Am J Clin Exp Urol 2019;7(3):123-138
www.ajceu.us /ISSN:2330-1910/AJCEU0096357

*Original Article*

# Single-cell RNA-Seq analysis identifies a putative epithelial stem cell population in human primary prostate cells in monolayer and organoid culture conditions

Tara McCray[1*], Daniel Moline[2*], Bethany Baumann[1], Donald J Vander Griend[1,3], Larisa Nonn[1,3]

*[1]Department of Pathology, University of Illinois at Chicago, Chicago 60612, Illinois, USA; [2]Committee on Development, Regenerative, and Stem Cell Biology (DRSB), University of Chicago, Chicago 60637, Illinois, USA; [3]University of Illinois Cancer Center, Chicago 60612, Illinois, USA. [*]Equal contributors.*

Received April 30, 2019; Accepted May 30, 2019; Epub June 15, 2019; Published June 30, 2019

**Abstract:** Human primary prostate epithelial (PrE) cells represent patient-derived *in vitro* models and are traditionally grown as a monolayer in two-dimensional culture. It has been recently demonstrated that expansion of primary cells into three-dimensional prostatic organoids better mimics prostate epithelial glands by recapitulating epithelial differentiation and cell polarity. Here, we sought to identify cell populations present in monolayer PrE cells and organoid culture, grown from the same patient, using single-cell RNA-sequencing. Single-cell RNA-sequencing is a powerful tool to analyze transcriptome profiles of thousands of individual cells simultaneously, creating an in-depth atlas of cell populations within a sample. Organoids consisted of six distinct cell clusters (populations) of intermediate differentiation compared to only three clusters in the monolayer prostate epithelial cells. Integrated analysis of the datasets allowed for direct comparison of the monolayer and organoid samples and identified 10 clusters, including a distinct putative prostate stem cell population that was high in Keratin 13 (*KRT13*), Lymphocyte Antigen 6D (*LY6D*), and Prostate Stem Cell Antigen (*PSCA*). Many of the genes within the clusters were validated through RT-qPCR and immunofluorescence in PrE samples from 5 additional patients. KRT13+ cells were observed in discrete areas of the parent tissue and organoids. Pathway analyses and lack of EdU incorporation corroborated a stem-like phenotype based on the gene expression and quiescent state of the KRT13+ cluster. Other clusters within the samples were similar to epithelial populations reported within patient prostate tissues. In summary, these data show that the epithelial stem population is preserved in PrE cultures, with organoids uniquely expanding intermediate cell types not present in monolayer culture.

**Keywords:** Prostate, organoid, primary cell culture, 3D culture, single-cell RNA sequencing, 10X Genomics, ingenuity pathway analysis, immunofluorescent staining

## Introduction

The prostate epithelium has a high incidence of neoplastic disease with prostate cancer being the second most common epithelial cancer in men [1, 2]. *In vitro* culture models have produced valuable insights into the biology of the prostate, however, the limited number of cell lines from early and intermediate stages of disease presents a significant obstacle to unifying data with clinically relevant findings [3, 4]. An alternative strategy is patient-derived primary cell culture, which preserves patient heterogeneity [5-10]. Primary prostate cell culture can be a valuable tool for studying "normal" cells, but is restricted to selectively expand epithelial cells that display a homogenous, transit-amplifying phenotype and lack the luminal differentiation of prostate epithelium observed *in vivo* [11-14]. Organoids are three-dimensional (3D) structures grown in extracellular matrix that recapitulate many facets of prostate epithelial tissue morphology including structure and cell polarity [15-17]. Compared to their traditional two-dimensional (2D) monolayer counterparts, organoids can grow from a single stem or progenitor cell in the presence of charcoal stripped FBS and androgen to differentiate into both

basal and luminal epithelial populations [15, 16, 18].

The human prostate consists of stratified epithelial secretory glands surrounded by a fibromuscular stroma. The epithelial glands are composed of a basal layer, a secretory luminal layer, and a rare neuroendendocrine population [19, 20]. Recently, single-cell RNA-Seq analysis of prostate tissue revealed two additional cell populations within the human prostate epithelium that exhibit stem cell characteristics [21]. Single-cell RNA-Seq (scRNA-Seq) is a method that lends itself to the identification of cryptic sub-populations within a heterogeneous sample using an unbiased analysis of individual expression profiles of cells. This approach involves the isolation of single cells into microfluidic droplets containing oligonucleotide-covered gel beads that capture and barcode the transcripts. Transcripts are converted to cDNA, sequenced, and aligned by barcode using computational analysis to create an individual transcriptome library for each cell. Libraries are then clustered into distinct cell populations using dimensional reduction analysis [22-25].

Here we use scRNA-Seq to compare the subpopulations present within primary prostate cells and organoids from the same patient specimen and identify previously unknown subpopulations of epithelial cells grown *in vitro*. Cell populations were validated in additional patient samples by RT-qPCR and immunofluorescence microscopy.

## Materials and methods

### *Contact for reagent and resource sharing*

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Larisa Nonn (lnonn@uic.edu).

### *Primary prostate epithelial cells*

Human primary prostate cells were isolated and established from fresh male radical prostatectomy tissues. Radical prostatectomy patients consented prior to surgery and prostate tissue samples from benign regions of the peripheral zone were collected according to UIC Internal Review Board-approved protocol #2007-0694. A portion of tissue was reserved for histologic inspection by a board-certified pathologist to verify the region as benign. Tissue samples were collected, formalin-fixed, paraffin-embedded, and 5 μm sections were stained with hematoxylin and eosin. Remaining tissue was digested in collagenase/trypsin to produce a single cell suspension. Cells were grown in Prostate Cell Growth Media (Lonza, Basel, Switzerland) to select for epithelial cells. When ~70% confluent, cells were trypsinized to single cells, counted and cryopreserved into multiple aliquots. Epithelial purity was authenticated with RT-qPCR, confirming the expression of epithelial markers *KRT5*, *KRT8*, *KRT18* and *TP63*, and the lack of stromal marker *TIMP3*. Patient information is listed within **Table 1**.

### *Monolayer and organoid culture*

For standard monolayer culture, prostate epithelial (PrE) cells were thawed from primary passage into a collagen-coated dish and maintained in PrEGM (Lonza; Basal, Switzerland). For organoid culture, a separate aliquot of the same PrE cells was thawed and plated sparsely (500-5,000 cells per well) in a flat-bottom 96-well microplate into 33% growth factor reduced phenol red-free Matrigel (Corning Inc., Corning NY) on top of a solidified base layer of 50% Matrigel in media. 100 μL of organoids suspended in 33% Matrigel were maintained in keratinocyte serum-free media (Gibco, Thermo Fisher Scientific, Waltham, MA) supplemented with 5% charcoal-stripped fetal bovine serum and 10 nM dihydrotestosterone (DHT) as previously described by our group [26]. Media was refreshed every 2-3 days for all cultures. Monolayer cells were collected at ~70% confluent for endpoints and organoids were grown for 8-14 days as detailed in the figure legends. Brightfield images of organoids were captured at 10 × and 20 × magnification using the Evos FL Auto 2 imaging System (Thermo Fisher Scientific, Waltham, MA), shown in **Figures 1** and S4.

### *10 × single cell separation, library prep and sequencing*

Patient-matched epithelial cells were grown in monolayer and organoid culture as described above. Monolayer cells were collected by TrypLE (Gibco, Thermo Fisher Scientific, Waltham, MA) dissociation. Organoids were harvested at day 8 by Dispase (STEMCELL Technologi-

**Table 1.** Primary cell patient characteristics

| PrE ID | Pathology | Age | Race | Prostate region of origin | Endpoints |
|---|---|---|---|---|---|
| PrE1 | Benign, moderate chronic inflammation | 58 | AA | Peripheral zone, L5 Left Posterior (scRNAseq)<br>Peripheral Zone, L2 Left Posterior (FFPE block) | scRNAseq<br>RT-qPCR<br>Tissue staining |
| PrE2 | Benign | 68 | AA | Peripheral zone, L3 (whole mount and RNA)<br>Peripheral zone, L5 (FFPE block) | RT-qPCR,<br>Tissue staining,<br>ICC, whole mount ICC |
| PrE3 | Benign, mild chronic inflammation, atrophy | 63 | EA | Peripheral zone, L1 Left Anterior | RT-qPCR,<br>Tissue staining,<br>ICC, whole mount ICC |
| PrE4 | Benign, mild chronic inflammation, focal atrophy | 58 | AA | Peripheral zone, L3 Left Posterior | RT-qPCR,<br>Tissue staining,<br>ICC, whole mount ICC |
| PrE5 | Benign | 71 | AA | Peripheral zone, L5 Left anterior (whole mount, RNA)<br>Peripheral zone, L3 Right anterior (FFPE block) | Tissue staining, ICC,<br>whole mount ICC |
| PrE6 | Benign | 60 | AA | Peripheral zone, L4 Left Anterior | ICC, whole mount ICC<br>IF |

AA = African American; EA = White European American (non-Hispanic).

es, Vancouver, Canada) dissociation followed by a second dissociation to single cells using TrypLE. Cell number and viability were determined by a Trypan Blue exclusion assay quantified on a Cellometer Automated Cell Counter (Nexcelom, Lawrence MA). Both samples consisted of > 80% viable cells (**Table 2**) prior to proceeding with the 10X Genomics (Pleasanton, CA) protocol for 3' Transcript Capture and Single Cell Library Prep. Cell samples were loaded at a concentration to yield approximately 5 × $10^3$ total captured cells on a 10 × Chip A. GEM generation, RT, cleanup, cDNA amplification, fragmenting, end repair & A-tail prep, and sample index tagging were performed using the Chromium Single Cell 3' Library and Gel Bead Kit v2 per manufacturer's instructions. Libraries were labeled with a sample index and pooled for sequencing at 10 nM. Sample quantification and quality control were performed using Qubit Fluorometer (London, England) and TapeStation Bioanalyzer (Agilent Technologies, Santa Clara CA). Sequencing was run on the HiSeq 4000 (Illumina, San Diego CA) at the University of Illinois at Urbana Champaign (UIUC) DNA services. Samples were sequenced across 3 lanes of the HiSeq 4000, generating 100 base pair paired-end reads at a depth of 45,000 reads per cell. Leftover cells not used for scRNA-Seq were collected into TRIzol Reagent (Thermo Fisher, Waltham MA) and reserved for validation of the sequencing.

*Single cell rna-seq analysis*

Initial Read Alignment and Quality Control: Single-cell RNA sequencing samples were processed and aligned to Ensembl genome GRCh38 using the Cell Ranger 3.0.0 pipeline by UIUC DNA Services. The Cell Ranger output was loaded into Seurat pre-release v3.0 for clustering [22]. Cells with high mitochondrial features (> 8% of total mapped reads) were struck from the analysis to remove the influence of dead cells. A small number of cells with unusually high or low numbers of mapped reads were also removed from the dataset, as these outliers could be doublets or poorly-captured cells [22-24]. Cells removed in QC totaled 29% of the initial input, see **Table 2** for full QC. Individual genes related to the cell cycle or with uniquely low unique molecular identifier (UMI) counts within the context of the dataset had their variance regressed out to minimize their influence on variance-based clustering.

Principal Component Determination: A JackStraw resampling method was used to select statistically significant ($P < 0.05$) principal components, shown in Figure S1A and S1B [23]. These components were used to identify the distance between cells for a k-nearest neighbors calculation and construction of a shared nearest neighbor graph. Modularity (M) was used as a quantitative measure of the independence of individual networks in the t-SNE plot, using $M > 0.8$ as a cut-off to ensure reproducibility. Principal components used were 27 and 40 for monolayer and organoid culture respectively in the individual analyses provided in Figure S2. The integrated dataset was analyzed using 30 principal components to produce the t-SNE plot shown in **Figure 1**.

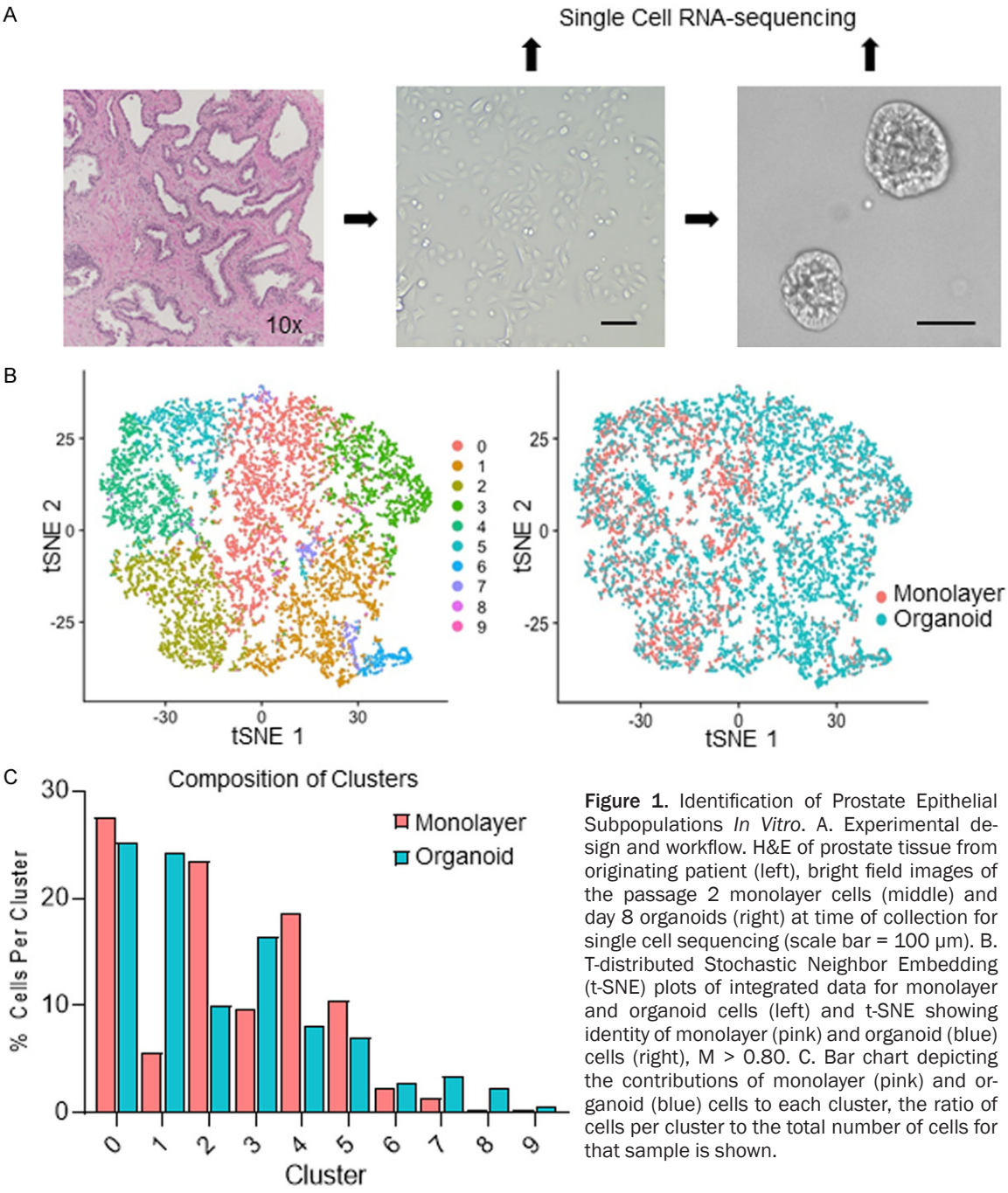


**Figure 1.** Identification of Prostate Epithelial Subpopulations *In Vitro*. A. Experimental design and workflow. H&E of prostate tissue from originating patient (left), bright field images of the passage 2 monolayer cells (middle) and day 8 organoids (right) at time of collection for single cell sequencing (scale bar = 100 μm). B. T-distributed Stochastic Neighbor Embedding (t-SNE) plots of integrated data for monolayer and organoid cells (left) and t-SNE showing identity of monolayer (pink) and organoid (blue) cells (right), M > 0.80. C. Bar chart depicting the contributions of monolayer (pink) and organoid (blue) cells to each cluster, the ratio of cells per cluster to the total number of cells for that sample is shown.

Identification of Clusters: Canonical correlative analysis was performed in Seurat to integrate the separate datasets and allow for direct comparison of populations between the monolayer and organoid samples [22]. Clusters were assigned identities based on their expression of previously reported epithelial markers listed in **Table 3**. Heat maps and dot plots were generated in Seurat, **Figure 2**. Highly expressed genes in each cluster are provided in Table S1.

Pseudotime analysis: Monocle version 2.10 was used for unsupervised pseudotime analysis of the organoid sample [27], Cell Ranger output was uploaded into Monocle and data was subset to exclude cells with low and high

**Table 2.** Quality metrics for 10 × input, sequencing and analysis

| Sample | Monolayer | Organoid |
|---|---|---|
| Viability at Collection | 85.60% | 81% |
| Aimed Recovery | 5,000 | 5,000 |
| TapeStation Yield (pmol/L) | 173000 | 121000 |
| Concentration (ng/μL) | 44 | 31.4 |
| Qubit Yield (ng/μL) | 88 | 82 |
| Achieved Cell Recovery | 5,194 | 7,422 |
| Mean Reads per Cell | 31,629 | 41,116 |
| Mean Genes per Cell | 3,569 | 3,783 |
| Reads Mapped Confidently to Genome | 87.10% | 86.90% |
| Number of Cells Remaining After Processing | 3,687 | 5,322 |

numbers of mRNAs, removing dead cells and doublets (cells with < 9000 or > 45000 captured transcripts). Clustering of cells was performed unguided without the influence of marker genes, using 27 principal components as shown in Figure S1. Monocle performed unsupervised selection of genes that define progress, performing dimensional reduction to produce a plot ordering cells in pseudotime, shown in **Figure 4**.

Integrated tissue analysis: Canonical correlative analysis was used to integrate the monolayer, organoid and a publicly available human prostate tissue data set (D17_FACS_filtered GSE_117403, [21]) to allow for direct comparison of populations between the *in vitro* samples and *in vivo* tissue. 30 principal components (Figure S2) were used to yield a t-SNE with M > 0.95 (**Figure 4**). Clusters were assigned identities based on their expression of previously reported epithelial markers listed in **Table 3**. A dot plot was generated in Seurat for genes highly expressed by each cluster, shown in Figure S7. Highly expressed genes in each cluster are provided in Table S1.

*RT-qPCR gene expression*

Multiple patient-derived epithelial cell cultures (**Table 1**) were grown as matched monolayer and organoid cultures as described above. Cells were stored in TRIzol Reagent before RNA isolation. Samples were homogenized by chloroform and RNA collected by alcohol precipitation and rehydration. RNA quantity and quality was determined by OD 260/280 and 260/230 on the NanoDrop Spectrophotometer (Thermo Fisher Scientific, Waltham MA). cDNA was synthesized with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Beverly Hills CA) and qPCR run on LightCycler (Roche Applied Science, Penzberg, Germany). RQ was calculated from ΔΔCT to the reference gene *RPL13A* [28], primers are listed in **Table 4**.

*Immunofluorescence and Immunohistochemistry*

Whole-mount immunocytofluorescent staining: Organoids were collected for whole-mount imaging by transferring to a chamber slide as previously described by our group [26] and monolayer cells were seeded from P2 and grown on a chamber slide. Organoids and cells were fixed in 4% paraformaldehyde. Organoids and cells were permeabilized with Triton-X 100 (Sigma Aldrich, St. Louis MO). Cells were incubated first with primary and next secondary antibodies for 1 hour at room temperature, each. Organoids were incubated with primary antibodies over two nights at 4°C and secondary antibodies overnight at 4°C. Primary antibodies included monoclonal rabbit anti-Cytokeratin 13 (ab92551, Abcam) diluted at 1:200, monoclonal mouse anti-E-cadherin (ab76055) diluted at 1:100, and polyclonal guinea pig anti-Keratin 8/18 (03-GP11, American Research Products, Inc.) diluted at 1:200. Secondary antibodies included goat anti-rabbit Alexa Fluor 488, Alexa Fluor 647 goat anti-mouse, and goat anti-guinea pig Alexa Fluor 568 (Life Technologies) were used at a 1:200 dilution. Cells were counterstained with Alexa Fluor™ 647 Phalloidin (Thermo Fisher Scientific, Massachusetts) and DAPI, and imaged on the Zeiss LSM 710 confocal microscope (ZEISS, Oberkochen Germany), monolayer cells were counterstained with DAPI alone when E-cadherin primary was used. Shown in **Figures 3** and S4.

Formalin fixed paraffin embedded staining: Organoids were formalin-fixed and paraffin-embedded as previously described by our group [26]. Briefly, organoids were dissociated in Dispase, resuspended in HistoGel™ (Thermo Fisher Scientific, Waltham MA), solidified at 4°C, fixed in 4% paraformaldehyde for 1 hour, transferred to 70% ethanol, and paraffin-em-

**Table 3.** Gene expression profiles for prostate cell types

| | Pan-Epi | Luminal | Basal | Intermed | Stem |
|---|---|---|---|---|---|
| Henry [21] | CD326, CD324, TACSTD2, ITGA6 | DPP4, KRT8, KRT18, GP2, KLK3, MSMB ACPP, KLK4, PLA2G2A, MT1E, KLK2, SMB, SOCS2, TSPAN8 | PDPN, KRT5, TP63, CD104, CD271, RGCC KRT14, DST, NOTCH4, LTBP2, DKK1, KRT15 | KRT19, KRT18, KRT14 | SCGB1A1, PSCA, KRT13, LCN2, LYPD3, SERPINB1, SCGB3A1, PIGR, WFDC2, FCGBP, CSTB, APOBEC3A |
| Zhang [40] | TACSTD2 | AR, KLK3, ALOX15B, ACPP, TOX3, FOLH1, OSTADPP4, CKK, PLA2G2A, MB, CWH43, TRPV6, ELOVL2, CPNE4, ANO7, POTEM, MUC2, LMAN1L, DNAJC12, ASRGL1, DLL4, DOCK11, GPR98, SYT7, INHB8, TBXAS1, SERHL2, NPTX2, GFPT2, PTPRN2, CSGALNACT1, ST8SIA1, VNN3, C2, TRPM8, RAMP1PDE8B, SPDEF, LTB, HLA-DMB, LOC286002, FGF13, DCDC2, KRT20, FBP1, SLC2A12, TSPAN8 | TP63, KRT5, KRT14, ITGA6KRT6A, BNC1, KRT34, FAT3, SYNE1, TNC, FGFR3, DKK3, COL17A1, CSMD2, CDH13, FJX1, MUM1L1, MMP3, DLK2, FLRT2, IL33, GIMAP8, PDPN, FHL1, VSNL1, NRG, IGFBP7, ERG, HMGA2, IL1A, NOTCH3, THBS2, TAGLNSPARC, FOXI1, MSRB3, NGFR, NIPAL4, ANXABL2, COL4A6, KCNQs, JAG2WNT7A, KCNMA1, LTBP2, JAM3, SH2D5, MRC2, SERPINB13, CNTAP3B, ARHgAP25, AEBP1, DLC1, SERPINF1, C16orf74, KIRREL | | |
| Hu [33] | | | | | KRT13, IGF2, CCL2, CARM1, LDH1A2, ALDH8A1, LIFRCAC15, PSCA, Nanog, SOX2, PCAT1, VEGFC Nestin, FOXA1, APRAC1 |
| Moad [34] | | | | | SCNN1A, PPL, CENPF, DDIT4, KLK11, DLK1, |
| Wang [41] | | KRT8, KRT18 | KRT19, GSTpi, KRT5, TP63, KRT14 | KRT19 | |
| Schmelz [32] | | TPJ1, KLK3 | KRT5 | KRT19 | KRT6A |

Markers observed in monolayer and organoid scRNA-Seq data are marked in red.

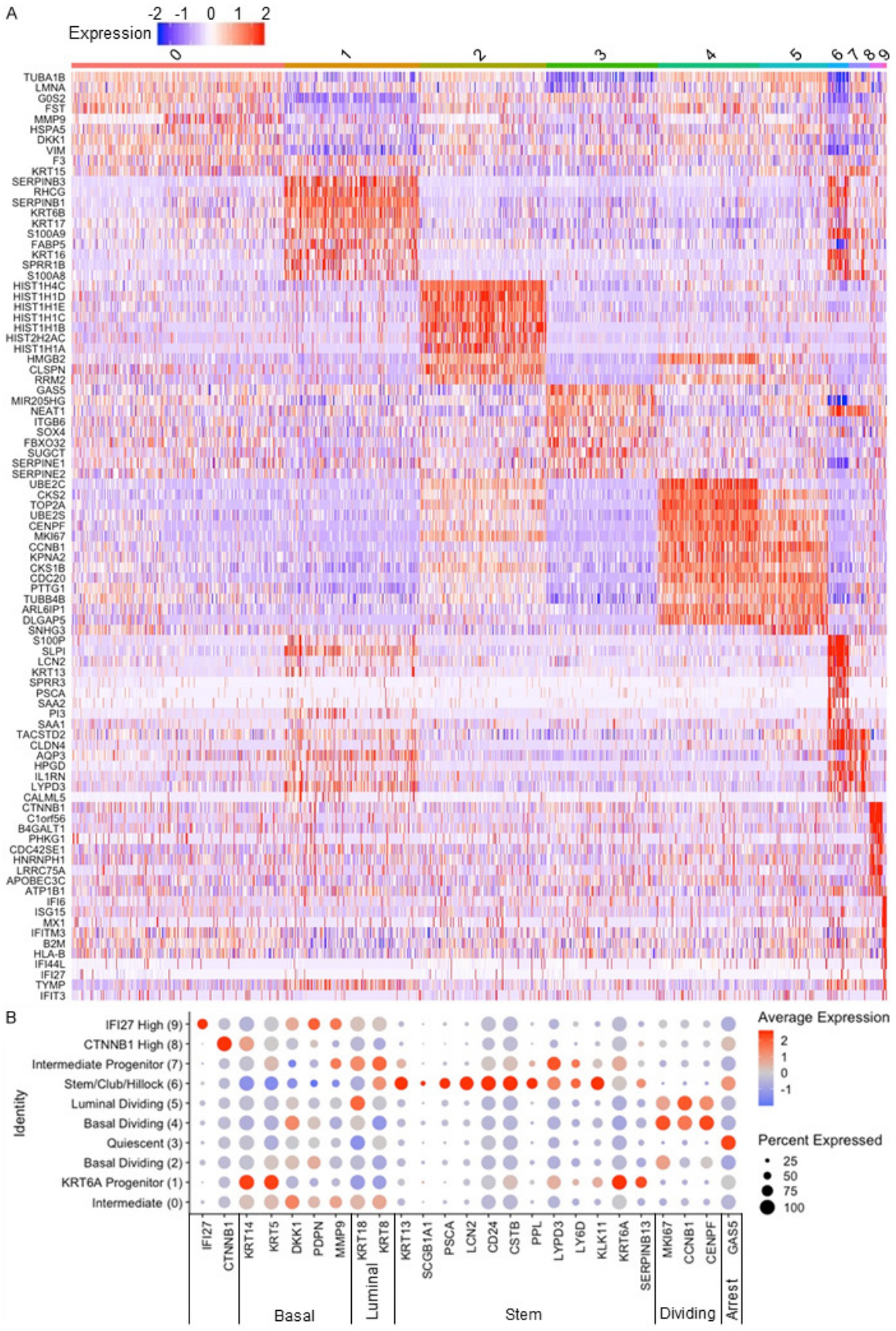
A
Expression
-2 -1 0 1 2
0 1 2 3 4 5 6 7 8 9
TUBA1B
LMNA
G0S2
FST
MMP9
HSPA5
DKK1
VIM
F3
KRT15
SERPINB3
RHCG
SERPINB1
KRT6B
KRT17
S100A9
FABP5
KRT16
SPRR1B
S100A8
HIST1H4C
HIST1H1D
HIST1H1E
HIST1H1C
HIST1H1B
HIST2H2AC
HIST1H1A
HMGB2
CLSPN
RRM2
GAS5
MIR205HG
NEAT1
ITGB6
SOX4
FBXO32
SUGCT
SERPINE1
SERPINE2
UBE2C
CKS2
TOP2A
UBE2S
CENPF
MKI67
CCNB1
KPNA2
CKS1B
CDC20
PTTG1
TUBB4B
ARL6IP1
DLGAP5
SNHG3
S100P
SLPI
LCN2
KRT13
SPRR3
PSCA
SAA2
PI3
SAA1
TACSTD2
CLDN4
AQP3
HPGD
IL1RN
LYPD3
CALML5
CTNNB1
C1orf56
B4GALT1
PHKG1
CDC42SE1
HNRNPH1
LRRC75A
APOBEC3C
ATP1B1
IFI6
ISG15
MX1
IFITM3
B2M
HLA-B
IFI44L
IFI27
TYMP
IFIT3
B
Identity
IFI27 High (9)
CTNNB1 High (8)
Intermediate Progenitor (7)
Stem/Club/Hillock (6)
Luminal Dividing (5)
Basal Dividing (4)
Quiescent (3)
Basal Dividing (2)
KRT6A Progenitor (1)
Intermediate (0)
IFI27 CTNNB1 KRT14 KRT5 DKK1 PDPN MMP9 KRT18 KRT8 KRT13 SCGB1A1 PSCA LCN2 CD24 CSTB PPL LYPD3 LY6D KLK11 KRT6A SERPINB13 MKI67 CCNB1 CENPF GAS5
Basal
Luminal
Stem
Dividing
Arrest
Average Expression
2
1
0
-1
Percent Expressed
25
50
75
100

**Figure 2.** Gene expression of the 10 clusters identified. A. Heat map for top 10 genes expressed by each of the 10 clusters found in the integrated dataset. B. Naming assignments for the clusters and genes of interest shown by dot plot. Red is highly expressed and blue is lowly expressed, the size of the dot indicates the percentage of cells in that cluster that are expressing the gene.

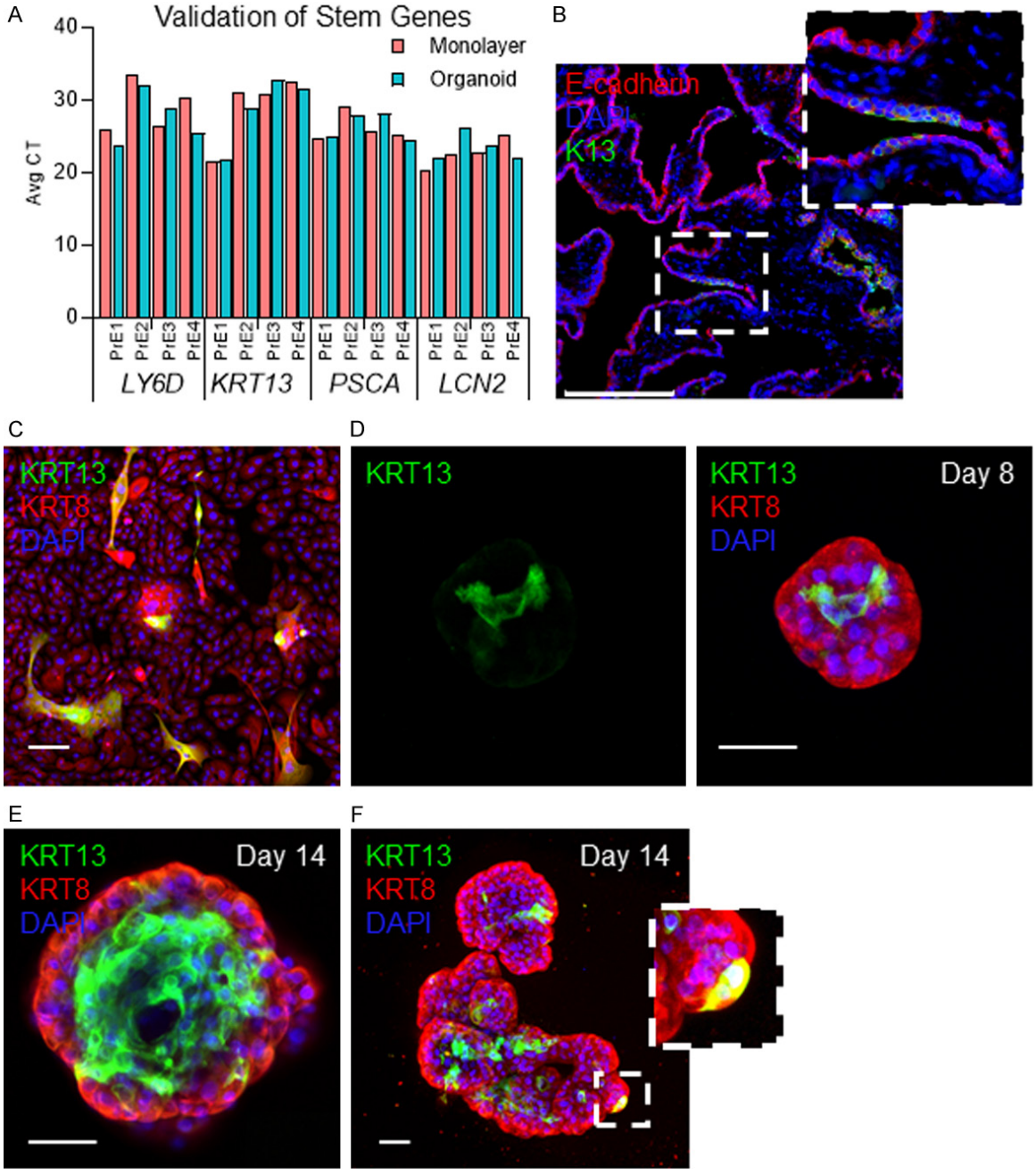


**Figure 3.** Presence of the KRT13+ population in tissue, monolayer cells and organoids. A. RT-qPCR analysis for stem cell markers in whole RNA extracts of monolayer and organoid samples derived from 4 patients. B. Immunofluorescent staining for KRT13, DAPI and E-cadherin on the scRNA-Seq patient tissue (scale bar = 200 μm). C. Immunocytochemistry of monolayer cells derived from patient PrE2 stained for KRT13, KRT8, and DAPI (scale bar = 50 μm). D. Whole-mount immunocytochemistry of day 8 organoid cells derived from patient PrE2 stained for KRT13, KRT8 and DAPI (scale bar = 50 μm). E, F. Whole-mount immunocytochemistry of day 14 organoid cells derived from patient PrE2 stained for KRT13, KRT8, and DAPI, inset (right) shows KRT8+/KRT13+ cell (scale bar = 50 μm).

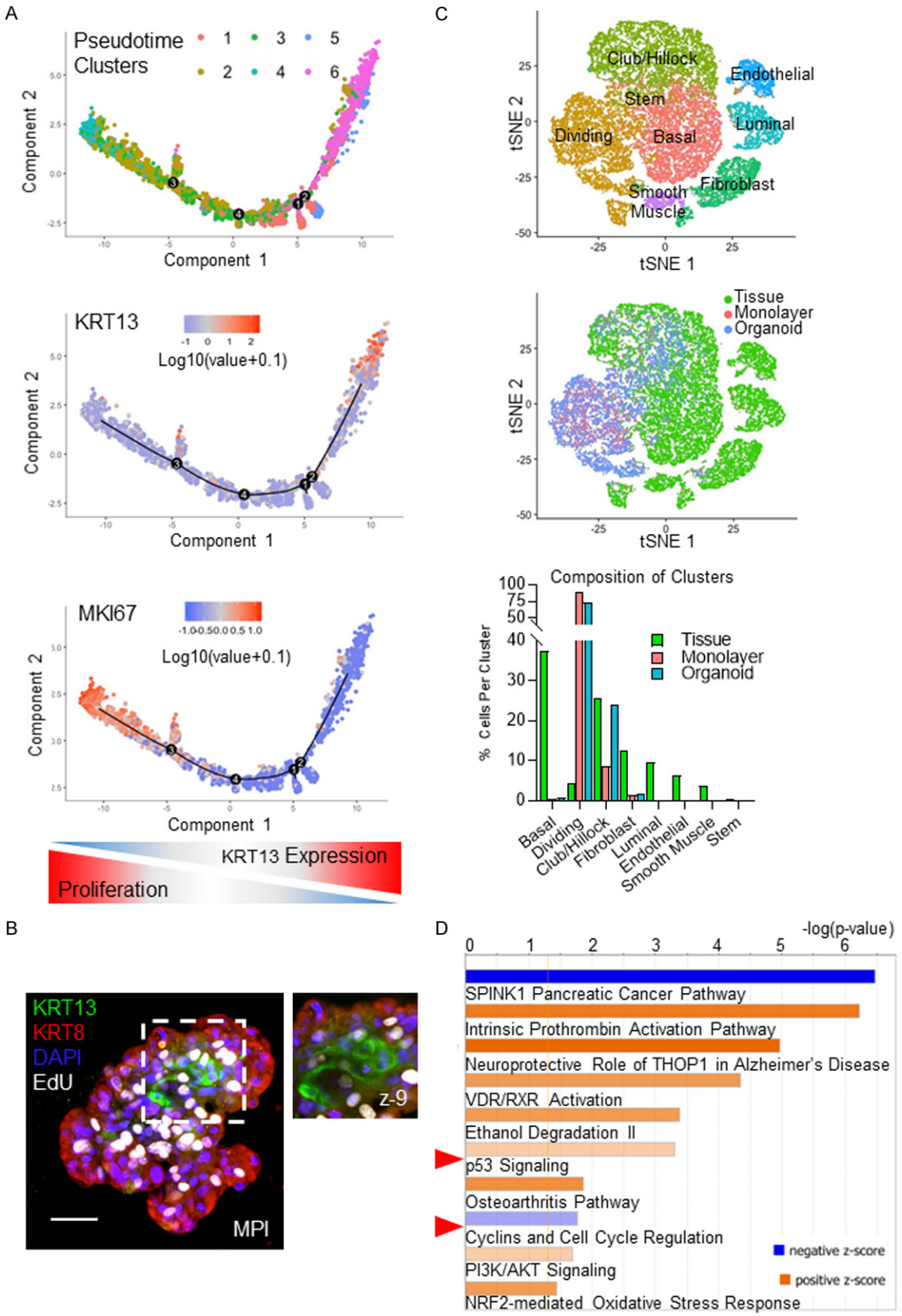


**Figure 4.** KRT13 Subpopulation is Quiescent. A. Pseudotime plots of organoid cell clusters (top), *KRT13* expression (middle) and *MKI67* expression (bottom). B. EdU incorporation and whole-mount immunocytochemistry of day 14

organoid cells derived from patient PrE3 stained for KRT13, KRT8 and DAPI (scale bar = 50 μm). C. T-SNE plot of integrated data sets shows 8 cluster identities (top), sample identities (middle), bar chart depicting the contribution of sample to each cluster (bottom). D. Ingenuity pathway analysis of genes highly expressed by cluster 6.

bedded. 5 μm sections were incubated overnight at 4°C with monoclonal rabbit anti-androgen receptor primary antibody (5153, Abcam, Cambridge UK) used at 1:400 dilution. AR was detected using the rabbit specific HRP/DAB (ABC) detection IHC kit (ab64261, Abcam, Cambridge UK) following the manufacturer's instructions. Slides were counterstained with hematoxylin and imaged on the EVOS FL Auto 2 (Thermo Fisher Scientific, Waltham MA). Shown in Figure S8.

*EdU proliferation assay*

Proliferating cells were visualized using Click-iT EdU AlexaFluor 647 Imaging Kit (Thermo Fisher Scientific, Waltham MA) as previously described [26, 29]. Fully formed day 12 organoids were pulsed with EdU overnight to incorporate into dividing cells. The next day EdU was washed off and cells were given a night to recover prior to staining. On day 14, organoids were collected and fixed using the whole mount protocol described above. To visualize EdU, the Click-iT EdU protocol was followed according to manufacturer specifications. Briefly, cells were permeabilized with Triton-X 100, washed and given the AlexaFluor 647 reaction cocktail. After EdU detection was complete, organoids were stained for KRT13, KRT8 and counterstained with DAPI as described above, shown in **Figure 4**.

*Pathway analysis*

Genes identified as highly expressed in each cluster compared to all other cells were uploaded into the IPA system and analyzed (QIAGEN Inc., Hilden Germany) [30]. Core analysis was performed using the Ingenuity Knowledge Base reference set. We report the enriched canonical pathways with predicted activation Z-scores > |1| and $p$-value < 0.05 (**Figures 4**, S5), and predicted upstream regulators (of class transcription regulator or ligand-dependent nuclear receptor) with predicted activation Z-scores > 2 and $p$-value < 0.05 (Figure S6).

*Data and software availability*

The single cell RNA-seq data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (GEO) and are accessible through GEO Series accession number GSE130318.

## Results

*Organoid cultures contain more distinct populations of prostate epithelial cells than monolayer cultures*

Patient-derived PrE cells can be grown in monolayer or organoid culture conditions as *in vitro* models of prostate cell biology and as useful tools for mechanistic studies. Here we compared the cell populations within these patient-derived models using scRNA-Seq analysis on monolayer epithelial cells and organoids from a single patient (**Figure 1A**). Seurat was used for clustering and analysis of the individual datasets separately, identifying 6 clusters in the organoids and only 3 within the monolayer cells (Figure S2) [24]. Integration of the monolayer and organoid datasets together increased the number of cells to allow for higher-granularity clustering with the same modularity cutoff, thus permitting greater separation of intermediate cell types and subpopulations [22]. Additionally, this integration allowed for the direct comparison of the two datasets on the same t-SNE plot to observe common cell populations across both culture conditions. Together, these factors informed our decision to use an integrated dataset. The resulting 10 clusters are shown on a t-SNE plot (**Figure 1B**) and the contribution of monolayer and organoid cells to each cluster is shown by bar graph (**Figure 1C**). Each cluster had varying representation of cells from both culture conditions, but clusters 1, 3, 8 and 9 were enriched in the organoid sample while cluster 2, 4 and 5 were enriched in the monolayer sample. Genes highly expressed by monolayer (*FST, IGFBP2, SFRP1*) or organoid (*INHBA, IGFBP3, SERPINB1*) cells were also tested by RT-qPCR in pooled sample of leftover scRNA-Seq RNA as a secondary validation to our data (Figure S3). This was supported using 3 additional patient samples, corroborating the trends established by our integrated scRNA-Seq comparison of the two culture conditions.

*Identification of epithelial populations*

The 10 integrated populations each expressed a unique set of biomarkers with some degree of overlap (**Figure 2A**). The cell type identity of each cluster was determined by cross-reference with previously reported gene expression profiles (**Table 3**) and is shown by dot plot (**Figure 2B**). Cluster 0 was identified as a group of quiescent cells expressing high levels of the basal marker *DKK1* [21], moderate levels of luminal markers *KRT8* and *KRT18*, as well as *MMP9*, a factor shown to be highly expressed by PC3 prostasphere cells compared to PC3 monolayer cells [31]. Cluster 1 is marked by the expression of *KRT6A* and *SERPINB13*, denoting its members as a previously described glandular epithelial cell with possible stem characteristics [32]. Cluster 2 contained cells expressing both the basal marker *PDPN* [21] as well as proliferation markers, indicating that this cluster contains proliferating basal cells. Cluster 3 exhibited high levels of cell cycle arrest gene *GAS5* as well as low expression of cell cycle progression genes including *CDC44* and *PCNA*, denoting its possible identity as a population of quiescent cells. This population was also notable for its low expression of epithelial markers, which could signify that these cells are in a relatively undifferentiated state. Cluster 4 constituted another proliferating basal population, marked by its expression of *MKI67* as well as the basal marker *DKK1*. Cluster 5 was also identified as a proliferating population, with high levels of the luminal marker *KRT18*. Cluster 6 exhibited high expression of numerous putative prostate epithelial stem cell markers including *KRT13, SERPINB1, LY6D, PSCA*, *KLK11* and *CSTB* [21, 33, 34]. These markers were also expressed by Cluster 7, with Clusters 6 and 7 differing by their expression levels of basal and luminal markers. Clusters 8 and 9 were difficult to identify as they both contain a very small number of cells. Cluster 8 highly expressed *CTNNB1* and Cluster 9 was enriched for cells expressing *IFI27* and *IFI6*, markers typically expressed by endothelial cells [21]. These markers indicate that Cluster 8 and Cluster 9 may be persisting stromal contaminants, however they also express *PDPN* and *MMP9* which are basal and PC3 prostasphere markers [21, 30].

We observed that the organoid culture condition was enriched for several cell populations when compared to the monolayer condition, such as clusters 1, 3, 7, 8 and 9 (**Figure 1C**). Cells in clusters 1 and 7 were identified as progenitor populations according to their suite of expressed biomarkers (**Figure 2B**). Clusters 8 and 9, although very small, were present in the organoid condition (**Figure 1C**) and expressed some stromal genes (**Figure 2B**), although whether these cells are surviving stroma from the patient or represent epithelial-mesenchymal transition could not be determined from our study. These findings indicate that organoid culture conditions are conducive to the survival and proliferation of cell populations that are underrepresented from samples cultured in monolayer conditions.

*Stem and progenitor populations found in both monolayer and organoid conditions*

Cluster 6 was of considerable interest as it contained markers of a previously reported putative prostate epithelial stem cell population: *SERPINB1, KRT13, LYPD3, PSCA, LY6D, CSTB, LCN2* [21, 33]. It also was marked by low expression of cell cycle genes and high expression of the cycle-arrest gene *GAS5*, implying that the cells are quiescent. Cluster 6 was also the only cluster to express *PSCA*. This expression profile is similar to that of the KRT-13 label-retaining prostate stem cell described by Hu *et al.* [33]. RT-qPCR for *LY6D, KRT13, PSCA*, and *LCN2* validated their expression in both monolayer and organoid samples derived from multiple patients (**Figure 3A**), corroborating the presence of populations observed via scRNA-Seq.

To visualize the stem cell population and confirm protein expression, we stained the original parent tissue, 2D cells and 3D cells for the marker KRT13 (**Figures 3B-F**, S4). Tissue expression of KRT13 in the patient sample used for the scRNA-Seq is shown in **Figure 3B**, and we observed rare islets of KRT13 expression similar to the pattern previously described [21]. This phenomenon was also observed in 2D cells derived from 5 separate patients (**Figures 3C**, S4). At day 8, and in some cases at day 14, the organoids showed one or few KRT13 positive cells (**Figures 3D**, S4B white arrow), consistent with the idea that an organoid is maintained by a single resident stem cell. At day 8, the time point used for scRNA-seq, 3D cells showed KRT13 expressed by a

**Table 4.** Primers used for quantitative real time PCR

| Target Gene | Primer sequence (5'-3') | PCR Product Size (bp) | Exons |
|---|---|---|---|
| RPL13A | F-GGAGCAAGGAAAGGGTCTTAG | | 8 |
| | R-GGTTGCTCTTCCTATTGGTCATA | | |
| LYPD3 | F-GATGCTCCCCGAACAAGATGA | 104 | 2/3 |
| | R-CAGCGAGAATTGTCCGTGGAT | | |
| | PrimerBank ID: 93004087C1 | | |
| LY6D | F-GCTCCCAGACGACATCAGAG | 168 | 1/2 |
| | R-TGTTCGTGGTCTTGCAGAAG | | |
| KRT13 | F-AGGTGAAGATCCGTGACTGG | 134 | 1/2 |
| | R-GATGACCCGGTTGTTTTCAA | | |
| PSCA | F-TGCTGCTTGCCCTGTTGAT | 216 | 1/3 |
| | R-CCTGTGAGTCATCCACGCA | | |
| | PrimerBank ID: 5031995A1 | | |
| LCN2 | F-ACAAAGACCCGCAAAAGATG | 128 | 2/3 |
| | R-GCAACCTGGAACAAAAGTCC | | |
| S100P | F-AAGGTGCTGATGGAGAAGGA | 163 | 1/2 |
| | R-ACTTGTGACAGGCAGACGTG | | |
| SERPINB1 | F-CTGGCGTTGAGTGAGAACAA | 143 | 2/3 |
| | R-TCAACCGTGTTGAAATGGAA | | |
| INHBA | F-GGAGGGCAGAAATGAATGAA | 95 | 2/3 |
| | R-AATCTCGAAGTGCAGCGTCT | | |
| IGFBP3 | F-GTCAACGCTAGTGCCGTCAG | 107 | 1/2 |
| | R-CGGTCTTCCTCCGACTCAC | | |
| FST | F-TCTGCCAGTTCATGGAGGAC | 106 | 1/2 |
| | R-TCCTTGCTCAGTTCGGTCTT | | |
| SFRP1 | F-CTACTGGCCCGAGATGCTTA | 169 | 1/2 |
| | R-GCTGGCACAGAGATGTTCAA | | |
| IGFBP2 | F-CCTCTACTCCCTGCACATCC | 79 | 3/4 |
| | R-CCCGTTCAGAGACATCTTGC | | |
| AR | F-CCAGGGACCATGTTTTGCC | | 1/2 |
| | R-CGAAGACGACAAGATGGACAA | | |
| KRT8 | F-GCTGGTGGAGGACTTCAAGA | 66 | 2/3 |
| | R-TCGTTCTCCATCTCTGTACGC | | |
| KRT18 | F-CACAGTCTGCTGAGGTTGGA | 110 | 6/7 |
| | R-CAAGCTGGCCTTCAGATTTC | | |
| KRT5 | F-ATCGCCACTTACCGCAAGC | 110 | 7/9 |
| | R-CCATATCCAGAGGAAACACTGC | | |

small population of cells (**Figure 3D**), confirming our scRNA-Seq population size. When staining was performed on day 14, we observed that the KRT13+ cells were located on the interior of the spheroid structure (**Figures 3E**, S4B yellow arrow) which may correlate to these stem cells generating the inner luminal cells. There were also instances of multiple clustered KRT13+ cells that may represent different stages of stem/progenitor hierarchy (**Figures 3F**, S4B).

*KRT13+ cells are quiescent and exhibit similar gene expression as KRT13+ cells derived from tissue*

Stem cells are known to rarely proliferate in tissue and undergo asymmetric division in culture to maintain their quiescent state [33]. To visualize cellular state, Monocle was used to create a differentiation trajectory of the organoid sample in pseudotime based off gene expression. Pseudotime analysis showed that *KRT13*-expressing cells were found in a different state than *MKI67*-positive cells (**Figure 4A**). To validate this, we performed an EdU incorporation assay on fully-formed organoids. Organoids were pulsed with EdU overnight to mark actively dividing cells followed by fixation and staining for KRT13. As expected, KRT13+ cells were EdU-negative, indicating that they proliferate slowly (**Figure 4B**). It has been shown that knockdown of KRT13 in undifferentiated prostasphere culture leads to diminished sphere formation and self-renewal [33]. Our findings provide further support for KRT13 as a quiescent prostate stem marker in the *in vitro* context and agree with the previously characterized role of KRT13+ cells in stem cell maintenance.

Recently, Henry *et al*. identified a KRT13+ "hillock" epithelial cell population within prostate tissue which may harbor stem characteristics *in vivo* [21]. Our monolayer and organoid data were integrated and compared with a publicly available human prostate single cell data set (GSE117403) from this study (**Figure 4C**). The *KRT13*+ hillock population was present in all three samples, supporting our identification of Cluster 6 in our monolayer and organoid integrated analysis. Of note, there was enrichment for this cluster in the organoid sample to comparable levels as what is observed in tissue. Additionally, integrated analysis with the pros-

tate tissue supported that clusters 8 and 9 were stromal cell types.

To understand what regulatory pathways may be active in the KRT13+ cells, we performed Ingenuity Pathway Analysis (IPA) on genes highly expressed by this cluster. IPA Canonical Pathways that were significantly enriched and activated (positive z-score) or inactivated (negative z-score) are shown (**Figure 4D**). Of note, p53 signaling was predicted to be activated (z-score = 1.00) while cyclins were predicted to be inactivated (z-score = -1.00). Canonical pathways for other clusters are shown on Figure S5. IPA upstream regulator analysis was used to identify the transcriptional regulators and nuclear receptors predicted to be active in Cluster 6 based upon genes upregulated in the cluster. KLF4 was projected to be active (Figure S6), which has recently been reported to regulate prostate stem cell homeostasis in mice and has been implicated with KRT13 expression [35]. Similarly, GLI1 and CTNNB1 also exhibited positive z-scores in this analysis, implying the activity of the Shh and Wnt morphogenic pathways; both pathways have been reported as significant contributors to prostate tissue regeneration in the murine context [36, 37].

## Discussion

To experimentally investigate prostate function and disease, researchers utilize cell lines, primary cells and organoid culture for *in vitro* modeling. While it is known that monolayer cultures contain rare populations of stem and progenitor cells, most 2D cells observed consist of transit-amplifying epithelial phenotypes [11-14]. Prostate epithelial organoids grow out from a single cell and differentiate into basal and luminal cell types while maintaining a resident stem/progenitor cell that originated the organoid [15, 16, 18, 33]. While these basic profiles of *in vitro* epithelial cells are known, recent utilization of scRNA-Seq analysis has identified potentially novel populations of prostate epithelial cells from human prostate tissues [21]. To our knowledge, scRNA-Seq analysis of human primary prostate organoid models has not been performed. Using an integrated analysis, we identified 10 cell populations in monolayer and organoid culture, thus identifying more cell types than previously reported *in vitro*. Our analyses revealed populations of cells at varying levels of differentiation along the basal and luminal lineages. More specifically, KRT13+ putative stem cell populations were identified in both monolayer and organoid conditions and had expression profiles similar to those previously reported for KRT13+ cells in tissue [21].

We identified three putative stem and progenitor populations by high *KRT13* or *KRT6A* expression, which have been previously described to mark regenerative cells in tissue and prostasphere culture [21, 32, 33]. The three populations also showed expression of *LY6D, LYPD3,* and *CSTB* and one specific cluster of the three had high expression of *PSCA*. *PSCA* and *KRT13* were previously described by Henry *et al.* to mark two previously unknown clusters of epithelial cells, termed “club” and “hillock”, based off of their similarity to immunomodulatory and progenitor-like cells found in the mouse lung, respectively [21]. In contrast to the tissue-isolated cells in that study, which had distinct populations that were either *PSCA*+ (club) or *KRT13*+ (hillock), the *in vitro* cells had one $PSCA^{High}/KRT13^{High}$ cluster and one $PSCA^{Low}/KRT13^{High}$ cluster. KRT13 protein expression did vary greatly between organoid samples at day 8 and day 14, with day 14 organoids displaying multiple KRT13+ cells (**Figures 3F**, S4B). The two $PSCA^{Low}$ clusters, 1 and 7, expressed stem markers *LY6D*, *LYPD3*, *KLK11,* and *CSTB* and also showed co-expression of basal and luminal markers. These cells may be more-committed basal and luminal progenitors that reside below the $KRT13^{High}/PSCA^{High}$ cells on the stem-hierarchy. The degree of *KRT13* expression has been shown to positively correlate with the level of stemness in other contexts, where differentiated daughters show less expression as they divide and differentiate [33]. This was observed in some organoids that showed double positive KRT8/KRT13 cells (**Figure 3F**) and is likely the explanation for the heterogeneity of KRT13 expression seen in the day 14 organoids (**Figures 3F**, S4).

KRT13+ hillock cells of the human prostate were named so because of their similarities to cells in the mouse lung that show a progenitor-like phenotype [21, 38]. In the organoids we observed some evidence of KRT13+ hillock cells exhibiting a capacity for branching morphogenesis due to their localization to the interior of branching organoids (Figure S4 red

arrow) which may be similar to a "hillock" region seen in patient tissue. Recently, FACS-sorted mouse prostate epithelial cells that were analyzed using Fluidigm qPCR showed *LY6D* expression in a population of organoid-forming cells found in both the luminal and basal compartment [39]. The LY6D+ cells formed solid, acinar or translucent organoids, similar to the organoid morphologies that we observed in our culture (**Figures 3E**, **3F**, S4). Human DLK1+ prostate basal cells have also been shown to form solid spheroids, spheroids with lumens and spheroids with tubules. *LY6D* and *KRT13* were co-expressed in population 6 and 7 in our 3D cells and had similar expression profiles to the reported DLK1+ cells [34]. *PPL* was highly expressed in cluster 6 and *KLK11* was highly expressed in clusters 1 and 6. These genes were published as being expressed by the proximal niches that maintain epithelial flow in human prostate and contain possible DLK1+ prostate stem cells [34]. Taking these reports together, there is substantial support for the KRT13+ cells to be a stem cell population in organoid culture and indicate that this model system would be useful for studying hillock biology and proximal niches *in vitro*.

It is important to note that the monolayer culture condition yielded numerous KRT13+ cells that exhibited a stretched morphology typical to terminally differentiated cells and notably different from other KRT13+ cell shapes (Figure S4, yellow arrow). These data introduce the possibility that KRT13 may be an excellent stem cell marker when used in combination with secondary markers like LY6D and PSCA, but on its own it could mark a broad population that contains epithelial stem cells along with a small population of unhealthy squamous cells or differentiated daughter cells in this context.

The lack of expression of luminal markers *AR* and *KLK3* in the organoids by scRNA-Seq should be interpreted with caution. A limitation of all single cell sequencing technologies is that it captures only 10-20% of transcripts per cell (the 10 × Chromium Single Cell 3' v2 Kit used here has a capture rate of 14-15%), thus absence of the gene in the analysis is not conclusive evidence that it is not expressed. We did detect low *AR* expression by RT-qPCR of whole organoids (Figure S7) and confirmed AR protein expression was present in cells of day 14 organoids (Figure S7). Growth of fully differentiated luminal cells with robust AR remains a challenge in the field of human prostate organoids.

Overall, we observed that both monolayer and organoid culture are capable of cultivating a rare population of putative stem-like cells that are marked by high expression of *KRT13*, *LY6D*, *LYPD3*, and *PSCA*. These cells are similar to cells found in the basal and luminal compartments of mouse prostate and human club/hillock regions, implying that we are describing the *in vitro* isolation and culture of a population of stem-like cells that reside in tissue. There were populations of intermediate cells that were overrepresented in the organoid condition, as well as a *KRT6A*+ progenitor population that was enriched in 3D. Our catalog of unique populations in these two *in vitro* models shows the preservation of a specific stem-like cell population as well as provides an in-depth atlas of the populations present in both monolayer and organoid models. This can serve as a valuable resource to the field, allowing for a deeper understanding of which cells are present in benign model systems and how they may change between *in vitro* and *in vivo* conditions.

**Acknowledgements**

We thank the UIC Biorepository members, Dr. Klara Valyi-Nagy and Alex Susma, and the urologists, Drs. Michael Abern, Daniel Moreira and Simone Crivallero, for facilitation of the tissue acquisition for the primary cell cultures. We thank the UIC Urology patients for donating their tissue to research. We thank Dr. Laura Nolden, Dr. Mark Gerber, and Melissa Leone at 10x with their help getting set up with the machine. We thank Dr. Alvaro Hernandez, Chris Wright, Jenny Zadeh, and Jessica Holmes and the UIUC DNA services team for help with sequencing and help with the Cell Ranger pipeline. We thank Magdalena Rogozinska and Dr. Gayatry Mohapatra for their help with the TapeStation quantification. We thank Gervais Henry from UT Southwestern for his help with Seurat and Monocle. We thank Dr. Ke Ma with UIC Fluorescence Imaging Core for assisting with confocal. This work was funded, in part, by the Department of Defense Prostate Cancer Research Program Health Disparities Idea Award PC121923 (Nonn), the University of Chicago MCB Training Grant T32 GM007183 (Moline), and the UIC Center for Clinical and

Translation Science Pre-doctoral Education for Clinical and Translational Scientists (PECTS) Program (McCray).

## Disclosure of conflict of interest

None.

**Address correspondence to:** Dr. Larisa Nonn, Department of Pathology, University of Illinois at Chicago, 840 S Wood St. Chicago, IL, Chicago 60612, Illinois, USA. Tel: 312-996-0194; E-mail: lnonn@uic.edu

## References

[1] Siegel R. Prostate cancer statistics 2017. Ca Cáncer Journal 2017; 67: 7-30.
[2] Lim KB. Epidemiology of clinical benign prostatic hyperplasia. Asian J Urol 2017; 4: 148-151.
[3] Sobel RE and Sadar MD. Cell lines used in prostate cancer research: a compendium of old and new lines--part 1. J Urol 2005; 173: 342-59.
[4] Sobel RE and Sadar MD. Cell lines used in prostate cancer research: a compendium of old and new lines--part 2. J Urol 2005; 173: 360-72.
[5] Peehl DM. Human prostatic epithelial cells. Cultures 2002; 8: 171-194.
[6] Peehl DM. Growth of prostatic epithelial and stromal cells in vitro. Methods Mol Med 2003; 81: 41-57.
[7] Peehl DM and Stamey TA. Serum-free growth of adult human prostatic epithelial cells. In Vitro Cell Dev Biol 1986; 22: 82-90.
[8] Niranjan B, Lawrence MG, Papargiris MM, Richards MG, Hussain S, Frydenberg M, Pedersen J, Taylor RA and Risbridger GP. Primary culture and propagation of human prostate epithelial cells. Methods Mol Biol 2013; 945: 365-382.
[9] Peehl DM, Wong ST and Stamey TA. Clonal growth characteristics of adult human prostatic epithelial cells. In Vitro Cell Dev Biol 1988; 24: 530-6.
[10] Wise C. Epithelial cell culture protocols-primary culture and propagation of human prostate epithelial cells (pg 365-382). 2002; 188: 365-382.
[11] Uzgare AR, Xu Y and Isaacs JT. In vitro culturing and characteristics of transit amplifying epithelial cells from human prostate tissue. J Cell Biochem 2004; 91: 196-205.
[12] Peehl DM. Are primary cultures realistic models of prostate cancer? J Cell Biochem 2004; 91: 185-195.
[13] Bühler P, Wolf P, Katzenwadel A, Schultze-Seemann W, Wetterauer U, Freudenberg N, Elsässer-Beile U. Primary prostate cancer cultures are models for androgen-independent transit amplifying cells. Oncol Rep 2010; 23: 465-70.
[14] Litvinov IV, Vander Griend DJ, Xu Y, Antony L, Dalrymple SL, Isaacs JT. Low-calcium serum-free defined medium selects for growth of normal prostatic epithelial stem cells. Cancer Res 2006; 85: 1-27.
[15] Karthaus WR, Iaquinta PJ, Drost J, Gracanin A, van Boxtel R, Wongvipat J, Dowling CM, Gao D, Begthel H, Sachs N, Vries RGJ, Cuppen E, Chen Y, Sawyers CL and Clevers HC. Identification of multipotent luminal progenitor cells in human prostate organoid cultures. Cell 2014; 159: 163-175.
[16] Drost J, Karthaus WR, Gao D, Driehuis E, Sawyers CL, Chen Y and Clevers H. Organoid culture systems for prostate epithelial and cancer tissue. Nat Protoc 2016; 11: 347-58.
[17] Clevers H. Modeling development and disease with organoids. Cell 2016; 165: 1586-1597.
[18] Chua CW, Shibata M, Lei M, Toivanen R, Barlow LJ, Bergren SK, Badani KK, McKiernan JM, Benson MC, Hibshoosh H and Shen MM. Single luminal epithelial progenitors can generate prostate organoids in culture. Nat Cell Biol 2014; 16: 951-61, 1-4.
[19] Long RM, Morrissey C, Fitzpatrick JM, William R. Prostate epithelial cell differentiation and its relevance to the understanding of prostate cancer therapies. Clin Sci (Lond) 2005; 108: 1-11.
[20] Toivanen R and Shen MM. Prostate organogenesis: tissue induction, hormonal regulation and cell type specification. Development 2017; 144: 1382-1398.
[21] Henry GH, Malewska A, Joseph DB, Malladi VS, Lee J, Torrealba J, Mauck RJ, Gahan JC, Raj GV, Roehrborn CG, Hon GC, Macconmara MP, Reese JC, Hutchinson RC, Vezina CM and Strand DW. A cellular anatomy of the normal adult human prostate and prostatic urethra. Cell Rep 2018; 25: 3530-3542, e5.
[22] Hoffman P, Satija R, Papalexi E, Smibert P and Butler A. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 2018; 36: 411-420.
[23] Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A and McCarroll SA. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 2015; 161: 1202-1214.
[24] Satija R, Farrell JA, Gennert D, Schier AF and Regev A. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 2015; 33: 495-502.

[25] Macosko EZ, Basu A, Regev A, Mccarroll Correspondence SA, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK and Mccarroll SA. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 2015; 161: 1202-1214.

[26] McCray T, Richards Z, Marsili J, Prins GS and Nonn L. Handling and assessment of human primary prostate organoid culture. J Vis Exp 2019.

[27] Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS and Rinn JL. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 2014; 32: 381-386.

[28] Livak KJ and Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001; 25: 402-408.

[29] Richards Z, McCray T, Marsili J, Zenner ML, Garcia J, Manlucu JT, Voisine C, Murphy AB, Prins GS, Murray M, Kajdacsy-Balla A, Abdulkadir SA and Nonn L. Prostate stroma increases the viability and maintains the branching phenotype of human prostate organoids. iScience 2019; 12: 304-317.

[30] Kramer A, Green J, Pollard J Jr and Tugendreich S. Causal analysis approaches in Ingenuity pathway analysis. Bioinformatics 2014; 30: 523-530.

[31] Portillo-Lara R and Alvarez MM. Enrichment of the cancer stem phenotype in sphere cultures of prostate cancer cell lines occurs through activation of developmental pathways mediated by the transcriptional regulator ΔNp63α. PLoS One 2015; 10: e0130118.

[32] Schmelz M, Moll R, Hesse U, Prasad AR, Gandolfi JA, Hasan SR, Bartholdi M and Cress AE. Identification of a stem cell candidate in the normal human prostate gland. Eur J Cell Biol 2005; 84: 341-354.

[33] Hu WY, Hu DP, Xie L, Li Y, Majumdar S, Nonn L, Hu H, Shioda T and Prins GS. Isolation and functional interrogation of adult human prostate epithelial stem cells at single cell resolution. Stem Cell Res 2017; 23: 1-12.

[34] Moad M, Hannezo E, Buczacki SJ, Robson CN, Simons BD, Correspondence RH, Wilson L, El-Sherif A, Sims D, Pickard R, Wright NA, Williamson SC, Turnbull DM, Taylor RW, Greaves L and Heer R. Multipotent basal stem cells, maintained in localized proximal niches, support directed long-ranging epithelial flows in human prostates multipotent basal stem cells, maintained in localized proximal niches, support directed long-ranging epithelial flows. Cell Rep 2017; 20: 1609-1622.

[35] Xiong X, Schober M, Tassone E, Khodadadi-Jamayran A, Sastre-Perona A, Zhou H, Tsirigos A, Shen S, Chang M, Melamed J, Ossowski L and Wilson EL. KLF4, a gene regulating prostate stem cell homeostasis, is a barrier to malignant progression and predictor of good prognosis in prostate cancer. Cell Rep 2018; 25: 3006-3020, e3007.

[36] Lee SH, Johnson DT, Luong R, Yu EJ, Cunha GR, Nusse R and Sun Z. Wnt/beta-catenin-responsive cells in prostatic development and regeneration. Stem Cells 2015; 33: 3356-3367.

[37] Peng YC and Joyner AL. Hedgehog signaling in prostate epithelial-mesenchymal growth regulation. Dev Biol 2015; 400: 94-104.

[38] Montoro DT, Haber AL, Biton M, Vinarsky V, Lin B, Birket SE, Yuan F, Chen S, Leung HM, Villoria J, Rogel N, Burgin G, Tsankov AM, Waghray A, Slyper M, Waldman J, Nguyen L, Dionne D, Rozenblatt-Rosen O, Tata PR, Mou H, Shivaraju M, Bihler H, Mense M, Tearney GJ, Rowe SM, Engelhardt JF, Regev A and Rajagopal J. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature 2018; 560: 319-324.

[39] Barros-Silva JD, Linn DE, Steiner I, Guo G, Ali A, Pakula H, Ashton G, Peset I, Brown M, Clarke NW, Bronson RT, Yuan GC, Orkin SH, Li Z and Baena E. Single-cell analysis identifies LY6D as a marker linking castration-resistant prostate luminal cells to prostate progenitors and cancer. Cell Rep 2018; 25: 3504-3518.

[40] Zhang D, Park D, Zhong Y, Lu Y, Rycaj K, Gong S, Chen X, Liu X, Chao HP, Whitney P, Calhoun-Davis T, Takata Y, Shen J, Iyer VR, Tang DG. Stem cell and neurogenic gene-expression profiles link prostate basal cells to aggressive prostate cancer. Nat Commun 2016; 7: 10798.

[41] Wang Y, Hayward S, Cao M, Thayer K, Cunha G. Cell differentiation lineage in the prostate. Differentiation 2001; 68: 270-9.

**Table S1.** Genes Highly Expressed by Monolayer, Organoid and Integrated Clusters

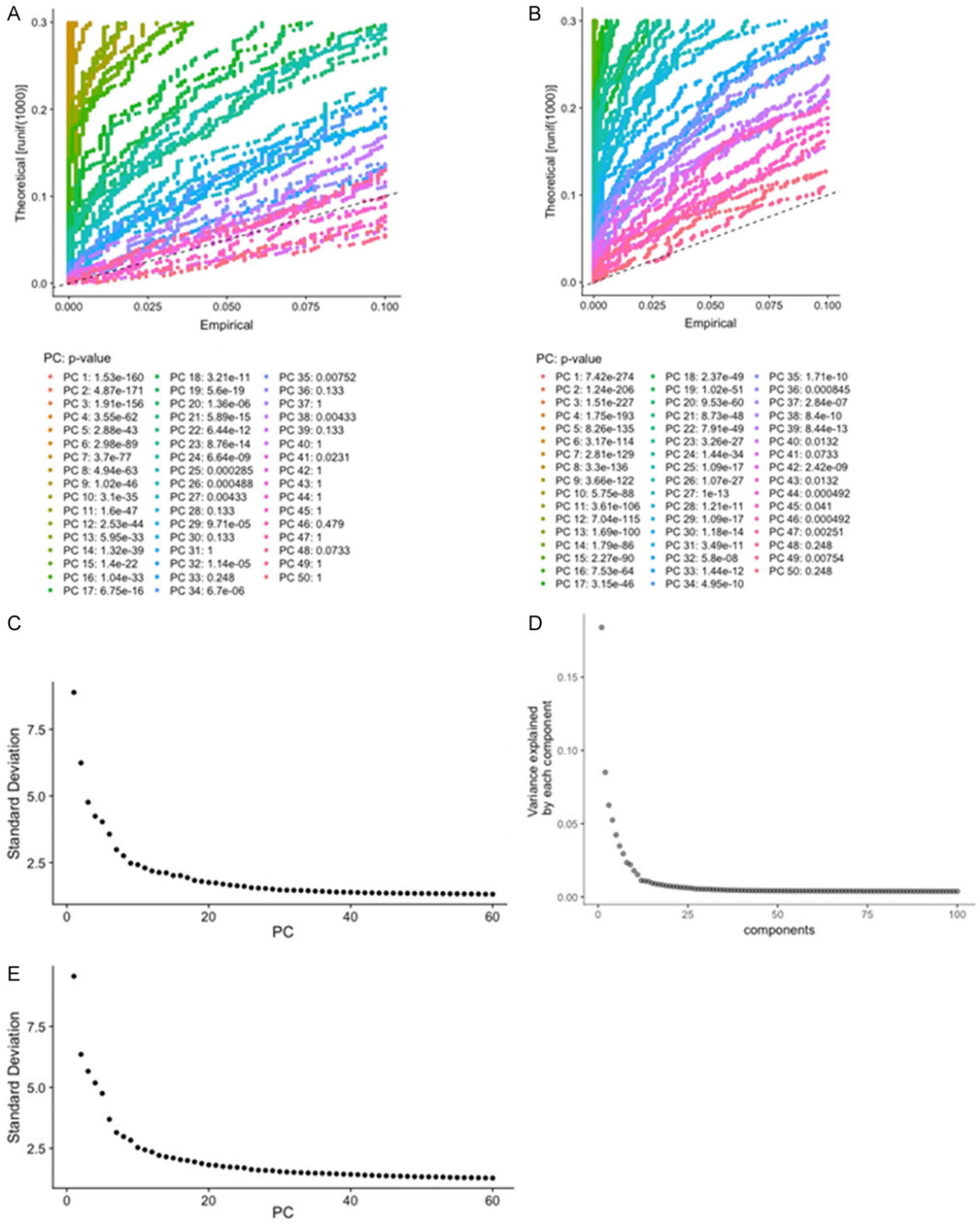


**Figure S1.** Principal component plots used to determine the significant number of components for Seurat and Monocle analyses. A. Jackstraw plot of PCs for monolayer cells. B. Jackstraw plot of PCs for organoid cells. C. Elbow plot of PCs for integrated monolayer and organoid cells. D. Elbow plot of PCs for organoid pseudotime clustering. E. Elbow plot of PCs for integrated monolayer, organoid, and tissue cells.

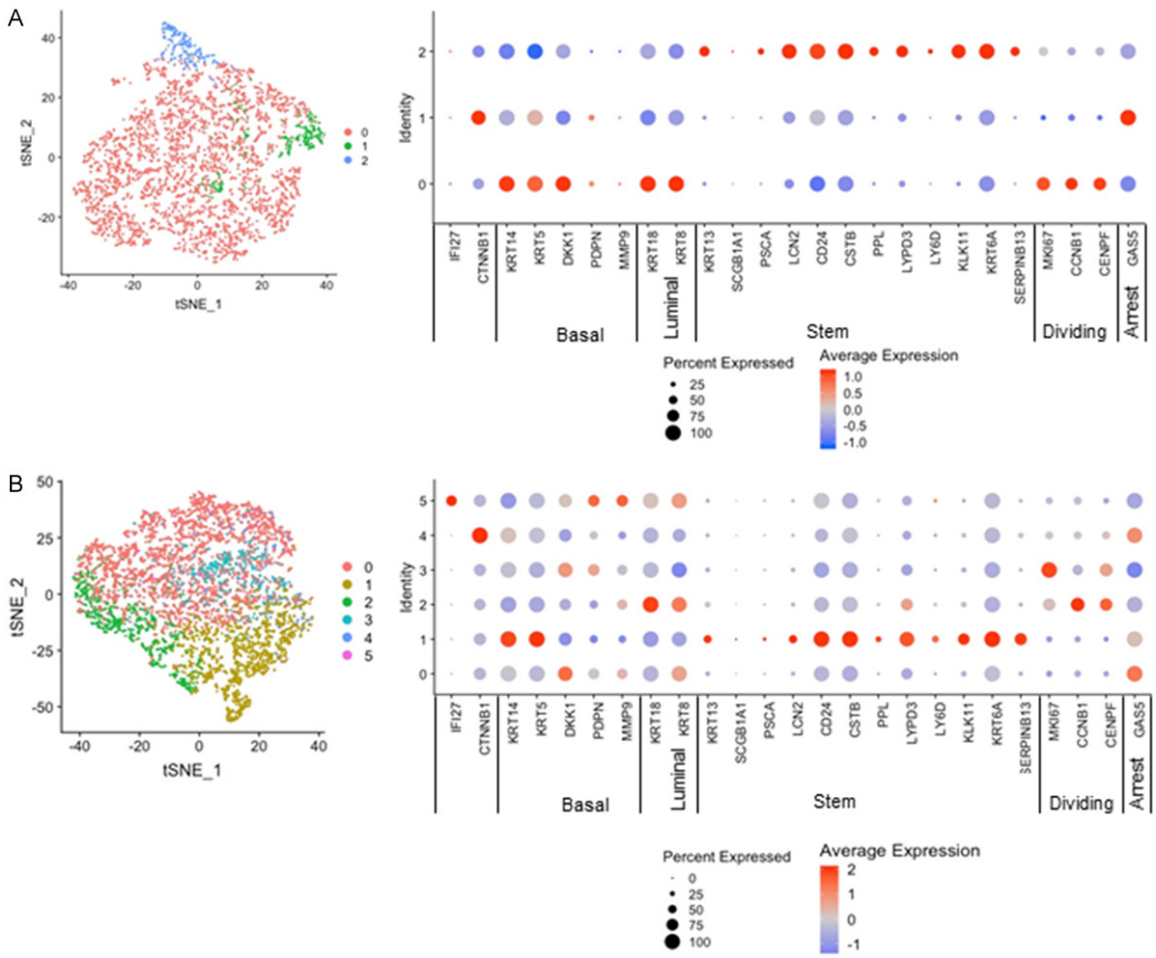


**Figure S2.** A, B. t-SNE plots (left) for unintegrated monolayer and organoid cells, respectively, M > 0.80 (left). Dot plot showing expression of genes of interest across the 3 clusters (right) (related to **Figure 1**).

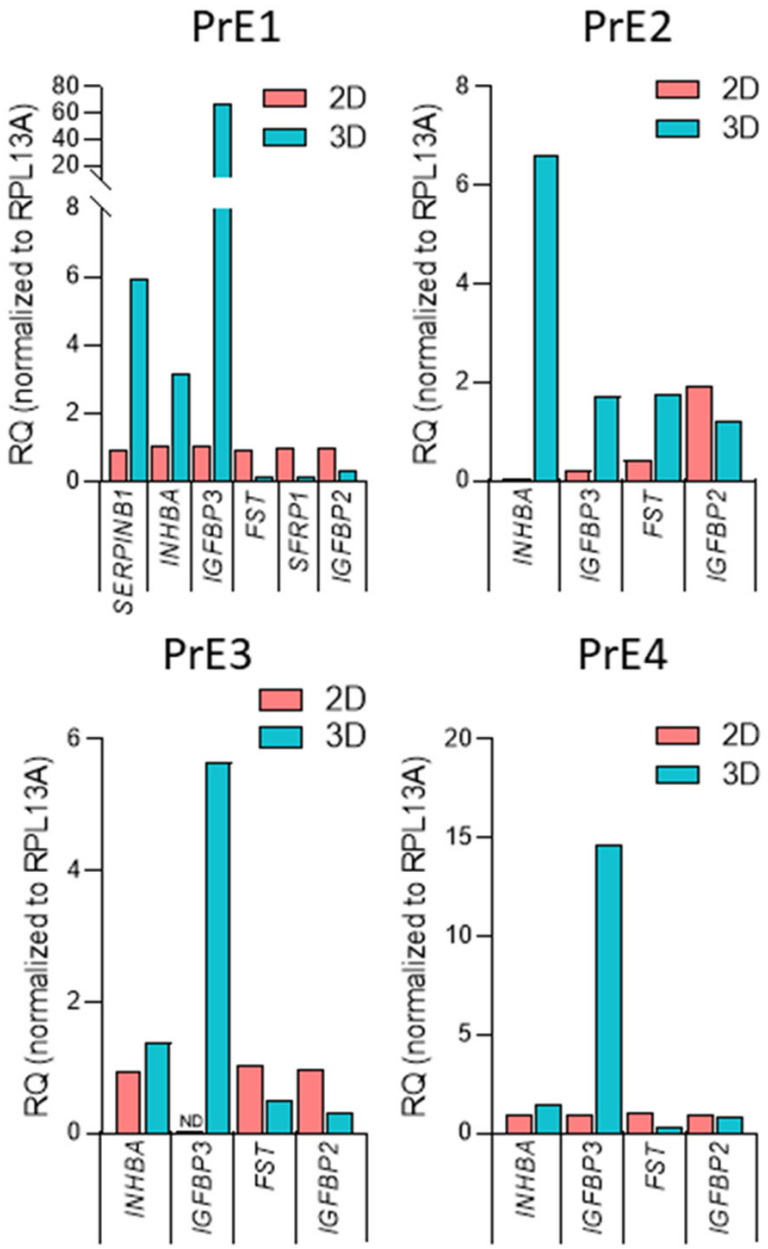


**Figure S3.** RT-qPCR validation of genes highly expressed by organoids or monolayer cells in pooled, non-single cell RNA samples derived from 4 patients (related to **Figure 2**). PrE1 was the patient used for scRNA-Seq.

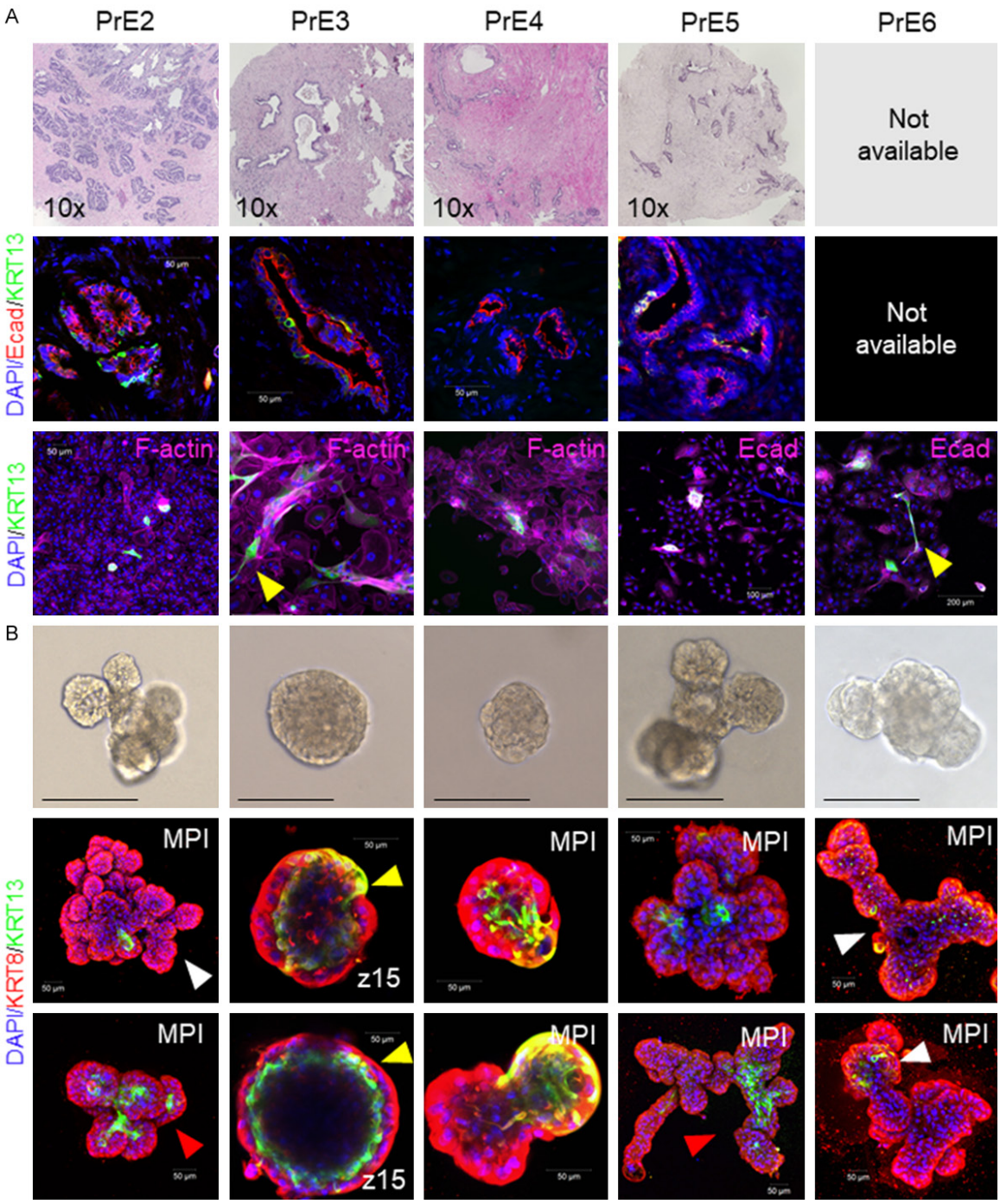


**Figure S4.** A. H&E of patient tissue (top), immunofluorescent staining of KRT13, DAPI and E-cadherin on patient tissue (middle, scale bar = 50 µm), immunocytochemistry of monolayer cells derived from patient tissue stained for KRT13, DAPI and Phalloidin or E-cadherin (bottom, scale bar = 50 µm). B. Bright field image of day 14 organoids derived from patient tissue (top, scale bar = 200 µm). Whole-mount immunocytochemistry of day 14 organoids stained for KRT13, DAPI, and KRT8, two examples shown per patient (bottom, scale bar = 50 µm). Related to **Figure 3**.

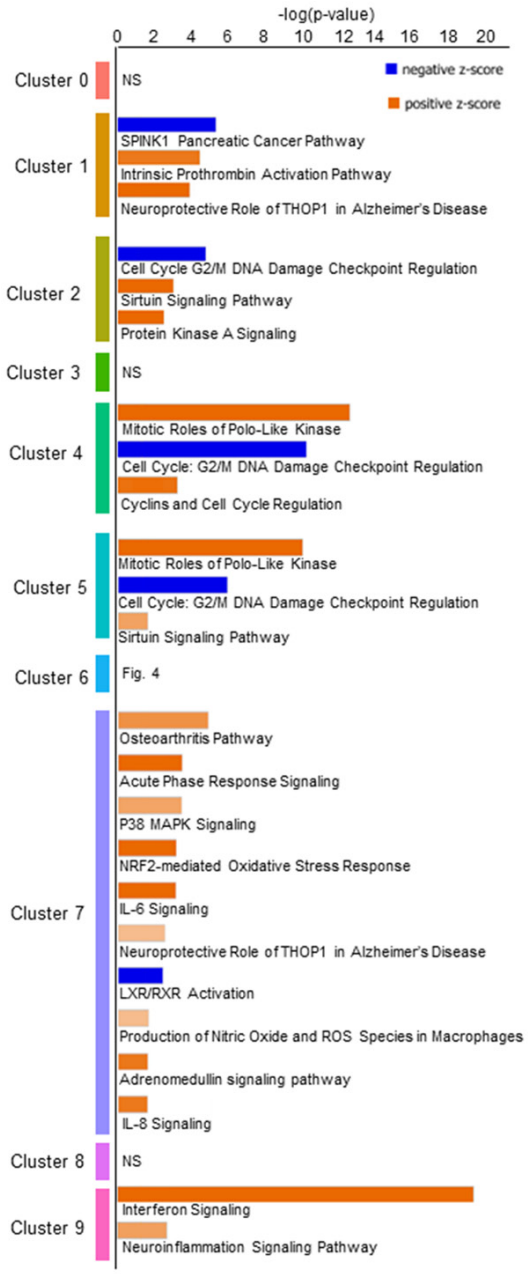


**Figure S5.** Ingenuity pathway analysis of genes highly expressed by each cluster, showing the IPA Canonical Pathways which were significantly enriched and activated (positive Z-score) or inactivated (negative Z-score). Cluster 0, 3 and 8 had ≤ 21 highly genes expressed in each cluster, so these clusters did not yield significant results for the IPA Canonical Pathways analysis (related to **Figure 4**).

Cluster 1

| Upstream Regulator | Z-Score | P-value |
|---|---|---|
| EHF | 3 | 2.65E-12 |
| SMARCA4 | 2 | 1.20E-03 |

Cluster 2

| Upstream Regulator | Z-Score | P-value |
|---|---|---|
| E2F3 | 2.828 | 1.84E-12 |
| RARA | 2.449 | 4.29E-06 |
| ESR1 | 2.449 | 7.96E-07 |

Cluster 4

| Upstream Regulator | Z-Score | P-value |
|---|---|---|
| FOXM1 | 4.125 | 4.06E-30 |
| ESR1 | 3.973 | 1.26E-12 |
| MITF | 3.873 | 3.44E-18 |
| RARA | 3.873 | 4.62E-14 |
| FOXO1 | 3.85 | 1.81E-19 |
| E2F3 | 3.606 | 2.26E-16 |
| TAL1 | 2.828 | 6.42E-06 |
| MYBL2 | 2.789 | 6.20E-20 |
| ATF6 | 2.236 | 3.32E-07 |
| CREB1 | 2 | 1.33E-05 |

Cluster 5

| Upstream Regulator | Z-Score | P-value |
|---|---|---|
| ESR1 | 3.538 | 6.04E-13 |
| RARA | 3.317 | 4.10E-10 |
| FOXM1 | 3.234 | 2.21E-18 |
| MITF | 3.162 | 1.43E-11 |
| FOXO1 | 2.961 | 7.59E-11 |
| E2F3 | 2.828 | 2.08E-11 |
| TAL1 | 2.449 | 1.54E-04 |
| MYBL2 | 2.2 | 1.58E-12 |

Cluster 6

| Upstream Regulator | Z-Score | P-value |
|---|---|---|
| EHF | 4.472 | 1.51E-22 |
| TP63 | 2.03 | 1.09E-13 |
| EZH2 | 2.316 | 4.74E-11 |
| PGR | 2.376 | 2.86E-09 |
| FOXO1 | 2.472 | 1.33E-07 |
| IRF2 | 2.646 | 3.16E-07 |
| NFKBIZ | 2.159 | 5.03E-07 |
| CTNNB1 | 2.226 | 1.02E-06 |
| TP53 | 2.538 | 3.91E-06 |
| KLF4 | 2.183 | 4.00E-06 |
| TP73 | 2.545 | 5.63E-06 |
| FOXO3 | 2.362 | 7.05E-06 |
| CEBPB | 2.009 | 1.02E-05 |
| SMARCA4 | 3.302 | 4.75E-05 |
| PPRC1 | 2.191 | 2.25E-04 |
| STAT1 | 2.183 | 5.38E-04 |
| RELA | 2.145 | 1.52E-03 |
| GLI1 | 2.417 | 2.23E-03 |
| SREBF1 | 2.162 | 2.80E-03 |
| EP300 | 2.219 | 3.47E-03 |
| POU2F2 | 2 | 8.51E-03 |

Cluster 7

| Upstream Regulator | Z-Score | P-value |
|---|---|---|
| NUPR1 | 3.162 | 9.87E-04 |
| EHF | 2.828 | 8.80E-09 |
| SMARCA4 | 2.813 | 1.15E-07 |
| GLI1 | 2.745 | 3.42E-07 |
| PGR | 2.36 | 1.97E-17 |
| ECSIT | 2.219 | 4.80E-07 |
| SP1 | 2.183 | 1.16E-10 |
| SPI1 | 2.065 | 3.65E-04 |
| FOXO1 | 2.057 | 6.91E-08 |

Cluster 9

| Upstream Regulator | Z-Score | P-value |
|---|---|---|
| IFNL1 | 4.298 | 4.92E-39 |
| IFNG | 4.274 | 6.17E-33 |
| PRL | 4.123 | 3.22E-32 |
| IFNA2 | 3.674 | 3.22E-32 |
| SPI1 | 3.317 | 7.44E-20 |
| TNF | 2.62 | 1.91E-14 |
| IRF3 | 2.385 | 6.38E-11 |
| SMARCA4 | 2.236 | 7.99E-07 |
| IFNL4 | 2.209 | 2.92E-12 |
| STAT1 | 2.098 | 3.11E-23 |

**Figure S6.** IPA upstream regulator analysis performed using the top genes expressed by each cluster compared to all other cells, regulator analysis predicts the transcriptional regulators and nuclear receptors that would be active in each cluster (related to **Figure 4**). Cluster 0, 3 and 8 had ≤ 21 highly genes expressed in each cluster, so these clusters did not yield results for the IPA upstream regulator analysis.

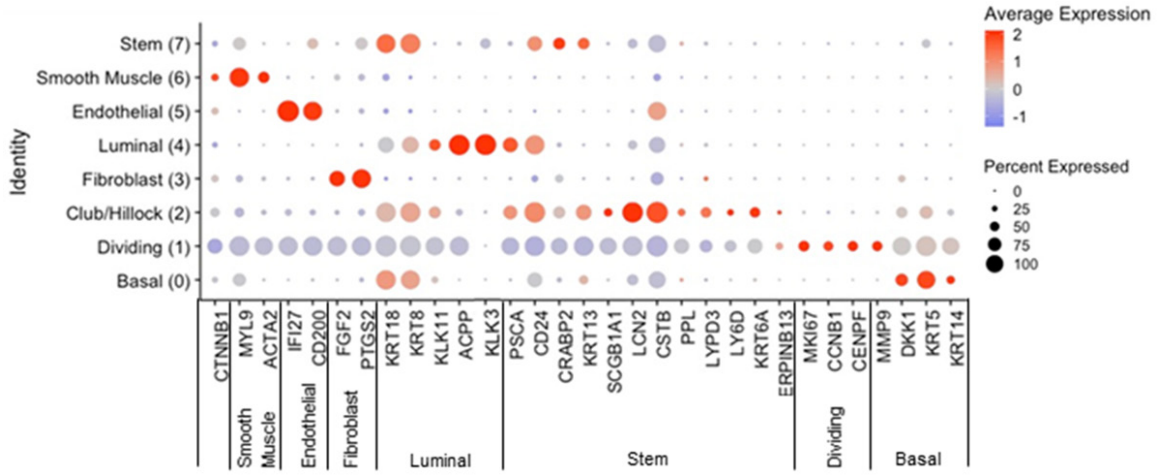


**Figure S7.** Dot plot showing expression of genes of interest across the 8 clusters for integrated monolayer cells, organoid cells and publically available patient tissue dataset used for cluster identification. Cluster names have been assigned by expression profile compared to tissue.

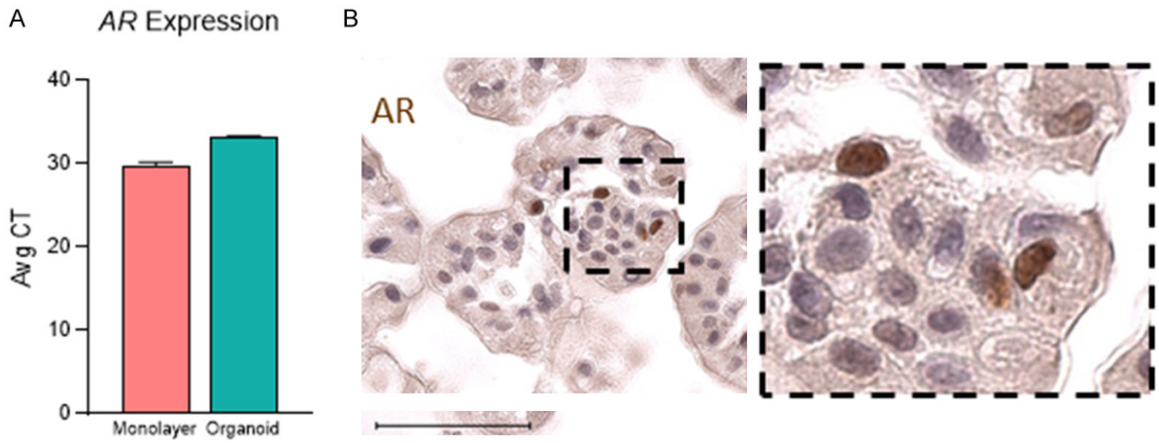


**Figure S8.** AR expression in organoid culture by RT-qPCR, (A) or by immunohistochemistry, (B) in day 14 PrE5 organoids.