nature medicine

Article

https://doi.org/10.1038/s41591-023-02314-7

The transcriptional and regulatory identity of erythropoietin producing cells

Received: 23 January 2022

Accepted: 17 March 2023

Published online: 27 April 2023

Check for updates

Bjørt K. Kragesteen[1,17] ✉, Amir Giladi[1,2,3,17], Eyal David[1,17], Shahar Halevi[1,17], Laufey Geirsdóttir[1], Olga M. Lempke[4], Baoguo Li[1], Andreas M. Bapst[4], Ken Xie[1], Yonatan Katzenelenbogen[1], Sophie L. Dahl[4], Fadi Sheban[1], Anna Gurevich-Shapiro[1,5,6], Mor Zada[1,6], Truong San Phan[1], Roberto Avellino[1], Shuang-Yin Wang[1], Oren Barboy[1], Shir Shlomi-Loubaton[1], Sandra Winning[7], Philipp P. Markwerth[8], Snir Dekalo[6,9], Hadas Keren-Shaul[10], Merav Kedmi[10], Martin Sikora[11], Joachim Fandrey[7], Thorfinn S. Korneliussen[11], Josef T. Prchal[12], Barak Rosenzweig[6,13], Vladimir Yutkin[14], Fernando Racimo[11,17], Eske Willerslev[11,17], Chamutal Gur[1,15,17,18], Roland H. Wenger[4,16,17,18] & Ido Amit[1,17,18] ✉

Erythropoietin (Epo) is the master regulator of erythropoiesis and oxygen homeostasis. Despite its physiological importance, the molecular and genomic contexts of the cells responsible for renal Epo production remain unclear, limiting more-effective therapies for anemia. Here, we performed single-cell RNA and transposase-accessible chromatin (ATAC) sequencing of an Epo reporter mouse to molecularly identify Epo-producing cells under hypoxic conditions. Our data indicate that a distinct population of kidney stroma, which we term Norn cells, is the major source of endocrine Epo production in mice. We use these datasets to identify the markers, signaling pathways and transcriptional circuits characteristic of Norn cells. Using single-cell RNA sequencing and RNA in situ hybridization in human kidney tissues, we further provide evidence that this cell population is conserved in humans. These preliminary findings open new avenues to functionally dissect *EPO* gene regulation in health and disease and may serve as groundwork to improve erythropoiesis-stimulating therapies.

All mammalian cells require oxygen for essential biochemical reactions. To meet the unceasing cellular demand for an oxygen supply, roughly 2–3 million erythrocytes are generated per second[1]. This process, known as erythropoiesis, is tightly regulated both systemically and locally; although all cells can sense oxygen deprivation (hypoxia), the kidney is the primary source of the endocrine hormone Epo in adults[2–4]. Epo is secreted into the plasma and subsequently binds to Epo receptor-expressing erythroid progenitors in the bone

[1]Department of Systems Immunology, Weizmann Institute of Science, Rehovot, Israel. [2]Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences, Utrecht, the Netherlands. [3]Oncode Institute, Utrecht, the Netherlands. [4]Institute of Physiology, University of Zurich, Zurich, Switzerland. [5]Division of Haematology, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel. [6]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. [7]Institute of Physiology, University of Duisburg-Essen, Essen, Germany. [8]Institute for Forensic Medicine, University Hospital Essen, Essen, Germany. [9]Urology Department, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel. [10]Department of Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot, Israel. [11]GLOBE Institute, University of Copenhagen, Copenhagen, Denmark. [12]Department of Medicine, University of Utah, Salt Lake City, UT, USA. [13]Department of Urology, Sheba Medical Center, Ramat Gan, Israel. [14]Department of Urology, Hadassah Medical Center, Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel. [15]Department of Medicine, Hadassah Medical Center, Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel. [16]National Centre of Competence in Research 'Kidney.CH', University of Zurich, Zurich, Switzerland. [17]These authors contributed equally: Bjørt K. Kragesteen, Amir Giladi, Eyal David, Shahar Halevi, Fernando Racimo, Eske Willerslev, Chamutal Gur, Roland H. Wenger, Ido Amit. [18]These authors supervised this work: Chamutal Gur, Roland H. Wenger, Ido Amit. ✉e-mail: bjort.kragesteen@weizmann.ac.il; ido.amit@weizmann.ac.il

marrow to promote their survival, proliferation and differentiation, thereby regulating erythrocyte levels in circulation and controlling the systemic oxygen supply[5]. Clinically, altered Epo production can be detrimental, leading to either anemia or erythrocytosis[6,7]. Recombinant Epo is widely used, especially in treating patients with end-stage renal disease (ESRD), in which Epo expression is lost despite pronounced tissue hypoxia[8]. However, prolonged Epo injections are associated with severe side effects due to unbalanced dosage and off-target activity, resulting in an increased risk of thrombosis, hypertension, tumor progression and death[9,10]. There is thus a pressing need for alternative therapeutic approaches that restore physiological Epo levels.

Surprisingly little is known about the cellular and molecular determinants of *Epo* gene regulation in vivo. Key studies in hepatocellular cancer cell lines and animal knockout models have demonstrated that the transcription factor (TF) hypoxia-inducible factor (HIF)-2α (encoded by the gene *Epas1*) is a master regulator of *Epo* transcription[11–13]. When oxygen is readily available (normoxia), HIF-2α is rapidly degraded by the oxygen-sensing enzyme prolyl-4-hydroxylase domain (PHD)2 (encoded by *Egln1*), whereas in hypoxia, HIF-2α accumulates in the nucleus and forms a complex with the constitutively expressed HIF-1β. Together, the HIF-2α–HIF-1β heterodimer binds to a distinct core DNA motif (RCGTG), also known as the hypoxia response element (HRE), regulating downstream transcription[14,15]. The *EPAS1* and *EGLN1* loci show evidence of strong evolutionary selection in humans and other mammalian species that are adapted to live in high-altitude hypoxic environments[16–19]. In the Tibetan population, protective variations in these loci have been associated with reduced erythrocytosis-related mortality and increased birth rates compared with recent Han Chinese migrations to the Tibetan plateau[20–22]. Interestingly, *Epas1* is highly expressed by various cell types that are incapable of Epo production, and it remains unclear which additional factors orchestrate *Epo* transcription in the kidney[23].

During the past decades, various renal cell types have been indirectly implicated as the source of Epo production, including fibroblasts, pericytes, telocytes, mesangial, endothelial, epithelial and renin-producing cells[24–28]. Some stromal surface markers have been identified, including CD73 (encoded by *Nt5e*), CD133 (encoded by *Prom1*) and platelet-derived growth factor receptor beta (PDGFRβ) (encoded by *Pdgfrb*), but these markers are not restricted to Epo-producing cells[27,29,30]. As of yet, no single cell type or group of cells has been unequivocally attributed to renal Epo production in adult humans or animal models under physiological conditions[31]. One possibility is that Epo expression is restricted to a specialized, yet uncharacterized, cell type, similar to insulin secretion by pancreatic β-cells[32]. Alternatively, Epo production may be associated with a transient cell state activated by a heterogenous group of cell types, as is the case for the cytokine tumor necrosis factor alpha[33]. Another scenario is that hypoxia may initially cause the differentiation of Epo-incapable precursor cells into Epo-capable cells before the Epo gene is accessible to induction mediated by HIF-2α–HIF-1β. Underlying these uncertainties are experimental challenges, mainly low basal Epo transcript levels, rendering it practically undetectable except for under severe hypoxic conditions[34]. To complicate matters, only a subset of cells capable of Epo production activate Epo at any given time upon hypoxia stimulus[34]. Moreover, HIF-α protein and consequently Epo messenger RNA, undergo rapid degradation upon reoxygenation[35]. Together, these conditions pose major challenges to the isolation and subsequent characterization of this key physiological pathway in health and disease, and limit the potential for more effective translation of Epo-based therapies[24].

Previous studies detected Epo-producing cells by applying in situ hybridization[35,36], but could not associate Epo production with a specific cell type. Past efforts focused on the development of genetically altered mouse models to indirectly label Epo-producing cells, however the initial knock-in of reporter genes into the *Epo* locus resulted in gene inactivation and anemia[37]. Reporter mice targeting other implicated genes, including *Col1a1*, *Cd68*, *Ren1*, *Pdgfrb* and *Foxd1*, exhibited broad labeling spectra unconfined to renal Epo-producing cells and failed to recapitulate the acute oxygen-dependent function of these cells[25,27,30,38,39]. The use of a constitutive Cre recombinase allowed for the permanent activation of a reporter, but it remained unknown whether labeled cells maintained their cellular identity and capacity to produce EPO[37]. Therefore, we previously developed a conditional Epo-CreERT mouse model (hereafter referred to as *Epo-tdT*$^+$) that permanently labels actively producing cells[24]. Under hypoxic conditions, this mouse exclusively labels Epo mRNA-positive interstitial cells in the corticomedullary border region of the kidney[24,35].

Here, by performing single-cell multiomic profiling of cells labeled with Epo-CreERT, we characterize the Epo-producing cells of the kidney and describe their unique molecular and epigenetic identity. We found that Epo production, similar to insulin production by pancreatic β-cells, is restricted to a single, unique and evolutionary conserved population of kidney fibroblast-like cells, which we name Norn cells, after the mythological Norse entities believed to govern human fates[40].

Results

Enrichment of Epo-producing cells in mouse kidney

Several organs have been reported to maintain certain levels of Epo production in adults (brain, testis, liver, spleen and bone marrow), but the indispensable source of circulating Epo in humans is the kidney[41]. Similarly, we found that *Epo* gene expression in response to hypoxia is mainly contained in the mouse kidney (Fig. 1a). Inhalation of 0.1% CO activated progressive functional hypoxemia and elicited a rapid pulse of *Epo* expression in the kidney, peaking at 4 h (Fig. 1b). Despite this evident source of Epo production in the hypoxic kidney, we could not identify a population of *Epo* mRNA-positive cells in existing normoxic kidney single-cell datasets[42–47]. Potential challenges for the investigation of Epo-producing cells include low *Epo* mRNA levels under normoxic conditions, the lack of a unique molecular signature or characteristic cell-surface markers and high autofluorescence of certain kidney populations. To develop a strategy for enriching and isolating Epo-producing cells, we performed indexed fluorescence-activated cell sorting (FACS) on labeled cells from the Epo-CreERT reporter mouse (*Epo-tdT*$^+$) (Fig. 1c). We calibrated and tested multiple sorting schemes to enrich for tdTomato$^+$ cells using massively parallel single-cell RNA sequencing (MARS-seq) as a readout of enrichment efficiency (Fig. 1c)[48,49]. Even in this carefully gated population, the majority of index sorted tdTomato$^+$ cells were negative for *Epo* mRNA by MARS-seq analysis (Fig. 1d). Single-cell transcriptional analysis of index sorted tdTomato$^+$ cells revealed a heterogeneous mixture of stromal, endothelial, immune and epithelial lineages based on their transcriptional profile (Fig. 2). We found that several cell types generated autofluorescence impurities interfering with both the fluorescein isothiocyanate (FITC; 495–519 nm) and phycoerythrin (574 nm) channels. We therefore further calibrated our gating strategy to exclude cells fluorescently labeled by lineage markers for endothelial, immune and epithelial lineages, which enabled an enrichment level of *Epo*-producing cells suitable for further molecular analysis (Fig. 1c,d).

Molecular profiling of Norn cells

To characterize Norn cellular identity, we performed single-cell RNA sequencing (scRNA-seq) on 35,312 renal cells from C57BL/6J and *Epo-tdT*$^+$ mice from normoxic and hypoxic experimental conditions to generate a Norn-enriched atlas (Fig. 2a and Extended Data Fig. 1a,b). We used the MetaCell package for scRNA-seq analysis to group normoxic and hypoxic cells into homogenous transcriptional entities termed metacells[50]. To obtain meaningful metacell annotations, we imposed a cellular reference derived from a previous scRNA-seq atlas on our data, representing a myriad of cell types characteristic of the mouse kidney (Fig. 2a, Extended Data Fig. 2a,b and Supplementary Tables 1

https://doi.org/10.1038/s41591-023-02314-7

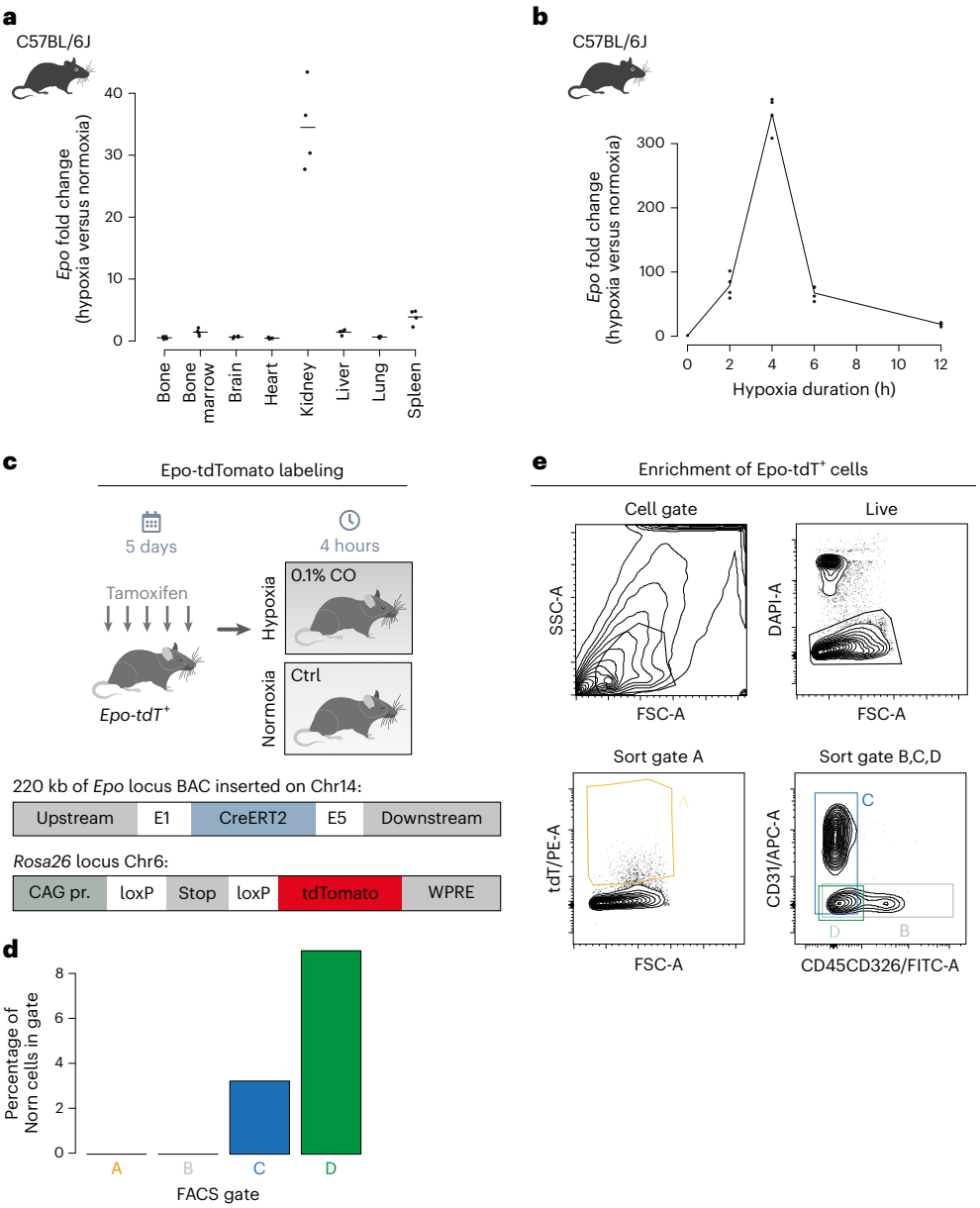


Fig. 1 | Detection and enrichment of renal Epo-producing cells. a, *Epo* expression analysis across tissues in hypoxic mice (0.1% CO, 4 h). *Epo* qPCR normalized to the *Actb* housekeeping gene. Data are shown as a fold change in hypoxia over normoxia ($n = 1$ independent experiment using tissues from $n = 1$ male mouse, $n = 4$ technical replicates). **b**, *Epo* expression during a hypoxia time course (0.1% CO, 4 h). Data were normalized to *Actb* and are shown as fold change in hypoxia over normoxia ($n = 1$ independent experiment using kidneys from $n = 1$ male mouse per time point, $n = 4$ technical replicates). **c**, Transgenic mice carry a bacterial artificial chromosome (BAC) insertion of the *Epo* locus with CreERT2 replacing *Epo* exons 2 to 4. The *Rosa26* locus harbors a tdTomato transgene that is driven by the CAG promoter, followed by a STOP cassette flanked by loxP sites. Tamoxifen and hypoxia treatment (0.1% CO, 4 h) activate tdTomato and permanently label the Epo-positive cells. **d**, Representative gating strategy to optimize enrichment of tdTomato$^+$ cells. Sort gate A contains live, tdTomato$^+$ cells. Sort gate B contains live, tdTomato$^+$CD31$^-$ cells. Sort gate C contains live, tdTomato$^+$CD45$^-$CD326$^-$ cells. Sort gate D contains live, tdTomato$^+$CD45$^-$CD326$^-$CD31$^-$ cells. **e**, Percentage enrichment of true tdTomato$^+$ cells in sort gates from **d**, as determined by single-cell RNA-seq combined with index-sorting analysis. APC-A, APC area; Ctrl, control; DAPI-A, DAPI area; FSC-A, forward scatter area; PE-A, PE area; SSC-A, side scatter area; WPRE, woodchuck hepatitis virus posttranscriptional regulatory element.

and 2)[43]. Using this approach, we identified the major renal cellular compartments, comprising various epithelial cells of the nephron (proximal tubule, collecting duct, distal convoluted tubule and loop of Henle), vascular endothelial cells, podocytes, as well as circulating and resident immune cells (monocytes, macrophages, T lymphocytes and neutrophils) (Fig. 2a). However, even with this detailed atlas, a large fraction of the kidney stromal compartment remained uncharacterized.

Annotation based on known stromal markers identified distinct clusters of fibroblast (*Pdgfra* and *Nt5e* encoding for CD73) and pericyte (*Acta2* and *Rgs5*) origins (Fig. 2a,b and Extended Data Fig. 2b)[31]. We report four major clusters of pericytes that can be characterized by expression of *Acta2*, *Rgs5*, *Mcam* and *Myh11*, with one population distinguished by the expression of renin (Ren1$^+$), part of the renin–angiotensin–aldosterone system involved in blood pressure and fluid control (Fig. 2b, lower)[51]. Interestingly, *Epo* transcripts were exclusively detected in a homogenous and specific cellular type defined as *Cxcl14*$^+$, *Col1a1*$^+$, *Dcn*$^+$, *Lpar1*$^+$, representing bona fide Norn cells (Fig. 2b and Extended Data Fig. 2b).

https://doi.org/10.1038/s41591-023-02314-7

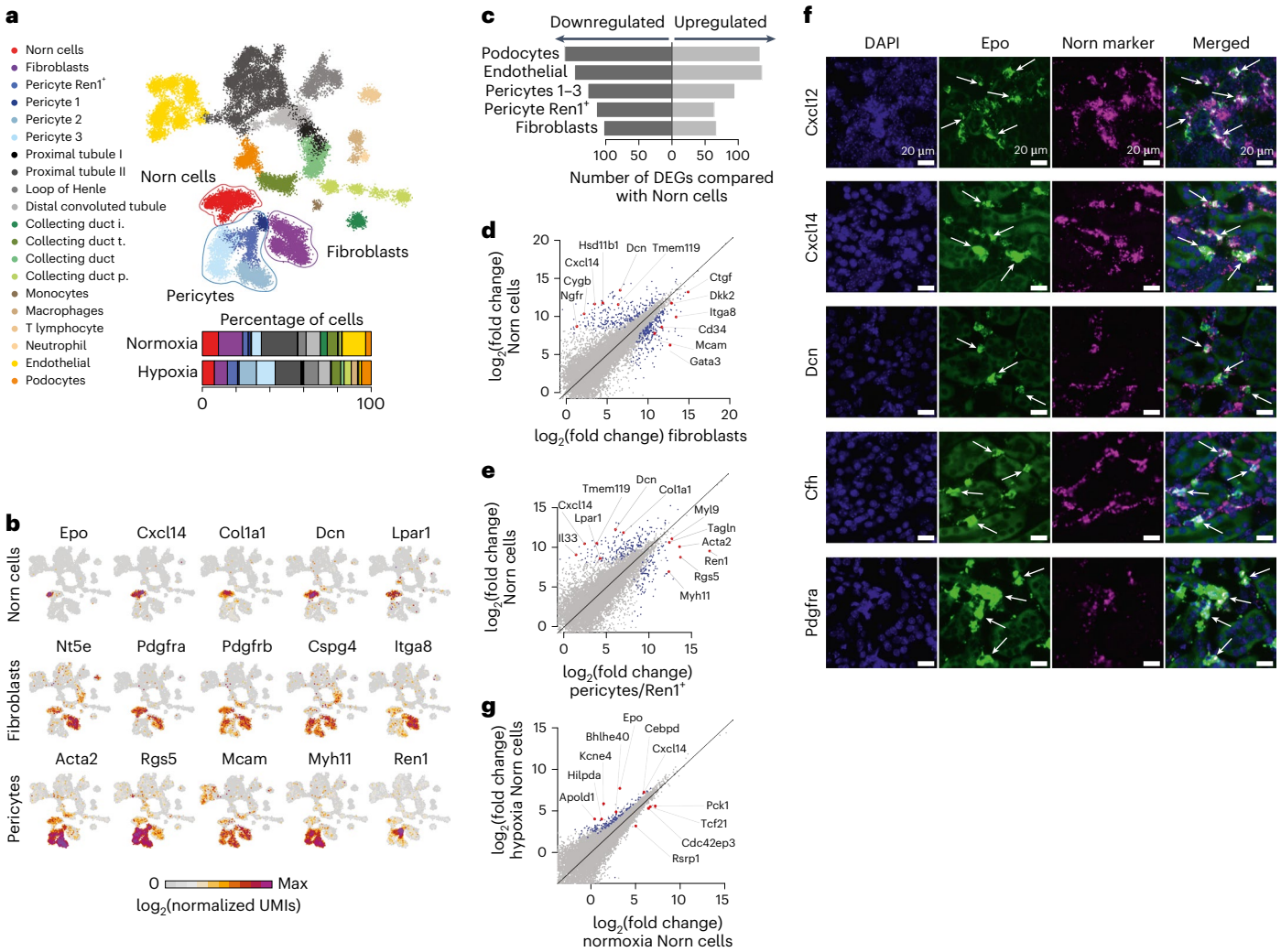


Fig. 2 | Epo-producing Norn cells are a distinct stromal cell type. a, Two-dimensional projection of 35,834 renal cells combined from normoxic and hypoxic (0.1% CO, 4 h) mice. Cells were grouped into 385 metacells, based on the MetaCell algorithm[50]. Dots represent single cells and colors denote cell-type annotation. The stack bar shows the percentage of renal cell types obtained in hypoxic (23,650 cells) and normoxic (12,184 cells) mouse experiments. Cells were pooled from $n = 22$ independent experiments with kidneys from $n = 52$ mice (33 males, 19 females). Collecting duct i., collecting duct intercalated; collecting duct t., collecting duct transient; collecting duct p., collecting duct principal. **b**, Two-dimensional projections of representative expression ($\log_2$(normalized UMI count)) and distribution of known marker genes for Norn cells, fibroblasts and pericytes. **c**, Summary of differential expression between Norn cells and selected cell populations. Differentially expressed genes (DEGs) were determined by $\log_2$(fold change) >1 and FDR adjusted $P < 0.0001$ (two-way chi-squared test). Data are shown as $\log_2$(sum of gene counts). **d**,**e**, Differential expression between Norn cells and (**d**) fibroblasts or (**e**) pericytes (pooled over all pericyte subsets). Data are shown as $\log_2$(size-normalized pooled expression) for equal numbers of cells. Blue dots denote significantly differential genes, red dots denote selected genes shown. **f**, RNA in situ hybridization of hypoxic mouse kidneys (0.1% CO, 4 h), $n = 1$ biological independent sample, $n = 3$ independent experiments. Representative images showing *Epo* and candidate Norn markers, from top: *Cxcl12*, *Cxcl14*, *Dcn*, *Cfh* and *Pdgfra*. Scale bar, 20 µm. Nuclei were stained with DAPI. **g**, Differential expression between mouse Norn cells derived from normoxic or hypoxic conditions (0.1% CO, 4 h). Data are shown as $\log_2$(size-normalized pooled expression) for equal numbers of cells.

Despite sharing expression of key fibroblast and pericyte markers such as *Pdgfra*, *Pdgfrb*, *Rgs5* and *Nt5e*, Norn cells are distinct from both fibroblasts and pericytes, exhibiting hundreds of differentially expressed genes compared with each of the other stromal populations (Fig. 2b–e and Supplementary Table 3). In particular, Norn cells showed abundant expression of *Hsd11b1* (encoding an enzyme regulating cortisol activity, relevant for glucocorticoid and anti-inflammatory signaling pathways), *Cxcl14*, *Cxcl12* (encoding ligands of Cxc chemokine system involved in angiostatic and angiogenic processes), *Cygb* (encoding a cytoglobin involved in intracellular oxygen transport) and *Cfh* (encoding a regulator of the complement pathway) absent from other kidney lineages (Fig. 2b–f). We further used mRNA fluorescent in situ hybridization (mRNA-FISH) to validate that Epo-expressing Norn cells are enriched for *Cxcl12*, *Cxcl14*, *Dcn*, *Cfh* and *Pdgfra* mRNA in hypoxic kidney (Fig. 2f and Extended Data Fig. 2c).

Building on these data, we managed to effectively isolate and sort for Epo-expressing Norn cells from wild-type C57BL/6J mice by using a carefully calibrated set of negative and positive markers, including the surface marker CD73 (Extended Data Fig. 2d)[52]. Although Norn cells are the only population exhibiting *Epo* expression, *Epo* mRNA was detected in only a fraction of hypoxic Norn cells (10.8%), highlighting the burst-like nature of Epo transcription[34]. Analysis of Norn cells from normoxic and hypoxic conditions revealed 134 differentially regulated genes (Supplementary Table 4), of which 49 are exclusive to Norn cells in the stromal compartment (Fig. 2g). In addition to Epo itself, this Norn-specific hypoxia module includes known hypoxia-regulated

genes such as *Hilpda* (Hypoxia Inducible Lipid Droplet Associated)[53] and *Bhlhe40* (Basic Helix-Loop-Helix Family Member E40) along with uncharacterized hypoxia-regulated genes such as *Kcne4* (Potassium Voltage-Gated Channel Subfamily E Regulatory Subunit 4)[54] (Fig. 2g). Compellingly, we found that the Norn TFs *Tcf21* and *Cebpd* are regulated in hypoxia (Fig. 2g). Together, our data support the existence of a specialized renal cell type capable of Epo production (Norn cells), which possesses a molecular identity discrete from any known cell type.

Norn cell regulatory landscape

To define the underlying distal and proximal DNA regulatory elements and TF circuits controlling Norn identity, we generated genomewide single-cell profiling of Epo-tdTomato$^+$ cells from hypoxic mice, simultaneously recovering mRNA and transposase-accessible chromatin (ATAC) regions from the same nuclei (Fig. 3a and Extended Data Fig. 1c,d). After alignment and quality controls, we obtained 3,306 and 3,862 cells containing ATAC-seq and RNA-seq reads, respectively, with 3,266 cells containing both RNA and ATAC signals. We projected the corresponding RNA-seq data on our MetaCell reference map, assigning cells their predefined cellular identity (Fig. 3b and Extended Data Fig. 3a,b), overall retrieving 188 (5.7%) Norn cells with ATAC-seq signal.

Clustering of open proximal chromatin regions 1 kb upstream to 100 bp downstream of the transcriptional start sites (TSS) showed that the different cell populations tend to share promoter peaks, denoting both poised and active promoters (Fig. 3c). By contrast, on clustering distal enhancer peaks, intronic or intergenic regions, cell-type-specific patterns emerged (Fig. 3d). This is in line with similar studies showing that cell-type-specific gene regulation is embedded within distal enhancers bound by cell-type-specific TFs[55,56]. From genetic studies, the upstream region (up to 10 kb from TSS) of *Epo* has been shown to contain the regulatory elements necessary for renal *Epo* expression[57,58]. However, the precise genomic sequences of *Epo* enhancers are unknown. We thus first searched for new candidate enhancer regions and identified five open chromatin regions upstream and downstream of the *Epo* gene, with two peaks 4.3 and 6.2 kb upstream of the *Epo* TSS that are specific to Norn cells (Fig. 3e). Moreover, these regions are highly conserved in mammals (Fig. 3e). This prompted us to perform a genomewide scan for candidate enhancers regulating the Norn transcriptional program. Analysis of unique peaks across renal cell populations identified 5,844 unique open chromatin regions in Norn cells, many of which are near genes expressed specifically in Norn cells, including *Epo*, *Prrx1*, *Il33* and *Cxcl14* (Fig. 3e, Extended Data Fig. 3c and Supplementary Table 5).

TFs cooperatively bind to specific DNA sequence motifs concentrated within enhancer regions to dictate cell-type-specific enhancer activities and target promoter output. We thus investigated which TF motif sequences are enriched within Norn-specific open chromatin regions using the HOMER package[59]. We identified more than 30 TF motifs significantly enriched in Norn-specific open chromatin regions (Extended Data Fig. 4) and matched the top hits with differential gene expression. Of those, the TCF21 motif is the most abundant (26.1%, $P = 1 \times 10^{-196}$) (Fig. 3f). Correspondingly, *Tcf21* expression is highly enriched in Norn cells (Fig. 3g). The C/EBP motif was identified in 13.1% ($P = 1 \times 10^{-82}$) Norn-specific peaks (Fig. 3f). Of the C/EBP family, *Cebpd* expression is highly enriched in Norn cells (Fig. 3g). Of note, both TCF21 and C/EBPδ motifs are present in the Norn *Epo* enhancer elements (Fig. 3e). The GATA6 motif is detected in 12.4% ($P = 1 \times 10^{-39}$) of unique Norn regions and its transcripts are highly enriched in Norn cells (Fig. 3f,g). Notably, we did not identify an enrichment of HIF motifs over the background within Norn-specific open chromatin regions using this analysis (2%–4%) (Fig. 3f). This motivated us to extend the analysis of DNA sequence motifs identified within the stromal compartment using the GimmeMotifs package, which implements a consensus database of known TFs[60]. Broadly, we find that the basic-helix-loop-helix (bHLH) family of TFs is highly enriched in Norn cells (Fig. 3h). Within the bHLH-e subfamily, the HRE motif is enriched in Norn-specific regions, albeit at a low prevalence (Fig. 3h). HIF-2α, which belongs to the bHLH-e family, is highly expressed in Norn cells, endothelial cells, pericytes and fibroblasts (Fig. 3g).

To find candidate target genes downstream of Norn TFs, we performed gene regulatory analysis using SCENIC on our normoxic and hypoxic MARS-seq datasets separately[61]. Tcf21 emerges as the singular strongly enriched TF in Norn cells in both normoxia and hypoxia (Extended Data Fig. 5a,b). The analysis identifies a dozen Tcf21 target genes, including the Norn markers *Dcn* and *Cxcl14*, as well as Norn-enriched genes *Hsd11b1* and *Igfbp3* (Extended Data Fig. 5c). Together, these data suggest that Tcf21 is a key transcriptional regulator of Norn cells.

In summary, we uncovered thousands of unique candidate enhancer regions that likely play a role in the cellular identity and function of Norn cells. These regions are enriched in motifs of the TFs TCF21, C/EBPδ and GATA6, suggesting this TF circuitry as a new component of HIF-2α–HIF-1β *Epo* gene regulation and governing Norn cellular identity.

Human Norn cells are the source of Epo production

We next asked whether Norn cell identity is conserved in the human kidney. To address this question, we performed a retrospective scan for evidence of Norn cells in published scRNA-seq datasets across platforms and developmental stages in mouse and human. We acquired a characteristic gene signature for each cell population identified in our MARS-seq data (Supplementary Tables 1, 2) and tested our cell-type identification scheme using these signatures. The Norn cell population was correctly mapped at 97.8% accuracy when tested on our murine MARS-seq data (Extended Data Fig. 6a, b). We subsequently screened five scRNA-seq datasets from adult and fetal mouse and human kidneys. Although these studies differ in the experimental procedures, Norn cells could be detected at varying frequencies. We identified 0.1% Norn cells in a dataset of whole adult mouse kidneys (Fig. 4a and Extended Data Fig. 6c)[44]. By contrast, we detected up to 7.6% Norn cells in a dataset derived from E18.5 fetal mouse whole kidneys (Fig. 4a and Extended Data Fig. 6d)[45]. We observed an even higher Norn cell frequency (19.9%) in a dataset of E18.5 fetal mouse enriched for the Foxd1 lineage, previously implicated as progenitors of Epo-producing cells (Fig. 4a and Extended Data Fig. 6e)[25]. In human datasets, we found that 3% and 15.8% of cells were highly correlated with the mouse Norn signature in adult and week 18 fetal kidneys, respectively (Fig. 4a and Extended Data Fig. 6f,g). Comparison of the core Norn signature, as well as Norn-specific receptors and TFs, showed conservation across developmental stages and between species (Fig. 4b and Extended Data Fig. 7a–e). In particular, the Norn core TF circuitry is largely conserved, including *TCF21*, *GATA6* and *TSHZ2*, whereas *CEBPD* is enriched in adult mice and humans but not in development (Fig. 4b). Importantly, Epo transcripts could not be detected in any of these normoxic datasets, including in the Norn cell population.

To derive a comprehensive human Norn profile, we next isolated cells from normoxic nontumor human kidney samples. We carefully calibrated tissue dissociation and tested several candidate surface markers for flow cytometry enrichment (Fig. 4c). We profiled 18,055 single cells and identified cells of the immune, vascular, epithelial and stromal compartment with 7.5% expressing the Norn signature, including *DCN* and *PDGFRA* (Fig. 4d,e and Extended Data Fig. 1e). To evaluate the relevance of these human Norn-like cells, we performed comparative gene set enrichment analysis (GSEA). We observed a transcriptional overlap between human Norn cells and their mouse counterparts, including extracellular matrix (ECM) glycoproteins, collagens, proteoglycans (matrisome) and epithelial to mesenchymal transition (Fig. 4f). We further examined Norn marker genes and found

https://doi.org/10.1038/s41591-023-02314-7

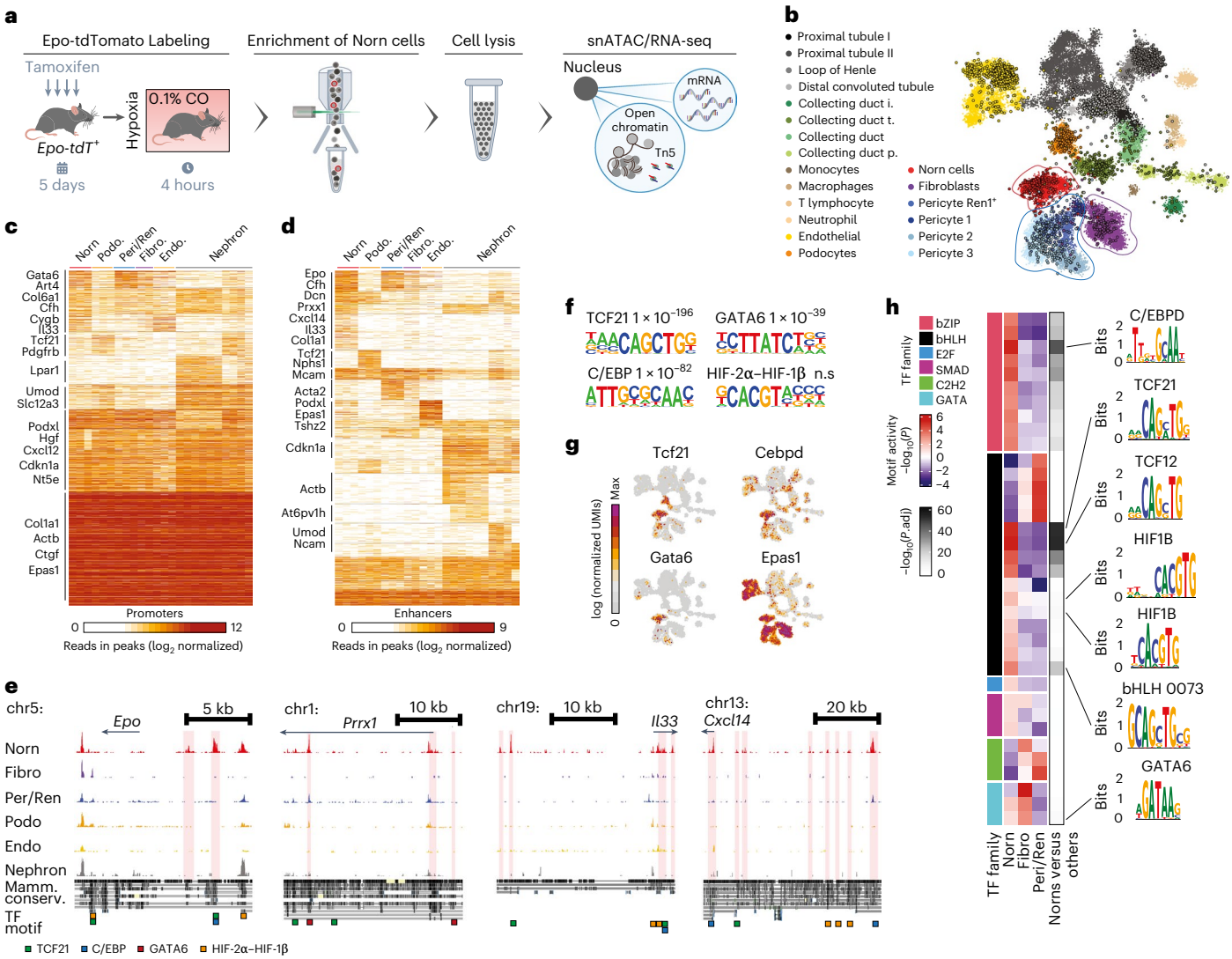


Fig. 3 | ATAC-seq profiling of Norn cells identifies genomewide and *Epo* locus-specific regulatory elements. a, Schematic of the experimental set-up for single nuclear ATAC-seq (snATAC-seq) and RNA-seq. $n = 1$ experiment, kidneys from $n = 5$ mice (2 females, 3 males). **b**, Two-dimensional projection of single nuclear RNA overlaid on the MARS-seq data. Single nuclei are represented as colored circles with black rims. **c**, Heatmap showing 29,021 ATAC peaks in promoters clustered with *k*-means ($n = 10$). Values indicate average read counts within a region. Each column represents a subsample of 50 cells from each population. Endo., endocyte; Fibro., fibroblast; Peri/Ren, pericyte/Ren1; Podo., podocyte. **d**, Heatmap showing 95,181 ATAC peaks in enhancer regions clustered with *k*-means ($n = 20$). Values indicate average read counts within a region. **e**, Pseudo-bulk of ATAC-seq peaks are shown. Norn-specific enhancer regions are shaded in transparent rose. Displayed peaks are at the locus-specific kb range. Scales of peaks are displayed from 0.5 to 100. The mammalian conservation (Mamm. conserv.) track is from the UCSC browser. TF motifs: TCF21, HRE, C/EBPδ and GATA6 are highlighted with colored boxes. **f**, Motif enrichment in all Norn-specific ATAC peaks is calculated using cumulative binomial distributions by HOMER package, correcting for multiple hypotheses using the Benjamini–Hochberg method[59]. Enrichment in $n = 5{,}844$ total target sequences over $n = 360{,}380$ total background sequences. TCF21 ($P = 1 \times 10^{-196}$) is present in 26.1% of Norn peaks. C/EBPδ ($P = 1 \times 10^{-82}$) is present in 13.1% of Norn peaks. GATA6 ($P = 1 \times 10^{-39}$) is present in 12.4% of Norn peaks. The HIF-2α–HIF-1β HRE motif was not significantly enriched. **g**, MARS-seq derived expression of *Tcf21*, *Cebpd*, *Gata6* and *Epas1* show enrichment in Norn cells. **h**, Motif enrichment within specific ATAC peaks in Norn, pericyte and fibroblast cells was derived using the GimmeMotif framework[60]. Heatmap data shown as $\log_{10}$. Left of heatmap: six TF families are shown as basic leucine zipper (bZIP), bHLH, E2F, SMAD, Cys2–His2 (C2H2) zinc fingers and GATA. Grayscale bar indicates FDR adjusted *P* value (two-way nonparametric Mann–Whitney *U*-test) for specificity in Norn cells compared with pericyte and fibroblasts. A selected consensus motif is shown on the right, with the relevant Norn-enriched factor on top of the motif.

that many are conserved across the species, identifying *DCN*, *CXCL12*, *CXCL14*, *HSD11B1*, *TMEM119*, and other genes as enriched in both mouse and human Norn cells compared with their respective fibroblasts (Fig. 4g). Similar to the normoxic mouse Norn cells, we could not detect *EPO* expression in human Norn cells under these conditions.

To define human Norn cells as Epo-producing cells in humans, we generated sections from the kidney of a 55-year-old female who died from CO poisoning and whose blood analysis revealed ~50% CO-bound hemoglobin, similar to the experimental conditions we induced in mice. We applied mRNA-FISH staining for these slices to detect *EPO* and other Norn and stromal marker mRNAs. In line with the mouse Norn data, *EPO* expression in this hypoxic kidney was highly associated with *DCN*, *TCF21*, *CXCL12* and *PDGFRA* expression, suggesting its containment solely within this compartment (Fig. 4h and Extended Data Fig. 7f).

In summary, Norn cells are the main Epo-producing cells in the mouse kidney and are molecularly and functionally conserved from mice to humans.

https://doi.org/10.1038/s41591-023-02314-7

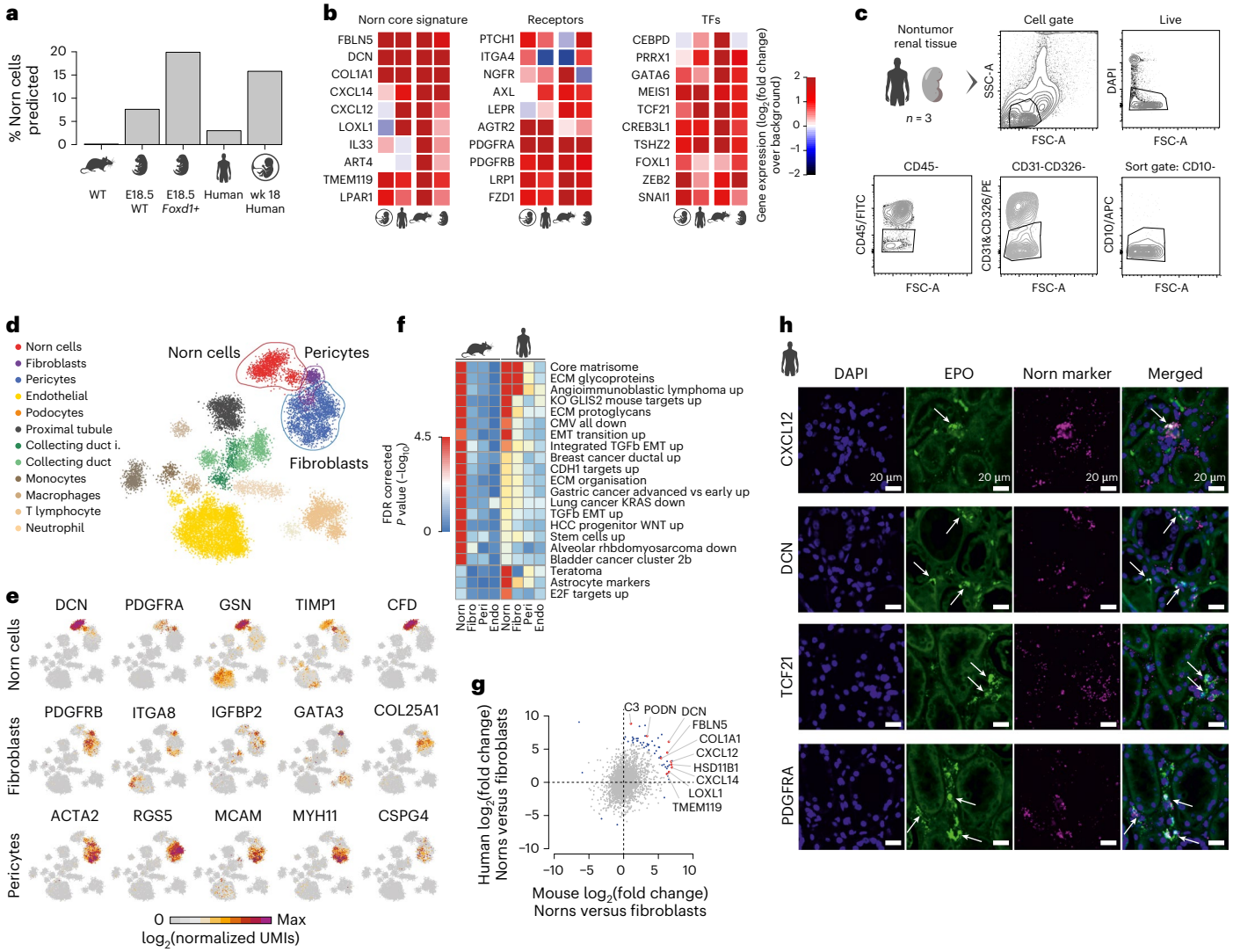


Fig. 4 | Norn cell molecular signature is conserved in human Epo-producing cells. a, Summary of Norn-like signature in kidney scRNA-seq datasets, from left: adult mouse C57BL/6J unenriched[44], fetal C57BL/6J mouse unenriched[45], fetal mouse *Foxd1*$^+$ enriched[46], adult human unenriched[42] and fetal week 18 human unenriched[45]. Values on the y axis indicate the percentage of Norn cells identified. WT, wild-type. **b**, Expression of genes found in core Norn gene signature, surface receptors and TFs. Data are shown as $\log_2$(fold change) in gene expression enrichment over background. **c**, Representative gating strategy to enrich for human Norn-like cells. **d**, Two-dimensional projection of 18,055 cells, grouped together into 71 metacells. Dots represent single cells and colors denote cell-type annotation. $n = 3$ independent experiments, $n = 3$ independent biological samples (2 males, 1 female). **e**, A two-dimensional projection of representative expression ($\log_2$(normalized UMI count)) and distribution of selected marker genes for human Norn-like cells, fibroblasts and pericytes. **f**, GSEA pathway analysis enrichment in mouse and human Norn-like cells. Data are shown as $-\log_{10}$(fold change) in enrichment score using a two-way Kolmogorov–Smirnov statistic test and FDR corrected *P* values with a threshold of 0.25. Rank was generated using fold over mean of cells separately for mouse and human, using in mouse $n = 2{,}703$ Norn cells, $n = 3{,}459$ fibroblasts, $n = 1{,}227$ pericyte Ren1 and $n = 4{,}716$ endothelial cells; and in human $n = 1{,}345$ Norn cells, $n = 548$ fibroblasts, $n = 2{,}473$ pericyte Ren1 and $n = 5{,}777$ endothelial cells. CMV, cytomegalovirus; E2F; EMT, epithelial-mesenchymal transition; HCC, hepatocellular carcinoma; KO, knock out; TGFb, transforming growth factor beta. **g**, Differential expression between mouse and human Norn cells and fibroblasts using two-way Mann–Whitney *U*-test with FDR corrected *P* values. Data are shown as $\log_2$(size-normalized pooled expression) for an equal number of cells. Blue dots denote significant genes, red dots denote selected conserved genes present in both mouse and human Norn cells. **h**, mRNA-FISH of human hypoxic kidney (50% CO-bound hemoglobin). $n = 3$ independent experiments, kidney from $n = 1$ human (female). Representative images of *EPO* together with candidate Norn markers. Scale bar, 20 µm. Nuclei were stained with DAPI.

Discussion

scRNA-seq assays have the extraordinary capacity to map the entire spectrum of cellular diversity within a given tissue. In the past decade, hundreds of millions of cells from a large variety of tissues, organisms and disease states have been profiled using single-cell genomic approaches. However, although effective in defining heterogeneity within cell types, relatively few new mammalian cell types with distinct functional roles have been uncovered so far, with airway ionocytes and thymic tuft cells representing exceptions to the rule[62,63]. Here, we undertook a targeted and functional approach to single-cell analysis, assigning a known biological function (Epo secretion) to an uncharacterized cell type. Similar approaches may be used to better understand other idiosyncratic physiological functions in their native cellular contexts.

Controversy has prevailed about the existence of a specialized renal Epo-producing cell type[24,31]. Our work exclusively assigns Epo production to Norn cells, a specialized, distinct and coherent cell

type that is maintained across species and development, independent of actual Epo production. Our single cell data show that basal Epo expression in Norn cells from normoxic mice is practically below detection levels in conventional count-based scRNA-seq techniques, in line with previous studies that detected only a few Epo-producing cells under normoxic conditions[24,35]. This poses a major challenge to detect Epo RNA in human kidneys, where hypoxia cannot be induced experimentally. This is compounded by the scarcity of hypoxic, noncancerous, human kidneys with high RNA quality, of which no scRNA-seq dataset is currently available. Nevertheless, building on our molecular profile of Norn-associated genes, together with formalin-fixed paraffin-embedded (FFPE) sections of a single hypoxic kidney sample, we could identify the human Norn equivalent and associate it with Epo production under hypoxia. Future studies should include larger cohorts and further functional experiments to validate our findings. Together, these findings pave the way to further study the role of this important cell type in human pathology and evolution.

Intriguingly, although most of the previously used markers to define renal Epo-producing cells, including CD73, COL1A1 and PDGFRβ, are also highly represented in the Norn scRNA-seq dataset, many Norn markers are also expressed to some extent in other kidney lineages. Therefore, the Norn cell population is best defined by the combination of multiple markers, providing an explanation why the Norn population has not been unequivocally recognized previously. Our work goes beyond known stromal markers and implicates Tcf21, Cebpd and Gata6 TFs as candidate master regulators of Norn cellular identity involved in regulating Epo production beyond the well characterized HIF-2α–HIF-1β circuit. Tcf21 has previously been shown to be essential in kidney development, where it is confined to mesenchymal cells of which Norn cells originate[64,65]. The TCF21 motif is the most abundant and enriched in Norn-specific regulatory regions and its target genes make up key Norn marker- and hypoxia-regulated genes. We also noted that the TCF21 motif has high similarities with the HIF-2α–HIF-1β HRE motif. It is interesting to speculate whether Tcf21 primes the Norn and Epo regulatory landscape for activation, allowing the HIF-2α–HIF-1β heterodimer to interact directly or indirectly with the DNA-bound TCF21, to generate a rapid burst of *Epo* and associated genes in response to hypoxia. Therefore, TCF21, together with other Norn-specific TFs, may sculpt the regulatory landscape that confines Epo expression to Norn cells. Future work using single-cell chromatin immunoprecipitation technologies will ultimately test these hypotheses.

Up to now, five members of a new class of HIF-PHD inhibitors, activating HIF TFs to promote endogenous Epo production, have been approved for the treatment of anemia in chronic kidney disease (CKD) patients with ESRD[8]. This suggests that functional Norn cells are present in CKD kidneys and become reactivated to produce Epo through the stabilization of HIF-2α in treated individuals[8,35]. With the Norn pathways and regulatory circuit now available, we can begin to define the mechanisms of Epo perturbations in CKD and how they may be reactivated to physiological levels by HIF-PHD inhibitors.

The discovery and molecular characterization of the insulin-secreting pancreatic β cells revolutionized medical research into diabetes, recently leading to cell-based therapeutic approaches[66]. Similarly, numerous attempts have been undertaken to generate renal Epo-producing cell lines, including from the reporter mouse models used here[24,67]. However, these attempts failed to produce cell lines with stable hypoxia-inducible *Epo* gene expression. The identified Norn cells and Norn TF circuit open new avenues to generate such cell lines, ultimately supporting the development of new drugs correcting the loss of Epo expression in ESRD.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-023-02314-7.

References

1. Orkin, S. H. Diversification of haematopoietic stem cells to specific lineages. *Nat. Rev. Genet.* **1**, 57–64 (2000).
2. Jacobson, L. O., Goldwasser, E., Fried, W. & Plzak, L. Role of the kidney in erythropoiesis. *Nature* **179**, 633–634 (1957).
3. Wenger, R. H. & Kurtz, A. Erythropoietin. *Compr. Physiol.* **1**, 1759–1794 (2011).
4. Lacombe, C. et al. Peritubular cells are the site of erythropoietin synthesis in the murine hypoxic kidney. *J. Clin. Invest.* **81**, 620–623 (1988).
5. Wu, H., Liu, X., Jaenisch, R. & Lodish, H. F. Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell* **83**, 59–67 (1995).
6. *Haemoglobin Concentrations for the Diagnosis of Anaemia and Assessment of Severity* (World Health Organization, 2011); https://apps.who.int/iris/bitstream/handle/10665/85839/WHO_NMH_NHD_MNM_11.1_eng.pdf
7. Prchal, J. T. Polycythemia vera and other primary polycythemias. *Curr. Opin. Hematol.* **12**, 112–116 (2005).
8. Haase, V. H. Hypoxia-inducible factor–prolyl hydroxylase inhibitors in the treatment of anemia of chronic kidney disease. *Kidney Int. Suppl.* **11**, 8–25 (2021).
9. Gangaraju, R. et al. Upregulation of thrombo-inflammatory pathways may contribute to increased thrombotic risk in polycythemia vera and essential thrombocythemia. *Blood* **128**, 3143 (2016).
10. Gordeuk, V. R. et al. Thrombotic risk in congenital erythrocytosis due to up-regulated hypoxia sensing is not associated with elevated hematocrit. *Haematologica* **105**, e87–e90 (2020).
11. Rankin, E. B. et al. Hypoxia-inducible factor-2 (HIF-2) regulates hepatic erythropoietin in vivo. *J. Clin. Invest.* **117**, 1068–1077 (2007).
12. Kapitsinou, P. P. et al. Hepatic HIF-2 regulates erythropoietic responses to hypoxia in renal anemia. *Blood* **116**, 3039–3048 (2010).
13. Paliege, A. et al. Hypoxia-inducible factor-2alpha-expressing interstitial fibroblasts are the only renal cells that express erythropoietin under hypoxia-inducible factor stabilization. *Kidney Int.* **77**, 312–318 (2010).
14. Semenza, G. L., Nejfelt, M. K., Chi, S. M. & Antonarakis, S. E. Hypoxia-inducible nuclear factors bind to an enhancer element located 3′ to the human erythropoietin gene. *Proc. Natl Acad. Sci. USA* **88**, 5680–5684 (1991).
15. Appelhoff, R. J. et al. Differential function of the prolyl hydroxylases PHD1, PHD2, and PHD3 in the regulation of hypoxia-inducible factor. *J. Biol. Chem.* **279**, 38458–38465 (2004).
16. Simonson, T. S. et al. Genetic evidence for high-altitude adaptation in Tibet. *Science* **329**, 72–75 (2010).
17. Yi, X. et al. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **329**, 75–78 (2010).
18. Huerta-Sánchez, E. et al. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* **512**, 194–197 (2014).
19. Witt, K. E. & Huerta-Sánchez, E. Convergent evolution in human and domesticate adaptation to high-altitude environments. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **374**, 20180235 (2019).
20. Moore, L. G., Young, D., McCullough, R. E., Droma, T. & Zamudio, S. Tibetan protection from intrauterine growth restriction (IUGR) and reproductive loss at high altitude. *Am. J. Hum. Biol.* **13**, 635–644 (2001).

21. Beall, C. M. et al. Natural selection on EPAS1 (HIF2α) associated with low hemoglobin concentration in Tibetan highlanders. *Proc. Natl Acad. Sci. USA* **107**, 11459–11464 (2010).
22. Lorenzo, F. R. et al. A genetic mechanism for Tibetan high-altitude adaptation. *Nat. Genet.* **46**, 951–956 (2014).
23. Wiesener, M. S. et al. Widespread hypoxia-inducible expression of HIF-2alpha in distinct cell populations of different organs. *FASEB J.* **17**, 271–273 (2003).
24. Imeri, F. et al. Generation of renal Epo-producing cell lines by conditional gene tagging reveals rapid HIF-2 driven Epo kinetics, cell autonomous feedback regulation, and a telocyte phenotype. *Kidney Int.* **95**, 375–387 (2019).
25. Kobayashi, H. et al. Distinct subpopulations of FOXD1 stroma-derived cells regulate renal erythropoietin. *J. Clin. Invest.* **126**, 1926–1938 (2016).
26. Obara, N. et al. Repression via the GATA box is essential for tissue-specific erythropoietin gene expression. *Blood* **111**, 5223–5232 (2008).
27. Chang, Y.-T. et al. DNA methyltransferase inhibition restores erythropoietin production in fibrotic murine kidneys. *J. Clin. Invest.* **126**, 721–731 (2016).
28. Loya, F., Yang, Y., Lin, H., Goldwasser, E. & Albitar, M. Transgenic mice carrying the erythropoietin gene promoter linked to lacZ express the reporter in proximal convoluted tubule cells after hypoxia. *Blood* **84**, 1831–1836 (1994).
29. Bussolati, B. et al. Renal CD133(+)/CD73(+) progenitors produce erythropoietin under hypoxia and prolyl hydroxylase inhibition. *J. Am. Soc. Nephrol.* **24**, 1234–1241 (2013).
30. Gerl, K. et al. Erythropoietin production by PDGFR-β(+) cells. *Pflug. Arch.* **468**, 1479–1487 (2016).
31. Broeker, K. A. E. et al. Different subpopulations of kidney interstitial cells produce erythropoietin and factors supporting tissue oxygenation in response to hypoxia in vivo. *Kidney Int.* **98**, 918–931 (2020).
32. Marchetti, P., Bugliani, M., De Tata, V., Suleiman, M. & Marselli, L. Pancreatic beta cell identity in humans and the role of type 2 diabetes. *Front. Cell Dev. Biol.* **5**, 55 (2017).
33. Laha, D., Grant, R., Mishra, P. & Nilubol, N. The role of tumor necrosis factor in manipulating the immunological response of tumor microenvironment. *Front. Immunol.* **12**, 656908 (2021).
34. Dahl, S.L. et al. Fate-mapping of erythropoietin-producing cells in mouse models of hypoxaemia and renal tissue remodelling reveals repeated recruitment and persistent functionality. *Acta Physiol. (Oxf.)* **234**, e13768 (2022).
35. Dahl, S. L., Bapst, A. M., Khodo, S. N., Scholz, C. C. & Wenger, R. H. Fount, fate, features, and function of renal erythropoietin-producing cells. *Pflug. Arch.* **474**, 783–797 (2022).
36. Shanks, J. H., Hill, C. M., Lappin, T. R. & Maxwell, A. P. Localization of erythropoietin gene expression in proximal renal tubular cells detected by digoxigenin-labelled oligonucleotide probes. *J. Pathol.* **179**, 283–287 (1996).
37. Yamazaki, S. et al. A mouse model of adult-onset anaemia due to erythropoietin deficiency. *Nat. Commun.* **4**, 1950 (2013).
38. Franke, K. et al. HIF-1α is a protective factor in conditional PHD2-deficient mice suffering from severe HIF-2α-induced excessive erythropoiesis. *Blood* **121**, 1436–1445 (2013).
39. Kobayashi, H., Davidoff, O., Pujari-Palmer, S., Drevin, M. & Haase, V. H. EPO synthesis induced by HIF-PHD inhibition is dependent on myofibroblast transdifferentiation and colocalizes with non-injured nephron segments in murine kidney fibrosis. *Acta Physiol.* **235**, e13826 (2022).
40. Bek-Pedersen, K. *Norns in Old Norse Mythology* (Dunedin Academic Press, 2013).
41. Suresh, S., Rajvanshi, P. K. & Noguchi, C. T. The many facets of erythropoietin physiologic and metabolic response. *Front. Physiol.* **10**, 1534 (2019).
42. Young, M. D. et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* **361**, 594–599 (2018).
43. Park, J. et al. Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**, 758–763 (2018).
44. Ransick, A. et al. Single-cell profiling reveals sex, lineage, and regional diversity in the mouse kidney. *Dev. Cell* **51**, 399–413 (2019).
45. Combes, A. N. et al. Single cell analysis of the developing mouse kidney provides deeper insight into marker gene expression and ligand-receptor crosstalk. *Development* **146**, dev178673 (2019).
46. England, A. R. et al. Identification and characterization of cellular heterogeneity within the developing renal interstitium. *Development* **147**, dev190108 (2020).
47. Hochane, M. et al. Single-cell transcriptomics reveals gene expression dynamics of human fetal kidney development. *PLoS Biol.* **17**, e3000152 (2019).
48. Jaitin, D. A. et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779 (2014).
49. Keren-Shaul, H. et al. MARS-seq2.0: an experimental and analytical pipeline for indexed sorting combined with single-cell RNA sequencing. *Nat. Protoc.* **14**, 1841–1862 (2019).
50. Baran, Y. et al. MetaCell: analysis of single-cell RNA-seq data using K-nn graph partitions. *Genome Biol.* **20**, 206 (2019).
51. Nishiyama, A. & Kim-Mitsuyama, S. New approaches to blockade of the renin–angiotensin–aldosterone system: overview of regulation of the renin–angiotensin–aldosterone system. *J. Pharmacol. Sci.* **113**, 289–291 (2010).
52. Pan, X. et al. Isolation and characterization of renal erythropoietin-producing cells from genetically produced anemia mice. *PLoS ONE* **6**, e25839 (2011).
53. Gimm, T. et al. Hypoxia-inducible protein 2 is a novel lipid droplet protein and a specific target gene of hypoxia-inducible factor-1. *FASEB J.* **24**, 4443–4458 (2010).
54. Mondéjar-Parreño, G. et al. Uncovered contribution of Kv7 channels to pulmonary vascular tone in pulmonary arterial hypertension. *Hypertension* **76**, 1134–1146 (2020).
55. Schödel, J. et al. High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. *Blood* **117**, e207–e217 (2011).
56. Orlando, I. M. C. et al. Distal and proximal hypoxia response elements cooperate to regulate organ-specific erythropoietin gene expression. *Haematologica* **105**, 2774–2784 (2020).
57. Semenza, G. L., Koury, S. T., Nejfelt, M. K., Gearhart, J. D. & Antonarakis, S. E. Cell-type-specific and hypoxia-inducible expression of the human erythropoietin gene in transgenic mice. *Proc. Natl Acad. Sci. USA* **88**, 8725–8729 (1991).
58. Hirano, I. et al. Renal anemia model mouse established by transgenic rescue with an erythropoietin gene lacking kidney-specific regulatory elements. *Mol. Cell. Biol.* **37**, e00451-16 (2017).
59. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
60. Bruse, N. & van Heeringen, S. J. GimmeMotifs: an analysis framework for transcription factor motif analysis. Preprint at *bioRxiv* https://doi.org/10.1101/474403 (2018).
61. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
62. Bornstein, C. et al. Single-cell mapping of the thymic stroma identifies IL-25-producing tuft epithelial cells. *Nature* **559**, 622–626 (2018).
63. Plasschaert, L. W. et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* **560**, 377–381 (2018).

64. Quaggin, S. E. et al. The basic-helix-loop-helix protein pod1 is critically important for kidney and lung organogenesis. *Development* **126**, 5771–5783 (1999).
65. Ide, S. et al. Transcription factor 21 is required for branching morphogenesis and regulates the Gdnf-axis in kidney development. *J. Am. Soc. Nephrol.* **29**, 2795–2808 (2018).
66. Helman, A. & Melton, D. A. A stem cell approach to cure type 1 diabetes. *Cold Spring Harb. Perspect. Biol.* **13**, a035741 (2021).
67. Bapst, A. M., Dahl, S. L., Knöpfel, T. & Wenger, R. H. Cre-mediated, loxP independent sequential recombination of a tripartite transcriptional stop cassette allows for partial read-through transcription. *Biochim. Biophys. Acta Gene Regul. Mech.* **1863**, 194568 (2020).

Methods

All human samples used for scRNA-seq complied with the Helsinki protocols 6370-19-SMC (Sheba-Telhashomer Medical Center), 0417-20-TLV (Ichilov Hospital) and 0297-22-HMO (Hadassah Medical Center) with local institutional review board approval from Weizmann Institute of Science (1568-1). Informed consent was obtained from all individuals. The collection and use of FFPE sections of human kidney samples by the Institute for Forensic Medicine, University Hospital Essen, Germany, was approved by the ethical commission of the Medical Faculty of the University Duisburg-Essen 2015 (15-6345-BO). Because a forensic sample was used and to preserve complete anonymization imposed by the ethical permit, we did not obtain consent from the individual or the family.

Data reporting

No statistical methods were used to predetermine sample sizes. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. We excluded the following data:

(1) For mapping of scRNA-seq reads, reads with multiple mapping positions were excluded.
(2) For 10x Genomics scRNA-seq, suspected doublets and cells with fewer than 200 unique molecular identifiers (UMIs) and more than 25% mitochondrial content were excluded: nFeature_RNA > 200 & nFeature_RNA < 5000 & percent.mt <25.
(3) For 10x Genomics scATAC-seq, the following exclusion was done: nFeature_RNA > 1000 & nFeature_RNA < 30000.
(4) For MARS-seq, suspected doublets and cells with fewer than 200 UMIs or more than 40% mitochondrial RNA content were excluded.

Mice

Wild-type C57BL/6 (B6) mice were purchased from Harlan Laboratories. Epo-CreERT2-tdTomato mice (Epo-tdTomato) have previously been described in detail[24]. Briefly, the mouse model enables permanent, conditional and exclusive labeling of Norn cells that expressed *Epo* during a 4-h hypoxia treatment following a 5-day course of tamoxifen treatment. The transgenic large BAC reporter construct includes Epo regulatory elements and neighboring genes. All animals were bred and housed in standard ventilated cages on a 12-h light and 12-h dark cycle and ambient temperatures and humidity according to the guidelines of Weizmann Institute of Science. All animals were housed in groups with a maximum of five mice per cage and had ad libitum access to food (Harlan Teklad, catalog no. 2918) and water. All experimental procedures were approved by the Institutional Animal Care and Use Committee, application numbers 01810220-2 and 11280219-2. For experiments, 10–18-week-old male and female mice were used. Because there is no evidence of sex-specific differences in Epo regulation, we included male and female mice and analyzed them together using MetaCell. Cells were pooled from $n = 22$ independent experiments and $n = 52$ mice (33 male, 19 female). In the quantitative PCR (qPCR) timeline, there were $n = 4$ technical replicates and n = 1 mouse (male) per time point. For qPCR detection of Epo in hypoxic mouse tissue there were $n = 4$ technical replicates and $n = 1$ mouse (male).

Human tissue

Nontumor kidney tissue was obtained from two male patients (68 and 83 years old) and one female patient (76 years old) undergoing partial or radical nephrectomy after signing an informed consent form. Because there is no evidence that there are sex-specific differences in Epo regulation, we included male and female patients and analyzed together using MetaCell.

The FFPE kidney sample used in this study was derived from a 55-year-old female who died in a fire (CO-Hb content of 50%) and who underwent a forensic autopsy.

Mouse hypoxia exposure and kidney dissociation

Mice were placed in an air-tight chamber with synthetic air containing 0.1% CO continuously flowing through the chamber for 4 h. In the time course experiment (Fig. 1b), mice were exposed to 0.1% CO for 2, 4, 6 or 12 h. Mice were killed and immediately perfused with cold PBS buffer. In the downstream process, we performed the tissue dissociation under normoxic conditions as follows. After isolation of the kidneys, the renal capsule was removed and minced with razor blades into 1-mm pieces. Each minced kidney was placed in 1 ml of DMEM medium containing collagenase IV (Worthington Biochemical, catalog no. LS004188, final concentration 0.25 mg ml^{-1}), neutral protease (Worthington Biochemical, catalog no. LS02106, final concentration 0.25 mg ml^{-1}) and DNase I (Roche, catalog no. 10104159001, final concentration 2,000 U ml^{-1}), and incubated in a thermomixer at 37 °C, 300 r.p.m. for 25–30 min. During this time, the suspension was passed ten times through a 18G syringe and eight times through a 21G syringe. The suspension was subsequently applied to tissue strainers (40 µm mesh size), 20 ml of MACS buffer (PBS containing 2% FBS and 5 mM EDTA pH 8.0) was added, and the cells centrifuged for 5 min at 4 °C, 320*g*. Fc-block (BD Biosciences, catalog no. BD553141, 1:200) was added for 17 min, followed by red blood cell lysis for 3 min on ice. MACS buffer was added to stop the lysis and the cells were centrifuged for 10 min at 4 °C, 300*g* and stained for FACS.

Human kidney sample collection and dissection

Nontumor kidney portions were collected in PBS on ice. Kidney samples were minced to 2–4 mm and dissociated according to the manufacturer's protocol for the Multi Tissue Dissociation Kit 1 (Miltenyi Biotech, catalog no. 130-110-201). Red blood cells were removed by incubating the sample with 2 ml of red blood cell lysis buffer for 5 min on ice. The sample was enriched to live cells using a Dead Cell Removal Kit (Miltenyi Biotech, catalog no. 130-090-101), centrifuged at 300*g* for 10 min, resuspended in MACS buffer (PBS containing 2% FBS and 5 mM EDTA pH 8.0) and stained for FACS.

Flow cytometry and sorting

Mouse cells were stained in MACS buffer using CD31-APC (Biolegend, catalog no. 102409, Clone 390; 1:400), CD326-FITC (Biolegend, catalog no. 118207, Clone G8.8; 1:600), CD45-FITC (eBioscience TY/11.8, catalog no. 11-0451-8, Clone 30-F11; 1:400) and CD73-APC-Fire750 (Biolegend, catalog no. 127221, 1:200). Following staining, cells were washed and resuspended in MACS buffer and sorted on FACS Symphony S6, ARIA III and FUSION cell sorters (BD Biosciences). Single cells were sorted into 384-well cell capture plates containing 2 µl of lysis solution and barcoded poly(T) reverse transcription (RT) primers for scRNA-seq as described previously[48]. Barcoded single-cell capture plates were prepared with a Bravo automated liquid handling platform. Four empty wells were kept in each 384-well plate as a no-cell control during data analysis. Immediately after sorting, each plate was spun down to ensure cell immersion into the lysis solution, snap frozen on dry ice and stored at −80 °C until further processing.

Human cells where stained in MACS buffer using CD326-PE (Biolegend, catalog no. 324206, Clone 9C4; 1:100), CD45-FITC (Biolegend, catalog no. 304006, Clone HI30; 1:100), CD31-PE (Biolegend, catalog no. 303105, Clone WM59; 1:100), CD10-APC (Biolegend, catalog no. 312210, Clone HI10a; 1:100) and CD235a-BV605 (Biolegend, catalog no. 740409, Clone GA-R2; 1:100). Following staining, cells were washed and resuspended in MACS buffer and bulk sorted into a cooled 1.7-ml Eppendorf tube containing 100 µl of MACS buffer using the FACS Symphony S6. For FACS analysis, we used the FACS Diva 7 and FlowJo v.10.4.2.

Massively parallel scRNA library preparation

Single-cell transcriptome libraries were prepared as previously described[48]. Briefly, mRNA from cells sorted into MARS-seq capture plates was barcoded, converted to complementary DNA and pooled

using an automated pipeline. The pooled sample was then linearly amplified by T7 polymerase-mediated in vitro transcription and the resulting mRNA was fragmented and converted into a sequencing-ready library by tagging the samples with pool barcodes and Illumina adapter sequences during ligation, followed by RT and PCR. Library quality and concentration were assessed and MARS-seq libraries were sequenced using an Illumina NextSeq 500 and Novaseq 6000.

Low-level processing and filtering

All RNA-seq libraries were sequenced at a mean depth of 17,500 reads per single cell. Sequences were mapped to the mouse genome (mm10), with the *Epo* 3′ untranslated region extended by 1 kb (chr5:137,481,020-137,485,852), demultiplexed and filtered as previously described[48], and a set of UMI that define distinct transcripts in single cells were extracted for further processing. We estimated the level of spurious UMIs in the data using MetaCell statistics on empty MARS-seq wells (mean noise 1.3%; Extended Data Fig. 4). Reads were mapped using HISAT (v.0.1.6)[68]; reads with multiple mapping positions were excluded. Reads were associated with genes if they were mapped to an exon, using the University of California Santa Cruz (UCSC) genome browser for reference. Exons of different genes that shared a genomic position on the same strand were considered a single gene with a concatenated gene symbol. Cells with fewer than 200 UMIs were discarded from the analysis. After filtering, cells contained a mean of 1,500 unique genes per cell. All downstream analysis was performed in R v.3.5.3.

Data processing and clustering using MetaCell

The MetaCell pipeline[69] v.0.3.41 was used to derive informative genes and compute cell-to-cell similarity, to compute Knn graph covers and derive the distribution of RNA in cohesive groups of cells (or metacells), and to derive strongly separated clusters using bootstrap analysis and computation of graph covers on resampled data. We used bootstrapping to derive robust clustering (500 iterations; resampling 70% of the cells in each iteration and clustering the co-cluster matrix with minimal cluster size set to 15). No further filtering or cluster splitting was performed on the metacells.

Annotating MARS-seq subsets using label transfer

We manually annotated our scRNA-seq dataset using published datasets. We relied on a label transfer approach, utilizing gene lists from Park et al.[43] that are characteristic of specific renal cell types. We summarized the $\log_2$(fold change) of each gene list in each metacell and derived a *z*-score. The *z*-score was adjusted to a minimum number of genes expressed. Annotation of the stromal compartment was done on the basis of known stromal markers[31] to identify distinct clusters of fibroblast (*Pdgfra* and *Nt5e*/CD73 antigen positive) and pericyte (*Acta2* and *Rgs5*) origins.

Deriving scores and mapping Norn cells in existing single-cell kidney datasets

To find the most differential and exclusive genes defining each cell population, we used $\log_2$(fold change) in gene expression between each individual cell population versus the rest of the populations. We derived the top 50 signature genes for each population (Supplementary Table 2) and calculated an associated gene expression value summing $\log_2$ values over the genes. We used this as input for the R-package 'cutpointr' with the parameter 'method' set to 'maximize_metric' and 'metric' set to accuracy constraints ('acc_constrain') to find the best cutoff threshold for classification. We derived scores over scRNA-seq datasets without previous knowledge of each cell's annotation. We calculated the sum of log-transformed size-normalized expression of all gene signatures. For each cell and signature, the fold change of its score over its predetermined cutoff threshold value was calculated. A cell was assigned a specific cell identity by choosing the score with highest fold increase over its cutoff value. Cells whose scores were all below cutoff values, were left unassigned (NA). To translate the gene signatures to human gene annotations, we used a previously described lift-over method[70].

Single-cell multiome analysis using the Chromium 10x Genomics platform

Pooled cells from $n = 5$ mice (3 males, 2 females) exposed to hypoxia (0.1% CO, 4 h) were bulk sorted on FACS Symphony S6, ARIA III and FUSION cell sorters (BD Biosciences) using a 130 nozzle and a flow rate of 2 (~2000 cells/second) for 3 h, and subsequently pooled. Some 75,000 cells were obtained and processed according to the 10x Genomics Chromium Single-Cell Multiome ATAC + Gene Expression low cell number protocol (CG000365). Lysis was performed for 7 min using diluted (1:10) lysis buffer to obtain intact nuclei.

Single-cell ATAC and RNA-seq libraries were prepared using the Chromium Single-Cell Multiome ATAC + Gene Expression platform (10x Genomics). Nuclei were prepared and counted to ensure quality and concentration. Nuclei were then transposed according to the manufacturer's protocol. The transposed nuclei suspension was loaded onto a Next GEM Chip J targeting 5,000 nuclei and run on a Chromium Controller instrument to generate gel beads-in-emulsion (GEM) emulsion (10x Genomics). Single-cell gene expression libraries, as well as single-cell ATAC-seq libraries, were generated according to the manufacturer's protocol using the Chromium Next GEM Single-Cell Multiome ATAC + Gene Expression Kit. Final libraries were quantified using a Next Library Quant Kit for Illumina (New England Biolabs) and high-sensitivity D1000 TapeStation (Agilent). Each library was sequenced separately on a NovaSeq 6000 instrument using an SP-100 cycles reagent kit (Illumina), targeting 25,000 reads per nuclei for ATAC-seq and a minimum of 20,000 reads per nuclei for gene expression.

Processing and analysis of 10x Genomics Chromium Single-cell Multiome ATAC + Gene Expression

Low-level read processing and alignment was performed with the Cell ranger ARC/2.0.0 tool, with the *Epo* 3′ untranslated region extended by 1 kb (mm10 chr5:137,481,020-137,485,852). Data from the reference atlas were integrated using label transfer from our reference model to the multiome RNA-seq data. In short, we used a Knn-classifier ($k = 50$) with Pearson's correlation as a similarity metric over the normalized gene features defined for the reference model. The distribution of cluster memberships over these *k*-neighbors was used to associate the new cell with a reference metacell (by majority voting) and to project the cell in two dimensions by weighted average of the linked reference clusters' mapped *x* and *y* coordinates. HOMER package[59] (v.4.11) was used to identify ATAC peaks pooled over all the ATAC-seq reads (command findPeaks -style factor). To identify differential peaks in promoter and enhancer chromatin regions, we divided each cluster into bins of 50 cells to account for variation in cluster size and pooled reads from each replicate as pseudo-bulk.

Figure 3c,d depicts pooled peak intensity across pseudo-bulk triplicates. Peaks within promoter (−1 kb to +100 bp around TSS, Fig. 3c) and enhancer (−50 kb to +50 kb around TSS barring promoter region; Fig. 3d) regions were assigned to their nearest gene.

We used DESEQ2 (v.1.30.1) to test for statistically significant differential peaks for each triplicate. DESEQ2 differential peaks upregulated in the Norn population were tested for enriched TF-binding motifs using motif finding algorithm findMotifGenome in the HOMER package[59].

For GimmeMotifs TF motif enrichment analysis, we used the unique differential peaks of Norn cells (4,424), fibroblasts (239) and pericytes (2199) selected from DESEQ2 for downstream motif enrichment analysis. Because the flanking sequences of known motifs are highly variable across various cell types, we used a consensus database of known TFs (gimme.vertebrate. v.5.0) as implemented in

the GimmeMotifs package (v.0.17.2)[60]. The maelstrom function was applied to calculate the differential enrichment of motifs between the different cell types with default parameters (all differential peaks were reduced or extended to 200 bp around the peak summit). Motif activity is defined as the positive or negative enrichment value of $-\log_{10}(P)$ for high or low rankings of the motif occurrence. We also applied a two-way nonparametric Mann–Whitney *U*-test with the null hypothesis of equal distribution in the motif log-odds ratio between two groups and we used the Benjamini–Hochberg method to calculate the adjusted *P* value. For heatmap visualization of the motifs, selected TFs expressed in more than 15% of the hypoxic Norn cells (from our MARS-seq data) are shown in Fig. 3g.

Gene Expression Single-Cell 10x Genomics

Human kidney cells were bulk sorted for 4,6-diamidino-2-phenylindole (DAPI), CD45, CD326, CD31 and CD10 cells using FACS Symphony S6 (BD Biosciences) with a 100 nozzle and flow rate of 1 (~1500 cells/second). Targeting for 12,000 cells per single-cell library, each sample was sorted for 50,000 to 100,000 cells. Cells were processed according to 10x Genomics Chromium Single-Cell 3′ v.3.1 Dual Index kit. Single-cell gene expression libraries were generated according to the manufacturer's protocol. Final libraries were quantified using Next Library Quant Kit for Illumina (New England Biolabs) and high-sensitivity D1000 TapeStation (Agilent). Each library was sequenced separately on a NovaSeq 6000 instrument using an SP-100 cycles reagent kit (Illumina), targeting 20,000 reads per cell.

Processing and analysis of 10x Genomics Chromium scRNA-seq

Samples were sequenced at 20,000 reads per cell. Samples were demultiplexed using cellranger v.6.1.2 and aligned to human reference genome refdata-cellranger-GRCh38-1.2.0 with the parameter 'include-introns'. Cells with fewer than 500 UMIs and a high mitochondrial percentage were filtered out. Data integration of the reference atlas was done using label transfer from our reference model. In short, we used a Knn-classifier ($k = 50$) with Pearson's correlation as a similarity metric over the normalized gene features defined for the reference model. The distribution of cluster memberships over these *k*-neighbors was used to associate the new cell with a reference metacell (by majority voting) and to project the cell in two dimensions by weighted average of the linked reference clusters' mapped *x* and *y* coordinates. Human genome (Hg38) was used as a reference genome.

Gene tracks and visualization

All gene tracks were visualized as bigWig files of the combined replicates normalized to 100,000 reads in UCSC browser.

Differential gene expression analysis

Differential gene expression analysis was performed on UMIs divided by the median UMI count using a Mann–Whitney *U*-test with FDR correction and at least onefold change difference between groups. For differential gene expression analysis in Fig. 4g, we first identified the shared conserved genes between species based on ensemble v95 Mouse and Human. Second, we grouped the genes between the species as the delta log of mean mouse expression versus the delta log of mean human expression and subtracted the groups.

GSEA analysis

We used GSEA v.4.2.3 for Linux. A sample ranked list was created using the fold change over the mean table. GSEA analysis was performed with the following parameters. Gene database: c2.all.v2022.1.Hs.symbols.gmt; number of permutation: 1,000; collapse: Gene Symbol; chip platform: Human_RefSeq_Accession_MSigDB.v2022.1.Hs.chip, chip or Mouse_Gene_Symbol_Remapping_Human_Orthologs_MSigDB.v2022.1.Hs.chip.

Quantitative PCR

mRNA was isolated using TRIzol and a micro RNA easy kit (Qiagen, catalog no. 217004), and reverse transcribed to cDNA using an AffinityScript RT kit (Agilent, catalog no. 600109). qPCR was performed using SYBR Green I Master Mix on a LightCycler480 SYBR Green I Master Mix (Roche). The same sample was measured in quadruplicate and normalized to *Actb*. One kidney from one mouse was used per technical quadruplicate. The *Actb* qPCR primer sequences were Actb_F GGAGGGGGTTGAGGTGTT and Actb_R TGTGCACTTTTAT TGGTCTCAAG. The *Epo* qPCR primer sequences were Epo_F TTG GTGATTTCAGCTGTTGC and Epo_R GGGTCAAGGAGCCATAGACA. Raw cycle threshold values and calculations are available in Supplementary Table 6.

mRNA-FISH

For mouse kidney, cryosections were pretreated according to the manufacturer's recommendations (RNAscope Multiplex Fluorescent v2 Assay; Advanced Cell Diagnostics) and as described previously[24,34]. The Epo probe (catalog no. 315501) consisted of 12 double Z probe pairs, targeting the region between 39 and 685 of mouse *Epo* mRNA. The Cfh probe (catalog no. 403671) consisted of 20 double Z probe pairs, targeting the region between 2,131 and 3,109 of mouse *Cfh* mRNA. The probes were hybridized for 2 h at 40 °C in a HybEZ oven (Advanced Cell Diagnostics), followed by signal amplification and signal detection using Opal 570 or Opal 650 fluorescent reagents (Akoya Biosciences). Sections were washed, counterstained with DAPI, mounted using ProLong Gold antifade mountant (Thermo Fisher Scientific, catalog no. P36930) and fluorescent signals were recorded using an Axio Scan. Z1a slide scanner (Zeiss).

For human kidney, FFPE sections were pretreated according to the manufacturer's recommendations (RNAscope Multiplex Fluorescent v2 Assay; Advanced Cell Diagnostics). The probes (Supplementary Table 7) were hybridized for 2 h at 40 °C in a HybEZ oven (Advanced Cell Diagnostics), followed by signal amplification and signal detection using Opal 570 or Opal 650 fluorescent reagents (Akoya Biosciences). Sections were washed, counterstained with DAPI, mounted using ProLong Gold antifade mountant (Thermo Fisher Scientific, catalog no. P36930) and fluorescent signals were recorded using an Axio Scan. Z1a slide scanner (Zeiss).

Imaging and image analysis

For FISH analysis, whole kidney slice images were converted to 8-bit grayscale images and whole slice analysis was performed with MATLAB R2018a (The Mathworks). Gray scale thresholds were arbitrarily selected and maintained throughout the experiment. After binarization, FISH-positive pixels and their colocalization were calculated. Image analyses were performed in a blinded fashion. Image quantifications are available in Supplementary Table 7.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Ethics statement

All mouse experimental procedures complied with relevant ethical protocols and were reviewed and approved by the Institutional Animal Care and Use Committee, application numbers 01810220-2 and 11280219-2, at Weizmann Institute of Science.

Data availability

Mouse single-cell RNA-seq and single-cell ATAC-seq data, and human single-cell RNA-seq data that support the findings of this study, were deposited in the Gene Expression Omnibus (GEO) under accession code GSE193321. Previously published scRNA-seq data that were re-analyzed here are available under the accession codes GSE114530 (ref. 47),

GSE129798 (ref. 44), EGAS00001002325GSE155794 (ref. 46) and GSE107585 (ref. 43), as well as from the European Genome-phenome Archive (EGA) under study IDs EGAS00001002325 (ref. 42). We used mouse genome mm10 and human genome Hg38 as reference genomes. ATAC-seq tracks are accessible on UCSC browser https://genome.ucsc.edu/s/bjoert/Mouse%20hypoxia%20kidney%20bulk%20scATAC%2Dseq%20from%20single%20nuclei.

Code availability

Metacell source code can be found at https://github.com/tanaylab/metacell. Source code to identify Norn cells in human single cell atlas can be found at https://github.com/AmitLab/Kidney-Norn-cells-identification. Source code used for transcription factor motif analysis can be found at https://github.com/vanheeringen-lab/gimmemotifs.

References

68. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
69. Giladi, A. et al. Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis. *Nat. Cell Biol.* **20**, 836–846 (2018).
70. Geirsdottir, L. et al. Cross-species single-cell analysis reveals divergence of the primate microglia program. *Cell* **179**, 1609–1622 (2019).
71. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).

Acknowledgements

We thank members of the Amit laboratory for critical discussions. We thank Y. Kuperman, L. Adler and S. Viukov for assisting with hypoxia experiments. We thank K. Pozyuchenko, G.H. Siloni, C. Padberg and T. Bajanowski for assisting with human kidney samples, and P. Spielmann and T. Knöpfel for technical support. We thank T.-M. Salame, E. Hagai and E. Kopitman for assisting with flow cytometry. This study was funded by European Union European Research Council Advanced (grant no. 101055341-TROJAN-Cell, I.A.); Deutsche Forschungsgemeinschaft (Project-ID 259373024 – TRR 167, I.A.), Israel Science Foundation (ISF; grant no. 1944/22, I.A.), the ISF Israel Precision Medicine Program (607/20, I.A.), Human Frontier Science Program (long-term postdoctoral fellow LT 000230/2019, B.K.K.), European Molecular Biology Organization (postdoctoral fellowship, ALTF 112-2022, A.G.), Villum Fonden (Young Investigator award project no. 00025300, F.R.) and by the Swiss National Science Foundation (grant no. 310030_184813 R.H.W.).

Author contributions

B.K.K. and I.A. conceived the project and designed the experiments. B.K.K. performed the experiments. A.G. and C.G. contributed to experimental and project design. L.G. contributed to experimental and project design and assisted with FACS experiments. A.M.B., Y.K., S.L.D., M.Z., O.B. and S.S.-L., assisted with experiments. A.G.-S., S.D., V.Y., J.F., S.W., P.P.M. and B.R. assisted with human kidney specimens. S.H. and C.G. contributed to the human experimental part. A.G., E.D. and K.X. analyzed the data. R.A. and S.-Y.W. contributed to the analysis. A.M.B., S.S.-L. and O.M.L. performed mRNA-FISH experiments and image analysis. B.L. and A.G. assisted with image analysis. F.R., E.W., J.T.P., M.S. and T.S.K. assisted with analysis. H.K.-S., F.S., T.S.P. and M.K. assisted the 10x multiome experiment. B.K.K., A.G. and I.A. wrote the manuscript.

Competing interests

All authors declare no competing interests.

Additional information

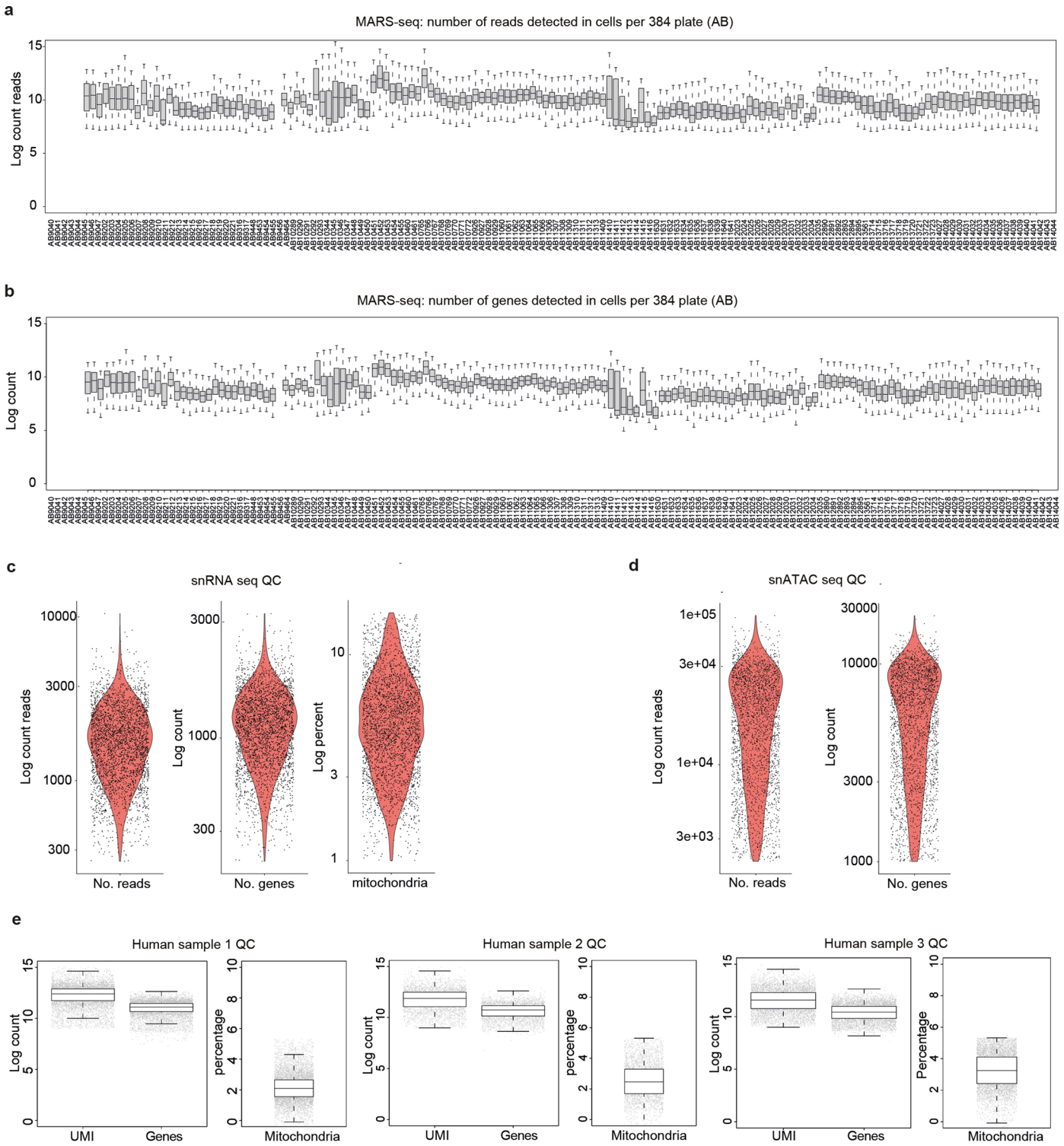
Extended data is available for this paper at https://doi.org/10.1038/s41591-023-02314-7.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-023-02314-7.

Correspondence and requests for materials should be addressed to Bjørt K. Kragesteen or Ido Amit.

Peer review information *Nature Medicine* thanks Katalin Susztak, Tammie Bishop and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Anna Maria Ranzoni, in collaboration with the *Nature Medicine* team.

https://doi.org/10.1038/s41591-023-02314-7

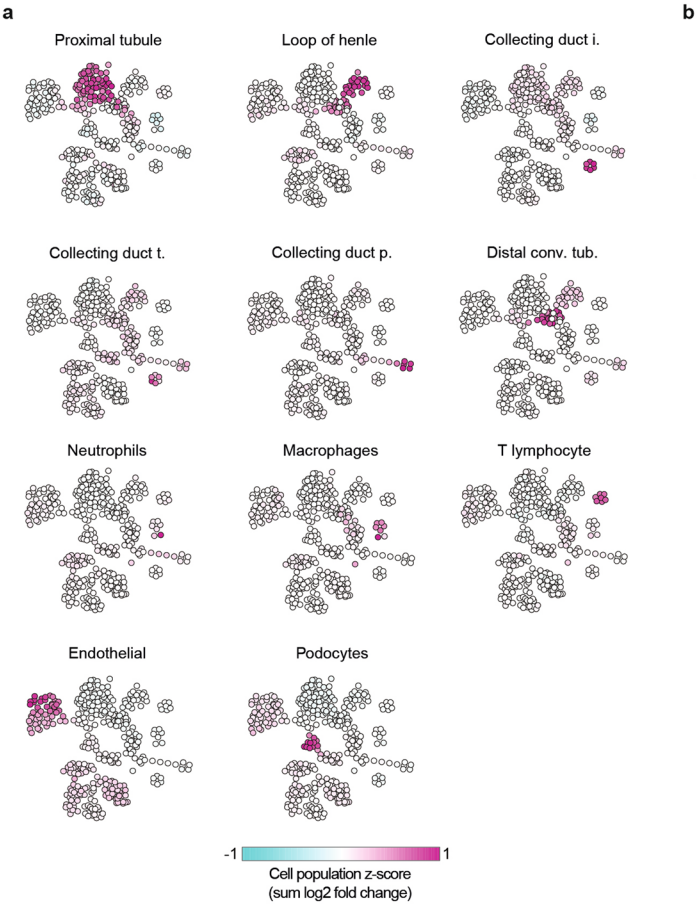


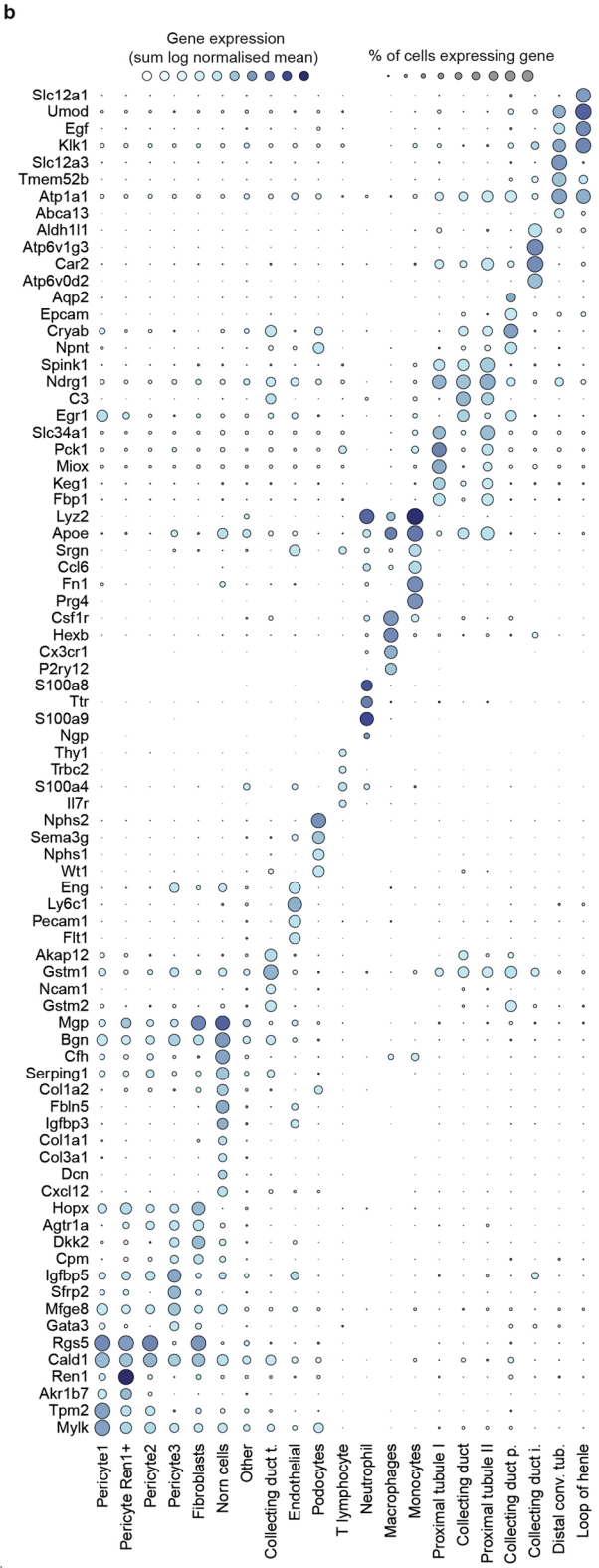
Extended Data Fig. 1 | See next page for caption.

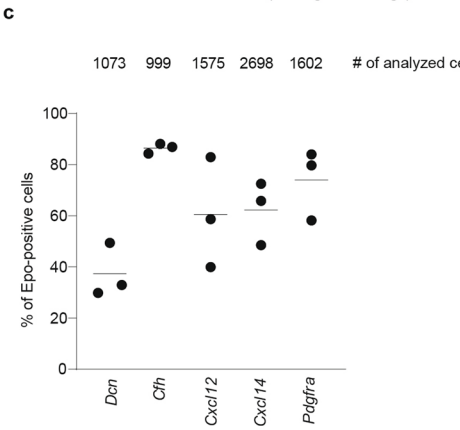
https://doi.org/10.1038/s41591-023-02314-7

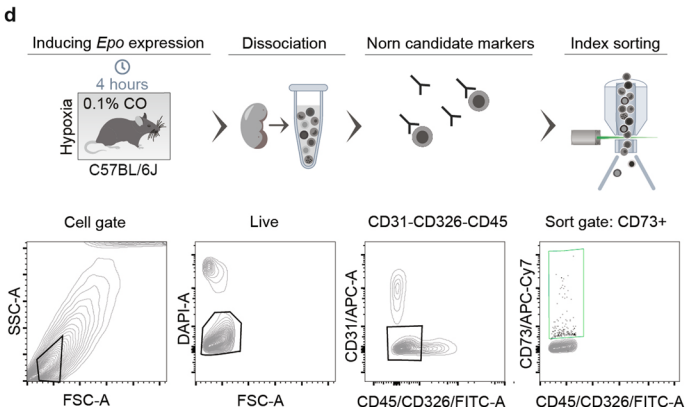
Extended Data Fig. 1 | QC of single cell data. a, **b**, QC for mouse renal MARS-seq showing number of reads and genes per plate, respectively. Box plots presents the third quartile (top of the box), median (centre lines) and first quartile (bottom of the box) of measurements. The whiskers represent 1.5 times the interquartile range from the top or bottom of the box. From n = 22 independent experiments, kidneys from n = 52 mice (33 males, 19 females), cells were sorted into 384 well plates. Nu,ber of QC cells per plate are as follows: AB9040 = 103, AB9041 = 122, AB9042 = 103, AB9043 = 107, AB9044 = 109, AB9045 = 124, AB9046 = 180, AB9047 = 177, AB9202 = 132, AB9203 = 12, AB9204 = 101, AB9205 = 28, AB9206 = 11, AB9207 = 16, AB9208 = 19, AB9209 = 182, AB9210 = 180, AB9211 = 203, AB9212 = 168, AB9213 = 211, AB9214 = 253, AB9215 = 249, AB9216 = 196, AB9217 = 198, AB9218 = 214, AB9219 = 182, AB9220 = 197, AB9221 = 174, AB9316 = 118, AB9317 = 116, AB9453 = 365, AB9454 = 298, AB9455 = 361, AB9456 = 361, AB9464 = 176, AB10289 = 12, AB10290 = 21, AB10291 = 91, AB10292 = 34, AB10293 = 28, AB10344 = 279, AB10345 = 299, AB10346 = 263, AB10347 = 259, AB10448 = 127, AB10449 = 35, AB10450 = 103, AB10451 = 327, AB10452 = 321, AB10453 = 342, AB10454 = 365, AB10455 = 361, AB10460 = 64, AB10461 = 291, AB10765 = 339, AB10766 = 318, AB10767 = 328, AB10768 = 277, AB10769 = 319, AB10770 = 297, AB10771 = 329, AB10772 = 301, AB10926 = 369, AB10927 = 362, AB10928 = 361, AB10929 = 357, AB11060 = 360, AB11061 = 355, AB11062 = 365, AB11063 = 364, AB11064 = 352, AB11065 = 330, AB11066 = 360, AB11306 = 356, AB11307 = 344, AB11308 = 335, AB11309 = 329, AB11310 = 320, AB11311 = 324, AB11312 = 353, AB11313 = 333, AB11409 = 16, AB11410 = 11, AB11411 = 26, AB11412 = 20, AB11413 = 43, AB11414 = 17, AB11415 = 67, AB11416 = 137, AB11630 = 295, AB11631 = 297, AB11632 = 326, AB11633 = 318, AB11634 = 285, AB11635 = 241, AB11636 = 293, AB11637 = 282, AB11638 = 279, AB11639 = 278, AB11640 = 269, AB11641 = 241, AB12023 = 260, AB12024 = 234, AB12025 = 333, AB12026 = 268, AB12027 = 185, AB12028 = 164, AB12029 = 335, AB12030 = 306, AB12031 = 315, AB12032 = 150, AB12033 = 50, AB12034 = 55, AB12035 = 9, AB12890 = 299, AB12891 = 292, AB12892 = 320, AB12893 = 311, AB12894 = 338, AB12895 = 336, AB13561 = 246, AB13714 = 323, AB13715 = 315, AB13716 = 288, AB13717 = 280, AB13718 = 327, AB13719 = 323, AB13720 = 260, AB13721 = 256, AB13722 = 257, AB13723 = 307, AB14027 = 322, AB14028 = 289, AB14029 = 345, AB14030 = 325, AB14031 = 333, AB14032 = 335, AB14033 = 302, AB14034 = 328, AB14035 = 342, AB14036 = 329, AB14037 = 338, AB14038 = 334, AB14039 = 300, AB14040 = 287, AB14041 = 204, AB14042 = 231, AB14043 = 177, AB14044 = 203. **c**, QC for 10x multiome mRNA showing number of reads, genes per plate and percentage mitochondria. N = 3,861 cells pooled from 5 mice (3 males, 2 females) exposed to hypoxia 0.1% CO, 4 hours. Each cells has a mean read of 14,537. **d**, QC for mouse renal 10x multiome ATAC data showing number of reads and genes per plate. N = 3,306 cells pooled from 5 mice (3 males, 2 females) exposed to hypoxia 0.1% CO, 4 hours, in one experiment. Per cell we obtained 10,666 ATAC median high-quality fragments. **e**, QC for human renal RNA data derived from 10x scRNAseq showing number of reads, genes, and percentage mitochondria per cell, per patient. Box plots presents the third quartile (top of the box), median (centre lines) and first quartile (bottom of the box) of measurements. The whiskers represent 1.5 times the interquartile range from the top or bottom of the box. N = 3 independent experiments, each experiment using kidney from different individual n = 3 biologically independent samples: human sample 1 = 8185 cells, human sample 2 = 3707 cells, human sample 3 = 6358 cells.

https://doi.org/10.1038/s41591-023-02314-7







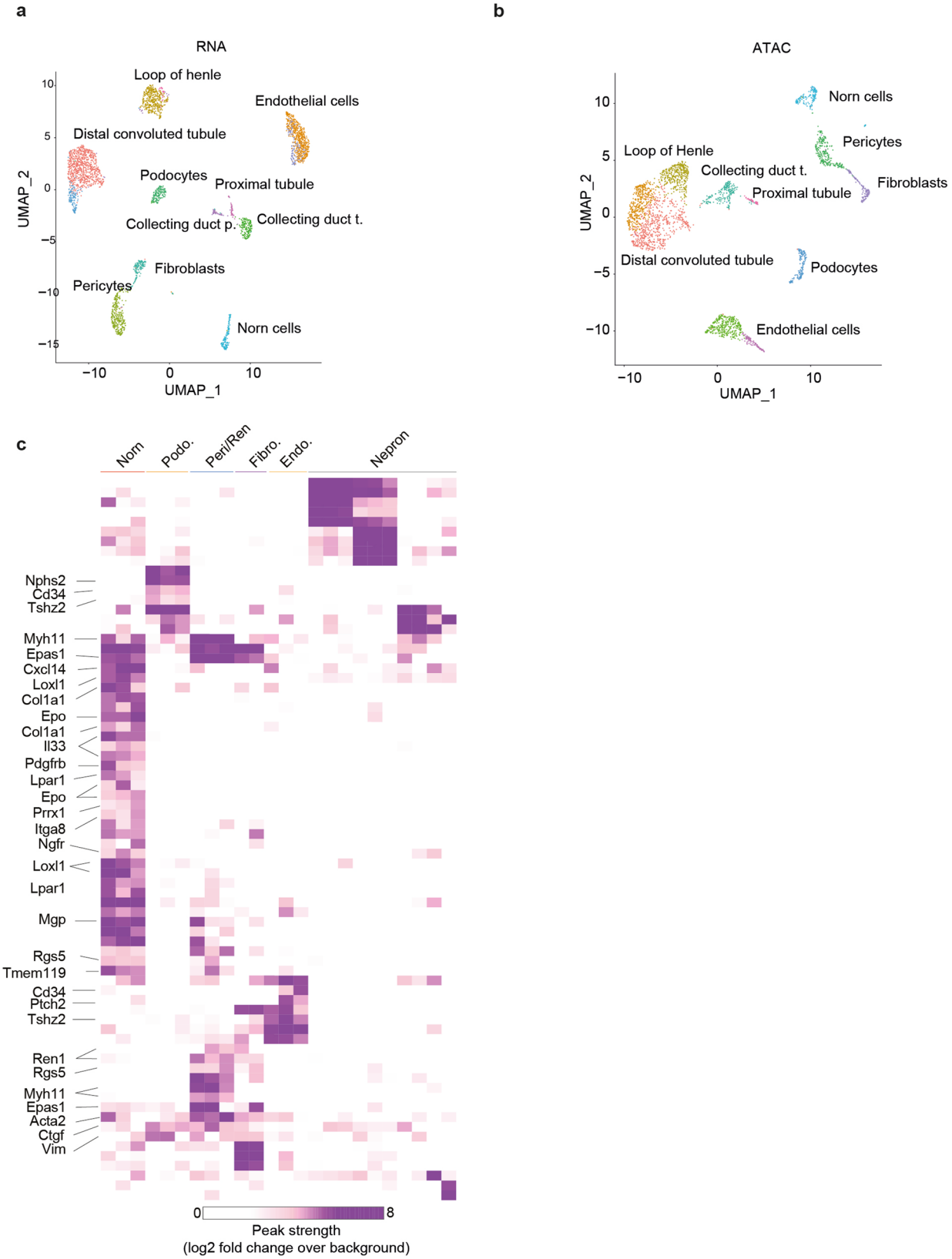


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41591-023-02314-7

Extended Data Fig. 2 | Identification and validation of Norn cell markers. **a**, Label transfer from Park et al.[43] to MARS-seq meta-cell clusters. Each coloured circle denotes a meta-cell. Colour scale denotes z-score of each gene signature across all meta-cells. Collecting duct i., collecting duct intercalated; collecting duct t., collecting duct transient; collecting duct p., collecting duct principal; Distal conv. Tub., distal convoluted tubule. **b**, Bubble plot depicting enrichment of selected marker genes used for cell labels. Coloured circles denote expression intensity (log_2 fold change over the mean). Size of circle represents the percentage of cells expressing the gene within cell type. **c**, Plot depicts quantification of double positive cells in mRNA FISH imaging of hypoxic (0.1%, 4 hours) kidney sections. N = 1 biological specimen (female), n = 3 independent experiments. The mean is shown. Quantification of cells are shown on top of the plot. Y-axis represents the percentage of *Epo* expressing cells that are double positive for tested Norn cell marker expression (x-axis). X-axis is showing Norn cell markers tested. **d**, Schematic of experimental procedure and representative gating strategy using CD73 to enrich for Norn cells in wildtype C57BL/6J mice. Linage depletion using CD31 (endothelial), CD326 (epithelial), and CD45 (immune) markers, and enrichment of CD73 positive cells.

https://doi.org/10.1038/s41591-023-02314-7



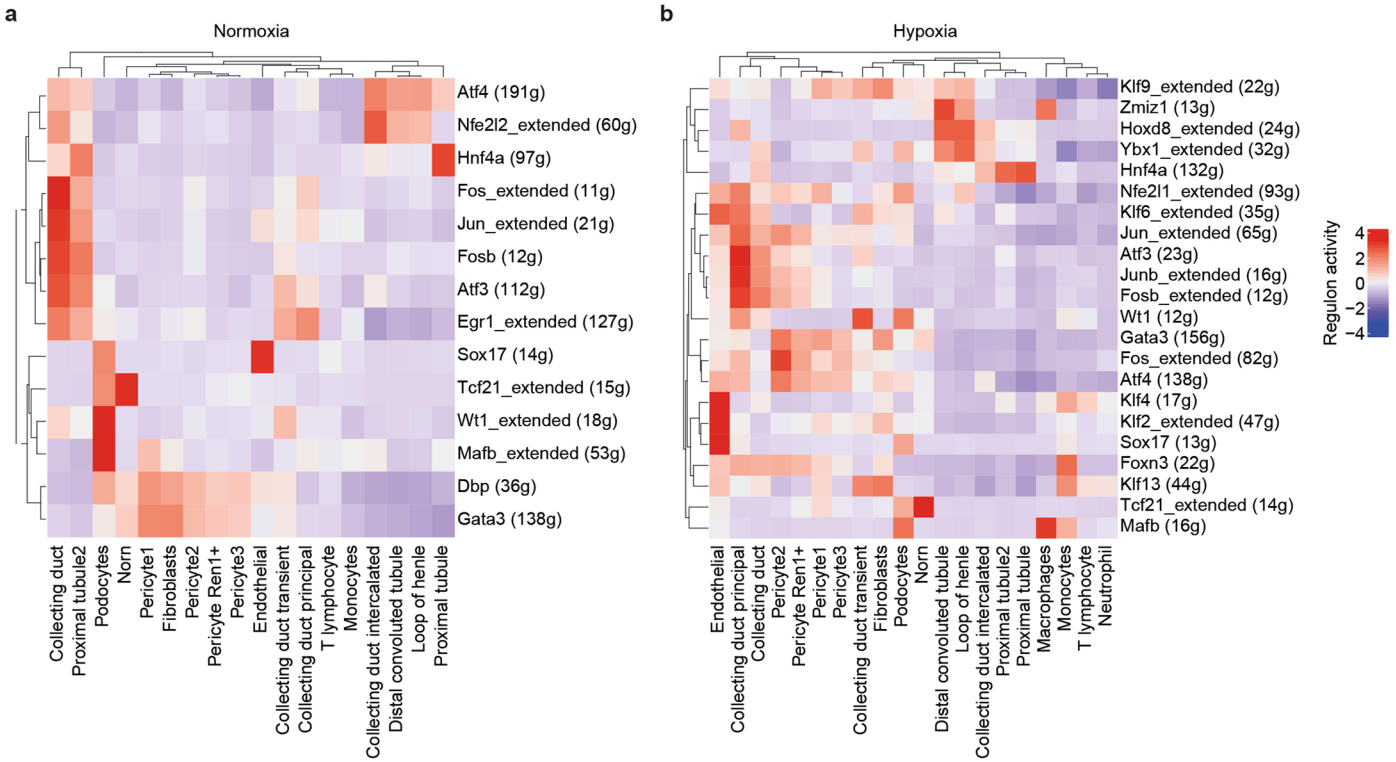
Extended Data Fig. 3 | Extended ATAC analysis. a, UMAP showing cell type clusters based on 10x multiome RNA using the Seurat tool[71]. Cell label transfer based on supplementary table 1. **b**, UMAP showing cell type clusters based on 10x multiome ATAC using Seurat tool. **c**, Heatmap representing sub clustering of selected ATAC-seq regions, and their associated genes.

https://doi.org/10.1038/s41591-023-02314-7

Rank	Motif	Name	P-value	# Target Sequences with Motif	% of Targets Seq. with Motif	# Background Sequences with Motif	% of Background Sequences with Motif
1		Tcf21(bHLH)/ArterySmoothMuscle-Tcf21-ChIP-Seq(GSE61369)/Homer	1e-196	1524.0	26.08%	42416.7	11.77%
2		Ap4(bHLH)/AML-Tfap4-ChIP-Seq(GSE45738)/Homer	1e-146	1689.0	28.90%	55995.8	15.54%
3		Atoh1(bHLH)/Cerebellum-Atoh1-ChIP-Seq(GSE22111)/Homer	1e-140	1455.0	24.90%	45709.4	12.69%
4		NeuroG2(bHLH)/Fibroblast-NeuroG2-ChIP-Seq(GSE75910)/Homer	1e-134	1801.0	30.82%	63173.4	17.53%
5		Tcf12(bHLH)/GM12878-Tcf12-ChIP-Seq(GSE32465)/Homer	1e-117	1317.0	22.54%	42362.4	11.76%
6		Myf5(bHLH)/GM-Myf5-ChIP-Seq(GSE24852)	1e-116	1067.0	18.26%	31195.3	8.66%
7		ZBTB18(Zf)/HEK293-ZBTB18.GFP-ChIP-Seq(GSE58341)/Homer	1e-113	815.0	13.95%	20960.5	5.82%
8		MyoG(bHLH)/C2C12-MyoG-ChIP-Seq(GSE36024)/Homer	1e-110	1368.0	23.41%	45819.9	12.72%
9		Olig2(bHLH)/Neuron-Olig2-ChIP-Seq(GSE30882)/Homer	1e-107	2001.0	34.24%	78062.2	21.67%
10		MyoD(bHLH)/Myotube-MyoD-ChIP-Seq(GSE21614)/Homer	1e-100	1062.0	18.17%	33068.5	9.18%
11		Ascl1(bHLH)/NeuralTubes-Ascl1-ChIP-Seq(GSE55840)/Homer	1e-86	1600.0	27.38%	61196.1	16.98%
12		CEBP(bZIP)/ThioMac-CEBPb-ChIP-Seq(GSE21512)/Homer	1e-82	768.0	13.14%	22311.1	6.19%
13		NeuroD1(bHLH)/Islet-NeuroD1-ChIP-Seq(GSE30298)/Homer	1e-69	1016.0	17.39%	35486.0	9.85%
14		AP-1(bZIP)/ThioMac-PU.1-ChIP-Seq(GSE21512)	1e-60	1033.0	17.68%	37884.4	10.51%
15		Fra2(bZIP)/Striatum-Fra2-ChIP-Seq(GSE43429)	1e-59	800.0	13.69%	26905.7	7.47%
16		Atf3(bZIP)/GBM-A TF3-ChIP-Seq(GSE33912)	1e-57	950.0	16.26%	34280.3	9.51%
17		JunB(bZIP)/DendriticCells-Junb-ChIP-Seq(GSE36099)/Homer	1e-56	857.0	14.66%	30020.2	8.33%
18		BATF(bZIP)/Th17-BA TF-ChIP-Seq(GSE39756)	1e-56	941.0	16.10%	34178.6	9.49%
19		Fra1(bZIP)/BT549-Fra1-ChIP-Seq(GSE46166)	1e-53	845.0	14.46%	30034.4	8.34%
20		E2A(bHLH)/proBcell-E2A-ChIP-Seq(GSE21978)	1e-49	1299.0	22.23%	53707.5	14.91%
21		SCL(bHLH)/HPC7-Scl-ChIP-Seq(GSE13511)	1e-45	3469.0	59.36%	180295.3	50.04%
22		HEB(bHLH)/mES-Heb-ChIP-Seq(GSE53233)	1e-45	1602.0	27.41%	70970.7	19.70%
23		Ptf1a(bHLH)/Panc1-Ptf1a-ChIP-Seq(GSE47459)	1e-45	2150.0	36.79%	101712.3	28.23%
24		RUNX(Runt)/HPC7-Runx1-ChIP-Seq(GSE22178)	1e-44	747.0	12.78%	26989.5	7.49%
25		Fosl2(bZIP)/3T3L1-Fosl2-ChIP-Seq(GSE56872)	1e-43	613.0	10.49%	20795.8	5.77%
26		RUNX-AML(Runt)/CD4+-PolII-ChIP-Seq(Barski_et_al.)/Homer	1e-39	717.0	12.27%	26473.8	7.35%
27		Gata6(Zf)/HUG1N-GATA6-ChIP-Seq(GSE51936)	1e-39	726.0	12.42%	27009.8	7.50%
28		Gata2(Zf)/K562-GATA2-ChIP-Seq(GSE18829)	1e-38	576.0	9.86%	19950.2	5.54%
29		ETS1(ETS)/Jurkat-ETS1-ChIP-Seq(GSE17954)	1e-37	1117.0	19.11%	47141.4	13.08%
30		Gata4(Zf)/Heart-Gata4-ChIP-Seq(GSE35151)	1e-36	784.0	13.42%	30460.5	8.45%

Extended Data Fig. 4 | Transcription-factor motif analysis. Transcription-factor motif analysis showing enrichment of motifs in accessible regions n = 5,844 peaks Norn specific ATAC-enhancer peaks vs n = 360380 total background sequences. P-value is calculated using cumulative binomial distributions.

https://doi.org/10.1038/s41591-023-02314-7

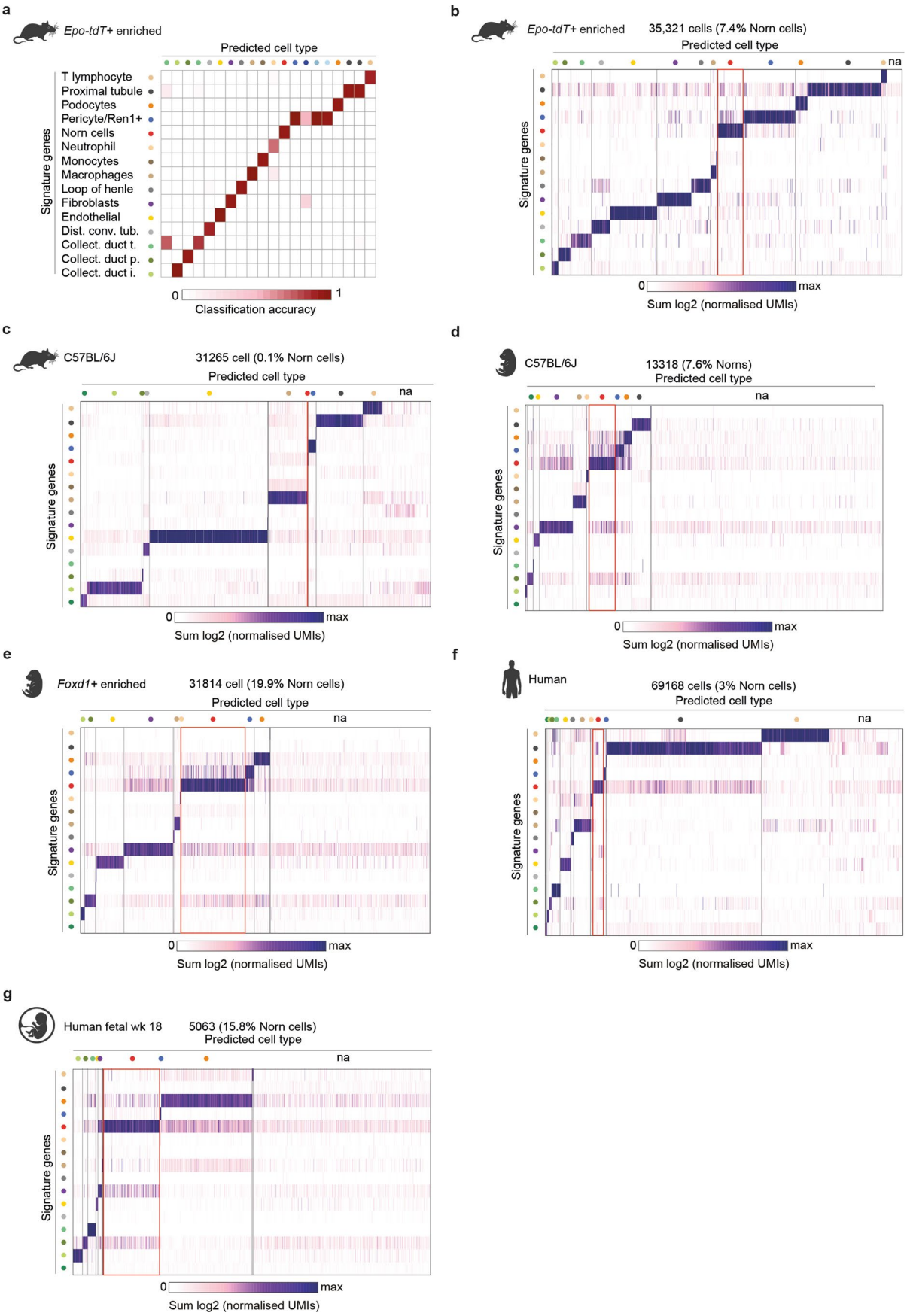


c

Tcf21 regulome (normoxia)		Tcf21 regulome (hypoxia)	
1	Cdkn1c	1	Cdkn1c
2	Cfh	2	Cfh
3	Cxcl14	3	Cxcl14
4	Dcn	4	Dcn
5	Fbln5	5	Fbln5
6	Gsn	6	Gsn
7	Hsd11b1	7	Hsd11b1
8	Igfbp3	8	Igfbp3
9	Mafb	9	Mafb
10	Myadm	10	Mgp
11	Nphs1	11	Nphs1
12	Scn7a	12	Serping1
13	Serping1	13	Thsd7a
14	Thsd7a		

Extended Data Fig. 5 | SCENIC analysis of Norn cells identifies Tcf21 regulome. a, Heatmap representing TF regulome enrichment from MARS-seq dataset derived from normoxic kidneys. **b**, Heatmap representing TF regulome enrichment in our MARS-seq dataset derived from hypoxic kidneys. **c**, Tcf21 regulome: List of downstream targets of Tcf21.

https://doi.org/10.1038/s41591-023-02314-7

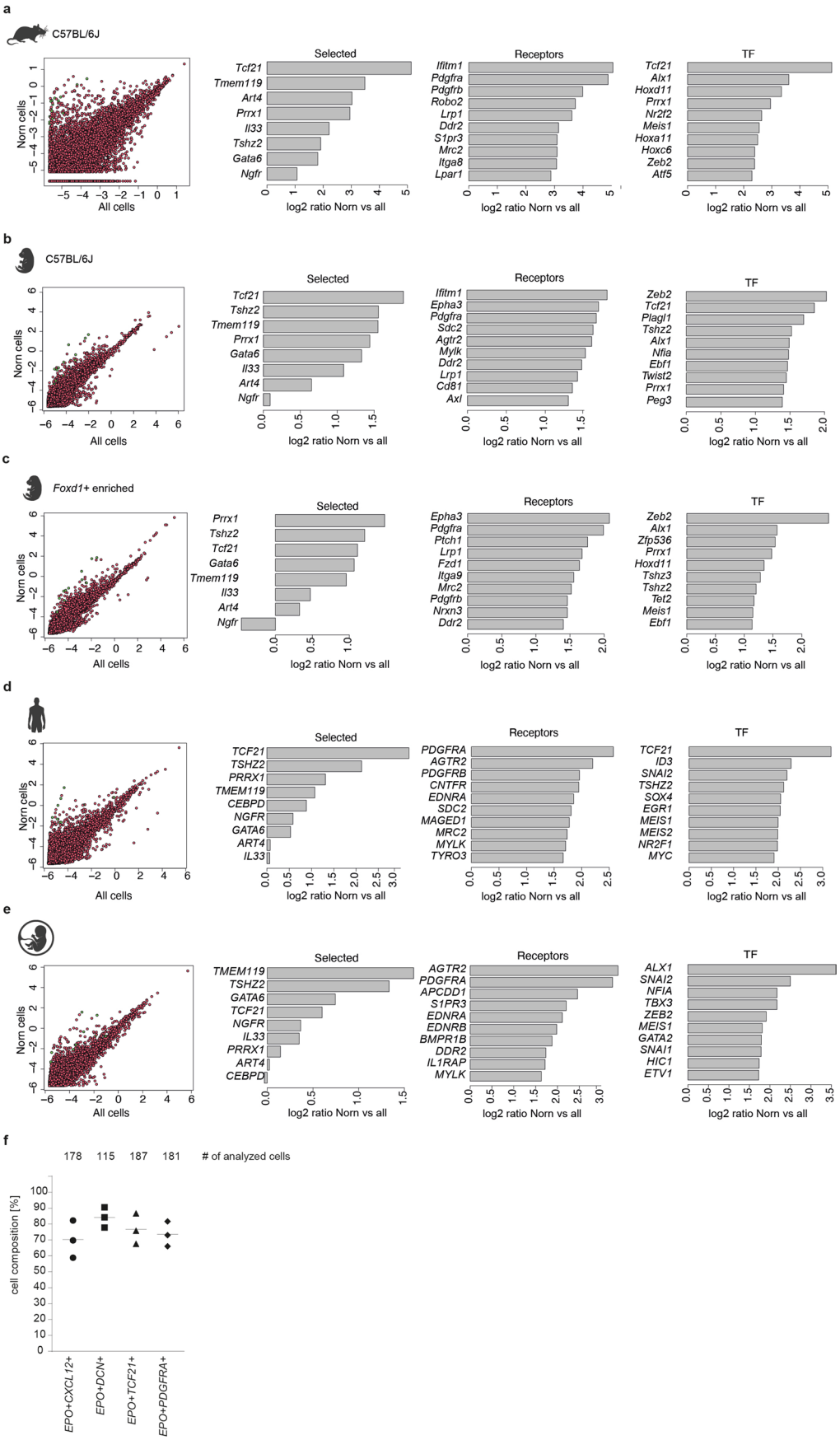


Extended Data Fig. 6 | See next page for caption.

https://doi.org/10.1038/s41591-023-02314-7

Extended Data Fig. 6 | Identification of Norn cells across development and species. a, Prediction of identified renal cell types, including Norn cells, in adult mouse kidney MARS-seq data (as shown in Fig. 4a). Heatmap showing sum $\log_2$ of normalised UMIs. **b**–**g** Heatmaps representing predicted cell types, with red box highlighting Norn-like cells: (**b**) Norn enriched MARS-seq renal dataset (7.4% Norn cells), (**c**) C57BL/6J adult mice renal scRNA-seq[44] (0.1% Norn cells), (**d**) E18.5 foetal kidney scRNA-seq data[45] (7.6% Norn cells, n = 3 individuals), (**e**) E18.5 *Foxd1* enriched foetal kidney scRNA seq data[46] (19.9% Norn cells), (**f**) adult human kidney scRNA seq data (3% Norn cells, n = 12 individuals)[42], and (**g**) foetal week 18 human kidney scRNA seq data (18.8% Norn cells, n = 1 individual)[47].

https://doi.org/10.1038/s41591-023-02314-7



Extended Data Fig. 7 | See next page for caption.

https://doi.org/10.1038/s41591-023-02314-7

Extended Data Fig. 7 | Conservation of Norn cell markers across development and species. a–**e**: Scatter plot depicting global log_2 size-normalised gene expression in Norn cells versus all other cell types. Green circles represent genes shown in 'selected' bar plot. Bar plots showing log_2 fold change of Norn cells versus all other cell types in selected genes, receptors, and TFs. **a**, C57BL/6J adult mice renal scRNA-seq[44]. **b**, E18.5 unenriched mouse foetal kidney scRNA-seq data[45]. **c**, E18.5 *Foxd1* enriched foetal kidney scRNA seq data[46]. **d**, Adult human unenriched kidney scRNA seq data[42]. **e**, Foetal week 18 human unenriched kidney scRNA seq data (18.8% Norn cells)[47]. **f**, Plot showing quantification of double positive cells in mRNA FISH imaging from hypoxic human kidney. N = 3 independent experiments from n = 1 biological specimen (female). The mean is shown. Y-axis is showing percentage of *EPO* expressing cells that are double positive for tested Norn cell marker expression (x-axis). X-axis is presenting Norn cell markers tested.

nature portfolio

Corresponding author(s): Ido Amit

Last updated by author(s): Mar 16, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No open-source or custom code was used to collect data for this paper.
Data analysis	No custom software was used to collect data. Metacell source code can be found at https://github.com/tanaylab/metacell. Source code to identify Norn cells in human single cell atlas can be found at https://github.com/AmitLab/Kidney-Norn-cells-identification. Source code used for transcription factor motif analysis can be found at https://github.com/vanheeringen-lab/gimmemotifs. Multiome ATAC-seq data analyses using Cell ranger ARC/2.0.0 tool. Motif identification were done using the HOMER package (version 4.11) and GimmeMotifs (v0.17.2). We used DESEQ2 (version 1.30.1) to test for statistically significant differential ATAC peaks. Single-cell RNAseq data were analysed with the published MetaCell package (version 0.3.41). For 10x scRNA-seq libraries, samples demultiplexed using cellranger/6.1.2. Reads were mapped using HISAT (version 0.1.6). For FACS analysis, we used the following software: FACSDiva 7 FlowJo 10.4.2 All data analysis was performed in R (version 3.5.3).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Mouse single-cell RNA-sequencing and single-cell ATAC-seq data, and human single-cell RNA-seq data that support the findings of this study, were deposited in the Gene Expression Omnibus (GEO) under accession code GSE193321. Previously published scRNA-seq data that were re-analysed here are available under the accession codes GSE11453047, GSE12979844, EGAS00001002325GSE15579446, and GSE10758543, as well as from the European Genome-phenome Archive (EGA) under study IDs EGAS0000100232542. We used mouse genome mm10 and human genome Hg38 as reference genomes. ATAC-seq tracks are accessible on UCSC browser https://genome.ucsc.edu/s/bjoert/Mouse%20hypoxia%20kidney%20bulk%20scATAC%2Dseq%20from%20single%20nuclei .

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to pre-determine sample sizes. Sample sizes were based on the experience in working with obtaining Norn cells from mouse kidneys. We obtained on average circa 50 QC Norn cells from each mouse. To obtain high quality representation of Norn cell heterogeneity in our MARS-seq data, we aimed for 2000 QC Norn cells. Based on our experience and since the identity and profile of Norn cells was reproducible across each experiment, we reasoned that using a single mouse, or pooling mice together, in a single experiment was sufficient to obtain high quality data in 10x multiome analysis, qRT-PCR, and RNA in situ hybridisation.
Data exclusions	We excluded data based on the following points. 1. Mapping of RNA-seq reads was done using HISAT (version 0.1.6) (Kim et al., 2015); reads with multiple mapping positions were excluded. This was parameter was pre-established. 2. 10x Genomics scRNA-seq, suspected doublets and cells with less than 200 UMIs were excluded: nFeature_RNA > 200 & nFeature_RNA < 5000 & percent.mt < 25). We followed the standard guidelines used for analysis of 10x genomics. 3. For 10x Genomics scATAC-seq following exclusion was done: nFeature_RNA > 1000 & nFeature_RNA < 30000. We followed the standard guidelines used for analysis of 10x genomics. 4. For MARS-seq, suspected doublets or more than 40% mitochondrial RNA content were excluded. This was parameter was pre-established. 5. For MARS-seq, cells with less than 200 UMIs were excluded. This parameter was determined after analysing hundreds of Norn cells and we observed that even with 200 UMIs we could obtain sufficient specific data in Meta-cell analysis.
Replication	1. For qPCR analysis, we performed 4 technical replicates per tissue derived from a single male mouse. For time point series, we performed 4 technical replicates per kidney derived from one animal (male) per time point. One experiment was done. We successfully replicated the data in the laboratory using 8% oxygen as hypoxia. 2. For 10x multiome, kidneys from 5 mice (3 males, 2 females) were pooled together to FACS bulk sort 75.000 cells. One experiment was done, since it was extremely challenging to obtain a sufficient number of cells for this demanding protocol. 3. For MARS-seq, cells were pooled from kidneys obtained from 52 mice (33 males, 19 females). N=22 independently performed experiments. Experiments were reproducible with variable enrichment of Norn cells (3-33% enrichment). 4. For 10x scRNAseq of human kidneys, we did separate runs of each indiviudal (3 human individuals in total, 2 males and 1 female). Experiments were reproducible with variable enrichment of Norn cells (5-20% enrichment). 5. For RNA in situ hybridisation, n=1 mouse kidney (male) or n=1 human kidney (female) was used to perform n=3 independent experiments. The experiments were reproduced successfully.
Randomization	The conclusions of this study, mainly the existence of Norn cells and their transcriptional and epigenetic landscapes, were derived from an aggregated analysis of the entire population of collected cells. We did not perform any statistical analysis regarding variation between individuals (mouse or human). Therefore, no randomization was applied during data collection and analysis.
Blinding	The conclusions of this study, mainly the existence of Norn cells and their transcriptional and epigenetic landscapes, were derived from an aggregated analysis of the entire population of collected cells. We did not perform any statistical analysis regarding variation between individuals (mouse or human). Therefore, no blinding was applied during data collection and analysis..

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Mouse: anti-mouse CD31-APC (Biolegend, Cat no. 102409, Clone 390; 1:400), anti-mouse CD326-FITC (Biolegend, Cat. no. 118207, Clone G8.8; 1:600), anti-mouse CD45-FITC (eBioscience TY/11.8, Cat no. 11-0451-8, Clone 30-F11; 1:400) and anti-mouse CD73-APC-Fire750 (Biolegend, Cat. no. 127221, 1:200)
Human: anti-human CD326-PE (Biolegend, Cat no. 324206, Clone 9C4; 1:100), anti-human CD45-FITC (Biolegend, Cat no. 304006, Clone HI30; 1:100), anti-human CD31-PE (Biolegend, Cat no. 303105, Clone WM59; 1:100), anti-human CD10-APC (Biolegend, Cat no. 312210, Clone HI10a; 1:100), and anti-human CD235a-BV605 (Biolegend, Cat no. 740409, Clone GA-R2; 1:100)

Validation

All antibodies used are commercially available and have been used in previous studies, such as PMID: 33176333. In addition, we validated FACS antibodies by performing a titration analysis and compared (separately) to an unstained sample. Only FACS antibodies that yielded high signal and separation in the flow cytometer (fold change > 10) were used for further experiments involving sorting and index sorting. We performed this for both mouse and human samples.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

For qPCR wild-type C57BL/6 (B6) 10-12 weeks male mice were used (purchased from Harlan Laboratories).
For MARS-seq C57BL/6 (B6) 10-12 week old male and female mice (purchased from Harlan Laboratories) as well as 10-18 week old Epo-CreERT2-tdTomato male and female mice (Epo-tdTomato+; Imeri et al 2019) were used. Epo-tdTomato+ mouse model enables permanent, conditional, and exclusive labelling of Epo producing Norn cells that expressed Epo during a 4-hour hypoxia treatment following a 5-day course of tamoxifen treatment. The transgenic large BAC reporter construct includes Epo regulatory elements and neighbouring genes.
All animals were bred and housed in standard ventilated cages on a 12-h light and dark cycles with temperatures of 18-24°C and humidity of 35-60% according to the guidelines of Weizmann Institute of Science and the IGMM and IGF animal facilities (RAM). All animals were housed in groups of a maximum five per cage with ad libitum access to food (Harlan Teklad, Cat. no. 2918) and water. Sex was not considered in the study design. Since there is no evidence that there are sex-specific differences in Epo regulation, we included male and female mice and analysed together using Metacell.

Wild animals

No wild animals were used

Field-collected samples

No field-collected samples were used

Ethics oversight

All mouse experimental procedures complied with relevant ethical protocols and were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC), application numbers 01810220-2 and 11280219-2, at Weizmann Institute of Science.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics

Non-tumour kidney tissue was obtained from two male (68 and 83 years old, respectively) and one female (76 years old) patients undergoing partial or radical nephrectomy after signing an informed consent form according to the Helsinki protocols 6370-19-SMC (Sheba-Telhashomer Medical Centre), 0417-20-TLV (Ichilov hospital), and 0297-22-HMO (Hadassah Medical Centre) with local IRB approval from Weizmann Institute of Science (1568-1). Exclusion criteria: infection.

Recruitment

Helsinki and IRB approval was obtained prior to recruitment. Patients scheduled to undergo radical or partial nephrectomy were recruited by informed consent. Since these patients often have very large tumours in the kidney, it is challenging to estimate the function and quality of the non-tumour portion of the kidney that we obtained. Moreover, the exact sample site and section obtained differed between patients. This may thus have impact the proportions of cells in our samples as well as the transcriptional profile of the cells. Sex was not considered in the study design.

Ethics oversight

Non-tumour kidney tissue was obtained from two male (68 and 83 years old, respectively) and one female (76 years old) patients undergoing partial or radical nephrectomy after signing an informed consent form according to the Helsinki protocols 6370-19-SMC (Sheba-Telhashomer Medical Centre), 0417-20-TLV (Ichilov hospital), and 0297-22-HMO (Hadassah Medical Centre) with local IRB approval from Weizmann Institute of Science (1568-1).

The collection and use of FFPE sections of human kidney samples by the Institute for Forensic Medicine, University Hospital Essen, Germany, was approved by the ethical commission of the Medical Faculty of the University Duisburg-Essen 2015 (15-6345-BO). The sample used in this study was derived from a 55-year-old female who died in a fire accident (CO-Hb content of 50%) and who underwent a forensic autopsy. Because a forensic sample was used and to preserve complete anonymization imposed by the ethical permit, we did not obtain consent from the individual or the family.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

GSE193321, https://genome.ucsc.edu/s/bjoert/Mouse%20hypoxia%20kidney%20bulk%20scATAC%2Dseq%20from%20single%20nuclei

Files in database submission

norn_atac_S1_L001_I1_001.fastq.gz
norn_atac_S1_L001_R1_001.fastq.gz
norn_atac_S1_L001_R2_001.fastq.gz
norn_atac_S1_L001_R3_001.fastq.gz
norn_S3_L001_I1_001.fastq.gz
norn_S3_L001_I2_001.fastq.gz
norn_S3_L001_R1_001.fastq.gz
norn_S3_L001_R2_001.fastq.gz
norn_S3_L002_I1_001.fastq.gz
norn_S3_L002_I2_001.fastq.gz
norn_S3_L002_R1_001.fastq.gz
norn_S3_L002_R2_001.fastq.gz
ATAC/barcodes.tsv.gz
ATAC/features.tsv.gz
ATAC/matrix.mtx.gz
Collecting_duct_principal_SC_cells.bed
Distal_convoluted_tubule_SC_cells.bed
Fibroblasts_SC_cells.bed
Norn_SC_cells.bed
Podocytes_SC_cells.bed
Collecting_duct_transient_SC_cells.bed
Endothelial_SC_cells.bed
Loop_of_henle_SC_cells.bed
Pericytes_Renin_SC_cells.bed
Proximal_tubule_SC_cells.bed
Sample1_features.tsv
Sample1_matrix.mtx
Sample1_barcodes.tsv
ShaharH_Gene_S2_L001_R1_001.fastq.gz
ShaharH_Gene_S2_L001_R2_001.fastq.gz
ShaharH_Gene_S2_L002_R1_001.fastq.gz
ShaharH_Gene_S2_L002_R2_001.fastq.gz
Bjort-Shahar-GEX_S1_L001_R1_001.fastq.gz
Bjort-Shahar-GEX_S1_L001_R2_001.fastq.gz
Bjort-Shahar-GEX_S1_L002_R1_001.fastq.gz
Bjort-Shahar-GEX_S1_L002_R2_001.fastq.gz
Bjort-Shahar-GEX_S1_L002_R2_001.fastq.gz
Sample2_barcodes.tsv
Sample2_features.tsv
Sample2_matrix.mtx
Hs_Kidney3_GEX_Bjort_S3_L001_R1_001.fastq.gz
Hs_Kidney3_GEX_Bjort_S3_L001_R2_001.fastq.gz
Hs_Kidney3_GEX_Bjort_S3_L002_R1_001.fastq.gz
Hs_Kidney3_GEX_Bjort_S3_L002_R2_001.fastq.gz
Sample3_features.tsv
Sample3_matrix.mtx
Sample3_barcodes.tsv
metatada_GEO.txt
AB10289_SB402_R1_01.fastq.gz
AB10289_SB402_R2_01.fastq.gz
AB10290_SB402_R1_01.fastq.gz
AB10290_SB402_R2_01.fastq.gz
AB10291_SB402_R1_01.fastq.gz
AB10291_SB402_R2_01.fastq.gz
AB10292_SB402_R1_01.fastq.gz

AB10292_SB402_R2_01.fastq.gz
AB10293_SB402_R1_01.fastq.gz
AB10293_SB402_R2_01.fastq.gz
AB10344_SB403_R1_01.fastq.gz
AB10344_SB403_R2_01.fastq.gz
AB10345_SB403_R1_01.fastq.gz
AB10345_SB403_R2_01.fastq.gz
AB10346_SB403_R1_01.fastq.gz
AB10346_SB403_R2_01.fastq.gz
AB10347_SB403_R1_01.fastq.gz
AB10347_SB403_R2_01.fastq.gz
AB10448_SB404_R1_01.fastq.gz
AB10448_SB404_R2_01.fastq.gz
AB10449_SB404_R1_01.fastq.gz
AB10449_SB404_R2_01.fastq.gz
AB10450_SB404_R1_01.fastq.gz
AB10450_SB404_R2_01.fastq.gz
AB10451_SB404_R1_01.fastq.gz
AB10451_SB404_R2_01.fastq.gz
AB10452_SB404_R1_01.fastq.gz
AB10452_SB404_R2_01.fastq.gz
AB10453_SB404_R1_01.fastq.gz
AB10453_SB404_R2_01.fastq.gz
AB10454_SB404_R1_01.fastq.gz
AB10454_SB404_R2_01.fastq.gz
AB10455_SB404_R1_01.fastq.gz
AB10455_SB404_R2_01.fastq.gz
AB10460_SB404_R1_01.fastq.gz
AB10460_SB404_R2_01.fastq.gz
AB10461_SB404_R1_01.fastq.gz
AB10461_SB404_R2_01.fastq.gz
AB10926_SB411_R1_001.fastq.gz
AB10926_SB411_R2_001.fastq.gz
AB10926_SB411_Undetermined_R1_001.fastq.gz
AB10926_SB411_Undetermined_R2_001.fastq.gz
AB10927_SB411_R1_001.fastq.gz
AB10927_SB411_R2_001.fastq.gz
AB10927_SB411_Undetermined_R1_001.fastq.gz
AB10927_SB411_Undetermined_R2_001.fastq.gz
AB10928_SB411_R1_001.fastq.gz
AB10928_SB411_R2_001.fastq.gz
AB10928_SB411_Undetermined_R1_001.fastq.gz
AB10928_SB411_Undetermined_R2_001.fastq.gz
AB10929_SB411_R1_001.fastq.gz
AB10929_SB411_R2_001.fastq.gz
AB10929_SB411_Undetermined_R1_001.fastq.gz
AB10929_SB411_Undetermined_R2_001.fastq.gz
AB11060_SB413_R1_01.fastq.gz
AB11060_SB413_R2_01.fastq.gz
AB11061_SB413_R1_01.fastq.gz
AB11061_SB413_R2_01.fastq.gz
AB11062_SB413_R1_01.fastq.gz
AB11062_SB413_R2_01.fastq.gz
AB11063_SB413_R1_01.fastq.gz
AB11063_SB413_R2_01.fastq.gz
AB11064_SB413_R1_01.fastq.gz
AB11064_SB413_R2_01.fastq.gz
AB11065_SB413_R1_01.fastq.gz
AB11065_SB413_R2_01.fastq.gz
AB11066_SB413_R1_01.fastq.gz
AB11066_SB413_R2_01.fastq.gz
AB11306_SB415_R1_01.fastq.gz
AB11306_SB415_R2_01.fastq.gz
AB11307_SB415_R1_01.fastq.gz
AB11307_SB415_R2_01.fastq.gz
AB11308_SB415_R1_01.fastq.gz
AB11308_SB415_R2_01.fastq.gz
AB11309_SB415_R1_01.fastq.gz
AB11309_SB415_R2_01.fastq.gz
AB11310_SB415_R1_01.fastq.gz
AB11310_SB415_R2_01.fastq.gz
AB11311_SB415_R1_01.fastq.gz
AB11311_SB415_R2_01.fastq.gz
AB11312_SB415_R1_01.fastq.gz
AB11312_SB415_R2_01.fastq.gz
AB11313_SB415_R1_01.fastq.gz

AB11313_SB415_R2_01.fastq.gz
AB11409_SB416_R1_01.fastq.gz
AB11409_SB416_R2_01.fastq.gz
AB11410_SB416_R1_01.fastq.gz
AB11410_SB416_R2_01.fastq.gz
AB11411_SB416_R1_01.fastq.gz
AB11411_SB416_R2_01.fastq.gz
AB11412_SB416_R1_01.fastq.gz
AB11412_SB416_R2_01.fastq.gz
AB11413_SB416_R1_01.fastq.gz
AB11413_SB416_R2_01.fastq.gz
AB11414_SB416_R1_01.fastq.gz
AB11414_SB416_R2_01.fastq.gz
AB11415_SB416_R1_01.fastq.gz
AB11415_SB416_R2_01.fastq.gz
AB11416_SB416_R1_01.fastq.gz
AB11416_SB416_R2_01.fastq.gz
AB11630_SB418_R1_01.fastq.gz
AB11630_SB418_R2_01.fastq.gz
AB11631_SB418_R1_01.fastq.gz
AB11631_SB418_R2_01.fastq.gz
AB11632_SB418_R1_01.fastq.gz
AB11632_SB418_R2_01.fastq.gz
AB11633_SB418_R1_01.fastq.gz
AB11633_SB418_R2_01.fastq.gz
AB11634_SB418_R1_01.fastq.gz
AB11634_SB418_R2_01.fastq.gz
AB11635_SB418_R1_01.fastq.gz
AB11635_SB418_R2_01.fastq.gz
AB11636_SB418_R1_01.fastq.gz
AB11636_SB418_R2_01.fastq.gz
AB11637_SB418_R1_01.fastq.gz
AB11637_SB418_R2_01.fastq.gz
AB11638_SB418_R1_01.fastq.gz
AB11638_SB418_R2_01.fastq.gz
AB11639_SB418_R1_01.fastq.gz
AB11639_SB418_R2_01.fastq.gz
AB11640_SB418_R1_01.fastq.gz
AB11640_SB418_R2_01.fastq.gz
AB11641_SB418_R1_01.fastq.gz
AB11641_SB418_R2_01.fastq.gz
AB12023_SB422_R1_01.fastq.gz
AB12023_SB422_R2_01.fastq.gz
AB12024_SB422_R1_01.fastq.gz
AB12024_SB422_R2_01.fastq.gz
AB12025_SB422_R1_01.fastq.gz
AB12025_SB422_R2_01.fastq.gz
AB12026_SB422_R1_01.fastq.gz
AB12026_SB422_R2_01.fastq.gz
AB12027_SB422_R1_01.fastq.gz
AB12027_SB422_R2_01.fastq.gz
AB12028_SB422_R1_01.fastq.gz
AB12028_SB422_R2_01.fastq.gz
AB12029_SB422_R1_01.fastq.gz
AB12029_SB422_R2_01.fastq.gz
AB12030_SB422_R1_01.fastq.gz
AB12030_SB422_R2_01.fastq.gz
AB12031_SB422_R1_01.fastq.gz
AB12031_SB422_R2_01.fastq.gz
AB12032_SB422_R1_01.fastq.gz
AB12032_SB422_R2_01.fastq.gz
AB12033_SB422_R1_01.fastq.gz
AB12033_SB422_R2_01.fastq.gz
AB12034_SB422_R1_01.fastq.gz
AB12034_SB422_R2_01.fastq.gz
AB12035_SB422_R1_01.fastq.gz
AB12035_SB422_R2_01.fastq.gz
AB12890_SB429_R1_01_1.fastq.gz
AB12890_SB429_R2_01_1.fastq.gz
AB12891_SB429_R1_01_2.fastq.gz
AB12891_SB429_R2_01_2.fastq.gz
AB12892_SB429_R1_01_1.fastq.gz
AB12892_SB429_R2_01_1.fastq.gz
AB12893_SB429_R1_01_2.fastq.gz
AB12893_SB429_R2_01_2.fastq.gz
AB12894_SB429_R1_01_1.fastq.gz

AB12894_SB429_R2_01_1.fastq.gz
AB12895_SB429_R1_01_2.fastq.gz
AB12895_SB429_R2_01_2.fastq.gz
AB13561_SB434_R1_01_2.fastq.gz
AB13561_SB434_R2_01_2.fastq.gz
AB13714_SB435_R1_01_1.fastq.gz
AB13714_SB435_R2_01_1.fastq.gz
AB13715_SB435_R1_01_2.fastq.gz
AB13715_SB435_R2_01_2.fastq.gz
AB13716_SB435_R1_01_1.fastq.gz
AB13716_SB435_R2_01_1.fastq.gz
AB13717_SB435_R1_01_2.fastq.gz
AB13717_SB435_R2_01_2.fastq.gz
AB13718_SB435_R1_01_1.fastq.gz
AB13718_SB435_R2_01_1.fastq.gz
AB13719_SB435_R1_01_2.fastq.gz
AB13719_SB435_R2_01_2.fastq.gz
AB13720_SB435_R1_01_1.fastq.gz
AB13720_SB435_R2_01_1.fastq.gz
AB13721_SB435_R1_01_2.fastq.gz
AB13721_SB435_R2_01_2.fastq.gz
AB13722_SB435_R1_01_1.fastq.gz
AB13722_SB435_R2_01_1.fastq.gz
AB13723_SB435_R1_01_2.fastq.gz
AB13723_SB435_R2_01_2.fastq.gz
AB14027_SB439_R1_01.fastq.gz
AB14027_SB439_R2_01.fastq.gz
AB14028_SB439_R1_01.fastq.gz
AB14028_SB439_R2_01.fastq.gz
AB14029_SB439_R1_01.fastq.gz
AB14029_SB439_R2_01.fastq.gz
AB14030_SB439_R1_01.fastq.gz
AB14030_SB439_R2_01.fastq.gz
AB14031_SB439_R1_01.fastq.gz
AB14031_SB439_R2_01.fastq.gz
AB14032_SB439_R1_01.fastq.gz
AB14032_SB439_R2_01.fastq.gz
AB14033_SB439_R1_01.fastq.gz
AB14033_SB439_R2_01.fastq.gz
AB14034_SB439_R1_01.fastq.gz
AB14034_SB439_R2_01.fastq.gz
AB14035_SB439_R1_01.fastq.gz
AB14035_SB439_R2_01.fastq.gz
AB14036_SB439_R1_01.fastq.gz
AB14036_SB439_R2_01.fastq.gz
AB14037_SB439_R1_01.fastq.gz
AB14037_SB439_R2_01.fastq.gz
AB14038_SB439_R1_01.fastq.gz
AB14038_SB439_R2_01.fastq.gz
AB14039_SB439_R1_01.fastq.gz
AB14039_SB439_R2_01.fastq.gz
AB14040_SB439_R1_01.fastq.gz
AB14040_SB439_R2_01.fastq.gz
AB14041_SB439_R1_01.fastq.gz
AB14041_SB439_R2_01.fastq.gz
AB14042_SB439_R1_01.fastq.gz
AB14042_SB439_R2_01.fastq.gz
AB14043_SB439_R1_01.fastq.gz
AB14043_SB439_R2_01.fastq.gz
AB14044_SB439_R1_01.fastq.gz
AB14044_SB439_R2_01.fastq.gz
AB9040_SB387_R1_01.fastq.gz
AB9040_SB387_R2_01.fastq.gz
AB9041_SB387_R1_01.fastq.gz
AB9041_SB387_R2_01.fastq.gz
AB9042_SB387_R1_01.fastq.gz
AB9042_SB387_R2_01.fastq.gz
AB9043_SB387_R1_01.fastq.gz
AB9043_SB387_R2_01.fastq.gz
AB9044_SB387_R1_01.fastq.gz
AB9044_SB387_R2_01.fastq.gz
AB9045_SB387_R1_01.fastq.gz
AB9045_SB387_R2_01.fastq.gz
AB9046_SB387_R1_01.fastq.gz
AB9046_SB387_R2_01.fastq.gz
AB9047_SB387_R1_01.fastq.gz

AB9047_SB387_R2_01.fastq.gz
AB9202_SB389_R1_01.fastq.gz
AB9202_SB389_R2_01.fastq.gz
AB9203_SB389_R1_01.fastq.gz
AB9203_SB389_R2_01.fastq.gz
AB9204_SB389_R1_01.fastq.gz
AB9204_SB389_R2_01.fastq.gz
AB9205_SB389_R1_01.fastq.gz
AB9205_SB389_R2_01.fastq.gz
AB9206_SB389_R1_01.fastq.gz
AB9206_SB389_R2_01.fastq.gz
AB9207_SB389_R1_01.fastq.gz
AB9207_SB389_R2_01.fastq.gz
AB9208_SB389_R1_01.fastq.gz
AB9208_SB389_R2_01.fastq.gz
AB9209_SB389_R1_01.fastq.gz
AB9209_SB389_R2_01.fastq.gz
AB9210_SB389_R1_01.fastq.gz
AB9210_SB389_R2_01.fastq.gz
AB9211_SB389_R1_01.fastq.gz
AB9211_SB389_R2_01.fastq.gz
AB9212_SB389_R1_01.fastq.gz
AB9212_SB389_R2_01.fastq.gz
AB9213_SB389_R1_01.fastq.gz
AB9213_SB389_R2_01.fastq.gz
AB9214_SB389_R1_01.fastq.gz
AB9214_SB389_R2_01.fastq.gz
AB9215_SB389_R1_01.fastq.gz
AB9215_SB389_R2_01.fastq.gz
AB9216_SB389_R1_01.fastq.gz
AB9216_SB389_R2_01.fastq.gz
AB9217_SB389_R1_01.fastq.gz
AB9217_SB389_R2_01.fastq.gz
AB9218_SB389_R1_01.fastq.gz
AB9218_SB389_R2_01.fastq.gz
AB9219_SB389_R1_01.fastq.gz
AB9219_SB389_R2_01.fastq.gz
AB9220_SB389_R1_01.fastq.gz
AB9220_SB389_R2_01.fastq.gz
AB9221_SB389_R1_01.fastq.gz
AB9221_SB389_R2_01.fastq.gz
AB9316_SB390_R1_01.fastq.gz
AB9316_SB390_R2_01.fastq.gz
AB9317_SB390_R1_01.fastq.gz
AB9317_SB390_R2_01.fastq.gz
AB9453_SB392_R1_01.fastq.gz
AB9453_SB392_R2_01.fastq.gz
AB9454_SB392_R1_01.fastq.gz
AB9454_SB392_R2_01.fastq.gz
AB9455_SB392_R1_01.fastq.gz
AB9455_SB392_R2_01.fastq.gz
AB9456_SB392_R1_01.fastq.gz
AB9456_SB392_R2_01.fastq.gz
AB9464_SB392_R1_01.fastq.gz
AB9464_SB392_R2_01.fastq.gz
AB10765_SB408_R1_01.fastq.gz
AB10765_SB408_R2_01.fastq.gz
AB10766_SB408_R1_01.fastq.gz
AB10766_SB408_R2_01.fastq.gz
AB10767_SB408_R1_01.fastq.gz
AB10767_SB408_R2_01.fastq.gz
AB10768_SB408_R1_01.fastq.gz
AB10768_SB408_R2_01.fastq.gz
AB10769_SB408_R1_01.fastq.gz
AB10769_SB408_R2_01.fastq.gz
AB10770_SB408_R1_01.fastq.gz
AB10770_SB408_R2_01.fastq.gz
AB10771_SB408_R1_01.fastq.gz
AB10771_SB408_R2_01.fastq.gz
AB10772_SB408_R1_01.fastq.gz
AB10772_SB408_R2_01.fastq.gz
AB10289.txt
AB10290.txt
AB10291.txt
AB10292.txt
AB10293.txt

AB10344.txt
AB10345.txt
AB10346.txt
AB10347.txt
AB10448.txt
AB10449.txt
AB10450.txt
AB10451.txt
AB10452.txt
AB10453.txt
AB10454.txt
AB10455.txt
AB10460.txt
AB10461.txt
AB10765.txt
AB10766.txt
AB10767.txt
AB10768.txt
AB10769.txt
AB10770.txt
AB10771.txt
AB10772.txt
AB10926.txt
AB10927.txt
AB10928.txt
AB10929.txt
AB11060.txt
AB11061.txt
AB11062.txt
AB11063.txt
AB11064.txt
AB11065.txt
AB11066.txt
AB11306.txt
AB11307.txt
AB11308.txt
AB11309.txt
AB11310.txt
AB11311.txt
AB11312.txt
AB11313.txt
AB11409.txt
AB11410.txt
AB11411.txt
AB11412.txt
AB11413.txt
AB11414.txt
AB11415.txt
AB11416.txt
AB11630.txt
AB11631.txt
AB11632.txt
AB11633.txt
AB11634.txt
AB11635.txt
AB11636.txt
AB11637.txt
AB11638.txt
AB11639.txt
AB11640.txt
AB11641.txt
AB12023.txt
AB12024.txt
AB12025.txt
AB12026.txt
AB12027.txt
AB12028.txt
AB12029.txt
AB12030.txt
AB12031.txt
AB12032.txt
AB12033.txt
AB12034.txt
AB12035.txt
AB12890.txt
AB12891.txt

AB12892.txt
AB12893.txt
AB12894.txt
AB12895.txt
AB13561.txt
AB13714.txt
AB13715.txt
AB13716.txt
AB13717.txt
AB13718.txt
AB13719.txt
AB13720.txt
AB13721.txt
AB13722.txt
AB13723.txt
AB14027.txt
AB14028.txt
AB14029.txt
AB14030.txt
AB14031.txt
AB14032.txt
AB14033.txt
AB14034.txt
AB14035.txt
AB14036.txt
AB14037.txt
AB14038.txt
AB14039.txt
AB14040.txt
AB14041.txt
AB14042.txt
AB14043.txt
AB14044.txt
AB9040.txt
AB9041.txt
AB9042.txt
AB9043.txt
AB9044.txt
AB9045.txt
AB9046.txt
AB9047.txt
AB9202.txt
AB9203.txt
AB9204.txt
AB9205.txt
AB9206.txt
AB9207.txt
AB9208.txt
AB9209.txt
AB9210.txt
AB9211.txt
AB9212.txt
AB9213.txt
AB9214.txt
AB9215.txt
AB9216.txt
AB9217.txt
AB9218.txt
AB9219.txt
AB9220.txt
AB9221.txt
AB9316.txt
AB9317.txt
AB9453.txt
AB9454.txt
AB9455.txt
AB9456.txt
AB9464.txt
metadata_f.txt

Genome browser session
(e.g. UCSC)

https://genome.ucsc.edu/s/bjoert/Mouse%20hypoxia%20kidney%20bulk%20scATAC%2Dseq%20from%20single%20nuclei

Methodology

Replicates	Pool of kidneys from 5 mice (2 females, 3 males) were FACS bulk sorted and used for a single experiment.
Sequencing depth	452,630,624 total reads, 10,666 ATAC Median high-quality fragments per cell, 1,159 GEX Median genes per cell
Antibodies	ATAC-seq data and thus antibodies not used.
Peak calling parameters	HOMER package (version 4.11) was used to identify ATAC peaks pooled over all the ATAC-seq reads (command findPeaks -style factor).
Data quality	Quality of cells were assessed by the Cell ranger ARC/2.0.0 tool. Number of peaks identified at FDR 5% and above 5-fold enrichment in our data: Collecting_duct_principal - number of putative peaks = 31903 (47 cells) Collecting_duct_transient - 69527 (106 cells) Distal_convoluted_tubule - 465968 (851 cells) Endothelial - 68227 (415 cells) Fibroblasts - 57009 (86 cells) Loop_of_henle - 291046 (415 cells) Norn - 85455 (147 cells) Pericytes_Renin - 100681 (261 cells) Podocytes - 82093 (191 cells) Proximal_tubule - 64503 (52 cells)
Software	Cell ranger ARC/2.0.0 tool, DESEQ2 version 1.30.1, HOMER version 4.11, and GimmeMotifs v0.17.2.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Mouse kidney: Mice were placed in an air-tight chamber and synthetic air containing 0.1% CO was continuously flowing through the chamber for 4 hours. In the time course experiment (Fig. 1b), mice were exposed to 0.1% CO for 2, 4, 6, or 12 hours. Mice were euthanized and immediately perfused with cold PBS. In the downstream process, we performed the tissue dissociation under normoxic conditions as follows: After isolation of kidneys, the renal capsule was removed, and minced with razor blades into 1 mm pieces. Each minced kidney was placed in 1 ml DMEM medium containing collagenase IV (Worthington Biochemical Cat no. LS004188, final concentration 0.25 mg/ml), neutral protease (Worthington Biochemical Cat no. LS02106, final concentration 0.25 mg/ml), and DNase I (Roche Cat. no. 10104159001, final concentration 2000 U/ml), and incubated in a thermomixer at 37°C, 300 RPM, for 25-30 min. During this time, the suspension was passed 10 times through a 18G syringe and 8 times a 21G syringe. The suspension was subsequently applied to tissue strainers (40 µm mesh size), 20 ml MACS buffer (PBS containing 2% FBS and 5 mM EDTA pH 8.0) was added, and the cells centrifuged for 5 min at 4° C, 320 x g. Fc-block (BD Biosciences Cat. no. BD553141, 1:200) was added for 17 min, followed by red blood cell lysis for 3 min on ice. MACS buffer was added to stop the lysis and the cells centrifuged for 10 min at 4°C, 300 x g and stained for FACS. Human kidney: Non-tumour kidney portions were collected into PBS on ice. Kidney samples were minced into 2-4mm and dissociated according to the manufacturer's protocol of Multi Tissue Dissociation Kit 1 (Miltenyi Biotech, Cat# 130-110-201). Red blood cells (RBC) were removed by incubating the sample with 2 ml RBC lysis buffer for 5 minutes on ice. The sample was enriched to live cells using Dead Cell Removal Kit (Miltenyi Biotech, Cat# 130-090-101), centrifuged at 300 x g for 10 minutes, resuspended in MACS buffer (PBS containing 2% FBS and 5 mM EDTA pH 8.0) and stained for FACS.
Instrument	FACS Symphony S6, ARIA III, and FUSION cell sorters (BD Biosciences).
Software	FACS was operated with FACSDiva 7 FACS data was analyzed with FlowJo 10.4.2.
Cell population abundance	Cells were sorted as single cells into 384-plates, processed and sequenced by the MARS-seq protocol. Small cells were discarded by a 200 detected molecule (UMI) cutoff.
Gating strategy	All gates were taken from the live cells (FSC-A/SSC-A). Relevant gates shown in Figure 1d and Extended data Figure 2c.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.