ANDROGEN ACTION IN CELL FATE AND COMMUNICATION DURING PROSTATE DEVELOPMENT AT SINGLE-CELL RESOLUTION

Dong-Hoon Lee[1¶], Adam W. Olson[1¶], Jinhui Wang[2], Won Kyung Kim[1], Jiaqi Mi[1], Hong Zeng[3], Vien Le[1], Joseph Aldahl[1], Alex Hiroto[1], Xiwei Wu[2], Zijie Sun[1*]

[1]Department of Cancer Biology, [2]Integrative Genomics Core, Cancer Center and Beckman Research Institute, City of Hope, Duarte, CA 91010; [3]Transgenic, Knockout and Tumor Model Center, Stanford University School of Medicine, Stanford, CA 94305

¶These authors contributed equally to the work

*Corresponding Author

Correspondence: Zijie Sun, Ph.D, MD, Beckman Research Institute and Cancer Center, Beckman Bldg, Room 2311, 1500 Duarte Road, Duarte, CA 91010, Tel: (626) 218-0955, Email: zjsun@coh.org

Key words: androgen signaling/prostate development/Wnt signaling/Sonic hedgehog signaling/mouse models.

Summary statement: Using newly generated mouse genetic tools, the authors demonstrate the dynamic regulation of mesenchymal androgen signaling as cellular niches controlling prostate early development.

ABSTRACT

Androgens/androgen receptor (AR) mediated signaling pathways are essential for prostate development, morphogenesis, and regeneration. Specifically, stromal AR-signaling has been shown to be essential for prostatic initiation. However, the molecular mechanisms underlying AR-initiated mesenchymal-epithelial interactions in prostate development remain unclear. Here, using a newly generated mouse model, we directly addressed the fate and role of genetically marked AR-expressing cells during embryonic prostate development. Androgen signaling-initiated signaling pathways were identified in mesenchymal niche populations at single cell transcriptomic resolution. The dynamic cell-signaling networks regulated by stromal AR were characterized in regulating prostatic epithelial bud formation. Pseudotime analyses further revealed the differentiation trajectory and fate of AR-expressing cells in both prostatic mesenchymal and epithelial cell populations. Specifically, the cellular properties of *Zeb1*-expressing progenitors were assessed. Selective deletion of AR signaling in a subpopulation mesenchymal rather than epithelial cells dysregulates the expression of the master regulators and significantly impairs prostatic bud formation. These data provide novel, high-resolution evidence demonstrating the important role of mesenchymal androgen signaling as cellular niches controlling prostate early development by initiating dynamic mesenchyme-epithelia cell interactions.

INTRODUCTION

The prostate develops from the endodermal urogenital sinus (UGS) that is derived from the caudal terminus of the hindgut called the cloaca (Cunha et al., 2018). Androgen signaling mediated through the androgen receptor (AR) is essential for prostate development (Gelmann, 2002). Mouse prostatic development initiates at embryonic day 17.5 (E17.5) from the UGS in response to rising levels of fetal testicular androgens (Cunha et al., 1987; Staack et al., 2003). Mutation of the *Ar* gene in the testicular feminized (Tfm) mice results in the complete absence of prostate development (Cunha and Chung, 1981). During embryogenesis, the AR is initially detected in the urogenital sinus mesenchyme (UGM) prior to the initiation of prostate budding and morphogenesis. Subsequently, its expression extends to the urogenital sinus epithelium (UGE) after the initiation of prostatic budding and branching morphogenesis (Cunha et al., 1987; Cooke et al., 1991; Takeda and Chang, 1991). Tissue recombination studies further demonstrated that mesenchymal, rather than epithelial, AR signaling plays a decisive role in inducing development of the prostatic epithelium through paracrine regulation (Cunha and Lung, 1978; Cunha, 1984). However, the cellular properties of AR-expressing cells and the mechanisms by which stromal androgen signaling initiates and regulates other pathways and regulators through mesenchymal-epithelial interactions during early prostatic development and morphogenesis remain unclear.

During early development, budding and branching morphogenesis takes place in a variety of organs, including the prostate (Ochoa-Espinosa and Affolter, 2012; Iber and Menshykau, 2013; Varner and Nelson, 2014). Despite biological differences in cellular branch structure and function between organs and species, the developmental programs

are largely conserved and are regulated by precise spatiotemporal epithelial-mesenchymal paracrine interactions (Thomson et al., 2002; Prins and Putz, 2008). Therefore, investigating mesenchymal androgen action in regulating prostatic budding and development will provide new and important insight to our understanding of cellular niches in branched organ development and their related human disorders.

The AR is a member of the nuclear hormone receptor superfamily (Chang et al., 1988). Androgen-induced AR transcriptional activity directly contributes to early prostate development and morphogenesis (Jenster et al., 1991). To trace AR-expressing cells and assess their functions during prostate development, we used gene-targeting approaches to generate a mouse $Ar^{IRES\text{-}Cre}$ allele, which enables us to genetically mark AR-expressing cells and trace their fate and function in early prostate development. Utilizing single-cell mRNA sequencing (scRNAseq) and other experimental approaches, we evaluated the cellular properties of AR-expressing cells at single cell resolution, and characterized the signaling pathways and factors that contribute to early prostatic development and morphogenesis in mouse UGS tissues. Trajectory analysis further explores the identity of prostatic progenitors and the regulatory role of AR in prostatic differentiation. These data provide novel, high resolution insight into cell fates and cell-cell interactions between prostatic mesenchyme and epithelium in early prostate development.

RESULTS

Genetically labeling AR-expressing cells using *Ar* $^{IRES\text{-}Cre}$ mice. Using gene targeting approaches, we generated a new mouse model, *Ar*$^{IRES\text{-}Cre}$ mice, by inserting *Cre* recombinase into the mouse *Ar* gene locus on the X chromosome through an engineering internal ribosome entry site (IRES) within the 3' untranslated region (Fig. 1A). This advanced mouse tools allows for *Ar* promoter directed expression of *Cre* recombinase wh.,ile maintaining endogenous *Ar* expression. Both male and female *Ar*$^{IRES\text{-}Cre}$ mice,

';?1B). The activity of *Cre* recombinase results in spontaneous recombination of the *floxed* reporter loci, generating a permanent genetic marker by switching from expression of membrane-bound tandem dimer Tomato (mT) to membrane-bound green fluorescent protein (mGFP) (Muzumdar et al., 2007) (Fig. 1B). Those genetically labeled cells will not only carry mGFP expression through their lifespan but will also pass it to their offspring, enabling us to trace the fate of AR-expressing cells and descendants.

Histological analysis of adjacent UGS sections prepared from male *R26R*$^{mTmG/+}$*:Ar*$^{IRES\text{-}Cre}$ embryos between the ages of E12.5 and E17.5 showed normal morphology of the UGS at each time point (Fig. 1C1-6). Co-staining of UGS tissues for vimentin, a mesenchymal cellular marker, and E-cadherin, an epithelial marker, visually separated epithelial and mesenchymal compartments in these sections (Fig. 1D1-6). Interestingly, *Cre*-induced mGFP expression first appears mainly within the UGE in E12.5 UGS tissues (Fig. 1E1). The expression of mGFP continues to increase in the UGE and, starting at E13.5, gradually reveals in the adjacent UGM (Fig. 1E2). The observation of mGFP expression in the UGE prior to that in the UGM is unexpected and

suggests that *Ar* transcription actually occurs first in UGE cells, differing from the results reported previously (Cunha et al., 1987; Cooke et al., 1991; Takeda and Chang, 1991). While the UGE appears almost completely mGFP positive by E14.5 and beyond, cells located in the protruding Wolffian duct remain mT positive through E17.5 (Blue arrows, Fig. 1E5-1G6). Immunofluorescent staining (IF) for mGFP showed a similar expression pattern as observed in mTmG assays at different time points (Fig. 1F1-6). Relatively weak AR staining appears within UGM and beyond UGE areas starting at E12.5 (Fig. 1G1 and 1I1). It gradually becomes stronger (Fig. 1G2-3), and extends to UGE areas, clearly visualized at E14.5 in both the UGE and UGM (Fig. 1G4-6). At E12.5 and E13.5, cytoplasmic staining of AR appeared in cells with mGFP and DAPI (white arrows, Fig. 1I1-2 and 1J1-2). However, uniform nuclear staining of AR clearly overlays with DAPI in mGFP positive UGE and UGM cells of E14.5 and elder UGS samples (Fig. 1I3-6 and 1J3-6). The above results showed that *Cre*-induced mGFP expression mimics the expression pattern of endogenous AR but appears stronger than AR expression in UGS tissues. Additionally, the observation of mGFP prior to AR expression in UGE areas at E12.5 suggests that different translation processes through the 5' cap and the IRES site take place for AR and *Cre* protein synthesis, respectively (Komar and Hatzoglou, 2011). Nevertheless, the current $R26R^{mTmG/+}$:$Ar^{IRES\text{-}Cre}$ model provides a new tool for detecting and tracing AR-expressing cells.

Assessing urogenital sinus cell populations using single-cell RNA sequencing. To assess the characteristics of AR-expressing cells in early prostate development, we performed single-cell RNA sequencing (scRNAseq) analyses using UGS tissues isolated from male $R26R^{mTmG/+}$:$Ar^{IRES\text{-}Cre}$ embryos (Fig. 2A). At E17.5, UGE and immediately

surrounding mesenchyme reveal positive staining for *Ar-Cre* induced mGFP expression, whereas much of the more distant stroma is mT positive, especially closer to the bladder (Fig. 2B). In addition, the Wolffian duct epithelial cells and their surrounding stromal cells, remain mT positive (blue arrow, Fig. 2B). Embryonic UGS tissues were dissected, and non UGS tissues, including the Wolffian duct, bladder, and urethra, were removed (Fig. 2C), and approximately 11,000 viable cells were collected and used for sequencing with 10x Genomics Chromium Single Cell 3' Solution (Fig. 2A). Following sequencing and alignment to the mm10 reference genome with an added mGFP sequence (Chen et al.), 10,478 cells were included after the initial quality control in the analysis using the Seurat package (Butler et al.). Cells then underwent filtering to remove potential empty droplets, doublets, and low-quality cells with high percentages of mitochondrial RNA. After this final filtering step, 9,440 cells with an average of 4,976 genes and 31,214 UMI counts per cell were used for future analyses (Fig. S1A-B). When visualized in two dimensions based on unbiased transcription profiling using t-Distributed Stochastic Neighbor Embedding (tSNE), twenty-one distinguished cellular clusters were identified (Fig. 2D). The cellular properties of these clusters were assessed using mesenchymal and epithelial cellular markers, including smooth muscle actin (*Acta2*), vimentin (*Vim*), and epithelial cell adhesion molecule (*Epcam*) (Fig. 2E and Fig. S2). Based on their expression patterns, epithelial and mesenchymal cell populations were determined and grouped (Fig. 2F). Seven stromal and four epithelial cell types were further identified (Fig. 2G-I) based on their transcription profiles in comparison with previously reported urogenital cellular markers and scRNAseq datasets from mouse and human prostate tissues (Georgas et al., 2015; Cunha et al., 2018; Henry et al., 2018; Kwon et al., 2019).

Five highly expressed genes representing specific cellular properties in each cluster were shown (Fig. 2I). Based on these cellular markers, four epithelial clusters were identified as urogenital sinus epithelial (UGE), bladder epithelial (BLE), Wolffian duct epithelial (WDE), and other epithelial (OE) cells. Most of the epithelial cells were in the UGE cluster with high expression of *Krt15*, *Krt4*, and *Tspan1*. These UGE cells also showed higher expression of *Ar* and *mGFP* compared to other epithelial cells (Fig. S2). The cells in the BLE cluster displayed high expression of *Upk3a*, *Upk2* and *Sptssb*, typical bladder cellular markers (Georgas et al., 2015; Habuka et al., 2015) (Fig. 2I), and low expression of *Ar* and *mGFP* compared to the UGE. The expression of *Pax8*, *Pax2*, and *Hoxb7*, WDE cell markers, showed in the WDE cluster (Narlis et al., 2007; Georgas et al., 2015) (Fig. 2B, Fig. S2A). Observation of small BLE and WDE cell clusters in the above sample is expected due to the inability to completely remove those tissues from the UGS during sample preparation. Lastly, the OE cell cluster contains about 60 cells with high expression of *Lrrn4*, *Gpm6a*, and *Muc16*, suggesting their mesothelial cell properties (Kanamori-Katayama et al., 2011).

Seven stromal cell types were identified in the analysis. Most of the stromal cells possess fibroblast properties with *Dcn*, *Fbln1*, and *Gsn* expression (Guerrero-Juarez et al., 2019; Karpus et al., 2019). Smooth muscle cells are the second largest stromal cell population featuring high expression of *Acta2*, *Tagln*, and *Myh11* (Brun et al., 2015). Other minor stromal clusters were identified as leukocytes (Henry et al., 2018), endothelial cells (Fontijn et al., 2014), peripheral glial cells and neurons (Jessen and Mirsky, 2005; Jacob et al., 2014), and myoblasts (Ganassi et al., 2018), representing similar cellular compositions of the UGS at this timepoint as reported previously

(Georgas et al., 2015). Among those stromal clusters, *Ar* and *mGFP* are highly expressed in fibroblasts, with some limited expression in smooth muscle cells (Fig. 2I and Fig. S2A). These data provide a comprehensive landscape of the cell populations and properties of the male embryonic urogenital sinus at single cell resolution.

Characterization of AR signaling in both urogenital epithelial and stromal cells. Using the transcription profiles generated with scRNAseq, we further assessed the expression of endogenous *Ar* and *Cre*-induced *mGFP* at single-cell resolution in embryonic UGS tissues at E17.5. Most AR-expressing cells appear in the fibroblast and UGE cell clusters, and the signals from the former reveal more intense than those from the latter (Fig. 3A-C). *Cre*-induced *mGFP* expression also appeared in both fibroblast and UGE cell clusters (Fig. 3B-C). The level of *mGFP* expression between the fibroblast and UGE cell clusters is comparable and is, in both cases, stronger than that of endogenous *Ar* (Fig. 3B-C). The different expression pattern between *mGFP* and endogenous *Ar* may be attributed to the fact that their transcription is regulated under different promoters, following the initial *Cre*-mediated activation of *mGFP* (Muzumdar et al., 2007). Analysis of cell distributions showed almost half of stromal and epithelial cells being *mGFP* and *Ar* positive (Fig. 3D). However, about a quarter of stromal and a third of epithelial cells appeared only *mGFP* positive, reflecting the transient nature of *Ar* expression in early prostate development. Few cells showed *Ar* expression without mGFP expression, suggesting *Ar* expression may have just initiated in those cells. Lastly, about a quarter of cells showed no detection of both *Ar* and *mGFP* expression in both the stroma and epithelia (Fig. 3D). Using IF approaches, we validated the expression of endogenous AR and mGFP as well as other cell markers in UGS tissues. AR and mGFP

double positive cells appear in both the UGE and UGM areas (Fig. 3F1-4). Interestingly, staining for mGFP in the UGE cells appears stronger than UGM cells (Fig. 3F1). AR and mGFP double positive cells also comprise prostatic buds (arrows, Fig. 3F3-4). The majority of UGE cells appear uniformly positive for E-cadherin and mGFP, however, weaker E-cadherin staining reveals in areas along the exterior epithelial layer, especially the protruding prostatic buds (Fig. 3G1-4). Co-staining of mGFP with vimentin was observed in the stromal cells directly surrounding the UGE cells (Fig. 3H1-4). Smooth muscle cells appear to be separated from the UGE by several layers of fibroblasts and exhibit limited overlay with mGFP (Grishina et al., 2005) (Fig. 3I1-4). The above co-IF results provide *in situ* expression patterns in support of the previous scRNAseq transcription profile analyses.

The cellular properties of urogenital epithelial and mesenchymal cells were further evaluated using both scRNAseq and immunostaining. *Krt8*, a luminal epithelial marker appears to be expressed uniformly throughout epithelial cell clusters (Fig. 3E), supported by relatively uniform positive staining for CK8 within the UGE (Fig. 3J1-2). *Krt5*, *Krt14* and *Trp63*, prostatic basal cell markers, are localized to only the upper portion of the UGE cluster in tSNE expression plots (Fig. 3E), and showed positive staining localized to the outer basal layer of the UGE and the prostatic buds (Fig. 3E, 3K1-2, 3L1-2 and 3M1-2). Staining of mGFP was also present in UGS areas, and overlaid with the above cellular markers. Interestingly, the expression of *Fkbp5* and *Serpinb2*, two AR downstream target genes, showed different patterns. The expression of *Fkbp5* appears in both mesenchymal and epithelial cells (Fig. 3E and 3N1-2) whereas *Serpinb2* is mainly expressed within the urogenital and bladder epithelial cells. Sonic

hedgehog (Shh) mainly reveals in the UGE and bladder cell clusters while *Gli1*, a downstream target gene of Shh, appears in fibroblasts and smooth muscle cells (Fig. 3E), implicating their reciprocal regulation in early prostate development (Podlasek et al., 1999). Positive staining for SHH was also observed in the UGE cells (Fig. 3O1-2). The above results from both tSNE expression and co-IF staining provide new and consistent insight into the cellular properties of AR-expressing cells in embryonic UGS tissues.

High resolution single-cell profiling of urogenital sinus epithelial cells. To assess androgen signaling initiated paracrine interactions between prostatic epithelial and mesenchymal cells, the epithelial cells were grouped and re-clustered, again yielding UGE, BLE, WDE, and OE cell clusters (Fig. 4A and Fig. S3A-C). Given the biological significance of UGE cells in prostate development, we further re-clustered these cells. Re-clustering resulted in four clusters that were identified from the innermost to the outermost layer of the UGE, including superficial, intermediate, basal, and budding UGE clusters based on urogenital epithelial cellular markers (Bhatia-Gaur et al., 1999; Grishina et al., 2005; Abler et al., 2011; Keil et al., 2012) (Fig. 4A-B and Fig. S3D-E). While nearly all cells within these clusters are *mGFP* positive, levels of *Ar* expression vary with the highest expression in the basal and intermediate UGE (Fig. 4C). Expression of prostatic budding markers, such as *Nkx3.1*, *Wif1* and *Bmp7*, were mainly localized to the budding UGE cells, as well as *Shh* and *Hoxd13*, a transcription regulator involved in prostatic differentiation (Javed and Langley, 2014) (Fig. 4D). Interestingly, *Etv4*, which interacts with Shh and Fgf signaling in branching morphogenesis (Lu et al., 2009; Herriges et al., 2015; Zhang et al., 2016), was also highly expressed in the budding UGE cells. A full list of DEGs specific to the budding UGE was generated (Table S1).

IHC analysis showed AR expression in both UGM and UGE areas but weaker expression presented in budding epithelial cells than adjacent stromal cells (Fig. 4E1-2). NKX3.1 and ETV4 expression appears selectively in the budding epithelial cells (Fig. 4E3-4) while the expression of BMP7, HOXD13, and WIF1 reveal in both budding epithelial and surrounding stromal cells (Fig. 4E5-7). Interestingly, staining for SHH appears noticeably weaker in the budding epithelial cells than in adjacent basal UGE cells (Fig. 4E8). A similar staining pattern has been observed in the early study (Podlasek et al., 1999). These data provide a clear breakdown of UGE cells at single-cell resolution during the onset of prostatic bud formation.

Trajectory analysis of urogenital epithelial differentiation. To explore the differentiation patterns of UGE cells we performed trajectory analysis using Monocle's reverse graph embedding (Qiu et al., 2017). Three well-defined branches with budding UGE and superficial UGE cells concentrated at two opposing branch tips were generated with Monocle analysis. Basal and intermediate UGE cells comprised most of the central region of the plot displaying an overall arch from the innermost to the outermost UGE cell types (Fig. 5A-B). A mixture of basal and intermediate UGE cells formed the third, shorter branch tip (Fig. 5A). Prostatic basal cells, particularly, *Zeb1* expressing basal cells, have been implicated to possess stem/progenitor cell properties and the ability to differentiate to prostatic luminal cells (Toivanen et al., 2016), (Wang et al., 2020). Consistent with this recent study (Wang et al., 2020), the expression of *Zeb1* appears concentrated at the end of the basal and intermediate UGE branch in the trajectory plot (Fig. 5C). Therefore, in the pseudotime plot, the point of origin was set as the small basal/intermediate branch displaying focal *Zeb1* expression (Fig. 5B-C). This group of

Zeb1 expressing cells also showed focal expression of other EMT related genes (Fig 5C-D). A list of DEGs specific to this cell population was identified (Table S2). From that population, two distinct differentiation paths leading to prostatic bud cells and superficial UGE were revealed (Fig. 5B-C). These differentiation paths similarly start with a reduction in EMT markers, although the budding epithelial cells display a slight recovery of EMT related genes, possibly contributing to their increased motility (Fig. 5D, right panel). *Nkx3.1*, *Etv4*, and *Wif1* expression appears concentrated at the budding UGE branch tip (Fig. 5C-D). Although slightly diffuse, *Hoxd13* and *Bmp7* showed strongest expression in the budding UGE (Fig. 5C-D). These data suggest a regulatory role of *Etv4*, *Wif1* and *Nkx3.1* in differentiated budding cells while *Hoxd13* and *Bmp7* may act in initiating basal UGE cell differentiation to prostatic buds (Fig. 5C-D). Co-IF staining showed clear nuclear staining of ZEB1 in the majority of UGM cells however, only varying cytoplasmic staining in UGE cells (Fig. 5E1 and 5E5). Epithelial ZEB1 staining was also overlaid slightly with CK8 but not Trp63 staining in UGE cells (arrows, Fig. 5E4-5 and 5E7-8). The above observations appear similar to previous reports showing cytoplasmic staining of ZEB1 in epithelial cells (Spaderna et al., 2006; Goscinski et al., 2015). Using SCENIC upstream regulator analysis (Aibar et al., 2017), we further assessed the driving factors contributing to UGE cell differentiation (Fig. S3F). The top 20 transcription factors (TFs) were identified in each UGE cell type, including Wnt (Sp5 and Tcf7) and Notch signaling related regulators (Zfp64 and Hey1), as well as ETV family members (Fig. S3F). Taken together, these data elucidate the differentiation status of UGE cells and explore the cell properties of urogenital epithelial progenitor cells at single cell resolution.

Characterizing cellular properties of urogenital sinus mesenchymal cells. We similarly re-clustered the stromal cells to gain deeper insight into their cell characteristics (Fig. 6A). Eight cell types were identified (Fig. 6B). The newly identified cell type, pericytes, were previously grouped within the smooth muscle cells (Fig. 2G). High expression of *Ar* and *mGFP* appeared primarily in fibroblasts (Fig. S4C). Since urogenital sinus fibroblasts are directly adjacent to the UGE and involved in regulating the epithelial-mesenchymal paracrine interactions during early prostate development (Grishina et al., 2005; Abler et al., 2011), we re-clustered them to further characterize their cellular properties. Of the nine resulting clusters, one showed high expression levels of cell proliferation related genes (Fig. 6C). Using cell-cycle regression, we removed those genes from the principal component analysis (PCA). Following this regression, re-clustering produced 8 fibroblast clusters (Fig. 6D-E) that were grouped into four subtypes, including progenitor-like (ProGf), Col14a1-expressing (Col14a1f), peri-epithelial (PeriEf), and other fibroblasts (OF) (Fig. 6F).

We first identified *Bmp7, Wif1*, and *Wnt5a* expressing fibroblasts following reports that they are directly surrounding prostatic buds and involved in early prostate development and morphogenesis (Grishina et al., 2005; Huang et al., 2009; Keil et al., 2012). The expression of *Bmp7, Wif1*, and *Wnt5a* appear mainly in clusters F4 and F5, thus identified as peri-epithelial fibroblasts (Fig. 6F-G), which also showed the highest level of *Ar* expression among the fibroblast clusters (Fig. 6F). In co-IF analyses, BMP7, WIF1, and WNT5A proteins were also detected in the epithelial buds and surrounding stromal cells and overlaid with AR (Fig. 6I'-K'). In addition, as observed in the scRNAseq, staining for AR appeared strongest in the UGM adjacent to the UGE while

the lowest expression levels were observed in the prostatic buds (Fig. 6I'-K'). The expression of BMP7, WIF1, and WNT5A along with elevated AR appeared in peri-epithelial fibroblasts as well as in the budding UGE cells, implicating their participation in epithelial-mesenchymal paracrine interactions during prostatic epithelial budding and development.

We further assessed the relationships between different fibroblast clusters during early prostate development using trajectory analysis. Four major branch tips were observed in the trajectory plot generated using Monocle (Fig. 6H). Their differentiation states were predicted using CytoTrace (Fig. S4D-F). Based on this, a group of potential progenitor fibroblasts, clusters F1 and F2, were identified at one of the branch tips, and set as the starting point for pseudotime analysis (Fig. 6H). The peri-epithelial fibroblasts were grouped together at another branch tip (Fig. 6H). A third branch tip featured a group of fibroblasts, primarily from cluster F3, with high expression of *Col14a1*, *Clec3b*, and *Anpep* (Fig. 6F, H), similar to Col14a1 expressing fibroblast populations reported in the lung and skin (Xie et al., 2018; Guerrero-Juarez et al., 2019). The identity of the fourth branch tip is unclear and possesses a mixture of the fibroblast subtypes. The expression of *Bmp7* and *Wif1* were mainly observed in the peri-epithelial fibroblast branch tip, while *Wnt5a, Fgfr2 Hoxd13* and *Gli1* expression appeared in all fibroblast branch tips but with focal expression on some individual tips (Fig. 6L). These trajectory analyses explored the cellular properties and differentiation status of prostatic fibroblasts.

Identifying molecular mechanisms underlying mesenchymal AR mediated regulatory pathways during early prostate development. To assess the regulatory role of the peri-epithelial fibroblasts in early prostate development, we examined the DEGs comparing the transcriptomes of peri-epithelial fibroblasts against the other fibroblasts (Fig. 7A and Table S3). Using Gene Set Enrichment Analysis (GSEA) with different databases, a number of enriched pathways, previously reported to directly associate with branching morphogenesis, were identified (Fig. 7B), including Wnt/β-catenin, Bmp, hedgehog, androgen, and retinoic acid related signaling pathways. The top 40 DEGs from the above analysis displayed elevated expression of peri-epithelial fibroblast markers, *Wif1, Bmp7* and *Wnt5a* along with Hedgehog downstream targets, *Foxf1, Cxcl14*, and *Dner* (Madison et al., 2009; Yu et al., 2009), and Wnt downstream targets, *Lef1* and *Tcf4* (Fig. 7A). Using IPA, a subset of up-regulated AR downstream genes was identified from the above DEGs, suggesting a regulatory role of AR in peri-epithelial fibroblasts (Fig. 7C). Using chromatin immunoprecipitation-quantitative polymerase chain reaction (ChIP-QPCR), we demonstrated an enrichment in the regulatory regions of AR downstream target genes (Nash et al., 2019), including *Fgfr2, Fgf7, Sgk1, Prkca*, and *Ets2* in AR antibody immunoprecipitated DNA samples isolated from E17.5 UGM cells in comparison to input samples (Fig. 7D). The transcriptional regulators enhanced in peri-epithelial fibroblasts were further identified using SCENIC (Aibar et al., 2017) master regulator analysis (Fig. 7E), including 7 Fox and Hox family TFs, five Wnt related TFs, including Lef1 and Tcf4, as well as Sox9 and Sox4. In addition, many of these TFs have been identified as AR associated proteins and co-regulators (http://androgendb.mcgill.ca/ARinteract.pdf; Fig. 7E). Given peri-epithelial fibroblasts'

proximity to the budding UGE cells, we characterized the interactomes between these two clusters to explore the paracrine interactions between the two cell types using SingleCellSignalR methods (Cabello-Aguilar et al., 2020). Predicted ligand-receptor interactions related to Fgf, Tgf-β, Shh, Wnt, and Notch signaling pathways appeared in both directions between peri-epithelial fibroblast and budding UGE cell clusters (Fig. 7F). These data demonstrate the important role of peri-epithelial fibroblasts in inducing prostate epithelial development and provide an atlas for further investigating the role of AR and other regulators in controlling prostate initiation and development.

Deletion of AR in urogenital sinus mesenchymal cells, but not epithelial cells, impairs prostate budding and development. Using mouse genetic tools, we further assessed the role of AR in UGM and UGE cells during early prostate. As illustrated in Fig. 8A and 8B, selective deletion of *Ar* expression and activation of mGFP expression in *Gli1*-expressing cells in the UGM, and in *Trp63* and *Krt8*-expressing cells in the UGE was achieved by administrating TM at E13.5 in $R26R^{mTmG/+}$:$Ar^{L/Y}$:$Gli1^{CreERT2/+}$, $R26R^{mTmG/+}$:$Ar^{L/Y}$:$Trp63^{CreERT/+}$, and $R26R^{mTmG/+}$:$Ar^{L/Y}$:$Krt8^{CreERT2/+}$ mice, respectively (Fig. 8B). Histological and IHC analyses showed normal prostate budding with positive staining of E-cadherin and NKX3.1 in UGS tissues isolated from E18.5 wild type mice (Fig. 8C1-3). IF staining of mGFP in UGS tissues of $R26R^{mTmG/+}$:$Gli1^{CreERT2/+}$, $R26R^{mTmG/+}$:$Trp63^{CreERT/+}$, and $R26R^{mTmG/+}$:$Krt8^{CreERT2/+}$ mice showed *Gli1*-expressing cells in the UGM, and *Trp63* and *Krt8*-expressing cells in different UGE compartments (Fig. 8D1-3), demonstrating their mesenchymal or epithelial properties, respectively. Examining E18.5 UGS tissues of $R26R^{mTmG/+}$:$Ar^{L/Y}$:$Gli1^{CreERT2/+}$ mice revealed no prostatic bud formation (Fig. 8E1-3'), consistent with previous reports (Le et al., 2020).

In contrast, no significant defects appeared in UGS tissues isolated from *Ar*$^{L/Y}$*:R26R*$^{mTmG/+}$*:Trp63*$^{CreERT/+}$, and *Ar*$^{L/Y}$*:R26R*$^{mTmG/+}$*:Krt8*$^{CreERT2/+}$ embryos (Fig. 8H1-1' and 8K1-1'). Positive staining of both E-cadherin and NKX3.1 also appear in the above UGS tissues (Fig. 8H2-3' and 8K2-3'). These data further demonstrate the significant role of mesenchymal AR signaling in prostate budding and development. Positive staining of LEF1, TCF4, HOXD13, and SOX9, which we demonstrated to be enhanced in peri-epithelial fibroblasts (Fig. 7A, C), were observed in UGM or both UGM and UGE cells of *Ar*$^{L/Y}$*:R26R*$^{mTmG/+}$*:Trp63*$^{CreERT/+}$ and *Ar*$^{L/Y}$*:R26R*$^{mTmG/+}$*:Krt8*$^{CreERT2/+}$ tissues (Fig. 8I1-4 and 8L1-4). Both FGF7 and FGFR2 also appear positive in the above UGM cells with VIM co-staining (Fig. 8J1-2 and 8M1-2). However, little to no staining of those proteins was detected in *Ar*$^{L/Y}$*:R26R*$^{mTmG/+}$*:Gli1*$^{CreERT2/+}$ UGS tissues (Fig. 8F1-4 and G1-2). Taken together, these data demonstrate selective deletion of AR in mesenchymal *Gli1*-expressing cells but not UGE cells impairs prostatic budding and diminishes the expression of key transcriptional and paracrine regulators in embryonic UGS tissues, which also provides proof-of-principle evidence supporting the scRNAseq data in this study.

DISCUSSION

Currently, the molecular mechanisms underlying fetal androgen-induced interactions with other signaling pathways in early prostate development remain unclear. Specifically, the identity of prostatic AR-expressing cells that convey androgen signaling and control early prostate induction and development are unknown. In this study, utilizing the novel mouse strain, *R26R*$^{mTmG/+}$*:Ar*$^{IRES-Cre/Y}$, we assessed AR-expressing cell properties and investigated the molecular basis of AR mediated signaling in initiating

prostatic mesenchymal-epithelial interactions during early prostate development. Interestingly, *Cre*-induced mGFP expression through endogenous AR transcription appeared in the E12.5 UGE rather than the UGM. This observation provides the experimental evidence that AR transcription actually occurs first in the UGE although the expression of endogenous AR proteins is first detected in the UGM at E13.5 (Cooke et al., 1991). This observation is also consistent with early *in vivo* hybridization study showing the AR transcription starts at E12.5 UGE and UGM cells (Crocoll et al., 1998). *Cre*-induced mGFP expression appears to mimic the expression pattern of endogenous AR but in a much more sensitive manner, which may be due to the high efficiency of *Cre*-mediated recombination and the different translational processes between the 5' cap and IRES site for synthesizing AR and Cre proteins (Komar and Hatzoglou, 2011). Therefore, the $Ar^{IRES\text{-}Cre}$ model is a valuable tool for tracing AR expressing cells when combined with reporter alleles, and for perturbation and manipulation of androgen signaling through the incorporation of other floxed alleles.

Using the $Ar^{IRES\text{-}Cre}$ model with scRNAseq and other experimental approaches, we assessed the male E17.5 UGS epithelial and mesenchymal cell types. The expression of endogenous *Ar* and *Cre*-initiated mGFP was mainly observed in the UGE and fibroblast cell clusters, suggesting the biological significance of androgen signaling in these cells. We further separated and re-clustered these cell populations individually to visualize them at the highest resolution. Based on highly specific expression of prostatic budding markers, including *Nkx3.1, Wif1* and *Bmp7,* we identified a cluster comprised of budding epithelial cells. In addition, the expression of *Etv4*, a key regulator in budding/branching processes in lung and kidney branching morphogenesis (Lu et al., 2009; Costantini, 2010;

Herriges et al., 2015) and in promoting epithelial-mesenchymal transition (EMT) and prostate cancer progression (Pellecchia et al., 2012; Aytes et al., 2013), was also identified as a novel marker of budding UGE cells. Using trajectory analysis with Monocle and CytoTrace software, we also explored the cellular identity of prostatic epithelial progenitors in E17.5 UGS tissues. A potential progenitor population comprised of *Zeb1*-expressing cells within basal/intermediate UGE clusters was identified, showing opposing pathways of differentiation leading to either budding or superficial UGE clusters (Wang et al., 2020). These results provide new insight into the cellular properties of budding and UGE progenitor cell populations, and provided unbiased characterization of the transcriptomes of these unique epithelial cells, helping to identify the signaling pathways and regulators acting in prostatic epithelial differentiation and development.

Here, we also characterized urogenital stroma to better understand mesenchymal AR signaling in regulating epithelial-mesenchymal paracrine interactions during prostate epithelial development. Since urogenital sinus fibroblasts are directly adjacent to the UGE, we specifically focused on peri-epithelial fibroblasts, which were shown to have the highest AR expression and expression of *Bmp7*, *Wif1*, and *Wnt5a*, regulators of prostatic budding (Grishina et al., 2005; Huang et al., 2009; Abler et al., 2011). Up-regulation of developmental signaling pathways including Wnt/β-catenin, Bmp, retinoic acid, hedgehog, and androgen signaling were identified in these fibroblasts based on the DEGs when compared with other fibroblasts. IPA analyses showed the upregulation of AR downstream target genes in these peri-epithelial fibroblasts. ChIP Q-PCR further demonstrated an enrichment of AR binding in the regulatory regions of AR target genes,

reported in AR ChIPseq data from P0 rat ventral and dorsolateral prostate lobes (Nash et al., 2019), in AR-antibody immunoprecipitated E17.5 UGM samples. Analysis of the interactome between the prostatic bud epithelia and peri-epithelial fibroblasts, using SingleCellSignalR (Cabello-Aguilar et al., 2020), further identified paracrine interactions related to Wnt, Bmp, Notch and Hedgehog signaling pathways. The above data provide the molecular basis for mesenchymal AR in regulating paracrine interactions between mesenchymal and epithelial cells at single cell resolution during prostate early development.

Lastly using mouse genetic tools, we further evaluated the role of epithelial and mesenchymal AR in prostate development. Consistent with previous reports (Le et al., 2020), impairment of prostatic budding appeared in E17.5 UGS tissues with *Gli1*-CreER induced *Ar* deletion but no significant defects showed in counterparts with *Krt8*- or *Trp63*-CreER induced AR deletion, further demonstrating the critical role of mesenchymal AR in prostate development. Interestingly, *Zeb1* expression was observed mainly in mesenchymal *Gli1*-expressing cells although epithelial *Zeb1* expressing cells with co-expression of *Krt8* also showed weak expression of *Gli1* in tSNE plots (Fig. S5A-F). These observations raise the possibility of either *Gli1*-expressing cells as the origin of *Zeb1*-expressing progenitors or Gli1 expression occurring in epithelial cells during embryonic prostate development. Therefore, more effort should be devoted to defining the cellular properties of *Gli1* and *Zeb1* expressing cells as prostatic stem cells and/or their niches cells in prostate development. Taken together, this study provides significant insight into the regulatory role of mesenchymal AR and provides a wealth of

data for future studies of epithelial-mesenchymal interactions in prostatic development and disease.

MATERIALS AND METHODS

Mouse generation, mating and genotyping. All animal procedures were approved by the City of Hope Institutional Animal Care and Use Committee (IACUC). *Rosa26*$^{mTmG/+}$ (*R26*$^{mTmG/+}$) reporter mice were kindly provided by Dr. Liqun Luo (Muzumdar et al., 2007). Experimental mice were generated by intercrossing *Ar*$^{IRES\text{-}Cre/X}$ female mice with *R26*$^{mTmG/+}$ male mice. *Gli1*^{CreER} mice were obtained from Jackson Laboratories (stocks 007913). *Ar*$^{Lox/Y}$ mice were obtained from Dr. Guido Verhoeven (De Gendt et al., 2004). *Trp63*^{CreERT2} and *Krt8*^{CreERT2} mice were generated as described previously (Lee et al., 2014) (Zhang et al., 2012). To elicit genetic recombination, mice were intraperitoneally injected with 125 μg/g body weight of tamoxifen (TM, Sigma) suspended in corn oil (Sigma) as previously reported (Lee et al., 2015). Genotyping was performed using the primers provided (Table S4). The targeting vectors for generating *Ar*$^{IRES\text{-}Cre}$ mice were designed based on the previous report for making *Ar*$^{IRES\text{-}PLAP\text{-}IRES\text{-}nLacZ}$ *mice* (Shah et al., 2004). The IRES-Cre fragment was generated using PCR approaches from the IRES-mCherry plasmid (Addgene #80139) and assembled accordingly. The fragment was introduced to the region after the stop codon of the mouse *Ar* gene. Three guide RNAs (gRNA) were designed at the above insertion site, one gRNA (AAGTGCCCAAGATCCTTTCT) was chosen after *in vitro* cutting efficiency validation. The gRNA was synthesized using *in vitro* transcription approaches, and CAS9 protein was purchased from PNA Bio. The above IRES-cre donor DNA fragment (dDNA) was cloned into the targeting vector and bracketed with left and right homologous arms,

(approximately 1kb each) which were generated by PCR using C57Bl6/J mouse genomic DNA as template with the following pairs of primers: AR-5F: TTCCAGTGGATGGGCTGAAAAATC, AR-5R: ACGCGTTCTTCACTGTGTGTGGAAATAGATGGGCTTGACTTTGCCAGAAAGG ATCTTGGGCACTTG, AR-3F: TTTGGAAACCCTAATACCC, AR-3R: CAAAGAGTCAGACCTTTCC). A point mutation was added by AR-5R primer at gRNA PAM to prevent gRNA cutting donor DNA. C57BL/6J mice (4-6 weeks old) were used for zygote collection. CD1 females were used for injected embryo recipients. Guide RNA, CAS9 protein and donor DNA were mixed before injection into mouse zygotes at a final concentration of 5 μg/μl gRNA, 15 μg/μl CAS9 protein, 3 ng/μl dDNA. Injected zygotes were implanted into CD1 foster mothers. The genomic DNA samples were isolated from the offspring and analyzed by genomic PCR and DNA sequencing. Confirmed *Ar* $^{IRES\text{-}Cre}$ founder mice were then backcrossed with wild type C57BL/6J for two to three generations to delete mosaicism resulted by Crispr/CAS9 based gene-editing approaches in mice.

Histology, immunostaining, mT/mG assays. Mouse tissues were fixed in 10% neutral-buffered formalin (American Master Tech Scientific) and processed in paraffin or processed to OCT following cryo-protection in 30% sucrose in 1X PBS pH 7.3 at 4°C overnight as previously described (Lee et al., 2015). Following embedding in paraffin or OCT tissue blocks were cut to 4 μm and 5 μm serial sections, respectively, and used for hematoxylin-eosin H&E staining for further histological analysis (Lee et al., 2015). Immunohistochemistry (IHC) was performed as previously described (Lee et al., 2015). In brief, IHC was performed following rehydration of tissue sections through a

decreasing ethanol gradient. Heat-induced epitope retrieval was performed using a microwave to boil slides in 0.01M citrate buffer (pH 6.0) followed by 15 min in 0.3% H_2O_2 in methanol. Next, tissue sections underwent blocking for 1 hr at room temperature in 5% goat serum in 1X PBS (pH 7.3) followed by incubation with primary antibodies in 1% goat serum in PBS overnight at 4°C. Tissue sections were then washed with PBS and incubated with Streptavidin ligated to horseradish peroxidase (Strep-HRP) (SA-5004, Vector Laboratories, 1:500 dilution) for 30 min and developed with 3,3'-diaminobenzidine (DAB) kit (SK-4100, Vector Laboratories). Tissue sections were counterstained in 5% Harris Hematoxylin and dehydrated through and increasing ethanol gradient. Coverslips were mounted using Permount Medium (SP15-500, Fisher Scientific).

Immunofluorescence staining was performed using similar procedures as described above for IHC, however, excluding the use of H_2O_2 in methanol. Slides were developed with fluorescent conjugated secondary antibodies and then mounted with coverslips using Vectashield Mounting Medium with DAPI (H-1200, Vector Laboratories). For detecting membrane-bound Tomato (mT) and membrane-bound green fluorescent protein (mGFP) signal, the staining was performed as previously described (Lee et al., 2015). Briefly sections from the OCT embedded tissues were washed with PBS, pH 7.3, and were developed with Vectashield Mounting Medium with DAPI (H-1200, Vector Laboratories). Antibodies used for both IHC and IF are listed (Table S5).

Microscope image acquisition. Images of H&E and immunohistochemistry were acquired on an Axio Lab A1 microscope using 5x, 10x, 20x and 40x Zeiss A-Plan objectives with a Canon EOS 1000D camera and using Axiovision software (Carl Zeiss,

Oberkochen, Germany). Images of immunofluorescent staining and mTmG signals were acquired on a Nikon ECLIPSE E800 Epi-Fluorescence Microscope using 20x and 40x Nikon Plan Fluor objectives with a QImaging RETIGA EXi camera and using QCapture software (QImaging).

Single-Cell RNA Sequencing Analysis. Female mice were placed with males overnight for mating with the following day considered embryonic day E0.5 if a vaginal plug was detected. Two individual sets of sc-RNA-seq experiments were performed using different littermates in this study. The urogenital sinuses of embryos were isolated at E17.5 and placed in DMEM with 10% FBS, 10nM DHT, 5% Nuserum and 25ug/ml insulin. The UGS was separated from the bladder at the bladder neck and distal portions of the urethra and Wolffian ducts were removed (Fig. 2C). Once trimmed, UGS tissues were dissociated to single cells by digestion using 1mg/mg collagenase at 37 °C for 90 min and TrypLE (Gibco) at 37 °C for 15min. Approximately 11,000 viable cells obtained from male *R26R* $^{mTmG/+}$*:Ar*$^{IRES\text{-}Cre}$ embryos (n=4) were used for sequencing. Library preparation was performed using 10x Genomics Chromium Single Cell 3' Solution with v3 chemistry following the manufacturer's protocol (10x Genomics). The library purity and size was validated by capillary electrophoresis using 2,100 Bioanalyzer (Agilent Technologies). The library quantity was measured fluorometrically using Qubit dsDNA HS Assay Kit from Invitrogen. The libraries were sequenced with a NovaSeq 6000 instrument (Illumina) to a depth of ~130k reads per cell. Raw sequencing data were processed using the 10x Genomics' Cell Ranger pipeline (version 3.1.0) to generate FASTQ files and aligned to the mm10 genome with an added mGFP sequence (Zhang et al., 1996) to generate gene expression counts. Following alignment and initial

quality control, a filtered feature bar coded matrix, including a total of 10,478 cells, was uploaded to R (3.6.1) using the Seurat package (3.1.3.9002) (Butler et al., 2018). Cells then underwent further filtering to remove potential empty droplets and doublets (1000 < nFeature_RNA < 9000) as well as low quality cells with high percentages of mitochondrial RNA (percent.mt < 15) (Fig. S1A and 1B). After this final filtering step, 9,440 cells with an average of 4,976 genes per cell and 31,214 UMI counts per cell were used for future analyses. Two individual sets of sc-RNA-seq experiments were performed using different littermates in this study, and one

For the visualization and analysis of scRNAseq results, normalized and scaled data were clustered using the top significant principal components (30) of 3000 highly variable genes with resolution of 0.5 using Seurat (Fig. S1C and 1D). Trajectory analysis and generation of pseudotime was performed by converting Seurat objects into CellDataSet format. Following conversion of the data to a CellDataSet, trajectory analysis was performed using the Monocle2 package (2.12.0) (Qiu et al., 2017) in R. Identification of stem/progenitor populations was supported through the use of CytoTRACE (https://cytotrace.stanford.edu/) (Gulati et al., 2020). Interactome (ligand-receptor) analysis and visualization were performed using the SingleCellSignalR package (0.0.1.6) in R (Cabello-Aguilar et al., 2020). Pathway analysis was performed using Gene Set Enrichment Analysis (GSEA 4.0.3). Lastly, single cell upstream regulator analysis was performed using Ingenuity Pathway Analysis (Qiagen, Version 51963813) and the SCENIC package (1.1.2.2) in R (Aibar et al., 2017).

Chromatin immunoprecipitation assay and qRT-PCR. The urogenital sinuses of embryos were isolated at E17.5. Mouse urogenital sinus mesenchyme (UGM) tissues were collected as reported previously (Xin et al., 2003). Briefly, the UGS was collected in 10 ml DMEM 10% FBS, 10 nM DHT and trimmed rostral and caudal parts of UGS tissues and the wolffian duct. Trypsin was removed carefully, and digestion was stopped with 250 ul of DMEM 20% FBS, 10 nM DHT followed by two washes with same amount of medium. The digested UGS was separated to UGE and UGM by fine needle under the microscope. Mouse UGM (n=6) tissues were minced and incubated with 1% formaldehyde for 15 min and quenched with 0.150 M glycine for 10 min. Samples were washed sequentially with cold PBS, and resuspended in cell lysis buffer (50 mM Tris-HCl (pH 8.0), 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.5% NP-40, and 0.25% Triton X-100), and then homogenized. The chromatin was sheared in nuclear lysis buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.5 mM EGTA, and 0.2% SDS) to an average size of 200–500 bp by sonication, and then diluted threefold in ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8.1, and 167 mM NaCl), and was subjected to immunoprecipitation by magnetic protein G beads (Invitrogen) conjugated with AR (ab74272, Abcam). Cross-links were reversed and chromatin DNA fragments were analyzed by real-time qPCR with specific primers (Table S6).

Statistical Analysis. Differentially expressed gene lists were determined using a Wilcoxon Rank Sum test, with genes showing p-value < 0.05 defined to be significant. As recommended by the GSEA User Guide, pathways with FDR < 0.25 were considered significant in exploratory GSEA pathway analysis.

Data Availability: The single cell mRNA sequencing raw data have been submitted to the GEO repository (GEO accession number GSE153701). You may also review them at: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE153701

Other data are available in the main text or supplementary materials. Source data underlying Figs. 3D, 7B, 7C, 7D and 7F are provided in the source data file.

Funding: This work was supported by the National Institutes of Health (R01CA070297, R01CA166894, and R01DK104941 to Z.S.)

Author Contributions:

Z.S., D.H.L and A.W.O: conceived the project and designed experimental approaches

J.W. and X.W. carried out sequencing procedures and processing.

Z.S. and H.Z. generated new mouse model.

D.H.L., J. W., W.K.K. V.L, A.H., J.M. and J.A. executed experimental procedures.

A.W.O and X.W. performed analysis of sequencing data.

All authors contributed to data analysis.

Z.S., A.W.O, D.H.L., and J.A. contributed to writing the manuscript.

Competing Interests: No competing interests declared

References:

Abler, L. L., Keil, K. P., Mehta, V., Joshi, P. S., Schmitz, C. T. and Vezina, C. M. (2011) 'A high-resolution molecular atlas of the fetal mouse lower urogenital tract', *Dev Dyn* 240(10): 2364-77.

Aibar, S., Gonzalez-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', *Nat Methods* 14(11): 1083-1086.

Aytes, A., Mitrofanova, A., Kinkade, C. W., Lefebvre, C., Lei, M., Phelan, V., LeKaye, H. C., Koutcher, J. A., Cardiff, R. D., Califano, A. et al. (2013) 'ETV4 promotes metastasis in response to activation of PI3-kinase and Ras signaling in a mouse model of advanced prostate cancer', *Proc Natl Acad Sci U S A* 110(37): E3506-15.

Bhatia-Gaur, R., Donjacour, A. A., Sciavolino, P. J., Kim, M., Desai, N., Young, P., Norton, C. R., Gridley, T., Cardiff, R. D., Cunha, G. R. et al. (1999) 'Roles for Nkx3.1 in prostate development and cancer', *Genes Dev* 13(8): 966-77.

Brun, J., Lutz, K. A., Neumayer, K. M., Klein, G., Seeger, T., Uynuk-Ool, T., Worgotter, K., Schmid, S., Kraushaar, U., Guenther, E. et al. (2015) 'Smooth Muscle-Like Cells Generated from Human Mesenchymal Stromal Cells Display Marker Gene Expression and Electrophysiological Competence Comparable to Bladder Smooth Muscle Cells', *PLoS One* 10(12): e0145153.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R. (2018) 'Integrating single-cell transcriptomic data across different conditions, technologies, and species', *Nat Biotechnol* 36(5): 411-420.

Cabello-Aguilar, S., Alame, M., Kon-Sun-Tack, F., Fau, C., Lacroix, M. and Colinge, J. (2020) 'SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics', *Nucleic Acids Res* 48(10): e55.

Chang, C. S., Kokontis, J. and Liao, S. T. (1988) 'Molecular cloning of human and rat complementary DNA encoding androgen receptors', *Science* 240(4850): 324-6.

Chen, C., Zhang, Q., Liu, S., Parajuli, K. R., Qu, Y., Mei, J., Chen, Z., Zhang, H., Khismatullin, D. B. and You, Z. (2015) 'IL-17 and insulin/IGF1 enhance adhesion of prostate cancer cells to vascular endothelial cells through CD44-VCAM-1 interaction', *Prostate* 75(8): 883-95.

Cooke, P. S., Young, P. and Cunha, G. R. (1991) 'Androgen receptor expression in developing male reproductive organs', *Endocrinology* 128(6): 2867-73.

Costantini, F. (2010) 'GDNF/Ret signaling and renal branching morphogenesis: From mesenchymal signals to epithelial cell behaviors', *Organogenesis* 6(4): 252-62.

Crocoll, A., Zhu, C. C., Cato, A. C. and Blum, M. (1998) 'Expression of androgen receptor mRNA during mouse embryogenesis', *Mech Dev* 72(1-2): 175-8.

Cunha, G. R. and Chung, L. W. (1981) 'Stromal-epithelial interactions--I. Induction of prostatic phenotype in urothelium of testicular feminized (Tfm/y) mice', *J Steroid Biochem* 14(12): 1317-24.

Cunha, G. R., Donjacour, A. A., Cooke, P. S., Mee, S., Bigsby, R. M., Higgins, S. J. and Sugimura, Y. (1987) 'The endocrinology and developmental biology of the prostate', *Endocr Rev* 8(3): 338-62.

Cunha, G. R. and Lung, B. (1978) 'The possible influence of temporal factors in androgenic responsiveness of urogenital tissue recombinants from wild-type and androgen-insensitive (Tfm) mice', *J Exp Zool* 205(2): 181-93.

Cunha, G. R., Vezina, C. M., Isaacson, D., Ricke, W. A., Timms, B. G., Cao, M., Franco, O. and Baskin, L. S. (2018) 'Development of the human prostate', *Differentiation* 103: 24-45.

Cunha, G.R. (1984) 'Androgenic effects upon prostatic epithelium are mediated via trophic influences from stroma', *Prog. Clin. Biol. Res.* 145: 81-102.

De Gendt, K., Swinnen, J. V., Saunders, P. T., Schoonjans, L., Dewerchin, M., Devos, A., Tan, K., Atanassova, N., Claessens, F., Lecureuil, C. et al. (2004) 'A Sertoli cell-selective knockout of the androgen receptor causes spermatogenic arrest in meiosis', *Proc Natl Acad Sci U S A* 101(5): 1327-32.

Fontijn, R. D., Favre, J., Naaijkens, B. A., Meinster, E., Paauw, N. J., Ragghoe, S. L., Nauta, T. D., van den Broek, M. A., Weijers, E. M., Niessen, H. W. et al. (2014) 'Adipose tissue-derived stromal cells acquire endothelial-like features upon reprogramming with SOX18', *Stem Cell Res* 13(3 Pt A): 367-78.

Ganassi, M., Badodi, S., Ortuste Quiroga, H. P., Zammit, P. S., Hinits, Y. and Hughes, S. M. (2018) 'Myogenin promotes myocyte fusion to balance fibre number and size', *Nat Commun* 9(1): 4232.

Gelmann, E. P. (2002) 'Molecular biology of the androgen receptor', *J Clin Oncol* 20(13): 3001-15.

Georgas, K. M., Armstrong, J., Keast, J. R., Larkins, C. E., McHugh, K. M., Southard-Smith, E. M., Cohn, M. J., Batourina, E., Dan, H., Schneider, K. et al. (2015) 'An illustrated anatomical ontology of the developing mouse lower urogenital tract', *Development* 142(10): 1893-908.

Goscinski, M. A., Xu, R., Zhou, F., Wang, J., Yang, H., Huang, R., Li, Y., Larsen, S. G., Giercksky, K. E., Nesland, J. M. et al. (2015) 'Nuclear, cytoplasmic, and stromal expression of ZEB1 in squamous and small cell carcinoma of the esophagus', *APMIS* 123(12): 1040-7.

Grishina, I. B., Kim, S. Y., Ferrara, C., Makarenkova, H. P. and Walden, P. D. (2005) 'BMP7 inhibits branching morphogenesis in the prostate gland and interferes with Notch signaling', *Dev Biol* 288(2): 334-47.

Guerrero-Juarez, C. F., Dedhia, P. H., Jin, S., Ruiz-Vega, R., Ma, D., Liu, Y., Yamaga, K., Shestova, O., Gay, D. L., Yang, Z. et al. (2019) 'Single-cell analysis reveals fibroblast heterogeneity and myeloid-derived adipocyte progenitors in murine skin wounds', *Nat Commun* 10(1): 650.

Gulati, G. S., Sikandar, S. S., Wesche, D. J., Manjunath, A., Bharadwaj, A., Berger, M. J., Ilagan, F., Kuo, A. H., Hsieh, R. W., Cai, S. et al. (2020) 'Single-cell transcriptional diversity is a hallmark of developmental potential', *Science* 367(6476): 405-411.

Habuka, M., Fagerberg, L., Hallstrom, B. M., Ponten, F., Yamamoto, T. and Uhlen, M. (2015) 'The Urinary Bladder Transcriptome and Proteome Defined by Transcriptomics and Antibody-Based Profiling', *PLoS One* 10(12): e0145301.

Henry, G. H., Malewska, A., Joseph, D. B., Malladi, V. S., Lee, J., Torrealba, J., Mauck, R. J., Gahan, J. C., Raj, G. V., Roehrborn, C. G. et al. (2018) 'A Cellular Anatomy of the Normal Adult Human Prostate and Prostatic Urethra', *Cell Rep* 25(12): 3530-3542 e5.

Herriges, J. C., Verheyden, J. M., Zhang, Z., Sui, P., Zhang, Y., Anderson, M. J., Swing, D. A., Zhang, Y., Lewandoski, M. and Sun, X. (2015) 'FGF-Regulated ETV Transcription Factors Control FGF-SHH Feedback Loop in Lung Branching', *Dev Cell* 35(3): 322-32.

Huang, L., Pu, Y., Hu, W. Y., Birch, L., Luccio-Camelo, D., Yamaguchi, T. and Prins, G. S. (2009) 'The role of Wnt5a in prostate gland development', *Dev Biol* 328(2): 188-99.

Iber, D. and Menshykau, D. (2013) 'The control of branching morphogenesis', *Open Biol* 3(9): 130088.

Jacob, C., Lotscher, P., Engler, S., Baggiolini, A., Varum Tavares, S., Brugger, V., John, N., Buchmann-Moller, S., Snider, P. L., Conway, S. J. et al. (2014) 'HDAC1 and HDAC2 control the specification of neural crest cells into peripheral glia', *J Neurosci* 34(17): 6112-22.

Javed, S. and Langley, S. E. (2014) 'Importance of HOX genes in normal prostate gland formation, prostate cancer development and its early detection', *BJU Int* 113(4): 535-40.

Jenster, G., van der Korput, H. A., van Vroonhoven, C., van der Kwast, T. H., Trapman, J. and Brinkmann, A. O. (1991) 'Domains of the human androgen receptor involved in steroid binding, transcriptional activation, and subcellular localization', *Mol Endocrinol* 5(10): 1396-404.

Jessen, K. R. and Mirsky, R. (2005) 'The origin and development of glial cells in peripheral nerves', *Nat Rev Neurosci* 6(9): 671-82.

Kanamori-Katayama, M., Kaiho, A., Ishizu, Y., Okamura-Oho, Y., Hino, O., Abe, M., Kishimoto, T., Sekihara, H., Nakamura, Y., Suzuki, H. et al. (2011) 'LRRN4 and UPK3B are markers of primary mesothelial cells', *PLoS One* 6(10): e25391.

Karpus, O. N., Westendorp, B. F., Vermeulen, J. L. M., Meisner, S., Koster, J., Muncan, V., Wildenberg, M. E. and van den Brink, G. R. (2019) 'Colonic CD90+ Crypt Fibroblasts Secrete Semaphorins to Support Epithelial Growth', *Cell Rep* 26(13): 3698-3708 e5.

Keil, K. P., Mehta, V., Branam, A. M., Abler, L. L., Buresh-Stiemke, R. A., Joshi, P. S., Schmitz, C. T., Marker, P. C. and Vezina, C. M. (2012) 'Wnt inhibitory factor 1 (Wif1) is regulated by androgens and enhances androgen-dependent prostate development', *Endocrinology* 153(12): 6091-103.

Komar, A. A. and Hatzoglou, M. (2011) 'Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states', *Cell Cycle* 10(2): 229-40.

Kwon, O. J., Zhang, Y., Li, Y., Wei, X., Zhang, L., Chen, R., Creighton, C. J. and Xin, L. (2019) 'Functional Heterogeneity of Mouse Prostate Stromal Cells Revealed by Single-Cell RNA-Seq', *iScience* 13: 328-338.

Le, V., He, Y., Aldahl, J., Hooker, E., Yu, E. J., Olson, A., Kim, W. K., Lee, D. H., Wong, M., Sheng, R. et al. (2020) 'Loss of androgen signaling in mesenchymal sonic hedgehog responsive cells diminishes prostate development, growth, and regeneration', *PLoS Genet* 16(1): e1008588.

Lee, D. K., Liu, Y., Liao, L., Wang, F. and Xu, J. (2014) 'The prostate basal cell (BC) heterogeneity and the p63-positive BC differentiation spectrum in mice', *Int J Biol Sci* 10(9): 1007-17.

Lee, S. H., Johnson, D. T., Luong, R., Yu, E. J., Cunha, G. R., Nusse, R. and Sun, Z. (2015) 'Wnt/beta-Catenin-Responsive Cells in Prostatic Development and Regeneration', *Stem Cells* 33(11): 3356-67.

Lu, B. C., Cebrian, C., Chi, X., Kuure, S., Kuo, R., Bates, C. M., Arber, S., Hassell, J., MacNeil, L., Hoshi, M. et al. (2009) 'Etv4 and Etv5 are required downstream of GDNF and Ret for kidney branching morphogenesis', *Nat Genet* 41(12): 1295-302.

Madison, B. B., McKenna, L. B., Dolson, D., Epstein, D. J. and Kaestner, K. H. (2009) 'FoxF1 and FoxL1 link hedgehog signaling and the control of epithelial proliferation in the developing stomach and intestine', *J Biol Chem* 284(9): 5936-44.

Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. and Luo, L. (2007) 'A global double-fluorescent Cre reporter mouse', *Genesis* 45(9): 593-605.

Narlis, M., Grote, D., Gaitan, Y., Boualia, S. K. and Bouchard, M. (2007) 'Pax2 and pax8 regulate branching morphogenesis and nephron differentiation in the developing kidney', *J Am Soc Nephrol* 18(4): 1121-9.

Nash, C., Boufaied, N., Badescu, D., Wang, Y. C., Paliouras, M., Trifiro, M., Ragoussis, I. and Thomson, A. A. (2019) 'Genome-wide analysis of androgen receptor binding and transcriptomic analysis in mesenchymal subsets during prostate development', *Dis Model Mech* 12(7).

Ochoa-Espinosa, A. and Affolter, M. (2012) 'Branching morphogenesis: from cells to organs and back', *Cold Spring Harb Perspect Biol* 4(10).

Pellecchia, A., Pescucci, C., De Lorenzo, E., Luceri, C., Passaro, N., Sica, M., Notaro, R. and De Angioletti, M. (2012) 'Overexpression of ETV4 is oncogenic in prostate cells through promotion of both cell proliferation and epithelial to mesenchymal transition', *Oncogenesis* 1: e20.

Podlasek, C. A., Barnett, D. H., Clemens, J. Q., Bak, P. M. and Bushman, W. (1999) 'Prostate development requires Sonic hedgehog expressed by the urogenital sinus epithelium', *Dev Biol* 209(1): 28-39.

Prins, G. S. and Putz, O. (2008) 'Molecular signaling pathways that regulate prostate gland development', *Differentiation* 76(6): 641-59.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A. and Trapnell, C. (2017) 'Reversed graph embedding resolves complex single-cell trajectories', *Nat Methods* 14(10): 979-982.

Shah, N. M., Pisapia, D. J., Maniatis, S., Mendelsohn, M. M., Nemes, A. and Axel, R. (2004) 'Visualizing sexual dimorphism in the brain', *Neuron* 43(3): 313-9.

Spaderna, S., Schmalhofer, O., Hlubek, F., Berx, G., Eger, A., Merkel, S., Jung, A., Kirchner, T. and Brabletz, T. (2006) 'A transient, EMT-linked loss of basement membranes indicates metastasis and poor survival in colorectal cancer', *Gastroenterology* 131(3): 830-40.

Staack, A., Donjacour, A. A., Brody, J., Cunha, G. R. and Carroll, P. (2003) 'Mouse urogenital development: a practical approach', *Differentiation* 71(7): 402-13.

Takeda, H. and Chang, C. (1991) 'Immunohistochemical and in-situ hybridization analysis of androgen receptor expression during the development of the mouse prostate gland', *J Endocrinol* 129(1): 83-9.

Thomson, A. A., Timms, B. G., Barton, L., Cunha, G. R. and Grace, O. C. (2002) 'The role of smooth muscle in regulating prostatic induction', *Development* 129(8): 1905-12.

Toivanen, R., Mohan, A. and Shen, M. M. (2016) 'Basal Progenitors Contribute to Repair of the Prostate Epithelium Following Induced Luminal Anoikis', *Stem Cell Reports* 6(5): 660-667.

Varner, V. D. and Nelson, C. M. (2014) 'Cellular and physical mechanisms of branching morphogenesis', *Development* 141(14): 2750-9.

Wang, X., Xu, H., Cheng, C., Ji, Z., Zhao, H., Sheng, Y., Li, X., Wang, J., Shu, Y., He, Y. et al. (2020) 'Identification of a Zeb1 expressing basal stem cell subpopulation in the prostate', *Nat Commun* 11(1): 706.

Xie, T., Wang, Y., Deng, N., Huang, G., Taghavifar, F., Geng, Y., Liu, N., Kulur, V., Yao, C., Chen, P. et al. (2018) 'Single-Cell Deconvolution of Fibroblast Heterogeneity in Mouse Pulmonary Fibrosis', *Cell Rep* 22(13): 3625-3640.

Xin, L., Ide, H., Kim, Y., Dubey, P. and Witte, O. N. (2003) 'In vivo regeneration of murine prostate from dissociated cell populations of postnatal epithelia and urogenital sinus mesenchyme', *Proc Natl Acad Sci U S A* 100 Suppl 1: 11896-903.

Yu, M., Gipp, J., Yoon, J. W., Iannaccone, P., Walterhouse, D. and Bushman, W. (2009) 'Sonic hedgehog-responsive genes in the fetal prostate', *J Biol Chem* 284(9): 5620-9.

Zhang, G., Gurtu, V. and Kain, S. R. (1996) 'An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells', *Biochem Biophys Res Commun* 227(3): 707-11.

Zhang, L., Zhang, B., Han, S. J., Shore, A. N., Rosen, J. M., Demayo, F. J. and Xin, L. (2012) 'Targeting CreER(T2) expression to keratin 8-expressing murine simple epithelia using bacterial artificial chromosome transgenesis', *Transgenic Res* 21(5): 1117-23.

Zhang, Y., Yokoyama, S., Herriges, J. C., Zhang, Z., Young, R. E., Verheyden, J. M. and Sun, X. (2016) 'E3 ubiquitin ligase RFWD2 controls lung branching through protein-level regulation of ETV transcription factors', *Proc Natl Acad Sci U S A* 113(27): 7557-62.

Figures

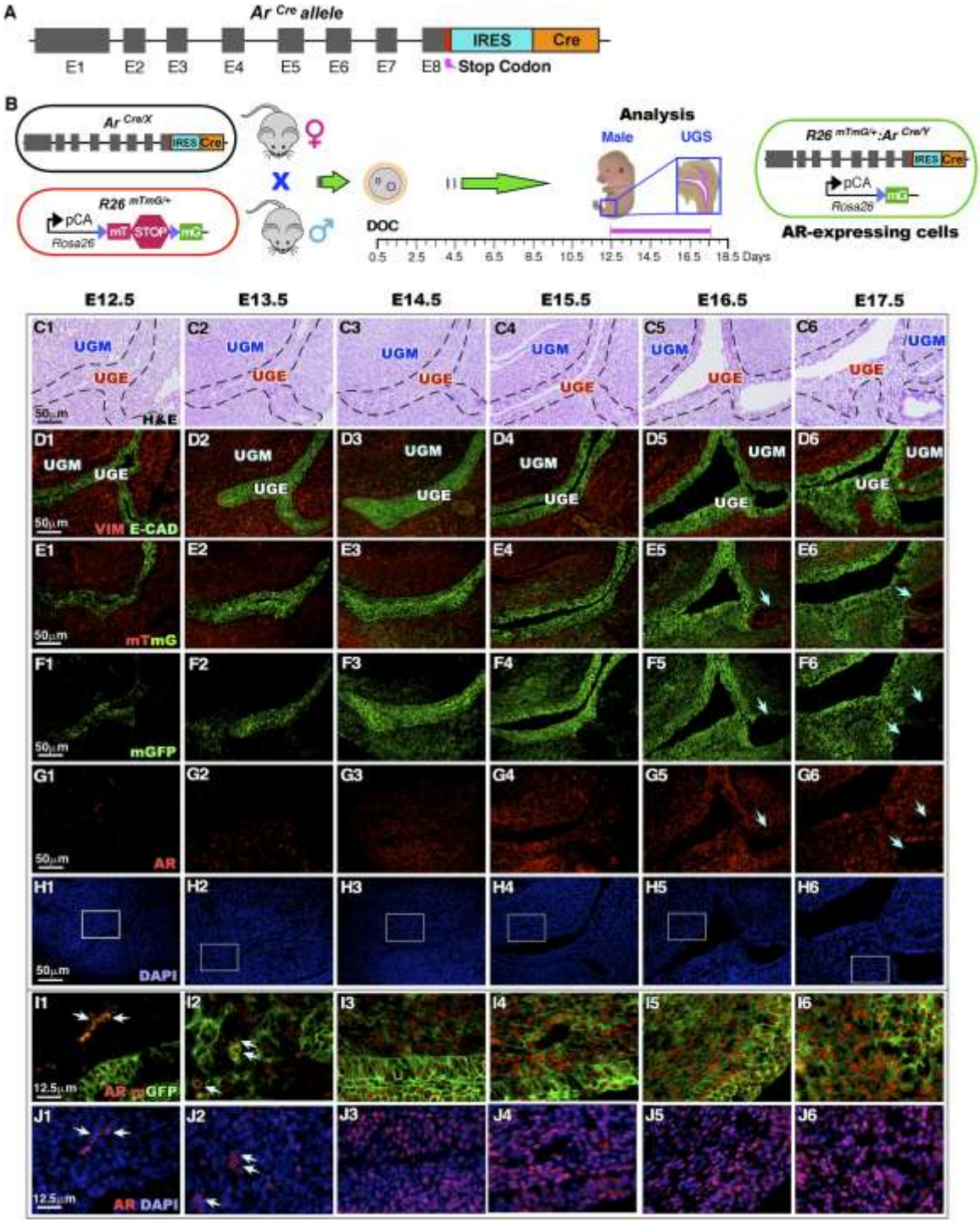

Fig. 1. Generation and characterization of *Ar*$^{IRES\text{-}Cre}$ mice. (A) Genetic construct of the targeted *Ar* allele displaying the inserted IRES and Cre sequences. (B) Schematics of the *Ar*$^{IRES\text{-}Cre}$ and *R26*$^{mTmG/+}$ alleles are shown in relation to the mating strategy for this experiment. Following the day of conception (DOC), a timeline is provided indicating the days of analysis as shown. A construct is displayed demonstrating the recombination event that will take place in *Ar* expressing cells resulting in a change from red to green fluorescence. (C**)** A total of four UGS samples at each time point were isolated and analyzed in this and following figures. Representative H&E images are displayed with dashed lines separating urogenital sinus epithelium and mesenchyme at the indicated time points. (D-J) Representative fluorescence imaging for the indicated proteins/antibodies. Blue arrows in images indicate mT positive Wolffian duct structures (E5-6, F5-6 and G5-6). White arrows indicate cytoplasmic staining from AR outside of the nucleus (I1-2, J1-2). All scale bars are sized as indicated in the images.

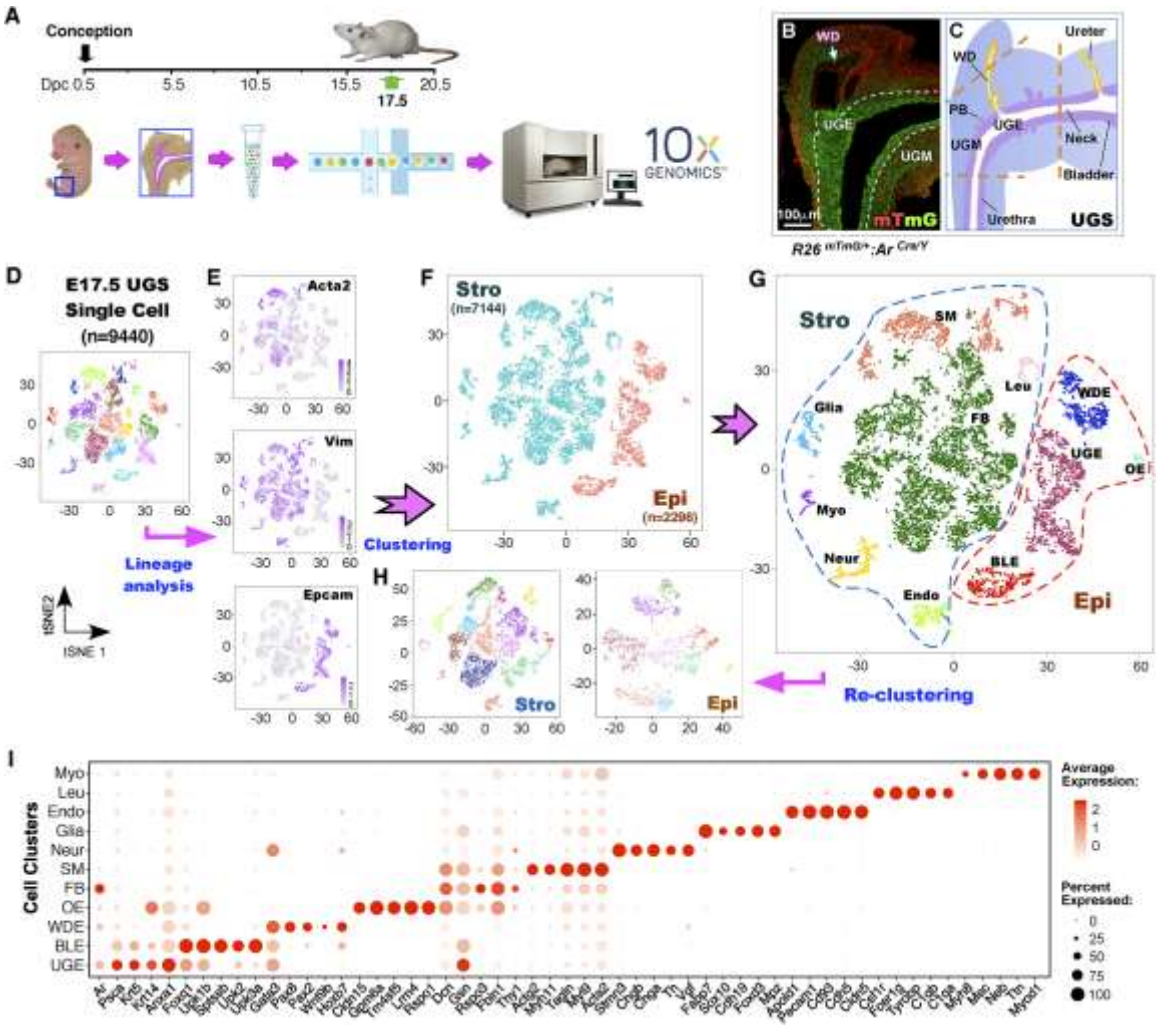

Fig. 2. Single-cell RNA sequencing of E17.5 male mouse UGS. (A) The timeline of the single cell sequencing experiment performed. (B) A representative mTmG assay image of a cross-section of a *R26*$^{mTmG/+}$:*Ar*$^{IRES\text{-}Cre}$ male UGS. The mT positive Wolffian duct epithelium is marked with a blue arrow. (C) An illustration depicting a male E17.5 UGS indicates, with dashed orange lines, roughly how the samples were trimmed prior to single cell sequencing (PB, Prostatic bud). (D) tSNE plot of single cells isolated from four male E17.5 urogenital sinuses. (E) Gene expression tSNE plots for the indicated epithelial and stromal cell marker genes. (F) Original clustering results grouped into either epithelial or stromal (non-epithelial) groups. (G) Identification of cell types as

indicated within the original clustering results. (H) Two tSNE plots show re-clustering results after separating the stromal and epithelial cell groups. (I) A dot plot of *Ar* as well as 5 highly specific genes for each of the cell types identified. All scale bars are sized as indicated.

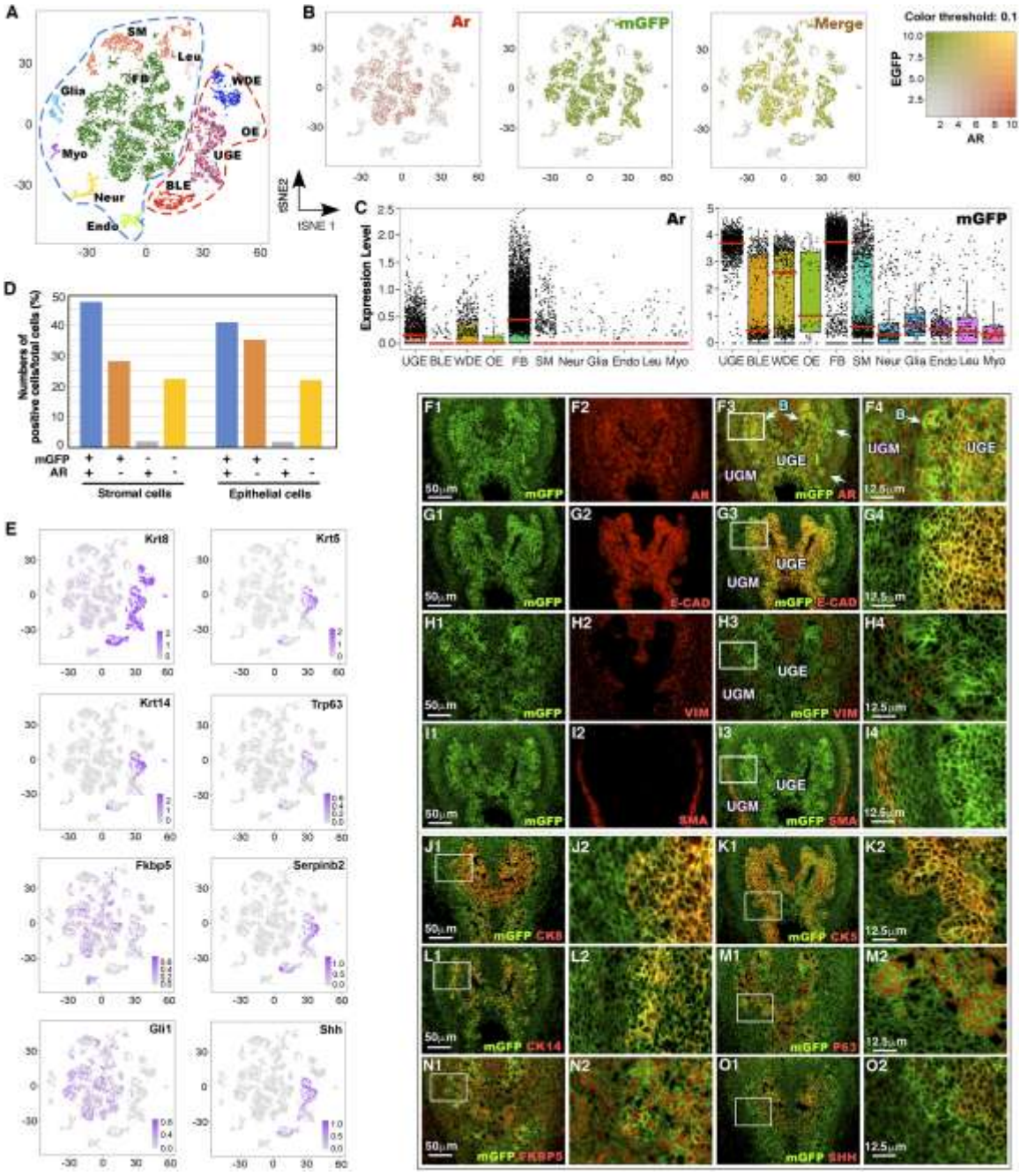

Fig. 3. Characterization of *Ar* expressing cells using scRNAseq. (A) tSNE plot of overall UGS tissue. (B) Blended tSNE expression plots displaying cell clusters with expression of *Ar* and *mGFP*. (C) Box plots representing scaled expression data for *Ar* and *mGFP*. Red lines mark the median expression value. (D) Bar chart corresponding to the percentage of stromal and epithelial cells with *Ar* and *mGFP* expression as indicated.

(E) Gene expression tSNE plots displaying expression patterns of the indicated genes. (F-O) Representative Co-IF images of male E17.5 UGS tissue stained for the indicated antibodies. Blue arrows indicate prostatic buds (F3-4). All scale bars are sized as indicated.

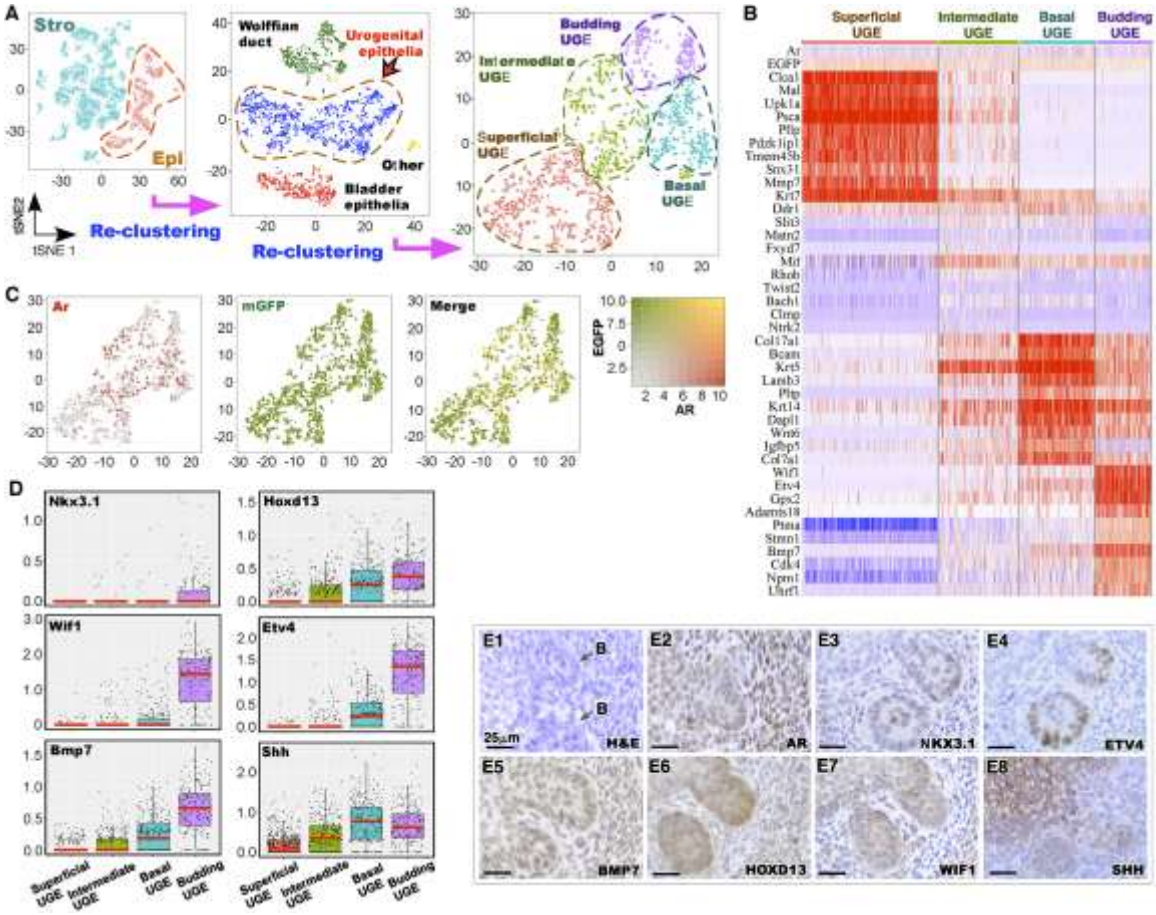

Fig. 4. Single cell RNA sequencing analysis of the urogenital sinus epithelium. (A) tSNE plots showing the separation and re-clustering of UGE cells. (B) Heatmap of the top 10 differentially expressed genes for each of the four UGE clusters. (C) Blended expression tSNE plots of *Ar* and *mGFP* expression in the UGE. (D) Box plots depicting scaled gene expression of the indicated genes. (E) Representative H&E and IHC images of prostatic buds stained for the indicated proteins. Prostatic buds labeled with arrows and letter "B" in E1. All scale bars are sized as indicated.

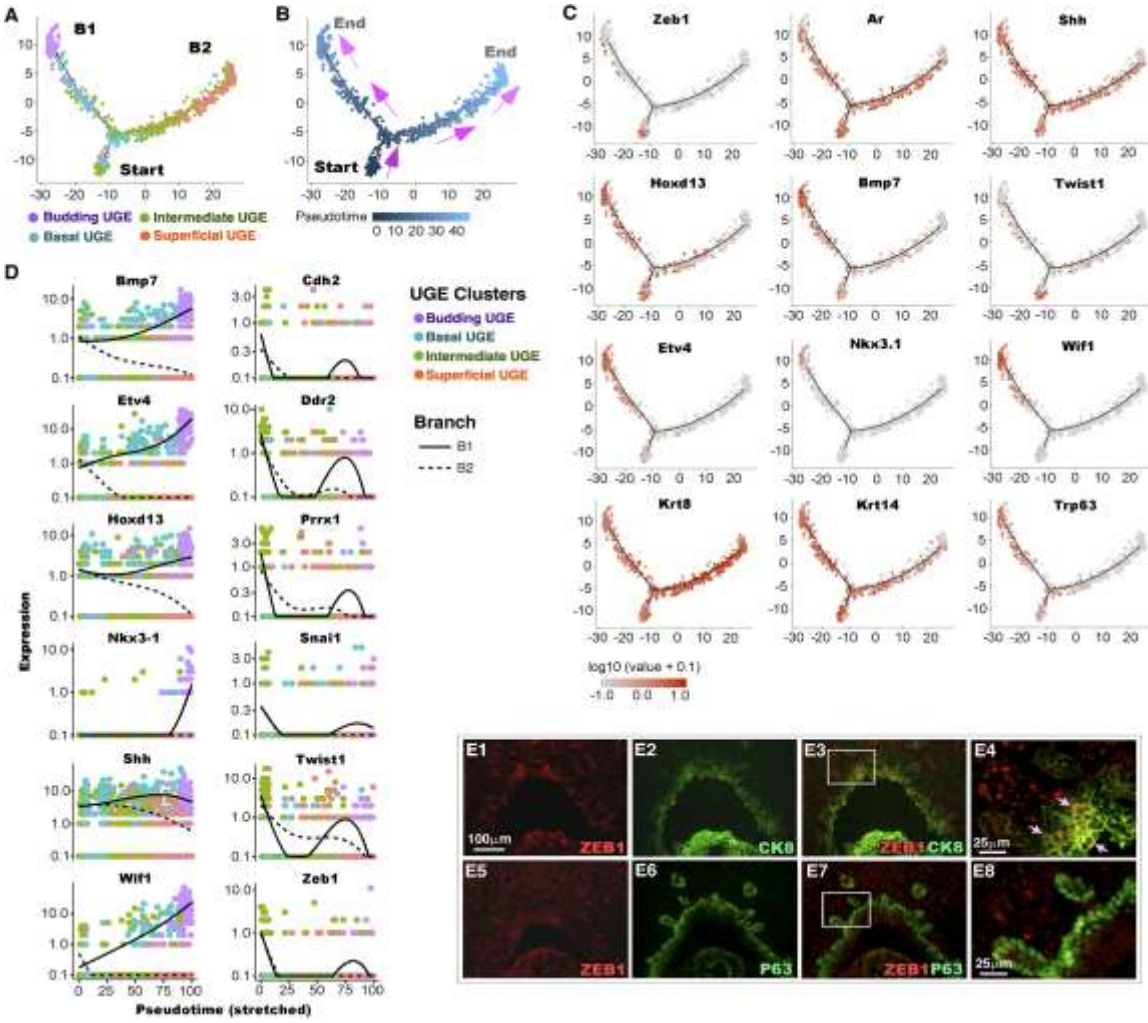

Fig. 5. Trajectory analysis of the urogenital sinus epithelium. (A) Trajectory analysis of UGE single cells. (B) Pseudotime plot displays a predicted directional path of differentiation between cell types as indicated. (C) Expression of the indicated genes in the trajectory plot. (D) Linear pseudotime expression plots for the indicated budding and EMT marker genes. Lines on each plot correspond to the path of differentiation moving from the start point (left) to the indicated branch tip (right). (E) Representative images of immunofluorescent staining for the indicated antibodies in E17.5 male UGS tissues. Pink arrows (E4) indicate cytoplasmic epithelial *Zeb1* staining. All scale bars are sized as indicated.

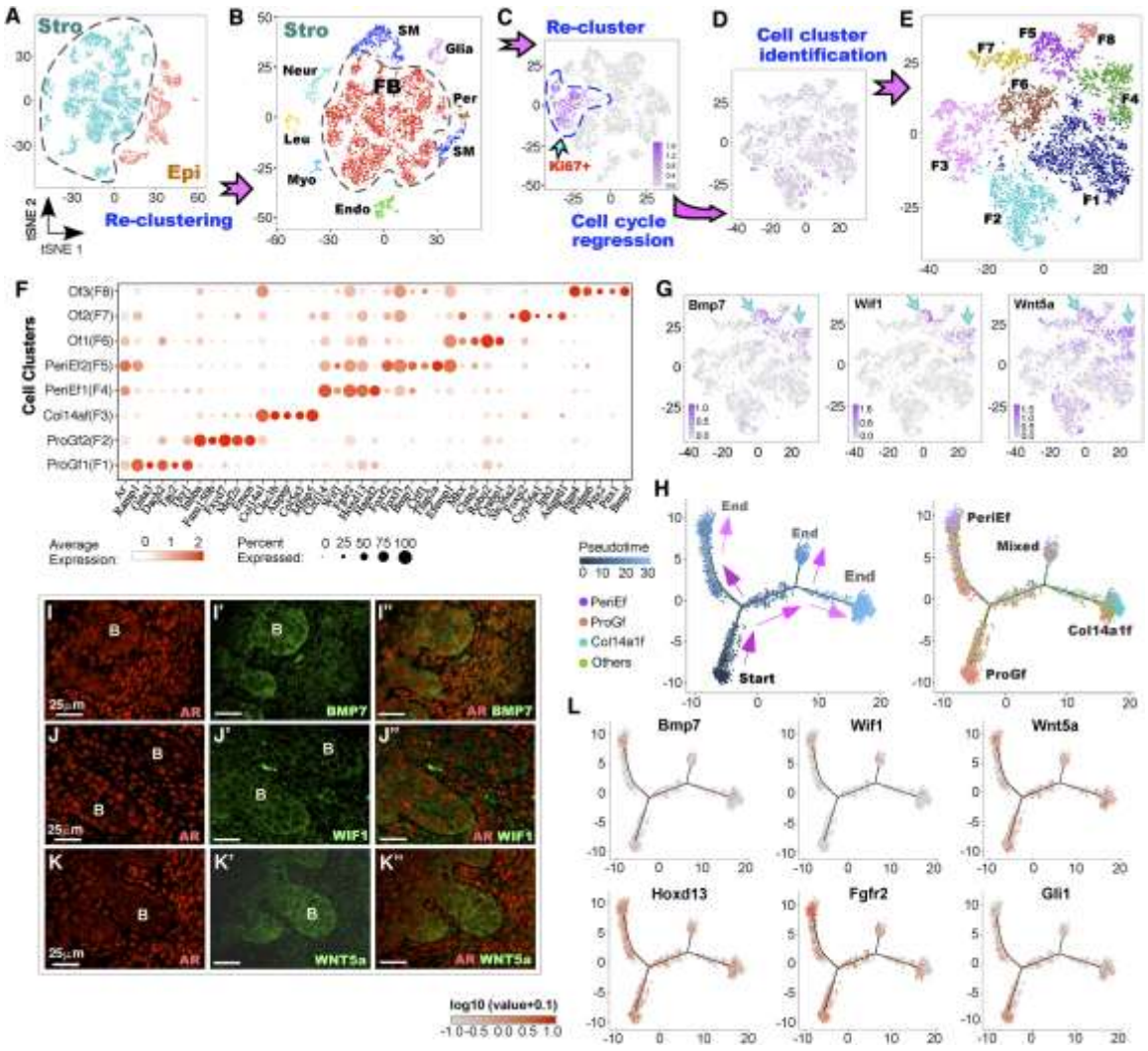

Fig. 6. Single cell RNA sequencing analysis of the urogenital sinus mesenchyme. (A) tSNE plot showing the stromal cell clusters selected for re-clustering. (B) Stromal re-clustering tSNE plot divided by cell type as indicated, showing selection of fibroblasts for re-clustering. (C) tSNE plot identifying *Mki67* expression following initial re-clustering of fibroblasts. (D) tSNE plot of *Mki67* expression, following cell cycle regression. (E) tSNE plot of final fibroblast clustering results used for analysis. (F) A dot plot of *Ar* as well as 5 genes highly specific to each fibroblast cluster. (G) Expression tSNE plots of the indicated peri-epithelial fibroblast markers. (I-K) Representative co-IF images of

male E17.5 UGS sections stained for the indicated antibodies. (H) Trajectory analysis of fibroblast cells. Pseudotime indicates a predicted pathway of differentiation between the fibroblast subtypes as indicated. (L) Expression plots of the indicated genes in pseudotime. All scale bars are sized as indicated.

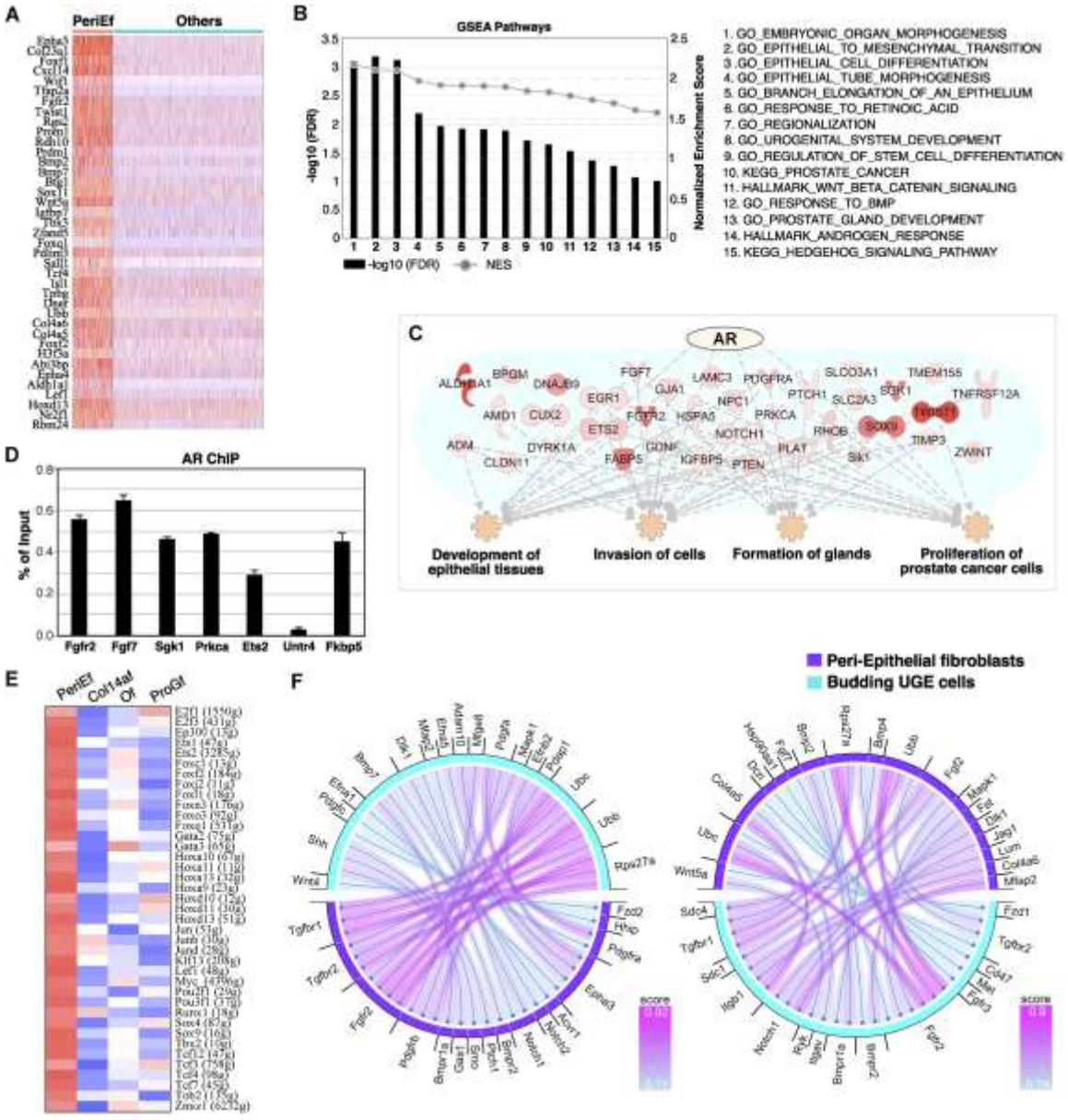

Fig. 7. Identification of potential mechanisms related to peri-epithelial AR signaling. (A) Heatmap of the top 40 differentially expressed genes specific to the peri-epithelial fibroblast clusters. (B) GSEA pathway analysis results comparing the peri-epithelial fibroblast clusters to the remaining fibroblasts. (C) AR downstream targets identified using IPA's upstream regulator analysis as being upregulated in peri-epithelial fibroblasts. Enriched pathways related to the AR downstream targets are indicated. (D) ChIP-qPCR

analyses were conducted using E17.5 UGM cells. Immunoprecipitated chromatin DNA fragments with AR antibody were analyzed by real-time qPCR with specific primers for the AR downstream target genes as labeled in the figure. (E) Transcription factors with enhanced activity in peri-epithelial fibroblasts relative to the remaining fibroblasts generated using SCENIC upstream regulator analysis. (F) Predicted ligand-receptor interactions between budding UGE cells and peri-epithelial fibroblasts as indicated, generated using SingleCellSignalR. Color scale corresponds to interaction scores for each interaction.

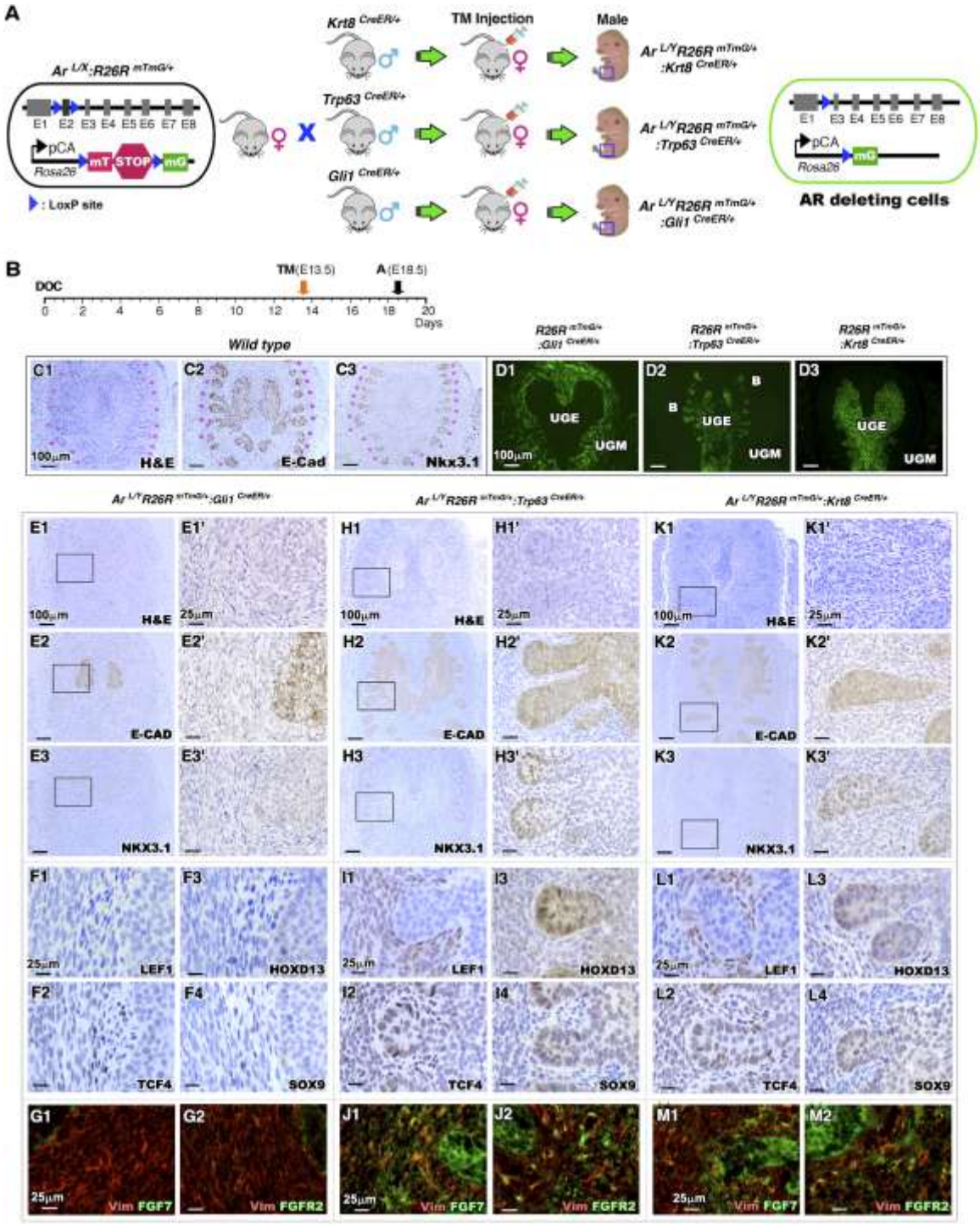

Fig. 8. Effects of selective *Ar* deletion on prostatic budding and expression of key developmental genes. (A) Schematic identifying gene constructs and mating strategies to yield *Cre* promoter specific *Ar* deletion in the indicated models. (B) Timeline for

tamoxifen injection and analysis of tissue samples following day of conception (DOC). (C) Representative H&E and IHC staining images of wild-type E18.5 male urogenital sinuses. (D) Representative fluorescence images of mGFP in male E18.5 urogenital sinus tissues from the indicated genotypes. (E-G) Representative images of the indicated staining of male E18.5 *R26R*$^{mTmG/+}$*:Ar*$^{L/Y}$*:Gli1*$^{CreERT2/+}$ UGS tissue. (H-J) Representative images of the indicated staining of male E18.5 *R26R*$^{mTmG/+}$*:Ar*$^{L/Y}$*:Trp63*$^{CreERT/+}$ UGS tissue. (K-M) Representative images of the indicated staining of male E18.5 *R26R*$^{mTmG/+}$*:Ar*$^{L/Y}$*:Krt8*$^{CreERT2/+}$ UGS tissue. All scale bars are sized as indicated.

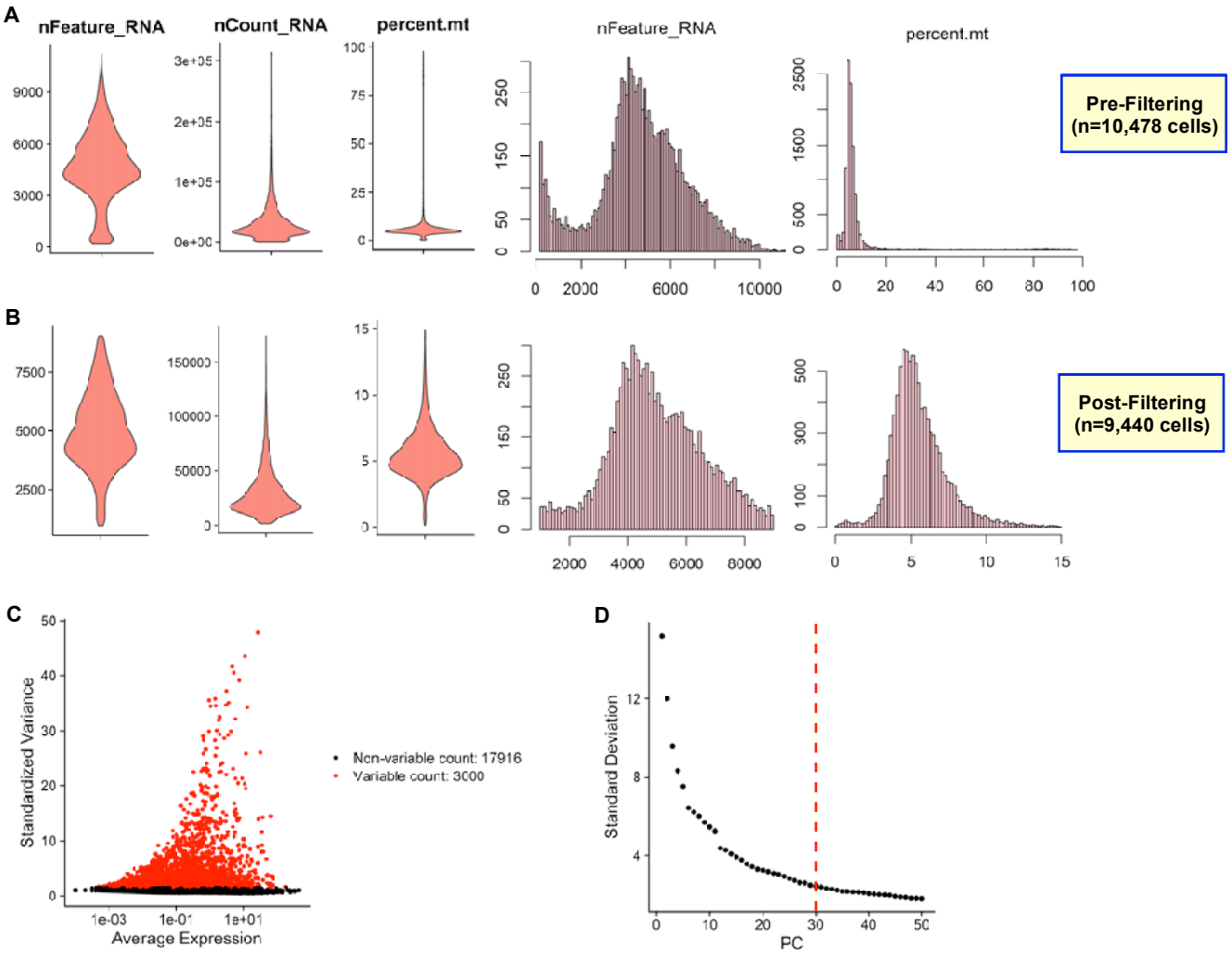

Figure S1. Filtering of single-cell RNA sequencing results in Seurat. **A** Violin and histogram plots corresponding to the number of features, counts and the percent of mitochondrial RNA in each cell pre-filtering. **B** Violin and histogram plots corresponding to the number of features, counts and the percent of mitochondrial RNA in each cell after filtering for 1000 < nFeature < 9000 and percent.mt < 15. **C** Plot of gene variability and selection of the top 3000 variable genes. **D** PCA elbow plot identifying a cutoff of 30 dims used for further analysis.

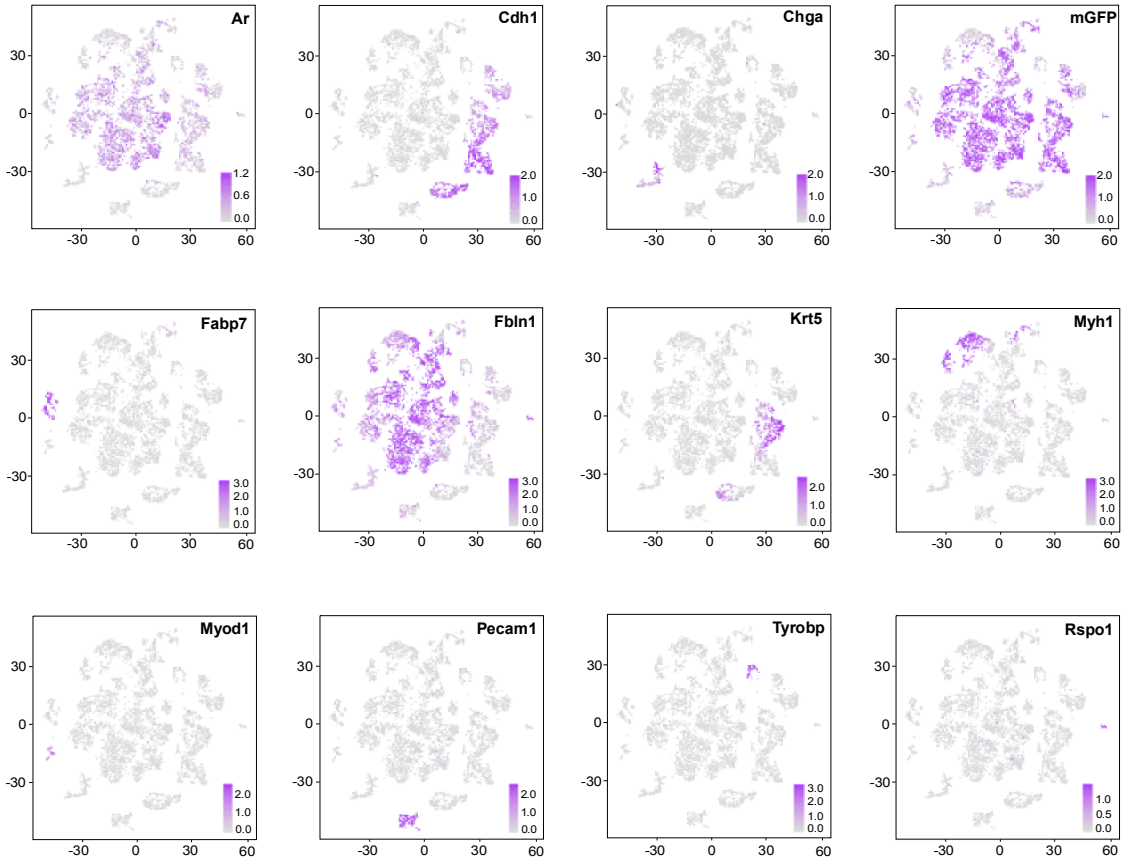

Figure S2. Expression profiling of the whole E17.5 male UGS. Different expression tSNE plots for the indicated genes from the original clustering of whole UGS tissue were shown above.

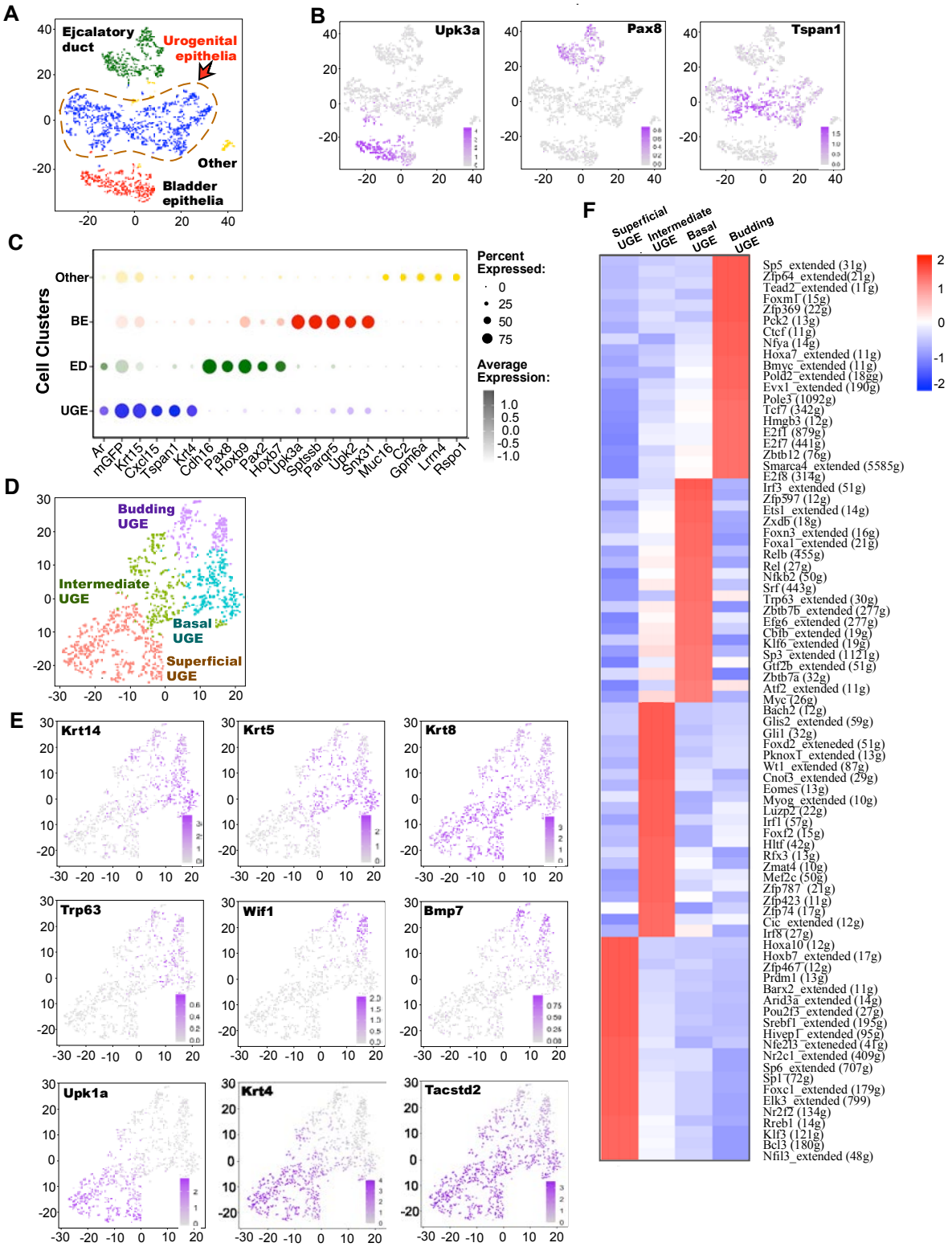

Figure S3. Cell identification and additional analysis of epithelial cells detected in scRNAseq. **A** tSNE plot of the indicated epithelial cells types. Arrow and dashed lines indicate the selection of the UGE for further analysis. **B** Expression tSNE plots of bladder, Wolffian duct and UGE markers (Upk3a, Pax8 and Tspan1 respectively). **C** Dot plot displaying Ar and mGFP expression as well as 5 highly specific genes expressed in each cell type. **D** tSNE plot of UGE re-clustering results. **E** Expression tSNE plots of the indicated genes in the UGE. **F** Heatmap of the top 20 enhanced transcription factors specific to each UGE subtype, generated using SCENIC.

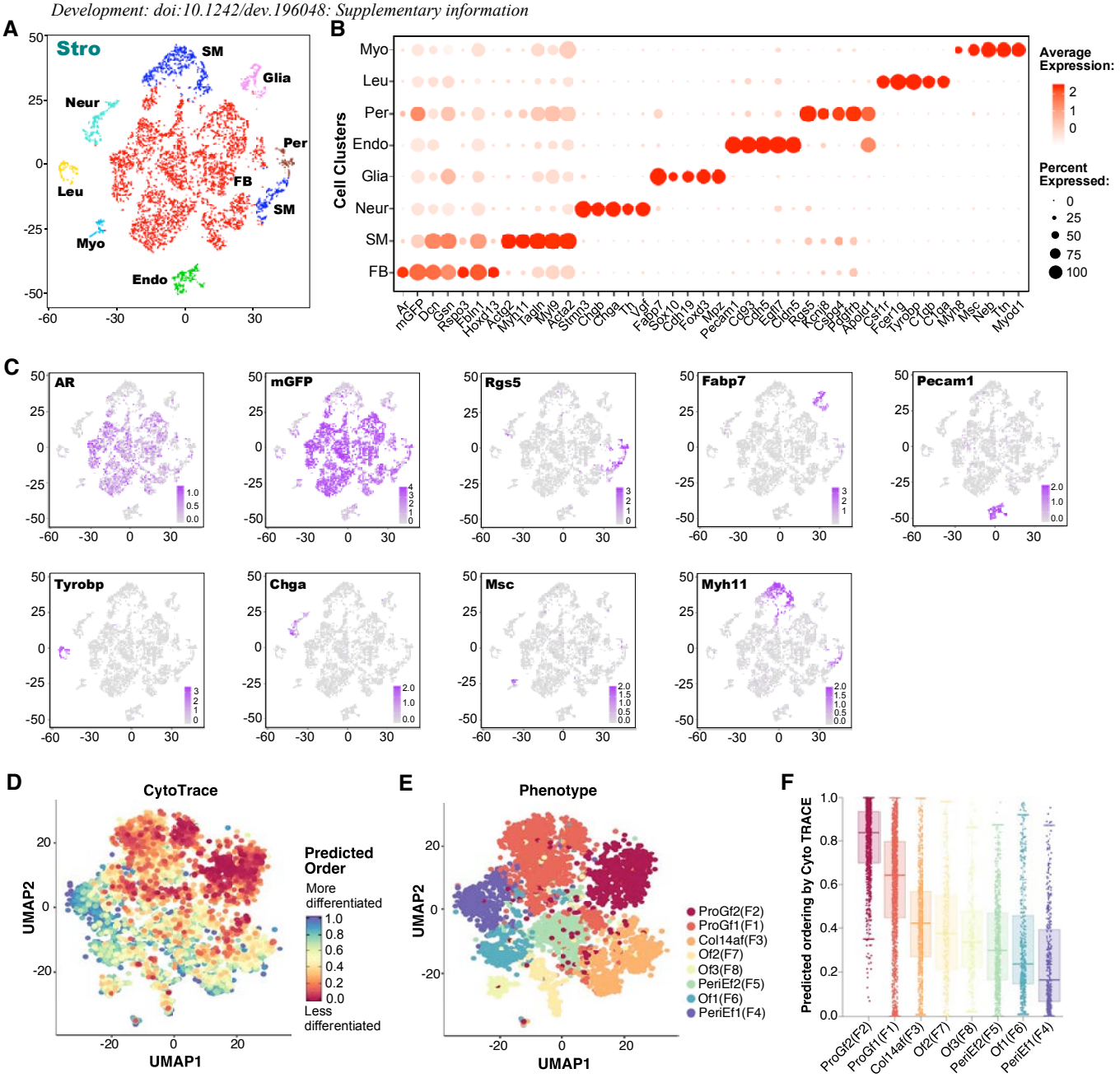

Figure S4. Stromal cell reculstering and assessment of fibroblast differentiation. **A** tSNE plot of stromal cell re-clustering grouped by cell type as indicated. **B** Dot plot displaying Ar and mGFP expression as well as 5 highly specific genes expressed in each cell type. **C** Expression tSNE plots of the indicated genes. **D-E** tSNE plots of the fibroblast clusters generated using CytoTrace, with predicted differentiation states displayed in the left image. **F** Box plots displaying the predicted ordering of cell differentiation for each cluster generated using CytoTrace.

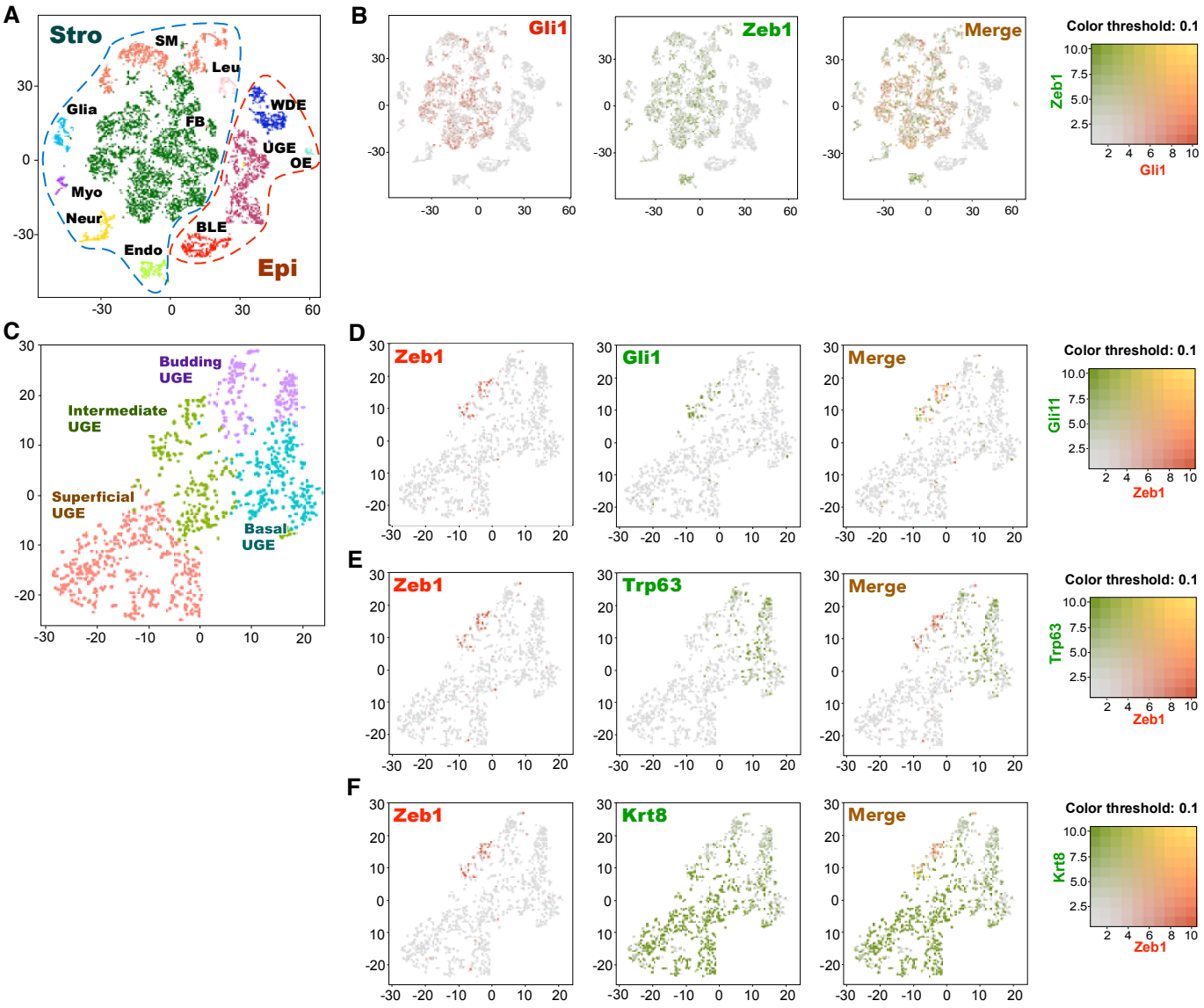

Figure S5. Characterized expression of Zeb1 in the UGS. **A** tSNE plot of the original clustering of the whole UGS grouped by cell type. **B** Blended expression plots of Gli1 and Zeb1 in the UGS. **C** tSNE plot of UGE clustering results grouped by UGE subtypes. **D-F** Blended expression plots of Zeb1 and the indicated genes in the UGE.

Table S1. DEG List: Budding UGE population vs remaining UGE cells

Click here to Download Table S1

Table S2: DEG List: Zeb1+ UGE population vs remaining UGE cells

Click here to Download Table S2

Table S3. DEG List: PeriE Fibroblast cluster vs remaining fibroblasts

Click here to Download Table S3