Stem Cell Reports, Volume *16*

Supplemental Information

The tumor suppressor WT1 drives progenitor cell progression and epithelialization to prevent Wilms tumorigenesis in human kidney organoids

Verena Waehle, Rosemarie Ungricht, Philipp S. Hoppe, and Joerg Betschinger

Waehle et al., Supplemental Figure 1

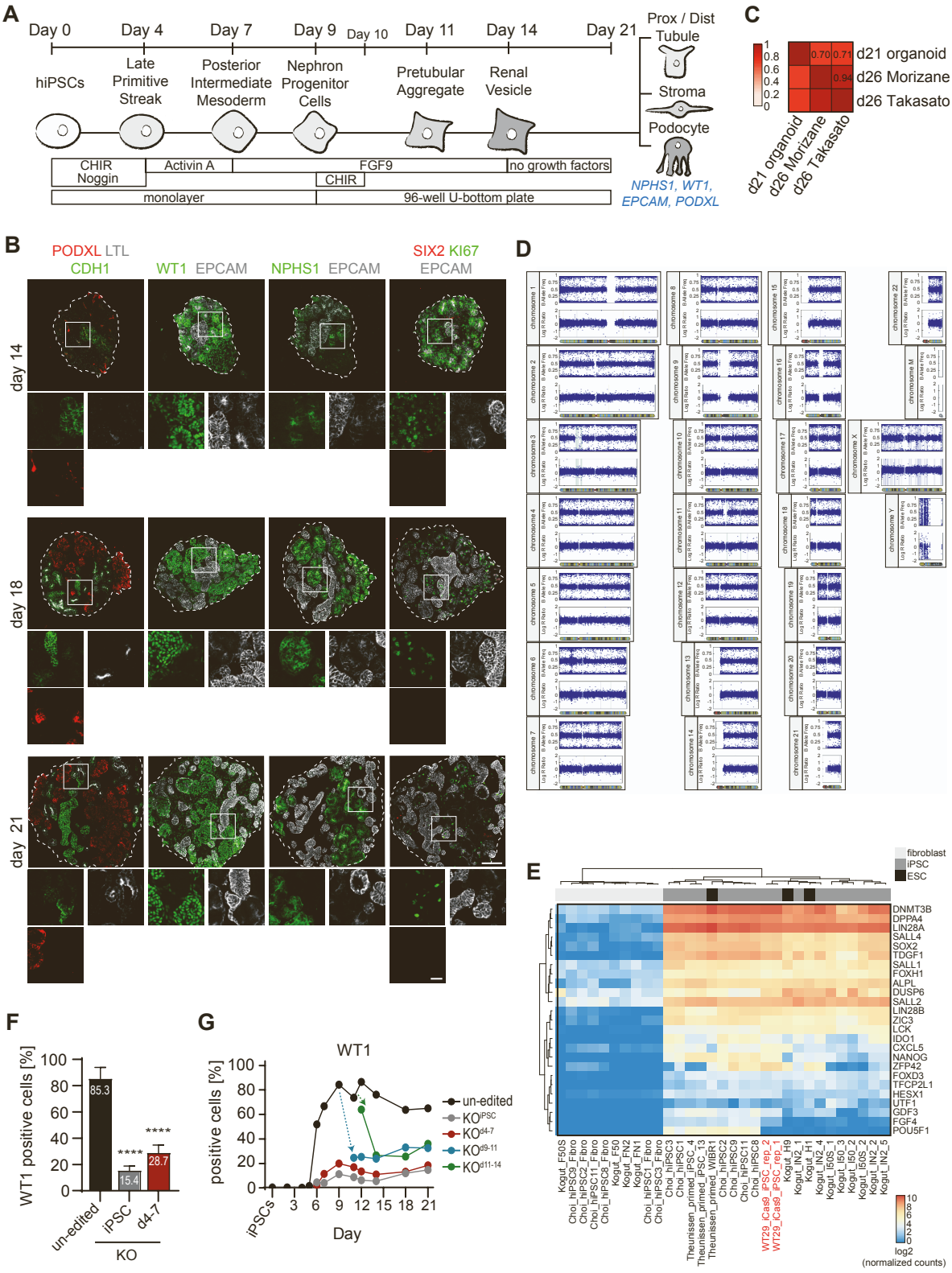

Figure S1: Related to Figure 1

A: Overview of the kidney organoid protocol adapted from (Morizane et al., 2015); Prox = proximal; Dist = distal; CHIR = CHIR99021.

B: Staining of the indicated markers in representative d14, d18 and d21 organoids (Scale bar: 100 µm). White boxed regions are shown as blow-ups (Scale bar: 25 µm).

C: Pearson correlation coefficients of pairwise comparisons between Log2 fold gene expression changes of indicated organoid samples relative to corresponding iPSC samples. Expression data for d26, Morizane and d26, Takasato are from (Wu et al., 2018), and the mean of 4 and 6 technical and biological replicates, respectively.

D: Karyogram of WT29-iCas9 iPSCs. Mean log R ratio (LRR) and B-allele frequency (BAF) are shown. Freq = Frequency

E: Unsupervised clustering of pluripotency marker gene expression in fibroblasts, human embryonic stem cells (ESCs) and iPSCs across different studies. The WT29-iCas9 iPSCs used in this work are highlighted in red.

F: Quantification of WT1-expressing cells in un-edited, KO^{iPSC} and KO^{d4-7} d9 NPCs by flow cytometry. Data presented as mean % of positive cells +/- SD from n=7 independent experiments.

G: Quantification of WT1-expressing cells by flow cytometry in un-edited, KO^{iPSC}, KO^{d4-7}, KO^{d9-11} cells and KO^{d11-14} organoids at the indicated time points, or in iPSCs (un-edited or KO^{iPSC}). Data is presented from one experiment. Arrows visualize reduction of WT1 protein expression upon induction of Cas9 from d9-11 and d11-14.

Waehle et al., Supplemental Figure 2

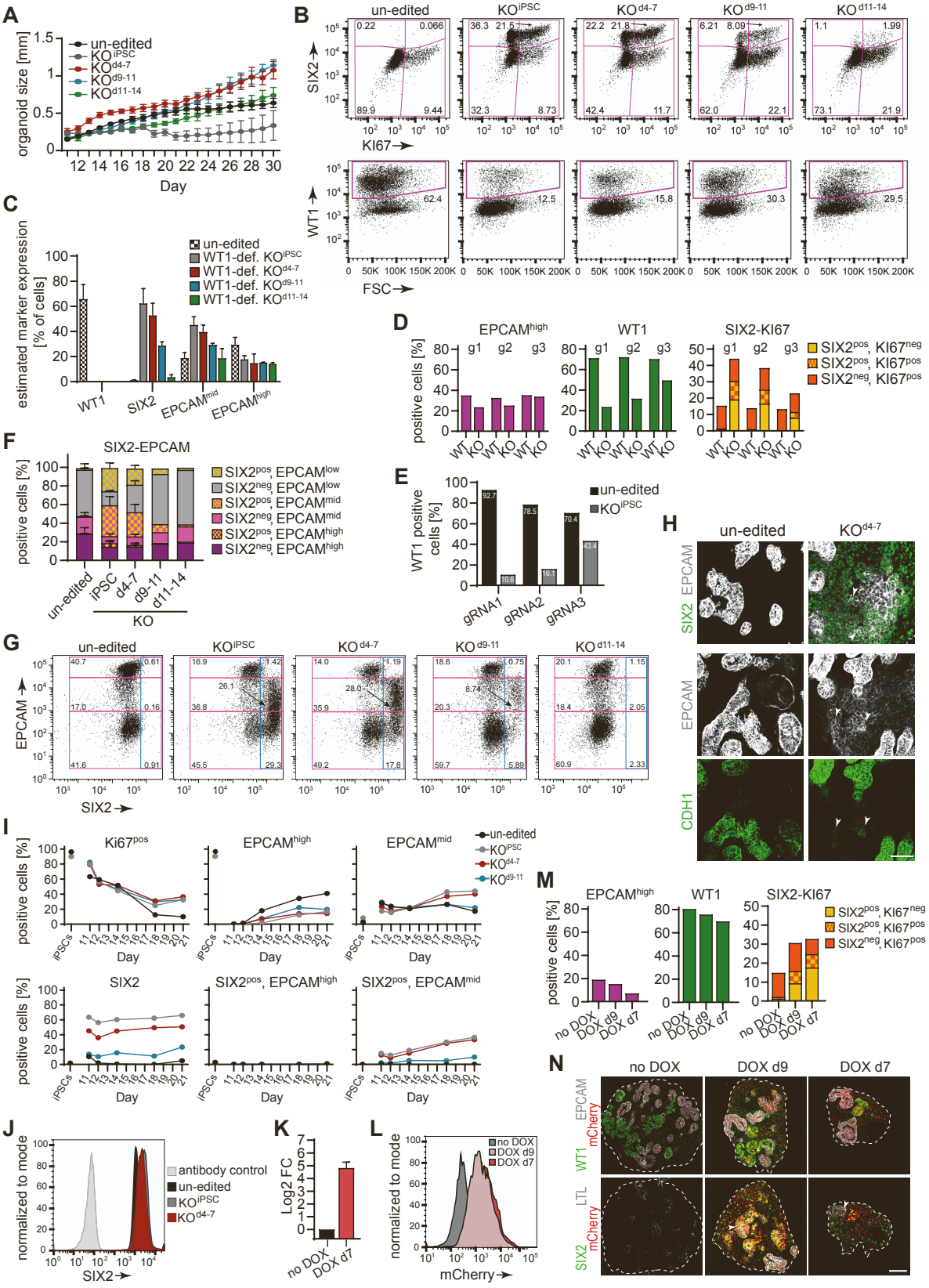

Figure S2: Related to Figure 1.

A: Growth of un-edited, KO^{iPSC}, KO^{d4-7}, KO^{d9-11} and KO^{d11-14} organoids between d11 and d30. Areas were derived from Incucyte® images and are presented as mean +/- SD for a minimum of n=10 organoids across all time points.

B: Representative flow cytometry dot plots showing SIX2 and KI67 staining (top) and WT1 staining (bottom) in d21 organoids as in **A**, and used for quantifications shown in **Figure 1A**. Pink boxes represent gates used for quantification and numbers indicate % of cells in the respective gates.

C: To gauge the cell type distribution of WT1-deficient and WT1-expressing cells in KO organoids, we estimated the percentage of WT1-deficient cells that express the indicated markers in KO^{iPSC}, KO^{d4-7}, KO^{d9-11} and KO^{d11-14} d21 organoids (see Supplemental Experimental Procedures) . Data is presented as estimated mean % of positive cells +/- SD in n=5 (un-edited, KO^{iPSC}, KO^{d4-7}) or n=2 (KO^{d9-11}, KO^{d11-14}) independent experiments. Values for un-edited d21 organoids are shown for reference and were taken from **Figure 1A,C**. WT1-def. = WT1-deficient cells

D: Quantification of subpopulations by flow cytometry of indicated markers in un-edited or KO^{iPSC} d21 organoids using indicated *WT1* gRNAs (g1, g2, g3). Data is presented from one experiment.

E: Quantification by flow cytometry of WT1-expressing cells in un-edited and KO^{iPSC} d9 NPCs using indicated *WT1* gRNAs. Data presented from one experiment.

F: Quantification of subpopulations by flow cytometry of indicated markers in d21 organoids as in **A**. Data is presented as mean % of positive cells +/- SD derived from n=5 (un-edited, KO^{iPSC}, KO^{d4-7}) and n=1 (KO^{d9-11}, KO^{d11-14}) independent experiments. Note that the SIX2-EPCAM quantification includes data presented in the d21 panel of **Figure 1C**. Two-sided student's t-test; *p*-value: ns >0.05; * ≤0.05; ** ≤0.01; *** ≤0.001; **** ≤0.0001.

G: Representative flow cytometry dot plots showing EPCAM and SIX2 staining in d21 organoids as in **A**. Pink boxes and associated numbers indicate % of cells that are $EPCAM^{high}$, $EPCAM^{mid}$ and $EPCAM^{low}$. Blue boxes and associated numbers indicate % of cells that are $SIX2^{pos}/EPCAM^{high}$, $SIX2^{pos}/EPCAM^{mid}$ or $SIX2^{pos}/EPCAM^{low}$.

H: Staining of indicated markers in representative un-edited and KO^{d4-7} d21 organoids. Images were enhanced to visualize EPCAM staining surrounding SIX2-positive cells (top) as well as overlap of EPCAM (middle) and CDH1 staining (bottom) in KO^{d4-7} organoids (white arrowheads). Note that images of SIX2-EPCAM staining are blow-ups derived from respective images in **Figure 1B**. Scale bar: 200 µm.

I: Quantification of subpopulations by flow cytometry of indicated markers in un-edited, KO^{iPSC}, KO^{d4-7} and KO^{d9-11} organoids at the indicated time points, or in iPSCs (un-edited or KO^{iPSC}). Data is presented from one experiment.

J: SIX2 expression in d9 NPCs by flow-cytometry.

K: qPCR analysis of *SIX2* transcription in d9 NPCs derived from mCherry::T2A::SIX2 iPSCs that were untreated (no DOX) or treated with DOX from d7 onwards. Data is presented as Log2 fold expression change relative to the untreated control, and shown as mean +/- SD derived from n=3 experiments.

L: mCherry expression in d21 organoids that were untreated (no DOX) or treated with DOX starting from the indicated time points to induce mCherry::T2A::SIX2 expression.

M: Quantification by flow cytometry of subpopulations expressing indicated markers in d21 organoids that were untreated (no DOX) or treated with DOX as in **L**. Data is shown as % of positive cells for one representative experiment.

N: Staining of indicated markers in representative d21 organoids that were untreated (no DOX) or treated with DOX as in **L**. mCherry represents induction of SIX2. White arrowheads indicate expression of SIX2 in LTL^{pos} tubules. Scale bar: 100 µm.

Waehle et al., Supplemental Figure 3

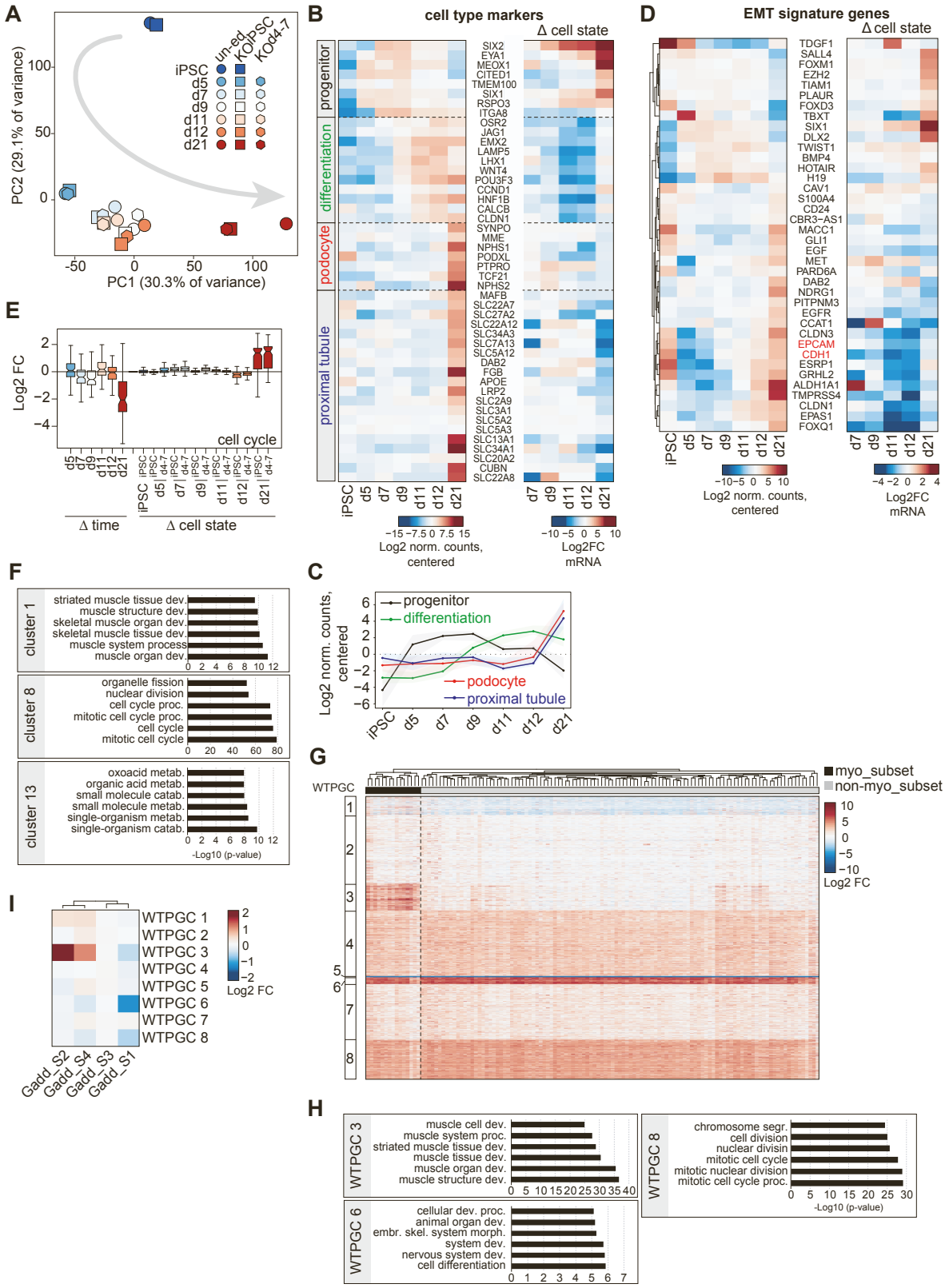

Figure S3: Related to Figures 2 and 3.

A: Principal component (PC) analysis of un-edited (un-ed.), KO^{iPSC} and $KO^{d4\text{-}7}$ samples at the indicated time points.

B: Heatmap of marker gene expression changes during organoid development (left) and in KO organoids (right). Note that Log2 fold changes relative to the mean in un-edited organoids at indicated time points (left) and Log2 fold changes of mean expression changes in KO^{iPSC} and $KO^{d4\text{-}7}$ organoids relative to un-edited control organoids at indicated time points (right) are shown. Cell types are indicated on the left.

C: Line representation of marker gene expression changes shown in **B**. Lines represent the medium and shades the lower and upper quartile.

D: Heatmap of gene expression changes of selected (see Supplemental Experimental Procedures) EMT marker genes as in **B**.

E: Quantification of Log2 fold expression changes of cell cycle genes and of contrasts as specified in **Figure 2A**. The cell cycle gene-set include G_1S- and G_2M-specific gene-sets (Liu et al., 2017).

F,H: The top six GO terms that are enriched in each of the indicated gene clusters is shown. catab. = catabolism; dev. = development; embr. = embryonic; metab. = metabolism; morph. = morphogenesis; proc. = process; segr. = segregation; skel. = skeletal.

G: k-means clustering of scaled Log2 fold expression changes in WT patients relative to control tissue of 2,100 genes (see Supplemental Experimental Procedures). Unsupervised clustering of patients identified a myogenic subset (myo_subset) in which Wilms tumor patient gene cluster (WTPGC) 3 is strongly induced.

I: Heatmap representation of mean-centered scaled Log2 fold expression changes of WTPGCs 1-8 in patient subsets defined by (Gadd et al., 2012).

Waehle et al., Supplemental Figure 4

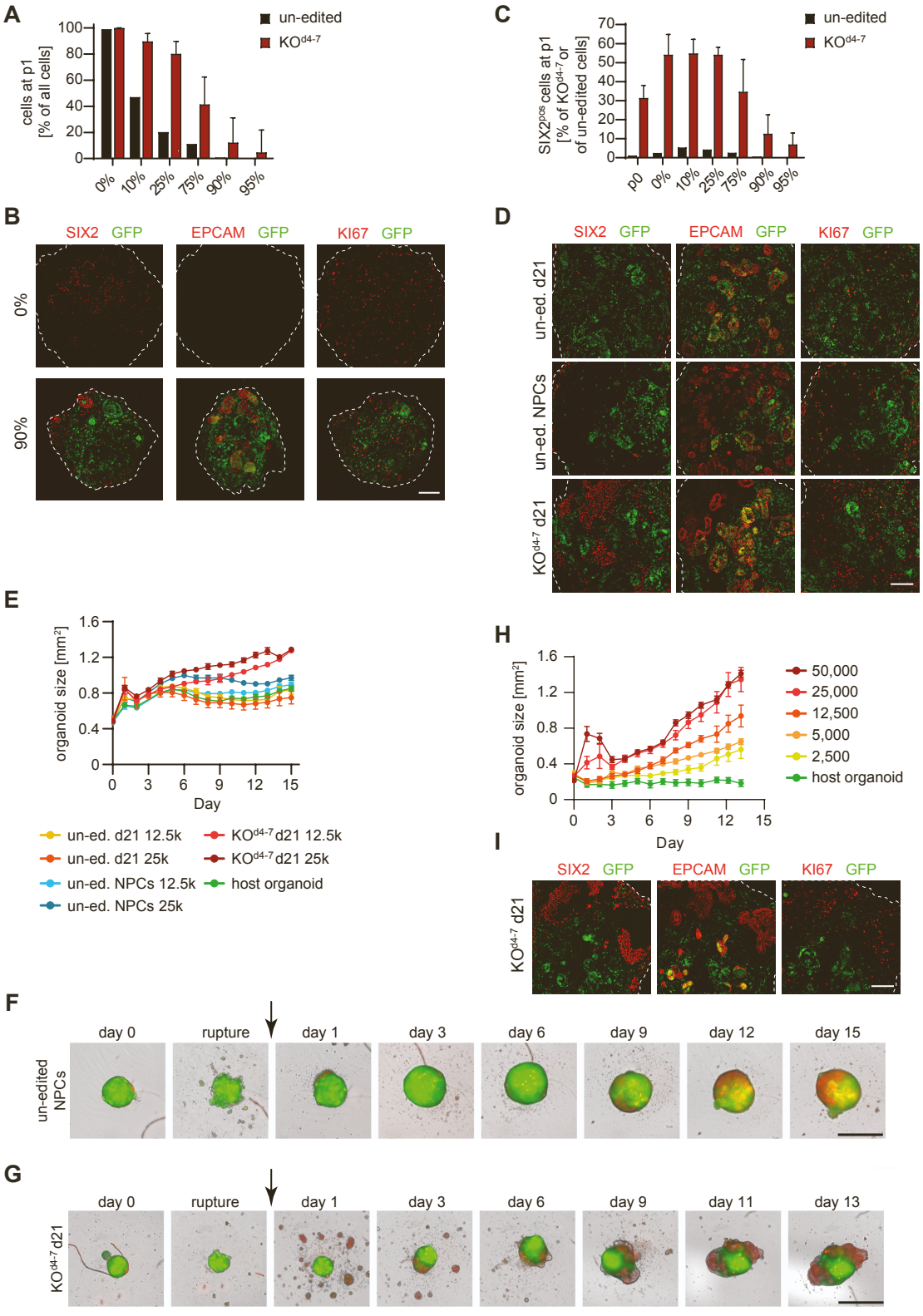

Figure S4: Related to Figure 4.

A,**C:** Quantification of, **A** all, and **C** SIX2^{pos} un-edited and KO^{d4-7} cells in chimeric organoids after passaging in the presence of wild-type GFP-expressing d9 NPCs at indicated ratios after passage 1. p0 indicates KO^{d4-7} d21 organoids that were used as starting material for passaging. Data is shown as mean % +/- SD for n=2 independent experiments. **A**: Percentages are relative to all cells of chimeric organoids (% of all cells). **C**: Percentages are relative to all mutant cells (% of KO^{d4-7} cells), and relative to all un-edited control cells (% of un-edited cells). In both cases, the GFP-expressing cells are excluded from the analysis.

B,**D**,**I:** Staining of the indicated markers, **B** at passage 3 and indicated mixing ratios, **D** 15 d after adding 12,500 RFP-expressing cells from un-edited d21 organoids (un-ed. d21, top), un-edited d9 NPCs (un-ed. NPCs, middle), or KO^{d4-7} d21 organoids (KO^{d4-7} d21, bottom) to ruptured wild-type GFP-expressing organoids in the presence of FGF9/CHIR, and **I** after adding 12,500 RFP-expressing cells from KO^{d4-7} d21 organoids to ruptured wild-type GFP-expressing organoids in the absence of growth factors. Scale bar: 100 µm.

E,**H:** Growth of organoids, **E** after adding 12,500 (12.5k) or 25,000 (25k) RFP-expressing cells to ruptured wild-type GFP organoids in the presence of FGF9/CHIR as specified in **D**, and **H** after adding the indicated numbers of RFP-expressing KO^{d4-7} d21 organoid cells to ruptured wild-type GFP organoids in the absence of growth factors. Areas were calculated from Incucyte® images and are presented as mean +/- standard error of the mean (SEM) for n=8 organoids per condition.

F,**G:** Images of organoids obtained after adding, **F** 12,500 RFP-expressing un-edited d9 NPCs to ruptured wild-type GFP organoids in the presence of FGF9/CHIR, and **G** 12,500 RFP-expressing cells from KO^{d4-7} d21 organoids to ruptured wild-type GFP organoids in the absence of growth factors. Images were recorded at the indicated time points using an Incucyte® system. Scale bar 1 mm.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Human iPSC culture

WT29 and WT29-iCas9 iPSCs (Ungricht et al., 2021) and their derivatives were cultured on Laminin (Biolaminin 521 LN; Biolamina #LN521) in mTeSR1 (Stem Cell Technologies; # 85850) plus 1% Penicillin-Streptomycin (Thermo Fisher #15140122). For passaging, cells were detached using TrypLE Express (Thermo Fisher #12604013), and single-cell suspensions were plated in mTeSR1 supplemented with 2 µM ROCK inhibitor (Y-27632 Dihydrochloride Tocris #1254).

Karyotyping

Using the service of Life&Brain Genomics (https://www.lifeandbrain.com/en/products-services/lb-genomics/), genomic DNA of WT29-iCas9 iPSCs was hybridized to the HumanOmni_EUR_HD_20021594_BeadChip array (Illumina), scanned on an Illumina iScan and analyzed using GenomeStudio V2.0.2 and CNV-Partition. The Karyogram is provided in Figure S1D.

Kidney organoid protocol

Adherent differentiation (d0-d9): 50'000-60'000 hiPSCs/cm^2 were plated in mTeSR1 supplemented with 2 µM ROCK inhibitor into Laminin-coated 6-well plates. After at least 6 h, medium was removed, cells gently washed with Dulbecco's Phosphate Buffered Saline (PBS; without magnesium and calcium; Thermo Fisher #14190169), and differentiation induced by adding basic differentiation medium (BDM; advanced RPMI 1640 (Thermo Fisher #12633012), 1% Glutamax (Thermo Fisher #35050038), 1% Penicillin-Streptomycin) supplemented with 8 µM CHIR99021 (Tocris #4423) and 5 ng/ml Noggin (Peprotech #120-10C) (= d0). Medium was changed after 2d. At d4, cells were visually inspected for the presence of contracting colonies with bright halo-like outlines, followed by a medium change to BDM containing 10 ng/ml Activin A (R&D Systems #338_AC). At d7, medium was changed to BDM supplemented with 10 ng/ml FGF9 (R&D Systems #273-F9). At d9, corresponding to the NPC state, cells were washed, dissociated from the cell culture dish using TrypLE and counted using a Vi-CELL™ XR Cell Viability Analyzer (Beckmann).

Organoid differentiation in suspension culture (d9 onwards): 25'000 or 50'000 cells were seeded in 150 µl of BDM containing 3 µM CHIR99021, 10 ng/ml FGF9 and 2 µM ROCK inhibitor into Corning® Costar® Ultra-Low Attachment 96-well round bottom plates (Sigma #CLS7007-24EA). Surplus NPCs were frozen in CryoStor® CS10 (Stem Cell Technologies #7930). To induce aggregation, the plate was briefly centrifuged at 90 g for 3 minutes (min). At d10 (d1 in suspension), 100 µl of medium was replaced with 150 µl of BDM plus 10 ng/ml FGF9. At d11 (d2), 100 µl of medium was replaced with 100 µl of BDM plus 10 ng/ml FGF9. At d14 (d5), d16 (d7) and d18 (d9), 100 µl of medium was replaced with 100 µl of BDM without growth factors. During culture beyond d21, medium was changed 3 times per week.

Throughout the study we compared the mutant edited organoids with wild-type un-edited control organoids that were generated side-by-side and analyzed at the same differentiation time points.

Organoid growth was determined using an Incucyte® system: Brightfield images were recorded every 24h, or as indicated, and organoid sizes were calculated. Object recognition parameters were manually defined for each experiment, and the detected objects validated.

Generation of single cell suspensions for flow cytometry or passaging of organoids

To generate single cell suspensions for flow cytometry or passaging, organoids were transferred into tubes using a cut P-1000 tip, and washed twice with PBS. A 1:1 mix of non-enzymatic cell dissociation solution (Thermo Fisher #13151014) and 0.25% Trypsin-EDTA (Thermo Fisher #25200056) was added for 10 mins at 37°C, and organoids were dissociated by pipetting up and down ten times. Trypsin was inactivated by adding 10% fetal calf serum (FCS; Bioconcept #2-01F36-I) in PBS, washed with 1% FCS and passed through a 50 µm filter (BD Biosciences #340632).

Generation and culture of chimeric organoids

For mixing with NPCs (Figure 4B,C, S4A-C), single cell suspensions of RFP-expressing un-edited or KO^{d4-7} d21 organoids were generated as described above, aggregated with freshly thawed GFP-expressing WT29-iCas9 (Ungricht et al., 2021) d9 NPCs at indicated ratios to a total of 50,000 cells, and plated into Ultra-Low Attachment 96-well round bottom plates in 150 µl BDM, supplemented with 3 µM CHIR99021, 10 ng/ml FGF9 and 2 µM ROCK inhibitor. Culture was continued as detailed in «Organoid differentiation in suspension culture» above. At d12 of suspension culture, organoids were dissociated into single cell suspensions that were used for flow cytometry analysis and for passaging by aggregating with freshly thawed GFP-expressing WT29-iCas9 d9 NPCs at respective ratios and subjecting to suspension culture. This was repeated for up to four passages. After each passage, a minimum of six organoids was processed for cryosectioning and immunofluorescence staining, as described below.

For mixing into d21 organoids (Figure 4D, S4E-I), GFP-expressing WT29-iCas9 d21 organoids were mechanically ruptured by pipetting them up and down five times in a P-200 tip. RFP-expressing un-edited and KO^{d4-7} d21 organoids were dissociated into single cell suspensions as described above, and un-edited RFP-expressing d9 NPCs were freshly thawed. After cell counting, indicated cell numbers were added to ruptured organoids, plates briefly spun to induce aggregation, and culture in BDM resumed as indicated, either in the presence or absence of 3 µM CHIR99021 and 10 ng/ml FGF9 according to «Organoid differentiation in suspension culture» above.

Molecular biology

The coding sequence of SIX2 was amplified from human iPSC cDNA, T2A sequences and Gateway cloning sites added by polymerase chain reaction (PCR), and recombined into pDONR221 using Gateway technology (Thermo Fisher #11789020 and #11791020). The expression vector was generated by recombining with a PiggyBac pPB-TRE-mCherry-DEST-rTA-HSV-neo expression destination vector. *WT1* gRNA-encoding vectors are derived from pRSI16-U6-sh-UbiC-TagRFP-2A-Puro (Cellecta #SVSHU616-L).

Oligonucleotide sequences:

SIX2 (gateway)	attB1-fwd	ggggacaagtttgtacaaaaaagcaggcttcaccATGTCCATGCTGCCCACCTTCG
	attB2-rev	ggggaccactttgtacaagaaagctgggtcCTAGGAGCCCAGGTCCACGAGG
T2A sequence	fwd	GAGGGCAGAGGAAGTCTCCTAACATGCGGTGACGTGGAGGAGAATCCTGGCCCA

fwd = forward sequence; rev = reverse sequence

gRNA sequences:

WT1	#1	TGTGTTTGCAGCCACAGCAC
WT1	#2	GGTGTGGCAGCCATAGACCG
WT1	#3	GCTGCCGGTGCAGCTGTCGG

Generation of transgenic WT29 iPSCs for SIX2 overexpression

Transgenic WT29 iPSCs were generated by transfecting PiggyBac expression vector and pBase (Villegas et al., 2019) using Lipofectamin Stem™ (Thermo Fisher #STEM00015) in OptiMem Reduced Serum Medium (Thermo Fisher #319850629), and selected for stable integration in the presence of 100 µg/ml G418 (Thermo Fisher #10131027). Inducible cells were further selected by exposing to 1 µg/ml Doxycycline (DOX; Clonetech #631311) for 48 h and purifying the 30% of cells with mCherry expression closest to the median of the population. Cell sorting was performed on a BD FACSAria™ Fusion Cell Sorter.

Lentivirus production and human iPSC *transduction*

Lentiviruses were produced in HEK293T cells. Prior to transfection, HEK293T cells were seeded onto collagen I-coated 6-well tissue culture plates (BD biosciences #346400) in packaging medium (DMEM (Thermo Fisher #11965), supplemented with 10% FCS and 1% Non-essential amino acids (Thermo Fisher #11140050)). The next day, cells were transfected with *WT1* gRNA-encoding vectors and Cellecta packaging mix (Cellecta #CPC-K2A) using the TransIT™293 transfection reagent (Mirus Bio #MIR 2700) in OptiMem Reduced Serum Medium (Thermo Fisher #31985062). The next day, medium was changed to 1 ml of packaging medium. After 3d, the virus-containing supernatant was collected, filtered through a 50 µm filter and stored at -80°C.

For virus titration, WT29-iCas9 cells were seeded into 6-well plates (200,000 cells / well) in mTeSR1 medium supplemented with ROCK inhibitor. After 7 hours (h), different volumes of viral supernatant were added to the cells. After 3d with daily mTeSR1 medium changes, cells were detached and RFP fluorescence was measured by flow cytometry. Based on this titration, WT29-iCas9 cells were transduced at a multiplicity of infection of 0.5, and infected cells were selected with puromycin (Thermo Fisher #A11138-03) for 6d.

RNA isolation, cDNA synthesis, qPCR

RNA was isolated using RNeasy Mini Kits (Qiagen #74104) and RNase-Free-DNase Sets (Qiagen #79256) according to the manufacturer's instructions, and concentration determined using a NanoDrop (Thermo Fisher).

cDNA was generated from at least 400 ng of total RNA using SuperScript III Reverse Transcriptase (Thermo Fisher # 18080044) using oligodT priming. 25 ng of cDNA was subjected to qPCR on a Step One Plus™ Real-Time PCR System (Thermo Fisher) using the TaqMan Fast Universal PCR Master Mix (Thermo Fisher # 4364103). Expression of SIX2 was quantified in technical duplicates using Universal Probe Library (UPL, Roche) Probe 88 together with custom-designed primer pairs: Fwd: ggcaagtcggtgttaggc, Rev: ggctggatgatgagtggtct, and multiplexing with a *GAPDH* probe (Thermo Fisher # 4326317E) . Cycle threshold (CT) values were normalized to *GAPDH* and to controls (ΔΔCT method).

Flow cytometry of mCherry-expression

Single live cell suspensions were washed with 1% FCS in PBS and resuspended in flow cytometry buffer (2% FCS and 1 mM ethylenediaminetetraacetic acid (EDTA; Thermo Fisher #AM9260G) in PBS). Flow cytometry was performed on a BD LSRFortessa™ and data analyzed using FlowJo™ software.

Antibodies for flow cytometry of fixed cells

EPCAM-Alexa Fluor (AF) 647 (Abcam #ab239273, 1:200); KI67-FITC (eBisoscience #11-5698-82, 1:200); SIX2 (Proteintech #11562-1-AP, 1:100); WT1 (Abcam #ab89901, 1:200); donkey-anti-rabbit AF488 (Thermo Fisher #A-21206, 1:500); donkey-anti-rabbit AF 647 (Thermo Fisher #A-31573, 1:500).

Estimating RFP expression in chimeric organoids

RFP-expression in fixed *WT1* KO cells of chimeric organoids (Figure 4B, S4A,B) by flow cytometry was ambiguous. We therefore determined the fraction of mutant cells in chimeric organoids by using opposing GFP-expression in wildtype cells. Heterogeneity of GFP expression was corrected for by comparing to the GFP distribution in GFP-expressing WT29-iCas9 host cells.

Immunofluorescence staining of Cryosections

Cryosections were washed with PBS for 10 mins at room temperature (RT). When using biotinylated LTL, slides were incubated with blocking/permeabilization buffer (1% BSA and 0.2% Triton X-100 in PBS) for 15 mins, then with blocking/permeabilization buffer containing four drops per ml of Streptavidin Block solution (Streptavidin/Biotin Blocking Kit; Vectorlabs #SP-2002) for 15 mins, and then with blocking/permeabilization buffer containing four drops per ml of Biotin Block solution for 15 mins. When using un-biotinylated primary antibodies, slides were instead incubated with blocking/permeabilization buffer for 30 mins at RT. After a quick wash in PBS, slides were incubated with primary antibodies in 1% BSA in PBS for 1 h at RT. Afterwards, slides were washed twice with PBS for each 10 mins, incubated with secondary antibodies and Hoechst 33342 diluted into 1% BSA in PBS for 1 h at RT, and after two additional washes with PBS for 10 mins each, mounted in ProLong™ Diamond Antifade Mountant (Thermo Fisher #P36970).

Antibodies were: CDH1 (BD Biosciences #610181, 1:200 and Thermo Fisher #MA514458, 1:100); EPCAM-AF647 (Abcam #ab239273, 1:200); Hoechst 33342 (Thermo Fisher #H3570; 1:10000); KI67 (SolA15, eBisoscience #14-5698-82, 1:200) (KI67 / KI67-FITC (SolA15, eBisoscience #14-5698-82, 1:200); LTL-Biotinylated (Vectorlabs B-1325, 1:500); NPHS1 (R&D Systems #AF4269, 1:60); PODXL (R&D Systems #AF1658, 1:500); SIX2 (Proteintech #11562-1-AP, 1:100); WT1 (Abcam #ab89901; 1:200); donkey-anti-rabbit AF 488 (Thermo Fisher #A-21206, 1:500); donkey-anti-rabbit AF 647 (Thermo Fisher #A-31573, 1:500); donkey-anti-goat AF 594 (Thermo Fisher #A-11058, 1:500); donkey-anti-goat AF 647 (Thermo Fisher #A-21447, 1:500); donkey-anti-sheep AF 488 (Thermo Fisher # A-11015, 1:500); donkey-anti-mouse AF 488 (Thermo Fisher #A-21202, 1:500); donkey-anti-rat AF 647 (Abcam #150155, 1:500); Streptavidin Fluorescent Dye 633-I (Abnova #U0295, 1:500).

Estimating the cell type distribution of WT1-deficient cells in KO organoids

In mosaic KO organoids, the observed percentage of cells expressing a particular cell type marker $\boldsymbol{p}_{marker}(mosaic\ organoid)$ equals $\boldsymbol{wt}_{marker} \cdot \boldsymbol{p}_{marker}(wildtype) + (1 - \boldsymbol{wt}_{marker}) \cdot \boldsymbol{p}_{marker}(KO)$, where $\boldsymbol{wt}_{marker}$ is the fraction of WT1-positive wild-type cells in mosaic organoids expressing this marker, and $\boldsymbol{p}_{marker}(wildtype \text{ or } KO)$ is the percentage of all wild-type or WT1-deficient (KO) cells in mosaic organoids expressing this marker.

To estimate $\boldsymbol{p}_{marker}(KO)$ (Figure S2C), we assume: (1) that $\boldsymbol{wt}_{marker}$ is invariable across cell types and equals the fraction of wild-type cells in podocytes defined by WT1 staining in mosaic and corresponding un-edited control organoids: $\boldsymbol{wt}_{marker} \approx \frac{\boldsymbol{p}_{WT1}\,(mosaic\ organoid)}{\boldsymbol{p}_{WT1}\,(unedited\ organoid)}$ (Figure 1A); and (2) that the propensity for differentiation towards a particular cell type is a cell-intrinsic

characteristic, such that $\boldsymbol{p}_{marker}(wildtype)$ equals the cell type distribution in un-edited control organoids: $\boldsymbol{p}_{marker}(wildtype) \approx \mathbf{p}_{\text{marker}}(\text{unedited organoid})$ (Figure 1A,C).

$$\boldsymbol{p}_{marker}(KO) \approx \frac{\boldsymbol{p}_{marker}(mosaic\ organoid) - \frac{\boldsymbol{p}_{WT1}(mosaic\ organoid)}{\boldsymbol{p}_{WT1}(unedited\ organoid)} \cdot \boldsymbol{p}_{marker}(unedited\ organoid)}{1 - \frac{\boldsymbol{p}_{WT1}(mosaic\ organoid)}{\boldsymbol{p}_{WT1}(unedited\ organoid)}}$$

Bioinformatics

RNA isolation of three independent biological replicates was performed as described above and RNA-seq libraries prepared using the TruSeq mRNA Library preparation kit (Illumina #20020595). RNA sequencing was performed on an Illumina HiSeq2500 machine (50 bp single-end reads). RNA-seq reads were aligned to the human hg38 genome using *qAlign* from the Bioconductor package QuasR (Gaidatzis et al., 2015) with default parameters except for *aligner="Rhisat2"* and *splicedAlignment=TRUE*. Alignments were quantified for known UCSC genes obtained from the TxDb.Hsapiens.UCSC.hg38.knownGene package using *qCount* (Table S1). Raw counts were further divided by the total number of counts in all genes in each sample , and multiplied with the minimum total counts across samples in order to normalize for different library sizes.

To evaluate selected pluripotent marker gene expression in pluripotent stem cells, the RNAseq data of the samples described in Figure S1E were downloaded from the datasets GSE82765 (Theunissen et al., 2016), GSE72311 (Choi et al., 2015) and GSE97265 (Kogut et al., 2018). Gene expression levels were determined as described above, normalized and Log2 transformed with a pseudocount of 1.

Principal component analysis (PCA) (Figure S3A) using normalized read counts of merged replicates and considering the top 30% variable genes was performed using the prcomp function in R.

Differential gene expression was determined using edgeR (Robinson and Oshlack, 2010). For heatmap visualization (Figure 2A), 7'626 genes (Table S1) were considered that were significantly regulated during control organoid formation (absolute Log2 fold gene expression change at d5, d7, d9, d11, d12 or d21 relative to iPSCs greater than Log2(3) with a false discovery rate (FDR) smaller than 0.001) or upon *WT1* KO (absolute Log2 fold expression changes in KO^{iPSC} or $KO^{d4\text{-}7}$ organoids relative to control organoids at any time point greater than Log2(3) with a FDR smaller than 0.001).

For comparison with published kidney organoid gene expression (Wu et al., 2018) (GSE118184) (Figure S1C), scRNAseq reads of 218 iPSCs and of 25120 (Morizane protocol) and 82024 (Takasato protocol) cells from d26 organoids were summed and normalized, and Log2 fold expression changes relative to iPSCs were calculated using a pseudocount of 1. Correlation coefficients are based on 12817 genes detected in (Wu et al., 2018) and the RNA-seq dataset reported in this work. Pearson correlation coefficients were calculated using R's *cor* function.

For generation of gene-set overlap scores (Figure 2C), we first calculated the odds-ratio for each pairwise comparison between embryonic kidney gene sets (Lindström et al., 2018) and gene clusters 1-17, using the *fisher.test* function in R. To correct for biases introduced by different gene-set sizes, each observed odd-ratio was then normalized by calculating a Z-score: $z\_score = \frac{obs - mean\_rand}{sd\_rand}$, where obs is the observed odds-ratio for a given pairwise comparison, and mean_rand and sd_rand are the mean and standard-deviation of 100 randomized odd-ratios, obtained from equal-sized sets of randomly selected genes.

Analyses of enriched gene sets (Figure S3F,H and Table S1,S2) was performed using DAVID (Huang et al., 2008) and selecting GOTERM_BP_ALL.

Single-cell RNAseq (scRNAseq) datasets of week 16 fetal kidney scRNAseq (Hochane et al., 2019) (GSE114530) were integrated with and contrasted to the results of this study. For the analysis and visualization of the scRNAseq data (Figures 2D) we filtered out genes detected in < 1% of the cells as well as the abundantly expressed and noisy ribosomal protein genes. From the remaining fraction, only the top 5% overdispersed genes were selected as input for the downstream dimensionality reduction and dataset integration methods according to a mean-variance trend fit using a semi-parametric approach (Zheng et al., 2017).The coordinates of the first 32-principal components were used to obtain the 2D tSNE representation of the data. For visualization purposes per-cell gene expression values were subjected to k-nearest neighbor smoothing (k=64) and normalized against a random set of 2000 detected genes in order to control for artefactual expression gradients.

For transcriptional changes in kidney cancer patients (Figure 3, S3G-I), Wilms tumor (WT; Gadd et al., 2017), Kidney Chromophobe Carcinoma (KIRC; TGCA; Davis et al., 2014), Kidney Papillary Cell Carcinoma (KIRP; The Cancer Genome Atlas Research Network, 2016), and Kidney Clear Cell Carcinoma (KIRC; TCGA; Creighton et al., 2013) datasets (TARGET-WT, TCGA-KICH, TCGA-KIRP, and TCGA-KIRC) were downloaded from GDC (https://portal.gdc.cancer.gov) using the TCGAbiolinks package available from Bioconductor. Data sets were normalized, and Log2 fold expression changes for each sample calculated over the mean of the respective control samples using a pseudocount of 1. Wilms tumor expression affymetrix data (Gadd et al., 2012) (GSE31403) were downloaded using *getGEO* from the Bioconductor package GEOquery, Log2 transformed, in case of multiple probe sets per gene the probe set with the highest average value across samples selected, and normalized to the mean expression in all patients (mean-centered). Only patients assigned to subsets S1 – S4 (Gadd et al., 2012) in this dataset were considered.

Wherever indicated, Log2 fold gene expression changes were scaled using the *scale* function with the parameters *center = FALSE* and *scale = TRUE* (scaled Log2 fold expression changes) to allow comparison between datasets from different studies.

Pairwise Pearson correlation coefficients in Figure 3B were calculated using R's *cor* function, and are of 24920 genes and between scaled mean Log2 fold expression changes in d21 KO^{iPSC} and $KO^{d4\text{-}7}$ organoids and scaled mean Log2 fold expression changes in TARGET-WT, TCGA-KICH, TCGA-KIRP and TCGA-KIRC patients.

For heatmap visualization in Figure S3G, 2100 genes with a scaled Log2 fold expression change of greater than 3 in at least one of the patients were considered (Table S2).

Cell type markers (Figure S3B) were selected from the literature, while EMT signature genes (Figure S3D) are the 38 genes in the EMTome database (Vasaikar et al., 2021) that have literature support (Vasaikar et al., 2021) and showed an absolute Log2 fold expression change between not-edited d21 and d9 organoids of greater than 2.

Unsupervised clusterings in Figures S1E (Log2 normalized counts), 3A (Log2 fold expression changes), and 3B,C,D (scaled Log2 fold expression changes) were performed using the *aheatmap* function from the Bioconductor package NMF, and in Figures 3D, S3G (scaled Log2 fold expression changes) and 3E (Log2 fold expression changes) using the Heatmap function from the Bioconductor package ComplexHeatmap. We named the patient cluster identified by unsupervised clustering in Figure S3G "myo_subset" as it correlated with strong expression of WTPGC 3 genes that are enriched for muscle genes (Figure 3D) and muscle terms (Figure S3H).

SUPPLEMENTAL REFERENCES

Choi, J., Lee, S., Mallard, W., Clement, K., Tagliazucchi, G.M., Lim, H., Choi, I.Y., Ferrari, F., Tsankov, A.M., Pop, R., et al. (2015). A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. Nat Biotechnol *33*, 1173–1181.

Gadd, S., Huff, V., Huang, C.-C., Ruteshouser, E.C., Dome, J.S., Grundy, P.E., Breslow, N., Jennings, L., Green, D.M., Beckwith, J.B., et al. (2012). Clinically Relevant Subsets Identified by Gene Expression Patterns Support a Revised Ontogenic Model of Wilms Tumor: A Children's Oncology Group Study. Neoplasia *14*, 742-756.

Gaidatzis, D., Lerch, A., Hahne, F., and Stadler, M.B. (2015). QuasR: quantification and annotation of short reads in R. Bioinformatics *31*, 1130–1132.

Hochane, M., Berg, P.R. van den, Fan, X., Bérenger-Currias, N., Adegeest, E., Bialecka, M., Nieveen, M., Menschaart, M., Lopes, S.M.C. de S., and Semrau, S. (2019). Single-cell transcriptomics reveals gene expression dynamics of human fetal kidney development. Plos Biol *17*, e3000152.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res *37*, 1–13.

Kogut, I., McCarthy, S.M., Pavlova, M., Astling, D.P., Chen, X., Jakimenko, A., Jones, K.L., Getahun, A., Cambier, J.C., Pasmooij, A.M.G., et al. (2018). High-efficiency RNA-based reprogramming of human primary fibroblasts. Nat Commun *9*, 745.

Lindström, N.O., Brandine, G.D.S., Tran, T., Ransick, A., Suh, G., Guo, J., Kim, A.D., Parvez, R.K., Ruffins, S.W., Rutledge, E.A., et al. (2018). Progressive Recruitment of Mesenchymal Progenitors Reveals a Time-Dependent Process of Cell Fate Acquisition in Mouse and Human Nephrogenesis. Dev Cell *45*, 651-660.e4.

Liu, Z., Lou, H., Xie, K., Wang, H., Chen, N., Aparicio, O.M., Zhang, M.Q., Jiang, R., and Chen, T. (2017). Reconstructing cell cycle pseudo time-series via single-cell transcriptome data. Nat Commun *8*, 22.

Morizane, R., Lam, A.Q., Freedman, B.S., Kishi, S., Valerius, M.T., and Bonventre, J.V. (2015). Nephron organoids derived from human pluripotent stem cells model kidney development and injury. Nat Biotechnol *33*, 1193–1200.

Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol *11*, R25.

Theunissen, T.W., Friedli, M., He, Y., Planet, E., O'Neil, R.C., Markoulaki, S., Pontis, J., Wang, H., Iouranova, A., Imbeault, M., et al. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. Cell Stem Cell *19*, 502–515.

Ungricht, R., Guibbal, L., Lasbennes, M.-C., Orsini, V., Beibel, M., Waldt, A., Cuttat, R., Carbone, W., Basler, A., Roma, G., et al. (2021). Genome-wide screening in human kidney

organoids identifies novel aspects of nephrogenesis. bioRxiv, doi: 10.1101/2021.05.26.445745

Vasaikar, S.V., Deshmukh, A.P., Hollander, P. den, Addanki, S., Kuburich, N.A., Kudaravalli, S., Joseph, R., Chang, J.T., Soundararajan, R., and Mani, S.A. (2021). EMTome: a resource for pan-cancer analysis of epithelial-mesenchymal transition genes and signatures. Brit J Cancer *124*, 259–269.

Villegas, F., Lehalle, D., Mayer, D., Rittirsch, M., Stadler, M.B., Zinner, M., Olivieri, D., Vabres, P., Duplomb-Jego, L., Bont, E.S.J.M.D., et al. (2019). Lysosomal Signaling Licenses Embryonic Stem Cell Differentiation via Inactivation of Tfe3. Cell Stem Cell *24*, 257-270.e8.

Wu, H., Uchimura, K., Donnelly, E.L., Kirita, Y., Morris, S.A., and Humphreys, B.D. (2018). Comparative Analysis and Refinement of Human PSC-Derived Kidney Organoid Differentiation with Single-Cell Transcriptomics. Cell Stem Cell *23*, 869-881.e8.

Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat Commun *8*, 14049.