Unveiling The Origin of Human High-Grade Serous "Ovarian" Cancer Using Bulk and Single-Nucleus RNA Sequencing

Qian Li[1,2,#], Keren Cheng[3,#,*], Lili Sun[2], Wei Yan[4,5,6,*]

[1]Department of Medical Genetics, China Medical University, Shenyang, Liaoning, 110122, China.

[2]Department of Gynecological Oncology, Cancer Hospital of China Medical University/ Liaoning Cancer Hospital & Institute, Shenyang, Liaoning, 110042, China.

[3]Department of Obstetrics and Gynecology, Center for Reproductive Medicine, the Fourth Affiliated Hospital of School of Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University, Yiwu, China,322000

[4]The Lundquist Institute for Biomedical Innovation at Harbor-UCLA, Torrance, CA 90502, USA.

[5]Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, USA.

[6]Current Address: School of Molecular Biosciences, Washington State University, Pullman, WA 99164, USA

[#]Co-first authors

*Correspondence: Keren.Cheng@zju.edu.cn or wei.yan@wsu.edu

Contact Lead: Dr. Wei Yan, wei.yan@wsu.edu

In Brief

This report describes comprehensive cell atlases of the human fallopian tube and high-grade serous “ovarian” cancer (HGSOC), the relationships among major cell subtypes within the human fallopian tube epithelium (FTE), and the most likely pathways by which FTE progenitor cells are transformed into HGSOC cells. The data will serve as an important resource and open new avenues for further investigation into diagnostic, therapeutic, and preventive measures for FTE cell-derived HGSOC.

Highlights

1. $LGR5^+/PGR^+$ cells in human FTE represent stem cells, mostly quiescent and capable of differentiating into $OVGP1^+/RNPC3^+$ progenitors. These progenitors, in turn, differentiate into secretory and ciliated cells to maintain cellular homeostasis in the FTE.
2. Progenitor cells, morphologically resembling secretory cells, are abundant in all segments of the fallopian tube, and those in the fimbriae are the most likely precursors of HGSOC.
3. Stem or progenitor cells can be transformed into cancerous cells following two distinct trajectories, leading to the formation of immunoreactive and mesenchymal cancer cell types common in HGSOC.

ABSTRACT

Secretory cells in the fallopian tube epithelium (FTE) have been suspected to be the origin of high-grade serous “ovarian” cancer (HGSOC). However, due to the high heterogeneity of the FTE cells, the true identity of the FTE cells that are transformed into HGSOC cells and the spatiotemporal molecular events leading to carcinogenesis remain unknown. Our integrative analyses, utilizing both bulk and single-nucleus transcriptomics, revealed that the FTE cells are indeed highly heterogeneous, consisting of multiple subtypes of secretory and ciliated cells in addition to a small stem cell pool. In the adult fallopian tube, the $LGR5^+/PGR^+$ stem cells, often situated in the basal layer of the FTE, can proliferate and differentiate into $OVGP1^+/RNPC3^+$ progenitors, constituting the predominant cell types in all segments of the fallopian tube. These progenitors can quickly differentiate into ciliated or mature secretory cells to replenish the lost FTE cells when needed. Exposure to certain oncogenic factors may cause the progenitors to undergo carcinogenesis along two differentiation pathways, forming mesenchymal-like and immunoreactive cancerous cells, commonly seen in HGSOC. Key transcription factors and genetic changes in critical genes were also identified during carcinogenesis. The present study provides comprehensive cell atlases of the human fallopian tube and HGSOC and helps gain insight into the molecular events leading to the transformation of FTE progenitors into HGSCO cells. (217 words)

KEYWORDS: *Ovarian cancer; fallopian tube; stem cell; progenitor; multiciliogenesis; tumorigenesis; differentiation; cancer microenvironment.*

INTRODUCTION

High-grade serous carcinoma represents the most common type of ovarian cancer, accounting for ~75% of all epithelial ovarian cancer cases. It is also the most lethal histotype of ovarian cancer, with a 10-year survival rate of <30%. While histologically different from typical ovarian epithelial cells, cancerous cells in high-grade serous "ovarian" cancer (HGSOC) resemble epithelial cells of the Müllerian duct-derived internal reproductive tracts, leading to the hypothesis that the fallopian tube may be the source of HGSOC. Indeed, phenotypic, genetic, and proteomic analyses have revealed similarities between HGSOC and fallopian tube epithelial (FTE) cells, especially those from the infundibulum of the fallopian tube.[1,2] While secretory cells are considered the main source of HGSOC, signatures of ciliated cells have also been found in HGSOC, giving rise to the postulation that there exist FTE progenitors, which are capable of differentiating into ciliated cells to replenish the damaged ciliated cells or entering carcinogenesis, leading to the formation of serous intraepithelial carcinoma which eventually spread to the ovary.[3,4]

The fallopian tube can be divided into four distinct segments from the ovary to the uterus: the infundibulum, ampulla, isthmus, and uterotubal junction. The infundibulum's proximity to the ovary subjects this portion to inflammatory factors liberated from oocyte expulsion during the monthly ovulatory events,[5] leading to frequent tissue injury, repair, and remodeling within the epithelial layer.[6] The fallopian tube possesses regenerative capacities to replenish the injured epithelial cells to maintain proper cellular homeostasis.[7,8] The small round cells occasionally seen at the basal layer of FTE (also called intercalary cells) have been postulated as the FTE stem

cells, which can both self-renew and differentiate into ciliated cells or secretory cells (also referred to as "Club" or Peg cells) to fulfill the replenishment function.[6,9] However, the molecular signature of the FTE stem cell remains elusive.

Single-nucleus RNA sequencing (snRNA-seq) allows for in-depth transcriptomic profiling using nuclei extracted from frozen samples, with sequencing deferred until pathological validation.[10] It offers distinct advantages, including minimized dissociation bias, compatibility with frozen specimens, elimination of dissociation-induced transcriptional stress responses,[11] and a heightened efficacy in capturing intricate cell subtypes.[12] Furthermore, snRNA-seq parallels single-cell RNA sequencing (scRNA-seq) in terms of sensitivity and classification of cell subtypes.[13] In the present study, we performed both bulk RNA-seq and snRNA-seq using human fallopian tubes and HGSOC samples collected from patients. We identified the $LGR5^+/PGR^+$ cells as the FTE stem cells that either self-renew or differentiate into the progenitor cells ($OVGP1^+/RNPC3^+$), which can further differentiate into normal FTE cells (i.e., secretory and ciliated cells) under physiological conditions, or get transformed into cancerous cells leading to HGSOC.

RESULTS

Dual origins of HGSOC identified by bulk RNA-seq

It has been hypothesized that HGSOC originates either from the fallopian tube or ovarian epithelium. To test the hypothesis, we first performed bulk RNA-seq analyses to explore transcriptomic correlations among the HGSOC, fallopian tube and ovary (Figure 1A). With the patients' consent, twenty-six specimens, including four cancer-free ovaries,

four cancer-free fallopian tubes with each dissected into three segments (fimbriae/infundibulum, ampulla, and isthmus), and ten HGSOC samples. The cancer-free ovary and fallopian tube samples were collected from five patients undergoing surgery due to hysteromyoma or endometrial polyps, whereas the cancerous tissues were from patients diagnosed with epithelial ovarian cancer (EOC) and subsequently confirmed to have HGSOC based on pathology (Table 1).

In our bulk RNA-seq analyses, each specimen yielded between 38 and 65 million high-quality reads (Figure S1A). We first performed principal component analyses and observed three distinct clusters, representing the three sample types (ovary, fallopian tube and HGSOC) analyzed (Figure 1B). The fallopian tube segments showed a close transcriptomic resemblance, and the biological replicates of the fallopian tube and ovarian specimens were similar (Figures 1B and S1B). Of interest, while eight EOC samples were clustered together, two (EOC-3 and -4) grouped with the ovaries, suggesting their potential ovarian origin. Despite transcriptomic heterogeneity among the eight HGSOC samples clustered together, these HGSOC samples appeared closer to the fallopian tube than the ovarian samples, based on hierarchical correlation and pair-wise Pearson coefficient assays (Figures 1C and S1B). We, therefore, categorized EOC-3 and EOC-4 as ovary-derived HGSOC (EOC-O) and the remaining HGSOC samples as fallopian tube-derived HGSOC (EOC-F) (Figure 1C).

By comparing the differentially expressed genes (DEGs) between EOC-F and EOC-O, we identified that the fallopian tube-specific genes (e.g., *FOXJ1*, *SOX17*, *PODXL*, *MUC16*, *WFDC2*) were upregulated in EOC-F, while the ovarian-specific genes (*FOXL2*, *ARX*, *KIT*, *STAR*, *WFIKKN2*) were upregulated in EOC-O (Figure S1C). Gene

ontology (GO) enrichment analyses of these DEGs revealed pathways such as "cellular division," "cell adhesion," and "spermatogenesis" in EOC-F, whereas "male gonad development," "female gonad development," and the "WNT signaling pathway" were predominant in EOC-O (Figure S1C). The fact that eight out of ten HGSOC samples analyzed in this study are likely derived from the fallopian tubes supports the notion that most of the HGSOC are derived from the fallopian tube.

DEGs identified among the EOC-F, FT, and ovary samples were grouped into seven genesets based on GO term enrichment analyses (Figure 1D and Table S1). Geneset 1, predominantly upregulated in the ovarian specimens, was enriched in male gonad development, ovarian follicle development, and the transforming growth factor-beta receptor signaling pathway. Geneset 2 showcased an enrichment in immune response and positive regulation of gene expression in normal fallopian tube and ovary samples. Notably, key transcription factors, including *HAND2, CITED2, OSR2, GATA2,* and *ALDH1A2,* known to play pivotal roles in ovarian and fallopian tube development, were among the upregulated genes. Geneset 3 captured the presence of smooth muscle components within the tissues. Geneset 4 was marked by fallopian tube-specific transcriptomic signatures, including those related to pronephros development, mesodermal cell fate specification, and embryonic skeletal system morphogenesis, with expression of key transcription factor *PAX2* and various *HOX* genes. Of interest was geneset 5, which was enriched with multicellular organism development and the regulation of transcription from the RNA polymerase II promoter. This set was prominently upregulated in both fallopian tubes and EOC-F, with key transcription factors like SPDEF, POU5F1, FOXJ1, HOXB2 and HOXB4 notably upregulated.

Geneset 6 is highlighted by cilium movement and flagellated sperm motility, showcasing the expression of cilia and flagella-associated proteins, outer dynein arm docking complexes and tektins. Geneset 7, reflecting cancer-specific metabolism, rapid cell cycle and keratinization of epithelium, was enriched with proteolysis, cell division, and keratinization pathways (Figures 1D &1E, Table S1).

By contrasting EOC-F, fallopian tube and ovary samples in a pairwise fashion (Figures S1D, S1E &S1F), we discerned that genes related to cell division, chemokine-mediated signaling pathways (CXCL9, CXCL10, CXCL11/CXCR3 axis, CXCL6/CXCR2 axis), inflammatory responses (IL1/IL1RAP axis, GAL-3 and NR1H4), apoptotic processes (IL24, TNFAIP8), and antigen presentation via MHC class II (CTSS induced by TNF-α/IFN-γ, CTSV9 induced by NF-κB) were augmented in EOC-F. These genes are known to bolster inflammation and tissue fibrosis, reflecting the multifaceted characteristics of EOC-F. In addition, immune-related genes HLA-DMA and HLA-DRA, which induce chemotaxis and promote the differentiation of T cells by Notch signaling, widely participate in immunocyte recruitment, impact tumor progression and modulate the tumor microenvironment,[14,15] were conspicuous. Overall, these findings underscored the pivotal roles of these genes in orchestrating inflammation, immune responses and the tumor microenvironment involved in tumorigenesis and progression of HGSOC. Taken together, our bulk RNA-seq data support the notion that HGSOCs have dual origins, i.e., the ovary and fallopian tube, and most HGSOCs are derived from the fallopian tube.

Identification of a large pool of progenitors capable of differentiating into secretory or ciliated cells in the fallopian tube epithelium

While bulk RNA-seq offers a broad transcriptomic overview, it lacks the nuanced resolution to discern the complexity inherent to individual cell subtypes and their transcriptomic dynamics contributing to the origin of HGSOC. Even though several previous studies have performed single-cell RNA-seq analyses of human fallopian tubes,[7,16–19] the definition of FTE stem cells, the distinction between various secretory cell subtypes, and the identity of mature secretory and ciliated cells remain controversial. To solve these problems, we collected human cancer-free fallopian tubes from three perimenopausal women undergoing hysterectomy with bilateral salpingo-oophorectomy due to benign uterine diseases, sectioned them into infundibulum/fimbriae, ampulla, and isthmus based on morphology, then performed single-nucleus RNA-seq (snRNA-seq) analyses.

Employing the 10X Genomics platform for snRNA-seq, we analyzed 87,920 cells (post-initial quality control) with a range of 36,000-63,000 mean reads per cell and a median of ~2,000 genes per cell from 9 specimens, spanning three fallopian tube segments: fimbriae/infundibulum, ampulla and isthmus (Figure S2A). After performing cell cycle scoring and regressing out the effects of cell cycle, mitochondrial genes and library size for subsequent analyses, our data did not reveal cell clustering predicated on cell cycles, a routine quality control (QC) step in single-cell data analyses (Figure S2B). Utilizing the uniform manifold approximation and projection (UMAP) for embedding, the various fallopian tube segments displayed an even distribution,

precluding the necessity for clustering based on sample preparation (Figure S2C). The QC assays demonstrated the high quality of our snRNA-seq datasets.

We first combined all fallopian tube samples and conducted initial cell clustering and annotations, leading to the identification of five cell types, including secretory cells, ciliated cells, smooth muscle cells, endothelial cells and macrophages (Figure S2D) based on well-documented marker genes (Figure S2E). By analyzing the differentially expressed genes and enriched GO terms (Figure S2F and Table S2), we discovered that secretory cells likely contain the potential stem cells because genes known critical for adult stem cell homeostasis in hair follicle development, metanephric epithelium development, and regulation of transcription from RNA polymerase II promoter, were among the most upregulated one (Figure S2F). Intriguingly, some genes are markers for adult stem cells, for example, *SOX17, LGR5, TFCP2L1* and *HES1,* which are highly expressed in the secretory cell cluster. Endothelial cell cluster exhibited hallmarks of angiogenesis, vasculogenesis and cell adhesion – quintessential characteristics of this cell lineage. The ciliated cell cluster demonstrated enrichment in cilium assembly, epithelial cilium movement, and outer dynein arm assembly. In the context of the fallopian tube, smooth muscle cells, vital to its function, were enriched in actin cytoskeleton organization and smooth muscle contraction. Meanwhile, macrophages displayed enhanced inflammatory responses, differentiation patterns, and positive modulation of cytokine production (Figure S2F). Moreover, a trend was observed where the number of secretory cells was the highest in the isthmus, decreasing towards the ovary, whereas ciliated cells increased in number from the isthmus to the fimbriae (Figure S2G).

We then collected all data points from 34,267 FTE cells to identify cell subtypes and determine the transcriptomic dynamics in the FTE. Utilizing Louvain for re-clustering, we segregated the FTE cells into 12 distinct clusters, as visualized using UMAP (Figure 2A). Cells were subsequently partitioned with ward.D2, based on mean expression across clusters, resulting in five discernible subgroups (Figure 2B). Based on the top marker genes and insights from existing literature, we designated cell cluster 12 as FTE stem cells (*LGR5*$^{+}$*, PGR*$^{+}$*, PODXL*$^{+}$*, SOX17*$^{+}$). Clusters 2 and 6, marked by elevated expression of *PAX8, ZBTB16* and *EGR1*, were annotated as mature secretory cells, while clusters 1, 4, 5, and 8 were identified as ciliated cells (*CAPS*$^{+}$*, CFAPs*$^{+}$). Significantly, we postulated that the remaining clusters (Clusters 3, 7, 9, 10 and 11) represented progenitor cells characterized by the pronounced expression of *OVGP1, CRISP3, RNPC3* and *BDP1* and the absence of *LGR5, PGR, PODXL, PAX8, CAPS* and *FOXJ1* expression (Figure 2C).

Leveraging RNA velocity analyses, we discerned that these progenitor cells were not fully mature, while harboring some secretory traits, and retained the potential to differentiate into ciliated cells (Figure 2D). Using PAGA to analyze developmental trajectory and connectivity among cell clusters, we found that FTE stem cells (Cluster 12) were located at the bottom tip, and progenitor cells (Clusters 3, 7, 9, 10 and 11) straddled between ciliated cells and secretory cells located at distinct trajectory termini, consistent with our RNA velocity findings (Figure 2E). We then used PHATE to better visualize cell developmental trajectories. Our findings revealed that cluster 12 remained at a terminal tip, while the other cells branched off in two directions: one leading to secretory cells and the other to ciliated cells (Figure 2F). Following the establishment of

these trajectories, we employed PHATE embedding to fit the principal curve and then computed the pseudotime based on these trajectories (Figure 2G). Key marker gene expressions were then projected to validate our trajectory analyses (Figure S2H). These trajectory analyses revealed a dynamic continuum encompassing the fallopian tube epithelial stem cell pool and progenitor cells with the potential to differentiate into either mature secretory cells or ciliated cells.

Based on the cell developmental trajectories, we analyzed the transcriptomic dynamics transitioning from stem cells to either secretory or ciliated cells (Figures 2H and 2I). Within the trajectory delineating FTE stem cell differentiation into secretory cells, we observed a declining expression of *LGR5, SPDEF, NR2F1* and *AR* in congruence with pseudotime. Notably, genes responsive to estradiol (*POSTN, MAP1B, PTGS2, and EZH2)* experienced a transient up-regulation, likely reflecting post-ovulation hormonal modulations that propel stem cells or progenitors towards the secretory cell phenotype. Mature secretory cells exhibited an amplified expression of *EGR1, PAX2, PAX8, CD74,* and *ZBTB16* (Figure 2H and Table S3). Conversely, transitioning from FTE stem cells to ciliated cells coincided with a decrease in *LGR5* and *LAMA5* expressions concurrent with pseudotime. As cells initiate differentiation, certain genes responsive to xenobiotic stimulus and cell adhesion were transiently up-regulated in the middle of the differentiation trajectory. Ciliated cells, positioned at the end of the trajectory, were characterized by elevated expressions of *CFAPs* and *Dynein*, consistent with active cilial assembly and movement (Figure 2I and Table S4). By calculating the proportional distribution of the annotated epithelial cell types, we found the FTE stem cells reside predominantly in the isthmus, coexisting with progenitor cells and collectively

representing approximately 2/3 of the total epithelial cell population (Figure S2I & S2J). In contrast, the stem cells were scarce in the fimbriae and ampulla, but progenitors constituted roughly 1/3 (Figure S2I & S2J) of the total epithelial population in these two segments.

Immunohistochemical staining of the ciliated cell marker TUBB4B and the secretory cell marker PAX8 across the isthmus, ampulla and fimbria (Figure S3A) revealed an incline in the density of TUBB4B^{+} ciliated cells and a decline in the density of PAX8^{+} secretory cells from the isthmus to the fimbria (Figure S3B), aligning with our snRNA-seq data (Figure S2G, S2I&S2J). Additionally, our immunofluorescent staining confirmed that the basal cells, with the typical small and round morphology, showed prominently nuclear expressions of LGR5 and PGR, consistent with their identity as the FTE stem cells (Figures 2J, 2K and S3C). Consistent with our snRNA-seq data (Figure S2J), our immunostaining also showed that the LGR5+/PGR+ stem cells were predominantly localized within the isthmus, followed by the ampulla, with a conspicuous absence in the fimbria (Figures 2J, 2K and S3C). Taken together, our snRNA-seq data reveal that progenitor cells (*OVGP1*+) represent the most abundant cell type, accounting for ~38% of the total FTE cells and that the progenitor cells are derived from the FTE stem cells, which are those small round basal cells, with the potential to develop into either ciliated (*CFAPs*$^{+}$*; DNAs*$^{+}$) or mature secretory (*PAX8*+) cells.

High inter-patient variability and heterogeneity in HGSOC cell composition

Given the intra-tumor heterogeneity and infiltration of non-cancer cells, we collected 15,322 cells from HGSOC snRNA-seq datasets (Figure S3D) for further analyses after

removing the batch effect and regressing out the effects of cell cycle and library size (Figure S3E) using LIGER (Linked Inference of Genomic Experimental Relationships).[20] As expected, marked heterogeneity was observed across patients (Figures S3F & S3G).

Within HGSOC samples, we delineated 13 cell clusters through Louvain clustering, as visualized on UMAP (Figure 3A), including ten cancer cell clusters and three non-cancer cell clusters (Figure 3B). The non-cancer cell clusters were annotated as macrophage, cancer-associated fibroblasts (CAFs) and endothelial cells. Cancer cell clusters were classified into proliferative, mesenchymal, ciliated-like, epithelial-to-mesenchymal transition (EMT) and MHC (immunoreactive) subtypes (Figure 3C). Analyses of the most upregulated genes in each cluster (Figure 3D and Table S5) revealed that genes instrumental in cell division and cell cycle are enriched in proliferative cancer cells, whereas the mesenchymal cell type exhibited overexpression of genes pertinent to the negative regulation of cell migration alongside key transcription factors like SOX17, MECOM and PAX8. While cilium-like cancer cells expressed genes involved in cilium assembly and movement, EMT cancer cells predominantly expressed genes associated with cell migration and adhesion, signifying their ongoing transition from an epithelial to a mesenchymal phenotype. MHC (immunoreactive) cancer cells were predominantly characterized by antigen processing and peptide or polysaccharide antigen presentation via MHC class II (Figure 3D and Table S5). We also assessed several HGSOC marker genes recently identified (Figure S3H).[21]

Furthermore, upon re-clustering the malignant cells (with minimal representation from HGSOC_1) and visualizing their subtypes on UMAP (Figure 3E, 3F), a strong patient-derived batch effect was discernible even post-LIGR-based batch effect

mitigation. Intriguingly, the mesenchymal subtype resembled the proliferative subtype (Figure 3G), leading us to postulate that these cell subtypes might function analogously to tumor stem cells. Cilium-like cancer cells, albeit sparse, were universally present across all four HGSOC samples (Figure 3G). All five cancer cell subtypes express high levels of PAX8 and MUC16 (Figure S4A), which are normally expressed in the FTE (Figure 2C), supporting the origin of HGSOC from the FTE cells. Indeed, both PAX8 and MUC16 are abundantly expressed in HGSOC cells (Figure S4B, S4C). MKI67, exhibiting wide and high expression in the fallopian tube, underscores the active proliferative state of FTE cells (Figure 2J). TUBB4B, while prevalent in HGSOC, is also a hallmark of ciliated cells in the FTE (Figure 2C), attesting to the presence of ciliated cell-like cancer cells (Figure S4A & S4B). MHC (immunoreactive) cancer cells predominantly expressed Clusterin (CLU) and specific immune cell genes (Figures 3H, S4C, S4D and Table S6). MUC16 and *ARGLU1* exhibited up-regulation in the EMT subtype, suggesting cellular migratory characteristics (Figures 3H, S4A, S4D and Table S6). MHC (immunoreactive) cancer cells also demonstrated CD74 expression, a trait observed in normal FT, albeit at diminished levels (Figures 3H, S4A, S4D and Table S6). The mesenchymal subtype was characterized by Collagens (COL1A1, COL1A2) and Versican (VCAN) expressions, epitomizing extracellular matrix features and tissue connectivity, likely constituting the tumor's core mass (Figures 3H, S4C, S4D and Table S6). Taken together, our snRNA-seq data reveal that HGSOCs display a high degree of inter-patient variability in tumor cell composition and prominent heterogeneity in the subtypes of malignant cells.

The FTE progenitor cells as the origin of HGSOC

Both previous reports and our data support the notion that most of the HGSOC originates from FTE cells.[7,16] However, it remains unclear which specific FTE cells deviate from their normal developmental program and get transformed into cancerous cells. By merging 11,028 HGSOC cells and 34,267 FTE cells, we attempted to scrutinize the relationship between normal FTE and HGSOC cells (Figure 4). The large enough number of cells analyzed on the same sequencing platform at single-cell resolution allowed us to include potential rare cell types which may be crucial for connecting the normal FTE cells and those undergoing carcinogenic transformation. Using Force-directed embedding in ForceAtlas2,[22] we identified 15 discernible cell clusters (Figure 4A). Using Slingshot, we discovered four developmental lineages, including those from epithelial stem cells (Cluster 11) to secretory cells (Cluster 1), ciliated cells (Cluster 3) and cancerous cells (Clusters 12 and 13), respectively. We also confirmed these trajectories using PAGA (Figure 4B). Supplementing RNA velocity to the Force-directed graph embedding, we discovered that clusters 0/4/8/11 (fallopian epithelial stem cells and progenitors) represented a potential starting state and developed into Cluster 1 (secretory cells), Cluster 3 (ciliated cells) and Cluster 12/13 (cancer cells) (Figure 4C). Furthermore, pseudotime analyses revealed concordant results with those from RNA velocity and PAGA trajectory assays (Figure 4D). In addition to normal FTE cell developmental trajectories (Cluster 11 to Cluster 3 or 1), two cancer cell trajectories (Cluster 11 to Cluster 12/13) were also identified.

Hierarchical clustering analyses categorized the 15 cell clusters into three primary groups: ciliated, secretory, and cancer cells (Figure 4E). In addition to the

epithelial stem cells (Cluster 11), the secretory subgroup contained progenitor cells (Clusters 4, 0 and 8) and mature secretory cells (Cluster 1), collectively classifiable as general secretory cells. Although classified as one of the five ciliated cell subtypes, Cluster 14 appeared to exhibit characteristics of HGSOC cells based on their position in the developmental trajectory (Figures 4B, C, and D), appearance in the cancerous tissue (Figure 4F) and gene expression signature (Figure 4G). Similarly, progenitors in Cluster 4, present in all segments of the fallopian tube, also displayed cancerous traits (Figures 4F and G), suggesting that the FTE progenitors are a potential source of HGSOC. Of significance, identification of the ciliated cell-like cancer cells (Cluster 14) is consistent with previous reports showing that HGSOC tumor cells retain some of the ciliated cell attributes characteristic of their progenitor lineage.[7,16]

Among cancer cells, four major subtypes were discerned, including mesenchymal-like subtype (Clusters 10 with upregulated *VEGFA and NEAT1*), EMT-like subtype (Cluster 13 with upregulated *SPARC* and *COL1A*), proliferative subtype (Cluster 9 with upregulated *TOP2A, ASPM* and *MKI67*) and immunoreactive subtype (Cluster 12, inflammatory-like cells, labeled as MHC with upregulated expression of *CD74* and *HLA-DRA*). It is of note that while the majority of Cluster 9 consisted of cancer cells, it also incorporated cells from the fimbriae, ampulla and isthmus, indicating that the progenitors within each segment of the fallopian tube can all be transformed into cancerous cells (Figure 4F).

Although increasing lines of evidence support the notion that HGSOC is derived from FTE cells, it remains unknown how the FTE cells become cancerous cells at molecular levels. By exploring the expression patterns of the top 5,000 variable genes

along the two cancer cell trajectories mapped to pseudotime, we attempted to gain insight into the molecular events during FTE cell transformation into HGSOC cells (Figure 4H and I). From epithelial stem cells (Cluster 11) to immunoreactive cancer cells (Cluster 12), genes associated with the Notch signaling pathway and cell fate determination were first downregulated at the beginning of pseudotime, followed by upregulation of genes mediating anti-apoptotic processes, cilium assembly, and peptide antigen assembly via the MHC II protein complex during the intermediate or transformative phase of pseudotime. In the terminal stages of pseudotime, genes associated with EMT, cellular response to tumor necrosis, and collagen fibril organization exhibited elevated expression (Figure 4H and Table S7). During the trajectory from epithelial stem cells (Cluster 11) to mesenchymal-like cancer cells (Cluster 13), the early pseudotime phase marked a decline in genes modulating stem cell division. Toward the end of this trajectory, genes promoting angiogenesis, integrin-mediated signaling, and cellular responses to fibroblast growth became dominant, echoing the canonical transcriptomic signatures characteristic of tumor proliferation, invasion and angiogenesis (Figure 4I and Table S8).

Broadly construed, stem cells and progenitors tend to differentiate into ciliated cells in physiological environments (Figure 2). During this normal differentiation, certain cilium-related genes either transiently increase or demonstrate a brief upregulation followed by a decline (Figure 2I and Table S4). In contrast, under carcinogenic conditions, genes mediating cellular responses to tumor necrosis and integrin signaling become upregulated (Figures 4H, 4I). Such aberrant cellular environments also activate the cell cycle and metabolic shifts, favoring aerobic respiration (Figure 4I). This

metabolic reprogramming, synonymous with the Warburg effect, potentially facilitates an environment conducive to tumor growth by optimizing nutrient acquisition, proliferation, metastasis and drug resistance mechanisms.[23,24] Overall, our findings pinpoint FTE progenitors located at fimbriae as the primary source of cancer cells. Under the aegis of inflammatory stimuli or cellular injury, these progenitors undergo transformation and become cancer stem cell-like entities *via* two distinct trajectories, eventually culminating as either immunoreactive or mesenchymal-like HGSOC cells.

Key transcriptional factors driving the differentiation from FTE progenitors into normal epithelial or HGSOC cells

We used pure cancer and FTE cells to conduct single-cell regulatory network inference and clustering (SCENIC) analyses (Figure 5A). Notably, the highest-ranking transcription factors (TFs) in ciliated cells were identified as NR2F1, FOXJ1 and HOXC4. In contrast, TPI1, NR2F6, and MAFB manifested distinct specificity in cancer cells. Secretory cells, progenitors, and stem cells exhibited a commonality of TFs (Figure 5B), reinforcing the dynamic cell fate interplay among these subsets, corroborating our preceding findings. We delineated several TFs potentially influencing cell fate during normal and anomalous fallopian epithelium maturation, subsequently categorizing regulon specificity based on AUC scores (Figure 5C). NR2F6, a newly identified superfamily member of TFs in multiple types of cancer,[25,26] can potentially modulate diverse biological events in response to hormones or nutrients. NR2F6 might represent a novel therapeutic target for HGSOC intervention. Distinctively, SPDEF manifested specificity in progenitors and stem cells, further emphasizing its role as a pivotal

regulator across diverse cancers (Figure 5C).[27,28] Historically, SPDEF was recognized as an integral component of the Notch signaling cascade.[29] Given the shared motif sequence among the SOX family, our identification of SOX5, rather than SOX17, was unsurprising. We further corroborated these findings by analyzing overlap patterns through the AUC score (bottom row) and gene expression levels (top row) of regulons (Figure 5D). By this criterion, NR2F6 emerged as a central regulator in HGSOC, whereases SPDEF featured prominently in progenitors and stem cells. Through the prediction of gene regulatory networks within each cell subtype, we identified SOX5 and SPDEF as master regulators in stem cells and progenitors, PAX2, TBX2 and TCF7L2 in mature secretory cells, HOXC4 and ZNF214 in ciliated cells, NR2F6 and TPI1 in HGSOC cells (Figure 5E).

DNA Copy Number Variations in HGSOC

To elucidate potential genomic changes in HGSOC, we analyzed copy number variations (CNVs) using the inferCNVs tool,[30,31] leveraging normal FTEs as a reference benchmark (Figure 6A). We documented the presence of both gain and loss of expression spanning extensive genomic scales. Utilizing the Hidden Markov Model (HMM), we identified 138 CNV regions. These regions were further categorized via a six-state CNV model delineating the following CNV levels: State1 denoting a complete loss (0), State2 reflecting a loss of a single copy (0.5x), State3 being neutral (1x), State4 indicating the addition of one copy (1.5x), State5 highlighting the addition of two copies (2x), and State6, which essentially serves as a placeholder for > 2x copies but is modeled as 3x (3x) (Figure 6B and Table S9).

Aggregating the distributions across these CNV regions revealed that neutral variants constituted ~39%, gains of copies represented ~34%, and loss of copies comprised ~27%. Predominantly, these events were localized to Chromosome 1, with the most pronounced alterations observed on Chromosome 6 (Figure 6C). Our analyses further authenticated that the aberrant up-regulation of *S100A6, MUC16, COL1A1, COL1A2, CD74* and *WFDC2* in HGSOC could be attributed to the amplification of copy numbers. Conversely, the diminished expression of *CLU, PODXL* and *PGR* likely results from the deletion of copy numbers (Figure 6D). These insights could provide a rationale for the preferential differentiation of progenitors into HGSOC cells rather than terminally differentiated cells.

Pivotal Role of Collagen Signaling in the Tumor Microenvironment of HGSOC.

Within the tumor microenvironment (TME) of HGSOC, cancer-associated fibroblasts (CAFs) emerged as dominant regulators of other cell subtypes (Figures S5A, S5B). A comprehensive assessment illuminated the collagen signaling pathway as the paramount extracellular matrix (ECM) regulator, influencing tissue architecture, cellular adhesion, EMT, and tumorigenesis (Figures S5C, S5D).[32,33] Further, our analyses underscored the intricate interactions between tumor and non-tumor cell subtypes of HGSOC, mediated by key collagen signaling pathway ligands and receptors. Notably, ITGAV and ITGB8, encoding integrin receptors for fibronectin, were co-expressed in both HGSOC tumor cells and CAFs. This mutual expression pattern suggests a synergistic TME response orchestrated by these cell types. Moreover, genes encoding specific collagen types (*COL1A1, COL1A2, COL4A1, COL4A2, COL6A1, COL6A2* and

COL6A3) displayed upregulated expression in mesenchymal-like and cilium-like cancer cells and CAFs, highlighting their pivotal roles in modulating ECM remodeling within the TME (Figure S5E).

DISCUSSION

It is becoming increasingly clear that HGSOC has a dual origin: ovarian surface epithelium and FTE, with the latter being the majority, and our bulk RNA-seq data support this notion. Given numerous DEGs identified between the two types of HGSOC and their known physiological functions in the normal fallopian tube and ovary, it is feasible to develop diagnostic biomarkers to distinguish the two origins. Based on marker genes identified in our snRNA-seq analyses, it is also possible to further classify the fallopian tube-derived HGSOCs into various cancer subtypes, which may facilitate developing early detection, precision treatment, and tumor prevention strategies.

While increasingly accepted, the notion that HGSOC originates from FTE secretory cells remains confusing because it implies that some of the FTE secretory cells would undergo carcinogenesis and become cancerous cells. The secretory cells are terminally differentiated cells, which should not develop into tumor cells easily. Histologically, the secretory cells are heterogeneous, and many so-called secretory cell markers, e.g., PAX8 and OVGP1, have been found to be expressed in a subset of secretory cells but absent in other subsets of secretory cells. Recent single-cell transciptomes of fallopian tube have clearly demonstrated that the FTE secretory cells are indeed heterogeneous, and several subtypes of the secretory cells have been defined.[18] However, it remains unclear which subtype of secretory cells tends to

undergo carcinogenesis, why this specific subtype is prone to carcinogenesis, and what roles these cells play in normal FTE. To address these questions, we employed both bulk RNA-seq and snRNA-seq to dissect the intricate cellular homeostasis in normal FTE and to trace the specific FTE cell subtypes that undergo transformation and become cancerous cells. The large number of single cells used in this study for transcriptomic analyses allowed us to obtain a comprehensive human fallopian tube cell atlas with higher resolution, compared to all previous studies.[19,34,35] The latest cell developmental trajectory analyses adopted in this study made it possible to dissect the intricate relationship among FTE stem cells and various subtypes of secretory and ciliated cells.

We've discovered that the $LGR5^+/PGR^+$ basal small round cells are the FTE stem cells, characterized by their relative quiescence and attenuated MKI67 expression. Similar to mature secretory cells, the number of FTE stem cells gradually decreases from the isthmus to the ampulla and infundibulum/fimbriae. The FTE stem cell population is capable of differentiation into secretory and ciliated cells *via* intermediate progenitors, substantiating a dynamic interchange amongst stem cells, progenitors and secretory cells, thereby preserving fallopian epithelial homeostasis.

The stem cells ($LGR5^+/PGR^+$) are scarce in the FTE but can produce a large pool of progenitor cells ($OVGP1^+$). Although histologically resembling secretory cells due to a lack of cilia, these FTE progenitors are in a poised state and can differentiate into either ciliated ($CFAPs^+$) or mature secretory ($PAX8^+$) cells when needed. It is noteworthy that the progenitors represent the most abundant cell type in the FTE, accounting for ~38% of the total FTE cells. It is well-established that stem cell-derived

progenitors with the potential to differentiate into one or multiple functional cell types tend to undergo carcinogenesis under certain conditions.[36,37] It is noteworthy that despite the very low number of stem cells, the progenitors are abundant in fimbriae, probably reflecting more active replenishing due to the high turnover rate of FTE cells due to damages caused by monthly ovulatory events.

Drawing from comprehensive gene expression and active chromatin datasets, Lawrenson K. et al. show evidence that a preponderance of HGSOCs, particularly those of the mesenchymal type, are descendants of fallopian tube secretory epithelial cells (FTSECs).[38] Experiments by Nakamura K. et al., encompassing the inactivation of dominant negative p53 and the upregulation of KRAS, c-myc, and PI3K-Akt, emphatically validate that selective gene modulations can shepherd immortalized FTSECs, suggesting that two patterns of three gene alterations (p53/KRASV12/c-Myc and p53/KRASV12/PI3K-AKT) can develop immortalized FTSECs into early stage HGSOC *in vitro.*[39] Here, we report that fallopian tube-derived HGSOC originates from $OVGP1^{+}$/$RNPC3^{+}$ (Cluster 4) progenitors in the epithelium of infundibulum/fimbriae, which possess dual differentiation prowess towards both tumor and ciliated cells.

Historically, HGSOCs have been parsed into four main tumor subtypes: differentiated, proliferative, mesenchymal and immunoreactive.[40] We, however, based on the discrete functional attributes, classify HGSOC into five cellular subgroups, including proliferative, mesenchymal-like, ciliated-like, EMT and MHC subgroups. Data from the Australian Ovarian Cancer Study (AOCS) illustrate the ciliated subtype's preponderance within low-grade tumors (grade 1),[41] while the EMT-High (mesenchymal)

subtype bears significantly poor prognostic implications, as gleaned from eight independent datasets.[42–44]

The two distinct carcinogenic trajectories identified in the present study revealed the intricate interplay of immune infiltrates and fibroblasts within the TME. The prevalence of immuno-reactive cells potentially correlates with heightened tumor-infiltrating lymphocyte (TIL) abundance, portending favorable prognosis and augmented immunotherapeutic responsiveness.[45] Notably, the "immunoreactive" and "mesenchymal" reflected the expression programs of macrophage and CAFs, respectively, rendering them more as indicators of tumor composition than distinct tumor cell phenotypes.[40]

Cancer-associated fibroblasts (CAFs) and immune cells constitute integral facets of the TME. Upon activation, CAFs can enhance tumor proliferation, angiogenesis, invasion, metastasis, extracellular matrix (ECM) reconfiguration, and even confer chemoresistance through various molecular pathways. Many previous studies have underscored the pivotal interplay between CAFs and neoplastic cells in tumorigenesis and progression.[46,47] Our investigations confirmed that the trajectory of tumorigenesis is characterized by salient transcriptomic hallmarks, encapsulating tumor proliferation, invasion and angiogenesis. In response to hypoxic stimuli, CAFs secrete an array of angiogenic mediators, notably VEGF, leading to the downregulation of endothelial cell adhesion molecules, such as ICAM-1/2 and VCAM-1.[48] Tumor-associated macrophages (TAMs), bifurcated into the M1 and M2 phenotypes, display divergent functionalities. M1 macrophages execute their antineoplastic activity in the tumor immune microenvironment (TIME) predominantly through mediating antibody-dependent

576 cytotoxicity coupled with the generation of ROS and tumor necrosis factor (TNF).[49]
577 Conversely, M2 macrophages foster tumorigenesis by augmenting angiogenesis,
578 immunosuppression, tumor cell invasion, metastasis, and ECM dynamics.[50,51] The
579 aberrant neoplastic collagen matrix beckons TAMs and reciprocally enables TAMs to
580 modulate collagen deposition, culminating in escalated matrix rigidity and hastened
581 ECM metamorphosis.[52] Our data spotlighted the collagen signaling pathway as the
582 paramount cellular communication conduit, potentially elucidating the modus operandi
583 by which CAFs orchestrate the immunosuppression of TME, thus facilitating the
584 translation of CAF-centric therapeutic interventions into clinical trials.

585 Compared to other populations, individuals harboring BRCA mutations exhibit a
586 markedly elevated oncogenic predisposition. Approximately 20% of HGSOCs are
587 tethered to germline or somatic BRCA mutations.[53] A prospective study articulated that
588 by age 80, BRCA1 and BRCA2 mutation carriers bear 44% and 17% cumulative ovarian
589 cancer risks, respectively.[54] In fallopian tube specimens from BRCA mutation-positive
590 individuals undergoing Risk-Reducing Salpingo-Oophorectomy (RRSO), isolated serous
591 tubal intraepithelial carcinoma (STIC) lesions were discerned in 2.3% of cases, whereas
592 p53 signatures were present in a quarter of the cohort.[55] Clinically, STIC lesions
593 coexisted with HGSC in 11% to 61% of cases, with a prevalence surpassing 60%, even
594 in serous ovarian cancers devoid of BRCA linkage.[56] This ubiquity has ensconced STIC
595 as a precursor lesion to HGSC. While previous studies predominantly leveraged
596 fallopian tube samples from HGSOC-afflicted individuals, our approach pivoted towards
597 samples from oncologically unscathed women. Additionally, our findings elucidated that
598 the genesis of HGSOC from the fallopian tubes was a common phenomenon. Notably,

this occurrence exhibited no correlation with variables such as family history of tumors (designated as a high-risk population), menstrual history, marital and reproductive records, or a history of hormone therapy (a recognized high-risk factor).

In conclusion, our data provide evidence that fallopian tube-derived HGSOCs originate from $OVGP^+/RNPC3^+$ (Cluster 4) progenitor cells. Intriguingly, this is the inaugural identification of $LGR5^+/PGR^+$ epithelial stem cells, poised in a relative quiescence within the fallopian tube epithelium. HGSOCs encompass a ciliated cell-like subtype, which reciprocally interacts with MHC-like and EMT subtypes. Normal development programs ensuring proper differentiation from FTE stem cells to progenitors and then to mature secretory cells or ciliated cells, can be derailed by carcinogenic factors, e.g., chronic inflammation and cell damage due to exposure to harmful chemicals released during ovulation, etc., causing malignant transformation into MHC-like and mesenchymal-like cancer cell subtypes, which, in turn, interact with the TME, catalyzing tumor initiation and progression.

STAR ★ Methods

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit Anti-OVGP1 antibody	Bioss	Cat#bs-17571R
Rabbit Anti-Ki67 antibody	Bioss	Cat#bsm-60738R
Rabbit Anti-LGR5 antibody	Bioss	Cat#bs-20747R
Rabbit Anti-PGR antibody	Bioss	Cat#bs-23376R
Acetylated-Tubulin Rabbit pAb	Abcam	Cat#ab179484
Anti-PAX8 Rabbit pAb	Abcam	Cat#ab53490
Rabbit Anti-CA125 antibody	Bioss	Cat#bs-0091R
Biological Samples		
Surgical specimen from patients with ovarian cancer	The Cancer Hospital of China Medical University	N/A
Surgical specimen from benign patients	The Cancer Hospital of China Medical University	N/A
Critical Commercial Assays		
Trizol reagent kit	Invitrogen, Carlsbad, CA,USA	15596026CN
iQ™ SYBR® Green Supermix kit	Bio-Rad	1708886
UltraSensitive™ SP (Mouse/Rabbit) IHC Kit	MXB, Fujian, CA	KIT-9720
Deposited Data		
Bulk RNA seq	This paper	GSE223426
snRNA seq	This paper	GSE270250
Software and Algorithms		
Cell Ranger	10X Genomics; https://support.10xgenomics.com/	v7.1, RRID:SCR_017344

	single-cell-gene-expression/software/	
	overview/welcome	
R package Seurat	https://cran.r-project.org/web/packages/Seurat/index.html	v.4.3.0,RRID:SCR_016341
R package MAST	https://bioconductor.org/packages/release/bioc/html/MAST.html	v1.30.0
R package SCENIC	https://github.com/aertslab/SCENIC	V1.1.2, RRID:SCR_017247
R package edgeR	https://bioconductor.org/packages/release/bioc/html/edgeR.html	v3.34.1, RRID:SCR_012802
RcisTarget	https://www.bioconductor.org/packages/release/bioc/html/RcisTarget.html	v1.24.0, RRID:SCR_024860
Scanpy	https://scanpy.readthedocs.io/en/stable/	v1.8.1, RRID:SCR_018139
PAGA	https://github.com/theislab/paga	
velocyto	http://velocyto.org/velocyto.py/tutorial/index.html	V0.17, RRID:SCR_018167
scVelo	https://scvelo.readthedocs.io/en/stable/	v.0.2.2, RRID:SCR_018168
PHATE	https://phate.readthedocs.io/en/latest/	V0.3.0
Cytoscape	https://cytoscape.org/	v3.9.1, RRID:SCR_003032
infercnv	https://github.com/broadinstitute/infercnv	V1.3.3, RRID:SCR_021140
DAVID	https://david.ncifcrf.gov/	v2023q4, RRID:SCR_001881
In-house analysis	https://github.com/chengkeren/fallopiantube	NA

616

Data and code availability

The accession numbers generated in this study are GSE223426 for bulk RNA-seq and GSE270250 for snRNA-seq analysis. The codes are available at the GitHub repository: https://github.com/chengkeren/fallopiantube). Ethics approval was obtained in all participating countries and, where required, in all participating centers. All patients provided written informed consent.

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Wei Yan (weiyanucla@ucla.edu). This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human Fallopian Tube and Ovarian Cancer Samples

This study was approved by the Institutional Review Board (IRB) at Liaoning Cancer Hospital and the Cancer Hospital of China Medical University (Shenyang, P.R. China) (Protocol#: 20220245). All participating patients provided informed consent. For bulk RNA-seq analyses, HGSOC samples were collected from 10 patients (#1-10) who were diagnosed with EOC and underwent surgery. Benign ovaries and fallopian tubes came from patients (#11-15) who underwent total hysterectomy combined with bilateral salpingo-oophorectomy for benign uterine diseases, e.g., endometrial polyps, hysteromyomas, etc. (Table 1). For snRNA-seq analyses, the 13 samples were

collected from 9 patients, including 5 (#1-5) with hysteromyoma and adenomyosis, and 4 (#6-9) with EOC/HGSOC (Table 2).

METHOD DETAILS

Tissue processing

After removal from the patients, the cancer tissue was immediately rinsed in cold PBS and dissected into several portions: one for fixation in 10% formalin overnight, followed by preparation of paraffin blocks for immunohistochemical analyses, and the rest were snap-frozen in LN2, followed by storage at -80°C until RNA isolation for bulk RNA-seq or preparation of single nuclei for snRNA-seq. After removal, the human fallopian tubes were immediately rinsed in cold PBS, followed by dissection under a dissection scope into three portions: isthmus (proximal), ampulla (enlarged part), and fimbria (distal). The dissected fallopian tube segments were snap-frozen in LN_2, followed by storage at -80°C until RNA isolation for bulk RNA-seq or preparation of single nuclei for snRNA-seq.

Bulk RNA-seq library preparation and analysis

Total RNA was extracted using Trizol (Invitrogen, Carlsbad, CA, USA). RNA quality and quantity were assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). mRNA was enriched by oligo(dT) beads, and rRNA was removed by Ribo-ZeroTM Magnetic Kit (Epicentre, Madison, WI, USA). The enriched mRNA was fragmented into short fragments, followed by reverse transcription into cDNA using random primers. Second-strand cDNA was synthesized by DNA polymerase I, and the

cDNA fragments were purified using a QiaQuick PCR extraction kit (Qiagen, Venlo, The Netherlands), followed by end repair, poly(A) addition, and ligation to Illumina sequencing adapters. The ligation products were then size selected, followed by PCR amplification. The libraries were sequenced on Illumina NovaSeq6000 (PE150) at Genedenovo Biotechnology Co., Ltd (Guangzhou, China).

Raw fastq files were demultiplexed with bcl2fastq2 (v.2.20.0.422). Barcodes and adapters were removed with Trimmomatic (v.0.32).[57] Fastq files were mapped to the UCSC human reference genome (GRH38) using STAR (2.7.10a).[58] The raw gene count table was generated with featurecounts,[59] and weakly expressed genes were filtered with edgeR with the filterByExpr function with default parameters. Briefly, the raw counts were normalized to library size, and then genes with counts per million above 10 were included in downstream analysis. DEGs were analyzed with edgeR (v.3.36.0) with log2fold change > 1, p-value < 0.05, and FDR < 0.05. Reads per kilobase per million (RPKM) values were calculated in edgeR,[60] and the gene length was obtained from the UCSC table browser.[61] Downstream data analyses and visualization were performed with R (v.4.1.0). Hierarchical clustering was performed with hclust in R (v.4.1.0).

snRNA-seq analysis

Preparation of single nuclei from frozen human HGSOC and fallopian tube samples and constructions of 3' single-cell gene expression libraries were conducted using the 10x Genomics Chromium system (Next GEM v3.1) following the manufacturer's protocol. The libraries were sequenced with ~200 million PE150 reads per sample on Illumina NovaSeq at LC-BIO Co., Ltd (Hangzhou, China).

Raw data were demultiplexed with the mkfastq command in Cell Ranger(v7.1.0) to generate Fastq files. Trimmed sequence files were mapped to the reference genome for humans (GRCh38) provided by 10X Genomics using Cell Ranger mkfastq. Read counts were obtained by using cellranger count[62]. Secondary data analyses were performed in Python with Scanpy(v1.8.1)[63] or in R(v.4.1.0) with the Seurat(v.4.3.0),[64] SeuratObject(v4.1.3), SeuratWrappers(v0.3.2),[64] tidyverse(v.1.3.2),[65] destiny(v3.10.0),[66] pheatmap (v1.0.12), plot3D (v1.4) and Matrix (v.1.5-1) packages and Excel (Microsoft). UMI count tables were first loaded into R by using the Read10x function in Seurat, and Seurat objects were built from each sample. Cells with less than 200 genes, an aberrant high gene count above 7000, or total mitochondrial genes above 15% were filtered out. Samples were combined using Seurat merge function. The effects of mitochondrial genes, cell cycle genes, were regressed out by SCTransform during normalization in Seurat. Batch effect was removed using Harmony. Mitochondrial genes and cell cycles genes were excluded during Cell Clustering, dimensional reductions and trajectory analysis. Cells were clustered by a shared nearest neighbor (SNN) modularity optimization-based clustering algorithm in Seurat. Clusters were annotated on the basis of previously characterized marker gene expression by using the FeaturePlot function and the gene expression matrix file. Then, annotated cell clusters were used for downstream analyses. Dimensional reductions were performed with top 3000 highly variable genes and first 20 principal components by using Seurat. Differentially expressed genes (DEGs) in different clusters were calculated by MAST or area under curve (AUC) integrated in Seurat findallmarkers funtion, with a threshold of avgLog2FC > 0.25, p-value < 0.01. For LogNormalize Expression (LogExp) in Seurat, genes for each

cell were divided by the total counts for that cell and multiplied by the cal factor (10,000), then natural-log transformed using log1p.

Developmental trajectories of cells were simulated with first 30 PCs and 30 diffusion components by DiffusionMap in destiny and scanpy.[63] Trajectory principal lines were fitted by DiffusionMap.[66] Pseudotime was calculated by dpt in destiny. Single-cell pseudotime trajectories were constructed using PAGA,[67] and PHATE.[68]

For Pseudobulk differential expression analysis of scRNA-seq data, DEGs between two groups in scatterplot were identified by edgeR 3.34.1 with applying quasi-likelihood approach (QLF) and the fraction of detected genes per cell as covariate[60]. The DEGs were defined as the genes with FDR < 0.05, p-value < 0.05, and Log2-fold change above 1. Expression value was normalized by edgeR using log2(CPM+1) (CPM = Counts Per Million). Cell Cycle was analyzed using CellCycleScoring in Seurat. Data was visualized by ggplot2 and pheatmap. Genes in heatmap were hierarchically clustered by measuring euclidean distance, scaled by row, then visualized using pheatmap. Gene ontology enrichment was analyzed by DAVID v6.8.[69]

RNA velocity

The velocyto command line interface was used to generate spliced/unspliced expression matrices from the 10x Genomics Chromium samples.[70] RNA velocity analysis was then performed by using scVelo v.0.2.2.[71]

Regulatory network inference using SCENIC

We used single-Cell regulatory network inference and clustering method (SCENIC) to analyze transcription factors (TFs) regulatory networks, from co-expression of transcription factor and candidate target genes from RcisTarget database.[72] After defining the key TFs, the network was drew by using Cytoscape (v3.9.1).[73]

Single-cell copy number analysis

Initial CNVs (CNV) were estimated by sorting the analyzed genes by their chromosomal location and applying a moving average to the relative expression values, with a sliding window of 100 genes within each chromosome, as described.[74] To avoid considerable impact of any particular gene on the moving average, we limited the relative expression values to [-3,3] by replacing all values above 3 by a ceiling of 3, and replacing values below -3 by a floor of -3. This was performed only in the context of CNV estimation. This initial analysis is based on the average expression of genes in each cell relative to the other cells and therefore does not have a proper reference to define the baseline. We thus defined the gene expression clusters annotated as ciliated epithelial cells, progenitor, secretory cells, FTE stem cells by gene expression as the nonmalignant cells, and used the average CNV estimate at each gene across those cells as the baseline. The final CNV estimate of cell i at position j was defined as: the 6 states correspond to the following CNV events: State 1; 0x, complete loss. State 2; 0.5x, loss of one copy. State 3; 1x, neutral. State 4; 1.5x, addition of one copy. State 5; 2x, addition of two copies. State 6; 3x, essentially a placeholder for >2x copies but modeled as 3x.

Immunohistochemistry and Immunofluorescence staining

Immunohistochemical staining was conducted as described.[75] Tissue samples were fixed in 10% buffered formalin (Fisher Healthcare) with gentle rocking overnight at room temperature. After dehydration, tissues were embedded in paraffin, serially sectioned at 5 μm thickness with a microtome (Leica) and placed on Superfrost Microscope glass slides. Paraffin sections were then de-paraffinized with xylene. Antigens were retrieved with 10mM Sodium Citrate for 35 min at 90° C and then for 15 min at room temperature. The staining and incubation procedure for paraffin sections was with the following modifications: the blocking solution was 5% normal donkey serum in PBST; the primary antibody incubation was performed overnight at 4° C; and slides were washed with PBS six times after each incubation. Slides were mounted in Vectashield mounting medium for imaging. The dilution rates of primary antibodies used in this study are as follows: anti-LGR5 (1:3000), anti-PGR (1:400), anti-Ki67 (1:2000), anti-OVGP1 (1:200), anti-CRISP3 (1:200).

DECLARATION OF INTERESTS

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

Visualization, methodology and statistics, K.C., W.Y.; writing-original draft, Q.L., K.C.; writing-revise, review and editing, W.Y., K.C.; conceptualization, design and supervision, W.Y.; acquisition of specimens, Q.L., C.Y.W.; acquisition of clinical data, Q.L.; histological and pathological guidance, L.L.S.

ACKNOWLEDGMENTS

We thank LC-BIO Co. Ltd. (Hangzhou, China) for providing the Chromium 10X genomics single-nucleus transcriptome sequencing services and Genedenovo Biotechnology Co. Ltd. (Guangzhou, China) for providing the Illumina NovaSeq 6000 sequencing services for bulk RNA-seq.

REFERENCES

1. Gong, T.-T., Guo, S., Liu, F.-H., Huo, Y.-L., Zhang, M., Yan, S., Zhou, H.-X., Pan, X., Wang, X.-Y., Xu, H.-L., et al. (2023). Proteomic characterization of epithelial ovarian cancer delineates molecular signatures and therapeutic targets in distinct histological subtypes. Nat Commun *14*, 7802. https://doi.org/10.1038/s41467-023-43282-3.

2. Coscia, F., Watters, K.M., Curtis, M., Eckert, M.A., Chiang, C.Y., Tyanova, S., Montag, A., Lastra, R.R., Lengyel, E., and Mann, M. (2016). Integrative proteomic profiling of ovarian cancer cell lines reveals precursor cell associated proteins and functional status. Nat Commun *7*, 12645. https://doi.org/10.1038/ncomms12645.

3. Bhattacharya, R., Ghosh, A., and Mukhopadhyay, S. (2023). High-grade serous ovarian carcinoma, the “Achiles’ hill” for clinicians and molecular biologists: a molecular insight. Mol Biol Rep *50*, 9511–9519. https://doi.org/10.1007/s11033-023-08760-3.

4. Colvin, E.K., and Howell, V.M. (2020). Why the dual origins of high grade serous ovarian cancer matter. Nat Commun *11*, 1200. https://doi.org/10.1038/s41467-020-15089-z.

5. King, S.M., Hilliard, T.S., Wu, L.Y., Jaffe, R.C., Fazleabas, A.T., and Burdette, J.E. (2011). The impact of ovulation on fallopian tube epithelial cells: evaluating three hypotheses connecting ovulation and serous ovarian cancer. Endocrine-Related Cancer *18*, 627–642. https://doi.org/10.1530/ERC-11-0107.

6. Paik, D.Y., Janzen, D.M., Schafenacker, A.M., Velasco, V.S., Shung, M.S., Cheng, D., Huang, J., Witte, O.N., and Memarzadeh, S. (2012). Stem□Like Epithelial Cells Are Concentrated in the Distal End of the Fallopian Tube: A Site for Injury and Serous Cancer Initiation. STEM CELLS *30*, 2487–2497. https://doi.org/10.1002/stem.1207.

7. Dinh, H.Q., Lin, X., Abbasi, F., Nameki, R., Haro, M., Olingy, C.E., Chang, H., Hernandez, L., Gayther, S.A., Wright, K.N., et al. (2021). Single-cell transcriptomics identifies gene expression networks driving differentiation and tumorigenesis in the human fallopian tube. Cell Reports *35*, 108978. https://doi.org/10.1016/j.celrep.2021.108978.

8. Yuan, S., Wang, Z., Peng, H., Ward, S.M., Hennig, G.W., Zheng, H., and Yan, W. (2021). Oviductal motile cilia are essential for oocyte pickup but dispensable for sperm and embryo transport. Proc. Natl. Acad. Sci. U.S.A. *118*, e2102940118. https://doi.org/10.1073/pnas.2102940118.

9. Zhu, M., Iwano, T., and Takeda, S. (2020). Fallopian Tube Basal Stem Cells Reproducing the Epithelial Sheets In Vitro—Stem Cell of Fallopian Epithelium. Biomolecules *10*, 1270. https://doi.org/10.3390/biom10091270.

10. Slyper, M., Porter, C.B.M., Ashenberg, O., Waldman, J., Drokhlyansky, E., Wakiro, I., Smillie, C., Smith-Rosario, G., Wu, J., Dionne, D., et al. (2020). A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. Nat Med *26*, 792–802. https://doi.org/10.1038/s41591-020-0844-1.

11. Wu, H., Kirita, Y., Donnelly, E.L., and Humphreys, B.D. (2019). Advantages of Single-Nucleus over Single-Cell RNA Sequencing of Adult Kidney: Rare Cell Types and Novel Cell States Revealed in Fibrosis. JASN *30*, 23–32. https://doi.org/10.1681/ASN.2018090912.

12. Wiegleb, G., Reinhardt, S., Dahl, A., and Posnien, N. (2022). Tissue dissociation for single-cell and single-nuclei RNA sequencing for low amounts of input material. Front Zool *19*, 27. https://doi.org/10.1186/s12983-022-00472-x.

13. Ding, J., Adiconis, X., Simmons, S.K., Kowalczyk, M.S., Hession, C.C., Marjanovic, N.D., Hughes, T.K., Wadsworth, M.H., Burks, T., Nguyen, L.T., et al. (2020). Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. Nat Biotechnol *38*, 737–746. https://doi.org/10.1038/s41587-020-0465-8.

14. Tokunaga, R., Zhang, W., Naseem, M., Puccini, A., Berger, M.D., Soni, S., McSkane, M., Baba, H., and Lenz, H.-J. (2018). CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation – A target for novel cancer therapy. Cancer Treatment Reviews *63*, 40–47. https://doi.org/10.1016/j.ctrv.2017.11.007.

15. Wu, L., Awaji, M., Saxena, S., Varney, M.L., Sharma, B., and Singh, R.K. (2020). IL-17-CXC Chemokine Receptor 2 Axis Facilitates Breast Cancer Progression by Up-Regulating Neutrophil Recruitment. Am J Pathol *190*, 222–233. https://doi.org/10.1016/j.ajpath.2019.09.016.

16. Hu, Z., Artibani, M., Alsaadi, A., Wietek, N., Morotti, M., Shi, T., Zhong, Z., Santana Gonzalez, L., El-Sahhar, S., KaramiNejadRanjbar, M., et al. (2020). The Repertoire of Serous Ovarian Cancer Non-genetic Heterogeneity Revealed by Single-Cell Sequencing of Normal Fallopian Tube Epithelial Cells. Cancer Cell *37*, 226-242.e7. https://doi.org/10.1016/j.ccell.2020.01.003.

17. Lengyel, E., Li, Y., Weigert, M., Zhu, L., Eckart, H., Javellana, M., Ackroyd, S., Xiao, J., Olalekan, S., Glass, D., et al. (2022). A molecular atlas of the human postmenopausal fallopian tube and ovary from single-cell RNA and ATAC sequencing. Cell Reports *41*. https://doi.org/10.1016/j.celrep.2022.111838.

18. Ulrich, N.D., Shen, Y., Ma, Q., Yang, K., Hannum, D.F., Jones, A., Machlin, J., Randolph, J.F., Smith, Y.R., Schon, S.B., et al. (2022). Cellular heterogeneity of human

fallopian tubes in normal and hydrosalpinx disease states identified using scRNA-seq. Developmental Cell *57*, 914-929.e7. https://doi.org/10.1016/j.devcel.2022.02.017.

19. Brand, J., Haro, M., Lin, X., Rimel, B.J., McGregor, S.M., Lawrenson, K., and Dinh, H.Q. (2024). Fallopian tube single cell analysis reveals myeloid cell alterations in high-grade serous ovarian cancer. iScience *27*, 108990. https://doi.org/10.1016/j.isci.2024.108990.

20. Liu, J., Gao, C., Sodicoff, J., Kozareva, V., Macosko, E.Z., and Welch, J.D. (2020). Jointly defining cell types from multiple single-cell datasets using LIGER. Nat Protoc *15*, 3632–3662. https://doi.org/10.1038/s41596-020-0391-8.

21. Garsed, D.W., Pandey, A., Fereday, S., Kennedy, C.J., Takahashi, K., Alsop, K., Hamilton, P.T., Hendley, J., Chiew, Y.-E., Traficante, N., et al. (2022). The genomic and immune landscape of long-term survivors of high-grade serous ovarian cancer. Nat Genet, 1–12. https://doi.org/10.1038/s41588-022-01230-9.

22. Jacomy, M., Venturini, T., Heymann, S., and Bastian, M. (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. PLOS ONE *9*, e98679. https://doi.org/10.1371/journal.pone.0098679.

23. Jiang, P., Du, W., and Wu, M. (2014). Regulation of the pentose phosphate pathway in cancer. Protein & Cell *5*, 592–602. https://doi.org/10.1007/s13238-014-0082-8.

24. Yang, T., Shu, X., Zhang, H.-W., Sun, L.-X., Yu, L., Liu, J., Sun, L.-C., Yang, Z.-H., and Ran, Y.-L. (2020). Enolase 1 regulates stem cell-like properties in gastric cancer cells by stimulating glycolysis. Cell Death Dis *11*, 870. https://doi.org/10.1038/s41419-020-03087-4.

25. Wang, L., Long, H., Zheng, Q., Bo, X., Xiao, X., and Li, B. (2019). Circular RNA circRHOT1 promotes hepatocellular carcinoma progression by initiation of NR2F6 expression. Mol Cancer *18*, 119. https://doi.org/10.1186/s12943-019-1046-7.

26. Liu, J., Li, T., and Liu, X.-L. (2017). DDA1 is induced by NR2F6 in ovarian cancer and predicts poor survival outcome. Eur Rev Med Pharmacol Sci *21*, 1206–1213.

27. Cheng, X.-H., Black, M., Ustiyan, V., Le, T., Fulford, L., Sridharan, A., Medvedovic, M., Kalinichenko, V.V., Whitsett, J.A., and Kalin, T.V. (2014). SPDEF Inhibits Prostate Carcinogenesis by Disrupting a Positive Feedback Loop in Regulation of the Foxm1 Oncogene. PLoS Genet *10*, e1004656. https://doi.org/10.1371/journal.pgen.1004656.

28. Noah, T.K., Lo, Y., Price, A., Chen, G., King, E., Washington, M., Aronow, B.J., and Shroyer, N.F. (2013). SPDEF Functions as a Colorectal Tumor Suppressor by Inhibiting β-Catenin Activity. Gastroenterology *144*, 1012-1023.e6. https://doi.org/10.1053/j.gastro.2013.01.043.

29. Lo, Y.-H., Chung, E., Li, Z., Wan, Y.-W., Mahe, M.M., Chen, M.-S., Noah, T.K., Bell, K.N., Yalamanchili, H.K., Klisch, T.J., et al. (2017). Transcriptional Regulation by ATOH1 and its Target SPDEF in the Intestine. Cellular and Molecular Gastroenterology and Hepatology *3*, 51–71. https://doi.org/10.1016/j.jcmgh.2016.10.001.

30. De Falco, A., Caruso, F., Su, X.-D., Iavarone, A., and Ceccarelli, M. (2023). A variational algorithm to detect the clonal copy number substructure of tumors from scRNA-seq data. Nat Commun *14*, 1074. https://doi.org/10.1038/s41467-023-36790-9.

31. Shlien, A., and Malkin, D. (2009). Copy number variations and cancer. Genome Medicine *1*, 62. https://doi.org/10.1186/gm62.

32. Cui, X., Shan, T., and Qiao, L. (2022). Collagen type IV alpha 1 (COL4A1) silence hampers the invasion, migration and epithelial–mesenchymal transition (EMT) of gastric cancer cells through blocking Hedgehog signaling pathway. Bioengineered *13*, 8972–8981. https://doi.org/10.1080/21655979.2022.2053799.

33. Huang, H., Wright, S., Zhang, J., and Brekken, R.A. (2019). Getting a grip on adhesion: Cadherin switching and collagen signaling. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1866*, 118472. https://doi.org/10.1016/j.bbamcr.2019.04.002.

34. Xu, J., Fang, Y., Chen, K., Li, S., Tang, S., Ren, Y., Cen, Y., Fei, W., Zhang, B., Shen, Y., et al. (2022). Single-Cell RNA Sequencing Reveals the Tissue Architecture in Human High-Grade Serous Ovarian Cancer. Clinical Cancer Research *28*, 3590–3602. https://doi.org/10.1158/1078-0432.CCR-22-0296.

35. Johnson, B.K., Rhodes, M., Wegener, M., Himadewi, P., Foy, K., Schipper, J.L., Siwicki, R.A., Rossell, L.L., Siegwald, E.J., Chesla, D.W., et al. (2022). Single-cell Total RNA Miniaturized sequencing (STORM-seq) reveals differentiation trajectories of primary human fallopian tube epithelium. Preprint at bioRxiv, https://doi.org/10.1101/2022.03.14.484332 https://doi.org/10.1101/2022.03.14.484332.

36. Siegel, R.L., Miller, K.D., and Jemal, A. (2016). Cancer statistics, 2016. CA: A Cancer Journal for Clinicians *66*, 7–30. https://doi.org/10.3322/caac.21332.

37. Cook, D.P., and Vanderhyden, B.C. (2022). Transcriptional census of epithelial-mesenchymal plasticity in cancer. Science Advances *8*, eabi7640. https://doi.org/10.1126/sciadv.abi7640.

38. Lawrenson, K., Fonseca, M.A.S., Liu, A.Y., Segato Dezem, F., Lee, J.M., Lin, X., Corona, R.I., Abbasi, F., Vavra, K.C., Dinh, H.Q., et al. (2019). A Study of High-Grade Serous Ovarian Cancer Origins Implicates the SOX18 Transcription Factor in Tumor Development. Cell Reports *29*, 3726-3735.e4. https://doi.org/10.1016/j.celrep.2019.10.122.

39. Nakamura, K., Nakayama, K., Ishikawa, N., Ishikawa, M., Sultana, R., Kiyono, T., and Kyo, S. (2018). Reconstitution of high-grade serous ovarian carcinoma from primary fallopian tube secretory epithelial cells. Oncotarget *9*, 12609–12619. https://doi.org/10.18632/oncotarget.23035.

40. Izar, B., Tirosh, I., Stover, E.H., Wakiro, I., Cuoco, M.S., Alter, I., Rodman, C., Leeson, R., Su, M.-J., Shah, P., et al. (2020). A single-cell landscape of high-grade serous ovarian cancer. Nat Med *26*, 1271–1279. https://doi.org/10.1038/s41591-020-0926-0.

41. Yoshihara, K., Tajima, A., Yahata, T., Kodama, S., Fujiwara, H., Suzuki, M., Onishi, Y., Hatae, M., Sueyoshi, K., Fujiwara, H., et al. (2010). Gene Expression Profile for Predicting Survival in Advanced-Stage Serous Ovarian Cancer Across Two Independent Datasets. PLoS ONE *5*, e9615. https://doi.org/10.1371/journal.pone.0009615.

42. Ganzfried, B.F., Riester, M., Haibe-Kains, B., Risch, T., Tyekucheva, S., Jazic, I., Wang, X.V., Ahmadifar, M., Birrer, M.J., Parmigiani, G., et al. (2013). curatedOvarianData: clinically annotated data for the ovarian cancer transcriptome. Database *2013*. https://doi.org/10.1093/database/bat013.

43. Konecny, G.E., Wang, C., Hamidi, H., Winterhoff, B., Kalli, K.R., Dering, J., Ginther, C., Chen, H.-W., Dowdy, S., Cliby, W., et al. (2014). Prognostic and Therapeutic Relevance of Molecular Subtypes in High-Grade Serous Ovarian Cancer. JNCI: Journal of the National Cancer Institute *106*. https://doi.org/10.1093/jnci/dju249.

44. Tan, T.Z., Miow, Q.H., Huang, R.Y., Wong, M.K., Ye, J., Lau, J.A., Wu, M.C., Bin Abdul Hadi, L.H., Soong, R., Choolani, M., et al. (2013). Functional genomics identifies five distinct molecular subtypes with clinical relevance and pathways for growth control in epithelial ovarian cancer. EMBO Mol Med *5*, 1051–1066. https://doi.org/10.1002/emmm.201201823.

45. Axelrod, M.L., Cook, R.S., Johnson, D.B., and Balko, J.M. (2019). Biological Consequences of MHC-II Expression by Tumor Cells in Cancer. Clin Cancer Res *25*, 2392–2402. https://doi.org/10.1158/1078-0432.CCR-18-3200.

46. Sharma, P., Goswami, S., Raychaudhuri, D., Siddiqui, B.A., Singh, P., Nagarajan, A., Liu, J., Subudhi, S.K., Poon, C., Gant, K.L., et al. (2023). Immune checkpoint therapy—current perspectives and future directions. Cell *186*, 1652–1669. https://doi.org/10.1016/j.cell.2023.03.006.

47. Chen, Y., McAndrews, K.M., and Kalluri, R. (2021). Clinical and therapeutic relevance of cancer-associated fibroblasts. Nat Rev Clin Oncol *18*, 792–804. https://doi.org/10.1038/s41571-021-00546-5.

48. De Francesco, E.M., Lappano, R., Santolla, M.F., Marsico, S., Caruso, A., and Maggiolini, M. (2013). HIF-1α/GPER signaling mediates the expression of VEGF induced by hypoxia in breast cancer associated fibroblasts (CAFs). Breast Cancer Res *15*, R64. https://doi.org/10.1186/bcr3458.

49. Shapouri□Moghaddam, A., Mohammadian, S., Vazini, H., Taghadosi, M., Esmaeili, S., Mardani, F., Seifi, B., Mohammadi, A., Afshari, J.T., and Sahebkar, A. (2018). Macrophage plasticity, polarization, and function in health and disease. Journal Cellular Physiology *233*, 6425–6440. https://doi.org/10.1002/jcp.26429.

50. Allavena, P., Sica, A., Garlanda, C., and Mantovani, A. (2008). The Yin□Yang of tumor□associated macrophages in neoplastic progression and immune surveillance. Immunological Reviews *222*, 155–161. https://doi.org/10.1111/j.1600-065X.2008.00607.x.

51. Mantovani, A., Allavena, P., Sica, A., and Balkwill, F. (2008). Cancer-related inflammation. Nature *454*, 436–444. https://doi.org/10.1038/nature07205.

52. Varol, C. (2019). Tumorigenic Interplay Between Macrophages and Collagenous Matrix in the Tumor Microenvironment. In Collagen Methods in Molecular Biology., I. Sagi and N. A. Afratis, eds. (Springer New York), pp. 203–220. https://doi.org/10.1007/978-1-4939-9095-5_15.

53. Ledermann, J.A., Drew, Y., and Kristeleit, R.S. (2016). Homologous recombination deficiency and ovarian cancer. European Journal of Cancer *60*, 49–58. https://doi.org/10.1016/j.ejca.2016.03.005.

54. Kuchenbaecker, K.B., Hopper, J.L., Barnes, D.R., Phillips, K.-A., Mooij, T.M., Roos-Blom, M.-J., Jervis, S., Van Leeuwen, F.E., Milne, R.L., Andrieu, N., et al. (2017). Risks of Breast, Ovarian, and Contralateral Breast Cancer for *BRCA1* and *BRCA2* Mutation Carriers. JAMA *317*, 2402. https://doi.org/10.1001/jama.2017.7112.

55. Stanciu, P.I., Ind, T.E.J., Barton, D.P.J., Butler, J.B., Vroobel, K.M., Attygalle, A.D., and Nobbenhuis, M.A.E. (2019). Development of Peritoneal Carcinoma in women diagnosed with Serous Tubal Intraepithelial Carcinoma (STIC) following Risk-Reducing

Salpingo-Oophorectomy (RRSO). J Ovarian Res *12*, 50. https://doi.org/10.1186/s13048-019-0525-1.

56. Chen, F., Gaitskell, K., Garcia, M., Albukhari, A., Tsaltas, J., and Ahmed, A. (2017). Serous tubal intraepithelial carcinomas associated with high□grade serous ovarian carcinomas: a systematic review. BJOG *124*, 872–878. https://doi.org/10.1111/1471-0528.14543.

57. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

58. Dobin, A., and Gingeras, T.R. (2015). Mapping RNA-seq Reads with STAR. Curr Protoc Bioinformatics *51*, 11.14.1-11.14.19. https://doi.org/10.1002/0471250953.bi1114s51.

59. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

60. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140. https://doi.org/10.1093/bioinformatics/btp616.

61. Chen, Y., Lun, A.T.L., and Smyth, G.K. (2016). From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. F1000Res *5*, 1438. https://doi.org/10.12688/f1000research.8987.2.

62. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat Commun *8*, 14049. https://doi.org/10.1038/ncomms14049.

63. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biology *19*, 15. https://doi.org/10.1186/s13059-017-1382-0.

64. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573-3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

65. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. JOSS *4*, 1686. https://doi.org/10.21105/joss.01686.

66. Angerer, P., Haghverdi, L., Büttner, M., Theis, F.J., Marr, C., and Buettner, F. (2016). destiny: diffusion maps for large-scale single-cell data in R. Bioinformatics *32*, 1241–1243. https://doi.org/10.1093/bioinformatics/btv715.

67. Wolf, F.A., Hamey, F.K., Plass, M., Solana, J., Dahlin, J.S., Göttgens, 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. https://doi.org/10.1186/s13059-019-1663-x.

68. Moon, K.R., van Dijk, D., Wang, Z., Gigante, S., Burkhardt, D.B., Chen, W.S., Yim, K., Elzen, A. van den, Hirn, M.J., Coifman, R.R., et al. (2019). Visualizing structure and transitions in high-dimensional biological data. Nature Biotechnology *37*, 1482–1492. https://doi.org/10.1038/s41587-019-0336-3.

69. Sherman, B.T., Hao, M., Qiu, J., Jiao, X., Baseler, M.W., Lane, H.C., Imamichi, T., and Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res *50*, W216–W221. https://doi.org/10.1093/nar/gkac194.

70. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498. https://doi.org/10.1038/s41586-018-0414-6.

71. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol *38*, 1408–1414. https://doi.org/10.1038/s41587-020-0591-3.

72. Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nature Methods *14*, 1083–1086. https://doi.org/10.1038/nmeth.4463.

73. Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., Christmas, R., Avila-Campilo, I., Creech, M., Gross, B., et al. (2007). Integration of biological networks and gene expression data using Cytoscape. Nat Protoc *2*, 2366–2382. https://doi.org/10.1038/nprot.2007.324.

74. Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq

highlights intratumoral heterogeneity in primary glioblastoma. Science *344*, 1396–1401. https://doi.org/10.1126/science.1254257.

75. Cheng, K., Seita, Y., Moriwaki, T., Noshiro, K., Sakata, Y., Hwang, Y.S., Torigoe, T., Saitou, M., Tsuchiya, H., Iwatani, C., et al. (2022). The developmental origin and the specification of the adrenal cortex in humans and cynomolgus monkeys. Science Advances *8*, eabn8485. https://doi.org/10.1126/sciadv.abn8485.

FIGURE LEGENDS

Figure 1. Bulk RNA-seq analyses unveil transcriptomic correlation among human HGSOCs, fallopian tubes and ovaries

(A) Overview of the study design and workflow.

(B) Principal component (PC) analyses of 26 human samples analyzed by bulk RNA-seq, including 10 HGSOC, 4 ovarian, and 12 segmented fallopian tube samples. EOC, epithelial ovarian cancer; FTF, fallopian tube fimbria; FTA, fallopian tube ampulla; FTI, fallopian tube isthmus; OV, ovary.

(C) Dendrogram showing transcriptomic distance among all 26 samples analyzed by bulk RNA-seq based on ward.D2 method.

(D) Heatmap showing the top 2,000 variable genes and the transcriptomic signatures of EOC-Fs (Fallopian-derived EOCs), fallopian tubes, and ovaries. Color bar showing Z-score normalized gene expression. Red is for higher expression and blue is for lower expression.

(E) Gene ontology (GO) term enrichment analyses of seven genesets representing the transcriptomic signatures of EOC-Fs (Fallopian-derived EOCs), fallopian tubes, and ovaries.

See also Figures S1, Tables S1.

Figure 2. snRNA-seq analyses reveal human FTE cell subpopulations and the development trajectory

(A) UMAP representation of 12 distinct FTE cell clusters revealed by graph-based Louvain clustering. Each color represents a cell cluster, and the color scheme remains the same in all subsequent panels (B-F).

(B) Dendrogram showing hierarchical clustering of the twelve subpopulations identified in human FTE. Cells were partitioned using ward.D2 based on the average expression of each cell cluster.

(C) Bubble plot showing average expression levels and prevalence of major markers used to annotate the FTE cell types. Color indicating averaged gene expression and the size of circle showing the percentage of cells expressing corresponding gene in the cell cluster.

(D) RNA velocity analyses of all FTE cell subpopulations to predict the developmental trajectory. Each color represents a cell cluster. Arrows pointing the potential developmental directions.

(E) PAGA analyses of developmental trajectory and connectivity among 12 FTE subpopulations. Each color represents a cell cluster, and the thickness of the lines represents correlations between FTE cell clusters.

(F) Visualization of the developmental trajectories using PHATE. Each point represents a cell, and each color stands for a cell sub-cluster. Branches show different differentiation directions or fates.

(G) Pseudotime of the developmental trajectories. Pseudotime was calculated based on the principal curve by PHATE embedding regardless of sub-clusters.

(H and I) Pseudotime heat maps showing the transcriptomic dynamics from FTE stem cells to secretory cells (H) and ciliated cells (I) in human FTE. Color bar showing the Z-score normalized gene expression.

(J and K) Immunofluorescence showing the localization and co-expression of LGR5, PGR and Ki67 in the isthmus and ampulla segments of human FTE. White, DAPI; Blue, Ki67; Green, PGR; Red, LGR5. Yellow arrows point to the small round basal stem cells. Scale bars, 10μm.

See also Figures S2, Tables S3 and S4.

Figure 3. snRNA-seq analyses uncover a high degree of variation and cellular heterogeneity of HGSOCs.

(A) UMAP representation of 13 distinct cell subclusters identified in HGSOC samples by graph-based Louvain clustering. Each color represents a cell cluster.

(B) Dendrogram showing hierarchical clustering of the 13 subclusters identified in HGSOC samples. The HGSOC cells were partitioned using ward.D2 based on the average expression of each cell cluster. CAFs, cancer-associated fibroblasts; EMT, epithelial-mesenchymal transition; MHC, major histocompatibility complexes.

(C) UMAP representation of 8 cell subgroups identified in HGSOC samples. Each color represents a subgroup of cells in HGSOC samples.

(D) Heatmap showing the transcriptomic signatures for each of the eight HGSOC cell subtypes. Enriched GO terms in each of the 8 subtypes of HGSCO cells are presented next to the heatmap.

(E and F) UMAP display of cell subclusters and their identity in all malignant cells of HGSOC samples.

(G) Calculated correlation of transcriptomic distance of pure malignant cell subgroups identified in HGSOC samples. Color density represents correlation coefficient size.

(H) Average expression levels and prevalence of major markers used to annotate the malignant cell types in HGSOC samples.

See also Figures S3 and Table S5.

Figure 4. snRNA-seq analyses predict the trajectory leading to the transformation from normal FTE cells to HGSOC cells

(A) Combination of cancer cells of HGSOC and all FTE cells by force-directed graph visualization identified 15 cell clusters that were further divided into three big partitions for predicting potential cell lineages using Slingshot.

(B) Cell trajectories built by PAGA. Each color represents a cell cluster, and the line thickness indicates the degree of correlation among all cell clusters.

(C) RNA velocity analyses based on force-directed graph embedding, with the arrows pointing to the developmental directions.

(D) Pseudotime analyses of all tumor and FTE cells regardless of subclusters. The differentiation from 0 to 1 is represented by blue to yellow.

(E) Dendrogram showing hierarchical clustering of the 15 subclusters identified in combined HGSOC tumor and FTE cells, which were partitioned using ward.D2 based on the average expression of each cell cluster.

(F) Histogram showing the proportional composition of 15 cell subtypes in the three segments of human fallopian tubes (Isth: isthmus; Amp: ampulla; Fimb: fimbria) and HGSOC.

(G) Bubble plot showing expression of top marker genes in the 15 cell clusters identified.

(H and I) Pseudotime heatmaps showing the transcriptomic dynamics from FTE stem cells to immunoreactive-type tumor cells (cluster 12) (H) and mesenchymal-type tumor cells (cluster 13) (I).

See also Table S7 and S8.

Figure 5. SCENIC analyses of snRNA-seq data.

(A) Single-cell regulatory network inference and clustering (SCENIC) analyses of the combined HGSOC tumor and FTE cells, with five cell subtypes annotated.

(B) SCENIC analyses of transcription factor (TF) regulatory networks. Heatmap was drawn using Cytoscape, and the rows (representing genes) and columns (indicating cell subtypes) were ordered using unsupervised hierarchical clustering based on Euclidian distances. Each module represents a cell.

(C) Specific regulons in the five cell subtypes are listed based on AUC score.

(D) Overlapping between gene expression (the top panels) and AUC score (the bottom panels) of the regulons.

(E) Genes regulatory networks with predicted master regulators in different cell subtypes.

Figure 6. The copy number variation of HGSOC.

(A) Significant copy number variants (CNVs) detected in HGSOC, with FTE cells as reference.

(B) CNV regions detected using the Hidden Markov Model (HMM). The blue-to-red color transition represents the CNV region's state from 0 to 6. Each row represents a region.

(C) Pie chart showing the proportional distribution of the six CNV states (top panels) and histogram indicating the frequency of the six states of CNV regions on each chromosome (bottom panel). State1 is 0x: complete loss, State2 is 0.5x: loss of one copy, State3 is 1x: neutral, State4 is 1.5x: addition of one copy, State5 is 2x: addition of two copies, State6 is 3x: essentially a placeholder for >2x copies but modeled as 3x..

(D) Heatmaps showing the gain or loss of gene copy numbers.

See also Table S9.

Figure 7. Schematic summary of the major findings of the present study.

Figure S1-S5

Figure S1. Transcriptomic correlation among human HGSOCs, fallopian tubes, and ovaries revealed by Bulk RNA-seq.

(A) Box graph showing expression distribution, based on FPKM values, in each of the 26 samples analyzed by bulk RNA-seq.

(B) Heatmap showing pair-wise Pearson correlation coefficient among the 26 samples.

(C, D, E and F) Scatter plot of the DEGs between fallopian-derived EOC (EOC_F) and ovarian-derived EOC (EOC_O) (C), EOC_F and fallopian tube (D), EOC_F and ovary (E), fallopian tube and ovary (F), and top enriched GO terms.

Figure S2. Proportional composition of major cell subtypes and functional enrichment in human fallopian tubes uncovered by snRNA-seq.

(A) snRNA-seq analyses of human fallopian tubes led to 87,920 cells that passed through the initial quality control, 36,000-63,000 mean reads per cell, and ~2,000 median genes per cell.

(B and C) UMAP representation of cell cycle scoring of all FTE cells after removing cell cycle effects, colored by (B) cell cycle state and (C) samples.

(D) UMAP representation of five major cell types identified in human fallopian tubes.

(E) UMAP representation of the expression of selected cell subgroup marker genes.

(F) Heatmap showing top genes and enriched gene ontology (GO) terms in each of the five major cell types in human fallopian tubes. Expression of genes is Z-score normalized.

(G) Histogram showing the proportional distribution of five major cell types in three segments of human fallopian tubes.

(H) Expression of seven key marker genes projected to PHATE-based trajectory analyses.

(I) Histograms showing the proportional distribution of subtypes in three segments of human fallopian tubes.

(J) Histogram showing proportional distribution of 12 subclusters in three segments of the human fallopian tubes.

See also Table S2.

Figure S3. Prediction and identification of cell subtypes of the fallopian tube and HGSOC.

(A and B) Immunohistochemical analyses showing the localization (A) and semi-quantitative results (B) of TUBB4B (red arrow) and PAX8 (yellow arrow) in three segments of human fallopian tubes. Scale bars, 100μm. The statistical results are presented as means ± SEM. Student's t-test. ns ≥ 0.05; *p < 0.05; **p < 0.01; ***p < 0.001. Isth: isthmus, Amp: ampulla; Fim: fimbria.

(C) Immunofluorescence analysis showing the localization and co-expression of LGR5, PGR, and Ki67 in the FTE of fimbria. White, DAPI; Blue, Ki67; Green, PGR; Red, LGR5. Scale bars, 50μm.

(D) Quality control data of 4 HGSOC samples in scRNA-seq analyses.

(E and F) UMAP representation of cell cycle scoring of all HGSOC cells after removing cell cycle effects, colored by (E) cell cycle state and (F) samples.

(G) Pearson correlation among four HGSOC samples.

(H) UMAP representation of the expression of selected cell subgroup marker genes.

Figure S4. Specificity and heterogeneity of pure HGSOC tumor cells in snRNA-seq.

(A and B) Immunohistochemical analyses showing localization (A) and relative abundance (B) of PAX8, TUBB4B, MUC16, and Ki67 in HGSOC and isthmus of the fallopian tube. Scale bars, 20μm. The statistical results are presented as means ± SEM. Student's t-test; ns ≥ 0.05; **p< 0.01; ****p< 0.0001. Isth: isthmus; HGSOC: High-grade serous ovarian cancer.

(C) Heatmap showing top genes and enriched gene ontology (GO) terms in each of the five major HGSOC tumor cell subtypes. Expression of genes is Z-score normalized.

(D) UMAP representation of the expression of selected HGSOC tumor cell subgroup marker genes.

See also Table S6.

Figure S5. Intercellular connections analysis in HGSOC.

(A and B) Ligand-receptor connections between each cancer and non-cancer cell subgroup of HGSOC, including number of interactions (A) and weights/strength of interactions (B). Lines indicate intercellular connections between ligands and receptors, and the line thickness is proportional to the number of ligand-receptor pairs.

(C) Overview of selected interactions of ligand-receptor pairs in outgoing and incoming signaling patterns**.**

(D) Ligand-receptor connections between each cancer and non-cancer cell subgroup of HGSOC in the collagen signaling pathway network.

(E) Violin plot showing marker genes of collagen signaling pathways in HGSOC tumor and non-tumor cell subgroups.

SUPPLEMENTARY DATASETS: see separate Excel documents

Table S1. Top 1000 variable genes among different samples, related with Figure 1DE.

Table S2. Markers identified among different cell types in human fallopian tube snRNA-seq, related with Figure S2F.

Table S2. Top 2000 variable genes along the secretory trajectory, related with Figure 2H.

Table S4. Top 2000 variable genes along the ciliated trajectory, related with Figure 2I.

Table S5. Markers identified in HGSOC snRNA-seq, related with Figure 3D.

Table S6. Markers identified among subtypes of HGSOC cancer, related with Figure S4C.

Table S7. Top 200 variable genes along the trajectory from cluster 11 to cluster 12, related with Figure 4H.

Table S8. Top 200 variable genes along the trajectory from cluster 11 to cluster13, related with Figure 4I.

Table S9. HMM predicted CNVs regions, related with Figure 6B.

Figure 1

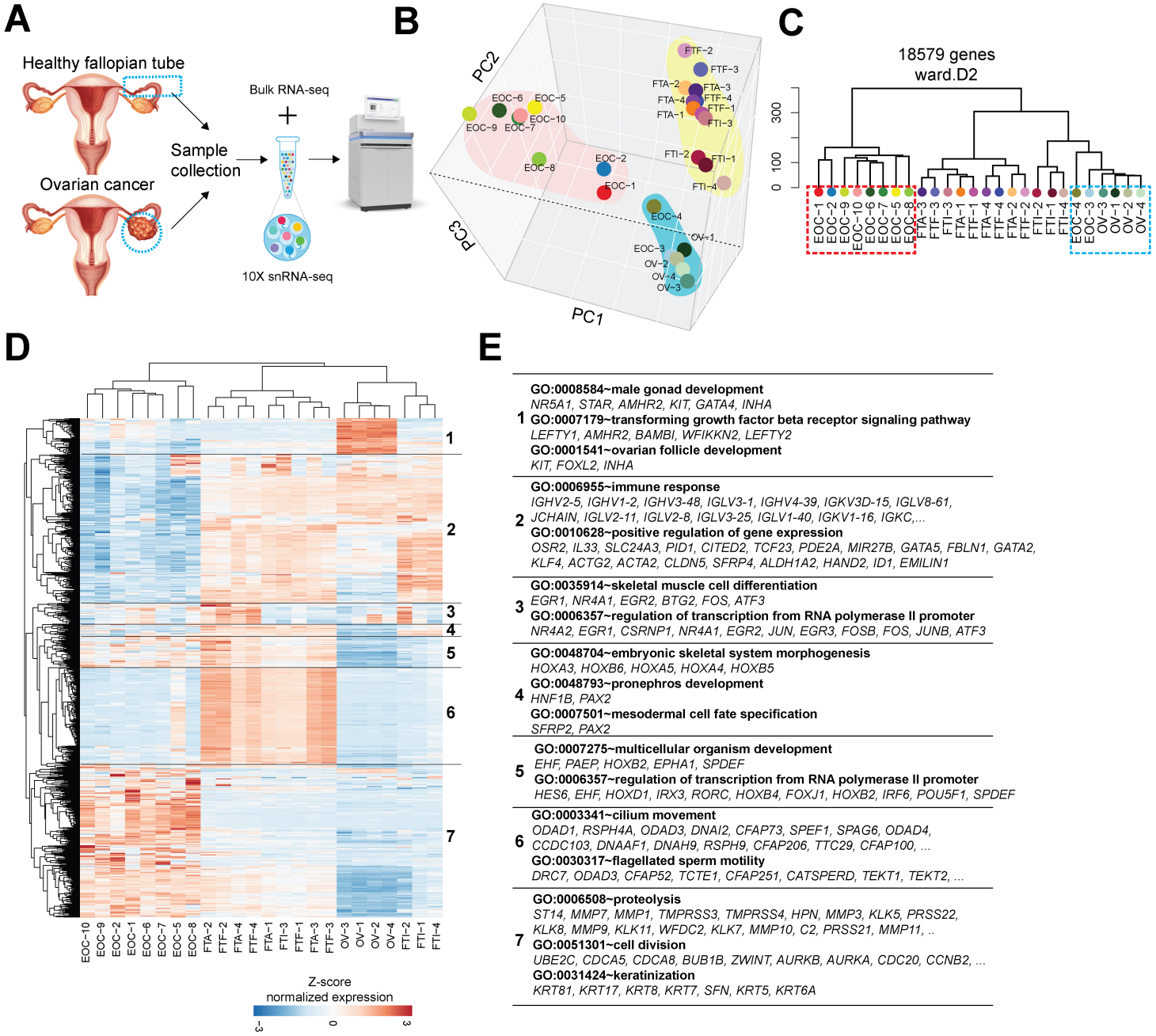

A
Healthy fallopian tube
Ovarian cancer
Sample collection
Bulk RNA-seq
+
10X snRNA-seq
B
PC2
PC3
PC1
EOC−6
EOC−5
EOC−10
EOC−9
EOC−7
EOC−8
EOC−2
EOC−1
EOC−4
EOC−3
OV−1
OV−2
OV−4
OV−3
FTF−2
FTF−3
FTA−2
FTA−3
FTA−4
FTF−4
FTF−1
FTA−1
FTI−3
FTI−2
FTI−1
FTI−4
C
18579 genes
ward.D2
300
100
0
EOC−1
EOC−2
EOC−9
EOC−10
EOC−6
EOC−7
EOC−5
EOC−8
FTA−3
FTF−3
FTI−3
FTA−1
FTF−1
FTA−4
FTF−4
FTA−2
FTF−2
FTI−2
FTI−1
FTI−4
EOC−4
EOC−3
OV−3
OV−1
OV−2
OV−4
D
1
2
3
4
5
6
7
EOC−10
EOC−9
EOC−2
EOC−1
EOC−6
EOC−7
EOC−5
EOC−8
FTA−2
FTF−2
FTA−4
FTF−4
FTA−1
FTI−3
FTF−1
FTA−3
FTF−3
OV−3
OV−1
OV−2
OV−4
FTI−2
FTI−1
FTI−4
Z-score
normalized expression
−3
0
3
E
1
GO:0008584~male gonad development
NR5A1, STAR, AMHR2, KIT, GATA4, INHA
GO:0007179~transforming growth factor beta receptor signaling pathway
LEFTY1, AMHR2, BAMBI, WFIKKN2, LEFTY2
GO:0001541~ovarian follicle development
KIT, FOXL2, INHA
2
GO:0006955~immune response
IGHV2-5, IGHV1-2, IGHV3-48, IGLV3-1, IGHV4-39, IGKV3D-15, IGLV8-61, JCHAIN, IGLV2-11, IGLV2-8, IGLV3-25, IGLV1-40, IGKV1-16, IGKC,...
GO:0010628~positive regulation of gene expression
OSR2, IL33, SLC24A3, PID1, CITED2, TCF23, PDE2A, MIR27B, GATA5, FBLN1, GATA2, KLF4, ACTG2, ACTA2, CLDN5, SFRP4, ALDH1A2, HAND2, ID1, EMILIN1
3
GO:0035914~skeletal muscle cell differentiation
EGR1, NR4A1, EGR2, BTG2, FOS, ATF3
GO:0006357~regulation of transcription from RNA polymerase II promoter
NR4A2, EGR1, CSRNP1, NR4A1, EGR2, JUN, EGR3, FOSB, FOS, JUNB, ATF3
4
GO:0048704~embryonic skeletal system morphogenesis
HOXA3, HOXB6, HOXA5, HOXA4, HOXB5
GO:0048793~pronephros development
HNF1B, PAX2
GO:0007501~mesodermal cell fate specification
SFRP2, PAX2
5
GO:0007275~multicellular organism development
EHF, PAEP, HOXB2, EPHA1, SPDEF
GO:0006357~regulation of transcription from RNA polymerase II promoter
HES6, EHF, HOXD1, IRX3, RORC, HOXB4, FOXJ1, HOXB2, IRF6, POU5F1, SPDEF
6
GO:0003341~cilium movement
ODAD1, RSPH4A, ODAD3, DNAI2, CFAP73, SPEF1, SPAG6, ODAD4, CCDC103, DNAAF1, DNAH9, RSPH9, CFAP206, TTC29, CFAP100, ...
GO:0030317~flagellated sperm motility
DRC7, ODAD3, CFAP52, TCTE1, CFAP251, CATSPERD, TEKT1, TEKT2, ...
7
GO:0006508~proteolysis
ST14, MMP7, MMP1, TMPRSS3, TMPRSS4, HPN, MMP3, KLK5, PRSS22, KLK8, MMP9, KLK11, WFDC2, KLK7, MMP10, C2, PRSS21, MMP11, ..
GO:0051301~cell division
UBE2C, CDCA5, CDCA8, BUB1B, ZWINT, AURKB, AURKA, CDC20, CCNB2, ...
GO:0031424~keratinization
KRT81, KRT17, KRT8, KRT7, SFN, KRT5, KRT6A

Figure 2

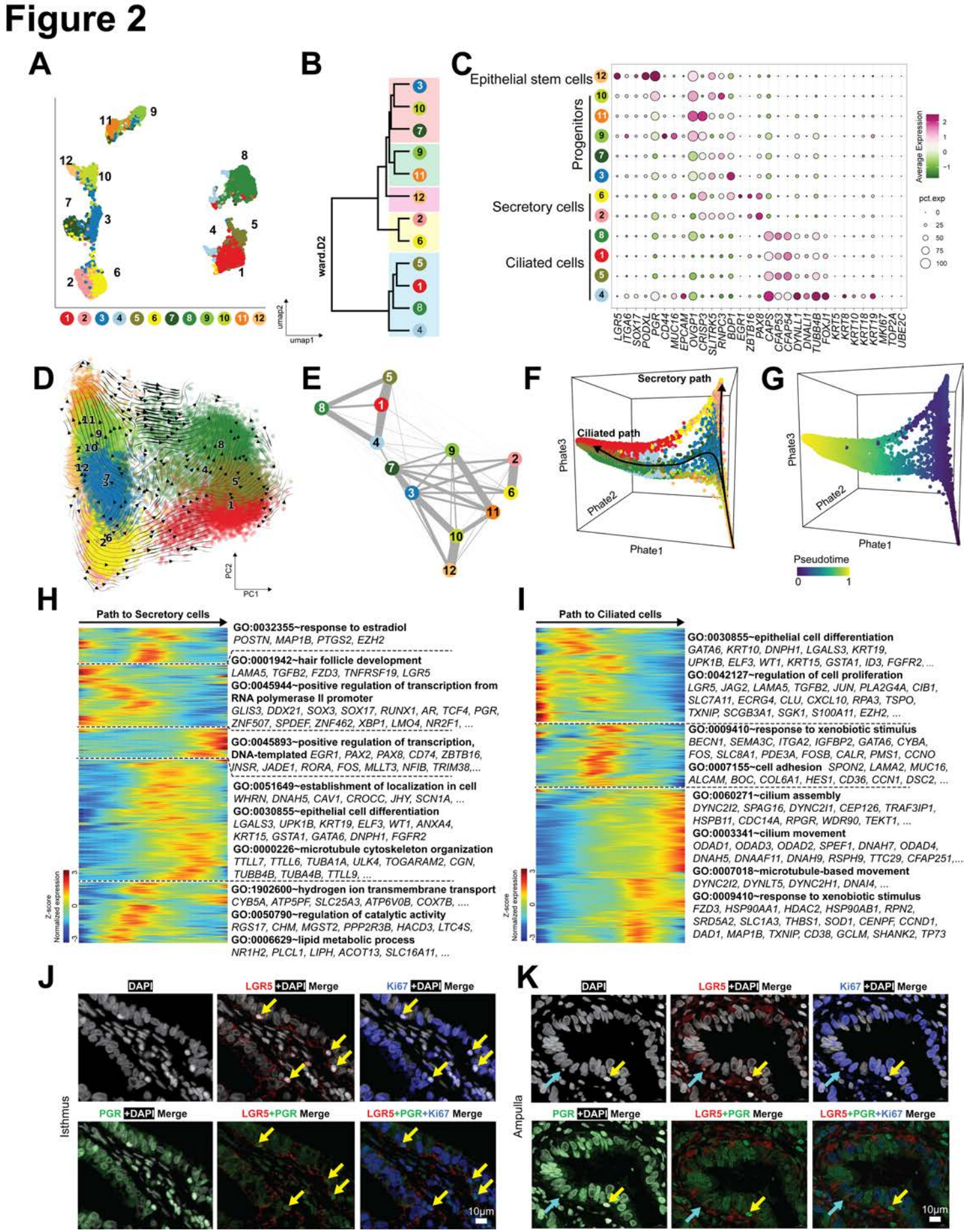


Figure 2
A
B
C
D
E
F
G
H
I
J
K
umap1
umap2
ward.D2
Epithelial stem cells
Progenitors
Secretory cells
Ciliated cells
Average Expression
pct.exp
LGR5 ITGA6 SOX17 PODXL PGR CD44 MUC16 EPCAM OVGP1 CRISP3 SLITRK5 RNPC3 BDP1 EGR1 ZBTB16 PAX8 CAPS CFAP53 CFAP54 DYNLL1 DNALI1 TUBB4B FOXJ1 KRT5 KRT8 KRT10 KRT18 KRT19 MKI67 TOP2A UBE2C
PC1
PC2
Secretory path
Ciliated path
Phate1
Phate2
Phate3
Pseudotime
Path to Secretory cells
GO:0032355~response to estradiol
POSTN, MAP1B, PTGS2, EZH2
GO:0001942~hair follicle development
LAMA5, TGFB2, FZD3, TNFRSF19, LGR5
GO:0045944~positive regulation of transcription from RNA polymerase II promoter
GLIS3, DDX21, SOX3, SOX17, RUNX1, AR, TCF4, PGR, ZNF507, SPDEF, ZNF462, XBP1, LMO4, NR2F1, ...
GO:0045893~positive regulation of transcription, DNA-templated EGR1, PAX2, PAX8, CD74, ZBTB16, INSR, JADE1, RORA, FOS, MLLT3, NFIB, TRIM38,...
GO:0051649~establishment of localization in cell
WHRN, DNAH5, CAV1, CROCC, JHY, SCN1A, ...
GO:0030855~epithelial cell differentiation
LGALS3, UPK1B, KRT19, ELF3, WT1, ANXA4, KRT15, GSTA1, GATA6, DNPH1, FGFR2
GO:0000226~microtubule cytoskeleton organization
TTLL7, TTLL6, TUBA1A, ULK4, TOGARAM2, CGN, TUBB4B, TUBA4B, TTLL9, ...
GO:1902600~hydrogen ion transmembrane transport
CYB5A, ATP5PF, SLC25A3, ATP6V0B, COX7B, ...
GO:0050790~regulation of catalytic activity
RGS17, CHM, MGST2, PPP2R3B, HACD3, LTC4S,
GO:0006629~lipid metabolic process
NR1H2, PLCL1, LIPH, ACOT13, SLC16A11, ...
Path to Ciliated cells
GO:0030855~epithelial cell differentiation
GATA6, KRT10, DNPH1, LGALS3, KRT19, UPK1B, ELF3, WT1, KRT15, GSTA1, ID3, FGFR2, ...
GO:0042127~regulation of cell proliferation
LGR5, JAG2, LAMA5, TGFB2, JUN, PLA2G4A, CIB1, SLC7A11, ECRG4, CLU, CXCL10, RPA3, TSPO, TXNIP, SCGB3A1, SGK1, S100A11, EZH2, ...
GO:0009410~response to xenobiotic stimulus
BECN1, SEMA3C, ITGA2, IGFBP2, GATA6, CYBA, FOS, SLC8A1, PDE3A, FOSB, CALR, PMS1, CCNO
GO:0007155~cell adhesion SPON2, LAMA2, MUC16, ALCAM, BOC, COL6A1, HES1, CD36, CCN1, DSC2, ...
GO:0060271~cilium assembly
DYNC2I2, SPAG16, DYNC2I1, CEP126, TRAF3IP1, HSPB11, CDC14A, RPGR, WDR90, TEKT1, ...
GO:0003341~cilium movement
ODAD1, ODAD3, ODAD2, SPEF1, DNAH7, ODAD4, DNAH5, DNAAF11, DNAH9, RSPH9, TTC29, CFAP251,...
GO:0007018~microtubule-based movement
DYNC2I2, DYNLT5, DYNC2H1, DNAI4, ...
GO:0009410~response to xenobiotic stimulus
FZD3, HSP90AA1, HDAC2, HSP90AB1, RPN2, SRD5A2, SLC1A3, THBS1, SOD1, CENPF, CCND1, DAD1, MAP1B, TXNIP, CD38, GCLM, SHANK2, TP73
Z-score
Normalized expression
DAPI
LGR5 +DAPI Merge
Ki67 +DAPI Merge
PGR +DAPI Merge
LGR5+PGR Merge
LGR5+PGR+Ki67 Merge
Isthmus
Ampulla
10μm

Figure 3

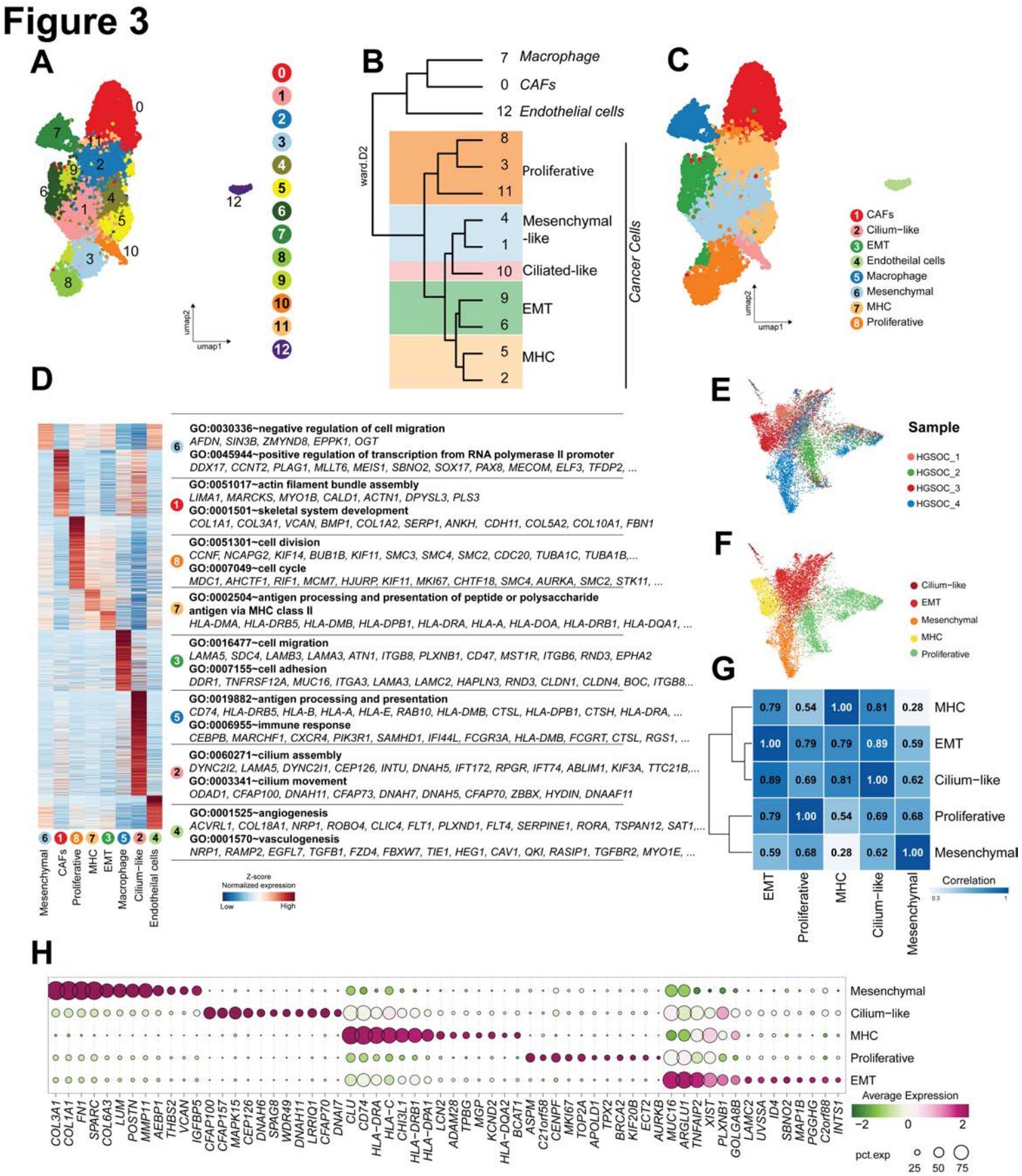


Figure 3
A
B
C
D
E
F
G
H
umap1
umap2
ward.D2
7 Macrophage
0 CAFs
12 Endothelial cells
Proliferative
Mesenchymal -like
Ciliated-like
EMT
MHC
Cancer Cells
1 CAFs
2 Cilium−like
3 EMT
4 Endotheilal cells
5 Macrophage
6 Mesenchymal
7 MHC
8 Proliferative
Sample
HGSOC_1
HGSOC_2
HGSOC_3
HGSOC_4
Cilium−like
EMT
Mesenchymal
MHC
Proliferative
GO:0030336~negative regulation of cell migration
AFDN, SIN3B, ZMYND8, EPPK1, OGT
GO:0045944~positive regulation of transcription from RNA polymerase II promoter
DDX17, CCNT2, PLAG1, MLLT6, MEIS1, SBNO2, SOX17, PAX8, MECOM, ELF3, TFDP2, ...
GO:0051017~actin filament bundle assembly
LIMA1, MARCKS, MYO1B, CALD1, ACTN1, DPYSL3, PLS3
GO:0001501~skeletal system development
COL1A1, COL3A1, VCAN, BMP1, COL1A2, SERP1, ANKH, CDH11, COL5A2, COL10A1, FBN1
GO:0051301~cell division
CCNF, NCAPG2, KIF14, BUB1B, KIF11, SMC3, SMC4, SMC2, CDC20, TUBA1C, TUBA1B,...
GO:0007049~cell cycle
MDC1, AHCTF1, RIF1, MCM7, HJURP, KIF11, MKI67, CHTF18, SMC4, AURKA, SMC2, STK11, ...
GO:0002504~antigen processing and presentation of peptide or polysaccharide antigen via MHC class II
HLA-DMA, HLA-DRB5, HLA-DMB, HLA-DPB1, HLA-DRA, HLA-A, HLA-DOA, HLA-DRB1, HLA-DQA1, ...
GO:0016477~cell migration
LAMA5, SDC4, LAMB3, LAMA3, ATN1, ITGB8, PLXNB1, CD47, MST1R, ITGB6, RND3, EPHA2
GO:0007155~cell adhesion
DDR1, TNFRSF12A, MUC16, ITGA3, LAMA3, LAMC2, HAPLN3, RND3, CLDN1, CLDN4, BOC, ITGB8...
GO:0019882~antigen processing and presentation
CD74, HLA-DRB5, HLA-B, HLA-A, HLA-E, RAB10, HLA-DMB, CTSL, HLA-DPB1, CTSH, HLA-DRA, ...
GO:0006955~immune response
CEBPB, MARCHF1, CXCR4, PIK3R1, SAMHD1, IFI44L, FCGR3A, HLA-DMB, FCGRT, CTSL, RGS1, ...
GO:0060271~cilium assembly
DYNC2I2, LAMA5, DYNC2I1, CEP126, INTU, DNAH5, IFT172, RPGR, IFT74, ABLIM1, KIF3A, TTC21B,...
GO:0003341~cilium movement
ODAD1, CFAP100, DNAH11, CFAP73, DNAH7, DNAH5, CFAP70, ZBBX, HYDIN, DNAAF11
GO:0001525~angiogenesis
ACVRL1, COL18A1, NRP1, ROBO4, CLIC4, FLT1, PLXND1, FLT4, SERPINE1, RORA, TSPAN12, SAT1,...
GO:0001570~vasculogenesis
NRP1, RAMP2, EGFL7, TGFB1, FZD4, FBXW7, TIE1, HEG1, CAV1, QKI, RASIP1, TGFBR2, MYO1E, ...
Mesenchymal
CAFs
Proliferative
MHC
EMT
Macrophage
Cilium−like
Endotheilal cells
Z-score
Normalized expression
Low
High
0.79 0.54 1.00 0.81 0.28 MHC
1.00 0.79 0.79 0.89 0.59 EMT
0.89 0.69 0.81 1.00 0.62 Cilium−like
0.79 1.00 0.54 0.69 0.68 Proliferative
0.59 0.68 0.28 0.62 1.00 Mesenchymal
EMT
Proliferative
MHC
Cilium−like
Mesenchymal
Correlation
0.3
1
Mesenchymal
Cilium−like
MHC
Proliferative
EMT
COL3A1 COL1A1 FN1 SPARC COL6A3 LUM POSTN MMP11 AEBP1 THBS2 VCAN IGFBP5 CFAP100 CFAP157 MAPK15 CEP126 DNAH6 SPAG8 WDR49 DNAH11 LRRIQ1 CFAP70 DNAI7 CLU CD74 HLA−DRA HLA−C CHI3L1 HLA−DRB1 HLA−DPA1 LCN2 ADAM28 TPBG MGP KCND2 HLA−DQA2 BCAT1 ASPM C21orf58 CENPF MKI67 TOP2A APOLD1 TPX2 BRCA2 KIF20B ECT2 AURKB MUC16 ARGLU1 TNFAIP2 XIST PLXNB1 GOLGA8B LAMC2 UVSSA ID4 SBNO2 MAP1B PGGHG C2orf88 INTS1
Average Expression
−2
0
2
pct.exp
25
50
75

Figure 4

Figure 4

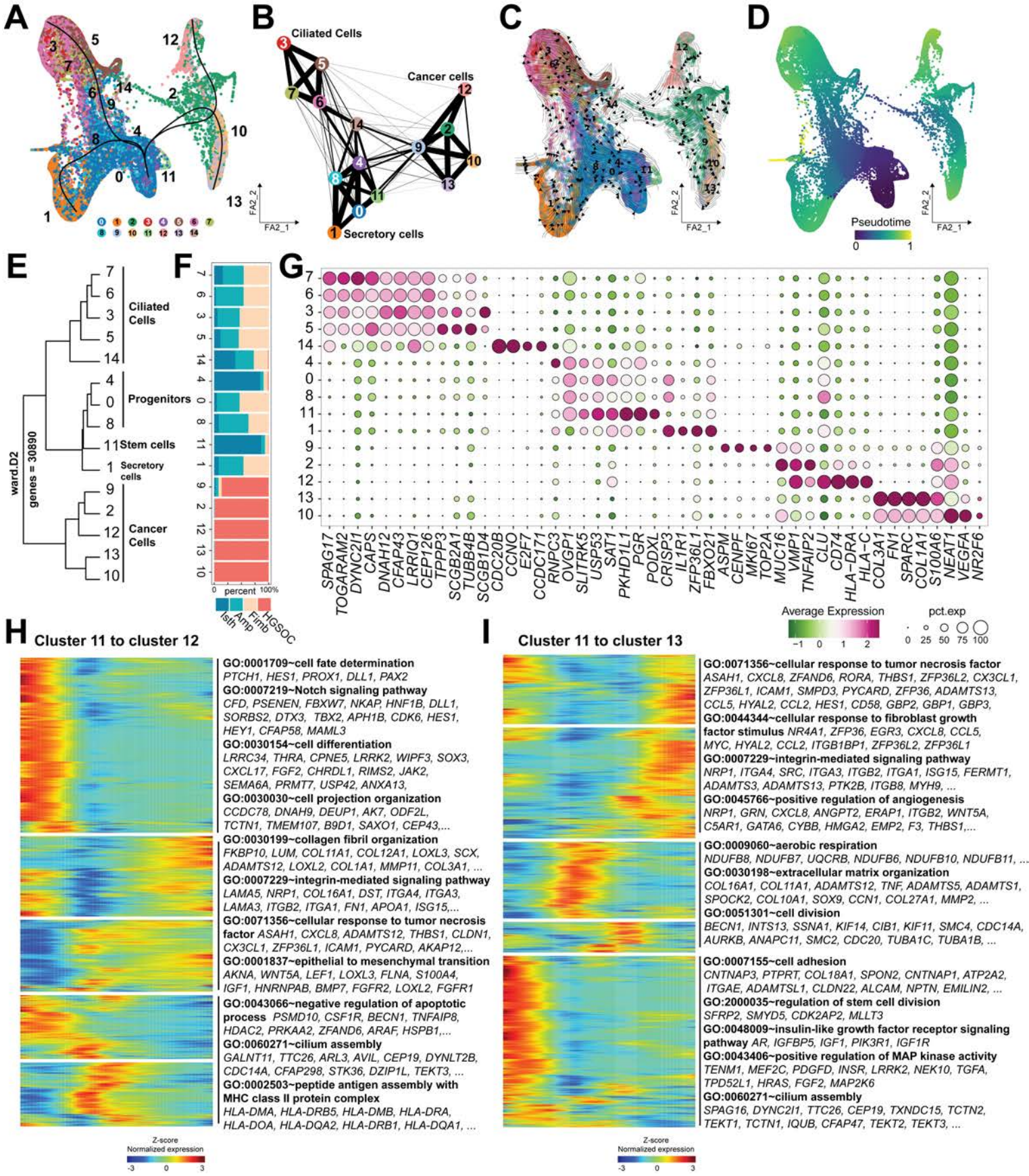


Figure 5

Figure 5

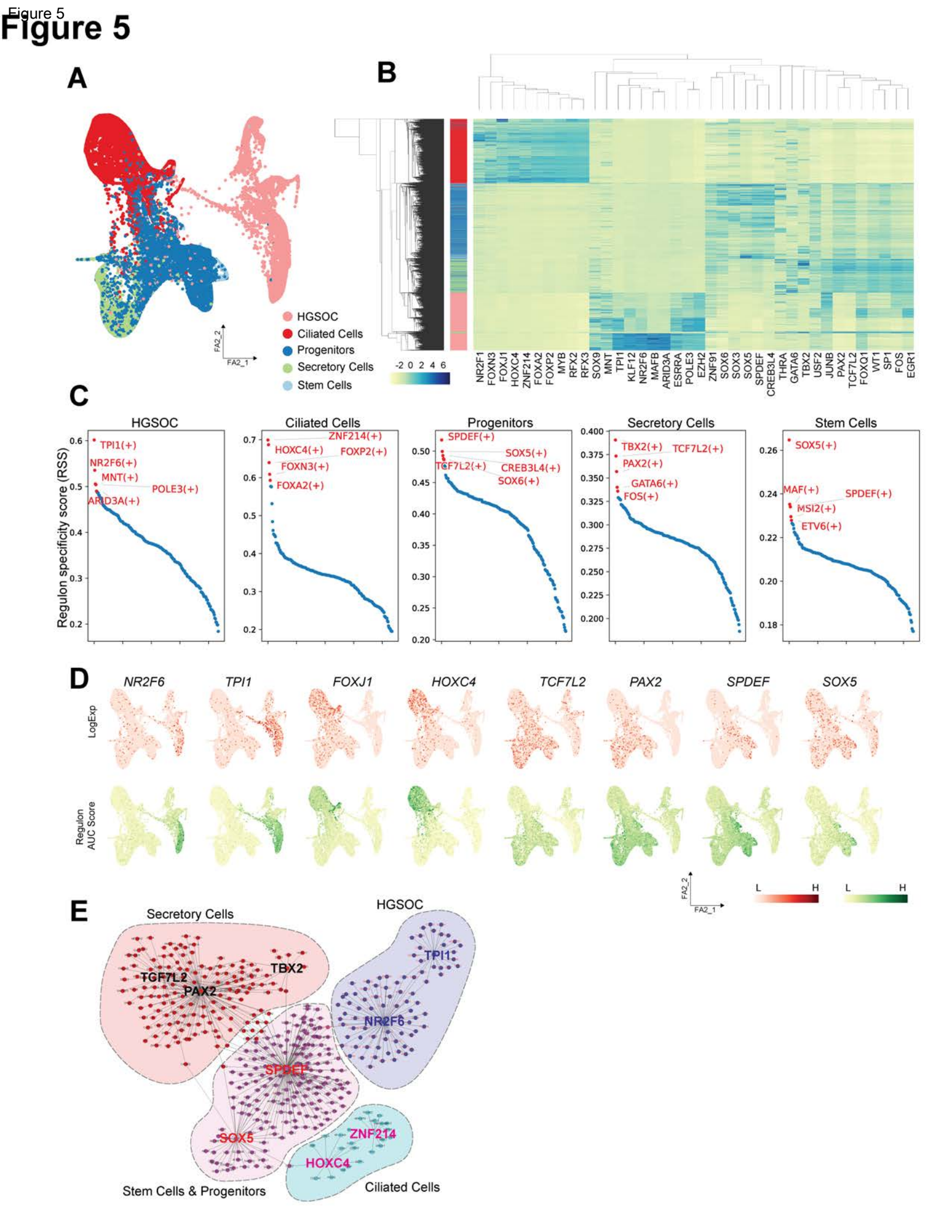


Figure 6

Figure 6

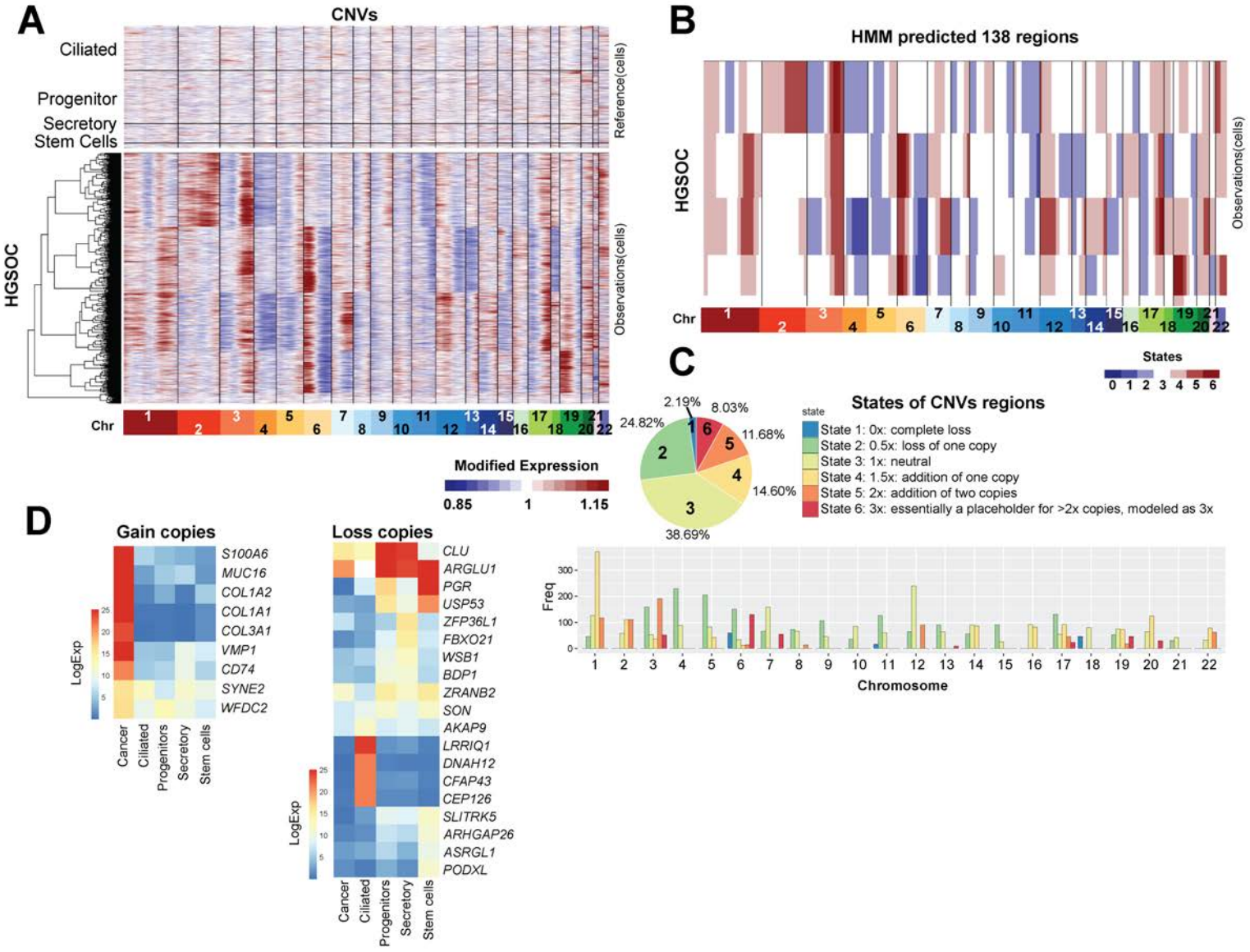
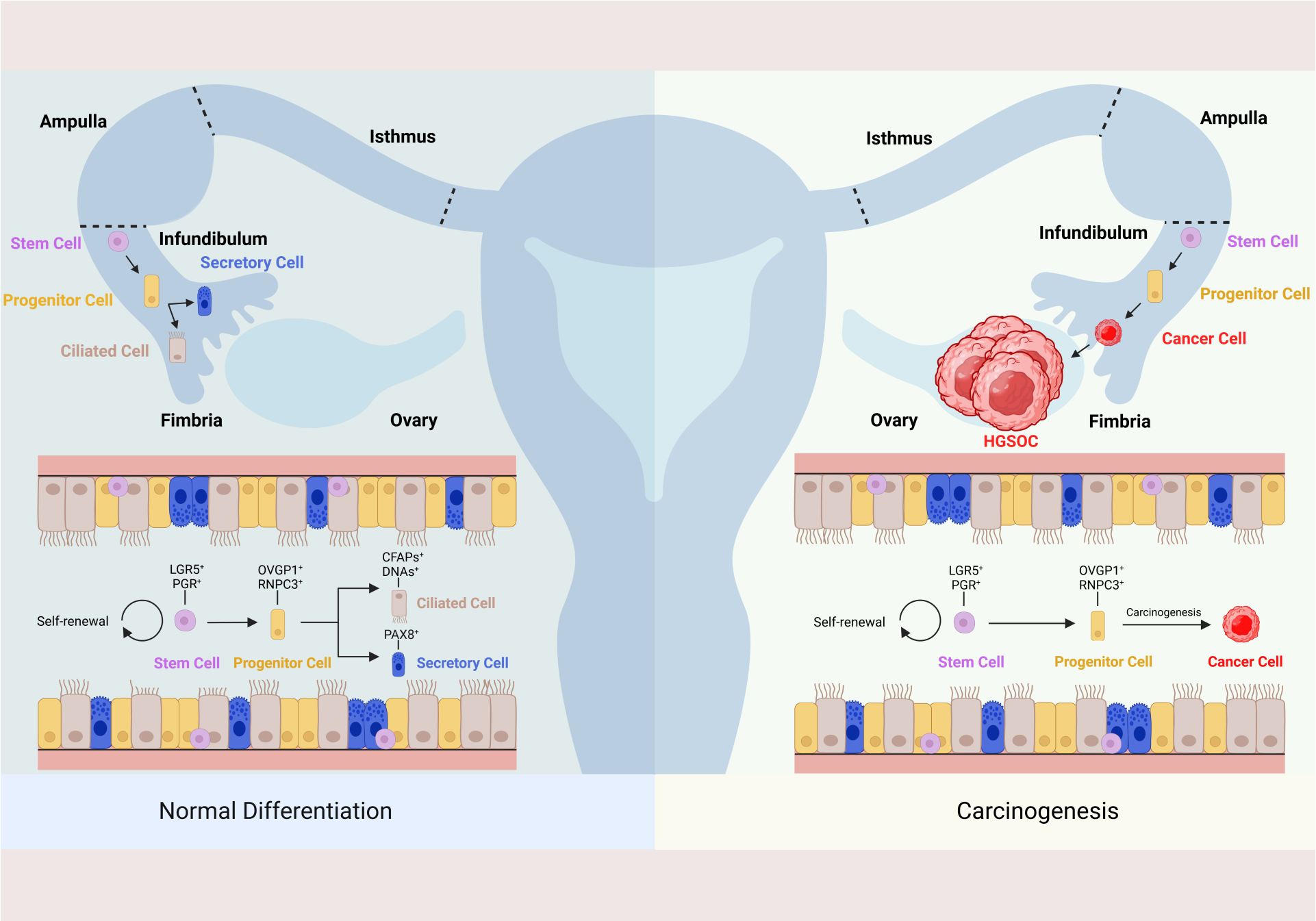


Figure 7


Ampulla
Isthmus
Stem Cell
Infundibulum
Secretory Cell
Progenitor Cell
Ciliated Cell
Fimbria
Ovary
LGR5+
PGR+
OVGP1+
RNPC3+
CFAPs+
DNAs+
Ciliated Cell
Self-renewal
PAX8+
Stem Cell
Progenitor Cell
Secretory Cell
Normal Differentiation
Isthmus
Ampulla
Infundibulum
Stem Cell
Progenitor Cell
Cancer Cell
Ovary
Fimbria
HGSOC
LGR5+
PGR+
OVGP1+
RNPC3+
Self-renewal
Carcinogenesis
Stem Cell
Progenitor Cell
Cancer Cell
Carcinogenesis