RUNX1 marks a luminal castration resistant lineage established at the onset of prostate development

AUTHORS

Renaud Mevel[1], Ivana Steiner[2], Susan Mason[3], Laura Galbraith[3], Rahima Patel[1], Muhammad ZH Fadlullah[1], Imran Ahmad[3,4], Hing Y. Leung[3,4], Pedro Oliveira[5], Karen Blyth[3,4], Esther Baena[2,6], Georges Lacaud[1,7]

AFFILIATIONS

[1] Cancer Research UK, Stem Cell Biology Group, Cancer Research UK Manchester Institute, The University of Manchester, Alderley Park, Alderley Edge, Macclesfield, SK10 4TG, UK.

[2] Cancer Research UK, Prostate Oncobiology Group, Cancer Research UK Manchester Institute, The University of Manchester, Alderley Park, Alderley Edge, Macclesfield, SK10 4TG, UK.

[3] Cancer Research UK Beatson Institute, Bearsden, Glasgow, G61 1BD, UK.

[4] Institute of Cancer Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Bearsden, Glasgow, G61 1QH, UK.

[5] Department of Pathology, The Christie NHS Foundation Trust, Manchester, UK.

[6] Belfast-Manchester Movember Centre of Excellence, Cancer Research UK Manchester Institute, The University of Manchester, Alderley Park SK10 4TG, UK.

[7] Correspondence: georges.lacaud@manchester.ac.uk

Phone: +44 (0) 161 446 6058, Fax: +44 (0) 161 446 3109

ABSTRACT

The development of castration-resistant prostate cancer is a major complication of androgen-deprivation therapy. However, the origin and identity of these resistant cells remains controversial. Here, we report that the transcription factor RUNX1 marks a specific subpopulation of proximal luminal cells (PLCs), enriched in the periurethral region of the developing and adult mouse prostate, and distinct from the previously identified NKX3.1^{+} luminal castration resistant cells. Using scRNA-seq profiling and genetic lineage tracing, we show that RUNX1^{+} PLCs are unaffected by androgen deprivation, and do not contribute to the regeneration of the distal luminal compartments. Furthermore, we demonstrate that a transcriptionally similar RUNX1^{+} population emerges at the onset of embryonic prostate specification to populate the proximal region of the ducts. Collectively, our results reveal that RUNX1^{+} PLCs is an intrinsic castration-resistant and self-sustained lineage that emerges early during prostate development and provide new insights into the lineage relationships of the prostate epithelium.

INTRODUCTION

Prostate is a glandular organ of the mammalian male reproductive system and Prostate Cancer (PCa) is the most frequently diagnosed cancer in men. One limitation in its treatment is the difficulty to accurately discriminate between indolent and more aggressive forms of the disease. Therefore, a greater understanding of PCa, as well as normal prostate biology would be transformative in improving current diagnostic and treatment modalities.

In mice, prostate development starts during embryogenesis at embryonic day (E) 15.5-16.5 with the emergence of the first prostatic buds from the rostral end of the urogenital sinus (UGS) (Bhatia-Gaur et al., 1999; Georgas et al., 2015; Keil et al., 2012; Toivanen & Shen, 2017). These initial buds grow into the surrounding mesenchyme to develop postnatally and through puberty into a branched ductal network organized in distinct pairs of lobes, known as the anterior prostate (AP), dorsolateral prostate (DLP) and ventral prostate (VP) (Sugimura et al., 1986a). Each lobe has distinct branching patterns, histopathological characteristics, and is thought to contribute differently to the physiological function of the prostate. The differentiated epithelium of the adult prostate gland is mainly composed of basal and luminal cells, interspersed with rare neuroendocrine cells (Shen & Abate-Shen, 2010; Toivanen & Shen, 2017; Y. Wang et al., 2001). Luminal cells form a layer of polarized tall columnar cells that depend on androgen signaling and produce the prostatic secretions. Basal cells act as a supportive layer located between the luminal cells and the surrounding stroma.

In PCa patients, androgen deprivation therapy (ADT) is frequently used to target androgen-dependent tumor cells. However, the emergence of castration-resistant PCa cells, proposed to either originate from intrinsically castration-resistant subpopulations or to be

acquired *de novo*, is a major limitation of this treatment (Shen & Abate-Shen, 2010; Zhang et al., 2018). Therefore, understanding the origin and identity of these resistant cells both in the normal prostate and PCa represents an essential step towards improving PCa treatments. In mice, surgical castration-induced prostate involution has proven an invaluable tool to identify progenitor castration-resistant cell populations, characterized by their ability to survive in the absence of androgens, and to fully regenerate an intact adult prostate after re-administration of testosterone (Barros-Silva et al., 2018; Kwon et al., 2016; McAuley et al., 2019; Tsujimura et al., 2002; B. Wang et al., 2015; X. Wang et al., 2009; Yoo et al., 2016). Indeed, despite being mostly quiescent under homeostatic conditions, the prostate gland encompasses incredible plasticity. Such plasticity has also been shown in defined experimental conditions to stimulate regenerative properties of epithelial subpopulations, including transplantations (Barros-Silva et al., 2018; Burger et al., 2005; Lawson et al., 2007; Lukacs et al., 2010; Richardson et al., 2004; X. Wang et al., 2009; Xin et al., 2005; Yoo et al., 2016), injury repair (Horton et al., 2019; Kwon et al., 2014; Toivanen et al., 2016), and organoid assays (Chua et al., 2014; Höfner et al., 2015; Karthaus et al., 2014). In addition, several studies have proposed that progenitor populations with distinct physiological roles and regenerative capacity reside at different locations within the prostate (Burger et al., 2005; Crowell et al., 2019; Goldstein et al., 2008; Goto et al., 2006; Kwon et al., 2016; Leong et al., 2008; McNeal, 1981; Tsujimura et al., 2002). However, the precise cellular hierarchy and how it is established during development remains controversial.

RUNX transcription factors (TF) are master regulators of lineage commitment and cell fate (Mevel et al., 2019). In particular, RUNX1 is essential for the ontogeny of the hematopoietic system and alterations of RUNX1 have been associated with a broad spectrum of hematological

malignancies. Interestingly, increasing evidence implicates RUNX1 in the biology and pathology of hormone-associated epithelia (Lie-A-Ling et al., 2020; Riggio & Blyth, 2017; Scheitz & Tumbar, 2013), including breast (Browne et al., 2015; Chimge et al., 2016; Ferrari et al., 2014; van Bragt et al., 2014), uterine (Planagumà et al., 2004, 2006), ovarian (Keita et al., 2013), and prostate cancers (Banach-Petrosky et al., 2007; Scheitz et al., 2012; Takayama et al., 2015). Despite the documented importance of RUNX TFs and reports of RUNX1 in PCa, its expression in the normal prostate gland during development, homeostasis, and regeneration has not been explored.

In this study, we found that *Runx1* marks a discrete subset of luminal cells located in the proximal region of the prostatic ducts. Using mouse models, combined with lobe-specific single-cell transcriptomic profiling of adult, castrated, and developing prostates, we show that RUNX1^{+} proximal luminal cells represent a distinct lineage established at the onset of prostate development, displaying intrinsic castration-resistant and self-sustaining properties.

RESULTS

RUNX1 marks a subpopulation of prostate proximal luminal cells (PLCs)

We initially sought to characterize the expression pattern of *Runx1* in adult mouse prostate. While RUNX1 was detected in basal cells at multiple spatial locations, its expression was specifically high in a subset of luminal cells found in the proximal region of all three lobes, also known as periurethral (Figures 1A, B; Figure 1–figure supplement 1A, B). Sections were co-stained with NKX3.1, a master regulator of prostate development broadly expressed in luminal cells. Using quantitative image-based cytometry (QBIC), we found that RUNX1 and NKX3.1 had a largely mutually exclusive expression pattern, with a sharp transition from RUNX1^{+} NKX3.1^{-} to RUNX1^{-} NKX3.1^{+} cells in the proximal region (Figures 1A, B; Figure 1–figure supplement 1A, B). These proximal luminal cells had a unique histological profile, with a compact organization, intense nuclear hematoxylin staining, and increased nuclear-to-cytoplasmic ratio (Figure 1–figure supplement 1C). In contrast, distal luminal cells had a large cytoplasm with intense pink eosin staining, likely reflecting their secretory function. These observations suggest that RUNX1 marks a subset of proximal luminal cells, distinct from the abundant NKX3.1^{+} luminal population lining the rest of the prostate epithelium.

The proximal site of the prostate has been proposed to be enriched in cells with stem/progenitor properties (Goldstein et al., 2008; Kwon et al., 2016; Tsujimura et al., 2002; Yoo et al., 2016). In order to study the regenerative potential of *Runx1* expressing cells *ex vivo*, we took advantage of isoform-specific fluorescent reporter mouse models of *Runx1* (Draper et al., 2018; Sroczynska et al., 2009). *Runx1* expression is controlled by two promoters, P1 and P2, that respectively drive the expression of the *Runx1c* and the *Runx1b* isoform (Mevel et al.,

2019). We found that *Runx1* expression in the prostate was exclusively mediated by the proximal P2 promoter, in up to 30% of all epithelial EPCAM$^+$ prostate cells (Figures 2A-C). Flow-cytometry profiling confirmed the enrichment of P2-*Runx1*:RFP in both basal (EPCAM$^+$ CD49f^{high}) and luminal (EPCAM$^+$ CD24^{high}) lineages of the proximal compared to the distal prostate (Figures 1C, D; Figure 1–figure supplement 2D). Mirroring our QBIC spatial analysis (Figure 1–figure supplement 1B), P2-*Runx1*:RFP was also detected in a large fraction of the VP epithelium (Figure 1–figure supplement 2D).

We therefore used the P2-*Runx1*:RFP mouse line to isolate *Runx1* positive (RFP$^+$) and negative (RFP$^-$) epithelial cells from the basal and luminal compartments of all three prostate lobes and evaluated their regenerative potential in organoid culture assays (Drost et al., 2016) (Figure 1E). The proximal and distal regions of the AP were analysed separately. Proximal RFP$^+$ luminal cells of the AP consistently displayed higher Organoid Formation Capacity (OFC) than the RFP$^-$ fraction (Figure 1F). Luminal RFP$^+$ sorted cells of the DLP and VP also had a greater OFC than RFP$^-$ cells (Figure 1–figure supplement 3A). In contrast, no significant differences in OFC were observed between basal enriched subsets and distal luminal RFP$^+$ and RFP$^-$ cells. Brightfield assessment revealed that virtually all organoids had a ‘solid’ aspect, except for the predominantly ‘hollow’ organoids derived from proximal RFP$^+$ luminal cells (Figure 1–figure supplement 3B). To further characterize their lineage potential, we classified organoids into three types based on the expression of specific lineage markers: unipotent “basal-like” Keratin 5$^+$ (K5$^+$), unipotent “luminal-like” Keratin 8$^+$ (K8$^+$), or multipotent K5$^+$ K8$^+$ (Figures 1G, H; Figure 1–figure supplement 3C). Interestingly, AP proximal luminal RFP$^+$ derived organoids were predominantly small unipotent K8$^+$, while the remainder fraction mainly gave larger unipotent

137 $K5^+$ organoids (Figures 1H; Figure 1–figure supplement 3D). Few multipotent $K5^+$ $K8^+$ organoids
138 were also identified in nearly all populations.

139 Together, our results show that RUNX1 marks a specific subset of proximal luminal cells
140 (PLCs), and that its expression in the prostate is mediated by the P2 promoter. $RUNX1^+$ PLCs
141 have a particular predisposition to form unipotent $K8^+$ hollow organoids, suggesting a lineage
142 bias towards the luminal identity, and highlighting differences within the luminal compartment
143 of proximal and distal regions.

144

145 ***Runx1* expressing cells are enriched in the castrated prostate epithelium**

146 In mice, androgen-deprivation can be modelled by surgical castration which leads to
147 prostate regression and enriches for castration-resistant cells (Toivanen & Shen, 2017; Zhang et
148 al., 2018). This process is accompanied by the death of luminal androgen-dependent cells and a
149 small proportion of basal cells (English et al., 1987; Sugimura et al., 1986b). To track changes in
150 *Runx1* expression following androgen withdrawal, we surgically castrated P2-*Runx1*:RFP mice
151 and harvested tissue ≥ 4 weeks post-surgery (Figure 2A). While intact prostates contained 22.8
152 ± 6.0% RFP^+ epithelial cells, their frequency increased to 87 ± 6.0% following castration (Figures
153 2B, C). High RUNX1 levels were no longer restricted to the proximal region, and RFP was
154 detected in virtually all basal cells of the AP, DLP and VP, as well as more than 75% of the
155 luminal castration-resistant cells (Figures 2D; Figure 2–figure supplement 4A). RUNX1
156 expressing cells often co-expressed TROP2 (Figure 2E), known to be widely expressed in
157 castrated prostate epithelium (Goldstein et al., 2008; X.-D. Wang et al., 2007). Several
158 castration-resistant luminal populations have been identified in mice (Barros-Silva et al., 2018;

Kwon et al., 2016; McAuley et al., 2019; Tsujimura et al., 2002; B. Wang et al., 2015; X. Wang et al., 2009; Yoo et al., 2016), including rare castration-resistant *Nkx3-1* expressing cells (CARNs). Accordingly, we observed low, but detectable, levels of NKX3.1 in some luminal cells, but only occasional RUNX1$^+$ NKX3.1$^+$ luminal cells in the distal regions of the castrated prostate (Figures 2D; Figure 2–figure supplement 4B, C). Importantly, the clear transition from RUNX1$^+$ to NKX3.1$^+$ cells identified in the proximal luminal layer of intact mice was conserved after castration (Figure 2D, ii).

Together, these results show that RUNX1 is expressed in the majority of the castration-resistant cells. The RUNX1$^+$ NKX3.1$^-$ subset identified in the proximal luminal epithelium of the intact prostate remain NKX3.1$^-$ following castration, supporting the notion that RUNX1$^+$ PLCs constitute a distinct lineage from distal NKX3.1$^+$ cells.

scRNA-seq profiling of *Runx1*$^+$ and *Runx1*$^-$ cells in individual lobes of the intact and castrated prostate

To further characterize the RUNX1$^+$ and RUNX1$^-$ fractions residing at different anatomical locations of the prostate, we performed droplet-based single cell (sc)RNA-seq. We sorted EPCAM$^+$ RFP$^+$ and RFP$^-$ cells from individually dissected lobes of intact and castrated prostates isolated from P2-*Runx1*:RFP reporter mice (Figure 3A). Sorted populations were multiplexed using MULTI-seq lipid-tagged indices to minimize technical confounders such as doublets and batch effects (McGinnis, Patterson, et al., 2019). We retrieved a total of 3,825 prostate epithelial cells from all sorted populations, with a median of 2,846 genes per cell (see Methods; Figure 3–figure supplement 5, 6A-G). We identified 9 *in silico* computed clusters

expressing canonical epithelial, basal, and luminal markers (Figure 3–figure supplement 6H-J). A large population of basal cells was annotated by merging 3 tightly connected subclusters broadly expressing *Krt5*, *Krt14,* and *Trp63* (Figures 3B-D; Figure 3–figure supplement 6E, J). Luminal populations expressed surprisingly heterogeneous levels of canonical luminal markers such as *Cd26*/*Dpp4, Cd24a*, *Krt8, and Krt18* (Figure 3–figure supplement 6I). We annotated those distinct clusters as Luminal-A (Lum-A), Lum-B, Lum-C, Lum-D, Lum-E and Lum-F (Figure 3B). Differential gene expression analysis revealed genes strongly associated with each luminal subpopulation (Figures 3C, 3D; Figure 3–figure supplement 7A; Supplementary file 2).

Initially, we sought to evaluate the effect of androgen withdrawal on lobe-specific cellular heterogeneity. Lum-A/B/C/D were largely enriched in luminal cells originating from intact prostates, whereas Lum-E/F contained mainly castrated luminal cells (Figures 3E; Figure 3–figure supplement 7B). However, a small fraction of intact cells was also identified in Lum-E/F, while some castrated cells were found in Lum-A/B/C/D clusters. Interestingly, Lum-A/C/F mainly contained VP cells, while Lum-B/D/E had a majority of AP and DLP cells, indicating that the lobular identity of luminal cells in the intact prostate is conserved following castration (Figures 3F; Figure 3–figure supplement 7C). This suggests that a subset of intact Lum-A/C undergoes partial reprogramming during castration-induced regression and gives rise to the Lum-F cluster. Similarly, surviving Lum-B/D may predominantly reprogram into Lum-E cells upon castration. In contrast, castrated basal cells were minimally affected by androgen-deprivation and clustered together with intact basal cells (Figure 3E). Overall, our dataset highlights the broad transcriptional reprogramming occurring predominantly in luminal cells upon androgen deprivation.

***Runx1* expressing luminal cells are transcriptionally similar to castration-resistant cells**

We next specifically focused our attention on RUNX1$^+$ luminal cells. The Lum-D cluster predominantly consisted of AP derived RFP$^+$ cells, as well as a small number of RFP$^+$ DLP and VP cells (Figure 3F, H; Figure 3–figure supplements 6E, 7B, 7C). High *Runx1* expression in Lum-D correlated with higher levels of *Tacstd2*/*Trop2*, Ly6 family members as well as *Runx2* (Figures 3D, G; Figure 3–figure supplements 7D, E). In contrast, *Runx1* was barely detected in clusters Lum-B/C which expressed high levels of *Nkx3-1* while Lum-A cells expressed low levels of both *Runx1* and *Nkx3-1*. These results suggest that the Lum-D cluster corresponds to the distinct RUNX1$^+$ luminal cells identified in the proximal region of all three prostate lobes (Figure 1).

To further characterize the specificities of those populations, we performed gene ontology analysis. In line with the secretory role of distal luminal cells, clusters Lum-A/B/C were enriched in enzymatic activity and protein synthesis functions. In contrast, the Lum-D cluster was enriched in terms related to epithelial developmental processes, similar to Lum-E/F (Figure 3–figure supplements 8A-I). This was supported by partition-based graph abstraction (Wolf et al., 2019), which uncovered a strong degree of connectivity between the mainly intact Lum-D and castrated Lum-E population (Figure 3B). Additionally, the Lum-D cluster contained a small, but defined, subpopulation of castrated epithelial cells, suggesting the preservation of its identity upon androgen deprivation (Figures 3E, F). In this population, we found very few genes significantly differentially expressed between intact and castrated cells (n=103; Supplementary file 3). As expected, androgen-regulated genes including *Psca* and *Tspan1* were downregulated in the castrated subset, while strong contributors of the Lum-D identity such as *Tacstd2*/*Trop2*,

Krt4 and *Runx1* did not vary (Figure 3–figure supplement 8J). These observations further support the hypothesis that Lum-D/RUNX1^{+} PLCs maintain their identity following androgen-deprivation.

Overall, our single-cell transcriptomic analysis highlighted a vast degree of heterogeneity within and between the luminal compartments of both intact and castrated mouse prostates. The tight transcriptional relationship observed between high *Runx1* expressing clusters Lum-D and Lum-E/F suggest that the Lum-D population, which corresponds to PLCs, may contain intrinsically castration-resistant luminal cells

Lineage-tracing of *Runx1* expressing cells establishes the intrinsic castration resistant properties of the proximal luminal lineage

To determine if RUNX1^{+} PLCs were enriched in castration-resistant cells, we combined prostate regression-regeneration assays with genetic lineage-tracing using *Runx1*$^{\text{mER-CRE-mER}}$*Rosa*$^{\text{Flox-stop-Flox-RFP}}$ mice (Luche et al., 2007; Samokhvalov et al., 2007), henceforth *Runx1*$^{\text{CRE-stop-RFP}}$ (Figure 4A). Using this model, we could genetically label an average of 4.70 ± 2.8% prostate epithelial *Runx1* expressing cells with RFP upon tamoxifen injection (Figures 4B, C; Figure 4–figure supplement 9A). This corresponded to 0.54 ± 0.2‰ of the total epithelium (Figure 4E). Consistent with the expression pattern of *Runx1*, the majority of labelled cells were located in the proximal region of the prostate (Figure 4C), and co-expressed Keratin 4 (K4) (Figure 4–figure supplements 9D, E), previously found enriched in Lum-D cells (Figure 3D).

Following surgical castration, we found that the absolute number of RFP^{+} marked cells remained stable (Figure 4–figure supplements 9C, D). However, the frequency of RFP^{+} cells in

247 the epithelial compartment increased by ~4.3 fold (Figures 4E; Figure 4–figure supplement 9B)
248 indicating that *Runx1* expressing cells have an enhanced capacity to survive castration
249 compared to *Runx1* negative cells. Next, we investigated whether these intrinsically castration-
250 resistant *Runx1* expressing cells were involved in epithelial regeneration upon testosterone
251 addback (Figure 4B, bottom). Surprisingly, only 0.71 ± 0.2‰ RFP^+ epithelial cells were found in
252 the regenerated prostate, which was comparable to the intact state (Figures 4E; Figure 4–figure
253 supplements 9B-D). Although the majority of RFP^+ clones consisted of single cells, we did
254 observe a minor ~2-fold increase in the frequency of larger clones (2-4 cells) after regeneration,
255 highlighting a modest contribution of RFP labeled cells during prostate regeneration (Figures 4F,
256 G). We found that most RFP marked cells were luminal $K8^+$ in intact, castrated, and regenerated
257 prostates (Figures 4F, H), with only a few basal $K5^+$ RFP^+ cells detected in distal areas (Figure
258 4F). Strikingly, more than 90% of all RFP^+ cells remained negative for NKX3.1 in all experimental
259 arms (Figure 4I).

260 Thus, these results indicate that RFP^+ cells, including PLCs, are mostly unaffected by
261 fluctuations in androgen levels during regression-regeneration assays. RUNX1 expression marks
262 intrinsically castration-resistant luminal cells that do not contribute to the expansion of luminal
263 $NKX3.1^+$ cells during prostate regeneration.

264

265 ***Runx1* marks proximal cells during prostate development**

266 Given the singular identity of proximal luminal *Runx1* expressing cells in the adult
267 prostate, we then asked if this luminal lineage was already emerging during prostate
268 development. At E18.5, once the first prostate buds have emerged, RUNX1 was mainly found in

the $K8^{high}$ inner layers of the stratified urogenital epithelium (UGE) (Figure 5A). Interestingly, these cells also co-expressed K4 (Figure 5–figure supplement 10A), previously found in the Lum D population (Figure 3D), as well as LY6D, recently shown to mark a subset of adult luminal progenitors (Barros-Silva et al., 2018) (Figure 5–figure supplement 10B). In contrast, RUNX1 expression was low in $p63^+$ and $K5^+$ cells, either lining the outer UGE or found in the tips of premature $NKX3.1^+$ prostate buds (Figures 5A-C). At postnatal day 14 (P14), a prepubescent stage when most of the initial branching events have already occurred (Sugimura et al., 1986a; Tika et al., 2019), RUNX1 was broadly expressed in the proximal region (Figure 5D), mainly in $K4^+$ luminal cells and in some $K5^+$ or $p63^+$ cells (Figure 5–figure supplement 10C-E). Conversely, $NKX3.1^+$ cells were found in distal locations, largely distinct from $RUNX1^+$ cells. The specific spatial expression pattern of RUNX1 in proximal luminal cells, largely mutually exclusive with NKX3.1, suggests that these two transcription factors already mark distinct cellular lineages during embryonic prostate organogenesis.

To study the dynamic emergence of $RUNX1^+$ cells during prostate development, we utilized an explant culture system (Berman et al., 2004; Doles et al., 2005; Kruithof-de Julio et al., 2013; Lopes et al., 1996). Dissected E15.5 UGS were cultured for up to 7 days in the presence of dihydrotestosterone (Figures 5E, F). Bud formation was initiated within 2 days of culture (Figure 5G) and composed of a double positive $K5^+$ $K8^+$ stratified epithelium, partially diversifying by day 7 (Figure 5–figure supplement 11A, B). On day 0 (E15.5), RUNX1 was detected at the rostral end of the UGE, particularly within the inner layers of the stratified epithelium. After 1 day in culture, NKX3.1 expression emerged in $RUNX1^+$ cells located in the outer layers of the UGE, while defined budding was yet to be observed. On day 2, $NKX3.1^+$

prostate buds were evident and had reduced or absent RUNX1 expression. This pattern was conserved in the mature explant, in which distal tips were mainly NKX3.1^{+}, whereas the proximal area remained RUNX1^{+} (Figures 5G, H), and co-expressed LY6D and K4 (Figure 5–figure supplements 10C, D). Cellular proliferation marked by Ki67 was more substantial in distal regions, suggesting that most of the expansion did not take place in the RUNX1^{+} compartment (Figure 5–figure supplement 10E).

These results suggest that prostate budding originates from a subset of cells located in the outer layers of the stratified UGE, transiently marked by RUNX1 and NKX3.1. During embryonic prostate development, *Runx1* expression is already primarily confined to the proximal region of the prostatic ducts, in a distinct compartment from NKX3.1^{+} cells.

scRNA-seq of explant cultures reveals the specification of the proximal luminal lineage during embryonic prostate development

The characterization by immunostainings of continuous developmental processes is generally constrained to a small number of markers at a time. To further study the specification of RUNX1 and NKX3.1 lineages, we performed scRNA-seq on UGS explant cultures collected at successive time points: E15.5 (D0), day 1 (D1), day 3 (D3), and day 6 (D6) (Figure 6A). After data processing, 3,937 developing prostatic cells were retained, with a median of 3,608 genes per cell (see Methods; Figure 6–figure supplement 12).

Visualization of the dataset using a force-directed layout highlighted the progressive cellular diversification taking place from D0 to D6 (Figure 6B). Cellular populations were divided into 9 clusters, annotated C0 to C8 (Figures 6C-E). C0/C1 contained the majority of D0 and D1

313 derived cells, while C2-C8 emerged and expanded at later time points. Due to the primitive
314 nature of the UGE at these time points, the classical basal and luminal lineages were not fully
315 established yet (Figures 6F; Figure 6–figure supplements 13A-E; Supplementary file 4).
316 Nevertheless, C4-C6 had a more pronounced 'basal' identity compared to the other clusters.
317 *Krt5*/*Krt14* marked mainly C4, and additional basal markers including *Trp63, Dcn*, *Apoe,* or *Vcan*
318 were higher in C5/C6. Overall, known regulators of prostate development (Toivanen & Shen,
319 2017) displayed a variable expression pattern across the different clusters. For example, *Foxa1*
320 and *Shh* were strongly expressed in C0/C1, *Notch1* was higher in C3, and *Sox9* in C7 (Figure 6–
321 figure supplement 13C), highlighting the potential of this dataset to interrogate specific
322 features of prostate development.

323 Consistent with our previous results, *Runx1* was highly expressed in clusters having
324 lower *Nkx3-1* levels, including C0, C1, C2 and C4 (Figure 6G). To determine how these clusters
325 relate to differentiated prostate lineages, we interrogated population-specific gene signatures
326 previously identified in the adult (Figure 3). The 'Basal' signature was enriched across all
327 clusters, especially in C4/C6 (Figures 6I; Figure 6–figure supplement 13F, G). Strikingly, the
328 'Lum-D' derived signature was highly enriched in C2 compared to all the other adult luminal
329 population signatures, suggesting that the 'Lum-D' fate is determined early during prostate
330 development. The singular identity of C2 was characterized by genes previously found highly
331 expressed in the adult 'Lum-D' population, including *Tacstd2*/*Trop2*, *Krt4*, *Psca*, as well as *Ly6d*
332 and *Nupr1* (Figure 6H; Figure 6–figure supplement 13A).

333 Collectively, our scRNA-seq analysis show that adult 'Lum-D'/PLCs share strong
334 similarities with the unique C2 population identified in embryonic explant cultures. This

suggests that the distinct proximal luminal lineage is established at the very onset of prostate specification.

RUNX1$^+$ cells contribute to the establishment of the proximal luminal lineage during embryonic prostate development

To trace the fate of RUNX1$^+$ cells during embryonic prostate specification, we cultured UGS explants isolated from the *Runx1*$^{CRE\text{-}stop\text{-}RFP}$ lineage tracing model. We performed 2 pulses of tamoxifen treatment on day 0 and 1 of culture and analyzed the explants on day 2 and day 7 (Figure 7A). The majority of the RFP labeled cells were in the most proximal RUNX1$^+$ subset and rarely found in the distal area of the branches, where RUNX1$^-$ cells reside (Figures 7B, C). Accordingly, the proportion of RFP$^+$ RUNX1$^+$ cells remained stable between day 2 and 7 (Figure 7D). Also, the fraction of RFP$^+$ cells co-expressing p63 remained unchanged throughout the culture (Figure 7–figure supplements 14A-C), while a small fraction diversified into either K5$^+$ or K8$^+$ cells (Figure 7–figure supplements 14D, E). The scattered RFP$^+$ RUNX1$^-$ cells detected in distal branches by day 7 often co-expressed NKX3.1 (Figures 7E, F). Overall, this indicates that *Runx1* expressing cells only marginally contribute to the expansion of the NKX3.1 compartment (Figure 7G). Finally, we wondered whether RUNX1$^+$ cells contributed to the establishment of the proximal luminal lineage. We evaluated the proportion of RFP labelled cells co-expressing K4, previously identified as a marker of the developing C2 and adult Lum-D populations (Figures 3D and 6H). Interestingly, the fraction of K4$^+$ RFP labeled cells increased from 56.9 ± 10.6% to 74.1 ± 3.0% between day 2 and 7 (Figures 7F, G). There was also an increase of RFP$^+$ cells expressing *Nupr1*, another marker of the C2 cluster (Figure 7–figure supplements 14F-H). Taken

357 together, these results show that only a small subset of *Runx1* expressing cells contributes to

358 the expansion of NKX3.1^{+} lineage, found in the distal region of the developing prostatic buds.

359 Instead, the majority of *Runx1* expressing cells preferentially remain in the proximal region of

360 the premature buds, where the proximal luminal lineage is established.

DISCUSSION

In this study, we identified RUNX1 as a new marker of a luminal population enriched in the proximal region of the prostatic ducts. By combining scRNA-seq profiling and genetic lineage tracing of *Runx1* expressing cells, we show that RUNX1^{+} PLCs present in the intact prostate constitute a developmentally distinct and intrinsically castration-resistant luminal lineage. We propose that proximal and distal lineages are separate luminal entities from the earliest stages of prostate development. As such, our study provides novel insights into the cellular composition and developmental hierarchy of the mouse prostate epithelium.

Until the recent advances in single-cell technologies, the prostate epithelial hierarchy was mainly defined based on anatomical features of the basal and luminal layers, their histological characteristics and the expression of a small subset of markers (Shen & Abate-Shen, 2010; Toivanen & Shen, 2017). Here we present two comprehensive scRNA-seq dataset covering both the adult and the developing prostate. To our knowledge, this constitutes the first comprehensive single-cell atlas covering both intact and castrated adult mouse prostates, annotated by their lobe of origin. These datasets can be browsed interactively at http://shiny.cruk.manchester.ac.uk/pscapp/. In particular, our adult scRNA-seq dataset highlighted an extensive degree of cellular heterogeneity, in particular within the luminal epithelia. Several studies recently made similar observations either focusing on the AP (Karthaus et al., 2020), or the intact prostate (Crowley et al., 2020; Joseph et al., 2020). Integration of these multiple datasets will provide a more global view of the transcriptional landscape of the prostate epithelium.

Although mainly known as a master regulator of hematopoiesis, RUNX1 is increasingly implicated in hormone-associated epithelia including malignant conditions such as prostate cancer (Banach-Petrosky et al., 2007; Lie-A-Ling et al., 2020; Scheitz et al., 2012; Takayama et al., 2015). Here, we identified a subset of RUNX1$^+$ luminal cells located in the proximal region of the developing and adult prostate, referred to as RUNX1$^+$ PLCs, and corresponding to the Lum-D cluster identified in our adult scRNA-seq dataset. Of note, this subset appears to be the equivalent of the 'L2' (Karthaus et al., 2020) or 'LumP' (Crowley et al., 2020) clusters identified in recent studies. We demonstrate that these RUNX1$^+$ PLCs exhibit a greater organoid forming potential compared to the remaining luminal fraction, consistent with previous reports isolating similar proximal populations using different markers such as SCA-1, TROP2 or CD26 (Crowley et al., 2020; Goldstein et al., 2008; Karthaus et al., 2020; Kwon et al., 2016). Furthermore, RUNX1$^+$ PLCs predominantly formed unipotent K8$^+$ hollow organoids demonstrating their preferential commitment to the luminal fate. The greater clonogenicity of RUNX1$^+$ PLCs may in fact be linked to the gene expression profile of the corresponding Lum-D population, suggesting a more immature epithelial state, committed to the luminal lineage but not the secretory function of the prostate. It is therefore tempting to speculate that these cells act as a latent niche of 'facultative' stem cells (Clevers & Watt, 2018), primed to generate a structured prostatic epithelium under defined conditions.

Further characterization of RUNX1 expression in prostate development revealed a consistent expression pattern with the adult. RUNX1$^+$ luminal cells were restricted to the most proximal region of the developing prostate buds, both in embryos and UGS explant cultures. Our scRNA-seq of the developing prostate revealed a broad basal identity, supporting the

presence of multipotent basal progenitors during embryonic development (Ousset et al., 2012; Pignon et al., 2013), switching to unipotency postnatally (Tika et al., 2019). However, we observed a distinct cluster (C2) that strongly resembled the adult Lum-D population, suggesting an early branching event towards the proximal luminal fate at the onset of prostate development. Subsequent lineage tracing experiments indicated that *Runx1* expressing cells preferentially populate the emerging proximal luminal identity. It would be interesting to determine if the adult Lum-A, Lum-B, and Lum-C derive from multipotent-basal progenitors or from any specific clusters identified in the developing prostate. This appears to be the case at least for the adult Lum-D/RUNX1^{+} PLCs which already emerges during embryonic specification.

Our data also sheds a light on the regenerative potential of specific epithelial populations. Basal and luminal lineages have previously been shown to be largely self-sustained using generic basal and luminal *Cre* drivers (Choi et al., 2012; Ousset et al., 2012). However, whether distinct subpopulations of luminal cells contribute to the regeneration of the others remains poorly understood (X. Wang et al., 2009; Yoo et al., 2016). Our characterization of RUNX1^{+} PLCs and the detection of a wide variety of luminal populations in our adult prostate scRNA-seq data highlights the possible existence of multiple self-contained luminal populations. Indeed, *Runx1* driven genetic tracing experiments in regression-regeneration assays revealed that RUNX1^{+} PLCs did not contribute to the regeneration of distal NKX3.1^{+} cells. It was however evident that RUNX1^{+} PLCs are intrinsically castration resistant and capable of sustaining their own lineage in the regenerated prostate. Recently, it was proposed that prostate epithelial regeneration is driven by almost all luminal cells persisting in castrated prostates (Karthaus et al., 2020). Our results are compatible with this model, but we further demonstrate that not all

luminal subsets retain the same *in vivo* regenerative potential in response to androgen stimulation. Thus, we suggest that the model of self-sustained basal and luminal populations might be extended to individual luminal subpopulations. This hypothesis should be tested in the future using a more specific Lum-D *Cre* driver (e.g. *Krt4*/*Psca*). It will also be of interest to investigate the self-sustenance of other luminal compartments using Lum-A, Lum-B and Lum-C specific *Cre* drivers.

In light of the extensive contribution of RUNX transcription factors to developmental processes (Mevel et al., 2019), our study suggests a possible functional role for *Runx1* in specific lineages of the prostate epithelium, in particular in PLCs. Also, the broad expression of *Runx1* after castration suggest that it may be a valuable marker of castration-resistant cells. Functional links between RUNX factors and hormones, including AR have been previously reported (Chimge et al., 2016; Lie-A-Ling et al., 2020; Little et al., 2014; Riggio & Blyth, 2017; Takayama et al., 2015; van Bragt et al., 2014) and warrants further investigations. Along these lines, there seem to be an intriguing parallel between the low or absent expression of *Runx1* in prostate secretory cells with the documented downregulation of *Runx1* in secretory alveolar cells of the mammary gland epithelium (van Bragt et al., 2014). Overall, our results suggest that *Runx1*, but also *Runx2*, may be functionally involved in the development and maintenance of specific subpopulations of the prostate epithelium. It would now be of particular interest to further study the contribution of RUNX1 to prostate tumorigenesis. In light of the prognostic value of RUNX1 in triple-negative breast cancer (Ferrari et al., 2014), it would also be important to evaluate its biomarker potential in clinical cohorts of PCa samples.

In conclusion, we characterized the expression pattern of *Runx1* in the developing, normal and castrated mouse prostate. We observed that *Runx1* marks proximal luminal cells, which is a distinct luminal lineage emerging early during prostate specification, displaying intrinsic castration-resistant and self-sustaining properties. Our results therefore reveal strong intrinsic lineage differences within the luminal compartment. The emerging C2/Lum-D population retains a more embryonic-like program which could relate to their intrinsic castration-resistant potential and have broader relevance to cancer treatment. These results warrant future investigation of this luminal subset in the context of cancer development, tumor aggressiveness and treatment responses.

MATERIALS AND METHODS

Animal work

Animal experiments were approved by the Animal Welfare and Ethics Review Body (AWERB) of the Cancer Research UK Manchester Institute and conducted according to the UK Home Office Project Licence (PPL 70/8580). Genetic lineage-tracing experiments were performed at the Beatson Biological Services Unit (PPL 70/8645 & P5EE22AEE) and approved by the University of Glasgow AWERB. Mice were maintained in purpose-built facility in a 12-hour light/dark cycle with continual access to food and water.

Immunocompetent wild-type ICR (CD-1) mice were purchased from Envigo. P1-*Runx1*:GFP and P2-*Runx1*:RFP have been described previously (Draper et al., 2018; Sroczynska et al., 2009). Colonies were maintained on a ICR (CD-1) background. C57Bl/6J *Runx1*$^{mER\text{-}CRE\text{-}mER}$ (Samokhvalov et al., 2007) were provided by RIKEN (Japan). C57Bl/6J *Rosa26*$^{flox\text{-}stop\text{-}flox\text{-}tdRFP}$ mice (Luche et al., 2007) were acquired from the European Mouse Mutant Archive (EMMA). For all transgenic lines, routine genotyping was undertaken at weaning (3 weeks of age) by automated PCR genotyping (Transnetyx). For timed mating experiments, vaginal plug detection was considered as embryonic day (E) 0.5.

All animal procedures were performed on adult males at least 7 weeks of age. Surgical castration was carried out under aseptic conditions. For prostate regeneration assays, testosterone pellets (Belma Technologies) were implanted subcutaneously. For *in vivo* genetic lineage-tracing experiments, tamoxifen (Sigma, T5648) was resuspended in ethanol and diluted

in corn oil at a concentration of 10 mg/mL and administered via intra-peritoneal injections daily for 4 consecutive days using the following regimen: 3mg, 2mg, 2mg, 2mg.

Isolation of mouse prostate cells

All dissections were performed under a stereo microscope in sterile PBS. Dissociated murine prostate cells were obtained by digesting pre-minced prostate tissue for 1h at 37˚C in digestive medium prepared in prepared in ADMEM/F12 (Gibco), and containing 1mg/mL Collagenase Type I (ThermoFischer Scientific, #17018029), 1mg/mL Dispase II (ThermoFischer Scientific, #17105041), 10% Fetal Bovine Serum (Gibco), 1% Penicillin-Streptomycin-Glutamine (Sigma), and 10 µM Y-27632 dyhydrochloride (Chemdea, #CD0141). For embryonic urogenital sinuses (UGS), dissociation time was reduced to 30 min. Single cells were obtained after an additional 10 min incubation in TrypLE (Gibco) at 37˚C before mechanical dissociation with a syringe and needle (25G). Cells were then filtered through a 50 µm cell strainer.

Flow-cytometry and cell-sorting

Single cell suspensions were kept in Advanced DMEM/F-12 (Gibco) containing 5% FBS supplemented with 10 µM Y-27632. Cells were incubated for 10 min using unconjugated anti-mouse CD16/32 antibody (Biolegend, C93, #101301) at 4˚C prior to labelling with specific fluorochrome-labelled antibodies. Details of FACS reagents and antibodies are listed in Table 3. Cells were filtered through a 50 µm cell strainer prior to acquisition. Hoechst 33258 or Sytox blue (ThermoFischer Scientific) were used as viability stains. Single-cell suspensions were

analyzed on a Fortessa (BD Biosciences) and sorts were performed on a FACSAriaIII (BD Biosciences). FACS data were analyzed using FlowJo software (BD Life Sciences).

Organoid formation assays

In vitro organoid formation assays were performed as described in Drost et al., 2016. Single cells were resuspended in 40 µL drops of phenol red-free Cultrex RGF BME Type 2 (BME 2, Amsbio, #3533-005-02), and seeded in CellCarrier-96 Ultra Microplates (PerkinElmer, #6055302). Defined organoid culture medium was prepared with Advanced DMEM/F-12 (Gibco), supplemented with 10 mM Hepes (Sigma), Gutamax (Gibco), Penicillin/Streptomycin (Sigma), B27 (Life Technologies, 17504-044), 50 mg/mL EGF (PeproTech, #AF-100-15), 500 ng/mL R-spondin 1 (R&D Systems, #4645-RS), 100 ng/mL Noggin (R&D Systems, #6057-NG), 10 mM Y-27632 dyhydrochloride (Chemdea, #CD0141), 200nM A83-01 (Tocris Bioscience, #2939), 1.25 mM N-Acetylcystein (Sigma), and 1 nM Dihydrotestosterone (DHT, Sigma #730637). Medium was refreshed every 2-3 days, and organoid cultures were scored after 7 days.

UGS explant cultures

UGS explant cultures were performed as described previously (Kruithof-de Julio et al., 2013). Briefly, E15.5 embryos were obtained from timed matings. Urogenital sinuses (UGS) were isolated from the embryos and cultured using a Durapore Membrane Filter 0.65 µm (#DVPP02500) placed on a stainless-steel mesh for up to 7 days in Ham's F-12/DMEM (Gibco) supplemented with Insulin-Transferrin-Sodium Selenite Supplement (Roche) and 10 µM dihydrotestosterone (Sigma). Media were renewed every 2-3 days. For lineage-tracing

524 experiments, tamoxifen-induced labelling was performed using 0.5 µM 4-hydroxytamoxifen
525 (#T176, Sigma).
526
527 **Immunohistochemistry**
528 Prostate tissues were harvested and fixed in 10% buffered formalin for 24h. Fixed tissues
529 were processed using standard procedures and embedded in paraffin. Formalin-fixed paraffin-
530 embedded (FFPE) sections (4 µm) were cut and dried overnight at 37°C. Multiplexed
531 immunofluorescent stainings of FFPE sections were performed on an automated Leica BOND RX
532 platform using the Opal multiplexing workflow (PerkinElmer). In brief, sections were dewaxed,
533 and rehydrated, and endogenous peroxidase activity was quenched by 10 min pre-treatment
534 with 3% hydrogen peroxide diluted in TBS-T (Tris-Buffered Saline 0.05% Tween-20). Following
535 on-board heat-induced epitope retrieval with citrate buffer (pH 6.0) for 20 min, sections were
536 incubated for 10 min with 10% Casein (Vector Laboratories) diluted in TBS-T. Each staining cycle
537 included a primary antibody incubation for 30 min, followed by buffer washes, and 30 min
538 incubation with HRP labelled secondary antibodies (Table 2). After further washes, the
539 Tyramide labeled with a fluorophore (Opal 520, Opal 570 or Opal 650, PerkinElmer) was added
540 for a final 10 min. Subsequent antibody stainings were performed by repeating the same
541 procedure, separated by heat-mediated antibody denaturation using citrate buffer (pH 6.0) for
542 5 min at 95°C. Nuclei were counterstained with DAPI (Sigma) and slides were sealed using
543 ProLong Gold Antifade Mountant (ThermoFischer Scientific). *In situ* hybridization (ISH) to detect
544 *Nupr1* (ACD, LS 2.5 Mm-Nupr1 #434818) was done using the Multiplex Fluorescent detection kit
545 (ACD) on the automated Leica BOND RX platform following the manufacturer's instructions.

Pre-treatment was done using an EDTA based pH 9.0 epitope retrieval solution for 15 min at 88°C followed by 10 min protease incubation. After ISH, antibody staining was carried out using an anti-RFP antibody for 1h detected with EnVision HRP anti-rabbit secondary (Agilent) followed by incubation with Tyramide-conjugated Opal 570 (PerkinElmer) as described above. Anti-CDH1 antibody was applied for 1h and detected using an anti-goat Alexa Fluor 647 secondary antibody (ThermoFischer Scientific, #A-21447). Staining of frozen sections was performed as described previously (Thambyrajah et al., 2016). The list of antibodies used is available in Table 1.

Image acquisition and analysis

Whole-slide images were acquired on an Olympus VS120 slide scanner. Images were analyzed using QuPath v0.2 (Bankhead et al., 2017). Briefly, annotations were drawn manually to select areas of interest. Nuclear detection was achieved using the “cell detection” module on the DAPI channel. A classifier was then trained for each batch of images using the random forest algorithm, to detect the epithelial layers based on either CDH1 or K5/K8 stainings. Single-cell intensity measurements were analyzed using R (3.6.3). For Quantitative Imaged-Based Cytometry (QBIC), single-cell intensity measurements were log10 transformed and plotted using the ‘geom_hex’ function of the ggplot2 R package. QuPath was used to extract representative high-quality raw images of selected areas from whole slide images using the ‘Send region to ImageJ’ tool. Images used for publication were processed with ImageJ (NIH Image, Maryland, USA). Confocal images were acquired using a Leica TCS SP8 confocal

567 microscope and LAS X Leica software. Images of whole UGS explant culture were taken using a
568 Leica MZ FLIII microscope.
569
570 **Whole-mount immunofluorescent staining of organoids**
571 Whole-mount staining was adapted from Yokomizo et al., 2012. Organoids were fixed
572 directly in 96-well plates using 4% paraformaldehyde for 1h at 4°C. After 3 washes of 5 min in
573 PBS, organoids were incubated in PBS-BST, containing PBS, 1% milk, 1% BSA, 10% goat serum
574 (Agilent, #X090710), 0.4% Triton X-100. Pre-conjugated primary antibodies, K5 Alexa Fluor 647
575 (#ab193895, Abcam) and K8 Alexa Fluor 488 (#ab192467, Abcam) were diluted at 1/400 in PBS-
576 BST and incubated with the organoids overnight at 4°C on a rocking platform. After 3 washes of
577 1h in PBS-BST at 4°C, organoids were stained with DAPI at 2 µg/mL diluted in PBS-BST and
578 incubated for another 30 min at 4°C on a rocking platform. Images were acquired on an Opera
579 Phenix High Content Screening System using the 10x air and 20x water lenses. Quantitative
580 analysis was performed using the Harmony software on maximum projection images.
581
582 **scRNA-seq sample preparation**
583 A detailed description of the samples, replicates, and the corresponding cellular populations
584 used for each sequencing run is provided in Supplementary file 1. For the adult mouse prostate
585 dataset, AP, DLP, and VP lobes were micro dissected and pooled from P2-*Runx1*:RFP reporter
586 mice after dissociation. Single live $EPCAM^{+}$ cells from RFP^{+} and RFP^{-} fractions of each lobes were
587 sorted separately (containing a mix of $CD49f^{high}$ basal and $CD24^{high}$ luminal cells). For the UGS
588 explant culture dataset, the middle regions of the explants were micro dissected to enrich for

prostatic branching events and pooled by time point after dissociation. Single live EPCAM+ cells were sorted for each independent time point.

scRNA-seq sample multiplexing

Individually sorted populations were multiplexed using the MULTI-seq protocol (McGinnis, Patterson, et al., 2019). Reagents were kindly provided by Dr. Zev Gartner. In brief, after sorting, cells were washed once in cold serum- and BSA-free PBS. A lipid-modified DNA oligonucleotide and a specific sample barcode oligonucleotide were then mixed and added to the cells at a final concentration of 200 nM each, and incubated in cold PBS for 5 min. Each individual sample to be multiplexed received an independent sample barcode. Next, a common lipid-modified co-anchor was added at 200 nM to each sample to stabilize the membrane bound barcodes. After an additional 5 min incubation on ice, cells were washed 2 times with PBS containing 1% FBS 1% BSA in order to quench unbound barcodes. Samples were then pooled together and washed once with PBS 1% FBS 1% BSA. After cell counting, cells were loaded in a Chromium Single Cell 3' GEM Library & Gel Bead Kit v3 (10x Genomics).

scRNA-seq library preparation, sequencing and pre-processing

Gene expression (cDNA) libraries were prepared according to the manufacturer's protocol. MULTI-seq barcode libraries were separated from the cDNA libraries during the first round of size selection, and PCR amplified prior to sequencing according to the MULTI-seq library preparation protocol (McGinnis, Patterson, et al., 2019). For the adult mouse prostate dataset, cDNA libraries of 'run 1' and 'run 2' were sequenced on Illumina NovaSeq 6000 System, and

'run 3' was sequenced on Illumina HiSeq 2500. The UGS mouse prostate explant run was also sequenced on Illumina HiSeq 2500. Sequencing data of cDNA libraries were processed using Cellranger v3.1.0 and mapped onto mm10 mouse reference genome. Pre-processing of the MULTI-seq library fastq files was performed using the 'deMULTIplex' (v1.0.2) R package (https://github.com/chris-mcginnis-ucsf/MULTI-seq) to generate a sample barcode UMI count matrix. Detailed quality control metrics of each sequencing run are provided in Supplementary file 1.

Adult mouse prostate dataset analysis

Quality control and barcode demultiplexing of individual runs. Each run was pre-processed individually prior data integration. Cellranger outputs were loaded into the R package Seurat (v3.1.5). Cells were kept if they had more than 750 detected genes, less than 7500 UMIs and less than 10% mitochondrial transcripts. Sample barcodes were demultiplexed using the HTODemux function implemented in Seurat. Briefly, a negative binomial distribution was used to estimate the background levels based on *k*-means clustering of the barcode normalized counts. Barcodes with values above the 99% quantile were considered 'positive' for a given sample. Cells positive for more than one barcode were considered as 'doublets'. Doublets and negative cells were excluded for all downstream analyses. Thresholds were empirically adjusted to remove additional cells with possible ambiguous classification (Supplementary file 1). Of note, in both 'run 1' and 'run 2', a large number of cells were classified 'negative' due to the failed labelling of 'Bar3' (corresponding to 'Intact DLP RFP+' sample). For these runs, we used DoubletFinder (McGinnis, Murrow, et al., 2019) to remove predicted doublets missed out as a

consequence of the failed labeling of 'Bar3'. After classification, barcodes were represented in UMAP space to confirm the purity of the barcode assignment obtained for each sample (Figure 3–figure supplement 5A). We obtained a total of 4,499 cells from 3 independent experiments.

Integration, low dimensional embedding and clustering. Data aggregation was performed according to the standard integration procedure implemented in Seurat. In brief, each dataset was log normalized, and 3000 variable features were initially computed using the 'vst' method. For integration, 2000 features and 50 dimensions were used as anchors. Integrated data were scaled and the first 50 principal components (PC) were calculated for downstream analyses. Uniform Manifold Approximation and Projection (UMAP) (McInnes et al., 2018) was used for visualization. Graph-based louvain clustering was performed on a shared nearest neighbor graph constructed using 20 nearest neighbors for every cell, and a resolution of 0.4, which gave a reasonable segmentation of the data (Figure 3–figure supplement 5B,C). Extensive exploration of each cluster based on known marker genes was then carried out to subset prostate epithelial cells. We found 10 prostate epithelial clusters (*Epcam*, *Krt8*, *Cd24a*, *Spink1*, *Krt19*, *Tacstd2*, *Psca*, *Krt4*, *Tgm4*, *Nkx3-1*, *Pbsn*, *Msmb*, *Piezo2*, *Trp63*, *Krt5*, *Krt14*), 3 clusters of hematopoietic cells (*Vim*, *Ptprc*, *Cd74*, *Itgam*, *Cd3d*), 1 cluster of endothelial cells (*Pecam1*), 1 cluster of fibroblasts (*Vim*, *Col1a1*) and 1 cluster of mesonephric derivatives (*Svs2*, *Pax2*) (Figure 3–figure supplement 5D,E).

Analysis of prostate epithelial populations. The same dimension reduction approach described above was performed on the selected prostate epithelial clusters, using a resolution of 0.3 for

graph-based louvain clustering. We annotated 1 large population of basal cells by merging 3 subclusters highly expressing *Krt5*, *Krt14* and *Trp63* as we did not discuss the heterogeneity of the basal compartment in this study (Figures 3B-D; Figure 3–figure supplement 6F,I). We annotated the different luminal clusters expressing higher levels of *Cd26*/*Dpp4, Cd24a*, *Krt8* and *Krt18*, as Lum-A, Lum-B, Lum-C, Lum-D, Lum-E and Lum-F. Several genes specifically marked each cluster, including *Sbp*/*Spink1* in Lum-A, *Tgm4* in Lum-B, *Msmb* in Lum-C, *Psca*/*Krt4* in Lum-D, *Basp1*/*Lpl* in Lum-E, and *Crym* in Lum-F (Figures 3C, D; Figure 3–figure supplement 7A). Data were then imported in Scanpy (v1.4.6) to infer lineage relationships between cellular populations via partition-based graph abstraction (PAGA) implemented in the tl.paga function (Wolf et al., 2019). Briefly, a single cell neighborhood graph (n_neighbors = 50) was computed using the integrated principal components previously calculated in Seurat. PAGA was generated based on our annotated clusters. The final UMAP representation was generated using PAGA-initialised positions to better preserve the global topology of the data. All final data visualizations were generated in R.

Differential gene expression analysis and gene ontology. Differential gene expression analyses between clusters were performed using the MAST method (Finak et al., 2015) implemented in Seurat within the 'FindAllMarkers' and 'FindMarkers' functions. Testing was limited to genes detected in at least 25% of the tested populations (min.pct = 0.25) and showing at least ± 0.25 log fold change difference (logfc.threshold = 0.25). The 'g:GOSt' function of the gprofiler2 R package was used to perform functional enrichment analysis on gene ontology terms (GO:BP,

676 biological processes). Genes showing at least 0.50 log fold change enrichment in the group
677 tested were kept.
678
679 **UGS explant cultures dataset**
680 A similar strategy was applied for the analysis of the UGS explant culture dataset, with some
681 alterations described below.
682
683 *Quality control and barcode demultiplexing.* Cells were kept if they had more than 1000
684 detected genes, and less than 7.5% mitochondrial transcripts. Barcode classification was
685 performed as above, using the 90% quantile in 'HTODemux' (Figure 6–figure supplement 12A).
686 We obtained a total of 5,122 cells that passed quality control from the 4 time points.
687
688 *Low dimensional embedding and clustering.* The first 50 principal components and 20 neighbors
689 were used for UMAP visualization. Graph-based clustering was done using a resolution
690 parameter of 0.3. We noticed a strong effect of cell cycle using cell cycles genes defined in
691 Tirosh et al., 2016. This was particularly evident using the 'CellCycleScoring' function
692 implemented in Seurat (Figure 6–figure supplement 12B). To minimize the impact of cell cycle
693 on downstream analyses, the cell cycle scores were regressed out during data scaling. We
694 identified 6 main clusters, that we annotated based on the expression of several marker genes
695 (Figure 6–figure supplement 12C-E). We identified 2 clusters of developing mesonephric
696 derivatives (*Hoxb7*, *Wfdc2*, *Gata3*, *Sox17*, *Pax2*, *Pax8*, *Lhx1*), 1 cluster of developing bladder
697 urothelium (*Upk3a*, *Foxq1*, *Plaur*, *Krt7*, *Krt20*), 1 cluster of mesenchymal cells (*Vim*, *Col3a1*,

698 *Col1a1*, *Pdgfra*, *Zeb1*) and 1 cluster corresponding to the developing prostatic epithelium
699 (*Epcam*, *Krt8*, *Krt5*, *Krt14*, *Krt15*, *Shh*, *Hoxb13*, *Hoxd13*, *Nkx3-1*). We also identified one cluster
700 largely associated with hypoxia and cellular stress ontologies (Figure 6–figure supplement 12F).
701
702 *Analysis of the developing prostatic epithelium.* The same dimension reduction approach was
703 initially applied on the developing prostatic cluster. After graph-based clustering using a
704 resolution of 0.5, 10 clusters were identified and visualized via UMAP (Figure 6–figure
705 supplement 12G-J). We computed diffusion components using 'runDiffusionMap'
706 (ncomponents = 20, k = 20) implemented in the scater (v1.14.6) R package. We found the small
707 cluster C9 to be largely diverging from the remainder fraction in diffusion space, therefore it
708 was excluded for downstream analysis (Figure 6–figure supplement 12K). We then imported
709 the data in Scanpy and used the first 10 diffusion components to compute a neighborhood
710 graph (n_neighbors = 20) which was used for PAGA. We finally computed a force-direct layout
711 (ForceAtlas2) using PAGA-initialised positions.
712
713 *Analysis of gene set activity.* Gene signatures were generated from the list of differentially
714 expressed genes by keeping those showing at least 0.50 log fold change enrichment in each
715 given group. Gene lists were used as custom gene sets (Supplementary file 5) in the AUCell
716 (Aibar et al., 2017) R package (v1.8.0). Briefly, AUCell uses the Area Under the Curve to evaluate
717 the enrichment of a given gene set in each cell, in a ranking based manner. It outputs an AUC
718 score for each individual cell, which is used to explore the relative expression of the signature.

Per cell AUC scores of each signatures were overlayed on the dimension reduction layout and plotted as boxplots to visualize enrichments across the different cellular subsets.

Data availability

Raw sequencing files and processed gene expression matrices have been deposited in the NCBI Gene Expression Omnibus under the accession number GSE151944. The processed datasets for both mouse adult prostate and UGS prostate explant cultures can be accessed via a searchable R Shiny application available at http://shiny.cruk.manchester.ac.uk/pscapp/. All code used to process data and generate figures is available on a public GitHub repository at https://github.com/glacaud/prostate-scRNAseq.

Statistical analyses

Statistical analyses were performed using Graphpad/Prism (v8.4.2). Data are represented as mean ±SD. Two-tailed unpaired *t*-tests were used to compare means between two groups. Statistical significance was set at $p < 0.05$. For animal model studies, no statistical method was used to pre-determine the sample size. No randomization or blinding was used for *in vivo* studies.

ACKNOWLEDGEMENTS

We thank the laboratories' members for critical reading of the manuscript, in particular Dr. Michael Lie-a-ling, Dr. Alice Lallo and Dr. Catherine Winchester. We thank the staff at the Histology, Flow Cytometry, Advanced Imaging, Molecular Biology, and Breeding Unit Core facilities of CRUK Manchester Institute, as well as the CRUK Beatson Biological Services Unit and Flow Cytometry Core facility for technical support. We thank Dr. Zev Gartner for kindly providing the MULTI-seq reagents. We thank Dr. Kirsteen Campbell with assistance on the lineage tracing experiments. We thank Dr. Berenika Plusa and Dr. Roberto de la Fuente with assistance on timed matings experiments. We thank Professor Hans Jorg Fehling and the European Mouse Mutant Archive for providing the Rosa26$^{flox\text{-}stop\text{-}flox\text{-}tdRFP}$ mice.

AUTHORS CONTRIBUTION

Conceptualization, R.M., E.B., K.B., and G.L.; Methodology, R.M., E.B., K.B., and G.L.; Formal analysis, R.M.; Data curation, R.M.; Visualization, R.M.; Investigation, R.M., I.S., S.M., L.G., R.P., M.Z.F.; Writing – Original Draft, R.M.; Writing – Review & Editing, R.M., I.S., K.B., E.B., G.L.; Resources, P.O., H.Y.L and I.A.; Supervision, E.B., K.B. and G.L.; Project Administration, G.L.; Funding Acquisition, E.B., K.B. and G.L.

FUNDING

This research was supported by CRUK core funding C5759/A20971 (G.L. and E.B., CRUK Manchester Institute) and C596/A17196 (K.B., CRUK Beatson Institute). The funders had no role

757 in study design, data collection and interpretation, or the decision to submit the work for

758 publication.

759

760 **COMPETING INTERESTS**

761 The authors declare no competing interests.

REFERENCES

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., van den Oord, J., Atak, Z. K., Wouters, J., & Aerts, S. (2017). SCENIC: Single-cell regulatory network inference and clustering. *Nature Methods*, *14*(11), 1083–1086. https://doi.org/10.1038/nmeth.4463

Banach-Petrosky, W., Jessen, W. J., Ouyang, X., Gao, H., Rao, J., Quinn, J., Aronow, B. J., & Abate-Shen, C. (2007). Prolonged exposure to reduced levels of androgen accelerates prostate cancer progression in Nkx3.1; Pten mutant mice. *Cancer Research*, *67*(19), 9089–9096. https://doi.org/10.1158/0008-5472.CAN-07-2887

Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., James, J. A., Salto-Tellez, M., & Hamilton, P. W. (2017). QuPath: Open source software for digital pathology image analysis. *Scientific Reports*, *7*(1), 16878. https://doi.org/10.1038/s41598-017-17204-5

Barros-Silva, J. D., Linn, D. E., Steiner, I., Guo, G., Ali, A., Pakula, H., Ashton, G., Peset, I., Brown, M., Clarke, N. W., Bronson, R. T., Yuan, G.-C., Orkin, S. H., Li, Z., & Baena, E. (2018). Single-Cell Analysis Identifies LY6D as a Marker Linking Castration-Resistant Prostate Luminal Cells to Prostate Progenitors and Cancer. *Cell Reports*, *25*(12), 3504-3518.e6. https://doi.org/10.1016/j.celrep.2018.11.069

Berman, D. M., Desai, N., Wang, X., Karhadkar, S. S., Reynon, M., Abate-Shen, C., Beachy, P. A., & Shen, M. M. (2004). Roles for Hedgehog signaling in androgen production and prostate ductal morphogenesis. *Developmental Biology*, *267*(2), 387–398. https://doi.org/10.1016/j.ydbio.2003.11.018

784 Bhatia-Gaur, R., Donjacour, A. A., Sciavolino, P. J., Kim, M., Desai, N., Young, P., Norton, C. R.,
785 Gridley, T., Cardiff, R. D., Cunha, G. R., Abate-Shen, C., & Shen, M. M. (1999). Roles for
786 Nkx3.1 in prostate development and cancer. *Genes & Development*, *13*(8), 966–977.
787 Browne, G., Taipaleenmäki, H., Bishop, N. M., Madasu, S. C., Shaw, L. M., van Wijnen, A. J.,
788 Stein, J. L., Stein, G. S., & Lian, J. B. (2015). Runx1 is associated with breast cancer
789 progression in MMTV-PyMT transgenic mice and its depletion in vitro inhibits migration
790 and invasion. *Journal of Cellular Physiology*, *230*(10), 2522–2532.
791 https://doi.org/10.1002/jcp.24989
792 Burger, P. E., Xiong, X., Coetzee, S., Salm, S. N., Moscatelli, D., Goto, K., & Wilson, E. L. (2005).
793 Sca-1 expression identifies stem cells in the proximal region of prostatic ducts with high
794 capacity to reconstitute prostatic tissue. *Proceedings of the National Academy of*
795 *Sciences of the United States of America*, *102*(20), 7180–7185.
796 https://doi.org/10.1073/pnas.0502761102
797 Chimge, N.-O., Little, G. H., Baniwal, S. K., Adisetiyo, H., Xie, Y., Zhang, T., O'Laughlin, A., Liu, Z.
798 Y., Ulrich, P., Martin, A., Mhawech-Fauceglia, P., Ellis, M. J., Tripathy, D., Groshen, S.,
799 Liang, C., Li, Z., Schones, D. E., & Frenkel, B. (2016). RUNX1 prevents oestrogen-
800 mediated AXIN1 suppression and β-catenin activation in ER-positive breast cancer.
801 *Nature Communications*, *7*, 10751. https://doi.org/10.1038/ncomms10751
802 Choi, N., Zhang, B., Zhang, L., Ittmann, M., & Xin, L. (2012). Adult murine prostate basal and
803 luminal cells are self-sustained lineages that can both serve as targets for prostate
804 cancer initiation. *Cancer Cell*, *21*(2), 253–265. https://doi.org/10.1016/j.ccr.2012.01.005

805 Chua, C. W., Shibata, M., Lei, M., Toivanen, R., Barlow, L. J., Bergren, S. K., Badani, K. K.,
806 McKiernan, J. M., Benson, M. C., Hibshoosh, H., & Shen, M. M. (2014). Single luminal
807 epithelial progenitors can generate prostate organoids in culture. *Nature Cell Biology*,
808 *16*(10), 951–961, 1–4. https://doi.org/10.1038/ncb3047
809 Clevers, H., & Watt, F. M. (2018). Defining Adult Stem Cells by Function, not by Phenotype.
810 *Annual Review of Biochemistry*, *87*(1), 1015–1027. https://doi.org/10.1146/annurev-
811 biochem-062917-012341
812 Crowell, P. D., Fox, J. J., Hashimoto, T., Diaz, J. A., Navarro, H. I., Henry, G. H., Feldmar, B. A.,
813 Lowe, M. G., Garcia, A. J., Wu, Y. E., Sajed, D. P., Strand, D. W., & Goldstein, A. S. (2019).
814 Expansion of Luminal Progenitor Cells in the Aging Mouse and Human Prostate. *Cell*
815 *Reports*, *28*(6), 1499-1510.e6. https://doi.org/10.1016/j.celrep.2019.07.007
816 Crowley, L., Cambuli, F., Aparicio, L., Shibata, M., Robinson, B. D., Xuan, S., Li, W., Hibshoosh, H.,
817 Loda, M., Rabadan, R., & Shen, M. M. (2020). A single-cell atlas of the mouse and human
818 prostate reveals heterogeneity and conservation of epithelial progenitors. *BioRxiv*,
819 2020.05.29.123836. https://doi.org/10.1101/2020.05.29.123836
820 Doles, J. D., Vezina, C. M., Lipinski, R. J., Peterson, R. E., & Bushman, W. (2005). Growth,
821 morphogenesis, and differentiation during mouse prostate development in situ, in renal
822 grafts, and in vitro. *The Prostate*, *65*(4), 390–399. https://doi.org/10.1002/pros.20321
823 Draper, J. E., Sroczynska, P., Fadlullah, M. Z. H., Patel, R., Newton, G., Breitwieser, W., Kouskoff,
824 V., & Lacaud, G. (2018). A novel prospective isolation of murine fetal liver progenitors to
825 study in utero hematopoietic defects. *PLoS Genetics*, *14*(1), e1007127.
826 https://doi.org/10.1371/journal.pgen.1007127

827 Drost, J., Karthaus, W. R., Gao, D., Driehuis, E., Sawyers, C. L., Chen, Y., & Clevers, H. (2016).
828 Organoid culture systems for prostate epithelial and cancer tissue. *Nature Protocols*,
829 *11*(2), 347–358. https://doi.org/10.1038/nprot.2016.006
830 English, H. F., Santen, R. J., & Isaacs, J. T. (1987). Response of glandular versus basal rat ventral
831 prostatic epithelial cells to androgen withdrawal and replacement. *The Prostate*, *11*(3),
832 229–242.
833 Ferrari, N., Mohammed, Z. M. A., Nixon, C., Mason, S. M., Mallon, E., McMillan, D. C., Morris, J.
834 S., Cameron, E. R., Edwards, J., & Blyth, K. (2014). Expression of RUNX1 Correlates with
835 Poor Patient Prognosis in Triple Negative Breast Cancer. *PLoS ONE*, *9*(6), e100759.
836 https://doi.org/10.1371/journal.pone.0100759
837 Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A. K., Slichter, C. K., Miller, H. W.,
838 McElrath, M. J., Prlic, M., Linsley, P. S., & Gottardo, R. (2015). MAST: A flexible statistical
839 framework for assessing transcriptional changes and characterizing heterogeneity in
840 single-cell RNA sequencing data. *Genome Biology*, *16*(1), 278.
841 https://doi.org/10.1186/s13059-015-0844-5
842 Georgas, K. M., Armstrong, J., Keast, J. R., Larkins, C. E., McHugh, K. M., Southard-Smith, E. M.,
843 Cohn, M. J., Batourina, E., Dan, H., Schneider, K., Buehler, D. P., Wiese, C. B., Brennan, J.,
844 Davies, J. A., Harding, S. D., Baldock, R. A., Little, M. H., Vezina, C. M., & Mendelsohn, C.
845 (2015). An illustrated anatomical ontology of the developing mouse lower urogenital
846 tract. *Development (Cambridge, England)*, *142*(10), 1893–1908.
847 https://doi.org/10.1242/dev.117903

848 Goldstein, A. S., Lawson, D. A., Cheng, D., Sun, W., Garraway, I. P., & Witte, O. N. (2008). Trop2
849 identifies a subpopulation of murine and human prostate basal cells with stem cell
850 characteristics. *Proceedings of the National Academy of Sciences of the United States of*
851 *America*, *105*(52), 20882–20887. https://doi.org/10.1073/pnas.0811411106
852 Goto, K., Salm, S. N., Coetzee, S., Xiong, X., Burger, P. E., Shapiro, E., Lepor, H., Moscatelli, D., &
853 Wilson, E. L. (2006). Proximal Prostatic Stem Cells Are Programmed to Regenerate a
854 Proximal-Distal Ductal Axis. *STEM CELLS*, *24*(8), 1859–1868.
855 https://doi.org/10.1634/stemcells.2005-0585
856 Höfner, T., Eisen, C., Klein, C., Rigo-Watermeier, T., Goeppinger, S. M., Jauch, A., Schoell, B.,
857 Vogel, V., Noll, E., Weichert, W., Baccelli, I., Schillert, A., Wagner, S., Pahernik, S., Sprick,
858 M. R., & Trumpp, A. (2015). Defined Conditions for the Isolation and Expansion of Basal
859 Prostate Progenitor Cells of Mouse and Human Origin. *Stem Cell Reports*, *4*(3), 503–518.
860 https://doi.org/10.1016/j.stemcr.2015.01.015
861 Horton, C., Liu, Y., Yu, C., Xie, Q., & Wang, Z. A. (2019). Luminal-contact-inhibition of epithelial
862 basal stem cell multipotency in prostate organogenesis and homeostasis. *Biology Open*,
863 *8*(10). https://doi.org/10.1242/bio.045724
864 Joseph, D. B., Henry, G. H., Malewska, A., Iqbal, N. S., Ruetten, H. M., Turco, A. E., Abler, L. L.,
865 Sandhu, S. K., Cadena, M. T., Malladi, V. S., Reese, J. C., Mauck, R. J., Gahan, J. C.,
866 Hutchinson, R. C., Roehrborn, C. G., Baker, L. A., Vezina, C. M., & Strand, D. W. (2020).
867 Urethral luminal epithelia are castration-insensitive cells of the proximal prostate. *The*
868 *Prostate*. https://doi.org/10.1002/pros.24020

869 Karthaus, W. R., Hofree, M., Choi, D., Linton, E. L., Turkekul, M., Bejnood, A., Carver, B.,
870 Gopalan, A., Abida, W., Laudone, V., Biton, M., Chaudhary, O., Xu, T., Masilionis, I.,
871 Manova, K., Mazutis, L., Pe'er, D., Regev, A., & Sawyers, C. L. (2020). Regenerative
872 potential of prostate luminal cells revealed by single-cell analysis. *Science*, *368*(6490),
873 497–505. https://doi.org/10.1126/science.aay0267
874 Karthaus, W. R., Iaquinta, P. J., Drost, J., Gracanin, A., van Boxtel, R., Wongvipat, J., Dowling, C.
875 M., Gao, D., Begthel, H., Sachs, N., Vries, R. G. J., Cuppen, E., Chen, Y., Sawyers, C. L., &
876 Clevers, H. C. (2014). Identification of multipotent luminal progenitor cells in human
877 prostate organoid cultures. *Cell*, *159*(1), 163–175.
878 https://doi.org/10.1016/j.cell.2014.08.017
879 Keil, K. P., Mehta, V., Abler, L. L., Joshi, P. S., Schmitz, C. T., & Vezina, C. M. (2012). Visualization
880 and quantification of mouse prostate development by in situ hybridization.
881 *Differentiation; Research in Biological Diversity*, *84*(3), 232–239.
882 https://doi.org/10.1016/j.diff.2012.07.005
883 Keita, M., Bachvarova, M., Morin, C., Plante, M., Gregoire, J., Renaud, M.-C., Sebastianelli, A.,
884 Trinh, X. B., & Bachvarov, D. (2013). The RUNX1 transcription factor is expressed in
885 serous epithelial ovarian carcinoma and contributes to cell proliferation, migration and
886 invasion. *Cell Cycle (Georgetown, Tex.)*, *12*(6), 972–986.
887 https://doi.org/10.4161/cc.23963
888 Kruithof-de Julio, M., Shibata, M., Desai, N., Reynon, M., Halili, M. V., Hu, Y.-P., Price, S. M.,
889 Abate-Shen, C., & Shen, M. M. (2013). Canonical Wnt signaling regulates Nkx3.1
890 expression and luminal epithelial differentiation during prostate organogenesis.

Developmental Dynamics : An Official Publication of the American Association of Anatomists, *242*(10), 1160–1171. https://doi.org/10.1002/dvdy.24008

Kwon, O.-J., Zhang, L., Ittmann, M. M., & Xin, L. (2014). Prostatic inflammation enhances basal-to-luminal differentiation and accelerates initiation of prostate cancer with a basal cell origin. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(5), E592-600. https://doi.org/10.1073/pnas.1318157111

Kwon, O.-J., Zhang, L., & Xin, L. (2016). Stem Cell Antigen-1 Identifies a Distinct Androgen-Independent Murine Prostatic Luminal Cell Lineage with Bipotent Potential. *STEM CELLS*, *34*(1), 191–202. https://doi.org/10.1002/stem.2217

Lawson, D. A., Xin, L., Lukacs, R. U., Cheng, D., & Witte, O. N. (2007). Isolation and functional characterization of murine prostate stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(1), 181–186. https://doi.org/10.1073/pnas.0609684104

Leong, K. G., Wang, B.-E., Johnson, L., & Gao, W.-Q. (2008). Generation of a prostate from a single adult stem cell. *Nature*, *456*(7223), 804–808. https://doi.org/10.1038/nature07427

Lie-A-Ling, M., Mevel, R., Patel, R., Blyth, K., Baena, E., Kouskoff, V., & Lacaud, G. (2020). RUNX1 Dosage in Development and Cancer. *Molecules and Cells*, *43*(2), 126–138. https://doi.org/10.14348/molcells.2019.0301

Little, G. H., Baniwal, S. K., Adisetiyo, H., Groshen, S., Chimge, N.-O., Kim, S. Y., Khalid, O., Hawes, D., Jones, J. O., Pinski, J., Schones, D. E., & Frenkel, B. (2014). Differential effects of RUNX2 on the androgen receptor in prostate cancer: Synergistic stimulation of a gene

913 set exemplified by SNAI2 and subsequent invasiveness. *Cancer Research*, *74*(10), 2857–
914 2868. https://doi.org/10.1158/0008-5472.CAN-13-2003
915 Lopes, E. S., Foster, B. A., Donjacour, A. A., & Cunha, G. R. (1996). Initiation of secretory activity
916 of rat prostatic epithelium in organ culture. *Endocrinology*, *137*(10), 4225–4234.
917 https://doi.org/10.1210/endo.137.10.8828481
918 Luche, H., Weber, O., Nageswara Rao, T., Blum, C., & Fehling, H. J. (2007). Faithful activation of
919 an extra-bright red fluorescent protein in 'knock-in' Cre-reporter mice ideally suited for
920 lineage tracing studies. *European Journal of Immunology*, *37*(1), 43–53.
921 https://doi.org/10.1002/eji.200636745
922 Lukacs, R. U., Goldstein, A. S., Lawson, D. A., Cheng, D., & Witte, O. N. (2010). Isolation,
923 cultivation and characterization of adult murine prostate stem cells. *Nature Protocols*,
924 *5*(4), 702–713. https://doi.org/10.1038/nprot.2010.11
925 McAuley, E., Moline, D., VanOpstall, C., Lamperis, S., Brown, R., & Vander Griend, D. J. (2019).
926 Sox2 Expression Marks Castration-Resistant Progenitor Cells in the Adult Murine
927 Prostate. *Stem Cells (Dayton, Ohio)*, *37*(5), 690–700. https://doi.org/10.1002/stem.2987
928 McGinnis, C. S., Murrow, L. M., & Gartner, Z. J. (2019). DoubletFinder: Doublet Detection in
929 Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Systems*, *8*(4),
930 329-337.e4. https://doi.org/10.1016/j.cels.2019.03.003
931 McGinnis, C. S., Patterson, D. M., Winkler, J., Conrad, D. N., Hein, M. Y., Srivastava, V., Hu, J. L.,
932 Murrow, L. M., Weissman, J. S., Werb, Z., Chow, E. D., & Gartner, Z. J. (2019). MULTI-
933 seq: Sample multiplexing for single-cell RNA sequencing using lipid-tagged indices.
934 *Nature Methods*, *16*(7), 619–626. https://doi.org/10.1038/s41592-019-0433-8

McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv:1802.03426 [Cs, Stat]*. http://arxiv.org/abs/1802.03426

McNeal, J. E. (1981). The zonal anatomy of the prostate. *The Prostate*, *2*(1), 35–49. https://doi.org/10.1002/pros.2990020105

Mevel, R., Draper, J. E., Lie-a-Ling, M., Kouskoff, V., & Lacaud, G. (2019). RUNX transcription factors: Orchestrators of development. *Development*, *146*(17), dev148296. https://doi.org/10.1242/dev.148296

Ousset, M., Van Keymeulen, A., Bouvencourt, G., Sharma, N., Achouri, Y., Simons, B. D., & Blanpain, C. (2012). Multipotent and unipotent progenitors contribute to prostate postnatal development. *Nature Cell Biology*, *14*(11), 1131–1138. https://doi.org/10.1038/ncb2600

Pignon, J.-C., Grisanzio, C., Geng, Y., Song, J., Shivdasani, R. A., & Signoretti, S. (2013). P63-expressing cells are the stem cells of developing prostate, bladder, and colorectal epithelia. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(20), 8105–8110. https://doi.org/10.1073/pnas.1221216110

Planagumà, J., Díaz-Fuertes, M., Gil-Moreno, A., Abal, M., Monge, M., García, A., Baró, T., Thomson, T. M., Xercavins, J., Alameda, F., & Reventós, J. (2004). A differential gene expression profile reveals overexpression of RUNX1/AML1 in invasive endometrioid carcinoma. *Cancer Research*, *64*(24), 8846–8853. https://doi.org/10.1158/0008-5472.CAN-04-2066

956 Planagumà, J., Gonzalez, M., Doll, A., Monge, M., Gil-Moreno, A., Baró, T., García, A., Xercavins,
957 J., Alameda, F., Abal, M., & Reventós, J. (2006). The up-regulation profiles of
958 p21WAF1/CIP1 and RUNX1/AML1 correlate with myometrial infiltration in endometrioid
959 endometrial carcinoma. *Human Pathology*, *37*(8), 1050–1057.
960 https://doi.org/10.1016/j.humpath.2006.03.007
961 Richardson, G. D., Robson, C. N., Lang, S. H., Neal, D. E., Maitland, N. J., & Collins, A. T. (2004).
962 CD133, a novel marker for human prostatic epithelial stem cells. *Journal of Cell Science*,
963 *117*(16), 3539–3545. https://doi.org/10.1242/jcs.01222
964 Riggio, A. I., & Blyth, K. (2017). The enigmatic role of RUNX1 in female-related cancers—Current
965 knowledge & future perspectives. *The FEBS Journal*, *284*(15), 2345–2362.
966 https://doi.org/10.1111/febs.14059
967 Samokhvalov, I. M., Samokhvalova, N. I., & Nishikawa, S. (2007). Cell tracing shows the
968 contribution of the yolk sac to adult haematopoiesis. *Nature*, *446*(7139), 1056–1061.
969 https://doi.org/10.1038/nature05725
970 Scheitz, C. J. F., Lee, T. S., McDermitt, D. J., & Tumbar, T. (2012). Defining a tissue stem cell-
971 driven Runx1/Stat3 signalling axis in epithelial cancer. *The EMBO Journal*, *31*(21), 4124–
972 4139. https://doi.org/10.1038/emboj.2012.270
973 Scheitz, C. J. F., & Tumbar, T. (2013). New insights into the role of Runx1 in epithelial stem cell
974 biology and pathology. *Journal of Cellular Biochemistry*, *114*(5), 985–993.
975 https://doi.org/10.1002/jcb.24453

976 Shen, M. M., & Abate-Shen, C. (2010). Molecular genetics of prostate cancer: New prospects for
977 old challenges. *Genes & Development*, *24*(18), 1967–2000.
978 https://doi.org/10.1101/gad.1965810
979 Sroczynska, P., Lancrin, C., Kouskoff, V., & Lacaud, G. (2009). The differential activities of Runx1
980 promoters define milestones during embryonic hematopoiesis. *Blood*, *114*(26), 5279–
981 5289. https://doi.org/10.1182/blood-2009-05-222307
982 Sugimura, Y., Cunha, G. R., & Donjacour, A. A. (1986a). Morphogenesis of ductal networks in the
983 mouse prostate. *Biology of Reproduction*, *34*(5), 961–971.
984 https://doi.org/10.1095/biolreprod34.5.961
985 Sugimura, Y., Cunha, G. R., & Donjacour, A. A. (1986b). Morphological and histological study of
986 castration-induced degeneration and androgen-induced regeneration in the mouse
987 prostate. *Biology of Reproduction*, *34*(5), 973–983.
988 Takayama, K., Suzuki, T., Tsutsumi, S., Fujimura, T., Urano, T., Takahashi, S., Homma, Y.,
989 Aburatani, H., & Inoue, S. (2015). RUNX1, an androgen- and EZH2-regulated gene, has
990 differential roles in AR-dependent and -independent prostate cancer. *Oncotarget*, *6*(4),
991 2263–2276. https://doi.org/10.18632/oncotarget.2949
992 Thambyrajah, R., Mazan, M., Patel, R., Moignard, V., Stefanska, M., Marinopoulou, E., Li, Y.,
993 Lancrin, C., Clapes, T., Möröy, T., Robin, C., Miller, C., Cowley, S., Göttgens, B., Kouskoff,
994 V., & Lacaud, G. (2016). GFI1 proteins orchestrate the emergence of haematopoietic
995 stem cells through recruitment of LSD1. *Nature Cell Biology*, *18*(1), 21–32.
996 https://doi.org/10.1038/ncb3276

997 Tika, E., Ousset, M., Dannau, A., & Blanpain, C. (2019). Spatiotemporal regulation of
998 multipotency during prostate development. *Development*, dev.180224.
999 https://doi.org/10.1242/dev.180224
1000 Tirosh, I., Izar, B., Prakadan, S. M., Wadsworth, M. H., Treacy, D., Trombetta, J. J., Rotem, A.,
1001 Rodman, C., Lian, C., Murphy, G., Fallahi-Sichani, M., Dutton-Regester, K., Lin, J.-R.,
1002 Cohen, O., Shah, P., Lu, D., Genshaft, A. S., Hughes, T. K., Ziegler, C. G. K., … Garraway, L.
1003 A. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell
1004 RNA-seq. *Science*, *352*(6282), 189–196. https://doi.org/10.1126/science.aad0501
1005 Toivanen, R., Mohan, A., & Shen, M. M. (2016). Basal Progenitors Contribute to Repair of the
1006 Prostate Epithelium Following Induced Luminal Anoikis. *Stem Cell Reports*, *0*(0).
1007 https://doi.org/10.1016/j.stemcr.2016.03.007
1008 Toivanen, R., & Shen, M. M. (2017). Prostate organogenesis: Tissue induction, hormonal
1009 regulation and cell type specification. *Development (Cambridge, England)*, *144*(8), 1382–
1010 1398. https://doi.org/10.1242/dev.148270
1011 Tsujimura, A., Koikawa, Y., Salm, S., Takao, T., Coetzee, S., Moscatelli, D., Shapiro, E., Lepor, H.,
1012 Sun, T.-T., & Wilson, E. L. (2002). Proximal location of mouse prostate epithelial stem
1013 cells. *The Journal of Cell Biology*, *157*(7), 1257–1265.
1014 https://doi.org/10.1083/jcb.200202067
1015 van Bragt, M. P. A., Hu, X., Xie, Y., & Li, Z. (2014). RUNX1, a transcription factor mutated in
1016 breast cancer, controls the fate of ER-positive mammary luminal cells. *ELife*, *3*, e03881.
1017 https://doi.org/10.7554/eLife.03881

1018 Wang, B., Wang, X., Long, J. E., Eastham-Anderson, J., Firestein, R., & Junttila, M. R. (2015).
1019 Castration-Resistant Lgr5+ Cells Are Long-Lived Stem Cells Required for Prostatic
1020 Regeneration. *Stem Cell Reports*, *4*(5), 768–779.
1021 https://doi.org/10.1016/j.stemcr.2015.04.003
1022 Wang, X., Kruithof-de Julio, M., Economides, K. D., Walker, D., Yu, H., Halili, M. V., Hu, Y.-P.,
1023 Price, S. M., Abate-Shen, C., & Shen, M. M. (2009). A luminal epithelial stem cell that is a
1024 cell of origin for prostate cancer. *Nature*, *461*(7263), 495–500.
1025 https://doi.org/10.1038/nature08361
1026 Wang, X.-D., Wang, B.-E., Soriano, R., Zha, J., Zhang, Z., Modrusan, Z., Cunha, G. R., & Gao, W.-
1027 Q. (2007). Expression profiling of the mouse prostate after castration and hormone
1028 replacement: Implication of H-cadherin in prostate tumorigenesis. *Differentiation;*
1029 *Research in Biological Diversity*, *75*(3), 219–234. https://doi.org/10.1111/j.1432-
1030 0436.2006.00135.x
1031 Wang, Y., Hayward, S., Cao, M., Thayer, K., & Cunha, G. (2001). Cell differentiation lineage in the
1032 prostate. *Differentiation; Research in Biological Diversity*, *68*(4–5), 270–279.
1033 Wolf, F. A., Hamey, F. K., Plass, M., Solana, J., Dahlin, J. S., Göttgens, B., Rajewsky, N., Simon, L.,
1034 & Theis, F. J. (2019). PAGA: Graph abstraction reconciles clustering with trajectory
1035 inference through a topology preserving map of single cells. *Genome Biology*, *20*(1), 59.
1036 https://doi.org/10.1186/s13059-019-1663-x
1037 Xin, L., Lawson, D. A., & Witte, O. N. (2005). The Sca-1 cell surface marker enriches for a
1038 prostate-regenerating cell subpopulation that can initiate prostate tumorigenesis.

Proceedings of the National Academy of Sciences of the United States of America, *102*(19), 6942–6947. https://doi.org/10.1073/pnas.0502320102

Yokomizo, T., Yamada-Inagawa, T., Yzaguirre, A. D., Chen, M. J., Speck, N. A., & Dzierzak, E. (2012). Whole-mount three-dimensional imaging of internally localized immunostained cells within mouse embryos. *Nature Protocols*, *7*(3), 421–431. https://doi.org/10.1038/nprot.2011.441

Yoo, Y. A., Roh, M., Naseem, A. F., Lysy, B., Desouki, M. M., Unno, K., & Abdulkadir, S. A. (2016). Bmi1 marks distinct castration-resistant luminal progenitor cells competent for prostate regeneration and tumour initiation. *Nature Communications*, *7*, 12943. https://doi.org/10.1038/ncomms12943

Zhang, D., Zhao, S., Li, X., Kirk, J. S., & Tang, D. G. (2018). Prostate Luminal Progenitor Cells in Development and Cancer. *Trends in Cancer*, *4*(11), 769–783. https://doi.org/10.1016/j.trecan.2018.09.003

TABLES

Table 1. List of primary antibodies used for immunohistochemistry

Target	Provider	Catalog Number	Raised specie
RUNX1	Cell Signalling	8529	Rabbit
NKX3.1	Athenaes	AES-0314	Rabbit
CDH1	Thermo Fischer	AF748	Goat
TP63	Cell Signalling	M3562	Rabbit
K5	Abcam	ab52635	Rabbit
K8	Abcam	ab53280	Rabbit
K4	Abcam	Ab9004	Mouse
LY6D	Proteintech	17361-1-AP	Rabbit
TROP-2	R&D	AF1122	Goat
BrdU	Abcam	ab6326	Rat
Ki67	Abcam	ab15580	Rabbit
RFP	Rockland	600-402-379	Rabbit
RFP	MBL	PM005	Rabbit
GFP	MBL	598	Rabbit

1060 **Table 2. List of primary antibodies used for immunohistochemistry**

Antibody	Provider	Catalog Number
EnVision+/HRP Anti-Rabbit	Dako (Agilent)	K4003
EnVision+/HRP Anti-Rabbit	Dako (Agilent)	K4001
ImmPRESS HRP Anti-Goat	Vector Laboratories	MP-7405
ImmPRESS HRP Anti-Rat	Vector Laboratories	MP-7444
Donkey anti-Goat IgG 647	ThermoFischer Scientific	A-21447

1061

1062 **Table 3. List of antibodies used for Fluorescence-Activated Cell Sorting (FACS)**

Target	Provider	Catalog Number	Clone	Fluorophore
CD16/32	Biolegend	101301	93	none (Fc block)
CD45	Biolegend	62-0451-82	30-F11	SB436
EPCAM	Biolegend	118225	G8.8	BV421
EPCAM	Biolegend	118214	G8.8	APC
CD49f	Biolegend	313606	GoH3	FITC
CD49f	Biolegend	313616	GoH3	APC
CD24	BD Biosciences	744470	M1/69	BV786

1063

FIGURE SUPPLEMENTS

Figure 1—figure supplement 1. RUNX1 is enriched in the mouse prostate epithelium.

Figure 1—figure supplement 2. *Runx1* expression is mediated by the P2 promoter in the mouse prostate epithelium.

Figure 1—figure supplement 3. Characterization of P2-*Runx1*:RFP derived mouse prostate organoids.

Figure 2—figure supplement 4. Characterization of RUNX1 expression in the castrated mouse prostate.

Figure 3—figure supplement 5. Pre-processing of the scRNA-seq dataset of adult intact and castrated mouse prostates.

Figure 3—figure supplement 6. Characterization of the scRNA-seq prostate epithelial subset.

Figure 3—figure supplement 7. Characterization of the scRNA-seq prostate epithelial dataset.

Figure 3—figure supplement 8. Gene Ontology and differential expression analysis within the scRNA-seq prostate epithelial dataset.

Figure 4—figure supplement 9. Lineage tracing of RUNX1 expressing cells labelled in intact mice.

Figure 5—figure supplement 10. Characterization of RUNX1 expression during prostate development *in vivo*.

Figure 5—figure supplement 11. Characterization of RUNX1 expression during prostate development in UGS explant cultures.

Figure 6—figure supplement 12. Pre-processing of the scRNA-seq dataset of UGS explant cultures.

Figure 7—figure supplement 14. Lineage-tracing of RUNX1 expressing cells in UGS explants.

Figure 6—figure supplement 13. Characterization of the developing prostatic epithelium in the scRNA-seq dataset of UGS explant culture.

SUPPLEMENTARY FILES

Supplementary file 1. Quality control metrics and metadata of scRNAseq experiments

Supplementary file 2. Genes differentially expressed between adult clusters

Supplementary file 3. Genes differentially expressed in intact versus castrated Lum-D cells

Supplementary file 4. Genes differentially expressed between UGS explants clusters

Supplementary file 5. Gene lists used for the analysis of gene set activity

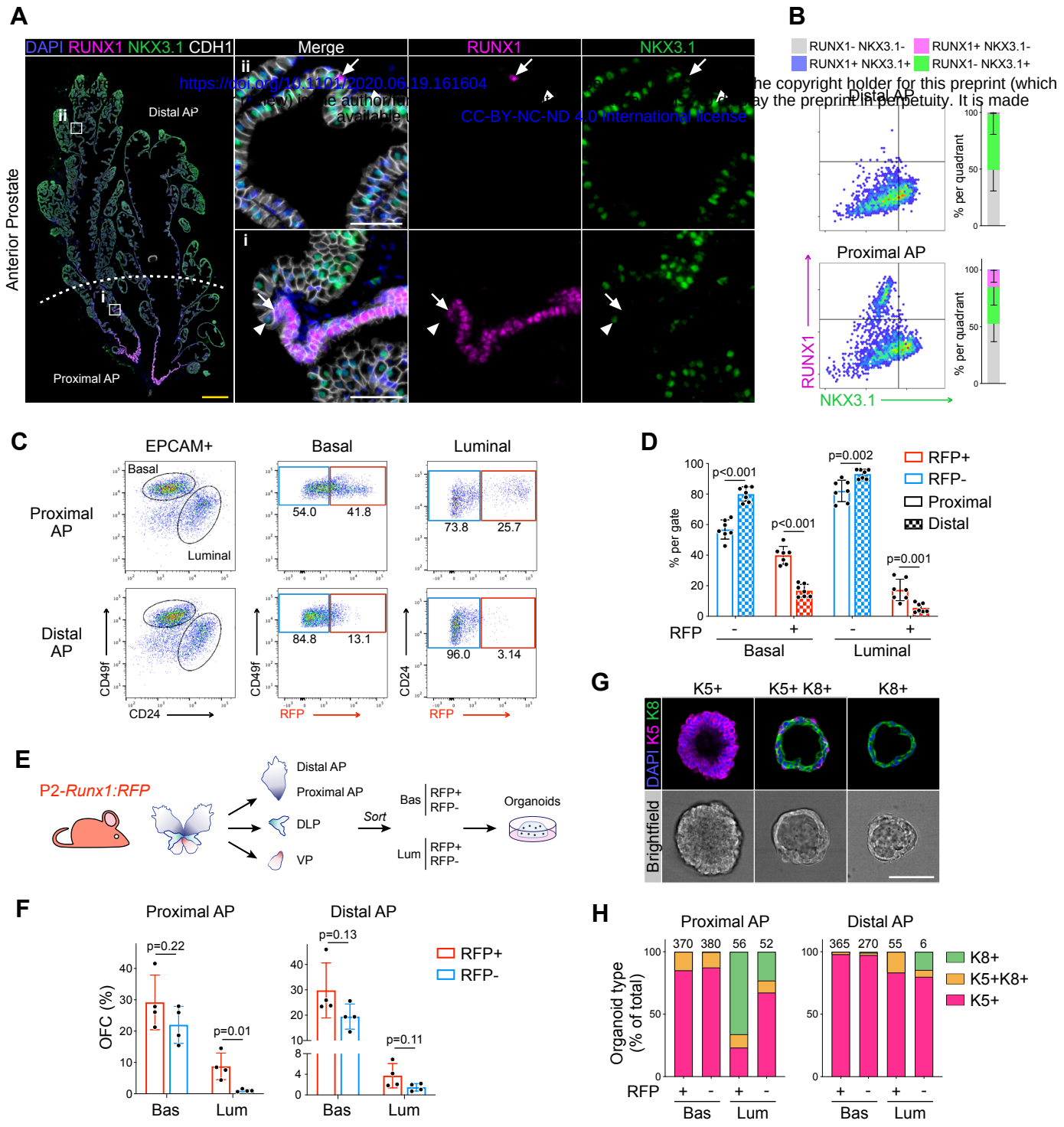

Figure 1. RUNX1 marks a subpopulation of mouse proximal prostate luminal cells (PLCs).

(**A**) Co-immunostaining of RUNX1, NKX3.1, CDH1 in the mouse Anterior Prostate (AP). Higher magnification images of (i) proximal AP and (ii) distal AP are shown. Arrows indicate RUNX1$^+$ NKX3.1$^-$ cells, arrowheads show RUNX1$^-$ NKX3.1$^+$ cells. Scale bars: 500µm (yellow) and 50µm (white). (**B**) Quantification of RUNX1 and NKX3.1 nuclear intensity (log10) in CDH1$^+$ epithelial cells by QBIC in proximal and distal AP. *n* = 6-8 mice. (**C, D**) Flow-cytometry analysis of P2-*Runx1*:RFP mice, and corresponding quantification of the percentages of RFP$^+$ and RFP$^-$ cells in the basal and luminal fractions of the proximal and distal AP. *n* = 7 mice. (**E**) Experimental strategy to grow organoids from sorted RFP$^+$ and RFP$^-$ cells from the basal (CD49f^{high}) and luminal (CD24^{high}) lineages of P2-*Runx1*:RFP mouse reporters. (**F**) Organoid Forming Capacity (OFC) of RFP$^+$ and RFP$^-$ basal and luminal sorted cells after 7 days in culture. *n* = 4 mice. (**G**) Whole-mount immunostaining of unipotent K5$^+$, unipotent K8$^+$ or multipotent K5$^+$ K8$^+$ organoids. Scale bar: 50µm. (**H**) Quantification of the type of organoids characterized by whole-mount immunostaining, as in (**G**). Numbers of organoids quantified are shown above the graph. *n* = 2 mice per group.

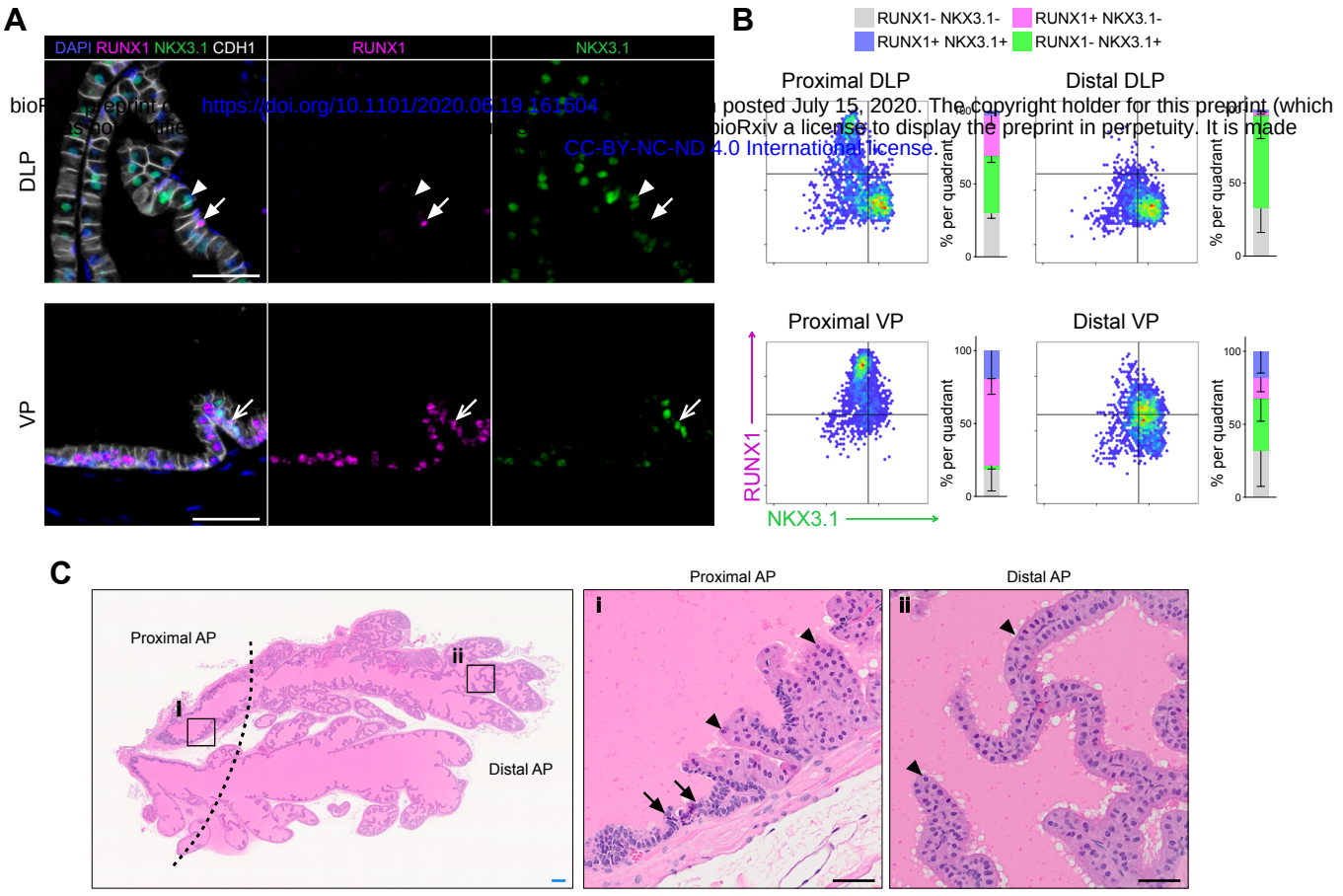

Figure 1—figure supplement 1. RUNX1 is enriched in the mouse prostate epithelium.

(**A**) Co-immunostaining of RUNX1, NKX3.1, CDH1 in the mouse DLP (top) and VP (bottom). Closed arrows indicate RUNX1^{+} NKX3.1^{-} cells, chevron arrows indicate rare RUNX1^{+} NKX3.1^{+} cells, arrowheads show RUNX1^{-} NKX3.1^{+} cells. Scale bars: 50µm. (**B**) Quantification of RUNX1 and NKX3.1 nuclear intensity (log10) in CDH1+ epithelial cells by QBIC in proximal DLP (*n* = 3), distal DLP (*n* = 8), proximal VP (*n* = 3) and distal VP (*n* = 8). (**C**) H&E staining of the mouse AP. Higher magnification images of (i) proximal AP and (ii) distal regions are shown. Arrows show typical proximal luminal cells with condensed nuclei and cytoplasm, arrowheads show typical distal secretory luminal cells with enlarged cytoplasm and strong pink eosinophilic staining. Scale bars: 200µm (blue) and 50µm (black).

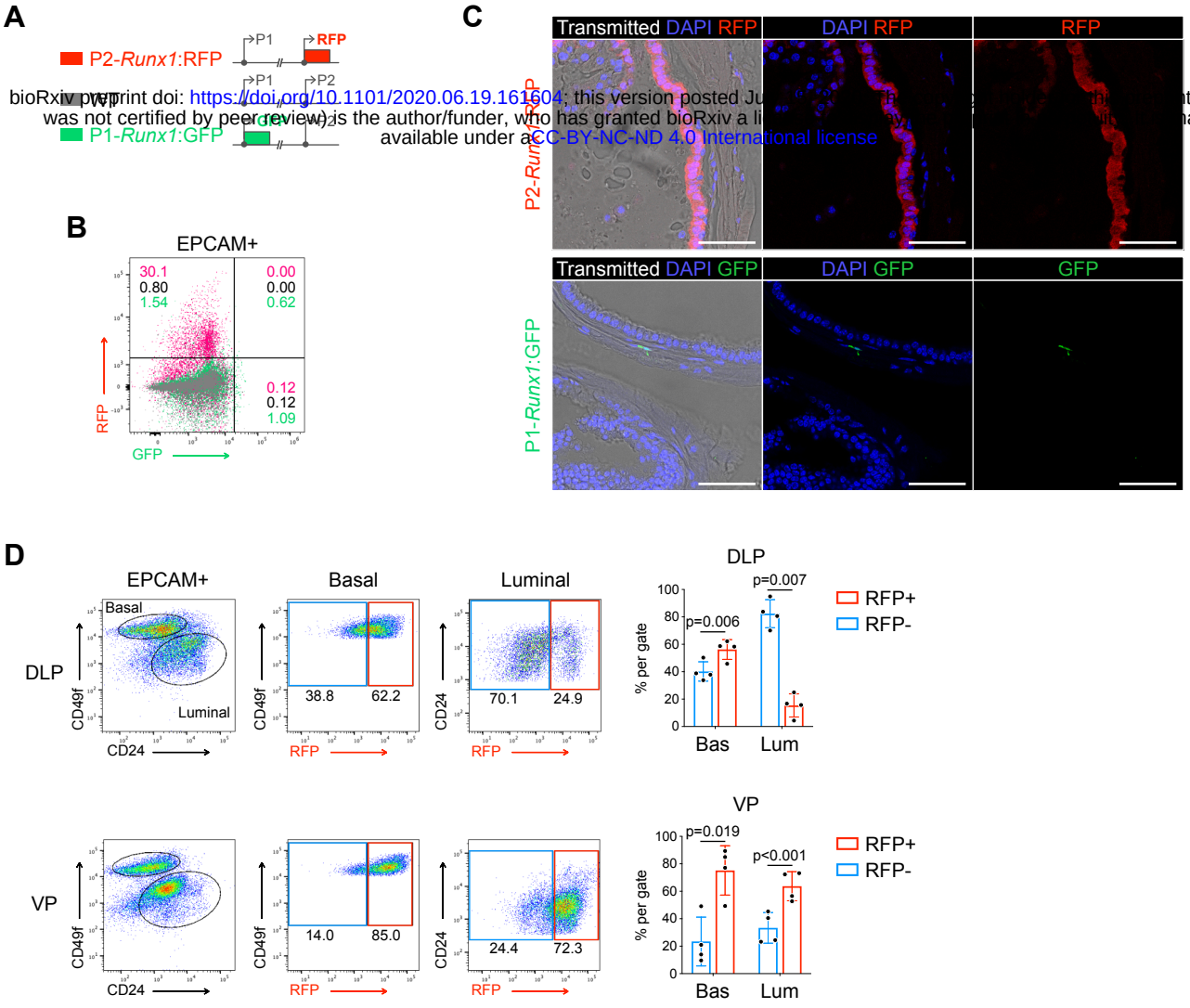

Figure 1—figure supplement 2. *Runx1* expression is mediated by the P2 promoter in the mouse prostate epithelium.

(**A**) Schematic diagrams of the P2-*Runx1*:RFP reporter allele (top), Runx1 WT allele (middle), and P1-*Runx1*:GFP reporter allele (bottom). In P2-*Runx1*:RFP or P1-*Runx1*:GFP mice, expression of RFP or GFP is directed by *Runx1* proximal P2 or distal P1 promoter respectively. WT *Runx1* expression is directed by the remaining WT allele. (**B**) Flow-cytometry analysis of the epithelial EPCAM$^+$ fraction of WT, P2-*Runx1*:RFP and P1-*Runx1*:GFP mice. Percentages are indicated in each quadrant gates. (**C**) Confocal images indicating P2-*Runx1*:RFP signal in the glandular epithelium (top), while P1-*Runx1*:GFP (bottom) activity is detected in the surrounding stroma. Scale bar: 50μm. (**D**) Flow-cytometry analysis of P2-*Runx1*:RFP mice, and corresponding quantification of the percentages of RFP$^+$ and RFP$^-$ cells in the basal and luminal fractions of DLP (n = 4 mice) and VP (n = 4 mice) lobes.

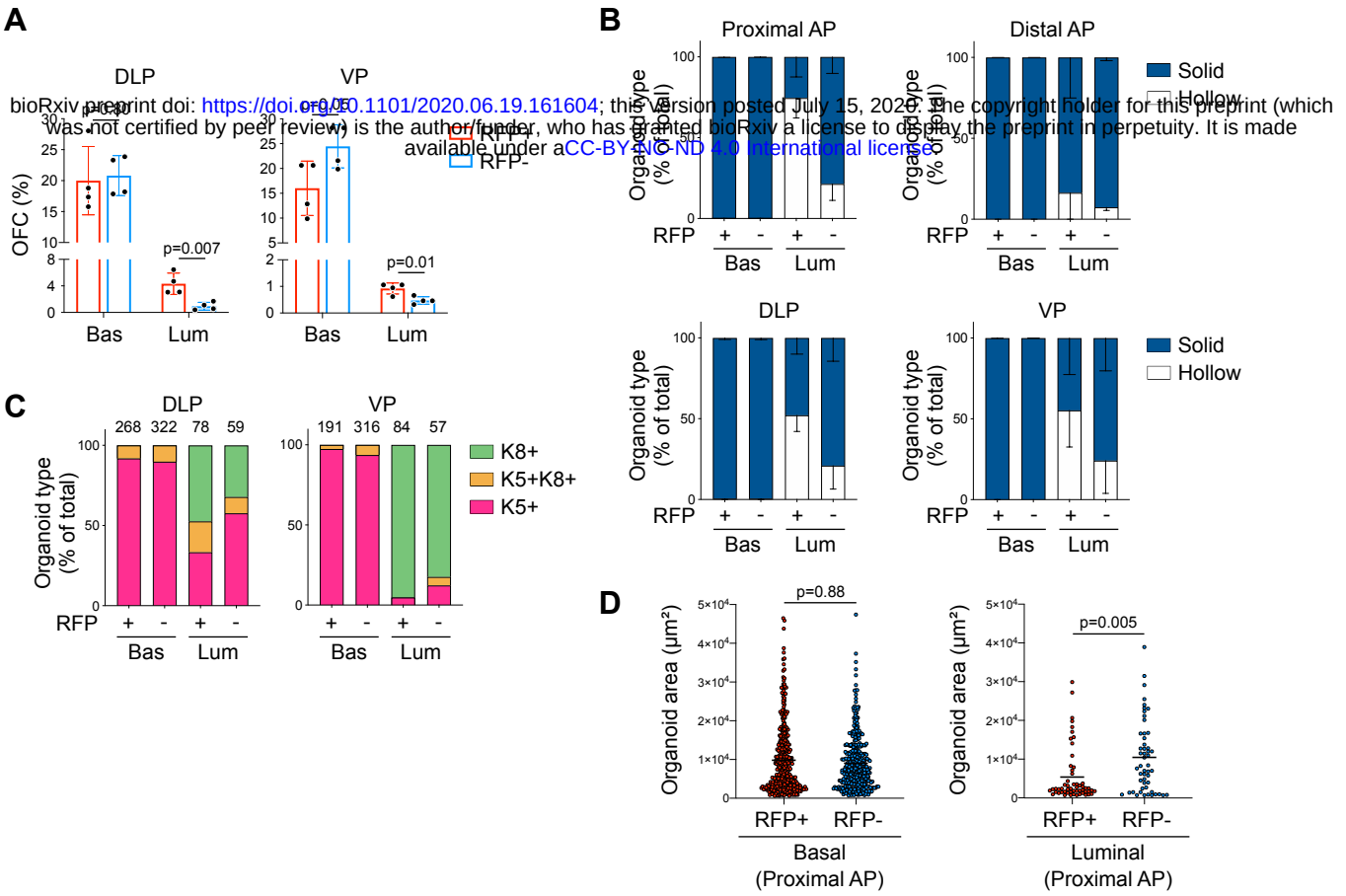

Figure 1—figure supplement 3. Characterization of P2-*Runx1*:RFP derived mouse prostate organoids.

(**A**) Organoid Forming Capacity (OFC) of RFP$^+$ and RFP$^-$ basal and luminal sorted cells from DLP (left) and VP (right) cultured for 7 days. *n* = 4 independent experiments. (**B**) Quantification of organoid morphologies determined by brightfield assessment. *n* = 4 mice. (**C**) Quantification of the type of organoids (pink: unipotent K5$^+$, green: unipotent K8$^+$, orange: multipotent K5$^+$ K8$^+$ organoids) characterized by whole-mount immunostaining. *n* = 2 mice. (**D**) Quantification of the organoid area estimated from maximum projection of stacked images from basal (left) and luminal (right) fractions of the proximal AP. Numbers of organoids quantified are shown above the graph. *n* = 2 mice.

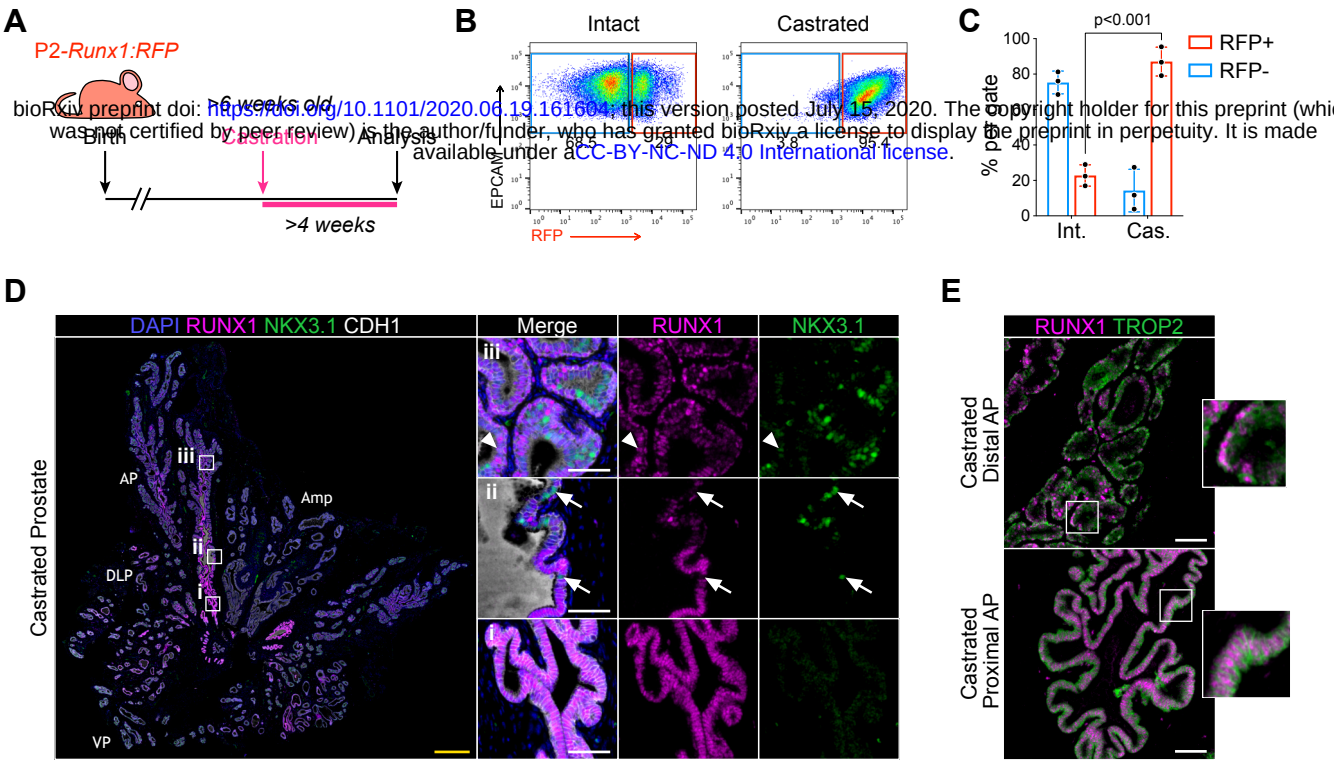

Figure 2. RUNX1 expressing cells are enriched in the castrated prostate epithelium.

(**A**) P2-*Runx1*:RFP reporter mice were surgically castrated between 6 to 12 weeks of age and analyzed at least 4 weeks post-castration. (**B**, **C**) Flow-cytometry analysis and corresponding quantification of the proportion of RFP$^+$ and RFP$^-$ cells in the EPCAM$^+$ fraction of intact and castrated prostates of P2-*Runx1*:RFP mice. *n* = 3 mice per group. Int: Intact, Cas: Castrated. (**D**) Co-immunostaining of RUNX1, NKX3.1, CDH1 in the castrated wild-type mouse prostate. Higher magnification images of (i) proximal, (ii) intermediate and (iii) distal AP are shown. Arrows indicate cells RUNX1$^-$ NKX3.1$^+$ cells, arrowheads show luminal cell co-stained for RUNX1 and NKX3.1. Amp: ampullary gland. Scale bars: 500µm (yellow) and 50µm (white). Int: Intact, Cas: Castrated. (**E**) Co-immunostaining of RUNX1 and TROP2 showing colocalization of the 2 markers in both proximal (bottom) and distal (top) castrated AP. Scale bars: 50µm (white).

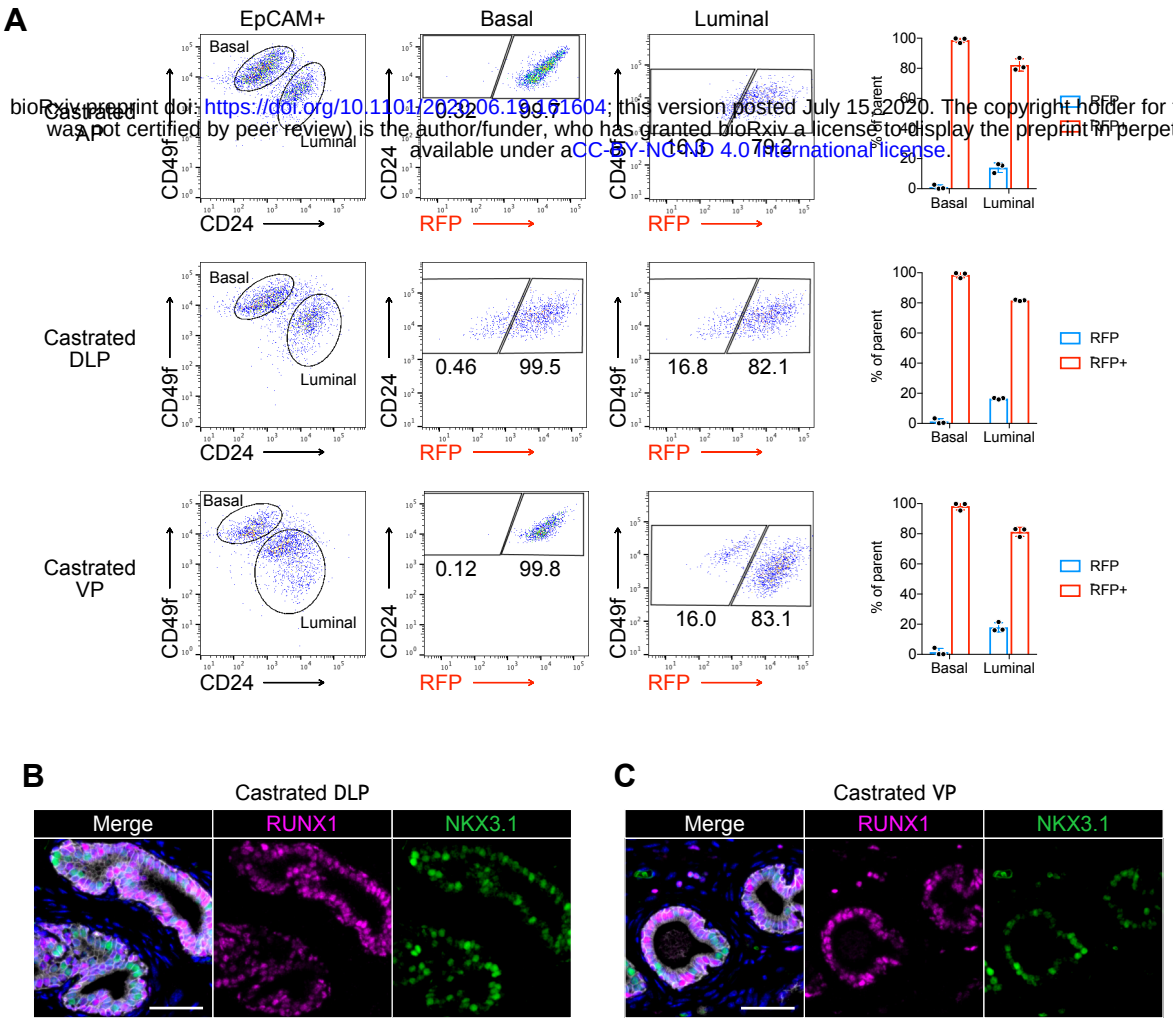

Figure 2—figure supplement 4. Characterization of RUNX1 expression in the castrated mouse prostate.

(**A**) Flow cytometry analysis and corresponding quantifications of the basal and luminal EPCAM^{+} fraction from AP, DLP and VP lobes of castrated P2-*Runx1*:RFP mice (n = 3). (**B, C**) Co-immunostaining of RUNX1, NKX3.1, CDH1 in castrated DLP (**B**) and VP (**C**) lobes. Scale bars: 50µm (white).

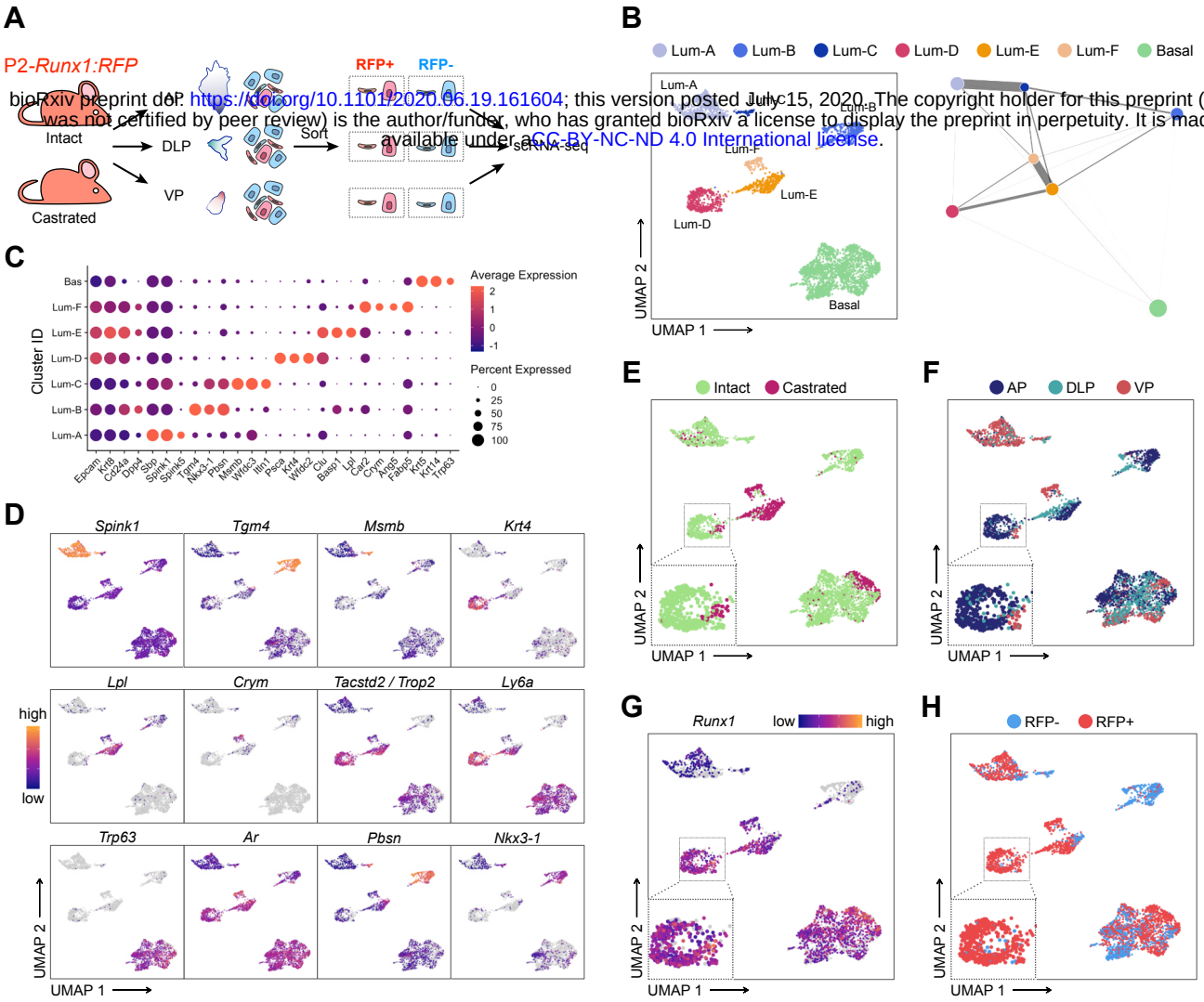

Figure 3. scRNA-seq profiling of intact and castrated *Runx1*⁺ cells reveals transcriptomic similarity between proximal luminal cells and castration-resistant cells.

(**A**) Experimental strategy for scRNA-seq on RFP⁺ and RFP⁻ cells individually dissected lobes of intact and castrated prostates isolated from P2-*Runx1*:RFP reporter mice. (**B**) UMAP visualization (left) and graph-abstracted representation (PAGA, right) of prostate epithelial cells (*n* = 3,825 cells from 3 independent experiments). Colors represent different clusters. In PAGA, clusters are linked by weighted edges that represent a statistical measure of connectivity. (**C**) Dot plot showing the expression of selected marker genes associated with each cluster. (**D-H**) UMAP visualization of prostate epithelial cells. Cells in **D** and **G** are colored by a gradient of log-normalized expression levels for each gene indicated. Cell colors in **E** represent the treatment of origin (intact, castrated), in **F** individual lobes of origin (AP, DLP, VP), and in (**H**) RFP FACS gate of origin (RFP⁺, RFP⁻).

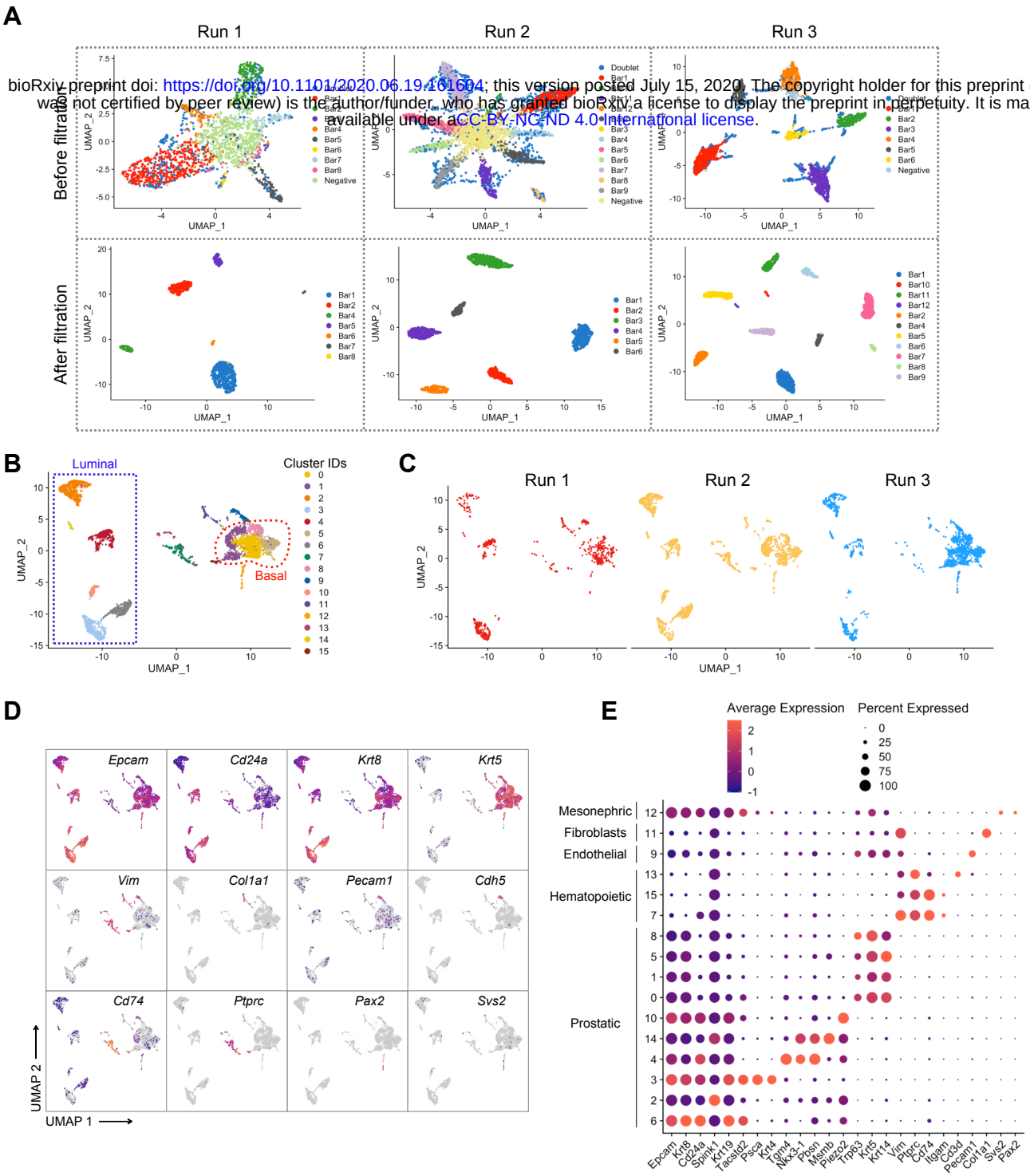

Figure 3—figure supplement 5. Pre-processing of the scRNA-seq dataset of adult intact and castrated mouse prostates.

(**A**) UMAP visualization in MULTI-seq barcodes space for each independent experiment (Run1: left, Run2: middle, Run 3: right) before (top) and after (bottom) quality control and barcode filtration (see Methods for details). (**B**) UMAP visualization of the integrated batch-corrected dataset (*n* = 4,499 cells from 3 independent experiments). Colors represent different clusters. (**C**) UMAP visualization of the integrated batch-corrected dataset split and colored by individual experiment (Run1: left, Run2: middle, Run 3: right). (**D**, **E**) Example key marker genes used for cell type characterization. (**D**) UMAP visualization of the integrated batch-corrected dataset. Cells are colored by a gradient of log-normalized expression levels for each gene indicated. (**E**) Dot plot showing the expression of selected marker genes associated with each cluster. Cells corresponding to the prostatic clusters were kept for downstream analyses.

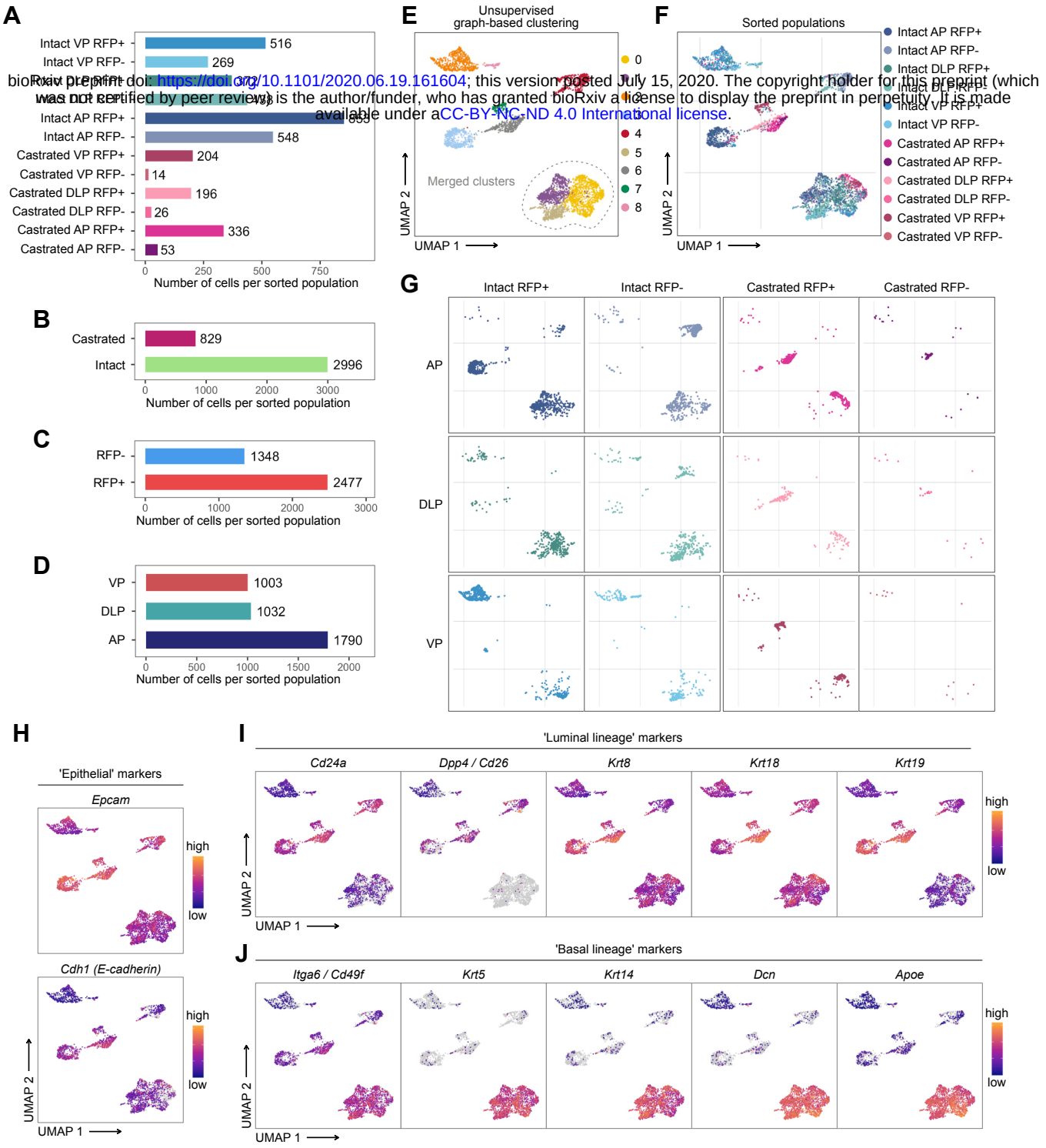

Figure 3—figure supplement 6. Characterization of the scRNA-seq prostate epithelial subset.

(**A**-**D**) Number of cells present in the prostate epithelial subset. In **A** colors correspond to phenotypic sorted populations, in **B** treatment of origin, in **C** RFP FACS gate of origin, in **D** individual lobes. (**E, G**) UMAP visualization of prostate epithelial cells. Cells in **E** are colored by clusters. Clusters 0, 1 and 5 were merged together. In **F** and **G** colors represent phenotypic sorted populations. (**H-J**) Cells are colored by a gradient of log-normalized expression levels for each gene. Canonical epithelial markers are shown in **H**, luminal lineage markers in **I**, and basal lineage markers in **J**.

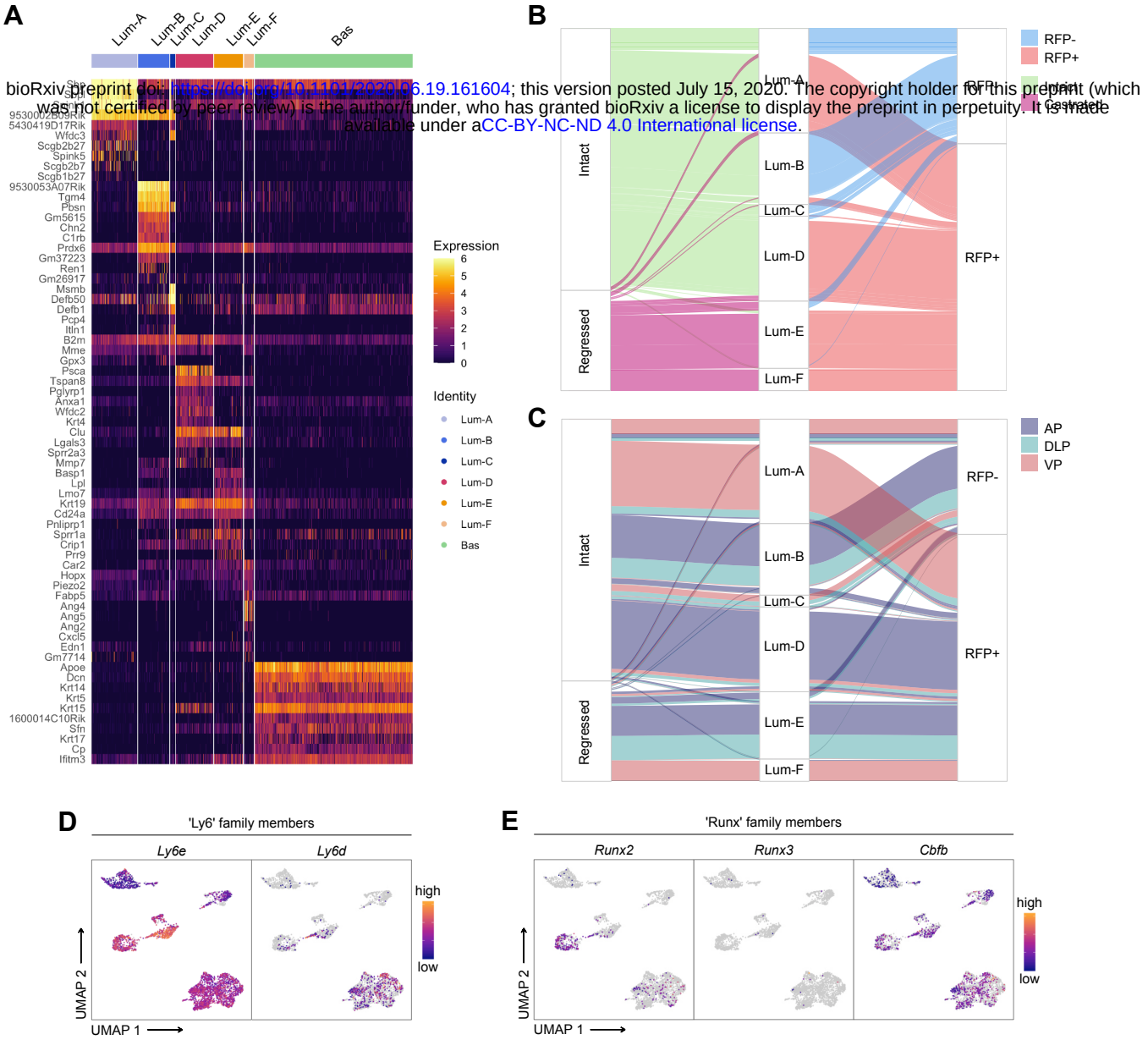

Figure 3—figure supplement 7. Characterization of the scRNA-seq prostate epithelial dataset.

(**A**) Heatmap of the 10 most differentially upregulated genes per cluster. (**B**, **C**) Sankey diagrams summarizing the proportion of cells and between phenotypic groups. (**D**, **E**) Cells are colored by a gradient of log-normalized expression levels for each gene. (**D**) shows the expression of Ly6 family members, and in **E** RUNX family members are shown.

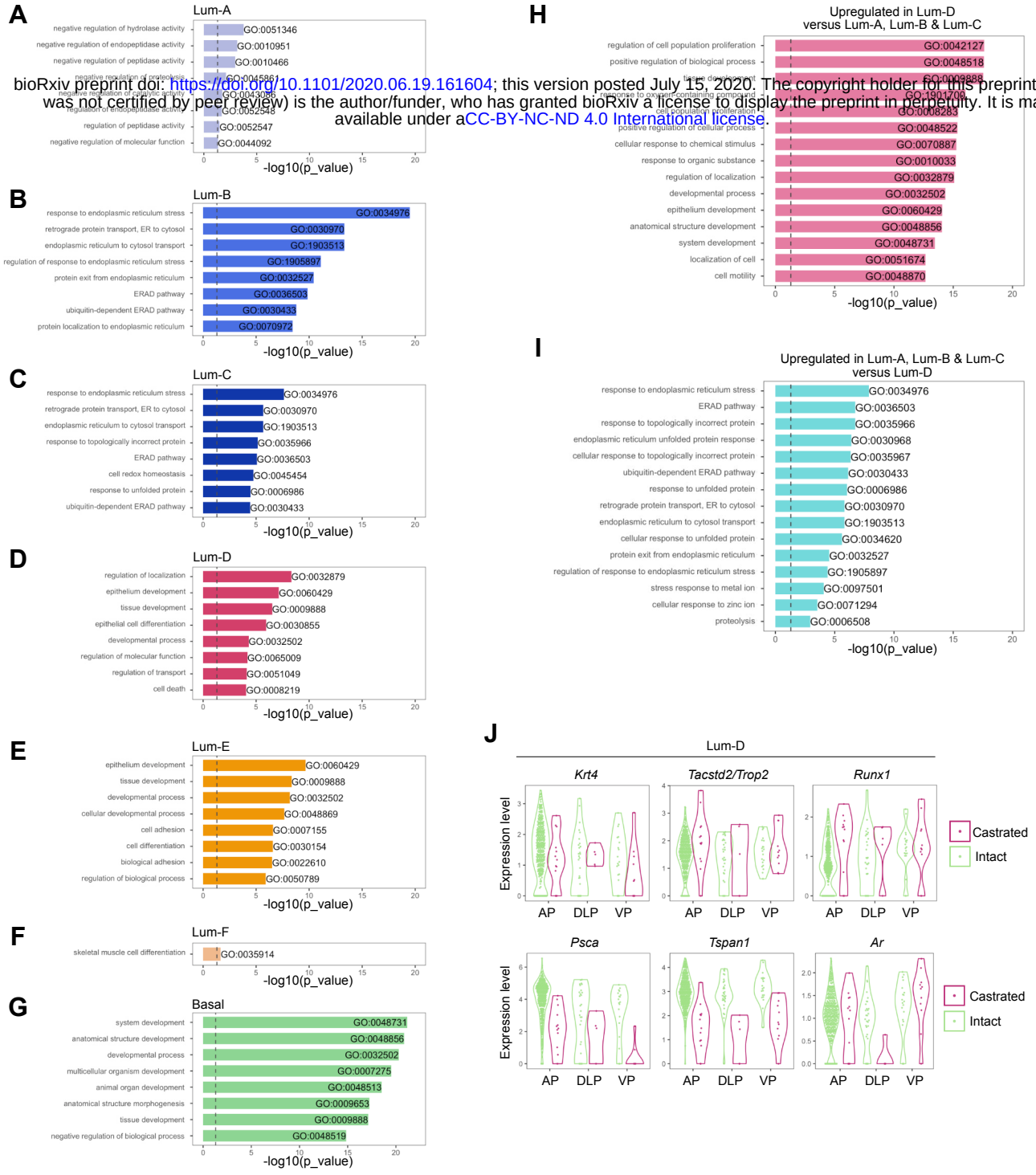

Figure 3—figure supplement 8. Gene Ontology and differential expression analysis within the scRNA-seq prostate epithelial dataset.

Bar plots of the 8 most significantly (g:Profiler adjusted p-value < 0.05) enriched gene ontology (GO) terms (GO:BP, Biological Processes) using the list of differentially upregulated genes specific to (**A**) Lum-A, (**B**) Lum-B, (**C**) Lum-C, (**D**) Lum-D, (**E**) Lum-E, (**F**) Lum-F, and (**G**) Basal clusters. (**H, I**) 15 most enriched GO terms (g:Profiler adjusted p-value < 0.05) for genes upregulated in the Lum-D cluster against combined Lum-A, Lum-B and Lum-C clusters in **H**. Results of the opposite comparison are shown in **I**. (**J**) Violin plots showing the expression levels of specific genes within the Lum-D cluster. Data are split and colored by treatment for each lobe. The expression of *Krt4*, *Tacstd2*/*Trop2* and *Runx1* does not vary upon treatment, while Androgen Receptor (AR) regulated genes such as *Psca* and *Tspan1* are downregulated after castration.

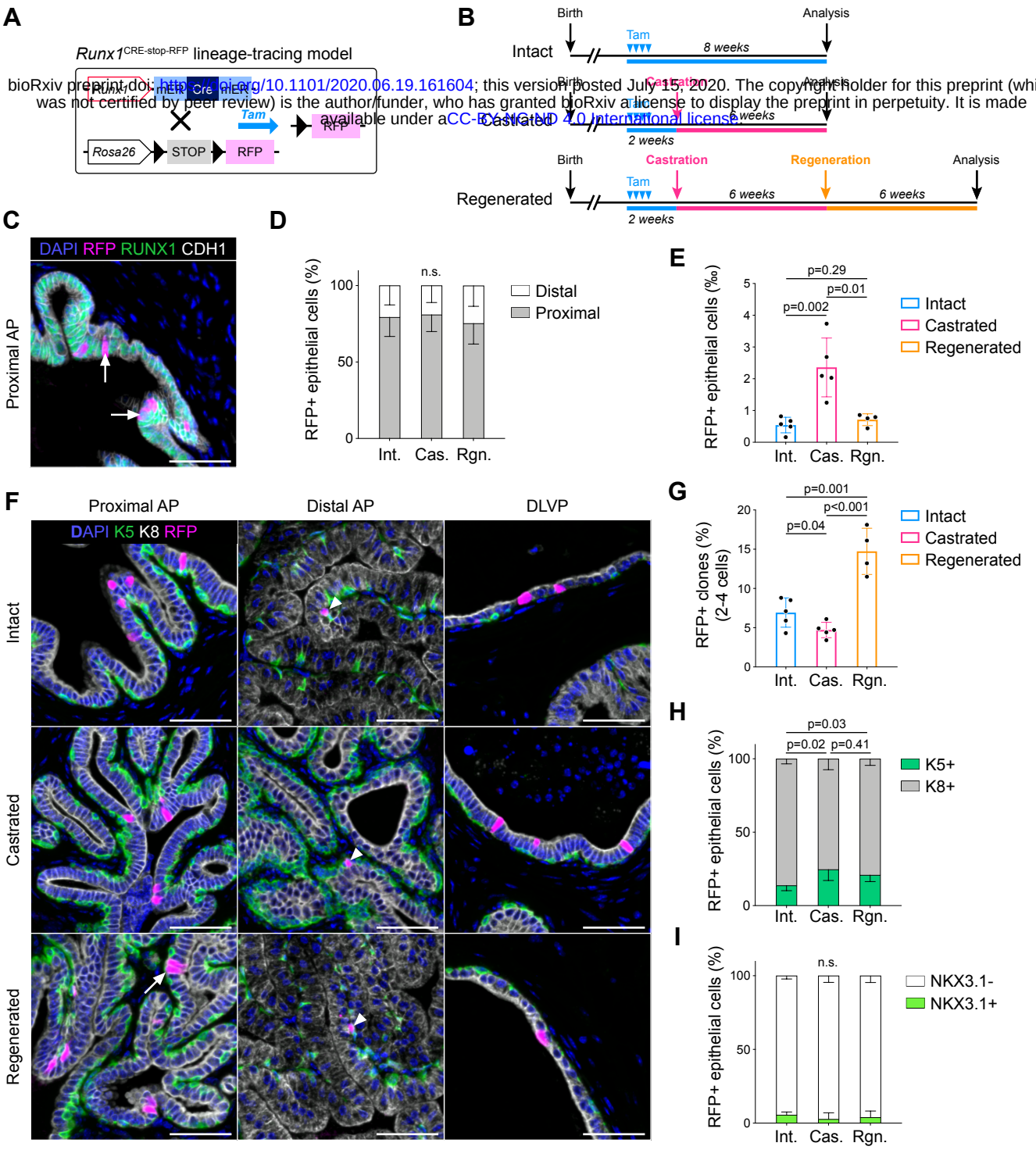

Figure 4. Lineage-tracing of *Runx1* expressing cells establishes the intrinsic castration-resistant properties of the proximal luminal lineage.

(**A**) Schematic summary of the genetic lineage tracing system employed. (**B**) Experimental strategy for lineage-tracing experiments. (**C**) Co-immunostaining of RFP, RUNX1, CDH1 in the proximal AP. Arrows indicate RFP labelled RUNX1$^+$ cells. Scale bar: 50µm. (**D**) Quantification of % of epithelial RFP$^+$ cells in proximal and distal regions of the prostate in intact (*n* = 5), castrated (*n* = 4) and regenerated (*n* = 4) mice. (**E**) Quantification of ‰ of epithelial RFP$^+$ cells in intact (*n* = 5), castrated (*n* = 5) and regenerated (*n* = 4) mice. (**F**) Co-immunostaining of RFP, K5, K8 in the proximal AP, distal AP, and DLVP (DLP + VP). Arrowheads indicate RFP labelled basal cells (K5$^+$) found in distal AP, the white arrow indicates a luminal (K8$^+$) RFP$^+$ clone made of 2 cells. Scale bar: 50µm. (**G**) Quantification of % of epithelial RFP$^+$ clones comprising between 2 to 4 cells in intact (*n* = 5), castrated (*n* = 5) and regenerated (*n* = 4) mice. (**H, I**) Quantification of % of RFP$^+$ cells being K5$^+$ or K8$^+$ in **H**, or NKX3.1$^+$ or NKX3.1$^-$ in **I,** in intact (*n* = 5), castrated (*n* = 5) and regenerated (*n* = 4) mice. Int: Intact, Cas: Castrated, Rgn: Regenerated.

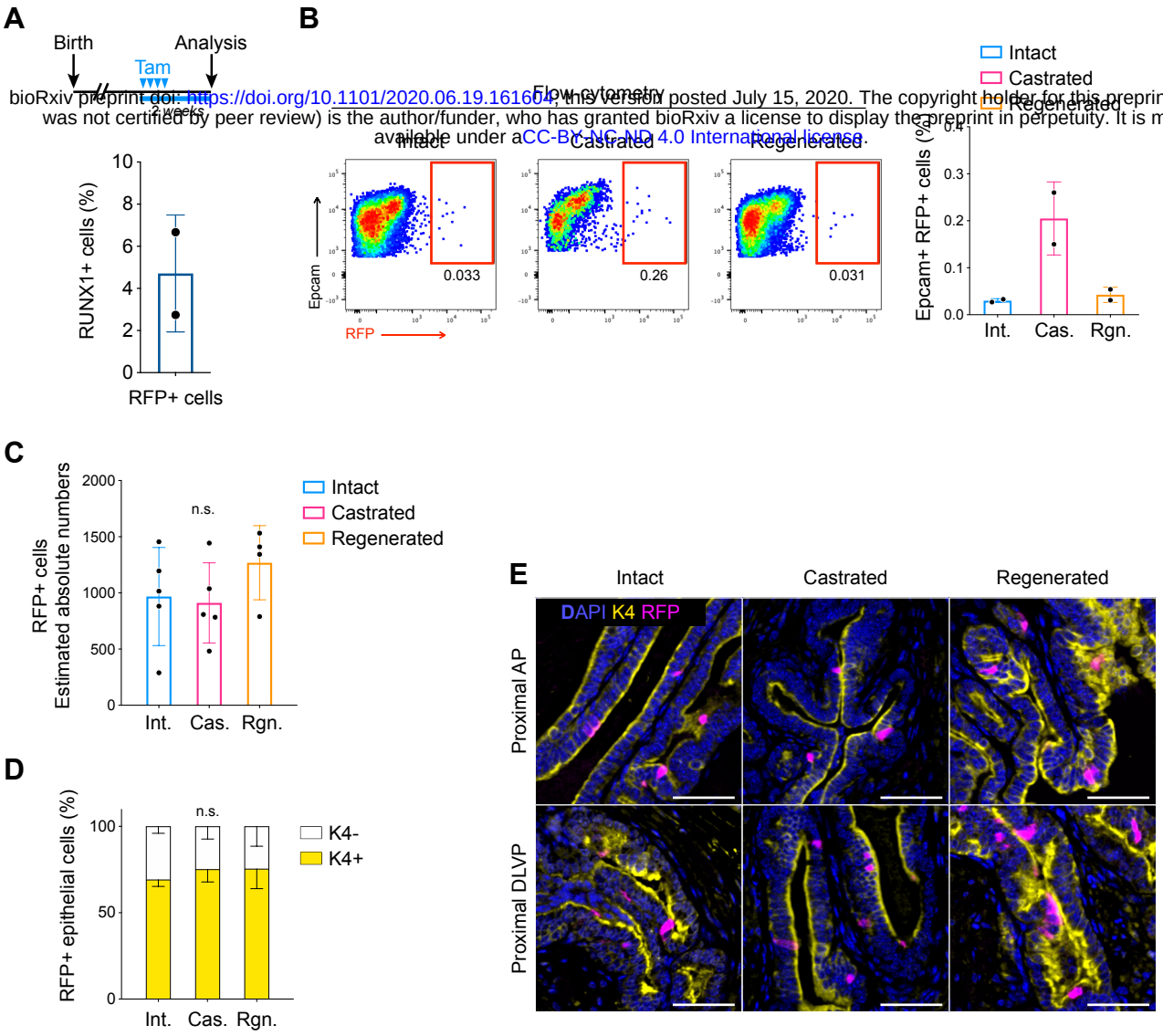

Figure 4—figure supplement 9. Lineage tracing of RUNX1 expressing cells labelled in intact mice.

(**A**) Scheme of the tamoxifen labelling strategy to evaluate the labelling efficiency of *Runx1* expressing cells. (**B**) Flow-cytometry analysis of intact (*n* = 2), castrated (*n* = 2), and regenerated (*n* = 2) *Runx1*$^{CRE\text{-}stop\text{-}RFP}$ mice, and corresponding quantification of the percentage of RFP$^+$ cells in the epithelial EPCAM$^+$ fraction. (**C**) Estimated absolute number of epithelial RFP$^+$ cells in intact (*n* = 5) and castrated (*n* = 5) and regenerated (*n* = 4) prostates based on IHC quantifications. Int: Intact, Cas: Castrated, Rgn: Regenerated. (**D**) Quantification of % of epithelial RFP$^+$ cells being K4$^+$ or K4$^-$ in intact (*n* = 3), castrated (*n* = 4) and regenerated (*n* = 5) mice. (**F**) Co-immunostaining showing the enrichment of K4$^+$ RFP labelled cells in the proximal region of the AP (top) and DLVP (bottom) prostate lobes. Scale bar: 50µm.

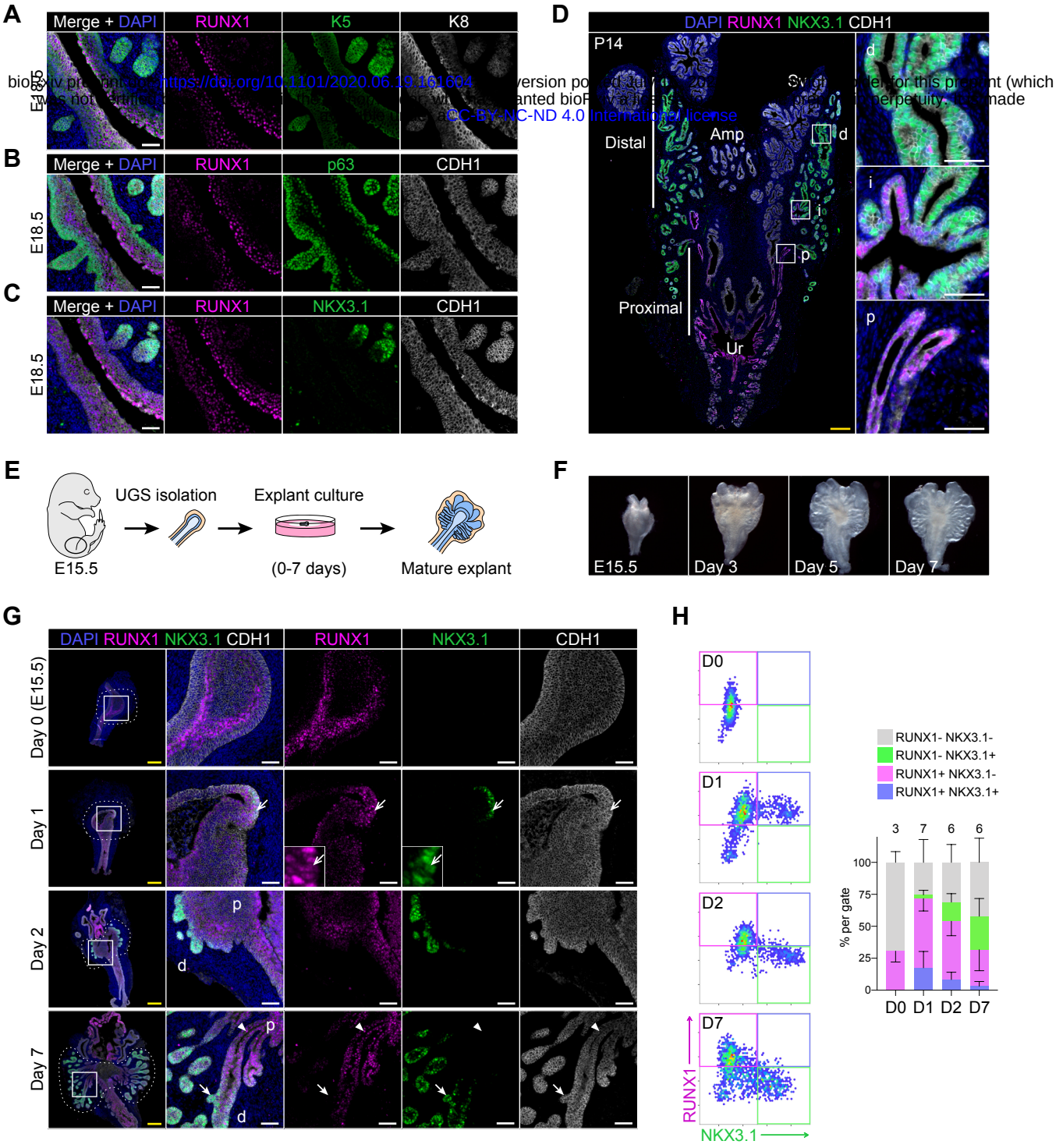

Figure 5. RUNX1 marks proximal cells during embryonic prostate development.

(**A-C**) Co-immunostainings of the mouse urogenital sinus at E18.5 for RUNX1, K5, K8 in **A**, RUNX1, p63, CDH1 in **B**, RUNX1, NKX3.1, CDH1 in **C**. Scale bar: 50µm. (**D**) Co-immunostainings of RUNX1, NKX3.1, CDH1 at postnatal (P) day 14. Higher magnification images of (p) proximal, (i) intermediate, and (d) distal regions are shown. Scale bars: 200µm (yellow) and 50µm (white). Amp: ampullary gland; Sv: seminal vesicles; Ur: urethra; p: proximal; i: intermediate; d: distal. (**E**) Scheme of the protocol to culture *ex vivo* explants of mouse UGS harvested at E15.5. (**F**) Representative images of UGS explants at E15.5 (day 0), day 3, day 5 and day 7 of culture showing the formation of premature prostate buds. (**G**) Co-immunostaining of RUNX1, NKX3.1, CDH1 in UGS explants harvested at day 0, day 1, day 2 and day 7. Higher magnification images of each square (left) are shown for each time point. Chevron arrows show RUNX1$^+$ NKX3.1$^+$ cells, closed arrows indicate RUNX1$^-$ NKX3.1$^+$ cells, arrowheads show RUNX1$^+$ NKX3.1$^-$ cells. Scale bars: 200µm (yellow) and 50µm (white). (**H**) Quantification of RUNX1 and NKX3.1 nuclear intensity (log10) in CDH1$^+$ epithelial cells of UGS explants by QBIC. Quantification was performed within the boundaries delimited in **G** by dotted lines, at day 0 (*n* = 3 explants), day 1 (*n* = 7 explants), day 2 (*n* = 6 explants) and day 7 (*n* = 6 explants).

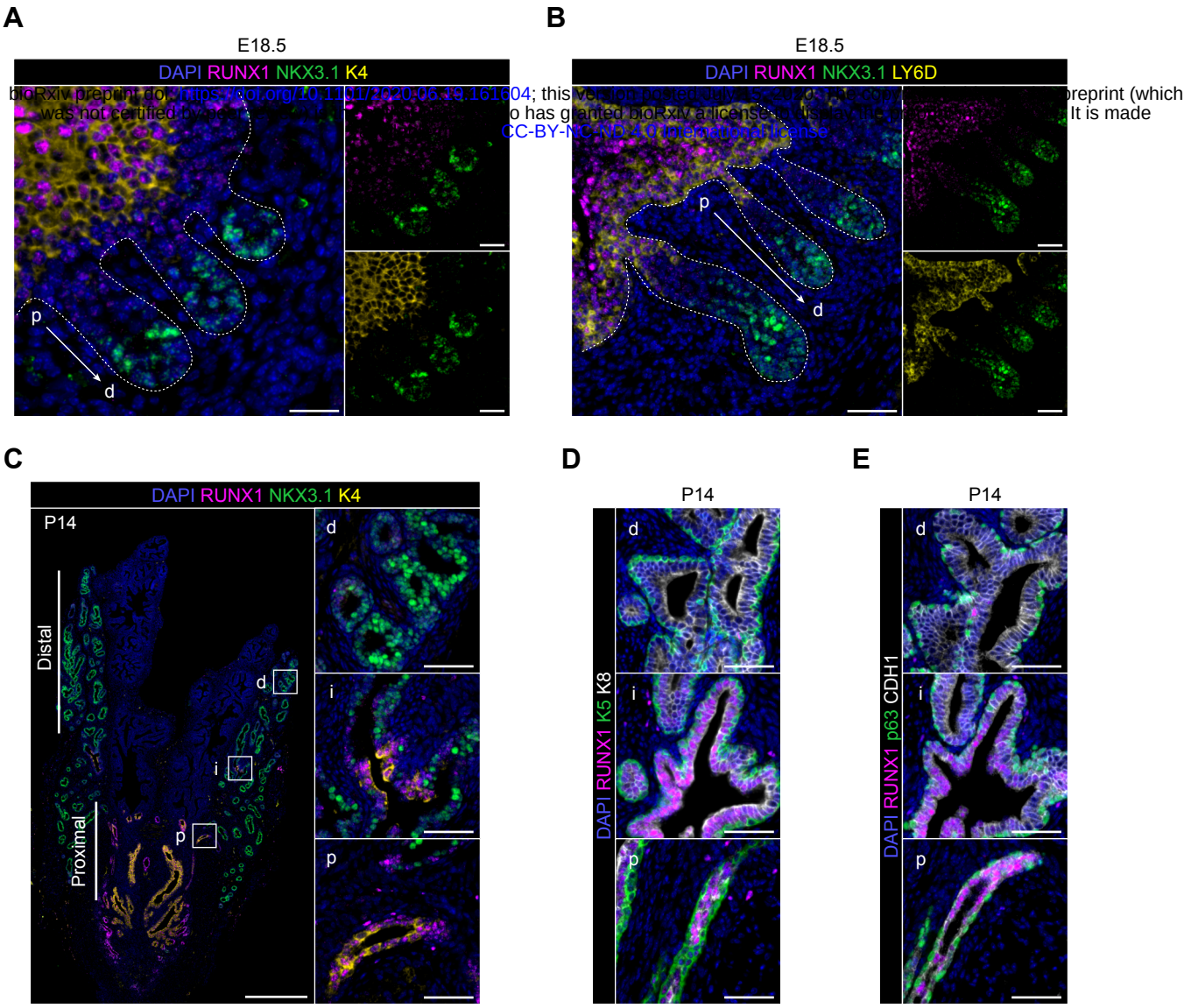

Figure 5—figure supplement 10. Characterization of RUNX1 expression during prostate development *in vivo*.

(**A**) Co-immunostainings of the mouse urogenital sinus at E18.5 for RUNX1, NKX3.1, K4. p: proximal, d: distal. Scale bar: 50µm. Dotted lines indicate the urogenital epithelium. (**B**) Co-immunostainings of the mouse urogenital sinus at E18.5 for RUNX1, NKX3.1, LY6D. p: proximal, d: distal. Scale bar: 50µm. (**C**) Co-immunostaining of RUNX1, NKX3.1, K4 at postnatal (P) day 14. Higher magnification images of (p) proximal, (i) intermediate, and (d) distal regions are shown. Scale bars: 200µm (yellow) and 50µm (white). (**D**, **E**) Co-immunostainings at postnatal P14 for RUNX1, K5, K8 in **B** RUNX1, p63, CDH1 in **C**. p: proximal, i: intermediate, d: distal. Scale bar: 50µm.

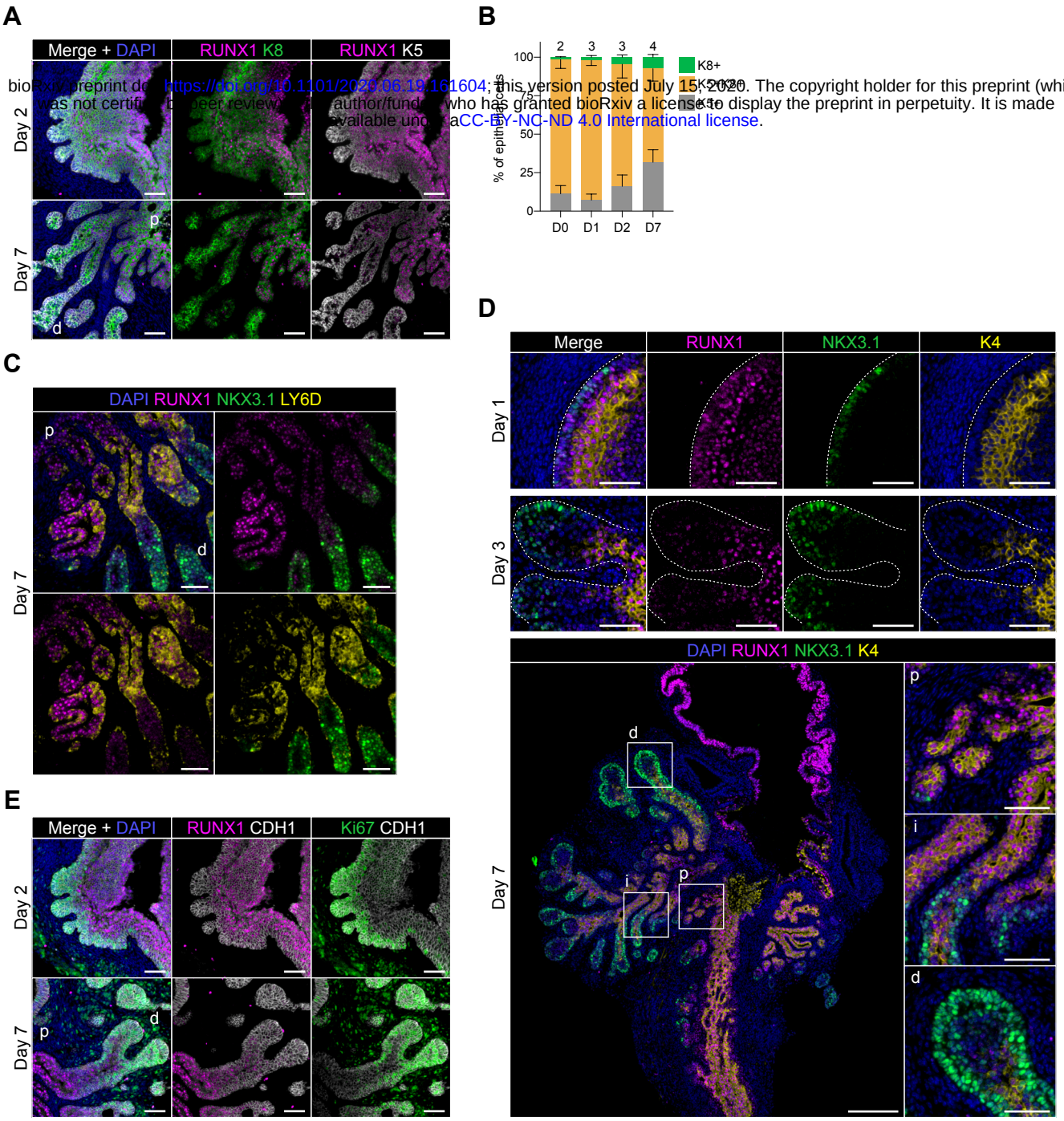

Figure 5—figure supplement 11. Characterization of RUNX1 expression during prostate development in UGS explant cultures.

(**A**) Co-immunostaining of RUNX1, K5, K8 in UGS explants harvested at day 2 and day 7. p: proximal, d: distal. Scale bar: 50µm. (**E**) Quantification of the percentage of $K5^+$, $K8^+$, $K5^+$ $K8^+$ cells during UGS explant cultures at day 0 (*n* = 2 explants), day 1 (*n* = 3 explants), day 2 (*n* = 3 explants), day 7 (*n* = 4 explants), showing a progressive reduction in the proportion of $K5^+$ $K8^+$ double positive cells. (**C**) Co-immunostaining of RUNX1, NKX3.1, LY6D in UGS explants harvested at day 7. p: proximal, d: distal. Scale bar: 50µm. (**D**) Co-immunostaining of RUNX1, NKX3.1, K4 in UGS explants harvested at day 1 (top), 3 (middle) and 7 (bottom). Higher magnification images of (p) proximal, (i) intermediate, and (d) distal regions are shown for day 7. Scale bars: 200µm (yellow) and 50µm (white). Dotted lines indicate the urogenital epithelium. (**E**) Co-immunostaining of RUNX1, Ki67, CDH1 in UGS explants harvested at day 2 and day 7. p: proximal, d: distal. Scale bar: 50µm.

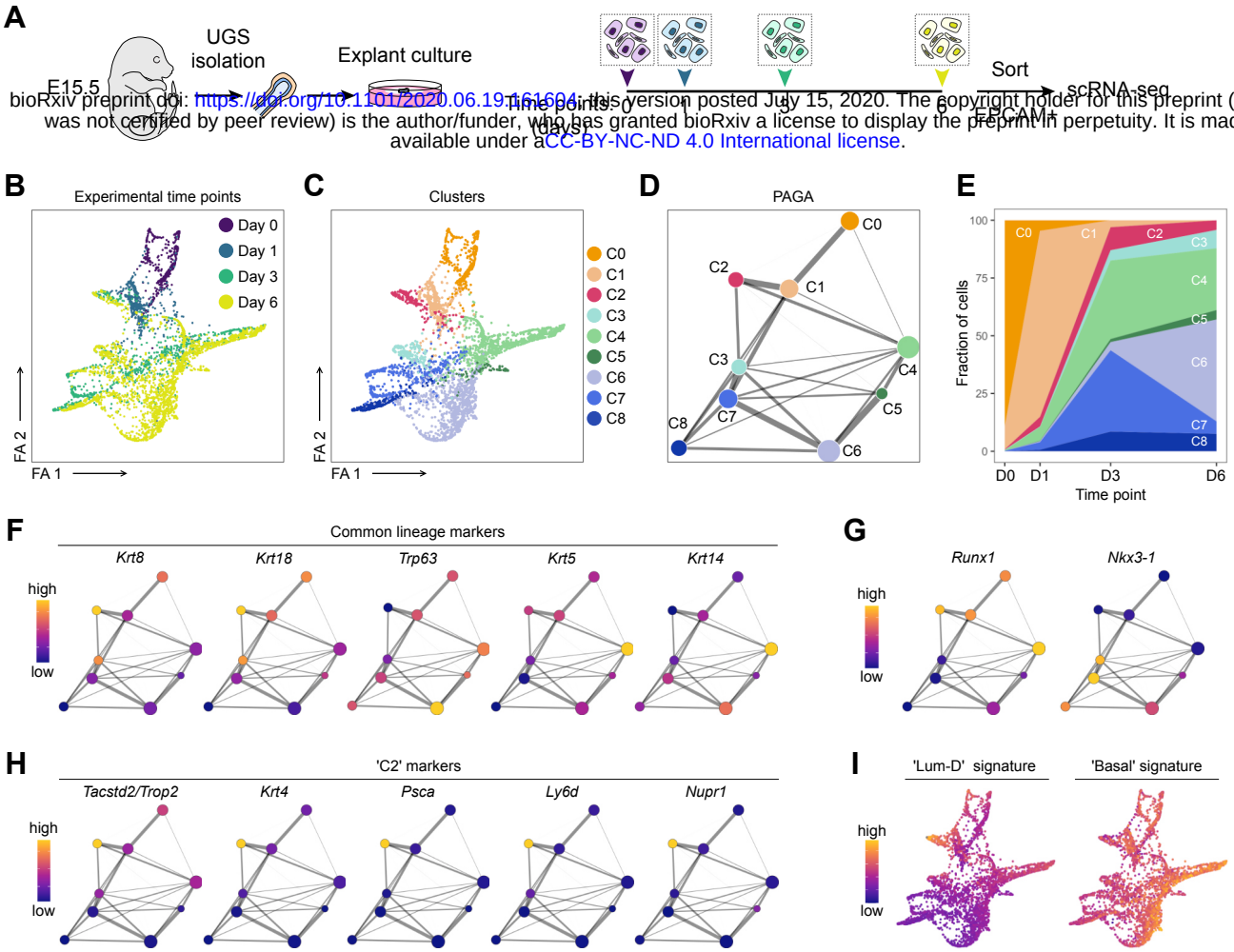

Figure 6. scRNA-seq of UGS explant cultures shows specification of the proximal luminal lineage during embryonic prostate development.

(**A**) Experimental strategy for scRNA-seq of UGS explant cultures at day 0, day 1, day 3 and day 6. (**B**, **C**) Force directed visualization of the developing prostatic epithelium in UGS explant cultures. In **B** cells are colored by experimental time points, and in **C** cells are colored by clusters. (**D**) PAGA representation of the clusters as in **C**. Weighted edges between cluster nodes represent a statistical measure of connectivity. (**E**) Fraction of cells per cluster at each experimental time point, displaying a progressive cellular diversification. (**F-H**) PAGA representations with cluster nodes colored by a gradient representing the mean log-normalized expression levels of each gene. (**I**) Force directed visualization of the developing prostatic epithelium in UGS explant cultures. Color gradient represents AUC scores per cell. Per-cell AUC scores were calculated using the 'AUCell' package. Gene signatures for 'Lum-D' (left) and 'Basal' (right) were generated using the list of differentially upregulated genes previously obtained from our adult mouse prostate clusters.

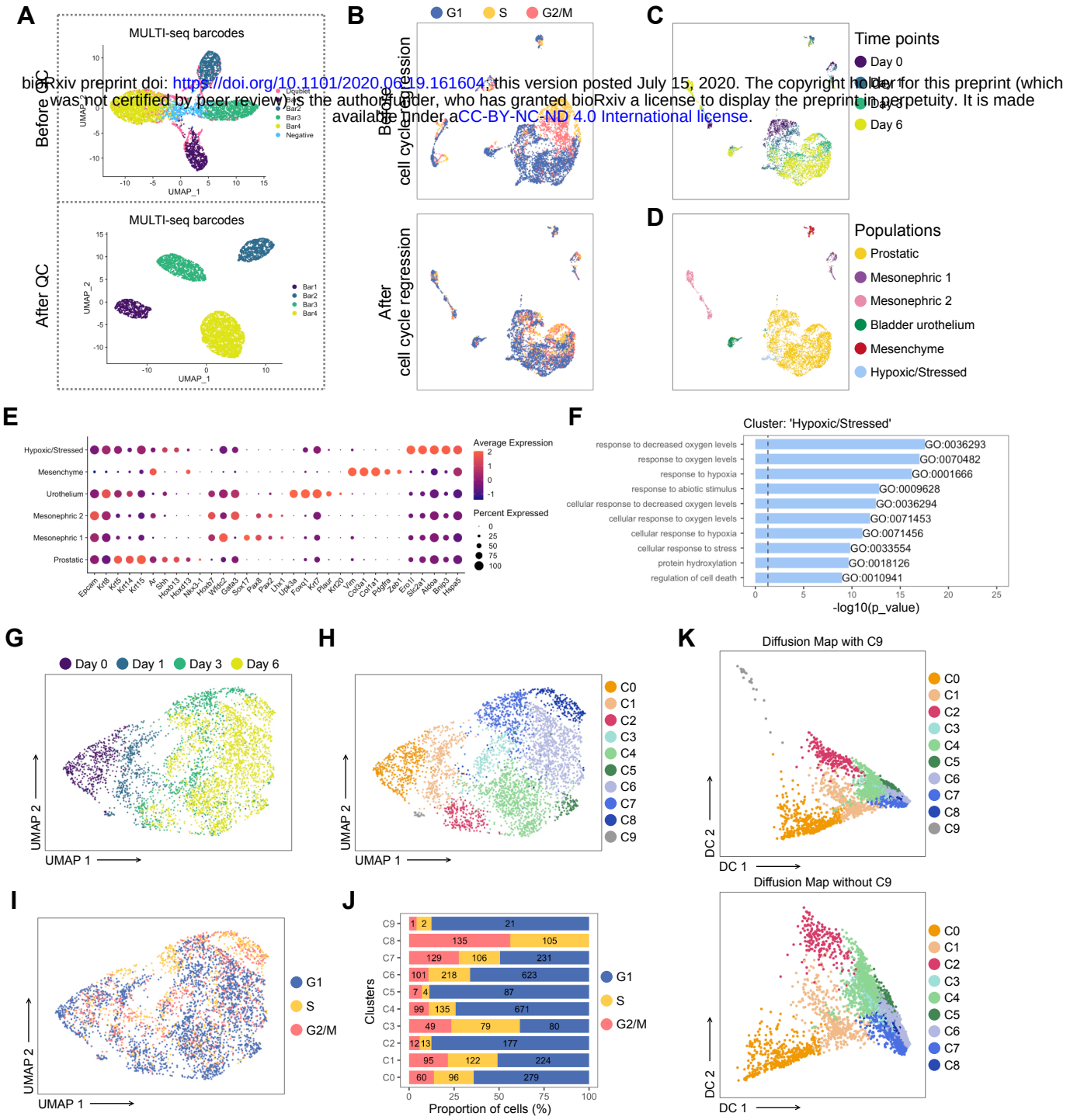

Figure 6—figure supplement 12. Pre-processing of the scRNA-seq dataset of UGS explant cultures.

(**A**) UMAP visualization in MULTI-seq barcodes space before (top) and after (bottom) quality control and barcode filtration (see Methods for details). (**B**) UMAP visualization of the dataset colored by estimated cell cycle phase (top). Due to the strong underlying impact on clustering, the cell cycle effect was regressed out (bottom). (**C, D**) UMAP visualization of the dataset colored in **C** by experimental time point and in **D** by labeled populations. (**E**) Dot plot showing the expression of selected marker genes associated with each labelled population. (**F**) Bar plots of the 10 most significantly (g:Profiler adjusted p-value < 0.05) enriched gene ontology (GO) terms (GO:BP, Biological Processes) on the differentially upregulated genes specific to the 'Hypoxic/Stressed' cluster. (**G**-**I**) UMAP visualization of the prostatic subset (from **D**), labelled by time points in **G**, clusters in **H**, estimated cell cycle phase in **I**. (**J**) Percentage of cells in the different estimated cell cycle phases in each clusters. (**K**) Diffusion maps representation of the prostatic subset using the first 2 diffusion components with (top) and without C9 (bottom). C9 was excluded for downstream analyses due to its outlying profile.

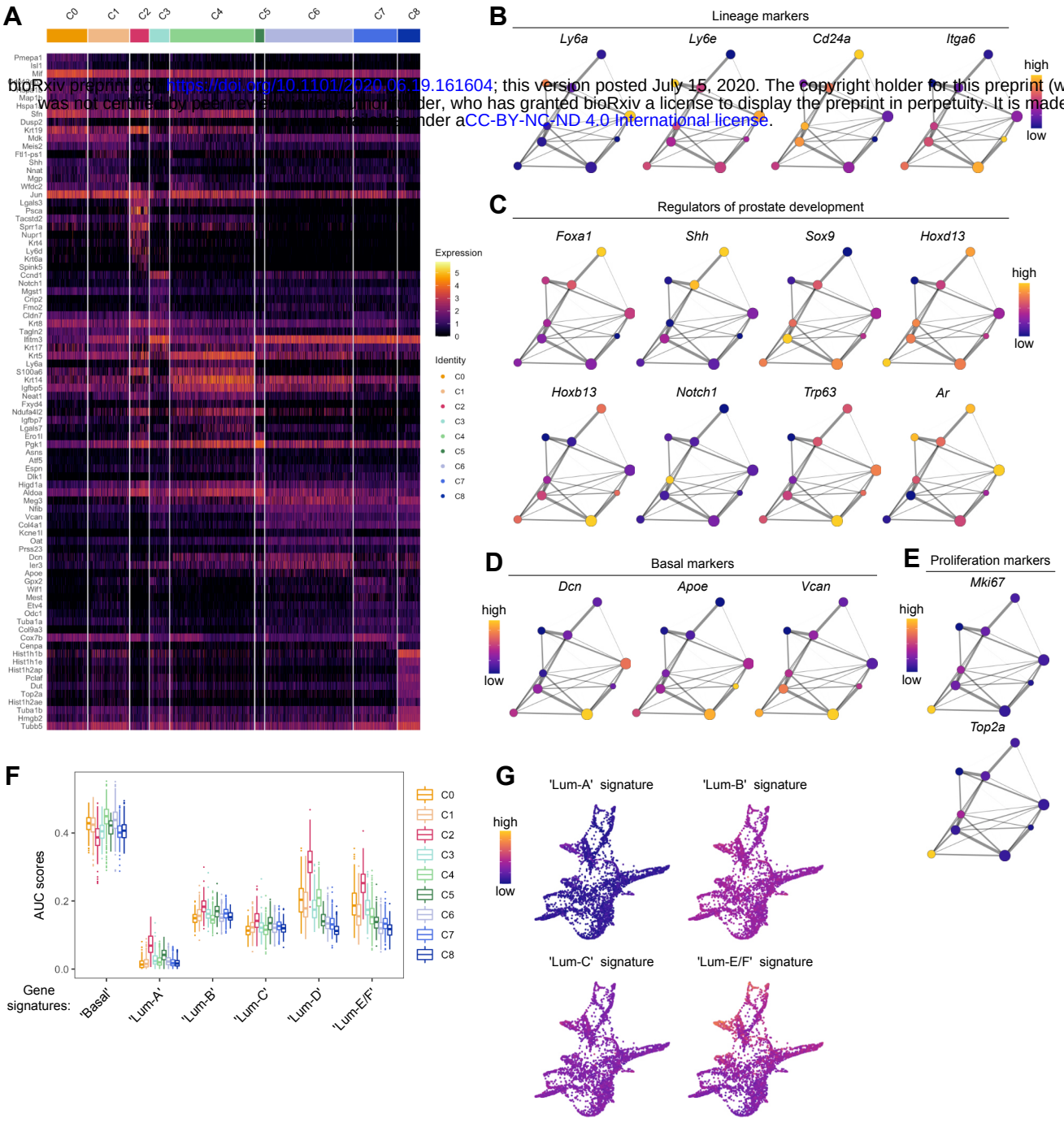

Figure 6—figure supplement 13. Characterization of the developing prostatic epithelium in the scRNA-seq dataset of UGS explant culture.

(**A**) Heatmap of the 10 most differentially upregulated genes per cluster. (**B**-**E**) PAGA representations with cluster nodes colored by a gradient representing the mean log-normalized expression levels of each genes indicated. (**B**) shows the expression of common prostate lineage markers, (**C**) shows known regulators of prostate development, (**D**) shows basal markers, (**E**) shows proliferation markers. (**F**) Box plots of per-cell AUC signature scores calculated using the 'AUCell' package for individual UGS explant clusters. Gene signatures were generated using the list of differentially upregulated genes previously obtained from our adult mouse prostate clusters. (**G**) Force directed visualization of the developing prostatic epithelium in UGS explant cultures. Color gradient represents AUC scores per cell.

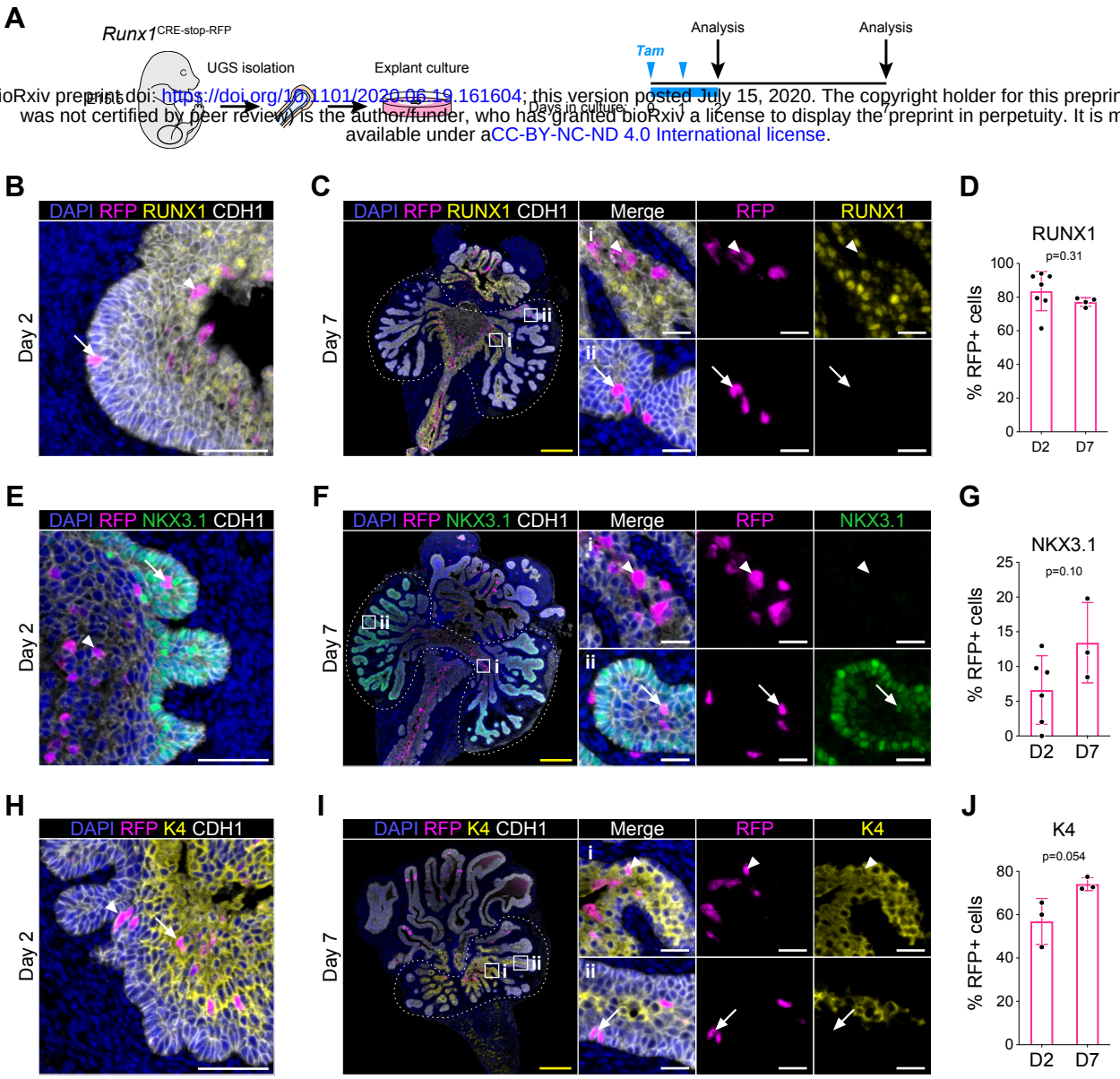

Figure 7. RUNX1⁺ cells contribute to the establishment of the proximal luminal lineage during embryonic prostate development.

(**A**) Strategy for lineage-tracing of RUNX1⁺ cells in UGS explant cultures. Tamoxifen was applied on day 0 and day 1 and washed out on day 2. (**B, C**) Co-immunostaining of RFP, RUNX1, CDH1 in UGS explants harvested at day 2 (**B**) and day 7 (**C**). Higher magnification images of proximal (i) and (ii) distal regions are shown for day 7. Arrows show RFP⁺ RUNX1-low cells, arrowheads show RFP⁺ RUNX1⁺ cells. Scale bars: 200µm (yellow) and 50µm (white). (**C**) Quantification of % of epithelial RUNX1⁺ cells in the RFP subset at day 2 (*n* = 7) and day 7 (*n* = 3) of UGS explant cultures. Quantification was performed within the boundaries delimited in **B** by dotted lines. (**E, F**) Co-immunostaining of RFP, NKX3.1, CDH1 in UGS explants harvested at day 2 (**E**) and day 7 (**F**). Higher magnification images of (i) proximal and (ii) distal regions are shown for day 7. Arrows show RFP⁺ NKX3.1⁺ cells, arrowheads show RFP⁺ NKX3.1⁻ cells. Scale bars: 200µm (yellow) and 50µm (white). (**G**) Quantification of % of epithelial NKX3.1⁺ cells in the RFP subset at day 2 (*n* = 6) and day 7 (*n* = 4) of UGS explant cultures. Quantification was performed within the boundaries delimited in **F** by dotted lines. (**H, I**) Co-immunostaining of RFP, K4, CDH1 in UGS explants harvested at day 2 (**H**) and day 7 (**I**). Higher magnification images of (i) proximal and (ii) distal regions are shown for day 7. Arrows show RFP⁺ K4⁻ cells, arrowheads show RFP⁺ K4⁺ cells. Scale bars: 200µm (yellow) and 50µm (white). (**J**) Quantification of % of epithelial K4⁺ cells in the RFP subset at day 2 (*n* = 3) and day 7 (*n* = 3) of UGS explant cultures. Quantification was performed within the boundaries delimited in **I** by dotted lines.

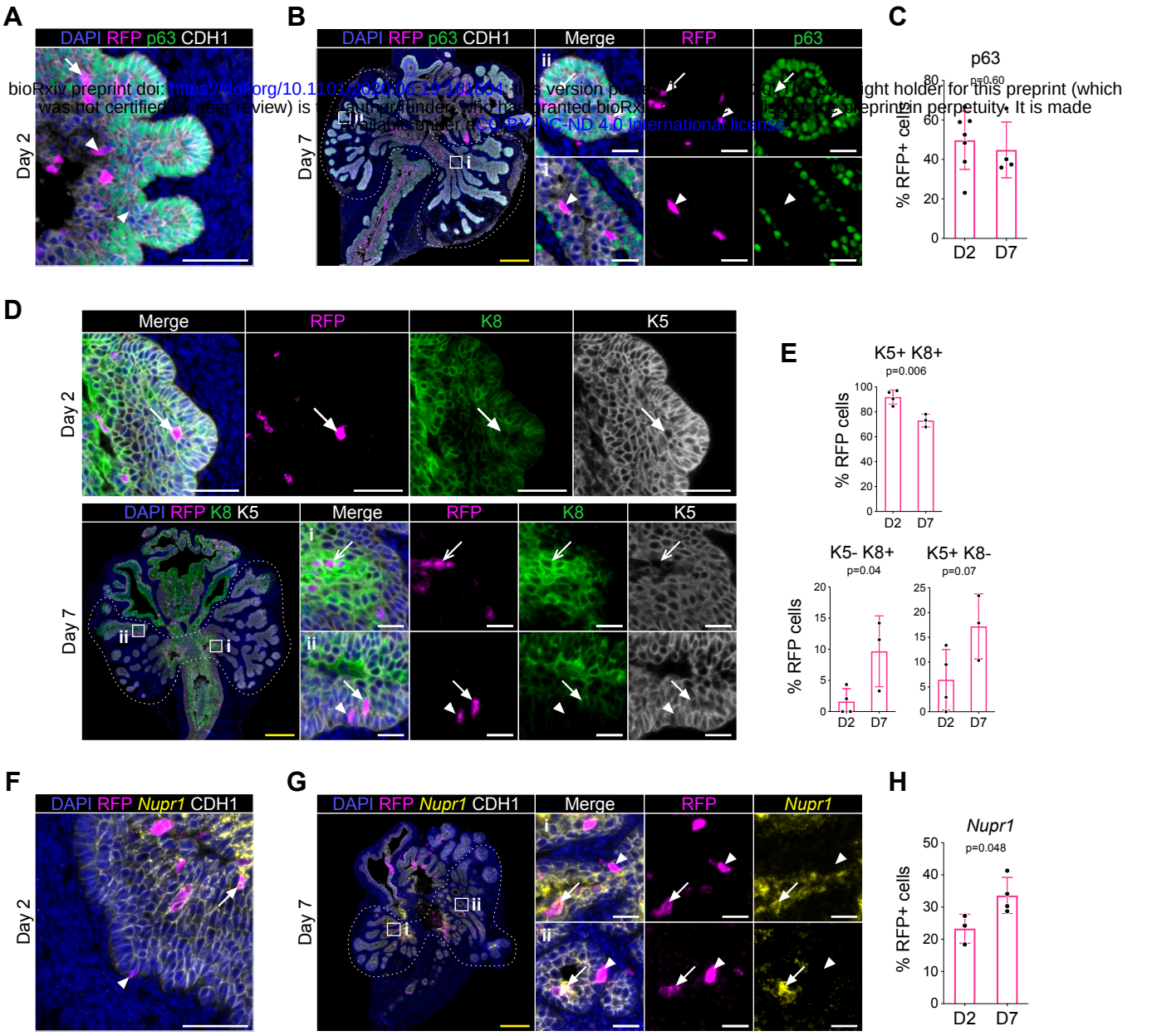

Figure 7—figure supplement 14. Lineage-tracing of RUNX1 expressing cells in UGS explants.

(**A, B**) Co-immunostaining of RFP, p63, CDH1 in UGS explants harvested at day 2 (**A**) and day 7 (**B**). Higher magnification images of (i) proximal and (ii) distal regions are shown for day 7. Arrows show RFP$^+$ p63$^+$ cells, arrowheads show RFP$^+$ p63$^-$ cells. Scale bars: 200µm (yellow) and 50µm (white). (**C**) Quantification of the percentage of epithelial RFP$^+$ p63$^+$ cells at day 2 (*n* = 7) and day 7 (*n* = 4) of UGS explant cultures. Quantification was performed within the boundaries delimited in **B** by dotted lines. (**D**) Co-immunostaining of RFP, K5, K8 in UGS explants harvested at day 2 (top) and day 7 (bottom). Higher magnification images of proximal (i) and (ii) distal regions are shown for day 7. Arrows show RFP$^+$ K5$^+$ K8$^+$ cells, chevron arrows show RFP$^+$ K5$^-$ K8$^+$, arrowheads show RFP$^+$ K5$^+$ K8$^-$ cells. Scale bars: 200µm (yellow) and 50µm (white). (**E**) Quantification of the percentage of epithelial K5$^+$ K8$^+$, K5$^-$ K8$^+$ cells and K5$^+$ K8$^-$ cells in the RFP subset at day 2 (*n* = 4) and day 7 (*n* = 3) of UGS explant cultures. Quantifications were performed within the boundaries delimited in **D** by dotted lines. (**F, G**) Co-immunostaining of RFP, *Nupr1* (mRNA), CDH1 in UGS explants harvested at day 2 (**F**) and day 7 (**G**). Higher magnification images of (i) proximal and (ii) distal regions are shown for day 7. Arrows show RFP$^+$ *Nupr1*$^+$ cells, arrowheads show RFP$^+$ *Nupr1*$^-$ cells. Scale bars: 200µm (yellow) and 50µm (white). **(J)** Quantification of % of epithelial K4$^+$ cells in the RFP subset at day 2 (*n* = 3) and day 7 (*n* = 4) of UGS explant cultures. Quantification was performed within the boundaries delimited in **G** by dotted lines.