as a marker of nascent GRs. Embryos from E10.5 and E13.5 were individually genotyped for sex by PCR. Embryos at E11.5 were pooled without being genotyped and used for downstream processing. In one experiment, GRs were dissected from a cynomolgus monkey embryo at E36 (CS18) and processed similarly to mouse GRs, as described below.

GRs were dissociated into single cells by incubation with 0.25% trypsin/PBS (T4799, Sigma) for approximately 10 min at 37°C followed by repeated pipetting. The resultant single cells were dispersed in 0.1 mg/ml polyvinyl alcohol/PBS (P8136, Sigma) and processed for the SC3-seq analysis. Single-cell cDNAs (E9.5, 180 cells; E10.5, 60 cells; E11.5, 120 cells; and E13.5 60 cells) were prepared as described previously (Ishikura et al., 2016; Nakamura et al., 2015). Briefly, cells were picked up manually into the lysis buffer containing tagged dT primer and ERCC spike-in RNA sequences, which allow to estimate the transcript copy numbers per cell. After cells were lysed by heating (90 s at 70°C), the reserve transcription buffer was added, and the cDNA strand was synthesized. The excess amount of tagged dT primers were subsequently digested by exonuclease. Next, the poly A tail was added at the end of the synthesized cDNA strand by a terminal deoxynucleotidyl transferase reaction. After the second cDNA strand with another tag was synthesized, double strand cDNAs were amplified by PCR using tag sequences. Amplified cDNAs were purified by using a 0.6 x volume of AMPure XP and eluted in 50 ul of Buffer EB (QIAGEN) and the quality was evaluated by qPCR. Only high quality cDNAs with high *Arbp* expression (Ct value < 19 by qPCR) were used for further selection. For cDNAs derived from E9.5 embryos, cells expressing *Wt1* (Ct value < 23) were selected for scRNA-seq analysis. For cDNAs derived from the remaining embryos, *Wt1*$^+$*Gata4*$^+$ (Ct

value < 23) cells for E10.5 (39 cells), various somatic cells (*Wt1*$^{+/-}$*Gata4*$^{+/-}$*Foxl2*$^{+/-}$*Mafb*$^{+/-}$, Ct value < 23 for positive criteria), E11.5 (88 cells) and E13.5 (35 cells) were used for sequencing analyses.

The cDNA libraries for Nextseq550 (Illumina) were constructed as described previously. Briefly, the amplified products were purified three times with an AxyPrep Mag PCR clean-up kit (Corning, MAGPCR-CL), then fragmented with a Covaris E220 sonicator (Covaris, E220). The fragmented ends were polished with T4 DNA polymerase and T4 polynucleotide kinase for 30 min, and subsequently purified with an AxyPrep Mag PCR clean-up kit. After addition of the Rd1SP and Rd2SP adaptors, the samples were PCR-amplified with the Nextera XT Index 1 (N7XX) and Index 2 (S5XX) primers for 11 cycles, then purified with an AxyPrep Mag PCR clean-up kit. The quality and quantity of the resultant cDNA libraries were evaluated with LabChip GX analysis (Perkin Elmer), a Qubit dsDNA HS assay kit (Thermo Fisher, Q32854) and a TaqMan qPCR assay using Thunderbird Probe qPCR mix (TOYOBO, QPS-101) and a TaqMan probe (Ac04364396; Applied Biosystems). The sequence data were acquired with a NextSeq 500 high output kit v2 (75 cycles) (Illumina, FC-404-2005).

All read data were converted into expression levels as described previously (Nakamura et al., 2015). Briefly, all reads were treated with cutadapt-1.3 to remove the V1 and V3 adaptor sequences and poly A sequences. The resulting reads of 30-bp and longer were mapped onto the mm10 genome (for mice) or MacFas5.0 (for cynomolgus monkeys) with tophat1.4.1/bowtie1.0.1 with the "-no-coverage-search" option. The mapped reads were then converted to expression levels (RPM) with cufflinks-2.2.0 with the "-compatible-hits-norm," "-no-length-correction," "-max-mle-iterations 50000" and "library-type fr-secondstrand" options and mm10 reference gene annotations with up to 10-kb extension at the 3′ end.

Secondary analysis of the SC3-seq data

RNA-seq data analysis was performed in R software version (v.1.2.5019) with the ggplot2 (v.3.3.2), gplots (v.3.0.3), qvalue (v.2.18.0), maptools (v.0.9-9), genefilter (v.1.68.0) and rgl (v.0.100.54) packages, and Excel (Microsoft). Low quality cells (mapped reads on the sense strand < 0.5 M reads, or 75% quantile gene expression < $\log_2$(RPM+1) = 2.5) were further filtered out. For E9.5 and E11.5 mouse embryo samples, female cells ($\log_2$ [RPM+1] > 4 for *Xist*) were selected for downstream analyses. For SC3-seq analysis of cynomolgus monkey embryos, low quality cells (mapped reads on the sense strand < 0.5 M reads) were filtered out.

UHC was performed with the hclust function with Euclidian distances and Ward distance function (ward.D2) by using genes (protein-coding genes or non-coding RNA) whose $\log_2$(RPM+1) values were > 6 (> 4 for cynomolgus monkey samples) in at least one sample. PCA was performed with the prcomp function without scaling. For identification of DEGs by multi-group comparison, the same approach described for 10x Chromium scRNA-seq was followed, except that FDR < 0.01 was used for the statistical significance call (multi-group and pairwise comparison), and the minimum mean abundance of $\log_2$ (RPM+1) > 2 was used (pairwise comparison). For the GO analysis, the DAVID webtool with the background list set to "Mus musculus" was used. Only enriched GO terms with p < 0.05 are shown.

Comparison of gene expression between mice and cynomolgus monkeys

For a comparison between cynomolgus monkeys and mice, first, a mouse-human annotation list was generated, and then the human-cynomolgus monkey list and mouse–human list were combined by using human gene identifiers, as described previously (Nakamura et al., 2016). This process resulted in 15415 genes in the cynomolgus monkey–human–mouse gene list (Table S6). In some analyses, we performed Z scaling for normalization of expression values between mouse and monkey data using the genescale function in the 'genefilter' package (Figures 7A and 7F). In other analyses, we simply isolated the DEGs between cell types in one platform/species and evaluated their expression in the other platform/species (Figures 7B–7D) or comparing DEGs between cell types in one platform to those in the corresponding cell types in the other platform (Figure 7E).

qPCR analysis

qPCR on amplified cDNA used for SC3-seq was performed with Power SYBR Green PCR Master mix (Life Technologies) with a CFX384 real-time qPCR system (Bio-Rad).

The gene expression levels were examined by calculation of ΔCt on $\log_2$ scale, normalized to the averaged Ct values of *PPIA* (*Ppia*) and *ARBP* (*Arbp*). The primer sequences used are listed in Table S5.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis was performed in Excel and RStudio (version 1.2.5019, http://www.rproject.org/). The number of biological replicates and the statistical test performed are indicated in the figure legends or elsewhere in the STAR Methods section. A boxplot shows median (center line), upper and lower quartiles (box limits) and 1.5x interquartile range (whiskers). All fluorescence and histologic images and flow cytometric data were representatives of at least two independent experiments with similar results obtained at each experiment, unless stated otherwise in figures.