

SCIENTIFIC REPORTS



OPEN

Whole Transcriptome Analysis of Renal Intercalated Cells Predicts Lipopolysaccharide Mediated Inhibition of Retinoid X Receptor alpha Function

Vijay Saxena¹, James Fitch², John Ketz³, Peter White^{1,2,4}, Amy Wetzel², Melinda A. Chanley³, John D. Spencer^{3,4}, Brian Becknell^{3,4}, Keith R. Pierce⁵, Sam W. Arregui¹, Raoul D. Nelson⁶, George J. Schwartz⁷, Victoria Velazquez⁸, Logan A. Walker⁹, Xi Chen¹⁰, Pearly Yan^{10,11}, David S. Hains¹ & Andrew L. Schwaderer¹

The renal collecting duct consists of intercalated cells (ICs) and principal cells (PCs). We have previously demonstrated that collecting ducts have a role in the innate immune defense of the kidney. Transcriptomics is an important tool used to enhance systems-level understanding of cell biology. However, transcriptomics performed on whole kidneys provides limited insight of collecting duct cell gene expression, because these cells comprise a small fraction of total kidney cells. Recently we generated reporter mouse models to enrich collecting duct specific PC and ICs and reported targeted gene expression of anti-microbial peptide genes. Here we report transcriptomics on enriched ICs and PCs and performed a pilot study sequencing four single ICs. We identified 3,645 genes with increased relative expression in ICs compared to non-ICs. In comparison to non-PCs, 2,088 genes had higher relative expression in PCs. IC associated genes included the innate interleukin 1 receptor, type 1 and the antimicrobial peptide (AMP) adrenomedullin. The top predicted canonical pathway for enriched ICs was lipopolysaccharide/Interleukin 1 mediated inhibition of Retinoid X Receptor alpha function and decreased Retinoid X Receptor expression was confirmed to occur 1-hour post experimental murine UTI in ICs but not in non-ICs.

The renal collecting duct contains intercalated cells (ICs) and principal cells (PCs). PCs express aquaporin 2 (AQP2), modulate water and electrolyte reabsorption while ICs express the B1 subunit of vacuolar H⁺-ATPase (V-ATPase-B1) and maintain acid-base homeostasis^{1,2}. We, and others, have demonstrated that ICs are involved in the renal bacterial defense³⁻⁷. The expansion of collecting duct function to include innate immunity indicates that IC and/or PC functions are more diverse than previously recognized.

¹Indiana University School of Medicine, Riley Children's Hospital, Indianapolis, Indiana, United States. ²The Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, Ohio, United States. ³The Research Institute at Nationwide Children's, Center for Clinical and Translational Research, Columbus, Ohio, and College of Medicine, Ohio State University, Columbus, Ohio, United States. ⁴Department of Pediatrics, College of Medicine, The Ohio State University, Columbus, Ohio, United States. ⁵Innate Immunity Translational Research Center, Children's Foundation Research Institute at Le Bonheur Children's Hospital, Memphis, Tennessee, United States. ⁶Division of Nephrology, Department of Pediatrics, University of Utah, Salt Lake City, Utah, United States. ⁷University of Rochester Medical Center, School of Medicine and Dentistry, Rochester, New York, United States. ⁸Research Institute at Nationwide Children's Hospital Flow Cytometry Core Laboratory, Columbus, Ohio, United States. ⁹Department of Physics, College of Arts and Sciences, The Ohio State University, Columbus, Ohio, United States. ¹⁰Genomics Shared Resource, The Ohio State University Comprehensive Cancer Center, Columbus, Ohio, United States. ¹¹Division of Hematology, Department of Internal Medicine, The Ohio State University Wexner Medical Center, Columbus, Ohio, United States. Correspondence and requests for materials should be addressed to V.S. (email: visaxena@iu.edu) or D.S.H. (email: dhains@iu.edu) or A.L.S. (email: Schwadea@iu.edu)

Received: 7 March 2018

Accepted: 23 November 2018

Published online: 24 January 2019

Transcriptomics is a critical component of systems-level understanding of cell biology⁸. However, analysis of collecting duct cell function by transcriptomics performed on whole kidneys or tubular segments is limited because the kidney is a conglomerate of many cell types. For instance, the kidney contains cells from the proximal tubule (PTCs), loop of Henle (LOH), distal convoluted tubule, connecting tubule, collecting duct, glomerulus along with vascular, interstitial and resident immune cells. Further, we had previously reported that ICs account for ~1% of cortical cells and ~2% of medullary cells while PCs accounted for ~2% of cortical cells ~20% of medullary cells in the murine kidney³. Therefore, IC and/or PC expression risks being diluted by numerous cell types. Past techniques to enrich collecting duct cells have included dissecting the relatively collecting-duct rich medulla from the collecting duct poor cortex or, pooling microdissected tubules from different nephron segments or use of cell cultures^{9–11}. While the collecting duct can be evaluated using these aforementioned methodologies, the distinct cell type (e.g. ICs and PCs) cannot be analyzed individually by dissecting out the medulla or tubular segments and cultured cells may not retain the phenotype of the targeted cell type.

Generation of transgenic mice expressing cre recombinase under the control of IC and PC specific promoters provides a modality for the evaluation of collecting duct innate immunity at the cellular level^{12,13}. We have reported on methodologies using two fluorescent reporter mice, V-ATPase B1-cre⁺tdTomato⁺ mice to label ICs and AQP2-cre⁺tdTomato⁺ mice to label PCs and then flow sorting to enrich viable tdTomato⁺ PCs and ICs for analysis¹⁴. The objective of this study is to identify distinct and overlapping transcriptome profiles associated with ICs and PCs.

Results

Quality control. To determine if there were any unexpected issues and to help ensure that the observed differences in expression were due to experimental conditions, a principal component analysis plot (PCA), volcano plot and log intensity ratio (M-value or MA) plot were generated for the groups: IC versus non-ICs (IC vs non-IC), PC versus non-PCs (PC vs non-PC) and ICs versus PCs (Supplemental Data S1). The PCA plots showed that the samples segregate by sample group, indicating that expression levels were not affected by something other than the intended treatment (like a batch effect). The volcano plots for global gene expression demonstrated a large number of statistically significant differentially expressed genes with more genes with lower relative expression ICs or PCs compared to non-ICs and non-PCs. Additionally, characteristic MA plot shapes were noted.

Relative enrichment of IC and PC cells. IC and PC cells were enriched from IC and PC reporter mice by enzymatic digestion of kidney followed by flow sorting of Tdtomato positive (presumed ICs and PCs) and negative (presumed non-ICs and non-PCs) cells then RNA-Seq was carried out to determine relative abundance for all expressed genes in IC and PC compared to non-ICs and non-PCs. Lineage markers known to be widely expressed by these cells were used to determine the degree IC and PCs were enriched. IC lineage markers included the following genes (corresponding proteins in parenthesis): *Atp6v1b1* (V-ATPase B1), *Slc4a1* (Anion exchanger 1), *Slc26a4* (pendrin), *P2ry14* (P2Y purinoceptor 14) and *Foxi1* (forkhead box i1)^{15–17}. PC lineage markers included *Aqp2*, *Elf2* (E74-like factor 2), *Scnn1a*, *Scnn1b* and *Scnn1g* (alpha, beta and gamma subunits respectively for the epithelial sodium channel, ENaC)^{18,19}. *Lrp2* (megalin) was used as a proximal tubule marker and *Umod* (Tamm-Horsfall protein) was used as a loop of Henle marker^{20,21}. The RNA-Seq groups lineage marker comparison results are presented in as IC/non-IC and PC/non-PC read counts in Fig. 1A and similar markers are presented as differentially expressed genes as heatmap in Fig. 1B.

IC versus non-ICs. *Genes enriched in ICs.* Compared to non-ICs, 3,645 genes had higher relative expression by ≥ 2 -fold with a p-value < 0.05 in ICs. The top 25 genes (excluding long non-coding genes) with highest fold change increase in ICs are presented in Table 1 and a heatmap of the top 25 differentially expressed genes is presented in Fig. 1C (left panel). The complete list of genes is presented at <https://hpcwebapps.cit.nih.gov/ESBL/Database/IC/index.html>. The top genes with increased relative expression in ICs included those previously reported to be relatively enriched in ICs compared to PCs including doublesex and mab-3 related transcription factor 2 (*Dmrt2*) and *Foxi1*, but also traditional principal markers such as *Aqp2*²².

Canonical pathways. Ingenuity analysis predicted key pathways associated with ICs. The top 10 canonical pathways are presented in Fig. 2A. The top predicted pathway in ICs was “interaction between lipopolysaccharide (LPS)/interleukin -1 (IL-1) mediated inhibition of Retinoid X receptor (RXR) function”. Key IC associated innate immune genes involved in this pathway are presented in Supplemental Data S2. Other involved pathways included wingless-type MMTV integration site family member (Wnt)/Ca⁺⁺, axonal guidance signaling and sphingosine-1 phosphate signaling.

Functional evaluation of LPS/IL-1 mediated inhibition of RXR function in IC versus non-ICs. Retinoid X Receptors have been implicated to play key role in regulating variety of cell function such as cell differentiation, immune response, and metabolism²³. One of its member, RXR which is a key nuclear receptor in metabolic processes has been shown to be down regulated during viral infections and protect host²⁴. Ingenuity analysis predicted top pathway (inhibition of RXR function) in ICs was functionally validated in murine model of UTI. When ICs were enriched from IC reporter mice 1 hr post transurethral inoculation of UPEC vs. saline control, ICs but not non-ICs significantly down regulated *Rxr α* mRNA expression with UPEC vs. saline (Fig. 2B). Our findings indicate that ICs may use this mechanism to protect the host from transurethral infections.

PC versus non-PCs. *Genes enriched in PCs.* Compared to non-PCs, 2088 genes had higher relative expression in PCs by ≥ 2 -fold and an adjusted p value < 0.05 . The 25 genes with highest fold change increase in PCs

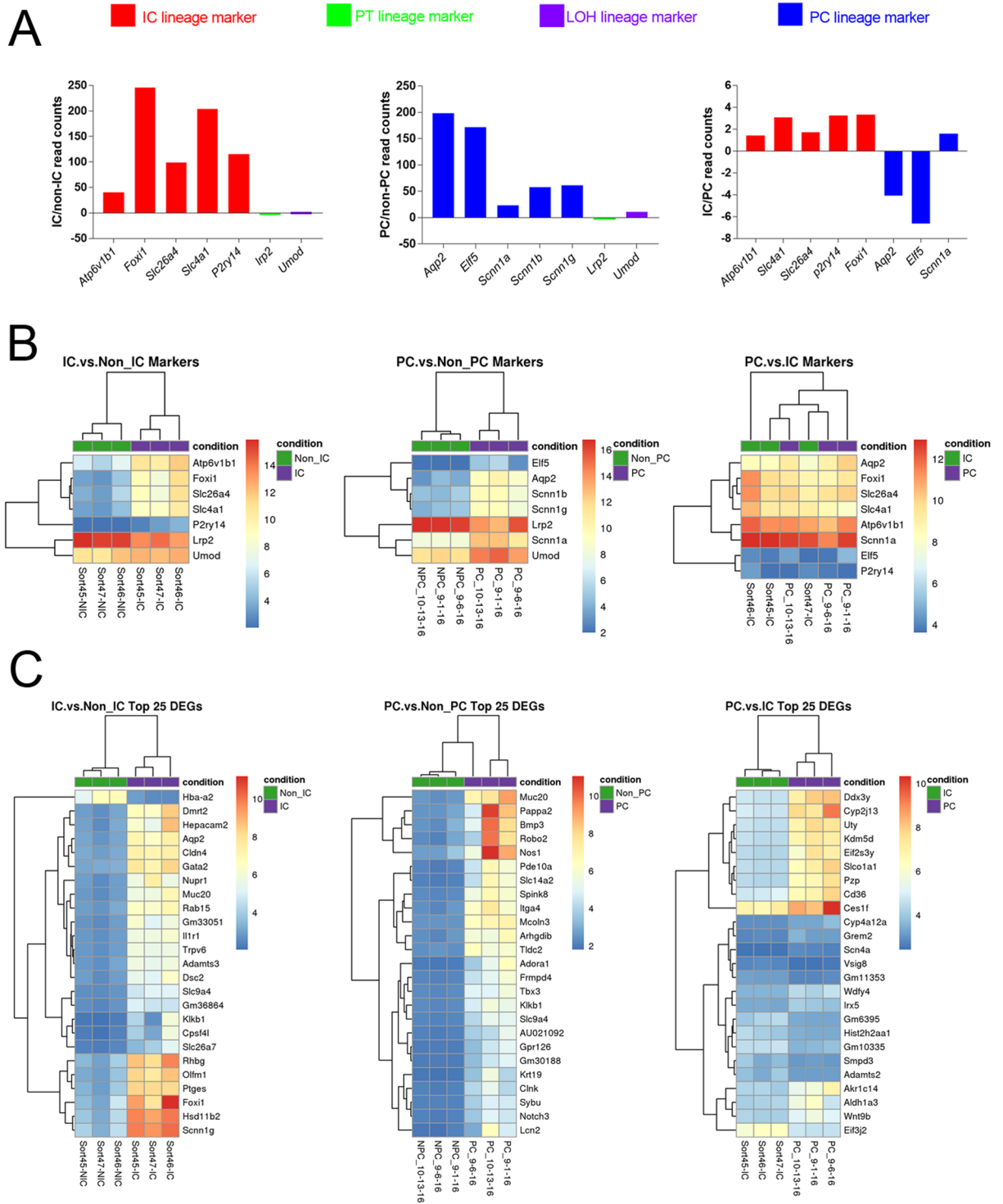


Figure 1. (A) Relative lineage marker fold change. IC lineage markers (left panel) mRNA expression was increased 41 to 246-fold in ICs compared to non-ICs, while PC lineage marker mRNA expression (middle panel) was increased 24 to 198-fold in PCs compared to non-PCs (middle panel). Of note, *Scnn1a* did not appear to be as enriched in PCs compared to other PC markers. In the direct IC to PC comparison (right panel) IC markers were around ≥ 1.5 -fold higher in ICs than PCs. PC markers were around ≥ 4 -fold higher in PCs compared to ICs with the exception of *Scnn1a* which was higher in ICs. (B) Heatmaps of the lineage markers are presented IC vs NICs (left panel), PC vs non-PCs (middle panel) and ICs vs PCs (right panel) showing relative expression. (C) Heatmaps of the 25 most differentially expressed genes in ICs vs non-ICs (left panel), PCs vs non-PCs (middle panel) and ICs vs PCs (right panel) are presented. The samples are grouped by dendrogram at the top of figure and genes are grouped by the dendrogram on the left side of the figure. Similarly regulated genes are grouped together. The color coded legend bar indicate normalized read counts that have been log₂ transformed. Red indicates higher expression while blue indicate lower expression.

Gene Symbol	Gene Description	Fold Change	p-Adjusted
<i>Dmrt2</i>	doublesex and mab-3 related transcription factor 2	1402.61	<0.001
<i>Muc20</i>	mucin 20	638.64	<0.001
<i>Aqp2</i>	aquaporin 2	497.92	<0.001
<i>Il1r1</i>	interleukin 1 receptor, type I	481.87	<0.001
<i>Trpv6</i>	transient receptor potential cation channel, subfamily V, member 6	446.16	<0.001
<i>Olfm1</i>	olfactomedin 1	423.47	<0.001
<i>Rab15</i>	RAB15, member RAS oncogene family	397.74	<0.001
<i>Nupr1</i>	nuclear protein transcription regulator 1	388.64	<0.001
<i>Hepacam2</i>	HEPACAM family member 2	381.31	<0.001
<i>Rhbg</i>	Rhesus blood group-associated B glycoprotein	367.17	<0.001
<i>Cldn4</i>	claudin 4	345.36	<0.001
<i>Klkb1</i>	kallikrein B, plasma 1	321.30	0.022
<i>Gata2</i>	GATA binding protein 2	317.62	<0.001
<i>Slc9a4</i>	solute carrier family 9 (sodium/hydrogen exchanger), member 4	311.26	<0.001
<i>Cpsf4l</i>	cleavage and polyadenylation specific factor 4-like	289.04	0.003
<i>Hsd11b2</i>	hydroxysteroid 11-beta dehydrogenase 2	275.03	<0.001
<i>Foxi1</i>	forkhead box i1	256.91	0.003
<i>Scnn1g</i>	sodium channel, non- voltage-gated 1 gamma	246.28	<0.001
<i>Ptges</i>	Prostaglandin E synthase	234.42	<0.001
<i>Dsc2</i>	desmocollin 2	228.71	<0.001
<i>Slc26a7</i>	solute carrier family 26, member 7	221.91	<0.001
<i>Adams3</i>	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 3	216.11	<0.001
<i>Egf12</i>	fibroblast growth factor 12	213.06	<0.001
<i>Rasd2</i>	RASD family, member 2	205.90	<0.001
<i>Kcns3</i>	potassium voltage-gated channel, delayed-rectifier, subfamily S, member 3	196.7	<0.001

Table 1. Top 25 gene with highest relative expression in ICs versus non-ICs.

are presented in Table 2. A heatmap of the 25 most differentially expressed genes is presented in Fig. 1B (middle panel). The complete list of genes is presented at <https://hpcwebapps.cit.nih.gov/ESBL/Database/PC/index.html>. Genes with higher expression in PCs compared to non-PCs included mucin 20 (*Muc20*), urea transporters including solute carrier family 14 member 2 (*Slc14a2*) and renal development genes such as bone morphogenetic protein 3 (*Bmp3*), and neurogenic locus notch homolog protein 3 (*Notch 3*).

Canonical pathways. Ingenuity software identified key pathways associated with PCs. The top 10 canonical pathways enriched in PCs are presented in Fig. 2C. The top pathway included axonal guidance signaling among others as thrombin signaling and CXCR4 signaling.

IC versus PC. In direct IC to PC comparison 123 genes were differentially expressed (adjusted $p < 0.05$ and fold-change ≥ 2). The complete list of genes is presented at <https://hpcwebapps.cit.nih.gov/ESBL/Database/PC/PC-IC%20Results.html>. ICs and PCs appeared more similar to each other than non-ICs or non-PCs respectively (Fig. 1B,C).

Differentially expressed genes in IC. Twenty genes with higher fold change in ICs vs. PCs are listed in Table 3 excluding predicted genes. A heatmap of top 25 differentially expressed genes in ICs vs. PCs is presented in Fig. 1C (right panel).

Differentially expressed genes in PC. Genes with higher relative expression in PCs compared to ICs ranked by fold change included water channels such as *Aqp4* (10.6-fold higher in PCs, $p = 0.0006$) along with proteins involved in tissue development, regeneration and cell polarity such as *Wnt9b* (16.5 fold, $p < 0.0001$). The top 20 genes with higher expression in PCs compared to ICs is presented in Table 4.

Innate immune genes. Innate immune genes were identified by keyword search such as “Toll like receptor” (*Tlr*), “nod like receptor” (*Nlr*), “interleukin receptor” (*Ilr*), “cluster of differentiation receptor” (*Cd*), “chemokine receptor”, “defensin”, “ribonuclease”, and “calcium binding proteins” from unfiltered comprehensive gene list (Supplemental Data S3). Most innate immune genes did not have significantly different relative expression between ICs and PCs when using adjusted p-values. Secretogranin 5 (*Scg5*), an antimicrobial peptide, had higher relative expression in ICs by RT-PCR validation (Supplemental Data S4).

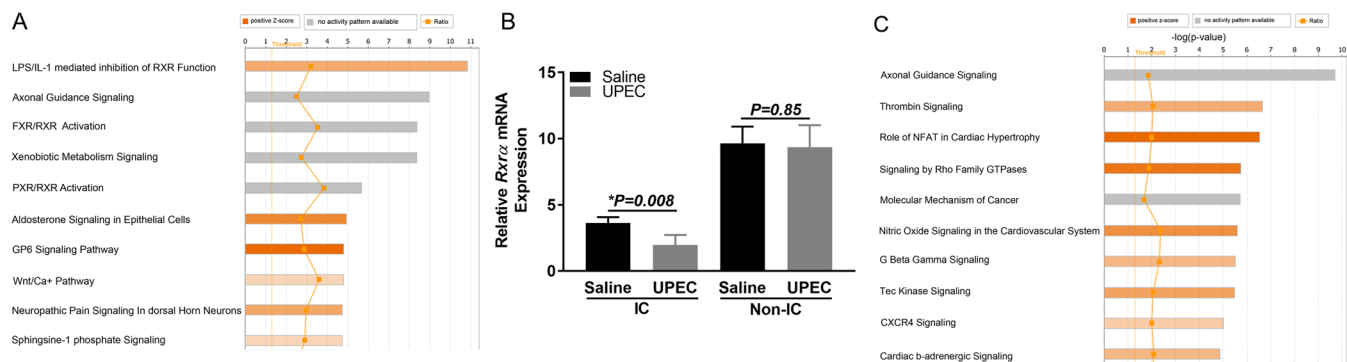


Figure 2. (A) The top 10 canonical pathways associated with IC mRNA expression are presented. (B) IC specific inhibition of *Rxrα* mRNA expression. IC reporter mice were infected with UPEC or saline and Tdt⁺ (IC) and Tdt⁻ (Non-IC) cells were enriched. *Rxrα* mRNA was measured normalized to GAPDH. IC cells showed significant inhibition of mRNA expression upon UPEC exposure (3.64 ± 0.44 in saline vs. 1.97 ± 0.85 in UPEC, $P = 0.008$) compared to NIC (9.63 ± 1.26 in saline vs. 9.34 ± 1.67 in UPEC, $P = 0.85$) cells. $N = 4$ experiments, each using a separate sort consisting of kidneys from 1–2 mice. (C) The top 10 canonical pathways associated with PC mRNA expression are presented.

Gene Symbol	Gene Description	Fold Change	p-Adjusted
<i>Muc20</i>	mucin 20	2739.04	<0.001
<i>Pappa2</i>	pappalysin 2	1294.57	<0.001
<i>Spink8</i>	serine peptidase inhibitor, Kazal type 8	723.69	<0.001
<i>Pde10a</i>	phosphodiesterase 10A	694.92	<0.001
<i>Robo2</i>	roundabout homolog 2 (Drosophila)	675.41	0.021
<i>Slc14a2</i>	solute carrier family 14 (urea transporter), member 2	667.44	<0.001
<i>Nos1</i>	nitric oxide synthase 1, neuronal	655.19	<0.001
<i>Bmp3</i>	bone morphogenetic protein 3	546.92	0.011
<i>Itga4</i>	integrin alpha 4	545.16	<0.001
<i>Mcoln3</i>	mucolipin 3	533.56	<0.001
<i>Adora1</i>	adenosine A1 receptor	525.07	<0.001
<i>Klkb1</i>	kallikrein B, plasma 1	409.54	<0.001
<i>Frmpl4</i>	FERM and PDZ domain containing 4	403.22	<0.001
<i>Tbx3</i>	T-box 3	396.27	<0.001
<i>Arhgdib</i>	Rho, GDP dissociation inhibitor (GDI) beta	362.66	<0.001
<i>Tlhc2</i>	TBC/LysM associated domain containing 2	326.32	<0.001
<i>Slc9a4</i>	solute carrier family 9 (sodium/hydrogen exchanger), member 4	326.09	<0.001
<i>Krt19</i>	keratin 19	313.32	<0.001
<i>AU021092</i>	expressed sequence AU021092	271.27	<0.001
<i>Clmk</i>	cytokine-dependent hematopoietic cell linker	269.25	<0.001
<i>Gpr126</i>	G protein-coupled receptor 126	259.86	<0.001
<i>Lcn2</i>	lipocalin 2	249.00	<0.001
<i>Sybu</i>	syntabulin (syntaxin-interacting)	246.59	<0.001
<i>Notch3</i>	notch 3	241.99	<0.001
<i>Prkar1b</i>	protein kinase, cAMP dependent regulatory, type I beta	219.15	

Table 2. Top 25 genes with highest relative expression in PCs versus non-PCs.

In situ and immunofluorescent evaluation of collecting duct marker expression and evaluation of cre lines. Since *Aqp2* expression (PC cell marker) was present in the enriched IC cells compared to non-IC cells by RNA-Seq (Table 1), we sought to confirm this intriguing finding. When *Aqp2* mRNA expression is overlaid with V-ATPase protein expression, a subset of V-ATPase positive cells (ICs) also appear to express *Aqp2* mRNA (Fig. 3). We counted 155 cells with positive V-ATPase expression from 2 kidneys. 9% ($n = 14$, 60X magnification) of V-ATPase positive cells also had *Aqp2* mRNA expression. On triple labeling for V-ATPASE E1, AQP2 and tdTomato (expression driven by AQP2 cre promoter), most but not all of the AQP2 positive cells expressed tdTomato (Fig. 4A) indicating that penetrance of cre was good, but not complete in this cre line as well.

Gene Symbol	Gene Description	Fold Change	p-Adjusted
<i>Smpd3</i>	sphingomyelin phosphodiesterase 3, neutral	25.09	0.016
<i>Hist2h2aa1</i>	histone cluster 2, H2aa1	17.56	<0.001
<i>Eif3j2</i>	eukaryotic translation initiation factor 3, subunit J2	15.20	<0.001
<i>Adams2</i>	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 2	13.65	0.031
<i>Dapl1</i>	death associated protein-like 1	8.70	0.043
<i>Rn7s2</i>	7S RNA 2	8.32	<0.001
<i>Rpl23a-ps3</i>	ribosomal protein L23A, pseudogene 3	7.71	0.040
<i>Rpl19-ps11</i>	ribosomal protein L19, pseudogene 11	6.34	<0.001
<i>Tpm3-rs7</i>	tropomyosin 3, related sequence 7	5.33	<0.001
<i>Nsa2-ps1</i>	NSA2 ribosome biogenesis homolog, pseudogene 1	5.18	<0.001
<i>Rasd1</i>	RAS, dexamethasone-induced 1	4.27	<0.001
<i>Adm</i>	adrenomedullin	4.01	0.016
<i>Erd1</i>	erythroid differentiation regulator 1	3.91	0.035
<i>Oaz1-ps</i>	ornithine decarboxylase antizyme 1, pseudogene	3.60	<0.001
<i>Ryr2</i>	ryanodine receptor 2, cardiac	3.20	0.006
<i>Tbx2</i>	T-box 2	2.48	<0.001
<i>A4galt</i>	alpha 1,4-galactosyltransferase	2.40	0.003
<i>Tbc1d1</i>	TBC1 domain family, member 1	2.16	0.001
<i>Gsta4</i>	glutathione S-transferase, alpha 4	2.06	0.005
<i>Anapc13</i>	anaphase promoting complex subunit 13	1.91	0.001

Table 3. Top 20 genes overexpressed in ICs compared to PCs.

We also evaluated whether tdTomato (expression driven by *VATPaseB1* promoter) co-labeled with V-ATPase B1 (Fig. 5A–D). Most cells co-labeled with the 2 antibodies but there were some cells that only labelled with tdTomato indicating some “leakiness” of the ATPV1B1-cre (Fig. 4B).

Single cell IC sequencing pilot study and comparison with recent study by Chen and colleagues. We sequenced 4 intercalated cells on the single cell level and presented the data at <https://hpcwebapps.cit.nih.gov/ESBL/Database/IC/index.html/>.

The 20 genes with the highest total expression in our single cell IC analysis are presented in Supplemental Data S5, please note this is total expression as opposed to the relative expression presented elsewhere since there were no comparison groups for this analysis. We compared our pooled and single cell data to the 2017 report by Chen and colleagues (Table 5)²². For our data, we used the normalized read counts from the IC versus PC comparison for the pooled cell presented in Table 5 because the normalized read counts will change somewhat depending on what other samples are normalized with, and each comparison used datum that was normalized with only the other samples included in that particular comparison.

Similar enrichment for PC genes, when pooled PCs were compared to pooled ICs, generally occurred in both datasets²². For example, there was 4.1 and 9.7-fold enrichment for *Aqp2* along with 6.6 and 8.3-fold enrichment for *Elf5* in our as well as Chen and colleague’s datasets respectively²². IC marker expression was 1.4 to 3.3-fold higher in our pooled ICs vs. pooled PCs analysis but 4.4 to 13.6-fold higher in the Chen and Colleagues analysis²². We did identify *Atp6v1e1* subunit expression on the single cell expression level in 4/4 cells and used this gene, along with the IC specific genes, *Fxyd2* and *Atp6v1g3*, to confirm their IC cell status^{22,25}. Additionally we did not find PC markers including *Elf5*, *Aqp2*, *Avpr2* or *Scnn1g* in our limited subset of ICs on single cell analysis²². An image of the integrated fluidic chip captured a tdTomato expressing single IC (Fig. 5A) and a summary of lineage marker expression is presented in Fig. 5B.

Innate immune gene expression (Table 5) was generally consistent across analysis in pooled ICs and PCs. Both our and Chen and colleague’s analysis revealed that expression for the Antimicrobial peptide (AMP) genes adrenomedullin (*Adm*), B-defensin 1 (*Defb1*) and secretogranin 5 (*Scg5*) was present in both pooled ICs and pooled PCs but the mean expression was higher in ICs²². Additionally, we were able to identify *Defb1* and *Scg5* in ICs on the single cell level. Lipocalin 2/NGAL (*Lcn2*) had higher expression in PCs than ICs in both pooled cell analysis²². The inflammatory signaling genes apolipoprotein E (*ApoE*), secreted phosphoprotein 1/osteopontin (*Spp1*), matrix metalloproteinase (*Mmp11*) and kallikrein-1 (*Klk1*) were identified in ICs at the single cell level. Toll like receptor 3 (*Tlr3*) had the highest expression in the TLR family in pooled ICs in both ours and Chen and colleague’s pooled analysis, but we did not identify this family of genes on the single cell level²².

Discussion

We have completed pooled RNA-Seq of IC and PC along with pilot single cell IC RNA-Seq to lend insight into the biological processes associated with murine renal collecting duct cells. This study complements and expands on our past studies evaluating the transcriptome profile of the collecting duct.

Our RNA-Seq IC component contains overlap compared to the prior proteomic analysis presented by Da Silva and colleagues²⁶. Of the 25 genes with the highest differential expression in ICs, several were also reported as

Gene Symbol	Gene Description	Fold Change	p-Adjusted
<i>Ddx3y</i>	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	1916.66	<0.001
<i>Cyp2j13</i>	cytochrome P450, family 2, subfamily j, polypeptide 13	1537.46	<0.001
<i>Uty</i>	ubiquitously transcribed tetratricopeptide repeat gene, Y chromosome	1448.34	<0.001
<i>Slc1a1</i>	solute carrier organic anion transporter family, member 1a1	1409.07	<0.001
<i>Kdm5d</i>	lysine (K)-specific demethylase 5D	1158.52	<0.001
<i>Eif2s3y</i>	eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked	1141.56	<0.001
<i>Pzp</i>	pregnancy zone protein	1141.01	<0.001
<i>Akr1c14</i>	aldo-keto reductase family 1, member C14	75.36	<0.001
<i>Cyp4a12a</i>	cytochrome P450, family 4, subfamily a, polypeptide 12a	51.90	0.013
<i>Grem2</i>	gremlin 2 homolog, cysteine knot superfamily (<i>Xenopus laevis</i>)	47.25	0.040
<i>Wdfy4</i>	WD repeat and FYVE domain containing 4	36.63	<0.001
<i>Cd36</i>	CD36 antigen	36.32	<0.001
<i>Aldh1a3</i>	aldehyde dehydrogenase family 1, subfamily A3	19.65	<0.001
<i>Wnt9b</i>	wingless-type MMTV integration site family, member 9B	16.46	<0.001
<i>Ces1f</i>	carboxylesterase 1F	15.66	<0.001
<i>Enox1</i>	ecto-NOX disulfide-thiol exchanger 1	11.51	0.040
<i>Aqp4</i>	aquaporin 4	10.60	<0.001
<i>Itga4</i>	integrin alpha 4	10.49	0.001
<i>Timd2</i>	T cell immunoglobulin and mucin domain containing 2	9.12	0.033
<i>Ugt2b38</i>	UDP glucuronosyltransferase 2 family, polypeptide B38	8.62	0.040

Table 4. Top 20 genes overexpressed in PCs compared to ICs.

enriched in the IC proteomic analysis including Kallikrein 1 (*Klk1*) and Hydroxysteroid 11-beta dehydrogenase 2 (*Hsd11b2*)²⁶.

There are also similarities and distinctions when our data is compared to past transcriptome collecting duct expression. Using microdissected rat tubules, Lee and colleagues identified high *Aqp2*, arginine vasopressin receptor 1A (*Avpr1a*), and *Defb1* expression in the cortical collecting duct and *Aqp2*, *Avpr2*, serum/glucocorticoid regulated kinase 1 (*Sgk1*) and *Defb1* expression in the medullary collecting duct⁹. Similarly, in our analysis *Defb1*, mRNA expression was enriched in both ICs and PCs, akin to our prior targeted RT-PCR analysis of AMP expression in enriched ICs and PCs¹⁴. The comparison between our data and data generated by Chen and colleagues provides interesting insight into collecting duct function and sequencing capabilities. Both research groups obtained relative enrichment for multiple collecting duct cell type markers. We were able to obtain several IC marker expression in the absence of PC cell marker expression on our single tdTomato + cells (presumed ICs) but not *Atp6v1b1* (Fig. 5). Therefore, our single cell transcriptomics while specific, may not be sensitive. Reasons for this diminished sensitivity may include, low cell number, lower read depth of the sequencing and/or lack of consistent mRNA expression between cells.

Table 1 (Top 25 genes enriched in ICs versus non-ICs) contains some PC associated genes such as *Aqp2* and *Scnn1g*, along with IC genes DMRT like family 2 (*Dmrt2*), Nuclear protein 1, transcriptional regulator 1 (*Nupr1*) and Arginine vasopressin receptor 1a (*Avpr1a*); while Table 2 (Top 25 genes enriched in PCs versus non-PCs) highlights historically IC associated genes, cytokine dependent hematopoietic cell linker (*Clnk*), as well as PC associated gene, mucolin 3 (*Mcoln3*)²². Although enriched (Fig. 1), our pooled collecting duct cell populations may not be absolutely pure and other kidney cell types may contribute to mRNA expression of some genes. This overlap between ICs and PCs might even be more pronounced when these cell types are compared to non-ICs and non-PCs. For example, in the IC versus PC comparison, a limited number of PCs in the ICs sort might result in relatively high PC associated gene expression in enriched ICs compared to the relatively collecting duct poor remainder of the kidney. Therefore, pooled IC and PC cell findings may be generalized to collecting duct cells rather than specific to ICs and PCs. Additionally mRNA expression might not always correlate with protein expression.

Another explanation for overlap between IC and PC genes in our sorted cells is overlap in the transcriptomes of ICs and PCs. For example, the historically PC associated gene *Aqp2* was identified in some ICs using sequencing on the single cell level by Chen and colleagues and by us using *in situ* hybridization (Fig. 3A)²⁵. We expectantly had 4.1-fold higher *Aqp2* expression in PCs compared to ICs (adjusted p-value 0.01) but also had this gene come up as enriched in ICs compared to non-ICs (Table 1). Similar results were noted in our targeted RT-PCR collecting duct cell analysis¹⁴. There is evidence from past studies which suggest that AQP2-expressing cells give rise to ICs^{27,28}. We performed *in situ* hybridization for *Aqp2* mRNA expression overlaid with immunostaining for V-ATPase protein expression. Approximately 9% of V-ATPase⁺ cells (presumed ICs) demonstrated *Aqp2* mRNA expression (Fig. 3C). We propose that an intermediary cell type with features of both PCs and ICs is present analogous to the Chen and colleagues which identified *Aqp2* and *Slc4a1* (AE1)/*Slc26a4* (pendrin) dual expressing hybrid cells and Park and colleagues which identified co-expression of AQP2 and V-ATPase B1 at the protein level^{22,29}. *Aqp2* expression in our TdTomato expressing cells from the V-ATPase cre-tdTomato mice is higher than would expected from these aforementioned hybrid cells and is likely due to some “leakiness” of the cre (Fig. 4B). Also some *Umod* (a loop of Henle marker) enrichment occurs in the PCs from the tdTomato driven by AQP2 cre,

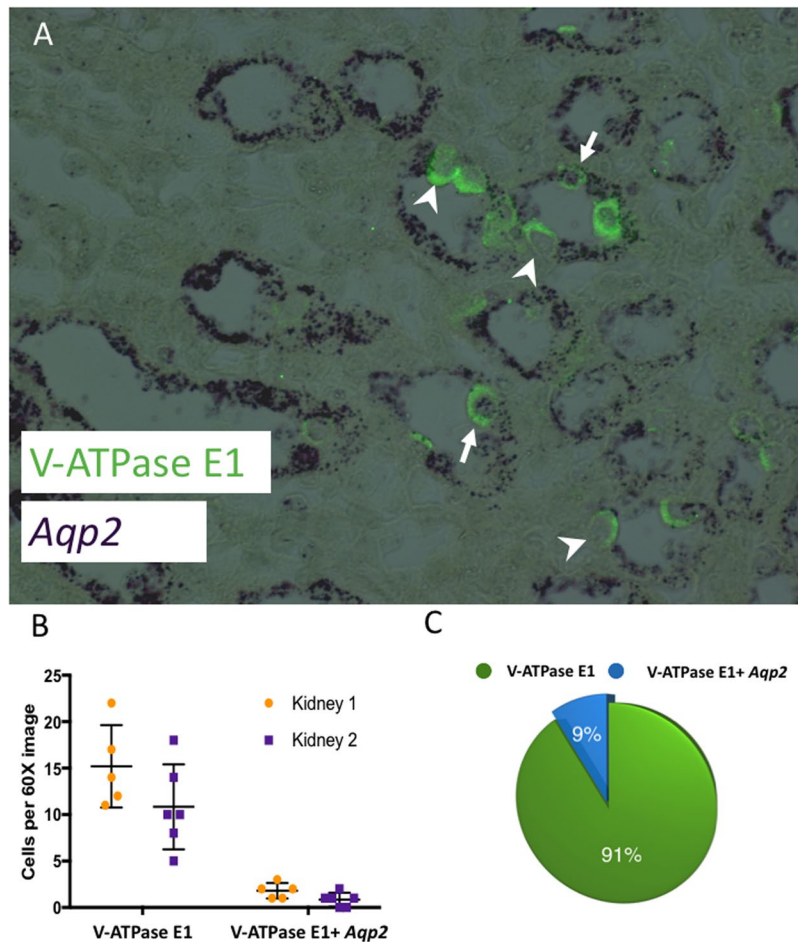


Figure 3. (A) Co-expression of *Aqp2* mRNA and V-ATPase E1 protein in wild type C57Bl/6 mice. Imaging of mouse kidney sections for *Aqp2* mRNA *in-situ* hybridization (purple) and overlaid with immunofluorescent protein labeling of V-ATPase E1 (green). The majority of ICs did not co-label for *Aqp2* mRNA (arrowheads) but a subset did (arrows), Magnification 60X. (B) Kidneys from 2 mice were evaluated and the cellular expression of V-ATPase E1 alone and both V-ATPase E1 and *Aqp2* expression was found to be similar between kidneys, p values 0.08 and 0.86 respectively. The horizontal bar represents the mean and the error bars represent the standard deviation. (C) Approximately 9% of cells with V-ATPase E1 protein expression also had expression for *Aqp2* mRNA.

indicating that this cre line also has some overlap with loop cells. Of note Chen *et al.* also reported collecting duct expression of *Umod*, but in ICs rather than PCs, so it is possible that some expression may be intrinsic rather than due to contamination²². It is also possible that there may be more overlap between the transcriptomes of ICs and PCs than previously realized and that the functional and structural differences between these cell types are due in part to post translational modifications. We also speculate that the PC enrichment was more efficient than IC enrichment because enrichment of the more prevalent cell type (PCs) conceptually results in less potential for contamination.

In 2011, we proposed that ICs “have a novel role in innate immunity” based on their expression of AMPs⁴. In this study, Ingenuity pathway analysis predicted LPS/IL-1 mediated inhibition of RXR function was the most activated canonical pathway. Decreased retinoid X receptor expression was confirmed to occur 1- hour post experimental murine model of UTI in ICs but not non-ICs. RXR is downregulated indirectly by interferon regulatory factor 3 (*Irf3*) and subsequently suppresses interferon genes²⁴. We identified *Irf3* in ICs on the single cell level, and Svanborg and colleagues have shown that *Irf3* SNPs lead to UTI susceptibility³⁰. The suppression of *Rxrα* at a 1-hr time point in a non-refluxing C57Bl/6 mouse may indicate an intercalated cell UTI response before a large bacterial burden is able to ascend into the kidneys³¹. Functional consequences of altered RXR expression on the bacterial defense of the kidney and urinary tract will be a key area of future research. Chassin *et al.*, reported that renal collecting duct cells bind to uropathogenic *E. coli* (UPEC) and activate TLR4 dependent and independent pathway⁷. Similarly, we earlier reported that upon UPEC (strain CFT073) challenge *in vitro* and *in vivo*, both ICs and PCs respond quickly by enhancing mRNA expression of anti-microbial peptides¹⁴. It is of note that in this study, IC associated pathways and LPS interaction are only predicted at the transcriptome level, although we validated the top pathway functionally at mRNA level. Nonetheless our analysis will provide a roadmap for future studies that identify the mechanisms by which ICs defend the kidney against bacterial invasion.

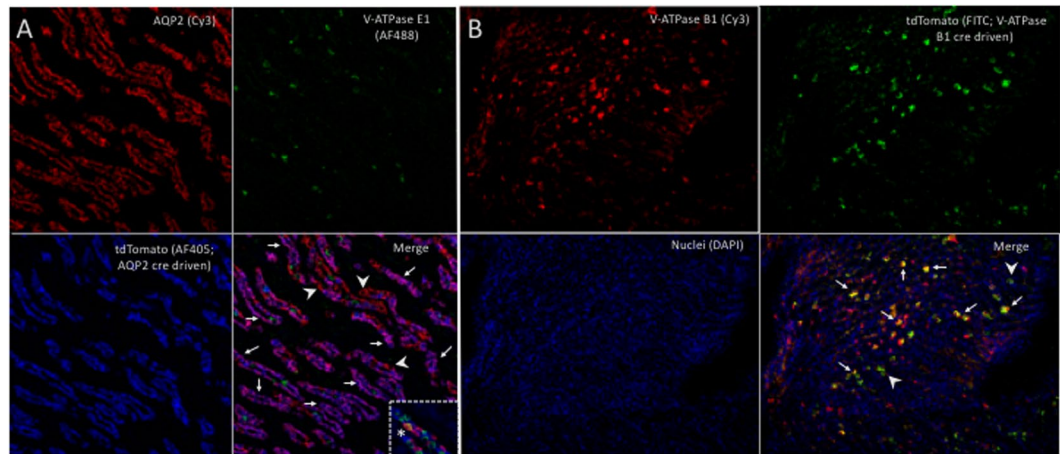


Figure 4. (A) Triple immunofluorescence labeling of a kidney section from AQP2-cre⁺tdTomato⁺ (PC reporter) mice showing: AQP2 (red cells, upper left); V-ATPase E1 (green cells, ICs, upper right); tdTomato (blue cells, lower left) staining in inner medulla. On the merged image (lower right) most AQP2 cells were also positive for tdTomato (purple cells, white arrows) while some were negative for tdTomato (red cells, white arrowheads). The dashed box contains a tubule imaged from another kidney section that contained an isolated cell that was positive for both AQP2 and V-ATPase E1 (Yellow cell, asterisk). Magnification 40X, confocal microscopy. (B) Triple immunofluorescence labeling of a kidney section from V-ATPase B1-cre⁺tdTomato⁺ (IC reporter) mice showing: V-ATPase B1 (red cells, upper left). tdTomato (green cells, upper right) and nuclei (blue, lower left) staining in inner medulla. On the merged image (lower right) most V-ATPase B1 positive cells were also tdTomato positive (yellow cells, white arrows) while a few cells were tdTomato positive but were V-ATPase B1 negative (green cells, white arrowheads). Magnification 20X, immunofluorescent microscopy.

We have previously reported that murine and human ICs express innate immune proteins, particularly AMPs in the ribonuclease A and defensin families⁶. This unbiased screen expanded the AMPs known to be expressed by the collecting duct expression. Collecting duct AMPs identified in this study included *Adm Scg5*. *Adm* is a multifunctional protein with direct antimicrobial activity against *Staphylococcus* and *E. coli*³². *Scg5* is potentially important in the kidney's anti-bacterial response because secretogranins have been reported to have antimicrobial activity³³.

The genes with the highest expression in our pilot study of sequenced single ICs (Supplemental Data S5) also lends insight into the innate immune function of these cells. The gene with the highest expression was Metastasis Associated Lung Adenocarcinoma Transcript 1 (*Malat1*). *Malat1* is a conserved long, noncoding RNA that has been demonstrated to be upregulated in LPS activated macrophages and interact with NF- κ B subunits to regulate the expression of pro-inflammatory cytokines³⁴. Some of the other genes highly expressed in the single ICs also have key innate immune functions. Kallikrein 1 (*Klk1*), the 4th most highly expressed gene, is a member of a gene family that has been demonstrated to regulate the antibacterial effects of AMPs by controlling activation of precursor proteins and influencing processing of AMPs to smaller peptides with alternate biological activity³⁵. Fms related tyrosine kinase 1 (*Ftl1*), the 6th most highly expressed gene, is also expressed by developing dendritic cells and is involved in Vascular endothelial growth factor mediated dendritic cell inhibition^{36,37}. Secreted phosphoprotein 1/osteopontin (*Spp1*), the 8th most highly expressed gene in ICs is also expressed by macrophages, neutrophils and dendritic cells³⁸. Prior research indicates that *Spp1* fosters cell adhesion, may facilitate phagocytosis and is involved in migration of neutrophils and dendritic cells^{39–41}. High expression of *Malat1*, *Klk1*, *Ftl1* and *Spp1* by ICs indicates that on a transcriptomic level, they contain key components of regulating the antibacterial response.

Our study did have limitations. For example, uromodulin was increased in PCs compared to ICs. This gene should be limited to the Loop of Henle (LOH), therefore it appears that the PC sort, but not IC sort may contain some LOH cells. We speculate that this is from some limitations in our mouse model, specifically some differences between cre induced tdTomato expression and AQP2 expression. Although tdTomato and AQP2 expression co-localized in most cells there were some cells with only AQP2 or tdTomato expression (Fig. 4B). We cannot exclude that cell dissociation has some effect on mRNA expression and/or selects a subset of cells that are able to remain viable through the process. These findings will need to be compared and contrasted to future studies that use other technologies such as laser capture microdissection. Our single cell IC sequencing pilot study evaluated 4 cells, a larger number of cells will be needed to compare expression between subtypes of ICs. Additionally, it is possible that murine and human gene expression may be different. We have seen this with Ribonuclease 7 (RNase7) which is a potent AMP, is expressed in human ICs but there is no mouse RNase7 orthologue is known⁴². Lastly, this study was designed to identify genes with differential expression between collecting duct cells and other renal cells. There may be important genes that are highly expressed in the collecting duct that were not highlighted in this study because they are also highly expressed in other nephron segments. Conversely genes with very low expression in one cell population and even lower expression in another cell population in another might be identified as enriched. For example, Hba-2a meets criteria for differentially expressed in ICs versus non ICs (Fig. 1C, left panel) because of expression in some non-ICs and virtually absent expression in ICs.

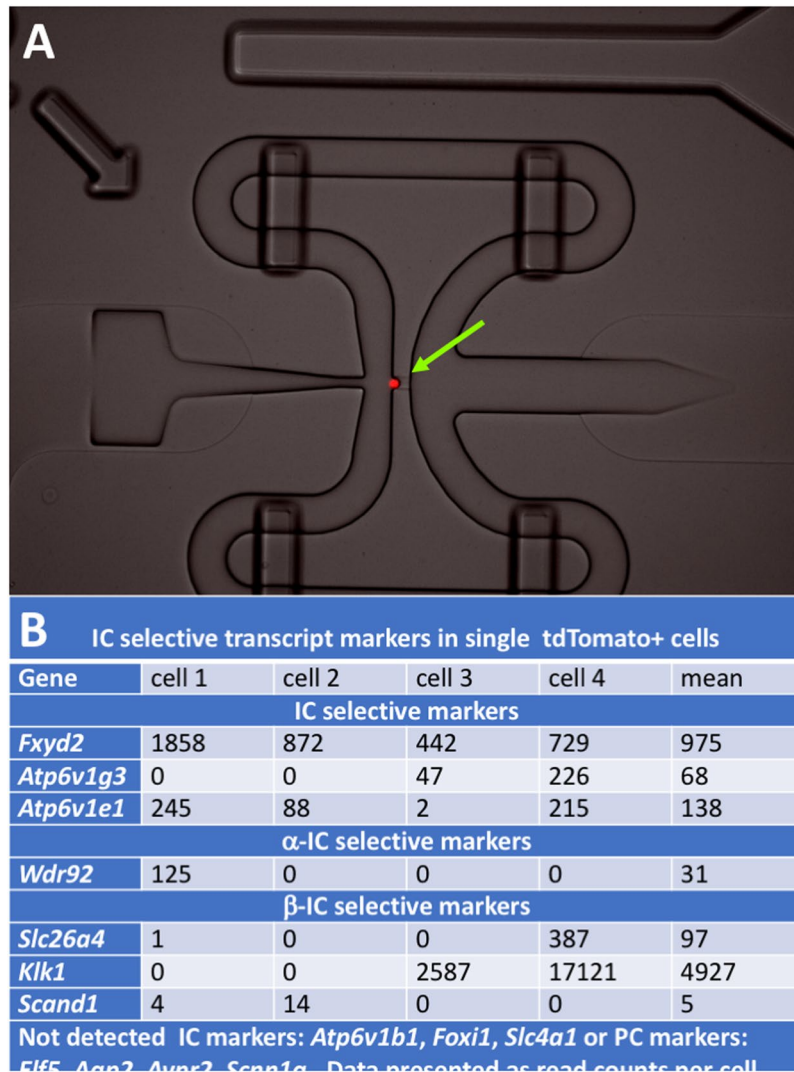


Figure 5. (A) Imaging of tdTomato⁺ single intercalated cell (IC) captured on medium size (10–17 μ m) IFC plate. Bulk ICs were sorted from IC reporter (V-ATPase-cre x tdTomato) mice with Flow sorting. Cells were pelleted, counted and suspended in Fluidigm cell suspension reagent at 3:2 ratio at 200 cells/ μ l concentration. Pooled cells were separated to single IC cells on Fluidigm C1 system and imaged under fluorescence microscope at 20X magnification. (B) Expression of IC markers. Base on the expression patterns of α -IC and β -IC markers, we estimate that cell 1 was a α -IC while cells 2–3 were β -ICs.

Our study shows potential limitations of single cell RNA-sequencing as several key cell associated genes were not detected (Fig. 5) this could be due to multiple factors such as sequencing depth, transcript per cell or stability of the single cell transcriptome.

Our study is also different from Chen and colleagues study because our study does not require *in vitro* labeling of cells. Chen and colleagues used a c-kit antibody which may affect the function of c-kit and subsequently cellular proliferation, downstream signal transduction and homologous dimerization^{22,43}. Secondly we focused on innate immune signaling where ingenuity analysis predicted LPS/IL-1 mediated inhibition of RXR function among the top canonical pathway of IC function. Lastly we experimentally validated the hybrid phenotype described by Chen *et al.* of V-ATPase⁺Aqp2⁺ cells by combined *in-situ* hybridization and immunofluorescence.

Overall, our study confirms that UPEC leads to decreased *Rxr α* expression in ICs but not ICs consistent with LPS/IL-1 mediated inhibition of RXR function pathway which was the canonical pathway predicted to be most associated with ICs.

Methods

Mouse. AQP2-cre (B6.Cg-Tg(Aqp2-cre)1Dek/J, (stock no. 006881, Jackson lab, Maine) and VATPaseB1-cre transgenic mice (kindly provided by Dr. Raoul D. Nelson, University of Utah, UT) (in which 7 kb B1 promoter drives the cre expression in ICs) were crossed to tdTomato-*loxP* homozygous mice (B6.Cg-Gt(ROSA)26Sor^{tm9(CAG-tdtomato)Hze/J} (stock no.007909, Jackson lab, Maine) and genotyped as previously described¹⁴. All animal research was approved by the Research Institute at Nationwide Children's Hospital and Indiana University IACUC committees (AR12-00035

Gene Symbol	[^] Our pooled IC Mean	[#] Chen <i>et al.</i> ²² pooled IC Mean [^]	[^] Our Single cell IC Mean (n = 4)	[^] Our pooled PC Mean	[#] Chen <i>et al.</i> ²² pooled PC Mean [^]
Antimicrobial peptide genes					
<i>Adm</i>	161.7	123.6	not detected	40.4	72.6
<i>Defb1</i>	8210.5	1973.5	598.2	5797.2	1263.2
<i>Scg5</i>	240.6	52.6	26.7	186.0	41.2
Bacteriostatic nutrient metal scavenger genes					
<i>Lcn2</i>	0.6	14.4	not detected	63.0	120.6
<i>S100g</i>	46394.7	4462.4	5267.5	18584.8	1229.6
<i>S100a11</i>	557.7	743.2	16	777.8	not listed
Phagocytosis/endocytosis genes					
<i>Cav1</i>	202.0	104.2	not detected	128.8	497.2
<i>Cav2</i>	450.0	388.7	not detected	213.6	759.4
<i>Rab7</i>	3710.2	314.7	8.25	2988.7	279.9
Inflammatory signaling genes					
<i>Irf3</i>	321.7	39.5	98	333.7	47.9
<i>Il1r1</i>	113.7	17.6	not detected	83.1	26.9
<i>ApoE</i>	1245.8	419.7	422.3	1766.0	2420.0
<i>Spp1</i>	18131	3524	1715.7	14659.5	1556.8
<i>Mmp11</i>	8	8.6	128.5	not detected	4.7
<i>Klk1</i>	124448	6039.2	4927	66958.6	899.8
Pattern recognition receptor genes					
<i>Tlr1</i>	1.3	6.2	not detected	22.5	13.3
<i>Tlr2</i>	nd	8.4	not detected	not detected	26.8
<i>Tlr3</i>	130.6	13.4	not detected	138.4	9.5
<i>Tlr4</i>	not detected	3.6	not detected	not detected	13.9
<i>Tril</i>	4.9	2.3	not detected	25.6	8.88
Inflammasome genes					
<i>Nlr3</i>	34.4	0.5	not detected	27.1	0.7
<i>Naip1</i>	not detected	0.1	not detected	not detected	not listed

Table 5. Innate immune cell marker comparison. [^]This RNASeq data was obtained using conditional mice expressing tdTomato on V-ATPase-B1 cre (presumed intercalated cells) and AQP2-cre (presumed principal cells). Data presented as mean normalized read count normalized using DESeq2. [#]This RNASeq data was collected from wild type C57BL/6 mice using DBA lectin (presumed principal cells) and c-kit (presumed intercalated cells) and C1 fluidigm for single cell RNAseq. Data presented as mean TPM (transcripts per kilobase Million) and was acquired from the supplemental data 03 (pooled cells, “All C-Kit(ICs) vs DBA(PCs) from Chen L and *et al.*²².

and 11333 respectively) and adhered to the “NIH Guide for the Care and Use of Laboratory Animals”. Mice were ~6–10 weeks of age were used. Male mice were used for sequencing experiments and female mice were used for the transurethral infection experiments. Experimental UTIs were induced in female C57BL/6 mice by transurethral inoculation with 1×10^8 uropathogenic E.coli (UPEC) as previously described. Mice were euthanized and kidney tissue obtained ~1 hour post UTI induction. The UTI induction experiment was completed at Indiana University for 1 mouse, the remainder of murine experiments were completed at Nationwide Children’s Hospital.

Flowsorting to enrich for renal intercalated and principal cells. Viable renal intercalated cells (IC) and principal cell (PC) were enriched from IC and PC reporter mice as we previously described¹⁴. This report along with the prior observed function of collecting duct as long as 8 hour following euthanasia of the animal supports the stable properties of collecting duct cells^{14,22,44}. Briefly, kidneys were harvested in cold PBS and immediately transferred to Accumax solution (cat.no. AM-105, Innovative Cell Technologies, CA) placed in C tubes (Miltenyi Biotec, CA). Cells were then rapidly disrupted to single cell suspension using gentleMACS dissociator (Miltenyi Biotec, CA) within 2 minutes. Renal resident immune cells separated by incubation with anti-mouse CD45-APC (Cat.no.17-0451, clone 30-F-11, eBiosciences, CA) for 30 min at 4°C. Doublets, clumps and dead cells were removed from the cell suspension and viable cells were obtained as previously described¹⁴. Enriched ICs and PCs came from three biological replicates/sorts, each from a distinct mouse

SMARTer stranded total RNA-sequencing. Following DNase treatment and cleanup of RNA from isolated cells using Zymo RNA cleanup and concentrator kit (Zymo research, CA) quality assessment was performed on Agilent 2100 bioanalyzer with RNA Pico chip kit (Agilent Technologies, CA). RNA integrity score ≥ 6 was used for cDNA library preparation from each sample. 10 ng of total RNA was used to generate RNA-Seq libraries using SMARTer stranded total RNA-Seq kit-Pico input mammalian kit (Clontech Laboratories Inc., CA). Briefly, total RNA was converted to cDNA using SMART[®] (Switching Mechanism at 5’ end of RNA template)

and locked nucleic acid (LNA) technology included as part of the template-switching oligo (TSO). Ribosomal cDNA (originating from rRNA) was then cleaved by ZapR in the presence of mammalian-specific R-probes. The library fragments originating from non-rRNA molecules were enriched via PCR amplification. Quality of the libraries were assessed via Agilent 2100 Bioanalyzer using DNA High Sensitivity Chip kit, and quantified using Kappa SYBR[®] Fast qPCR kit (KAPA Biosystems, MA). 151 bp sequence paired end reads were generated using the Illumina HiSeq 4000 platform (Illumina, CA).

RNA-Seq data analysis. On average, 50 million paired-end 151 bp RNA-Seq reads were generated for each sample (the range was 34 to 84 million). Each sample was aligned to the GRCm38.p4 assembly of the mouse reference from NCBI using version 2.5.0c of the RNA-Seq aligner STAR⁴⁵. Transcript features were identified from the GFF file provided with the GRCm38.p4 assembly from NCBI. Feature coverage counts were calculated using HTSeq (<http://www-huber.embl.de/users/anders/HTSeq/doc/count.html>). The raw RNA-Seq gene expression data was normalized and post-alignment statistical analyses performed using DESeq2 and custom analysis scripts written in R. Differentially expressed features were calculated using DESeq2⁴⁶. Maximum-likelihood estimates (MLE) were used to determine log₂ fold change which was converted to fold change⁴⁶. The differential expression results (including the p-values) we report for all comparisons were calculated by DESeq2, which uses “the procedure of Benjamini and Hochberg” for pooled cell analysis and for multiple testing adjustment, as specified in the DESeq2⁴⁶.

Comparisons of gene expression and associated statistical analysis were made between different conditions of interest using the normalized read counts. All fold change values are expressed as test condition/control condition, where values less than one are denoted as the negative of its inverse (note that there will be no fold change values between -1 and 1 , and that the fold changes of “1” and “ -1 ” represent the same value).

Comparison groups consisted of ICs versus non-ICs, PCs versus non-PCs and ICs versus PCs. The “IC vs. NIC” or “PC vs. NPC” worksheet shows significant differentially expressed features for the comparison of the IC group versus the NIC group. Innate immune genes were identified by keyword search (Supplemental Data S3) of the RNA-Seq results. Biological relevance of identified transcripts was evaluated using Ingenuity Pathway Analysis (IPA; Ingenuity Systems, Redwood City, CA. <http://www.ingenuitypathway.org>; version 14197757). Figures of canonical pathways and upstream regulators were made with the Ingenuity Pathway Analysis Software. Only transcripts whose absolute value of fold change is ≥ 2 and have an adjusted p value ≤ 0.05 were included in the Ingenuity pathway analysis and development of Tables 1–3.

Aqp2 mRNA *in-situ* hybridization. *Aqp2* mRNA was visualized in 4 μ m paraffin embedded mouse kidney sections using the QuantiGene ViewRNA ISH Tissue Assay according to the manufacturer’s protocol (Affymetrix, Santa Clara, CA). Prehybridization conditions were optimized with 5 minutes of incubation in pretreatment solution (90 °C) and 5 minutes of treatment with protease QF (40 °C). Sections were hybridized for 3 hours (40 °C) with a viewRNA Probe Set designed against *Aqp2* mRNA. The probes were then amplified (Pre Amplifier and Amplifier molecules), conjugated to alkaline phosphatase (Label Probe 1-AP), and bound with Fast Red Substrate. Finally, sections were counterstained with hematoxylin and imaged with the EVOS FL Auto Cell Imaging System (Life Technologies, Grand Island, NY). *In-situ* hybridized slides were incubated with anti-mouse V-ATPase E1 antibody (Sigma, St. Louis, MO) for 90 min at room temperature and then incubated with anti-chicken AF488 conjugated IgG secondary antibody (Jackson ImmunoResearch, PA) and mounted with DAPI nuclear stain (Vector Lab, CA). The V-ATPase E1 subunit, like the B1 subunit is specific for ICs in murine kidneys¹⁶. Slides were visualized under Keyence BZ9000 microscope (Keyence, Japan).

Targeted RT-PCR. Selected genes found to have higher relative expression in IC vs NIC or PC vs NPC RNA seq results were validated using real-time PCR. To do this IC and PC cells were sorted from respective reporter mice as previously described. RNA was prepared, with RNeasy plus micro kit (Qiagen, CA), quality was validated with Agilent Bioanalyzer and reverse transcribed to cDNA with high capacity reverse transcription kit (Applied Biosystems, CA) and amplified with specific primer sets (Sigma, MO). $\Delta\Delta$ CT method was used to calculate fold change in gene expression as previously described¹.

Functional validation of LPS-IL-1 mediated inhibition of RXRA function in Intercalated vs non-intercalated cells of UPEC infected mice. To functionally validate the top IC pathway (LPS/IL-1 mediated inhibition of RXR function) identified by ingenuity analysis, V-ATPase-cre⁺ Tdt⁺ (IC reporter) mice were transurethrally infected with 1×10^8 CFU UPEC or saline. After 1 hr, mice were sacrificed and kidney cell suspension was prepared. Tdt⁺ IC and Tdt⁻ non-IC cells were flowsorted as previously shown and *Rxra* mRNA expression was measured in sorted cells. Primer sequence of *Rxra* were as follows; *Rxra* Forward: TAACAGAGCTGGTGTCTAAG, Reverse *Rxra*: TTAGAGTCAGGGTTGAACAG.

Immunofluorescence. 4 μ m paraffin embedded kidney sections from AQP2-cre⁺/tdTomato^{fllox} and V-ATPase B1-cre⁺/tdTomato^{fllox} mice were stained with AQP2, V-ATPase B1, V-ATPase E1, and/or tdTomato antibodies to visualize the specificity of cre recombination. AQP2-cre⁺ mice: In this mice, to visualize AQP2, anti-mouse AQP2 antibody (Santa Cruz, CA), 1:200 dilution, 1 hr at 4 °C), to visualize V-ATPase, anti-mouse V-ATPase E1 antibody (1:500 dilution, Sigma, MO, overnight at 4 °C) were used, to visualize tdTomato, anti-RFP antibody (Rockland antibodies, PA) at 1:200 dilution was incubated overnight at 4 °C. Secondary antibody for AQP2 was Cy3 conjugated anti-goat IgG (1:300 dilution), for V-ATPase, AF488 conjugated anti-chicken IgG (1:300 dilution), for tdTomato, AF405 conjugated anti-rabbit IgG (1:300 dilution). Sections were imaged with Zeiss LSM 510 confocal microscope (Oberkochen, Germany). V-ATPase-cre⁺ mice: In this mice, to visualize, V-ATPase, goat anti-mouse V-ATPase B1/B2 antibody (E20) (sc21209, 1:50 dilution, Santa Cruz, CA), to visualize

TdTomato, anti-RFP-FITC conjugated antibody (ab34764, 1:200 dilution, abcam, MA) were incubated overnight at 4 °C. Secondary antibody for V-ATPase was Cy3 conjugated anti-goat IgG (1:300 dilution) was used for 1 hr at room temperature. All secondary antibodies were purchased from Jackson ImmunoResearch, PA. Sections were visualized under Keyence BZ-9000 microscope (Keyence, Japan).

Single cell capture, cDNA generation and sequencing. TdTomato⁺ cells flowsorted from IC reporter mice as described previously. Briefly kidneys were harvested and immediately transferred to Accumax solution (Innovative Cell technology, CA) and single cell suspension was rapidly prepared with GentleMACS dissociator (Miltenyi Biotec, CA) with Lung program 40 second and spleen program 60 seconds. TdTomato⁺ cells were flow-sorted. Sorted cells were pelleted and suspended to a final concentration of 200 cells/μl in sterile PBS. Cells were then combined with C1 Fluidigm suspension reagent at 3:2 ratio and loaded on a medium sized (for 10–17 μm diameter cells) Fluidigm IFC for single cell capture on the C1 system. Each IFC chamber was visually inspected under microscope to find wells with individual single cells and to avoid doublets or damaged cells (Fig. 5). Each cell was then washed, lysed and reverse transcribed using oligo dT primer and amplified by PCR automatically on C1 with SMART-seq v4 ultra low input RNA kit (Clontech, CA). cDNAs were harvested from the outlets of IFC and quantified by Qubit 2.0 fluorometer and DNA size was determined by Agilent 2100 Bioanalyzer using High-sensitivity DNA kit (Agilent, CA). RNA-Seq library was prepared with an input amount of 1 ng of cDNA from the Fluidigm Clontech step using the Nextera XT DNA library preparation kit (Illumina, CA) and sequenced on the Illumina HiSeq. 4000 platform (Illumina, CA) to a depth of 10 million paired reads per cell.

References

- Hasler, U. *et al.* Long term regulation of aquaporin-2 expression in vasopressin-responsive renal collecting duct principal cells. *The Journal of biological chemistry* **277**, 10379–10386, <https://doi.org/10.1074/jbc.M111880200> (2002).
- Brown, D., Hirsch, S. & Gluck, S. An H⁺-ATPase in opposite plasma membrane domains in kidney epithelial cell subpopulations. *Nature* **331**, 622–624, <https://doi.org/10.1038/331622a0> (1988).
- Hains, D. S. *et al.* Carbonic anhydrase 2 deficiency leads to increased pyelonephritis susceptibility. *American journal of physiology. Renal physiology*, <https://doi.org/10.1152/ajprenal.00344.2014> (2014).
- Spencer, J. D. *et al.* Ribonuclease 7 is a potent antimicrobial peptide within the human urinary tract. *Kidney Int* **80**, 174–180, <https://doi.org/10.1038/ki.2011.109> (2011).
- Spencer, J. D. *et al.* Human alpha defensin 5 expression in the human kidney and urinary tract. *PLoS one* **7**, e31712, <https://doi.org/10.1371/journal.pone.0031712> (2012).
- Paragas, N. *et al.* alpha-Intercalated cells defend the urinary system from bacterial infection. *J Clin Invest* **124**, 2963–2976, <https://doi.org/10.1172/JCI171630> (2014).
- Chassin, C. *et al.* Renal collecting duct epithelial cells react to pyelonephritis-associated Escherichia coli by activating distinct TLR4-dependent and -independent inflammatory pathways. *J Immunol* **177**, 4773–4784 (2006).
- St-Pierre, C. *et al.* Transcriptome sequencing of neonatal thymic epithelial cells. *Scientific reports* **3**, 1860, <https://doi.org/10.1038/srep01860> (2013).
- Lee, J. W., Chou, C. L. & Knepper, M. A. Deep Sequencing in Microdissected Renal Tubules Identifies Nephron Segment-Specific Transcriptomes. *J Am Soc Nephrol* **26**, 2669–2677, <https://doi.org/10.1681/ASN.2014111067> (2015).
- Robert-Nicoud, M. *et al.* Transcriptome of a mouse kidney cortical collecting duct cell line: effects of aldosterone and vasopressin. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 2712–2716, <https://doi.org/10.1073/pnas.051603198> (2001).
- Ryazanova, M. A. *et al.* The gene-expression profile of renal medulla in ISIAH rats with inherited stress-induced arterial hypertension. *BMC genetics* **17**, 151, <https://doi.org/10.1186/s12863-016-0462-6> (2016).
- Miller, R. L. *et al.* V-ATPase B1-subunit promoter drives expression of EGFP in intercalated cells of kidney, clear cells of epididymis and airway cells of lung in transgenic mice. *American journal of physiology. Cell physiology* **288**, C1134–C1144, <https://doi.org/10.1152/ajpcell.00084.2004> (2005).
- Nelson, R. D. *et al.* Expression of an AQP2 Cre recombinase transgene in kidney and male reproductive system of transgenic mice. *The American journal of physiology* **275**, C216–C226 (1998).
- Saxena, V. *et al.* Cell specific qRT-PCR of renal epithelial cells reveals a novel innate immune signature in murine collecting duct. *American journal of physiology. Renal physiology*, [ajprenal.00512.2016](https://doi.org/10.1152/ajprenal.00512.2016), <https://doi.org/10.1152/ajprenal.00512.2016> (2017).
- Azroyan, A. *et al.* Renal intercalated cells sense and mediate inflammation via the P2Y14 receptor. *PLoS One* **10**, e0121419, <https://doi.org/10.1371/journal.pone.0121419> (2015).
- Vedovelli, L. *et al.* Altered V-ATPase expression in renal intercalated cells isolated from B1 subunit-deficient mice by fluorescence-activated cell sorting. *American journal of physiology. Renal physiology* **304**, F522–F532, <https://doi.org/10.1152/ajprenal.00394.2012> (2013).
- Kurth, I. *et al.* The forkhead transcription factor Foxi1 directly activates the AE4 promoter. *Biochem J* **393**, 277–283, <https://doi.org/10.1042/BJ20051094> (2006).
- Yu, M. J. *et al.* Systems-level analysis of cell-specific AQP2 gene expression in renal collecting duct. *Proc Natl Acad Sci USA* **106**, 2441–2446, <https://doi.org/10.1073/pnas.0813002106> (2009).
- Rubera, I. *et al.* Collecting duct-specific gene inactivation of alphaENaC in the mouse kidney does not impair sodium and potassium balance. *J Clin Invest* **112**, 554–565, <https://doi.org/10.1172/JCI16956> (2003).
- Legouis, D. *et al.* Ex vivo analysis of renal proximal tubular cells. *BMC Cell Biol* **16**, 12, <https://doi.org/10.1186/s12860-015-0058-4> (2015).
- Hoyer, J. R., Sisson, S. P. & Vernier, R. L. Tamm-Horsfall glycoprotein: ultrastructural immunoperoxidase localization in rat kidney. *Lab Invest* **41**, 168–173 (1979).
- Chen, L. *et al.* Transcriptomes of major renal collecting duct cell types in mouse identified by single-cell RNA-seq. *Proc Natl Acad Sci USA* **114**, E9989–E9998, <https://doi.org/10.1073/pnas.1710964114> (2017).
- Roszer, T., Menendez-Gutierrez, M. P., Cedenilla, M. & Ricote, M. Retinoid X receptors in macrophage biology. *Trends Endocrinol Metab* **24**, 460–468, <https://doi.org/10.1016/j.tem.2013.04.004> (2013).
- Ma, F. *et al.* Retinoid X receptor alpha attenuates host antiviral response by suppressing type I interferon. *Nat Commun* **5**, 5494, <https://doi.org/10.1038/ncomms6494> (2014).
- Paunescu, T. G. *et al.* Vasopressin induces apical expression of caveolin in rat kidney collecting duct principal cells. *Am J Physiol Renal Physiol* **305**, F1783–F1795, <https://doi.org/10.1152/ajprenal.00622.2012> (2013).
- Da Silva, N. *et al.* Proteomic analysis of V-ATPase-rich cells harvested from the kidney and epididymis by fluorescence-activated cell sorting. *Am J Physiol Cell Physiol* **298**, C1326–C1342, <https://doi.org/10.1152/ajpcell.00552.2009> (2010).

27. Wu, H. *et al.* Aqp2-expressing cells give rise to renal intercalated cells. *J Am Soc Nephrol* **24**, 243–252, <https://doi.org/10.1681/ASN.2012080866> (2013).
28. Trepiccione, F. *et al.* A fate-mapping approach reveals the composite origin of the connecting tubule and alerts on “single-cell”-specific KO model of the distal nephron. *Am J Physiol Renal Physiol* **311**, F901–F906, <https://doi.org/10.1152/ajprenal.00286.2016> (2016).
29. Park, J. *et al.* Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**, 758–763, <https://doi.org/10.1126/science.aar2131> (2018).
30. Fischer, H. *et al.* Pathogen specific, IRF3-dependent signaling and innate resistance to human kidney infection. *PLoS Pathog* **6**, e1001109, <https://doi.org/10.1371/journal.ppat.1001109> (2010).
31. Murawski, I. J. *et al.* The C3H/HeJ inbred mouse is a model of vesico-ureteric reflux with a susceptibility locus on chromosome 12. *Kidney Int* **78**, 269–278, <https://doi.org/10.1038/ki.2010.110> (2010).
32. Allaker, R. P. *et al.* Mechanisms of adrenomedullin antimicrobial action. *Peptides* **27**, 661–666, <https://doi.org/10.1016/j.peptides.2005.09.003> (2006).
33. Shooshtarizadeh, P. *et al.* The antimicrobial peptides derived from chromogranin/secretogranin family, new actors of innate immunity. *Regul Pept* **165**, 102–110, <https://doi.org/10.1016/j.regpep.2009.11.014> (2010).
34. Zhao, G., Su, Z., Song, D., Mao, Y. & Mao, X. The long noncoding RNA MALAT1 regulates the lipopolysaccharide-induced inflammatory response through its interaction with NF- κ B. *FEBS Lett* **590**, 2884–2895, <https://doi.org/10.1002/1873-3468.12315> (2016).
35. Yamasaki, K. *et al.* Kallikrein-mediated proteolysis regulates the antimicrobial effects of cathelicidins in skin. *FASEB J* **20**, 2068–2080, <https://doi.org/10.1096/fj.06-6075com> (2006).
36. Dikov, M. M. *et al.* Differential roles of vascular endothelial growth factor receptors 1 and 2 in dendritic cell differentiation. *J Immunol* **174**, 215–222 (2005).
37. Laxmanan, S. *et al.* Vascular endothelial growth factor impairs the functional ability of dendritic cells through Id pathways. *Biochem Biophys Res Commun* **334**, 193–198, <https://doi.org/10.1016/j.bbrc.2005.06.065> (2005).
38. Wang, K. X. & Denhardt, D. T. Osteopontin: role in immune regulation and stress responses. *Cytokine Growth Factor Rev* **19**, 333–345, <https://doi.org/10.1016/j.cytogfr.2008.08.001> (2008).
39. McKee, M. D. & Nanci, A. Secretion of Osteopontin by macrophages and its accumulation at tissue surfaces during wound healing in mineralized tissues: a potential requirement for macrophage adhesion and phagocytosis. *Anat Rec* **245**, 394–409, [https://doi.org/10.1002/\(SICI\)1097-0185\(199606\)245:2<394::AID-AR19>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1097-0185(199606)245:2<394::AID-AR19>3.0.CO;2-K) (1996).
40. Koh, A. *et al.* Role of osteopontin in neutrophil function. *Immunology* **122**, 466–475, <https://doi.org/10.1111/j.1365-2567.2007.02682.x> (2007).
41. Weiss, J. M. *et al.* Osteopontin is involved in the initiation of cutaneous contact hypersensitivity by inducing Langerhans and dendritic cell migration to lymph nodes. *J Exp Med* **194**, 1219–1229 (2001).
42. Cho, S., Beintema, J. J. & Zhang, J. The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. *Genomics* **85**, 208–220, <https://doi.org/10.1016/j.ygeno.2004.10.008> (2005).
43. Liang, J. *et al.* The C-kit receptor-mediated signal transduction and tumor-related diseases. *Int J Biol Sci* **9**, 435–443, <https://doi.org/10.7150/ijbs.6087> (2013).
44. Grantham, J. J. & Burg, M. B. Effect of vasopressin and cyclic AMP on permeability of isolated collecting tubules. *Am J Physiol* **211**, 255–259, <https://doi.org/10.1152/ajplegacy.1966.211.1.255> (1966).
45. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21, <https://doi.org/10.1093/bioinformatics/bts635> (2013).
46. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, <https://doi.org/10.1186/s13059-014-0550-8> (2014).

Acknowledgements

The study was funded by NIDDK R01DK106286-01 and Nationwide Children’s Hospital (NCH) Internal funds. These studies were performed while ALS was at NCH. We acknowledge Dr. Mark Knepper’s assistance including facilitating use of the Epithelial Systems Biology Laboratory Database at the National Heart, Lung and Blood Institute as part of its Kidney Systems Biology Project. LAW is supported by an Ohio State University Pelotonia Undergraduate Research Fellowship.

Author Contributions

V.S. and A.L.S. wrote the main manuscript text. A.L.S., V.S. and J.F. prepared figures. V.S., A.L.S., D.S.H., R.D.N., V.V., P.Y., G.J.S. and D.S.H. were involved in experimental design, V.S., J.F., J.K., A.W., S.W.A., M.A.C., K.R.P., L.A.W. and X.C. performed experiments, V.S., J.F., J.K., P.W., A.W., J.D.S., B.B., A.L.S., L.A.W., X.C., P.W., G.J.S. and D.S.H. were involved in data interpretation. All authors reviewed the manuscript. A.L.S. and D.S.H. are corresponding authors.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-36921-z>.

Competing Interests: Dr. Schwaderer has consulted for Allena Pharmaceuticals on a topic unrelated to this manuscript.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019