Ovarian development is driven by early spatiotemporal priming of the coelomic epithelium

Authors: Cyril Djari[1,2,†], Chloé Mayère[1,2,†], Maëva Guy[1,2], Aitana Perea-Gomez[3], Laura Bellutti[4,5], Paul Barreau[1,2], Herta Ademi[1,2], Agathe Rozier[3], Anthony S. Martinez, Tyler J. Gibson[6], Françoise Kühne[1,2], Cassandre Guérin[3], Dagmar Wilhelm[7], Gabriel Livera[4,5], Jennifer McKey[6], Marie-Christine Chaboissier[3], Serge Nef[1,2,*].

Affiliations:

[1]Department of Genetic Medicine and Development, University of Geneva, 1211 Geneva, Switzerland.

[2]iGE3, Institute of Genetics and Genomics of Geneva, University of Geneva, Switzerland.

[3]Université Côte d'Azur, Inserm, CNRS, Institut de Biologie Valrose (iBV), 06108 Nice, France.

[4]Université Paris Cité, CEA, Genetic Stability Stem Cells and Radiation, Laboratory of Development of the Gonads, F-92265 Fontenay-aux-Roses, France.

[5]Université Paris-Saclay, CEA, Genetic Stability Stem Cells and Radiation, LDG/IRCM/IBFJ, F-92265 Fontenay-aux-Roses, France.

[6]Section of Developmental Biology, Department of Pediatrics, University of Colorado Anschutz Medical Campus, Aurora, CO, 80045, USA.

[7]Department of Anatomy & Physiology, University of Melbourne, VIC 3010, Australia

[†]These authors share first authorship

[*]Corresponding Author: Email: Serge.Nef@unige.ch

Abstract

Ovarian organogenesis relies on the coordinated specification of supporting and steroidogenic lineages from multipotent coelomic epithelium (CE) progenitors. However, it remains unclear how ovarian cellular diversity arises and whether CE progenitors are fate-biased before or after ingression into the gonad. We show that CE cells in fetal mouse and human ovaries are transcriptionally heterogeneous and spatially organized into subdomains already primed toward supporting or steroidogenic identities. CE priming is dynamic, influenced by proximity to the mesonephros, with a transient coexistence of both progenitor types before resolution toward a predominantly supporting-biased CE. In mice, delamination of primed CE cells seeds intragonadal niches that generate pre-granulosa and steroidogenic progenitors. We further demonstrate that fetal steroidogenic progenitors give rise to adult stromal and theca cells, and that granulosa cells have a dual origin from CE-derived and supporting-like cells. Together, these findings reveal a conserved, spatially encoded program of ovarian lineage specification.

Keywords: Ovarian development; coelomic epithelium; ovarian surface epithelium; granulosa cells; steroidogenic/ progenitors; theca cells; single-cell RNA-sequencing (scRNA-seq); spatial transcriptomics; lineage tracing; fate restriction.

Introduction

The sex of an individual is established at fertilization by the paternal contribution of either an X or a Y chromosome, resulting in an XX or XY embryo (Saunders and Veyrunes, 2021). This genetic sex directs the differentiation of the initially bipotential gonads into ovaries in XX embryos or testes in XY embryos. This process, known as gonadal sex determination, is fundamental to sexual development and fertility, as hormones secreted by the differentiated gonads subsequently regulate the development of primary and secondary sexual characteristics as well as gametogenesis. Sex determination occurs at 5–7 weeks in humans and at embryonic day (E)11.5 in mice, with a slight delay in XX embryos compared with XY embryos (Stevant et al., 2019; Wilhelm et al., 2025). Prior to sex determination, proliferation of multipotent coelomic epithelium (CE) progenitors in the emerging gonadal primordium is essential for gonadal formation (Karl and Capel, 1998; Mayere et al., 2022; Schmahl et al., 2000). CE progenitors divide asymmetrically, with some daughter cells remaining epithelial while others undergo epithelial-to-mesenchymal transition (EMT) and ingress into the underlying gonadal tissue. This process drives gonadal growth and establishes the somatic foundation of the developing gonad (Karl and Capel, 1998; Kusaka et al., 2010; Lin et al., 2017; Schmahl et al., 2000).

Multipotent CE progenitors generate most somatic cells in both XX and XY gonads and give rise to the two major somatic lineages of the gonad: the supporting and steroidogenic lineages (Ademi et al., 2022; DeFalco et al., 2011; Karl and Capel, 1998; Liu et al., 2015; Liu et al., 2016; Mayere et al., 2022; Mork et al., 2012; Schmahl et al., 2003; Stevant and Nef, 2018, 2019). In XX gonads, these lineages form pre-granulosa and theca/stromal cells, whereas in XY gonads, they generate Sertoli and Leydig cells, respectively. Despite their central importance, the mechanisms underlying CE progenitor diversification remain incompletely understood. A key unresolved question is whether the CE contains heterogeneous progenitor populations already biased toward supporting or steroidogenic fates, or whether lineage specification occurs only after delamination into the gonad. If such heterogeneity exists, how these progenitors are temporally regulated and spatially organized to generate the two somatic lineages remains unknown.

Another major unresolved question in ovarian biology is how granulosa and steroidogenic lineages arise during development. Pre-granulosa cells emerge in two waves, a medullary first wave (Mork et al., 2012; Niu and Spradling, 2020; Zheng et al., 2014) and a cortical second wave (Cai et al., 2020; Feng et al., 2016; Niu and Spradling, 2020; Rastetter et al., 2014; Stevant et al., 2019), but the dynamics and relative contributions of these populations remain unclear, particularly given recent evidence that XX supporting-like cells (SLCs) also generate a subset of

pre-granulosa cells (Mayere et al., 2022; McKey et al., 2022). These pre-granulosa cells subsequently differentiate into granulosa cells at the onset of folliculogenesis after birth in mice (Mork et al., 2012; Zheng et al., 2014). Similarly, although theca cells differentiate only after birth (Liu et al., 2015; Magoffin and Weitsman, 1994), the specification and maintenance of their fetal progenitors remain poorly defined. We refer to these cells as stromal progenitors (SPs), a population encompassing both steroidogenic progenitors and non-steroidogenic stromal derivatives. How these distinct progenitor pools are specified, spatially organized, and integrated into ovarian development therefore remains a central open question.

To address how ovarian cellular diversity arises upstream of sex determination and how granulosa and steroidogenic lineages first emerge, we combined single-cell transcriptomics in mouse and human with spatial transcriptomics, *in vivo* lineage tracing, and *ex vivo* functional assays. These approaches reveal that granulosa cells have a dual origin and that spatially patterned CE subpopulations are primed toward distinct fates and directly seed ovarian lineages through localized delamination, establishing intragonadal niches with important implications for ovarian development.

Results

Coelomic epithelium and supporting-like cells are conserved progenitors of ovarian somatic lineages in humans and mice

To investigate the specification and differentiation dynamics of supporting and stromal/steroidogenic lineages in humans and mice, we generated species-specific single-cell RNA-sequencing (scRNA-seq) atlases of ovarian development spanning post-conception weeks (PCW) 5-21 in humans and embryonic days (E)10.5 to post-natal day (P)56 in mice. These atlases were assembled from two human-focused studies (Garcia-Alonso et al., 2022; Lardenois et al., 2026) and three mouse single-cell datasets (Mayère et al., 2022; Niu and Spradling, 2020; Rossitto et al., 2022) (**Fig. S1A-F**). After annotation harmonization and selection of ovary somatic cells based on *GATA4*/*Gata4* expression, we generated two refined somatic sub-atlases comprising 156,408 human cells **(Fig 1A, B** and **Fig. S1C**) and 43,053 mouse cells **(Fig 1C, D** and **Fig. S1D**). In both species, these sub-atlases included coelomic and surface epithelial cells (CE/SE; *UPK3B*□/*Upk3b*$^{+}$), supporting-like cells (SLCs; *PAX8*□/*Pax8*$^{+}$), stromal progenitors (SPs; *PDGFRA*□ in human/*Nr2f2*$^{+}$ in mouse), as well as pre-granulosa cells. In humans, pre-granulosa cells segregated into two distinct sub-populations corresponding to the first- and second-wave pre-granulosa cells (preGran-I; *CYP19A1*$^{+}$ and preGran-II; *IRX3*$^{+}$, *LHX2*$^{+}$)**(Fig.**

S1E). Mouse pre-granulosa cells formed a continuous cell community merging first-wave (*Foxl2*$^{+}$) and second-wave (*Lgr5*$^{+}$) pre-granulosa (**Fig. S1F**).

Partition-Based Graph Abstraction with conductance scoring (cPAGA) revealed that, in both species, CE progenitors form a central developmental node strongly connected to SPs and second-wave pre-granulosa cells (preGran-II in humans; *Lgr5*□ pre-granulosa cells in mice) (**Fig. 1B, D** and **FigS1F**). In contrast, SLCs showed a preferential connectivity with first-wave pre-granulosa cells (pre-Gran-I in humans and *Foxl2*$^{+}$ pre-granulosa cells in mice). Stage-wise UMAP projection and cPAGA analyses further revealed a temporally ordered progression of the lineage relationships. In humans, CE cells were initially transcriptionally connected to SPs at 5-6 PCW, then transiently connected to both SPs and preGran-II during a dual phase at 7-9 PCW, before becoming exclusively connected to the supporting lineage from 10 PCW onward (**Fig. 1E** and **Fig. S1G**). In mouse, CE/SE cells were connected to both SPs and pre-supporting/supporting cell at early stages (E10.5 to E14.5), followed by exclusive CE–supporting connectivity from E16.5 onward, consistent with a dynamic progenitor role of CE/SE cells (**Fig. 1F** and **Fig. S1H**).

Cell proportion and proliferation analyses further supported this model (**Fig. 1G** and **H**). Despite high proliferative activity at early stages, CE/SE cells progressively declined in abundance over development. In contrast, the largely non-proliferative pre-Granulosa population expanded rapidly in both species, consistent with expansion driven predominantly by continuous differentiation and recruitment of CE-derived progenitors rather than by intrinsic proliferation. SPs exhibited complementary dynamics: following a possible initial contribution from CE/SE, they showed intermediate proliferation activity and an autonomous expansion, with proportions decreasing at early stages (PCW5-11 and E10.5-E12.5), before increasing again at later time points. SLCs display minimal proliferative activity and constitute a small population whose abundance progressively decreases over developmental time.

Altogether, these comparative analyses identify proliferative CE cells as a central progenitor pool for both supporting and stromal/steroidogenic lineages, contributing substantially to the expanding but largely non-proliferative pre-granulosa population, while SLCs represent a parallel progenitor source for pre-granulosa cells.

Coelomic epithelium progenitors are transcriptionally heterogeneous and dynamically biased toward supporting or stromal fates

To determine whether CE progenitor cells (*UPK3B*$^{+}$/*Upk3b*$^{+}$) are initially homogeneous or instead contain fate-biased subpopulations, we examined the expression of established markers of

pre-granulosa cells (*LHX2, BMP2/Bmp2, IRX3/Irx3, Lgr5, Runx1* and *Wnt6*) and stromal progenitors (*PDGFRA/Pdgfra*, *NR2F2/Nr2f2, WNT5A/Wnt5a, Arx* and *Tcf21*) (**Fig. 1I, K**, **Fig.S1E** and **F**). In both species, CE cells display heterogeneous expression of supporting and stromal markers, revealing that the CE is not a uniform pool of progenitors.

To assess whether this observation reflects the expression of a limited set of markers or instead indicates a broader transcriptional bias of CE progenitors toward a specific fate, we calculated supporting and stromal identity scores for each CE/SE cell. These scores were generated using regularized regression to identify the optimal set of transcripts that distinguishes supporting and SP populations from others, using differentiated supporting and SPs from PCW8 to 21 in humans and E12.5 to P5 in mice as reference populations (**Fig. S1I** and **J**). UMAP projection colored with supporting and SP identity scores revealed distinct CE/SE transcriptional states in both species: supporting-biased CE/SE cells with high supporting scores localized near the supporting cells, whereas SP-biased CE/SE cells with high SP scores clustered near the SP population (**Fig. 1J** and **L**).

We next examined whether CE transcriptional heterogeneity follows similar temporal dynamics in humans and mice. Initially, a high proportion of CE/SE cells expressed stromal-associated markers (*NR2F2/Nr2f2*, *WNT5A/Wnt5a*, *PDGFRA/Pdgfra*, *Arx* and *Tcf21)*, which progressively declined, while the proportion of cells expressing supporting associated markers (*LHX2, IRX3/Irx3, WNT6/Wnt6, Bmp2, Lgr5* and *Runx1*) followed an opposite pattern (**Fig. 1M** and **N**). Identity score analysis confirmed parallel developmental trajectories. Median SP identity scores peaked around the time of sex determination and declined rapidly thereafter (PCW5–10 in humans; E10.5–E12.5 in mice), whereas supporting scores increased from the beginning of sex determination (PCW6–10 in humans; E11.5–P5 in mice) (**Fig 1O** and **P**). Stage-specific UMAPs visualization further revealed that these fate biases are largely mutually exclusive at the single-cell level. In humans, CE cells progressed from a predominantly SP-biased state at PCW5-6, then entered a transient dual phase at PCW7–9, characterized by distinct subpopulations biased toward either supporting or SP fates, after which CE cells became predominantly supporting-biased from PCW10 onward (**Fig 1Q**). Mouse CE/SE dynamics closely mirrored this three-phase progression: CE cells initially displayed a stronger SP bias, then distinct supporting- and SP-biased groups at E12.5-E14.5 and finally became predominantly supporting-biased from E16.5 onward (**Fig 1R**).

Together, these analyses reveal that CE is transcriptionally heterogeneous and follows a conserved, temporally ordered program. In both humans and mice, CE progenitors transition from an early phase dominated by SP-biased cells, to a transitional phase characterized by mutually

exclusive subpopulations biased toward either supporting or SP fates, to a final supporting-restricted phase, progressively refining lineage potential while preserving supporting fate competence.

Spatial patterning of the ovarian epithelium defines biased surface domains

Having established the conserved temporal dynamics of CE transcriptomic biases in humans and mice, we next asked whether these biased states are spatially organized along the ovarian surface during development. We focused our spatial analyses on the mouse ovary, whose small size and ease of tissue acquisition allow comprehensive mapping of CE biased domains across the entire ovarian surface at high resolution. We used spatial transcriptomics data generated with 10x Genomics Visium HD on parasagittal sections of XX gonads at E12.5, E14.5, and E16.5 to map CE priming *in situ* at near-cellular resolution (Martinez et al., 2025) (**Fig. 2A**). To resolve spatial organization within the CE, we delineated 4 µm x 4 µm CE bins along the ovarian surface, reconstructed an anteroposterior pseudo-spatial axis, and computed supporting- and SP-identity scores for each bin. At E12.5, CE cells with SP bias were enriched at the extremities of the pseudo-spatial axis near the gonad-mesonephros junction, whereas supporting-biased CE cells were concentrated in the central region (**Fig. 2B** and **C**). This regionalization persisted at E14.5 but became less distinct by E16.5 as the stromal signature declined (**Fig. 2C**). Expression patterns of individual *Runx1* supporting marker and *Nr2f2* SP marker confirmed this organization (**Fig. 2D**). Heatmap visualization of genes differentially expressed in supporting (n = 300) versus stromal (n = 243) cells further confirmed this spatial segregation in the CE (**Fig. 2E, Fig. S2** and **Data S1).** In particular, stromal-biased CE cells at the extremities of the pseudo-spatial axis expressed *Tcf21*, *Pdgfra*, and *Wnt5a*, whereas supporting-biased CE cells in the central region expressed *Wnt6*, *Bmp2*, *Irx3*, *Lgr5*, and *Fst* (**Fig. S2)**. Thus, the molecular signatures that define SP- and supporting-biased CE states in scRNA-seq map onto spatially segregated domains along the ovarian surface.

To validate and refine these spatial domains at cellular resolution, we performed immunofluorescence (IF) on mouse XX gonads from E11.5 to E16.5 using the supporting marker RUNX1 and the stromal/steroidogenic marker NR2F2 (**Fig. 2F**). At E11.5, NR2F2 was broadly expressed across the CE, whereas RUNX1 was not yet detectable, consistent with an initial widespread stromal competence within the CE prior to spatial restriction. By E12.5 and E13.5, IF confirmed the emergence of distinct spatial domains of supporting- and SP-biased CE cells along the ovarian surface. On transverse sections, NR2F2 CE cells localized near the gonad-

mesonephros junction, whereas RUNX1^{+} CE cells were restricted to the central region. These regions also exhibited distinct epithelial morphologies: the NR2F2^{+} domain appeared pseudostratified and columnar, while in the RUNX1^{+} domain cells were cuboidal. Over time, the proportion of NR2F2□ CE cells progressively declined, while RUNX1□ CE cells expanded and eventually covered the entire ovarian surface by E16.5, at which point NR2F2 was almost undetectable in the CE (**Fig. 2G** and **H**).

Whole-mount 3D imaging of cleared E12.5 ovaries further revealed that NR2F2□ CE domains form caps at the anterior and posterior poles of the gonad, near the mesonephros, whereas the central CE region was predominantly RUNX1□ (**Fig. 2I**). These observations suggest that mesonephric proximity provides positional cues promoting SP bias and/or preventing supporting specification. To test the contribution of the mesonephros to this spatial patterning, we cultured E11.5 XX gonads for 48 hours under three conditions: with the endogenous mesonephros attached, after surgical removal of the mesonephros, or with an additional mesonephros juxtaposed ventrally to generate a one-gonad/two-mesonephroi complex (**Fig. 2J**). Under control conditions with a single attached mesonephros, the *ex vivo* pattern after 48 hours recapitulated the *in vivo* situation at E13.5, with NR2F2□ CE domains near the mesonephros and a central RUNX1□ CE domain (**Fig. 2K**). Upon mesonephros removal, NR2F2□ stromal-biased CE cells no longer formed continuous polar domains but instead appeared as scattered clusters across the ovarian surface, indicating loss of organized stromal-biased caps. Conversely, the presence of an additional mesonephros promoted the formation and expansion of NR2F2□ CE domains adjacent to each mesonephros and distorted the central RUNX1□ supporting-biased region.

Together, these findings demonstrate that CE fate bias is spatially encoded along the ovarian surface, with SP-biased CE enriched near the gonad-mesonephros interface and supporting-biased CE concentrated centrally. The dependence of NR2F2□ stromal-biased domains on mesonephric proximity suggests that mesonephric cues contribute to establish and/or maintain these spatially restricted CE subpopulations.

Fate-primed ovarian surface epithelium organizes intragonadal domains

As the developing ovarian surface epithelium is partitioned into SP-biased and supporting-biased domains, we next investigated whether these surface compartments seed intragonadal somatic territories. We performed combined *in situ* hybridization (ISH) and IF for the CE marker *Upk3b*, the stromal/steroidogenic NR2F2 and the supporting lineage marker *Bmp2* on XX gonads from E11.5 to E16.5 (**Fig. 3A**). *Upk3b* expression, which is specific to the CE, was used to delineate

the surface epithelium from the underlying intragonadal territories. We took advantage of the transient expression of *Bmp2* as a proxy for supporting lineage commitment, as *Bmp2* is expressed in early supporting cells from E11.5 to E13.5 and subsequently downregulated (**Fig. 1N** and **S1F**). At E11.5, NR2F2 was broadly expressed across the CE, whereas *Bmp2* expression was restricted to the central CE domain, indicating that positional bias emerges within the CE before granulosa cell differentiation. At E12.5 and E13.5, *Bmp2* expression remained confined to the central CE, similarly to RUNX1 (**Fig. 2F**). By E16.5, although overall *Bmp2* levels decreased, its expression extended across most of the CE surface. At E12.5 and E13.5, NR2F2□ SPs aggregated beneath NR2F2□ SP-biased CE, while *Bmp2*□ pre-granulosa cells localized preferentially beneath the central *Bmp2*□ CE domain. By E16.5, as the stromal-biased CE domain regressed, the underlying SP-enriched regions became less sharply demarcated, whereas the supporting-enriched intragonadal domain extended beneath the surface epithelium.

To refine these spatial relationships, we projected scRNA-seq annotations onto the Visium HD spatial dataset to label intragonadal bins (**Fig. 3B**). To evaluate the spatial correspondences between intragonadal cell types and biased-CE progenitors along the ovarian surface, we quantified the relative proportion of SPs and pre-granulosa cells along the CE pseudo-spatial axis (**Fig. 3C**). At E12.5, regions directly beneath SP-biased CE at the ovarian extremities were enriched for SPs, whereas regions underlying the central supporting-biased CE domain contained higher proportions of supporting cells. This complementary pattern persisted at E14.5, with SP-enriched territories remaining aligned with the anterior and posterior caps, and supporting-lineage territories occupying the central ovarian region. By E16.5, as the SP-biased CE domain regressed and the supporting-biased CE domain expanded, the intragonadal SP and supporting territories were less clearly regionalized. Spatial transcriptomics confirmed that intragonadal distributions of pre-granulosa cells and SPs align with the spatiotemporal organization of CE-biased progenitors.

To further characterize how biased CE cells can contribute to their corresponding intragonadal compartments, we examined LAMA1 expression to evaluate the integrity of the basal lamina underlying the CE domains (**Fig. 3D**). LAMA1 staining revealed focal discontinuities in the basal lamina beneath both SP-biased and supporting-biased CE domains, at their interface with the corresponding intragonadal SP and pre-granulosa territories. These focal breaks are consistent with localized epithelial-to-mesenchymal transitions that enable fate-biased CE cells to delaminate into the gonad.

Together, these observations support a model in which spatially patterned, fate-primed CE domains contribute to corresponding intragonadal somatic territories, coupling ovarian surface organization to the establishment of supporting and stromal/steroidogenic lineages.

Early *Wnt5a*+ coelomic epithelium is the major source of granulosa cells

Transcriptional and spatial analyses suggest that pre-patterned CE domains seed both pre-granulosa cells and SPs. However, the spatiotemporal parameters governing their emergence and differentiation remain poorly defined, including the relative contributions of CE progenitors to medullary versus cortical pre-granulosa populations, and whether fetal SPs ultimately differentiate into steroidogenic stromal and thecal cells. To directly test these lineage relationships, we performed inducible *in vivo* lineage tracing using *Wnt5a-rtTA;TetO-Cre;Rosa-tdTomato* mice (hereafter *Wnt5a*^{lin}; (Ademi et al., 2022)). Single-cell RNA-seq analyses revealed that at E11.5, *Wnt5a* is expressed in the CE and early SPs but not in pre-supporting and pre-granulosa cells, whereas by E16.5 its expression is restricted to late SPs (**Fig. 4A**). This temporal shift in *Wnt5a* expression enabled two complementary induction windows in *Wnt5a*^{lin} mice: an early window (E11.5-E12.5) to label CE and a late window (E15.5-E16.5) to selectively label fetal SPs (**Fig. 4B** and **C**).

Following doxycycline induction at E11.5-E12.5, tdTomato labeled 85% of the CE at E12.5, indicating efficient targeting of the CE (**Fig. 4D** and **E**). Quantification revealed that 25% of FOXL2□ pre-granulosa cells were tdTomato^{+} at E12.5, suggesting an early CE contribution (**Fig. 4F**). By E16.5, this proportion increased to 52%, predominantly in the cortex, and in adult ovaries, 65% of FOXL2□ granulosa cells were tdTomato□ (**Fig. 4D, F** and **G**). To distinguish the contribution of CE cells to first versus second waves of folliculogenesis, we analyzed ovaries at P5 after primordial follicle activation (**Fig 4H** and **I**). The first wave of folliculogenesis is initiated in the medulla, whereas subsequent waves arise from a cortical follicle pool. At P5, tdTomato□ granulosa cells were preferentially localized in the cortex and accounted for 92% of tdTomato□/FOXL2□ cells, consistent with a predominant CE origin for cortical granulosa cells of the second follicular wave (**Fig 4I**). Notably, a substantial contribution was also observed in the medulla, where 79% of FOXL2□ granulosa cells were tdTomato□. Together, these data show that early *Wnt5a*□ CE cells give rise to a major proportion of the granulosa lineage, particularly within the cortical ovary, which harbors the follicle reserve for the reproductive lifespan.

Late fetal *Wnt5a*$^{+}$ stromal progenitors give rise to steroidogenic stromal and theca cells

We next leveraged the restriction of *Wnt5a* expression to trace the fate of fetal SP. Doxycycline induction at E15.5-E16.5 selectively labeled NR2F2□ fetal SPs, with no tdTomato detected in CE or FOXL2□ pre-granulosa cells, confirming induction specificity (**Fig. 4J**). In adult ovaries, tdTomato□ cells were confined to the stromal compartment, largely composed of NR2F2^{+} cells, and excluded from the FOXL2^{+} follicular compartment. Moreover, *Wnt5a*$^{+}$ fetal SPs contribute to CYP11A1□ steroidogenic stromal and thecal cells (**Fig. 4K**).

Together, these stage-specific *Wnt5a* lineage-tracing experiments functionally disentangle CE and stromal progenitor fates: early CE cells give rise to the majority of cortical and medullar granulosa cells, consistent with the supporting-biased domains identified at the ovarian surface, whereas late fetal SPs persist as long-lived adult SPs and generate stromal and theca steroidogenic populations.

Supporting-like cells provide an additional source of granulosa cells

Cross-species transcriptomic analyses predicted that, in addition to CE-derived granulosa cells, SLCs are an additional source of pre-granulosa cells. In the human atlas, SLCs were preferentially connected to preGran-I of the first follicular wave (**Fig. 1A, B** and **E**). Similarly, in the mouse dataset, SLCs connect to *Foxl2*$^{+}$ pre-granulosa cells (**Fig. 1C, D** and **F**), indicating a potential conserved contribution of SLC to the supporting lineage. *Pax8* expression is restricted to SLCs and the rete ovarii, with no expression in CE or pre-granulosa cells (**Fig. 5A** and **B**). This pattern enables specific lineage tracing using *Pax8-Cre;Rosa-tdTomato* mice (hereafter *Pax8*^{lin}) (Mayere et al., 2022), allowing us to assess the contribution of *Pax8*□ SLCs to granulosa cell formation (**Fig. 5C, D** and **Fig. S3**).

At E12.5, tdTomato□ cells were confined to the SLC/rete region at the hilum and were absent from the CE (**Fig. 5E** and **F**). Notably, ~5% of FOXL2□ pre-granulosa cells were tdTomato^{+}, suggesting an early contribution of SLC to granulosa cells (**Fig. 5E** and **G**). By E16.5, this proportion increased to ~22%, with tdTomato^{+}/FOXL2□ pre-granulosa cells enriched near the rete ovarii and hilum. By adulthood, ~10% of FOXL2□ granulosa cells were tdTomato□, consistent with a decline of the SLC-derived population (**Fig. 5H**). Having established the temporal dynamics of the SLC contribution to granulosa cells, we examined how these SLC-derived granulosa cells are distributed across medullary and cortical compartments at P5 (**Fig. 5I**). At this stage, SLC-derived granulosa cells were present in both medullary and cortical regions in comparable proportions of FOXL2^{+} cells (**Fig. 5J**). A gradient of tdTomato□/FOXL2□ granulosa

cells extended from the hilum toward the center of the ovary, supporting a spatially organized contribution of SLC-derived cells.

Together with the *Wnt5a* lineage-tracing experiments, these findings demonstrate that granulosa cells derived from two spatially distinct progenitor pools: early *Wnt5a*□ CE progenitors generate the majority of cortical granulosa cells that form the long-lived follicle reserve, whereas SLCs represent an additional source of granulosa cells with a more limited contribution to the adult ovary. This dual origin, supported by both mouse and human scRNA-seq data, refine the classical model of an exclusively CE-derived granulosa lineage.

Discussion

During early gonadogenesis, proliferating CE cells generate most somatic cells in both XX and XY gonads. Classical dye labeling, lineage tracing, and single-cell transcriptomic studies have established CE progenitors as multipotent, giving rise to both supporting and steroidogenic lineages (Ademi et al., 2022; DeFalco et al., 2011; Karl and Capel, 1998; Liu et al., 2016; Mayere et al., 2022; Mork et al., 2012; Schmahl et al., 2003; Stevant and Nef, 2018, 2019). Yet, the mechanisms that establish these divergent fates, particularly the specification of steroidogenic and supporting progenitors, and the spatiotemporal parameters governing their emergence and differentiation, remain unclear and poorly understood. Recent work on adrenogonadal primordium separation revealed that CE cells are specified toward adrenal or gonadal fates before ingression (Neirijnck et al., 2023), suggesting that specification within the CE may precede integration into the gonad through spatial pre-patterning of progenitor identity.

Conserved and dynamic coelomic epithelium priming in human and mouse ovarian development

Although ovarian development in humans unfolds over a longer timescale, weeks rather than days as in mice, the dynamics of somatic lineage specification are remarkably conserved. Transcriptomic connectivity between somatic populations, whether assessed globally or across developmental stages, shows similar patterns in human and mouse, as do cell-type proportions and proliferation dynamics. The only difference concerns the early SP population, which in humans appears to separate early pre-granulosa cells from the CE. It remains unclear whether this transitional state is absent in mice or instead represents a transient population that is too short-lived to be captured. Our study also identifies transcriptionally and spatially distinct fate-biased CE domains with progenitor subpopulations overlying the developing ovary. CE cells co-

expressing either SP (*Arx/ARX, Pdgfra/PDGFRA, Tcf21/TCF21*) or supporting (*Bmp2, Irx3/IRX3, Runx1, Wnt6/WNT6, LHX2*) markers exhibit transcriptomic signatures resembling their respective target lineages, consistent with *bona fide* priming. At early stages, in both mouse (E11.5) and human (PCW5–6), most CE cells are multipotent and NR2F2□. In mice, a subset of these cells already expressing *Bmp2*, marking the onset of positional bias. We propose that a first wave of supporting cells is generated before or around E11.5 through progressive downregulation of progenitor markers (e.g., NR2F2, TCF21) and acquisition of supporting identity upon entry into the gonad. From E12.5 onward, CE cells become spatially organized: supporting-primed (*Upk3b*□*/Bmp2*□) cells expand across the ovarian surface, whereas steroidogenic-primed (*Upk3b*□/NR2F2□) cells remain restricted near the mesonephros. At these stages (E12.5-E14.5), supporting-primed CE progenitors generate additional supporting cells, while steroidogenic-primed cells give rise to SPs. By E16.5, CE cells are almost exclusively supporting-primed and contribute to pre-granulosa and surface epithelial cells.

Spatial transcriptomics and immunofluorescence analyses revealed that intraovarian SP and pre-granulosa cells are positioned directly beneath their corresponding CE subdomains, consistent with local delamination of primed progenitors seeding spatially distinct intragonadal niches. These findings establish a direct, spatially organized contribution of CE subpopulations to ovarian lineage formation. In humans, CE progenitors display comparable dynamic priming but over a longer timescale. The sequence progresses from an early steroidogenic bias to a transitional coexistence of both progenitor types, followed by a predominance of supporting-biased CE cells as ovarian development advances.

Our observations align with work by Bunce et al. (Bunce et al., 2023), who reported center-biased ingression of supporting progenitors in both XX and XY gonads, producing higher supporting cell density in the central domain. Using a *Runx1-EGFP* reporter, they demonstrated greater accumulation of supporting precursors in the gonadal center. Consistently, our NR2F2/*Bmp2* combined *in situ* hybridization (ISH) and IF at E11.5 revealed that while most CE cells are NR2F2□, faint *Bmp2*□/NR2F2^{+} CE cells are detectable in the central ovarian region. Together, these results support a model in which regional patterning of CE progenitors determines lineage allocation.

Mesonephros as a positional regulator of coelomic epithelial fate

In mice, from E12.5 onward, supporting-biased CE cells expand broadly across the ovarian surface, whereas stromal/steroidogenic-biased CE cells remain confined near the mesonephros.

This spatial segregation suggests that mesonephric proximity provides positional cues that favor stromal/steroidogenic bias and/or prevent supporting bias specification when both CE subpopulations coexist. Consistent with this hypothesis, *ex vivo* culture experiments show that CE cells cultured adjacent to the mesonephros display enhanced NR2F2 expression, indicating that local signals either promote steroidogenic identity and/or inhibit supporting fate specification. We propose a model in which mesonephros-gonad crosstalk establishes a spatial gradient that sustains SP priming in CE progenitors located near the mesonephros. These results suggest that local environmental signals play a critical role in the specification of somatic lineages, and that they should be taken into account in protocols designed to generate functional ovarian organoids from iPS cells. Future work integrating ligand-receptor inference with targeted perturbations will be crucial to identify factor(s) or signaling pathways mediating this interaction and to determine how these mechanisms influence ovarian development and pathophysiology.

Conserved sequence but sex-specific timing of steroidogenic lineage differentiation

In mammals, testicular and ovarian differentiation follow distinct temporal trajectories, with testis differentiating earlier and more rapidly (Nicol et al., 2022; Stevant and Nef, 2019; Wilhelm et al., 2013). This divergence is most evident in androgen-producing cells: fetal Leydig cells begin secreting androgens shortly after testis differentiation (E12.5 in mouse; 6-7 PCW in human), whereas ovarian theca cells emerge only late in fetal life and do not acquire full steroidogenic activity until after birth, most prominently at puberty, when they become LH-responsive during folliculogenesis (Barsoum and Yao, 2010; Fortune and Armstrong, 1977; Magoffin, 2005; Mannan and O'Shaughnessy, 1991; Richards et al., 2018). Our mouse and human data reveal a conserved trajectory in which a subset of CE cells acquires a steroidogenic-primed identity during a restricted developmental window and contributes directly to intraovarian SPs. These SPs proliferate actively and ultimately constitute roughly half of all ovarian cells in both species, mirroring the marked expansion of SPs observed in the developing mouse testis, where they are also predominant at birth (Ademi et al., 2022). In both human and mouse, ovarian SPs express canonical markers (*Arx/ARX*, *Pdgfra/PDGFRA*, and *Tcf21/TCF21*), which are also characteristic of testicular SPs (Brennan et al., 2003; Cui et al., 2004; Eliveld et al., 2019; Estermann et al., 2025; Ge et al., 2006; Kilcoyne et al., 2014; Miyabayashi et al., 2013; Perea-Gomez et al., 2025; Qin et al., 2008; Shen et al., 2021). Lineage tracing further showed that fetal *Wnt5a*□ SPs persist and give rise to steroidogenic theca and stromal cells in the ovary as well as FLCs in the testis

(Ademi et al., 2022). Thus, although the timing differs, the overall sequence of SP specification expansion, and differentiation is likely to be conserved across sexes and species.

A dual origin for granulosa cells

Our study challenges the long-standing view that granulosa cells derive exclusively from the CE and instead supports a convergent, dual-origin model in both mouse and human ovaries. By integrating single-cell transcriptomics, spatial mapping and lineage tracing, we show that granulosa cells arise from two transcriptionally and spatially distinct progenitor pools. CE-derived progenitors generate the majority of cortical granulosa cells that contribute to the long-lived follicle reserve, whereas SLCs give rise to an additional granulosa subpopulation with a more limited contribution to the adult granulosa cells. This dual origin adds an important layer of complexity to ovarian lineage architecture and helps reinterpret follicular heterogeneity. The developmental origin of granulosa cells likely influences how distinct granulosa subsets interact with oocytes, proliferate, and respond to paracrine, endocrine and metabolic signals throughout ovarian life.

Material & Methods

Single-cell RNA sequencing analysis

Human data analysis

Raw count matrices from Garcia-Alonso et al. (Garcia-Alonso et al., 2022) were downloaded from their authors'public archive and human single cell raw counts from Lardenois et al. (Lardenois et al., 2026) were provided directly by the corresponding author. Dimensionality reduction was performed using SCVI with batch correction by sample and processing laboratory or by processing laboratory only. A SCANVI model for cell type prediction was trained using the annotations from both Garcia and Lardenois datasets.

Human data annotation

To refine cell type predictions obtained with the SCANVI model, the dataset was subset by developmental time point. For each subset, UMAP embeddings were generated and clustering was performed. Small clusters composed of barcodes with low UMI counts and low numbers of detected genes were removed as low-quality cells. Remaining clusters were annotated based on marker gene expression, SCANVI predictions, and the annotations reported in the original publications.

Human data embedding

As correcting batch effect at the sample level can conceal biological relevant information such as developmental stage, we created neighbor graphs using intermediate correction strategies. Specifically, we generated SCVI latent spaces corrected either by sample or by laboratory and built the corresponding neighbor graphs. Graph connections were then interpolated between these two correction schemes to produce 10 intermediate correction steps. After visual inspection, the third interpolation step was selected for downstream visualization, as it preserved developmental stage structure while reducing batch effects. For visualization, gene expression levels were log-normalized (target sum = 10000).

Human data lineage analysis and scoring

Lineage relationships within the human somatic compartment were inferred using a modified version of PAGA (Wolf et al., 2019). Conductance-based PAGA (cPAGA) in which the connectivity matrix between groups of cells is measured using intergroup conductance in the neighbor graph. In short, conductance measures the proportion of edges that are connected to neighbors that between two groups over the total number of edges. CE cells Sup and SP scoring was performed by training an ElasticNet regressor on Sup and SP cell categories and predicting the score of CE cells.

Mouse data preprocessing

Demultiplexing, alignment, barcode filtering, UMI counting and gene mapping were performed as described in the original publications: Mayère et al., 2022 for E10.5, E11.5, E12.5, E13.5 and E16.5 (Mayere et al., 2022); Niu et Spradling, 2020 for E14.5, E18.5, P1 and P5 (Niu and Spradling, 2020); and Rossitto et al., 2022 for P56 (Rossitto et al., 2022). For each sample, low-quality or empty droplets were removed by filtering barcodes with low UMI counts, following the procedure described in (Mayere et al., 2022). Otherwise, they were removed by visual inspection of the UMI counts density curves. Doublet scores were predicted using Scrublet v0.2.3 (Wolock et al., 2019) via the Scanpy v1.9.8 (Wolf et al., 2018) wrapper. Doublet scores thresholds were defined by identifying the closest local minima to the main peak of scores distribution, and barcodes above this threshold were removed. Additional quality controls and filtering were performed using scanpy.pp.calculate_qc_metrics for mitochondrial and ribosomal content. Cells with >20% mitochondrial counts were excluded. Potential XY-contaminating barcodes were identified based on *Ddx3y* expression (discarded if >0). All samples were then merged into a single AnnData object retaining 68,366 cells and 19,961 genes in total: 6,798 (E10.5), 5,935

(E11.5), 9,230 (E12.5), 10,866 (E13.5), 9,555 (E14.5), 6,223 (E16.5), 3,186 (E18.5), 4,932 (P1), 5,281 (P5), and 6,360 (P56).

Dimensionality reduction and batch correction of mouse data.

Dimensionality reduction and batch correction were performed with scvi.model.SCVI from scVI-tools v1.3.0 (Gayoso et al., 2022) on raw counts data. The model was trained using the samples IDs as batch keys to correct for the batch effects, with the following hyperparameter: 100 latent_space, 3 hidden layers with 128 nodes each.

Mouse data counts normalization

For expression visualization, differential expression, scoring, and model training, counts were log-normalized. Specifically, for each cell, UMI counts for each gene were divided by the total UMI counts of the cell, multiplied by a factor of 10,000, and log-transformed using scanpy.pp.normalize_total and scanpy.pp.log1p respectively. Cell cycle scoring was performed with scanpy.tl.score_genes_cell_cycle using the gene set from (Tirosh et al., 2016) as input.

Mouse data clustering, UMAP embedding, and annotation

For the full atlas, neighbors were computed in the SCVI latent space using scanpy.pp.neighbors with default parameters. Clustering was performed with scanpy.tl.leiden at resolution 1. We used PAGA on subsampled (ncells=200) clusters to optimize the global UMAP topology, we generated and plotted the PAGA graph with default parameters (scanpy.tl.paga) and used PAGA positions as seeds for the UMAP embedding. UMAP embedding was computed with scanpy.tl.umap function with min_dist parameter set to 0.2.
scVI based dimensionality reduction and batch correction are highly reliable, at the cost of a substantial loss of temporal information. To account for the time-dependent annotations and avoid inconsistencies, we annotated the dataset at two levels. First, a coarse annotation across the full dataset was performed using known marker expression. Second, refined annotations providing greater accuracy and reducing inconsistencies were carried out at each developmental stage using the coarse annotation as a guide. Each timepoint was clustered and embedded independently using the same preprocessing and clustering pipeline, except that PAGA subsampling was reduced to 100 cells per cluster. Known stage-specific markers were used to annotate clusters. Because temporal subsets contain fewer cells, rare cell populations were sometimes difficult to isolate. In ambiguous cases, subclustering was performed with a lower Leiden resolution (0.15 or 0.25).

Additionally, both E10.5 and E11.5 subsets required a specific annotation strategy due to their lower differentiation state and fraction of gonadal cells. For the E10.5 subset, we retained the coarse annotation and Leiden clustering from the full atlas and defined the mesonephric mesenchyme (MM) population by their low *Gata4* and high *Gata2* expression. For the E11.5 subset, Leiden clustering performed either on the full atlas or on the E11.5 subset did not separate gonadal and extragonadal stromal cells. We therefore isolated all stromal cells ($Tcf21^{+}/Gata4^{+}/Gata2^{+}$) using this initial clustering. To discriminate gonadal from extragonadal cells, we applied a k-nearest neighbors (kNN)-based co-expression search to identify the ten genes strongly co-expressed with *Gata4* and *Gata2* respectively. After verifying the specificity of these gene sets, we computed expression scores using sc.tl.score_genes, scaled them with a MinMax scaler, and assigned each cell to the gonadal or extragonadal compartment based on the dominant score.

Sub-atlases

To focus on the two major somatic lineages of the ovary, we extracted a sub-atlas restricted to ovarian somatic populations. scVI latent space was not recomputed, but the same preprocessing and clustering strategy as in the full atlas was applied. PAGA threshold parameter was set to 0.2, and UMAP min_dist parameter set to 0.3. Additional subsets were generated by isolating ovarian somatic populations at each stage, following the same strategy used for the time-point subsets described above.

Scoring

To create SP and supporting scores, we trained an ElasticNet model in a one-versus-rest fashion on CE/SE, SLC, SP and supporting (pre-Gran, Gran) cells from E12.5 to P5. Model parameters were set to an L1 ratio of 0.5, alpha of 0.01, and tolerance of 1e-4. Model parameters were set to keep only positive features, and with an L1 ratio of 0.5, alpha of 0.01, and tolerance of 1e-4. To avoid bias from overrepresented cell types, we randomly subsampled each group to match the size of the smallest population prior to training. By comparing each group to the other cells, this approach attributes to each gene a coefficient to create a transcriptomic signature. Genes with high descriptive power are attributed to high coefficients. Consequently, the accuracy of this method depends heavily on the annotation quality and the cellular composition of the subset. For reliability, E10.5 and E11.5 cells were excluded since lineages are not yet committed, and somatic differentiation has not occurred at this stage. We also restricted the analysis to fetal and perinatal data and focused exclusively on ovarian somatic cells (CE/SE, SLC, supporting lineage and

steroidogenic/stromal lineage). All the other cell types (adrenosympathetic, erythrocytes, endothelial, germ, immune, mesonephric tubules, mesonephric mesenchyme, perivascular) were excluded from training set.

Differential expression analysis

Differential expression analysis was performed using sc.tl.rank_genes_groups and the Wilcoxon method.

Spatial transcriptomics analysis

Data collection

Detailed methods for generating the full spatial transcriptome dataset of the developing and adult mouse ovary are described in (Martinez et al., 2025). Briefly, ovaries from fetal CD-1 outbred mice were collected at E12.5, E14.5 and E16.5, transferred to cryomolds, embedded in OCT, flash frozen and stored at -80ºC. Cryoblocks were sectioned using a Leica cryostat to produce 10µm sections, which were stained with Hematoxylin and Eosin (H&E). H&E-stained slides were imaged at 20x, and the highest quality and most representative slide was selected for sequencing. Slides were processed at the University of Colorado Anschutz Medical Campus Genomics Shared resource (RRID: 021984) following the Visium HD Spatial Gene Expression Reagent Kits User Guide (CG000685 Rev B). Sections were hybridized overnight with 10x Genomics' whole transcriptome probe pairs. The prepared Visium HD slides were processed on the CytAssist instrument (10x Genomics) for spatial barcoding, followed by library preparation. Libraries were pooled and sequenced on an Illumina NovaSeq X Plus system (150bp x 10bp x 10bp x 150bp), generating approximately 500 million reads per capture area.

Bins annotation

Raw sequencing and imaging data from 10x Visium HD were processed with spaceranger (v3.1.2) as described in (Martinez et al., 2025). Resulting Zarr datasets were then processed with the spatialdata python library. Initial bin annotation was performed using a logistic regression model trained on single-cell RNA-seq data restricted to genes also detected in the spatial dataset. To prevent overfitting and reduce bias from over-represented classes, cells were randomly under-sampled to 100 cells per population using RandomUnderSampler. The LogisticRegression model was trained using the following parameters: saga solver, elasticnet regularization (l1_ratio=0.8), a stopping criteria tolerance 0.01, and inverse regularization strength 0.8. The model achieved excellent predictive performance, with AUC values of 1.0 for most cell classes, and 0.99 for

CE/SE, SP and SLC populations. Precision and recall were also high, with the lowest average score observed for SLCs (precision 0.85) and MM cells (recall 0.83).

Spatial bins were predicted at 2, 4 and 8µm resolutions using the trained model. Because individual bins can contain transcripts from multiple cell types (e.g. pre-Gran and germ cells), predictions were strengthened using the spatial context of each bin. Specifically, logistic regression prediction scores for each bin were adjusted based on the scores of its eight nearest neighbors. For example, if a bin showed mixed Germ and pre-Gran prediction scores, but its neighbors display strong Germ prediction score, the Germ score for that bin was increased accordingly. This smoothing step was implemented by computing the matrix multiplication between the neighbor connectivities and logistic regression prediction probability matrices. Finally, bins with low or ambiguous prediction confidence across all classes were also left unassigned. The confidence threshold was adapted to each resolution, as smaller bins contain fewer transcripts and therefore have lower prediction scores.

Surface epithelium path reconstruction.

Gonadal areas and associated bins were manually annotated using Napari. Each gonad was aligned using PCA applied to bins spatial coordinates. Since SLCs are progenitors of the rete cells and are located at the anterior pole of the gonad, the averaged SLC prediction score was used to orient each sample. Only gonadal sections with a visible gonad-mesonephros junction were retained for further analysis.

Bins annotated as coelomic epithelium (CE) were extracted and used to reconstruct a path along the ovarian surface. First, a spatial neighbor graph of CE bins was generated using scikit-learn's kneighbors_graph. Long connections were pruned to avoid artifacts caused by isolated bins or links between opposite CE poles. The shortest paths between all bin pairs were computed using NetworkX (all_pairs_dijkstra_path_length). Terminal points were defined as the pair of bins separated by the longest geodesic distance, and the shortest path between these endpoints was extracted using networkx.shortest_path. All CE bins were then projected onto this path using a KD-tree approach. For each CE bin, the nearest path bin was identified, and distances were computed as cumulative sums of distances from root bin along the path. For visualization, bins were ordered along the obtained pseudospace. Because spatial transcriptomic data are sparse, expression values were smoothed with a generalized additive model (linearGam function of pyGam package), and visualized as heatmaps or curves using seaborn.

Scoring of CE and Gene selection for heatmaps

An alternative scoring model was trained using genes detected in both spatial and single-cell RNA-seq datasets. CE bins were then scored using this model. Heatmaps were generated using differentially expressed genes with a log2 fold-change >1 and a corrected p-value<0.001 between supporting cells and SPs of the scRNAseq atlas. Differential expression was assessed with a rank-sum test (log2FC > 1, adj. $p < 0.01$).

Proportion of CE neighbors

For visualization, cell annotations were merged into two lineage categories: Supporting (pre-gran and pre-Sup.) and stromal (gSP, mSP, GM). For each CE bin in each gonad and developmental stage, the proportion of neighboring cells belonging to each lineage was calculated using k = 100 nearest neighbors. Profiles were then interpolated to normalize path length across gonads and averaged per timepoint for plotting.

Fetal gonad collection and *in vitro* culture

Fetal gonads were collected from C57BL/6 embryos obtained from timed matings (E0.5 = day of vaginal plug). Pregnant females were euthanized by cervical dislocation, embryos were dissected, and gonads were isolated at E11.5. As sex cannot be determined morphologically at this stage, embryos were genotyped by PCR detection of the *Sry* gene using the following primers: FW 5′-GTCAAGCGCCCCATGAATGCAT-3′ and RV 5′-AGTTTGGGTATTTCTCTCTGTG-3′. *Sry*-negative (female) gonads were used for *in vitro* culture.

Gonads were placed on Millicell-CM membrane inserts (0.4 µm) under three conditions: (i) with attached mesonephros, (ii) without mesonephros, or (iii) positioned between two mesonephroi. Explants were cultured for 48 h at 37 °C in 5% CO□ in DMEM/F12 supplemented with 10% FBS and gentamicin (0.04 mg/mL), with medium changed after 24 h. Tissues were then fixed in 4% PFA for immunofluorescence analyses.

Wild-type, transgenic mouse models and cell lineage tracing

All animal procedures conformed to institutional and European regulations and were approved by the Service de la Consommation et des Affaires Vétérinaires (SCAV) of the Canton de Genève (licenses number GE67-19, GE214-19, GE379a and GE35). Animals were maintained on a mixed genetic background under a 12-h light/12-h dark cycle. with ad libitum access to food and water.

The alleles used in this study were previously described: *Wnt5atm1(Tet-On 3G)Nef* , *(tetO)7-CMV-Cre (TetO:Cre)* *Pax8tm1.1(cre)Mbu/J* and *Gt(ROSA)26Sortm9(CAG-tdTomato)Hze* (Ademi et al., 2022; Bouchard et al., 2002; Madisen et al., 2010; Perl et al., 2002). For constitutive tracing of *Pax8*$^{+}$cells, we generated *Pax8*^{lin} mice (*Pax8tm1.1(cre)Mbu/J*$^{+/ki}$;*Gt(ROSA)26Sortm9(CAG-tdTomato)HzeR26RFP/*$^{ki/ki}$). For inducible lineage tracing of *Wnt5a+* cells, we generated *Wnt5a*^{lin} mice (*Wnt5atm1(Tet-On3G)Nef*$^{+/ki}$*;(tetO)7-CMV-Cre/*$^{+/ki}$*;Gt(ROSA)26Sortm9(CAG-tdTomato)HzeR26RFP/*$^{+/ki}$). For *Wnt5a*^{lin}, doxycycline (DOX) was delivered in drinking water (pH 6.5) at 2 mg/mL during E11.5-12.5 or E15.5-E16.5. For spatial transcriptome experiments, CD-1 outbred mice were housed in an AALAC-accredited barrier facility at the University of Colorado Anschutz Medical campus. All mice were handled and cared for following accepted National Institute of Health guidelines. All experiments were conducted with the approval of the University of Colorado Anschutz Medical Campus Institutional Animal Care and Use Committee (IACUC protocol # 01262). Embryos were staged by detection of a vaginal plug (E0.5), postnatal stages were assigned by day of birth (P0). Ovaries were collected at fetal and adult time points as indicated to follow the fate of *Pax8+* or *Wnt5a+* cells across development.

Tissue processing, immunofluorescence and *in situ* hybridization analyses

Embryonic samples were collected at defined stages, fixed in 4% (*w/v*) paraformaldehyde overnight, processed for paraffin embedding, and sectioned into 5 µm sections. Immunofluorescence and Hoechst 33342 (H3570, Invitrogen) staining were performed as described in (Tang et al., 2020) or (Ademi et al., 2022). Primary antibodies included: mouse anti-NR2F2 (PP-H7147-00, R&D, 1/200), rabbit anti-RUNX1 (ab92336, Abcam, 1/500), rabbit anti-LAMA1 (L9303, SIGMA, 1/200), mouse anti-ß-Catenin (BD-610153, BD Biosciences, 1/100), rabbit anti-FOXL2 (1:500, gift from Dagmar Wilhelm, University of Melbourne, Australia), mouse anti-αSMA (1:500, gift from Christine Chaponnier, University of Geneva, Switzerland), rabbit anti-PAX8 (1:500, 10336-1-AP, Proteintech), mouse anti-RFP (DsRed) (1:250, sc-390909, SantaCruz Biotechnology, USA), goat anti-RFP (1:50, 200-101-379, Rockland Immunochemicals Inc., USA), rabbit anti-RFP (1:250, 600-401-379, Rockland Immunochemicals Inc., USA) and rabbit anti-CYP11A1 (1:500, gift from Dagmar Wilhelm, University of Melbourne, Australia).
RNAscope *in situ* hybridization was performed using the RNAscope Multiplex Fluorescent Reagent Kit v2 (Advanced Cell Diagnostic) using probes for *Bmp2* (406661-C3) and *Upk3b* (568561) following the manufacturer's instructions.Fluorescent images were acquired on Axio

Imager M2 or Z1 microscope (ZEISS, Germany) equipped with Axiocam 702 mono camera or MRm camera (ZEISS, Germany), and or on a Zeiss LSM 780 confocal microscope with a 40x/1.2 W objective. Images were minimally adjusted for global levels in ZEN (ZEISS), processed in Fiji (Bethesda, MD, USA) and assembled using OMERO (https://www.openmicroscopy.org/omero/) or Affinity.

Immunofluorescence quantification and statistical analysis

Cell quantifications were performed using QuPath-0.5.1. Nuclei were detected by DAPI intensity, cells were classified as positive/negative based on mean signal intensity thresholds for FOXL2, tdTomato, NR2F2 or RUNX1 (thresholds set per experiment). For each condition, 2-5 biological replicates (whole-ovary sections) were analyzed. Results are reported as the percentage of tdTomato$^+$, NR2F2$^+$, or RUNX1$^+$ cells (Mean ± SD).

Statistical analyses were performed in GraphPad Prism (version 10). Pairwise comparisons used two-tailed Mann-Whitney U tests. Significance was defined as $P < 0.05$. n biological replicates and statistical details are provided in Figure legends.

Whole-mount immunofluorescence for confocal imaging

E12.5 ovary/mesonephros complexes were collected and fixed in 4% paraformaldehyde 30 min at room temperature, dehydrated through a methanol/PBS gradient (25%, 50%, 75%, 100% methanol), and stored overnight in 100% methanol at -20°C. Samples were then rehydrated through the reverse methanol/PBS gradient, permeabilized in PBS/0.1% Triton X-100 for 30 min, and blocked for 1 h in PBS/0.1% Triton X-100/10% horse serum. Samples were incubated overnight at 4°C with primary antibodies diluted in blocking solution: mouse anti-NR2F2 (PP-H7147-00, R&D, 1/200) and rabbit anti-RUNX1 (ab92336, Abcam, 1/500). After washes in PBS/0.1% Triton X-100, samples were incubated overnight at 4°C with fluorophore-conjugated secondary antibodies (1:500). Samples were mounted in CUBIC Reagent II containing 2% agarose and imaged on a spinning-disk confocal microscope using a 20x/0.8 objective. Images were minimally adjusted for global levels and processed in Imaris.

Additional resources

An interactive viewer will be available upon publication for further exploration [https://www.unige.ch/medecine/nef/datasets].

Acknowledgments

We thank Coralie Gobbo, Karim Hammad and Violaine Regard for technical assistance. We are grateful to the Bioimaging platform and to the animal facility at the Faculty of Medicine, University of Geneva, for their support. We also thank the members of the Nef laboratory for helpful discussions.

Funding

Swiss National Science Foundation grant 31003A_173070 (SN)
Swiss National Science Foundation grant 310030_200316 (SN)
Swiss National Science Foundation grant 10.001.251 (SN)
Département de l'Instruction Publique of the State of Geneva (SN)
Agence Nationale de la Recherche grant ANR-23-CE14-0012 HeteroSex (MCC)
Agence Nationale de la Recherche grant ANR-24-CE14-2476 Anaform (MCC)
CU Anschutz Medical Campus, Department of Pediatrics (JM)
National Institutes of Health grant #R00HD103778 (JM)
Agence Nationale de la Recherche grant ANR- 22-CE17-0059 (GL)
Electricité de France-V3.203 (GL)

Author Contributions (CRediT)

Conceptualization: CD, CM, SN; data curation: CD, CM, MG; formal analysis: CD, CM, MG; Investigation: CD, CM, AG, LB, PB, HA, AR, FK, CG; methodology: CD, CM, AG, LB; visualization: CD, CM, MG, AG; resources: AM, TG, JM, MCC, GL; funding acquisition: SN, MCC, JM, GL; supervision: SN; writing - original draft: CD, SN; writing - review & editing: CD, CM, MG, AG, SN, MCC, JM, DW.

Competing Interests

The authors declare no competing interests.

Data and materials availability

All data needed to evaluate the paper are present in the paper and in the Supplementary Materials. The scRNAseq datasets are available at NCBI Gene Expression Omnibus GEO: GSE184708.

References

Ademi, H., Djari, C., Mayere, C., Neirijnck, Y., Sararols, P., Rands, C.M., Stevant, I., Conne, B., and Nef, S. (2022). Deciphering the origins and fates of steroidogenic lineages in the mouse testis. Cell Rep *39*, 110935.

Barsoum, I.B., and Yao, H.H. (2010). Fetal Leydig cells: progenitor cell maintenance and differentiation. J Androl *31*, 11–15.

Bouchard, M., Souabni, A., Mandler, M., Neubuser, A., and Busslinger, M. (2002). Nephric lineage specification by Pax2 and Pax8. Genes Dev *16*, 2958–2970.

Brennan, J., Tilmann, C., and Capel, B. (2003). Pdgfr-alpha mediates testis cord organization and fetal Leydig cell development in the XY gonad. Genes Dev *17*, 800–810.

Bunce, C., Barske, L., Zhang, G., and Capel, B. (2023). Biased precursor ingression underlies the center-to-pole pattern of male sex determination in mouse. Development *150*.

Cai, H., Liu, B., Wang, H., Sun, G., Feng, L., Chen, Z., Zhou, J., Zhang, J., Zhang, T., He, M., *et al.* (2020). SP1 governs primordial folliculogenesis by regulating pregranulosa cell development in mice. J Mol Cell Biol *12*, 230–244.

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.

DeFalco, T., Takahashi, S., and Capel, B. (2011). Two distinct origins for Leydig cell progenitors in the fetal testis. Dev Biol *352*, 14–26.

Eliveld, J., van den Berg, E.A., Chikhovskaya, J.V., van Daalen, S.K.M., de Winter-Korver, C.M., van der Veen, F., Repping, S., Teerds, K., and van Pelt, A.M.M. (2019). Primary human testicular PDGFRalpha+ cells are multipotent and can be differentiated into cells with Leydig cell characteristics in vitro. Hum Reprod *34*, 1621–1631.

Estermann, M.A., Grimm, S.A., Kitakule, A.S., Rodriguez, K.F., Brown, P.R., McClelland, K., Amato, C.M., and Yao, H.H. (2025). NR2F2 regulation of interstitial cell fate in the embryonic mouse testis and its impact on differences of sex development. Nat Commun *16*, 3987.

Feng, L., Wang, Y., Cai, H., Sun, G., Niu, W., Xin, Q., Tang, X., Zhang, J., Wang, C., Zhang, H., *et al.* (2016). ADAM10-Notch signaling governs the recruitment of ovarian pregranulosa cells and controls folliculogenesis in mice. J Cell Sci *129*, 2202–2212.

Fortune, J.E., and Armstrong, D.T. (1977). Androgen production by theca and granulosa isolated from proestrous rat follicles. Endocrinology *100*, 1341–1347.

Garcia-Alonso, L., Lorenzi, V., Mazzeo, C.I., Alves-Lopes, J.P., Roberts, K., Sancho-Serra, C., Engelbert, J., Marečková, M., Gruhn, W.H., Botting, R.A., *et al.* (2022). Single-cell roadmap of human gonadal development. Nature *607*, 540–547.

Gayoso, A., Lopez, R., Xing, G., Boyeau, P., Valiollah Pour Amiri, V., Hong, J., Wu, K., Jayasuriya, M., Mehlman, E., Langevin, M., *et al.* (2022). A Python library for probabilistic analysis of single-cell omics data. Nature Biotechnology *40*, 163–166.

Ge, R.S., Dong, Q., Sottas, C.M., Papadopoulos, V., Zirkin, B.R., and Hardy, M.P. (2006). In search of rat stem Leydig cells: identification, isolation, and lineage-specific development. Proc Natl Acad Sci U S A *103*, 2719–2724.

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

Kilcoyne, K.R., Smith, L.B., Atanassova, N., Macpherson, S., McKinnell, C., van den Driesche, S., Jobling, M.S., Chambers, T.J., De Gendt, K., Verhoeven, G., *et al.* (2014). Fetal programming of adult Leydig cell function by androgenic effects on stem/progenitor cells. Proc Natl Acad Sci U S A *111*, E1924–1932.

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.

Lardenois, A., Suglia, A., Moore, C.L., Evrard, B., Noël, L., Rivaud, P., Besson, A., Toupin, M., Blévinal, J., Dumortier, C., *et al.* (2026). Single-cell exploration of gonadal somatic cell lineage specification during human sex determination. Dev Cell *61*, 400–415.e406.

Lin, Y.T., Barske, L., DeFalco, T., and Capel, B. (2017). Numb regulates somatic cell lineage commitment during early gonadogenesis in mice. Development *144*, 1607–1618.

Liu, C., Peng, J., Matzuk, M.M., and Yao, H.H. (2015). Lineage specification of ovarian theca cells requires multicellular interactions via oocyte and granulosa cells. Nat Commun *6*, 6934.

Liu, C., Rodriguez, K., and Yao, H.H. (2016). Mapping lineage progression of somatic progenitor cells in the mouse fetal testis. Development *143*, 3700–3710.

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., *et al.* (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat Neurosci *13*, 133–140.

Magoffin, D.A. (2005). Ovarian theca cell. Int J Biochem Cell Biol *37*, 1344–1349.

Magoffin, D.A., and Weitsman, S.R. (1994). Insulin-like growth factor-I regulation of luteinizing hormone (LH) receptor messenger ribonucleic acid expression and LH-stimulated signal transduction in rat ovarian theca-interstitial cells. Biol Reprod *51*, 766–775.

Mannan, M.A., and O'Shaughnessy, P.J. (1991). Steroidogenesis during postnatal development in the mouse ovary. J Endocrinol *130*, 101–106.

Martinez, A.S., Gibson, T.J., and McKey, J. (2025). Single-cell-scale spatial transcriptome of the developing and adult mouse ovary. bioRxiv, 2025.2009.2016.676655.

Mayere, C., Regard, V., Perea-Gomez, A., Bunce, C., Neirijnck, Y., Djari, C., Bellido-Carreras, N., Sararols, P., Reeves, R., Greenaway, S., *et al.* (2022). Origin, specification and differentiation of a rare supporting-like lineage in the developing mouse gonad. Sci Adv *8*, eabm0972.

McKey, J., Anbarci, D.N., Bunce, C., Ontiveros, A.E., Behringer, R.R., and Capel, B. (2022). Integration of mouse ovary morphogenesis with developmental dynamics of the oviduct, ovarian ligaments, and rete ovarii. Elife *11*.

Miyabayashi, K., Katoh-Fukui, Y., Ogawa, H., Baba, T., Shima, Y., Sugiyama, N., Kitamura, K., and Morohashi, K. (2013). Aristaless related homeobox gene, Arx, is implicated in mouse fetal Leydig cell differentiation possibly through expressing in the progenitor cells. PLoS ONE *8*, e68050.

Mork, L., Maatouk, D.M., McMahon, J.A., Guo, J.J., Zhang, P., McMahon, A.P., and Capel, B. (2012). Temporal differences in granulosa cell specification in the ovary reflect distinct follicle fates in mice. Biol Reprod *86*, 37.

Neirijnck, Y., Sararols, P., Kühne, F., Mayère, C., Weerasinghe Arachchige, L.C., Regard, V., Nef, S., and Schedl, A. (2023). Single-cell transcriptomic profiling redefines the origin and specification of early adrenogonadal progenitors. Cell Rep *42*, 112191.

Nicol, B., Estermann, M.A., Yao, H.H., and Mellouk, N. (2022). Becoming female: Ovarian differentiation from an evolutionary perspective. Front Cell Dev Biol *10*, 944776.

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 U S A *117*, 20015–20026.

Perea-Gomez, A., Bellido Carreras, N., Dhellemmes, M., Tang, F., Le Gallo, C., and Chaboissier, M.C. (2025). NR2F2 is required in the embryonic testis for fetal Leydig cell development. Elife *14*.

Perl, A.K., Wert, S.E., Nagy, A., Lobe, C.G., and Whitsett, J.A. (2002). Early restriction of peripheral and proximal cell lineages during formation of the lung. Proc Natl Acad Sci U S A *99*, 10482–10487.

Qin, J., Tsai, M.J., and Tsai, S.Y. (2008). Essential roles of COUP-TFII in Leydig cell differentiation and male fertility. PLoS ONE *3*, e3285.

Rastetter, R.H., Bernard, P., Palmer, J.S., Chassot, A.A., Chen, H., Western, P.S., Ramsay, R.G., Chaboissier, M.C., and Wilhelm, D. (2014). Marker genes identify three somatic cell types in the fetal mouse ovary. Dev Biol *394*, 242–252.

Richards, J.S., Ren, Y.A., Candelaria, N., Adams, J.E., and Rajkovic, A. (2018). Ovarian Follicular Theca Cell Recruitment, Differentiation, and Impact on Fertility: 2017 Update. Endocr Rev *39*, 1–20.

Rossitto, M., Déjardin, S., Rands, C.M., Le Gras, S., Migale, R., Rafiee, M.R., Neirijnck, Y., Pruvost, A., Nguyen, A.L., Bossis, G., *et al.* (2022). TRIM28-dependent SUMOylation protects the adult ovary from activation of the testicular pathway. Nat Commun *13*, 4412.

Saunders, P.A., and Veyrunes, F. (2021). Unusual Mammalian Sex Determination Systems: A Cabinet of Curiosities. Genes *12*, 1770.

Schmahl, J., Eicher, E.M., Washburn, L.L., and Capel, B. (2000). Sry induces cell proliferation in the mouse gonad. Development *127*, 65–73.

Schmahl, J., Yao, H.H., Pierucci-Alves, F., and Capel, B. (2003). Colocalization of WT1 and cell proliferation reveals conserved mechanisms in temperature-dependent sex determination. Genesis *35*, 193–201.

Shen, Y.C., Shami, A.N., Moritz, L., Larose, H., Manske, G.L., Ma, Q., Zheng, X., Sukhwani, M., Czerwinski, M., Sultan, C., *et al.* (2021). TCF21(+) mesenchymal cells contribute to testis somatic cell development, homeostasis, and regeneration in mice. Nat Commun *12*, 3876.

Stevant, I., Kuhne, 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 e3273.

Stevant, I., and Nef, S. (2018). Single cell transcriptome sequencing: A new approach for the study of mammalian sex determination. Mol Cell Endocrinol *468*, 11–18.

Stevant, I., and Nef, S. (2019). Genetic Control of Gonadal Sex Determination and Development. Trends Genet *35*, 346–358.

Tang, F., Richardson, N., Albina, A., Chaboissier, M.C., and Perea-Gomez, A. (2020). Mouse Gonad Development in the Absence of the Pro-Ovary Factor WNT4 and the Pro-Testis Factor SOX9. Cells *9*.

Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., *et al.* (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196.

Wilhelm, D., Perea-Gomez, A., Newton, A., and Chaboissier, M.C. (2025). Gonadal sex determination in vertebrates: rethinking established mechanisms. Development *152*.

Wilhelm, D., Yang, J.X., and Thomas, P. (2013). Chapter Three - Mammalian Sex Determination and Gonad Development. In Current Topics in Developmental Biology, P. Thomas, ed. (Academic Press), pp. 89–121.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol *19*, 15.

Wolf, F.A., Hamey, F.K., Plass, M., Solana, J., Dahlin, J.S., Gottgens, B., Rajewsky, N., Simon, L., and Theis, F.J. (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol *20*, 59.

Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Syst *8*, 281–291 e289.

Zheng, W., Zhang, H., Gorre, N., Risal, S., Shen, Y., and Liu, K. (2014). Two classes of ovarian primordial follicles exhibit distinct developmental dynamics and physiological functions. Hum Mol Genet *23*, 920–928.

91

Figure legends

Figure 1. Single-cell transcriptomic atlas of the developing human and mouse ovaries reveals fate-biased heterogeneity in the coelomic epithelium and dual origin of granulosa cells.

(A-D) UMAP representations of single somatic cells from human (n = 156,408) and mouse (n = 43,053) datasets. Human cells are colored by post-conceptional week (PCW) (A) and cell annotation (B); mouse cells are colored by embryonic and postnatal day (C) and cell annotation (D). In the annotated UMAPs, Conductance-based Partition-based Graph Abstraction (cPAGA) is overlaid to depict connectivity between somatic cell populations and across developmental stages. Edge width reflects the conductance score. Cell type abbreviations: CE, coelomic epithelium; SE, surface epithelium; SLC, supporting-like cells; pre-Sup, pre-supporting cells; pre-Gran, pre-granulosa cells; Gran, granulosa cells; SP, stromal progenitors; Stero, steroidogenic cells; SM, smooth muscle cells; PV, perivascular cells; Luteal, luteal cells.

(E and F) Stage-wise UMAP representations colored by cell type, with cPAGA connectivity between cell types in human (E) and mouse (F).

(G-H) Bar plots (left Y-axis) showing the proportion of CE/SE, pre-Gran, SP, and SLC populations across developmental stages in human (G) and mouse (H). Blue line plots (right Y-axis) indicate the proportion of proliferating cells (G2/M phase) within each population.

(I and K) UMAPs of human (I) and mouse (K) somatic cells colored by expression of CE markers (*UPK3B/Upk3b*), supporting markers (*IRX3/Wnt6*), and SP markers (*PDGFRA/Tcf21*).

(J and L) UMAP representations of human (J) and mouse (L) somatic cells. CE/SE cells are colored by scaled SP (turquoise channel) and supporting (red channel) identity scores. Legend squares indicate color channels, with SP scores mapped to the X-axis and supporting scores to the Y-axis. White diagonal lines indicate equivalent SP and supporting scores. White crosses represent the median score within CE cells.

(M and N) Heatmap showing the proportion of cells expressing SP or supporting markers genes in the CE/SE population across developmental stages in human (M) and mouse (N).

(O and P) Lineplot showing median supporting and SP identity scores with 95% confidence intervals across developmental stages in human (O) and mouse (P).

(Q and R) Stage-specific UMAP representations of human (Q) and mouse (R) somatic cells. CE/SE cells are colored according to SP (turquoise channel) and supporting (red channel) identity scores. Legend squares indicate color channels as above. White diagonal lines represent equivalent identity scores; white crosses mark medians.

Figure 2. Dynamic spatial patterning of the coelomic epithelium during ovarian development.

(A) Spatial transcriptomic experimental design using 10x Visium HD technology on parasagittal ovaries/mesonephros sections. Gonads and CE were annotated, CE paths reconstructed and gene expression/score computed along a pseudospace.

(B) Example of stromal and supporting identity score projected along the CE pseudo spatial axis of a developing ovary at E12.5.

(C) Lineplots of smoothed stromal- and supporting- identity scores along CE pseudospace at E12.5, E14.5 and E16.5 (Mean with 95% confidence intervals). In the pseudospace, the posterior pole of the ovary is set at 0 and the anterior pole at 1.

(D) Lineplots of smoothed *Nr2f2* and *Runx1* expression along CE pseudospace at E12.5, E14.5 and E16.5 (Mean with 95% confidence intervals). Posterior = 0; anterior = 1.

(E) Heatmap of gene expression along CE pseudo-space at E12.5 for differentially expressed genes between supporting-cells (pre-Gran, Gran) (n=300) and stromal progenitors (n=243) from the scRNA-seq atlas (log2FC > 1, adj. $p < 0.01$). Genes are ordered by pattern similarity.

(F) Immunofluorescence for RUNX1 (supporting) and NR2F2 (SP) in developing ovaries at E11.5, E12.5, E13.5 and E16.5. Dotted lines outline the CE; dashed line marks the boundary between the gonadal and mesonephric compartments.

(G and H) Quantification of NR2F2□ cells (G) or RUNX1^{+} cells (H) as a percentage of the CE cells at E11.5, E12.5 and E13.5 on transverse sections, and at E16.5 on longitudinal sections.

(I) Maximum intensity projection (left) and 3D model generated by isosurface segmentation (right) from whole-mount 3D imaging of E12.5 ovaries stained for NR2F2 and RUNX1.

(J) Experimental design of E11.5 XX gonad-mesonephros explant culture. Gonads were cultured for 48h either with the endogenous mesonephros attached (control), after surgical removal of the mesonephros, or with an additional mesonephros placed along the ventral side of the gonad to generate a one-gonad/two-mesonephroi configuration.

(K) Co-immunofluorescence for RUNX1 (supporting), NR2F2 (SP), and cytokeratin 7 (CK7, epithelial cells) in E11.5 ovaries +/- mesonephros after 2 days of culture.

Scale bars, 100µm.

Abbreviation: CE/SE, coelomic epithelium; SP, stromal progenitors. A, anterior pole; P, posterior pole. Mes, mesonephros.

Figure 3. Fate-primed ovarian surface epithelium organizes intragonadal domains

(A) Combined ISH/IF for *Upk3b* (CE/SE), *Bmp2* (supporting) and NR2F2 (stromal progenitor, SP) on developing ovaries at E11.5, E12.5, E13.5 and E16.5. White dotted lines: CE, white dashed lines: boundary between the gonad and the mesonephros, green dashe lines: intragonadal region enriched in NR2F2^{+} cells; magenta dashed lines: intragonadal region enriched in *Bmp2*$^{+}$ cells.

(B) Spatial representation of the 4µm bins of a representative E14.5 gonad colored by cell-type signature.

(C) Stacked-area plots showing the proportion of intragonadal neighbors (pre-granulosa, Stromal Progenitors (SP)) among the k=100 nearest spatial neighbors (k-NN) of CE bins along pseudospace. (D) Co-IF for β-catenin and LAMA1 at E11.5, E12.5, E13.5 and E16.5. Arrowheads indicate basal lamina discontinuities.

Scale bars, 100µm. Abbreviations: CE, coelomic epithelium; EC, erythrocytes; egEpith, extra-gonadal epithelium; End, endothelium; Germ, germ cells; pre-Gran, pre-granulosa cells; PV, perivascular cells; SLC, supporting-like cells; SP, stromal progenitors.

Figure 4. Stage-specific *Wnt5a* lineage tracing reveals CE-derived granulosa cells and fetal SP-derived steroidogenic cells

(A) Mean *Wnt5a* expression across mouse ovarian somatic compartments (CE/SE, SLC, Sup., Stromal) from embryonic day (E)10.5 to post-natal day (P)56.

(B) Schematic of the inducible *Wnt5a*^{lin} mouse model (*Wnt5-rtTA;TetO-Cre;R26-tdTomato*). Upon doxycycline (DOX) exposure, Tet-On3G induces Cre expression, excising the STOP cassette in the R26-LSL-*tdTomato* locus and resulting in permanent tdTomato labeling of *Wnt5a*$^{+}$ cells.

(C) Experimental design for *in vivo* lineage-tracing. Pregnant females were either exposed to DOX between E11.5 and E12.5 (early induction) and ovaries were collected at E12.5, E16.5, P5, and 8 weeks or exposed to DOX between E15.5 and E16.5 (late induction) and ovaries were collected at E16.5 and 8 weeks.

(D) Co-immunofluorescence (co-IF) for tdTomato and FOXL2 in *Wnt5a*^{lin} ovaries at E12.5 and E16.5.

(E) Quantification of tdTomato as a percentage of CE cells at E12.5 and E16.5.

(F and G) Quantification of tdTomato□ as a percentage of total FOXL2□ pre-granulosa cells at E12.5 and E16.5 (F), and as a percentage of total FOXL2□ pre-granulosa in growing follicles at 8 weeks (G).

(H) Co-IF for tdTomato and FOXL2 in *Wnt5a*^{lin} ovaries at P5.

(I) Quantification of tdTomato□ cells as a percentage of total FOXL2□ granulosa cells in the ovarian medulla (Me) and cortex (Co) at P5.

(J) Co-IF for tdTomato and NR2F2 in *Wnt5a*^{lin} ovaries at E16.5.

(K) Co-IF for tdTomato with FOXL2, NR2F2 (stromal cells), or CYP11A1 (steroidogenic cells) in *Wnt5a*^{lin} ovaries at 8 weeks.

White lines delineate the CE, red lines delineate the ovarian medulla, yellow lines delineate follicular granulosa cells of the antral follicles (AF), yellow dashed lines delineate the thecal compartment.

Arrowheads indicate tdTomato□/FOXL2□ double-positive cells. Dashed lines delineate the CE/SE. Scale bars, 100 μm. Error bars represent mean ± SD.

Abbreviations: CE, coelomic epithelium; SE, surface epithelium; SLC, supporting-like cells; pre-Sup, pre-supporting cells; pre-Gran, pre-granulosa cells; Gran, granulosa cells; GM, gonadal mesenchyme; SP, stromal progenitors; Stero, steroidogenic cells; SM, smooth muscle cells; PV, perivascular cells.

Figure 5. *Pax8*□ supporting-like cells represent an additional source of granulosa cells

(A) UMAP representation showing *Pax8* expression in ovarian somatic cells.

(B) Mean *Pax8* expression across mouse ovarian somatic compartments (CE/SE, SLC, Sup., Stromal) from embryonic day (E)10.5 to post-natal day (P)56.

(C) Schematic of the constitutive *Pax8*^{lin} mouse model (*Pax8:Cre;R26-tdTomato*). *Pax8:Cre* knock-in drives *Cre* recombinase expression under the control of *Pax8* endogenous promoter. Cre-mediated excision of the STOP cassette at the *R26-LSL-tdTomato* locus permanently labeled *Pax8*$^{+}$ cells and their derived cells with tdTomato.

(D) Experimental design for *in vivo* lineage-tracing. Ovaries were collected at E12.5, E16.5, P5, and 8 weeks.

(E) Co-immunofluorescence (co-IF) for tdTomato and FOXL2 in *Pax8*^{lin} ovaries at E12.5, E16.5 and 8 weeks. (F) Quantification of tdTomato□ cells as a percentage of the CE cells at E12.5 and E16.5.

(G and H) Quantification of tdTomato□ cells as a percentage of total FOXL2□ pre-granulosa cells at E12.5 and E16.5 (G), and as a percentage of total FOXL2□ pre-granulosa in growing follicles at 8 weeks (H).

(I) Co-IF for tdTomato and FOXL2 in *Pax8*^{lin} ovaries at P5.

(J) Quantification of tdTomato□ cells as a percentage of total FOXL2□ granulosa cells in the ovarian medulla (Me) and cortex (Co) at P5.

Arrowheads indicate tdTomato□/FOXL2□ double-positive cells. White lines delineate the CE. Scale bars, 100 μm. Error bars represent mean ± SD.

Abbreviations: CE, coelomic epithelium; SE, surface epithelium; SLC, supporting-like cells; pre-Sup, pre-supporting cells; pre-Gran, pre-granulosa cells; Gran, granulosa cells; GM, gonadal mesenchyme; SP, stromal progenitors; Stero, steroidogenic cells; SM, smooth muscle cells; PV, perivascular cells.

Supplementary Figure S1 related to Figure 1.

(A and B) UMAP representations of single cells from human (n = 281,325) (A) and mouse (n = 68,366) (B) datasets. Cells are colored by dataset, developmental stage (post-conceptional week (PCW) in human, embryonic day (E) and post-natal day (P) in mouse and cell type annotation.

(C and D) DotPlots of relevant marker genes used to annotate cell populations in human (C) and mouse (D). Dot size represents the percentage of cells expressing the marker gene within a population, and color scale indicates normalized expression level.

(E and F) UMAP colored by level of expression of representative supporting markers (*FOXL2/Foxl2*, *BMP2/Bmp2*, *CYP19A1*, *LHX2*, *Runx1*, *Lgr5)* and stromal markers (*NR2F2*, *Wnt5a*, *Arx*, *Pdgfra*) in human (E) and mouse (F) datasets.

(G and H) Barplots representing the weights of the top 50 genes of supporting or SP identity models in human (G) and mouse (H).

(I and J) Stage-wise UMAP representations colored by cell type with cPAGA connectivity between cell types in human (I) and mouse (J).

Cell type abbreviations: CE, coelomic epithelium; SE, surface epithelium; SLC, supporting-like cells; pre-Sup, pre-supporting cells; GM, gonadal mesenchyme; pre-Gran, pre-granulosa cells; Gran, granulosa cells; SP, stromal progenitors; Stero, steroidogenic cells; SM, smooth muscle cells; PV, perivascular cells; Luteal, luteal cells; MT, mesonephric tubules; MM, mesonephric mesenchyme; Germ, germ cells; AS, Adreno Sympathic ; End, endothelial cells; Eryt, erythrocytes; Immune, immune cells.

Supplementary Figure S2 related to Figure 2.

Lineplots of smoothed expression level of stromal markers (*Tcf21*, *Pdgfra*, *Arx* and *Wnt5a*) and supporting markers (*Wnt6*, *Bmp2*, *Irx3*, *Lgr5* and *Fst*) along CE pseudospace at E12.5, E14.5 and E16.5 (Mean with 95% confidence intervals). In the pseudospace representation, the posterior pole of the ovary is set at 0 and the anterior pole at 1.

Supplementary Figure S3 related to Figure 3.

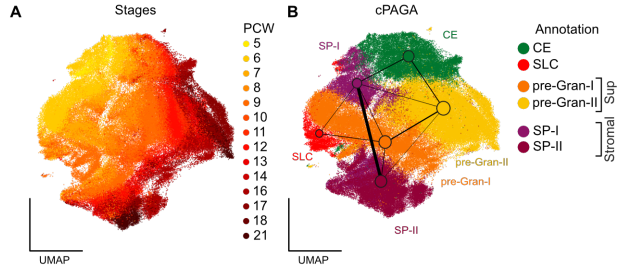
(A) Co-IF for GATA4 and PAX8 in E12.5 and E16.5 fetal ovaries. White line: ovarian domain, dashed line: mesonephric domain.

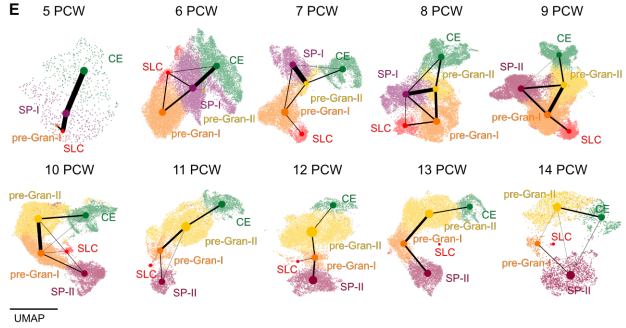
(B) IF for PAX8 in P0, P10 and adult *Pax8*^{lin} ovaries showing PAX8 expression restricted to the rete ovarii.

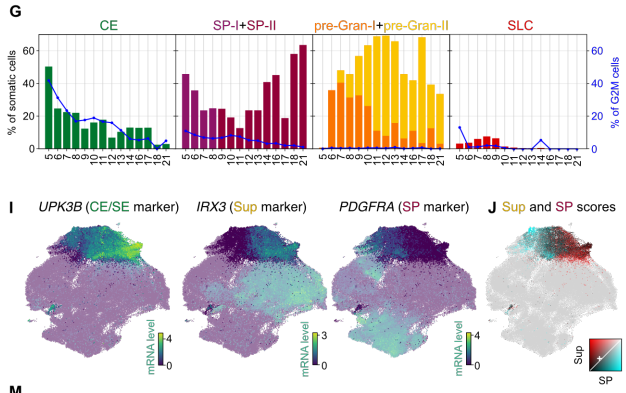
Figure 1

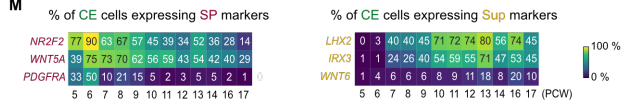
HUMAN somatic cells

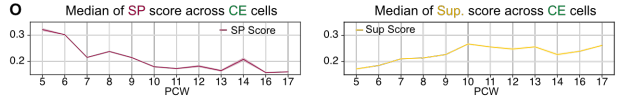
156,408 cells

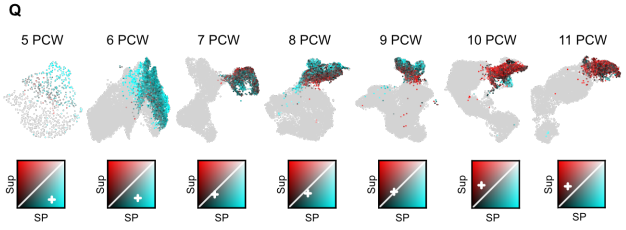






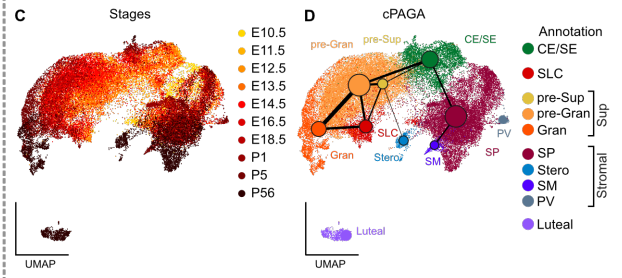


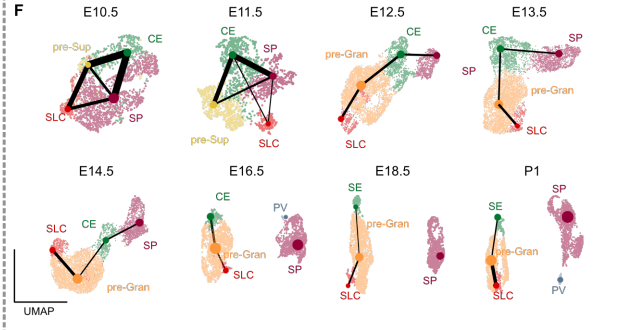


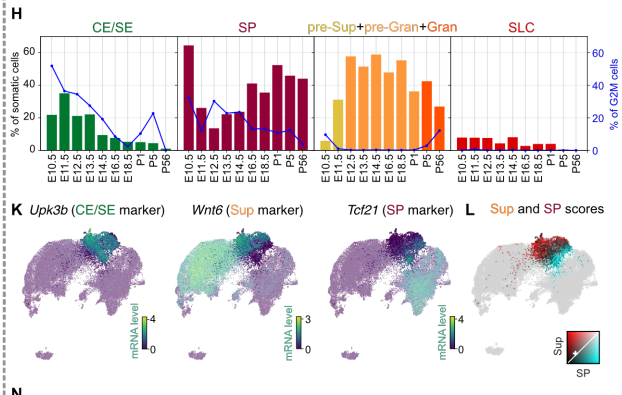


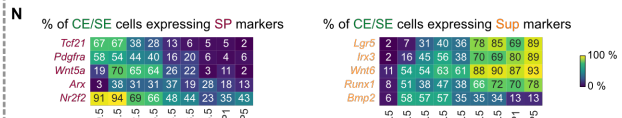
MOUSE somatic cells

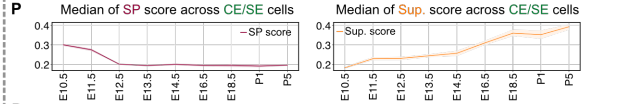
43,053 cells











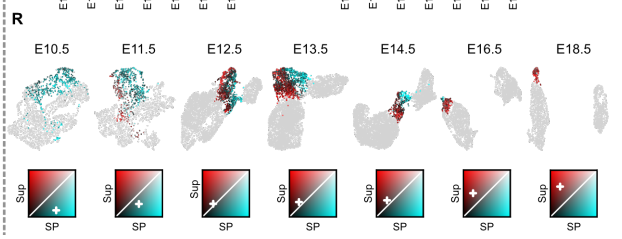


Figure 2

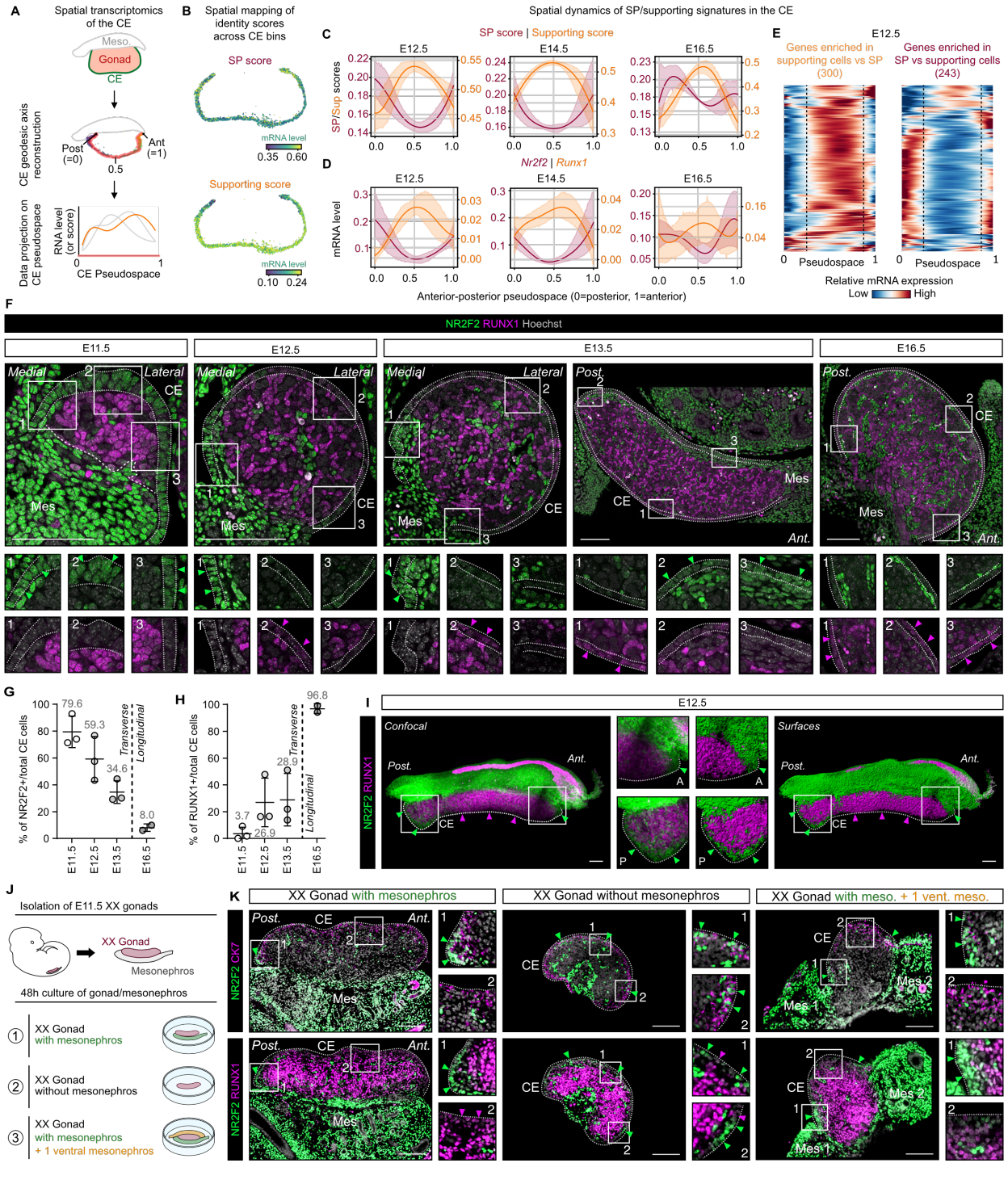

A
Spatial transcriptomics of the CE
Meso.
Gonad
CE
CE geodesic axis reconstruction
Post (=0)
Ant (=1)
0.5
Data projection on CE pseudospace
RNA level (or score)
0
1
CE Pseudospace
B
Spatial mapping of identity scores across CE bins
SP score
mRNA level
0.35 0.60
Supporting score
mRNA level
0.10 0.24
Spatial dynamics of SP/supporting signatures in the CE
C
SP score | Supporting score
E12.5
E14.5
E16.5
SP/Sup scores
D
Nr2f2 | Runx1
mRNA level
Anterior-posterior pseudospace (0=posterior, 1=anterior)
E
E12.5
Genes enriched in supporting cells vs SP (300)
Genes enriched in SP vs supporting cells (243)
Pseudospace
Relative mRNA expression
Low
High
F
NR2F2 RUNX1 Hoechst
E11.5
E12.5
E13.5
E16.5
Medial
Lateral
Post.
Ant.
CE
Mes
G
% of NR2F2+/total CE cells
79.6
59.3
34.6
8.0
Transverse
Longitudinal
E11.5
E12.5
E13.5
E16.5
H
% of RUNX1+/total CE cells
3.7
26.9
28.5
96.8
I
E12.5
Confocal
Surfaces
NR2F2 RUNX1
Post.
Ant.
CE
A
P
J
Isolation of E11.5 XX gonads
XX Gonad
Mesonephros
48h culture of gonad/mesonephros
① XX Gonad with mesonephros
② XX Gonad without mesonephros
③ XX Gonad with mesonephros + 1 ventral mesonephros
K
XX Gonad with mesonephros
XX Gonad without mesonephros
XX Gonad with meso. + 1 vent. meso.
NR2F2 CK7
NR2F2 RUNX1
Post.
Ant.
CE
Mes
Mes 1
Mes 2

Figure 3

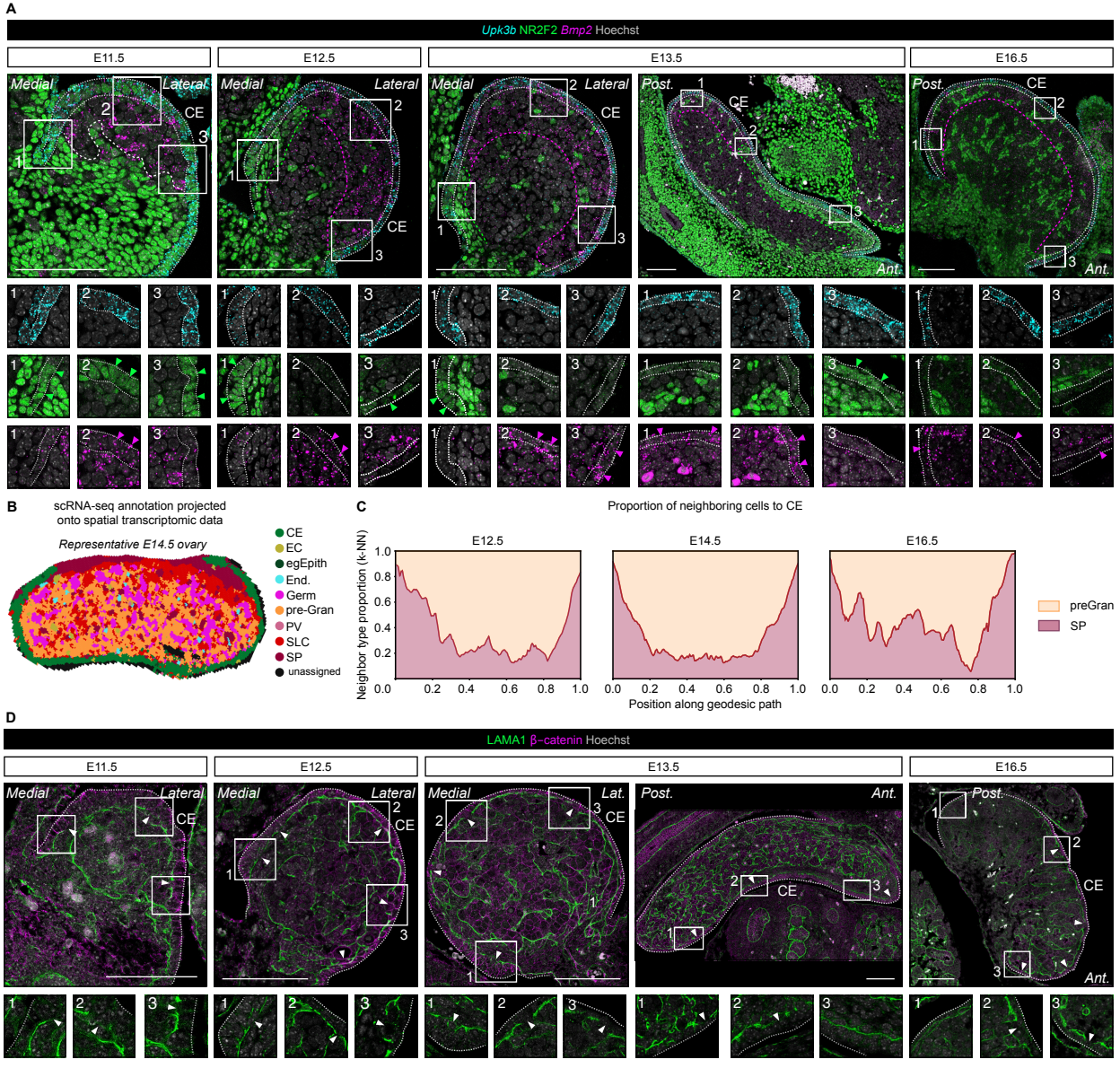

A
Upk3b NR2F2 Bmp2 Hoechst
E11.5
E12.5
E13.5
E16.5
Medial
Lateral
Post.
Ant.
CE
B
scRNA-seq annotation projected onto spatial transcriptomic data
Representative E14.5 ovary
CE
EC
egEpith
End.
Germ
pre-Gran
PV
SLC
SP
unassigned
C
Proportion of neighboring cells to CE
E12.5
E14.5
E16.5
Neighbor type proportion (k-NN)
Position along geodesic path
preGran
SP
D
LAMA1 β-catenin Hoechst
Lat.

Figure 4

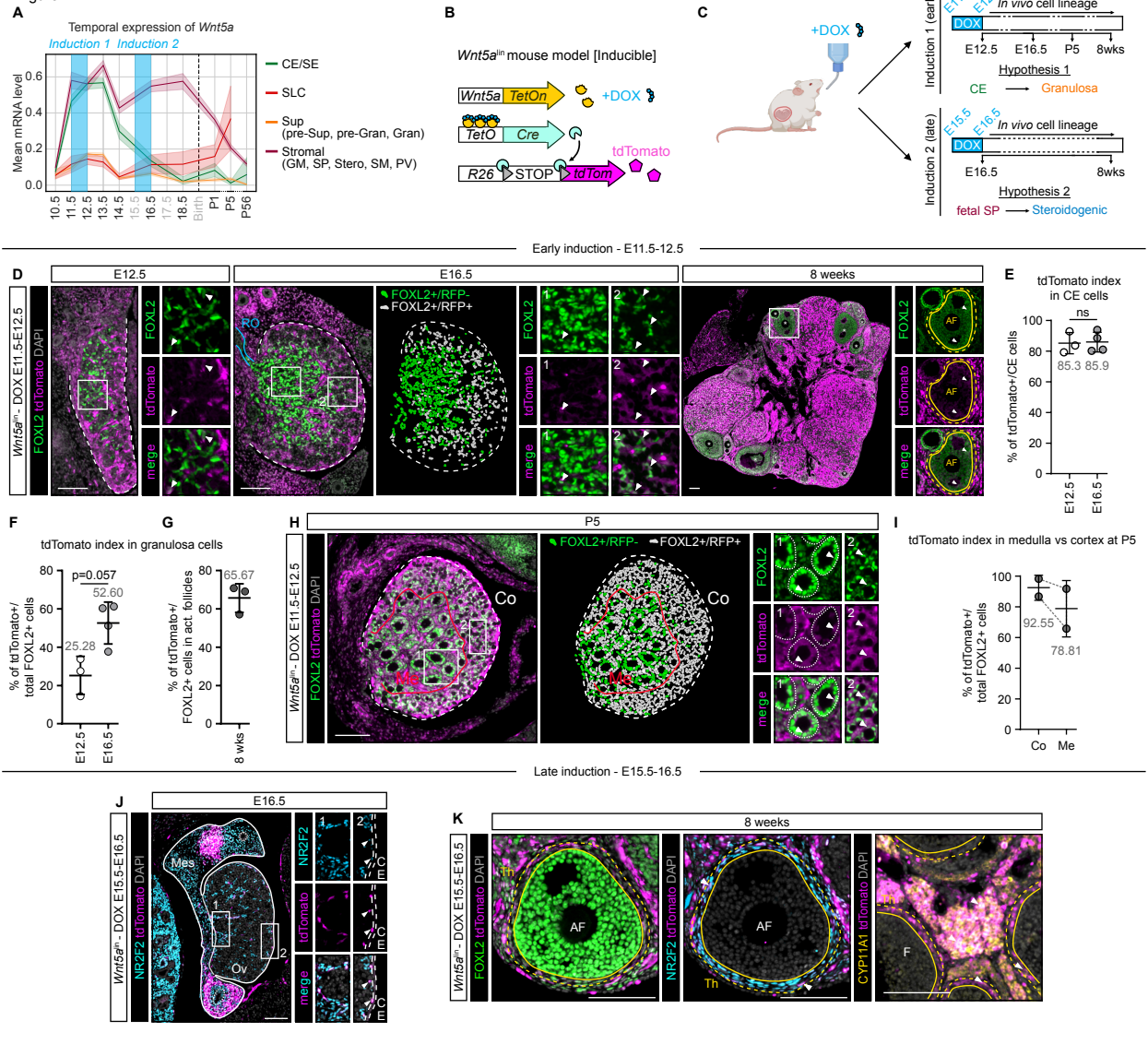

A
Temporal expression of Wnt5a
Induction 1
Induction 2
Mean mRNA level
CE/SE
SLC
Sup (pre-Sup, pre-Gran, Gran)
Stromal (GM, SP, Stero, SM, PV)
B
Wnt5a^lin mouse model [Inducible]
Wnt5a TetOn
+DOX
TetO Cre
R26 STOP tdTom
tdTomato
C
+DOX
Induction 1 (early)
In vivo cell lineage
DOX
E12.5 E16.5 P5 8wks
Hypothesis 1
CE → Granulosa
Induction 2 (late)
In vivo cell lineage
DOX
E16.5 8wks
Hypothesis 2
fetal SP → Steroidogenic
Early induction - E11.5-12.5
D
E12.5
E16.5
8 weeks
Wnt5a^lin- DOX E11.5-E12.5
FOXL2 tdTomato DAPI
FOXL2
tdTomato
merge
RO
FOXL2+/RFP-
FOXL2+/RFP+
AF
E
tdTomato index in CE cells
ns
% of tdTomato+/CE cells
85.3 85.9
E12.5
E16.5
F
tdTomato index in granulosa cells
% of tdTomato+/ total FOXL2+ cells
p=0.057
25.28
52.60
E12.5
E16.5
G
% of tdTomato+/ FOXL2+ cells in act. follicles
65.67
8 wks
H
P5
Wnt5a^lin- DOX E11.5-E12.5
FOXL2 tdTomato DAPI
Co
Me
FOXL2+/RFP-
FOXL2+/RFP+
I
tdTomato index in medulla vs cortex at P5
% of tdTomato+/ total FOXL2+ cells
92.55
78.81
Co
Me
Late induction - E15.5-16.5
J
E16.5
Wnt5a^lin- DOX E15.5-E16.5
NR2F2 tdTomato DAPI
Mes
Ov
NR2F2
tdTomato
mege
C
E
K
8 weeks
Wnt5a^lin- DOX E15.5-E16.5
FOXL2 tdTomato DAPI
Th
AF
NR2F2 tdTomato DAPI
CYP11A1 tdTomato DAPI
F

Figure 5

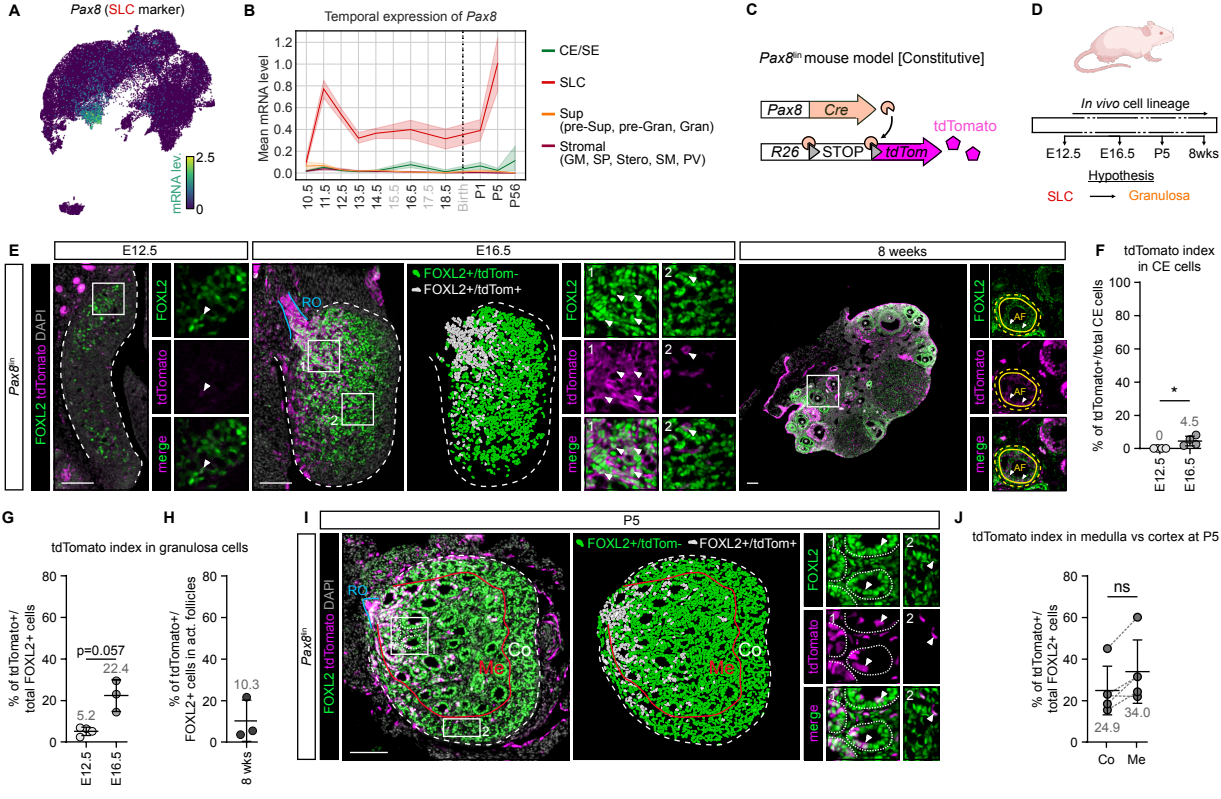

A
Pax8 (SLC marker)
mRNA lev.
2.5
0
B
Temporal expression of Pax8
Mean mRNA level
1.2
1.0
0.8
0.6
0.4
0.2
0.0
10.5
11.5
12.5
13.5
14.5
15.5
16.5
17.5
18.5
Birth
P1
P5
P56
CE/SE
SLC
Sup
(pre-Sup, pre-Gran, Gran)
Stromal
(GM, SP, Stero, SM, PV)
C
Pax8rtTA mouse model [Constitutive]
Pax8
Cre
R26
STOP
tdTom
tdTomato
D
In vivo cell lineage
E12.5
E16.5
P5
8wks
Hypothesis
SLC
Granulosa
E
E12.5
E16.5
8 weeks
Pax8rtTA
FOXL2 tdTomato DAPI
FOXL2
tdTomato
merge
RO
FOXL2+/tdTom-
FOXL2+/tdTom+
AF
F
tdTomato index
in CE cells
% of tdTomato+/total CE cells
100
80
60
40
20
0
*
0
4.5
E12.5
E16.5
G
tdTomato index in granulosa cells
% of tdTomato+/
total FOXL2+ cells
80
60
40
20
0
p=0.057
5.2
22.4
E12.5
E16.5
H
% of tdTomato+/
FOXL2+ cells in act. follicles
80
60
40
20
0
10.3
8 wks
I
P5
Pax8rtTA
FOXL2 tdTomato DAPI
RO
Co
Me
FOXL2+/tdTom-
FOXL2+/tdTom+
FOXL2
tdTomato
merge
J
tdTomato index in medulla vs cortex at P5
% of tdTomato+/
total FOXL2+ cells
80
60
40
20
0
ns
24.9
34.0
Co
Me

Figure S1

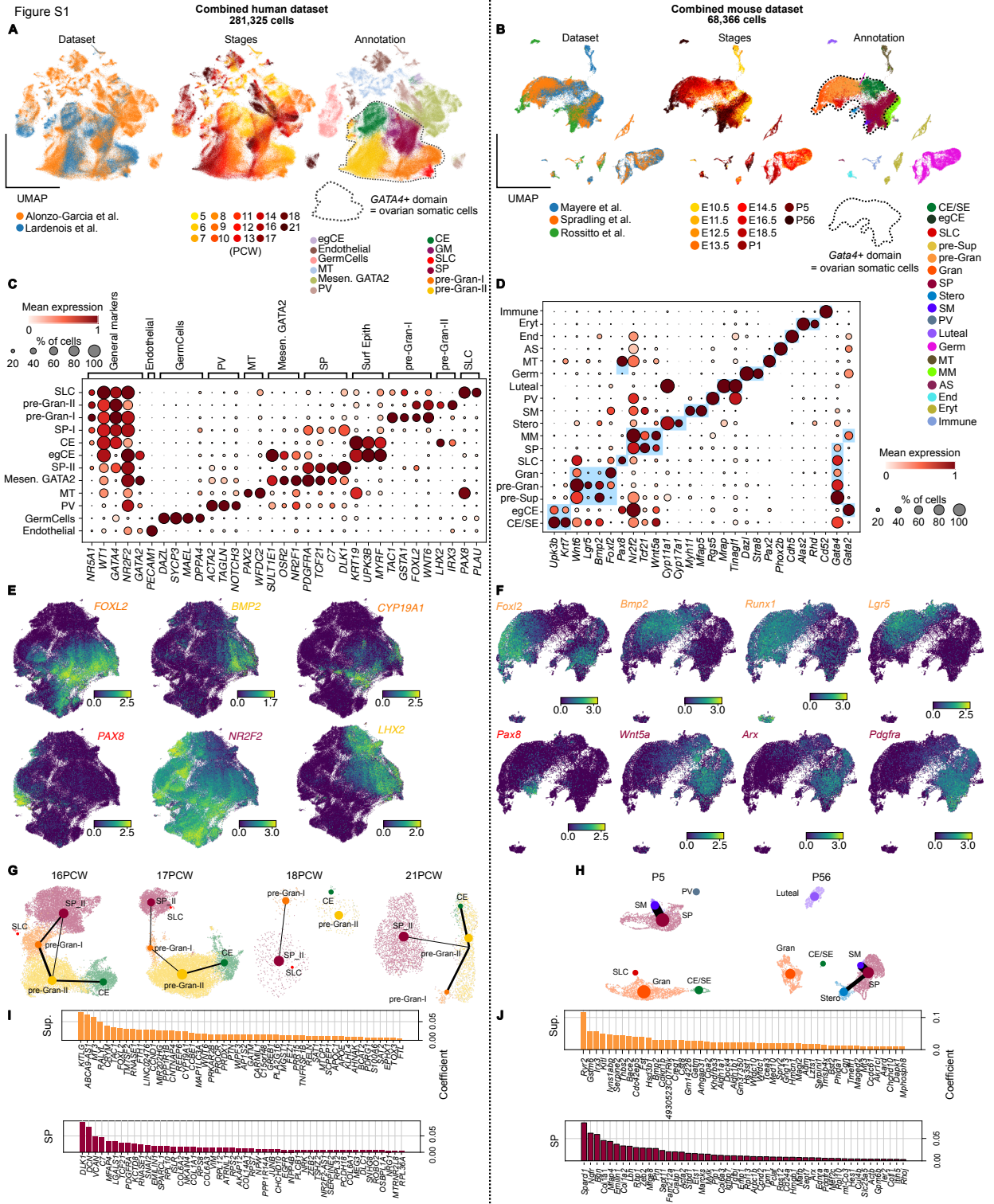

A
Combined human dataset
281,325 cells
Dataset
Stages
Annotation
UMAP
Alonzo-Garcia et al.
Lardenois et al.
5 6 7 8 9 10 11 12 13 14 16 17 18 21
(PCW)
GATA4+ domain
= ovarian somatic cells
egCE
Endothelial
GermCells
MT
Mesen. GATA2
PV
CE
GM
SLC
SP
pre-Gran-I
pre-Gran-II
B
Combined mouse dataset
68,366 cells
Dataset
Stages
Annotation
UMAP
Mayere et al.
Spradling et al.
Rossitto et al.
E10.5
E11.5
E12.5
E13.5
E14.5
E16.5
E18.5
P1
P5
P56
Gata4+ domain
= ovarian somatic cells
CE/SE
egCE
SLC
pre-Sup
pre-Gran
Gran
SP
Stero
SM
PV
Luteal
Germ
MT
MM
AS
End
Eryt
Immune
C
Mean expression
0
1
% of cells
20 40 60 80 100
General markers
Endothelial
GermCells
PV
MT
Mesen. GATA2
SP
Surf Epith
pre-Gran-I
pre-Gran-II
SLC
SLC
pre-Gran-II
pre-Gran-I
SP-I
CE
egCE
SP-II
Mesen. GATA2
MT
PV
GermCells
Endothelial
NR5A1 WT1 GATA4 NR2F2 GATA2 PECAM1 DAZL SYCP3 MAEL DPPA4 ACTA2 TAGLN NOTCH3 PAX2 WFDC2 SULT1E1 OSR2 NR2F1 PDGFRA TCF21 C7 DLK1 KRT19 UPK3B MYRF TAC1 GSTA1 FOXL2 WNT6 LHX2 IRX3 PAX8 PLAU
D
Immune
Eryt
End
AS
MT
Germ
Luteal
PV
SM
Stero
MM
SP
SLC
Gran
pre-Gran
pre-Sup
egCE
CE/SE
Upk3b Krt7 Wnt6 Lgr5 Bmp2 Foxl2 Pax8 Nr2f2 Tcf21 Wnt5a Cyp11a1 Cyp17a1 Myh11 Mfap5 Rgs5 Mfap Tnagl1 Dazl Stra8 Pax2 Phox2b Cdh5 Alas2 Rhd Rhcg Cd52 Gata4 Gata2
Mean expression
0
1
% of cells
20 40 60 80 100
E
FOXL2
BMP2
CYP19A1
PAX8
NR2F2
LHX2
0.0 2.5
0.0 1.7
0.0 2.5
0.0 2.5
0.0 3.0
0.0 2.0
F
Foxl2
Bmp2
Runx1
Lgr5
Pax8
Wnt5a
Arx
Pdgfra
0.0 3.0
0.0 3.0
0.0 2.5
0.0 2.5
0.0 3.0
0.0 3.0
G
16PCW
17PCW
18PCW
21PCW
SP_II
SLC
pre-Gran-I
pre-Gran-II
CE
H
P5
P56
PV
SM
SP
Luteal
Gran
CE/SE
SLC
Stero
I
Sup.
Coefficient
0.00 0.05
SP
0.00 0.05
J
Sup.
Coefficient
0.0 0.1
SP
0.00 0.05

Figure S2

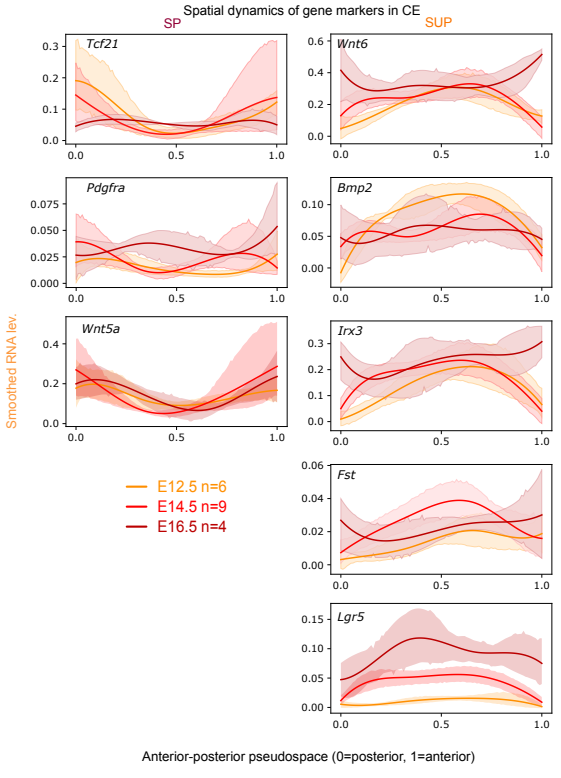
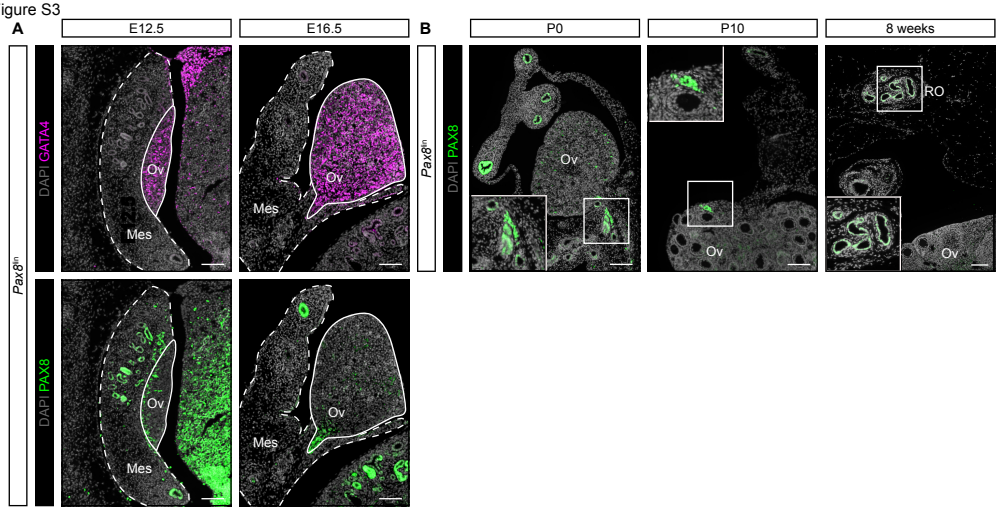

Spatial dynamics of gene markers in CE
SP
SUP
Tcf21
Wnt6
Pdgfra
Bmp2
Wnt5a
Irx3
Fst
Lgr5
Smoothed RNA lev.
E12.5 n=6
E14.5 n=9
E16.5 n=4
Anterior-posterior pseudospace (0=posterior, 1=anterior)

Figure S3


A
E12.5
E16.5
Pax8in
DAPI GATA4
DAPI PAX8
Ov
Mes
B
P0
P10
8 weeks
DAPI PAX8
RO