

Spatial single cell transcriptomic analysis informs tumor developmental hierarchy of DICER1 syndrome related sarcoma

Received: 13 January 2025

Accepted: 9 March 2026

Published online: 30 March 2026

 Check for updates

Felix K. F. Kommoss^{1,2,17}, Joyce Yu Han Zhang^{1,3,17}, Branden J. Lynch^{1,3,17}, Shary Yuting Chen^{1,3}, Janine Senz^{1,3}, Yana Moscovitz^{3,4}, Lesley A. Hill⁵, Ding Ma^{1,3}, R. Wilder Scott⁶, Jonathan Bush^{1,7}, Kenneth S. Chen⁸, Andrew Roth^{1,3}, Andreas von Deimling^{9,10}, William D. Foulkes^{11,12,13}, Gregg B. Morin^{4,14}, T. Michael Underhill⁵, Yemin Wang^{1,3,15,18} ✉ & David G. Huntsman^{1,3,16,18} ✉

DICER1-related tumor predisposition (DRTP), also known as DICER1 syndrome, encompasses a spectrum of malignancies mainly in children and young adults. Most are sarcomas, exhibiting histological and molecular similarities regardless of their anatomical origins, and only express the RNase IIIb domain-defective DICER1. To uncover their cellular origin and developmental hierarchy, we establish a lineage-traceable genetically engineered mouse model with controlled activation of hemizygous *Dicer1* RNase IIIb mutation in *Hic1*⁺ mesenchymal stromal cells. This causes renal tumors closely mirroring the developmental continuum of human DRTP sarcoma histologically and molecularly. Spatial single-cell transcriptomic analysis reveals a *Hic1*⁺*Pdgfra*⁺*Dpt*⁺*Pi16*⁺ fibroblastic progenitor population, corresponding to universal fibroblasts subjacent to transitional epithelium of renal collecting ducts, that can undergo rhabdomyoblastic differentiation or become proliferative sarcomatous cells. Investigation of patient samples identifies analogous cell states and developmental trajectories. This study uncovers a fibroblastic origin for DRTP sarcoma and provides a faithful mouse model for future mechanistic and translational investigation.

DICER1-related tumor predisposition (DRTP) (MONDO: 0100216), also known as DICER1 syndrome, is caused by germline pathogenic variants (PVs) in DICER1 that predisposes children and young adults to cancer^{1,2}. This includes several tumor types which arise at distinct anatomical locations, including lung, kidney, gynecologic tract, and central nervous system, among other rarer sites^{1–3}. Germline *DICER1* PVs typically inactivate the affected allele⁴. In DRTP, tumor development is usually associated with the occurrence of a hotspot missense mutation on the second *DICER1* allele, all of which result in the production of RNase IIIb catalytic domain-dead DICER1 protein⁵. This compound heterozygous state of DICER1 leads to impaired processing of pre-miRNAs and

subsequent loss of mature miRNAs derived from their 5p-strands (5p-miRNA)^{5–7}.

Although arising in different anatomical sites, it has long been noted that many DRTP cancers share key histomorphological features⁸. This includes subepithelial condensation of malignant mesenchyme (cambium layer), as well as frequent rhabdomyoblastic differentiation, resembling rhabdomyosarcoma (RMS). This overlap has raised important questions about whether these morphologically similar tumors arising at different sites constitute a common tumor entity. Recently, we addressed this hypothesis by studying a large series of DICER1 syndrome-associated cancers⁹. Our study identified

shared clinicopathological, genetic, and epigenetic (DNA methylation) features in a distinct group of DICER1 syndrome-associated mesenchymal neoplasms. This group encompasses well established tumor entities and comprises three clinically meaningful classes termed low-grade mesenchymal tumor with DICER1 alteration (LGMT DICER1), sarcoma with DICER1 alteration (SARC DICER1), and primary intracranial sarcoma with DICER1 alteration (PIS DICER1). While our findings suggest a common cellular origin for DRTP sarcoma, the underlying developmental biology and processes involved in DICER1-mediated tumor cell transformation remains largely unknown.

Genetically engineered mouse models (GEMM) are powerful tools for studying tumor development in a cell type specific manner. We have recently developed a conditional *Dicer1* RNase IIIb mutant allele (*Dicer1*^{+/fD1693N})¹⁰, which, upon Cre-mediated recombination, expresses the mouse p.D1693N mutant *Dicer1* (*Dicer1*^{D1693N}) that is equivalent to human p.D1709N mutant DICER1, one of the most common RNase IIIb mutations in human patients¹. With this tool strain and a *Dicer1*^{fl/fl} strain, we also developed a GEMM of DRTP that recapitulates the compound heterozygous *Dicer1* mutations seen in most DRTP-associated cancers¹⁰. Through an anti-Mullerian hormone receptor type 2 (Amhr2)-Cre driver, these *Dicer1* mutations were activated in stromal cells throughout the Mullerian tract, leading to the development of adenosarcomas, a cancer type well described in DRTP¹¹. Subsequently, Oikawa et al. developed a conditional human *DICER1* RNase IIIb hotspot mutation (p.G1809R), whose activation through an Albumin-driven Cre enabled the development of epithelial lesions in the liver¹². However, these and other models do not allow for controlling mutation activation and lineage-tracing, limiting their application to study tumor development continuum^{13,14}.

In this work, we cross our *Dicer1* tool strains with our recently generated CreERT2 mouse line-*Hic1*^{CreERT2}, which expresses Cre in mesenchymal stromal cells (MSCs) of most adult tissues upon tamoxifen treatment¹⁵. This inducible DRTP GEMM develops renal tumors that closely recapitulate the continuum of human DRTP sarcoma development. We leverage single-cell whole transcriptomics (scRNA-seq), a targeted spatial transcriptomics platform, and primary cell cultures to uncover the putative cellular origin and developmental trajectory of sarcomagenesis in this GEMM of DRTP sarcoma, which is found conserved in human disease.

Results

Development of a lineage-traceable and genetically true mouse model of DRTP sarcoma

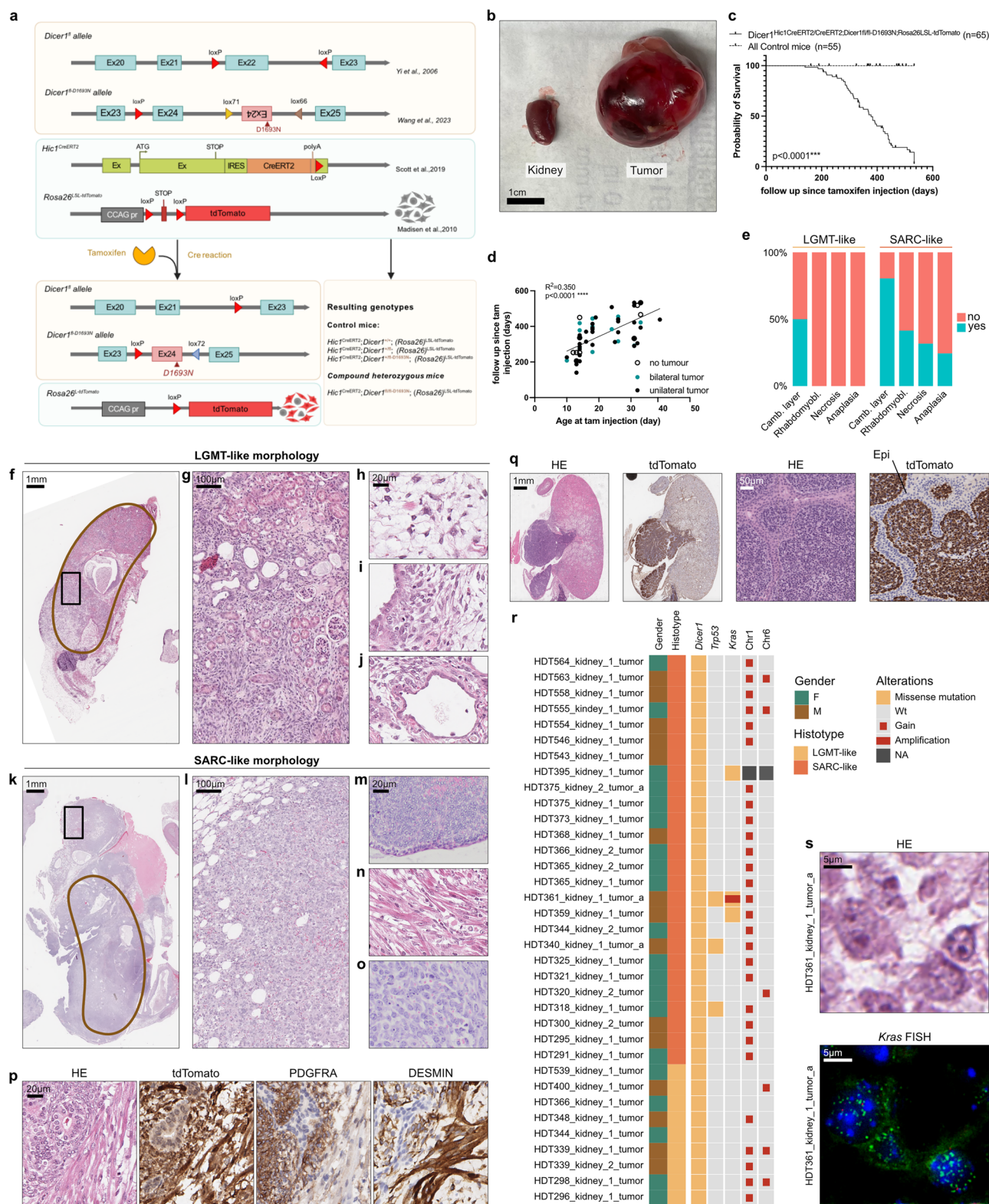
We utilized a tamoxifen-inducible *Hic1*^{CreERT2} driver, which marks quiescent mesenchymal progenitors in most vascularized tissues¹⁵, and the *Rosa26*^{LSL-tdTomato} strain, permanently labeling cells with tdTomato fluorescence upon Cre induction¹⁶, to develop a conditional and lineage-traceable *Dicer1* GEMM, *Hic1*^{CreERT2/CreERT2};*Dicer1*^{fl/fl};*D1693N*;*Rosa26*^{LSL-tdTomato} mice (referred to as HDT hereafter, Fig. 1a) that express hemizygous *Dicer1*^{D1693N} upon tamoxifen treatment in *Hic1*⁺ cells, recapitulating the compound heterozygous state of *DICER1* mutations in human tumors. Postnatal activation of *Dicer1* mutations in HDT mice, regardless of sex (Supplementary Fig. 1a), resulted in the development of unilateral or bilateral renal lesions (Supplementary Data 1). Macroscopic analysis revealed a tumor continuum that ranged from small tumors which maintain the integrity of the kidney structure to large masses that had completely overgrown the kidney (Fig. 1b), with 80% (52/65) penetrance and a median latency of 378 days (Fig. 1c and Supplementary Data 1). Different control genotypes did not lead to development of renal tumors and did not differ in overall survival (Supplementary Fig. 1b–f). While induction timing of *Dicer1* mutations did not correlate with HDT tumor penetrance or laterality (Supplementary Fig. 1g, h), late *Dicer1* mutation activation correlated with a prolonged tumor latency (Fig. 1d), suggesting that the number of cells amenable to mutant *DICER1*-driven oncogenesis decreases along

kidney development, consistent with the predominant childhood onset of these cancers in human patients.

Microscopically, renal HDT tumors closely mirrored human DRTP-associated mesenchymal neoplasms with tumor exhibiting a continuum from LGMT-like features to frank sarcomatous tumors showing features akin to SARC DICER1/PIS DICER1 (Fig. 1e and Supplementary Data 2). No correlation was observed between tumor histology and *Dicer1* mutation activation timing (Supplementary Fig. 1i, j). Notably, this continuum included histological features typically associated with cystic nephroma, which is classified under the LGMT DICER1 group, as well as anaplastic sarcoma of the kidney, categorized within the SARC DICER1 class⁹. In detail, this included small tumors, typically limited to the kidney's parenchyma (Fig. 1f, g), exhibiting hypocellular and primitive mesenchyme (Fig. 1h) with focal condensation adjacent to cystically dilated renal ducts (Fig. 1i, j). Conversely, large tumors that typically showed sarcomatous overgrowth of neighboring tissue consisted of sarcomatous mesenchyme with a pattern-less or fascicular growth (Fig. 1k, l). In these tumors, sub-epithelial aggregation of tumor cells underneath transitional cell epithelium (bona fide cambium layer) (Fig. 1m), prominent rhabdomyoblastic differentiation (Fig. 1n), and prominent anaplasia (Fig. 1o), often accompanied with necrosis, was noted. In line with human disease^{2,17}, tumors stained for PDGFRA and DESMIN, with PDGFRA marking areas of primitive tumor cells and DESMIN highlighting regions of rhabdomyoblastic differentiation (Fig. 1p and Supplementary Data 2). While the mesenchyme of all tumors was diffusely positive for lineage-tracing marker tdTomato, associated lining epithelium was negative for tdTomato throughout (Fig. 1q), indicating that it was not developed from the neoplastic lineage. Targeted sequencing of HDT tumors confirmed the expected *Dicer1* RNase IIIb (D1693N) mutation in all sequenced samples and revealed recurrent hotspot mutations in *Kras* and *Trp53*, which are commonly seen in human DRTP sarcoma^{9,18–21}, in 3/35 and 3/35 tumors, respectively (Fig. 1r and Supplementary Data 3). Furthermore, qPCR for *Kras* identified an amplification in one tumor with a concurrent *Kras* missense mutation, which was confirmed by fluorescence in-situ hybridization (FISH) (Fig. 1s and Supplementary Data 3). Notably, these mutations were only identified in neoplasms with SARC-like differentiation. Shallow whole-genome sequencing (sWGS) revealed gain of chromosome 1 in a significantly greater proportion (Fig. 1r and Supplementary Fig. 1k, l, Fisher's exact test, $p = 0.03051$, odds ratio = 8.41408, 95% Confidence Intervals = 1.2288, 78.5872) of SARC-like lesions (23/25) than LGMT-like lesions (5/9) (Supplementary Data 3), suggesting that aneuploidy may represent a secondary event associated with tumor progression. Overall, these results demonstrate that renal neoplasms from our HDT GEMM fully recapitulate the tumor development continuum of human DICER1 syndrome-associated mesenchymal neoplasms.

Organization of the *Hic1*⁺ mesenchymal stromal cell lineage in mouse steady state kidney

As pediatric tumors often mirror the normal development of affected organs, understanding the spatial composition of *Hic1*⁺ MSC lineage in normal mouse kidney is critical for establishing the tumor development trajectory of our renal HDT tumors. Notably, despite that *Hic1* effectively identifies various MSC populations, including fibroblasts and pericytes, alongside rare perineural cells in mouse skeletal muscle, heart, and skin^{15,22,23}, the characteristics of the *Hic1*⁺ MSC lineage in kidney (Supplementary Fig. 2) remains elusive. To address this, we analyzed renal tissues from control mice using scRNA-seq and targeted spatial transcriptomics (10× Genomics, Xenium, 379 genes; Supplementary Data 4 and Fig. 2a). ScRNA-seq analysis of normal mouse kidneys from 7 mice (post-natal week (PNW) 4–36, Supplementary Table 1) identified multiple clusters correlating to physiological kidney cell composition (Supplementary Fig. 3a–d)^{24–26}. Focused analysis of



tdTomato⁺ MSCs identified 6 clusters that were grouped into 5 cell types based on marker expression (Fig. 2b–d and Supplementary Fig. 3e–g). This included: (1) a cluster of mural cells (Mural, *Myh11*⁺, *Acta2*⁺, *Tagln*⁺), (2) two clusters of *Pdgfra*⁺ and *Dpt*⁺ fibroblasts, one characterized by *Pi16*^{high} and *Ly6a*^{high} (Uni Fibro *Pi16*^{high}) and the other by *Col15a1*^{high} and *Mfap4*^{high} (Uni Fibro *Col15a1*^{high}), and (3) two clusters of *Pdgfra*⁺ cells that were largely negative for *Dpt*, *Pi16* and *Col15a1*, one (Fibro) showing a signature resembling kidney interstitial fibroblasts

and mesangial cells (*Itga8*⁺, *Robo2*⁺, *Ptn*⁺) and the other (Cycle) enriched for cycling markers (*Cenpf*⁺, *Top2a*⁺, *Mki67*⁺).

Spatial transcriptomic analysis of 22 cored kidney regions of 10 kidneys and 2 whole kidney tissues from 9 normal control mice (Supplementary Data 5) corroborated the scRNA-seq data, revealing the cellular composition of normal kidneys (Supplementary Fig. 4a–e). Uni Fibro population was lower in aged (PNW: 81) vs young (PNW: 10) adult kidney of normal control mice (1.8% vs 2.3% of all cells identified,

Fig. 1 | Development of a lineage-traceable GEMM of DRTP sarcoma. **a** Scheme of HDT breeding strategy from 4 alleles: *Dicer1*^{fl}, *Dicer1*^{fl-D1693N}, *Hic1*^{CreERT2}, and *Rosa26*^{LSL-Tomato}. In HDT mice, tamoxifen injection leads to tdTomato expression and allows for lineage-tracing. Created in BioRender. Kommos, F. (2024) <https://BioRender.com/m44x473>. **b** Gross images of a normal kidney and renal mass from a HDT mouse. **c** Penetrance of tumor development in all control mice strains ($n = 55$) and HDT (*Dicer1*^{fl/fl-D1693N}) mice ($n = 65$). In this manuscript, n represents biological replicates unless specified otherwise. Comparison between two groups was done using the Log-rank (Mantel–Cox) test ($p = 0.000000000000763$). **d** Correlation of follow-up days since tamoxifen injection and age at tamoxifen (tam) injection. The linear relationship was analyzed with linear regression using the Wald test ($p = 0.00001393$). **e–o** histological features of renal tumors: small lesion confined to the renal parenchyma showing features akin to human low-grade mesenchymal tumours (LGMT) with DICER1 alteration (**f, g**), including hypocellular mesenchyme (**h**) and focal condensation adjacent to cystically dilated renal ducts (**i, j**). Large lesion overgrowing the renal parenchyma showing features akin to human sarcoma

(SARC) with DICER1 alteration (**k**), including sarcomatous mesenchyme (**l**), sub-epithelial aggregation of tumor cells (cambium layer, **m**), prominent rhabdomyoblastic differentiation (**n**), and anaplasia (**o**). **p** Tumors stain for PDGFRA, tdTomato, and DESMIN, with PDGFRA marking areas of primitive tumor cells and DESMIN highlighting regions of rhabdomyoblastic differentiation. **q** The mesenchyme of tumors, but not lining epithelium, is diffusely positive for tdTomato. **r** Summary of mutations identified from targeted sequencing, including *Dicer1* (Asp1693Asn), *Trp53* (Lys162Arg, Arg172His, and Phe110Val), and *Kras* (Gln61His and Gly12Arg, along with recurrent copy number alterations involving chromosomes 1 and 6 across 35 HDT tumor samples, with clinical and histological information included as an oncoplot to the side (relating to Supplementary Data 3). Histotype: Tumors with LGMT-like morphology, LGMT-like; tumors with SARC-like morphology, SARC-like. **s** FISH analysis validated *Kras* amplification in one HDT tumor. Representative micrographs from independent tumors were presented for (**f–p, q, s**); individual results of all cases are included in the Supplementary Data 3. HE Hematoxylin and Eosin. Source data are provided as a Source Data file.

Supplementary Fig. 4f). Targeted analysis of kidney MSCs (Fig. 2e) identified mural cells (Mural) mapping to pericytes and vascular smooth-muscle cells (*Rgs5*⁺, *Myh11*⁺, *Tagln*⁺), fibroblasts/mesangial cells (Fibro: *Bgn*⁺, *Cfh*⁺, *Itga8*⁺), and universal fibroblasts (Uni Fibro: *Bgn*⁺, *Lum*⁺, *Mfap4*⁺) that appeared to overlap with the Uni Fibro *Pil6*^{high} and Uni Fibro *Col15a1*^{high} populations by scRNA-seq analysis. In line with their expression profile, Fibro mapped to interstitial fibroblasts of cortex and medulla, as well as to rare mesangial cells, while Uni Fibro was predominantly found in the perivascular niche in the hilar region, the basal membrane niche subjacent to transitional cell epithelium, and the renal capsular surface (Fig. 2f–h). Although the resolution of the targeted gene panel did not allow for subclassification of Uni Fibro based on UMAP, *Pil6* expression was notably enriched in Uni Fibro within the perivascular niche, while *Mfap4* expression was high in Uni Fibro subjacent to transitional cell epithelium (Supplementary Fig. 5). Notably, a recent study of murine fibroblasts also identified two groups of *Dpt*⁺ fibroblasts that are present universally across various tissues (in this paper kidney was not studied)²⁷. The first group included *Dpt*⁺*Pil6*⁺ Uni Fibro that shared features with adventitial stromal cells²⁸, which are found in perivascular niches^{29,30}. The second group was composed of *Dpt*⁺*Col15a1*⁺ specialized fibroblasts and showed an association with the basement membrane. Therefore, our study identifies two types of fibroblastic progenitor populations originated from *Hic1*⁺ MSCs in the steady-state mouse kidney with *Pil6*^{high} Uni Fibro enriched within the perivascular niche and *Col15a1*^{high} Uni Fibro more associated with the subepithelial basement membrane of lining transitional epithelium.

Spatial cell composition of *Dicer1* renal sarcoma from HDT mice

Next, we characterized the transcriptomic landscape of 5 primary HDT renal tumors by scRNA-seq (Fig. 3a), one of which harbored a *Trp53* mutation in their corresponding formalin-fixed, paraffin-embedded (FFPE) tissue (Fig. 1r), revealing the presence of multiple cell lineages, such as transitional epithelial, mesenchymal, immune and endothelial cells (Supplementary Fig. 6). Unsupervised clustering of *tdTomato*⁺ MSCs (Supplementary Fig. 7a–c) revealed 10 distinct cell populations, all of which expressed some specific markers of steady kidney MSC lineages (Supplementary Fig. 7b). To further compare their developmental relatedness, we co-clustered scRNA-seq data of both HDT tumor and steady-state kidney samples (Supplementary Fig. 7d, e). This led to the identification of two clusters in HDT tumors corresponding to steady-state kidney mural cells, pericytes (Peri; *Rgs5*⁺) and vascular smooth muscle cells (SM Vasc; *Acta2*⁺), a cluster overlapping with steady-state kidney Uni Fibro that we termed Progenitor (Progen), and a ground population (Ground) closely resembling Fibro of steady-state kidney (Supplementary Fig. 7f and Fig. 3b–d). A Proliferative cluster (Prolif) enriched in genes related to cell proliferation, similar to

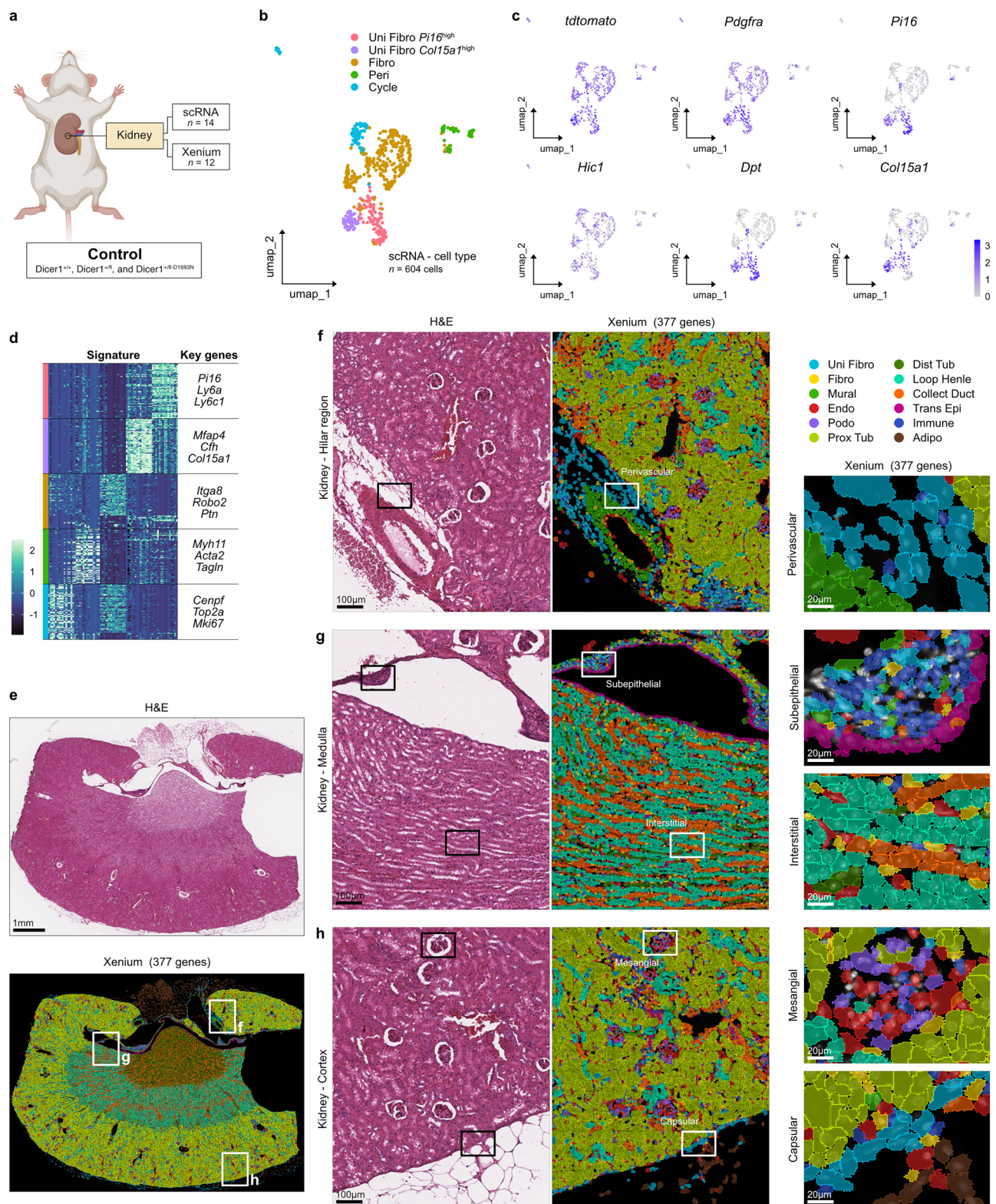
cycling cells of steady-state kidney, was also identified, some of which also expressed signature genes of Progen or Ground clusters, suggestive of their heterogenous nature. Moreover, HDT tumors also contained three populations not identified in steady-state kidney, including a cluster of perturbed-state fibroblasts (Fibro Perturb) and two myogenic populations. The Fibro Perturb cluster expressed a high level of *Pil6* with low *Dpt* and also expressed signature genes of perturbed-state fibroblasts (*Comp*⁺, *Cpxm2*⁺, and *Itgbl1*⁺) that were reported by Buechler et al.²⁷. The two additional clusters expressed genes linked to skeletal muscle development (e.g., *Pax7*) and committed muscle cells (e.g., *MyoG*), termed transiting-differentiated myogenic cells (TR Diff Myo) and differentiated myogenic cells (Diff Myo)³¹.

To spatially visualize these clusters, we performed targeted spatial transcriptomics on renal tumors ($n = 24$) and non-neoplastic kidneys ($n = 15$) harvested from HDT mice (Supplementary Data 5 and Supplementary Figure 8a–c). This revealed clusters corresponding to cell states in both steady-state kidney and tumors characterized by our scRNA-seq data (Supplementary Fig. 8d–g). Notably, the Progen cluster represented the dominant MSC population in LGMT-like tumors (Supplementary Fig. 8c and e), ubiquitously expressing *Mfap4* with a portion expressing *Dpt* and *Pil6* (Fig. 3e), indicating that this subepithelial cell cluster was composed of rare tumor initiating progenitors and abundant progenitor-like tumor cells. In SARC-like tumors, Progen cells were enriched in the subepithelial condensed cambium-like layer (Fig. 3f), where abundant Diff Myo and TR Diff Myo cells were often detected, suggestive of their lineage relatedness. Cells assigned to Ground were usually scattered within other cell states, e.g., Progen and Prolif clusters (Fig. 3g). The Ground and especially Prolif cluster cells showed frank sarcomatous differentiation (Fig. 3h).

Taken together, our scRNA-Seq and targeted spatial transcriptomic analysis not only revealed the spatial complexity of HDT tumors, but suggest that HDT tumors arise from perturbed differentiation of renal *Hic1*⁺ MSC progenitors subjacent to transitional epithelial cells of renal collecting ducts.

A *Dpt*⁺*Pil6*⁺ universal fibroblast lineage as the cell of origin for mouse *Dicer1* sarcoma

To identify the tumor development trajectory, we cultured *tdTomato*⁺ cells isolated from kidneys of control and HDT mice 4 weeks post tamoxifen treatment at 3-week-old. Bulk RNA-Seq analysis demonstrated that activation of *Dicer1* mutations led to increased expression of multiple Uni Fibro markers (*Dpt*, *Dcn*, *Pil6*, *Ly6a*, *Ly6c1*, *Col15a1*) and reduced expression of lineage markers for differentiated fibroblasts (*Itga8*) or pericytes (*Myh11*, *Acta2*) (Fig. 4a and Supplementary Fig. 9a), which were validated by real-time qRT-PCR (Supplementary Fig. 9b). This finding suggests that *Dicer1* mutation reprograms the multipotent



Hic1⁺ MSC, impairing the differentiation ability of their derivative primitive Uni Fibro, which primes them for subsequent malignant transformation towards a variety of cell states.

Pseudotime trajectory analysis from the *Dpt*⁺*Pi16*⁺ Uni Fibro-like population revealed 6 trajectories (Fig. 4b, c), all passing through the Progen population that was accompanied by a loss of Uni Fibro-specific markers (e.g., *Dpt*, *Pi16*, *Mfap4*) (Fig. 4d). Trajectories 2 and 3 led into Diff Myo (*Myog*⁺), passing through the “intermediate state” TR

Diff Myo cells (*Pax7*⁺), including a highly proliferative state of TR Diff Myo. Trajectories 1, 4, and 5 transitioned into the Ground state (*Itga8*⁺). Trajectory 6 transitioned into the Pericyte state (*Pdgfrb*⁺*Rgs5*⁺). To validate the lineage relationship, we conducted copy number (CN) analysis using scRNA-Seq datasets from HDT renal tumors, which revealed high CN changes in *tdTomato*⁺ mesenchymal cells, but not other cell types (Supplementary Fig. 9c), consistently with our finding using the sWGS analysis (Supplementary Fig. 1k, l). Analysis of normal

Fig. 2 | Organization of the *Hic1*⁺ mesenchymal stromal cell lineage in murine steady state kidney. **a** Steady-state kidney samples from control mice (*Hic1*^{CreERT2/CreERT2}; *Rosa26*^{LSL-tdTomato}, *Hic1*^{CreERT2/CreERT2}; *Dicer1*^{fl/+}; *Rosa26*^{LSL-tdTomato}, *Hic1*^{CreERT2/CreERT2}; *Dicer1*^{fl/+}; *Rosa26*^{LSL-tdTomato}) profiled by scRNA-seq ($n = 14$ kidneys from 7 mice in 3 batches with postnatal week [PNW] at 4, 10, and 33–36) and Xenium spatial transcriptomics ($n = 12$ kidneys from 9 mice with PNW 10 to PNW 81) analyses. Created in BioRender. Kommoss, F. (2024) <https://BioRender.com/m44x473>. **b** UMAP plot from scRNA-seq data ($n = 3$) of *tdTomato*⁺ mesenchymal stromal cells colored by cell type. Louvain method was used for clustering, and cell types are annotated based on transcripts representative of different lineages (relating to Supplementary Fig. 3f): Uni Fibro *Pil6*^{high}, universal fibroblast *Pil6*^{high}; Uni Fibro *Col15a1*^{high}, universal fibroblast *Col15a1*^{high}; Fibro, Fibroblast/mesangial

cell; Mural, mural cells; Cycle, cycling cells. **c** UMAP plots depicting expression of key fibroblast lineage markers. **d** Heatmap of top 30 defining transcripts, based on cell type annotation. Key genes of signature are highlighted. **e–h** Representative micrographs showing clusters identified from targeted spatial transcriptomics (related to Supplementary Fig. 4a–e) of 12 mouse kidneys from 9 normal control mice. Spatial transcriptomic cell segmentation shown was generated using all kidney samples. Fibro is present in the mesangium and interstitium of the renal cortex and medulla (**g** and **h**). Aggregation of Uni Fibro is noted in perivascular, subepithelial/basement membrane-associated, and capsular niches (**f–h**). Cell types are colored according to the legend. H&E Hematoxylin and Eosin. Source data are provided as a Source Data file.

control kidneys revealed minimal CN changes (Supplementary Fig. 9d), further confirming CN changes in HDT mesenchymal cells were tumor-specific. Focused analysis of *tdTomato*⁺ mesenchymal cells of HDT tumors identified CN changes in all *Hic1*⁺ clusters except Fibro Perturb and SM Vasc populations. In contrast, a sizable subset of Progen cells accumulated gains in chromosomes 1 and 6 (Fig. 4e). Notably, the Ground and Prolif populations exhibited near-universal gains in these same chromosomes. Additionally, only a minority of the TR Diff Myo and Diff Myo populations demonstrated gains in chromosomes 1 and 15, although these CN changes were more pronounced than those observed in Progen cells, suggesting that chromosome gains can occur early before their committed differentiation but are not essential for rhabdomyogenic differentiation. These findings support a model of complex tumor progression, in which *Dicer1* mutation perturbs the differentiation potential of *Hic1*⁺ MSC, priming *Dicer1*-mutant *Dpt*⁺*Pil6*⁺ Uni Fibro-like cells for oncogenic transformation via distinct mechanisms: a portion of these progenitors undergoes aberrant myogenic differentiation, whereas accumulation of CN alterations promotes tumorigenesis towards the Ground state, from which a subset progresses to proliferative sarcoma upon acquiring specific genomic changes (Fig. 4f). Consistent with this model and the recurrent occurrence of p53 mutations in both human and mouse tumors, deletion of p53 in *Dicer1*-mutant primary cells significantly accelerated cell growth (Fig. 4g). Furthermore, gene set enrichment analysis of differentially expressed genes among HDT tumor cell states revealed that the Prolif population was characterized by downregulation of p53-associated pathways and concomitant upregulation of MAPK signaling (Fig. 4h). Therefore, the identification of HDT tumor development trajectory may inform future therapeutic development.

Cell states of mouse *Dicer1* sarcoma are largely recapitulated in human DRTP sarcoma

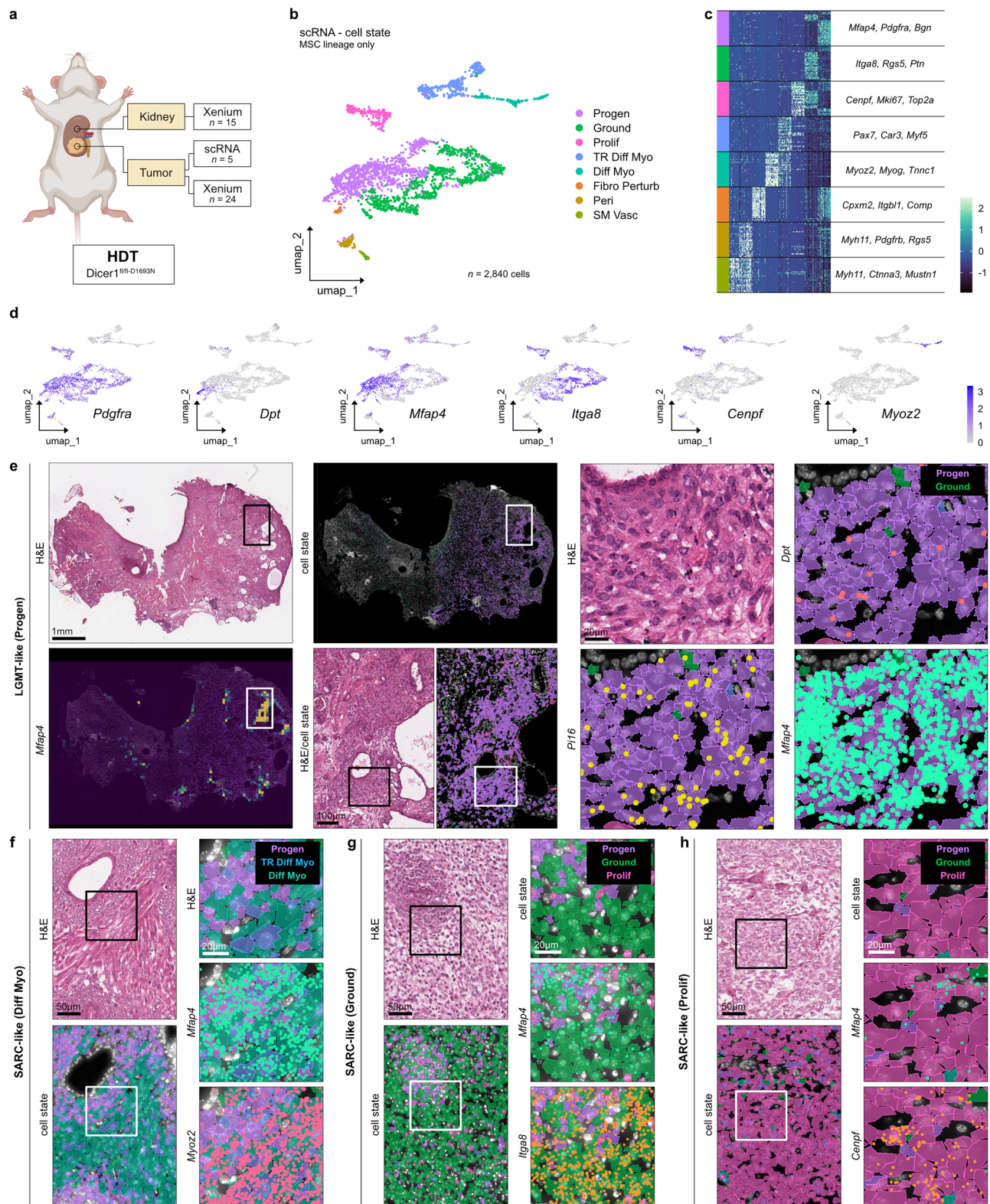
To determine whether cell states informed from the mouse *Dicer1* sarcoma are present in human DRTP sarcomas, we characterized the transcriptional landscape across 16 human DRTP mesenchymal tumors (4 LGMT DICER1, 11 SARC DICER1, and 1 PIS DICER1), including Cystic nephroma, pleuropulmonary blastoma (PPB) I/II/III, nasal chondroid mesenchymal hamartoma, DICER1-mutated embryonal RMS, and primary intracranial sarcoma (PIS DICER1), using a targeted spatial transcriptomics platform (Fig. 5a). Unsupervised clustering identified 13 Louvain clusters (Supplementary Fig. 10a, b). These clusters correspond to distinct cell populations, including non-neoplastic cell types (Supplementary Fig. 10c, d). Guided by lineage markers (e.g., *PDGFRA*) and spatial morpho-molecular correlation, we identified fibroblast lineage clusters (Supplementary Fig. 10e)³², which were grouped into cell states that largely recapitulated those identified in our mouse HDT tumors (Fig. 5b–d). These included a neoplastic progenitor cluster (Progen) expressing *PDGFRA* and universal fibroblast marker *DPT*, a proliferative cluster (Prolif) dominated by proliferation and cell cycle genes, two myogenic subpopulations expressing skeletal muscle-related genes (e.g., *DES*, *PROX1*) as well as

committed muscle cell markers (e.g., *MYBPC1*), termed transiting-differentiated myogenic cells (TR Diff Myo) and differentiated myogenic cells (Diff Myo). Additionally, a *PDGFRA*⁺ fibroblastic population lacking other signatures (Ground), and a distinct fibroblast cluster (Fibro) mapping to non-neoplastic tissue and marked by *PDGFRA*⁺ and *C7*⁺ were identified. Spatial analysis revealed Progen cells in all tumors at varying frequencies (Fig. 5e), often mapping to the cambium layer of LGMT and SARC DICER1 (Fig. 5f) as seen in our HDT renal tumors (Fig. 3f). Like HDT tumors, TR Diff Myo states were associated with increased proliferation (Fig. 5g). Bona fide rhabdomyoblastic differentiation (Diff Myo) co-localized with TR Diff Myo and the committed muscle marker *MYBPC1* (Fig. 5h). Ground and Prolif cells were abundant in SARC/PIS DICER1 and exhibited sarcomatous growth (Fig. 5i). In agreement with the elevated MAPK pathway in the highly proliferative state cells of HDT tumors (Fig. 4h), immunohistochemistry staining of human tumors revealed elevated p-ERK levels in SARC DICER1 compared to LGMT DICER1 (Supplementary Fig. 10f, g). These results alongside trajectory analysis (Supplementary Fig. 10h), closely mirrored spatial relationships observed in murine HDT tumors.

Human DRTP mesenchymal neoplasms closely mimic RMS clinically and morphologically. A recent scRNA-based meta-analysis identified a robust tripartite cell state landscape across various human RMS tumors, including Progenitor, Proliferative, and Differentiated metaprograms³³. This paper also inspired some of our cluster names. Our transcriptomics data showed that these RMS metaprograms are enriched in signatures present in DICER1 mesenchymal neoplasms (Supplementary Fig. 10i), highlighting their transcriptional relatedness.

Discussion

Patients with DRTP are predisposed to a variety of cancers, many of which are aggressive sarcomas, arising in diverse anatomical locations. Recent multi-omics studies have significantly advanced our understanding of the genetic underpinnings of these cancers and uncovered the similarities of DRTP sarcomas regardless of their anatomic sites of onset. However, despite the well-recognized mesenchymal nature of DRTP sarcomas, their cellular origin and developmental trajectory remain unknown. In this study, we demonstrate that expression of patient tumor-relevant hemizygous *Dicer1*^{D1693N} in *Hic1*⁺ MSC drove the development of renal sarcomas that closely resemble the human disease, both histomorphologically and genomically, which is consistent with the observations in tumors from our *Amhr2*^{+/cre}-driven *Dicer1* GEMM¹⁰. Importantly, with this lineage-traceable GEMM, we identified the spatial composition of *Hic1*⁺ MSCs in steady state normal mouse kidneys and discovered that upon activation of *Dicer1* mutations, *Hic1*⁺ MSC-derived universal fibroblast, marked by *Pdgfra*, *Dpt*, and *Pil6*, gained the potential to differentiate into a myogenic lineage and/or became proliferative sarcomatous cells. We further identified these cell states and their developmental hierarchy in human DRTP sarcomas. Thus, we uncovered universal fibroblasts as the putative cellular origin of mouse *Dicer1* sarcoma and deconvoluted their developmental trajectory, which appears conserved in human patients.



In steady-state murine kidneys, we show that *Dpt*⁺ universal fibroblasts are localized to the perivascular, capsular, and subepithelial basal membrane niches. This is particularly intriguing given that human DRTP sarcomas frequently arise in organs with epithelial lining, such as the lungs, kidneys, and uterus, and form a cambium layer (sarcoma botryoides)², yet they can also develop in non-epithelial tissues like the brain²⁰. Our data suggest that the anatomical predilection for DRTP sarcomas may be linked to the proposed cellular origin

within the universal fibroblastic niche, which spans basal membrane-associated, capsular, and perivascular compartments, with the perivascular localization potentially accounting for tumor development in non-epithelial tissues such as the brain. Specifically, the enrichment of transitional epithelial cells in our renal tumors, revealed by both scRNA-Seq and targeted spatial transcriptomic analysis (Supplementary Figs. 6d and 8c, d), strongly supports that these renal tumors arose from *Hic1*⁺ MSC progenitors subjacent to

Fig. 3 | *Hic1*⁺ mesenchymal stromal cell lineage in mouse *Dicer1* sarcoma.

a Samples from HDT mice (*Hic1*^{CreERT2/CreERT2}; *Dicer1*^{fl/fl-D1693N}; *Rosa26*^{LSL-tdTomato}) profiled by scRNA-seq and Xenium spatial transcriptomics analyses. Created in BioRender. Kommos, F. (2024) <https://BioRender.com/m44x473>. **b** UMAP plot from scRNA-seq data ($n = 5$) of *tdtomato*⁺ mesenchymal stromal cells colored by cell type. Louvain method was used for clustering, and cell types are annotated based on transcripts representative of different lineages (relating to Supplementary Fig. 6c): Progen progenitor, Ground ground (Fibro-like), Prolif proliferative, TR Diff Myo transiting-differentiated myogenic cells, Diff Myo differentiated myogenic cells, Fibro Perturb perturbed fibroblast, Mural mural cells. **c** Heatmap of top 30 signature defining transcripts, based on cell state annotation. Key genes of signatures are highlighted. **d** UMAP plots depicting expression of key markers. **e–h** Mapping of clusters identified from targeted spatial transcriptomics (10× Genomics, Xenium,

379 genes) of HDT tumors and non-neoplastic kidneys ($n = 24$ and 15 , respectively; related to Supplementary Fig. 8) to spatial domains of HDT tumors. Progen is abundant in tumors with LGMT-like morphology, shows primitive cytologic features, and expresses *Dpt*, *Pi16*, and *Mfap4* (**e**). In SARC-like tumors with rhabdomyoblastic differentiation, Progen colocalizes with TR Diff Myo and Diff Myo; expression of *Mfap4* and *Myo2* is shown (**f**). Ground corresponds to sarcomatous cells, which may be colocalized with Progen (**g**) and may progress into highly atypical/anaplastic and proliferative tumor cell populations (Prolif); expression of *Mfap4*, *Itga8*, and *Cenpf* is shown (**h**). Representative micrographs from independent tumors were presented. Spatial transcriptomic cell segmentation shown was generated using all independent samples ($n = 39$). H&E, Hematoxylin and Eosin. Source data are provided as a Source Data file.

transitional epithelial cells of renal collecting ducts. Future global spatial transcriptomics will help identifying the specific features of universal fibroblasts in the subepithelial basal membrane niche that distinguish them from those in other niches to enable tumorigenesis.

DRTP predominantly affects children and young adults, suggesting that the tumor susceptibility in DRTP may be influenced by the developmental stage. Mirroring this, we observed prolonged tumor latency in HDT mice with a late activation of *Dicer1*^{D1693N} mutation (Fig. 1d). Thus, earlier developmental windows may lead to the retention of certain universal fibroblasts prone to malignant transformation in the kidney. Moreover, despite that *Hic1* marks MSC populations across various tissue types, tumor development was restricted to kidneys in our HDT model. We hypothesize that while universal fibroblasts exist across a variety of tissues in mice²⁷, their specific developmental state in the kidney may influence the susceptibility to *Dicer1*-associated transformation. Future analysis is needed to address whether these mutant *Dicer1* transformable universal fibroblast populations are (1) declining along mouse kidney development and (2) uniquely present or more abundant in postnatal juvenile mouse kidneys than other organs due to delayed disappearance after embryogenesis. Given the potential sampling bias during kidney dissociation for scRNA-Seq, global spatial transcriptomic analysis is desired to profile mouse kidneys and other organs during their development continuum to identify such unique and rare cell population.

Notably, the Fibro Perturb and SM Vasc clusters appeared to be the non-active components of HDT renal tumors. The absence of lineage relationship between these clusters and the putative tumor cellular origin (*Dpt*⁺*Pi16*⁺ cells) (Fig. 4b) and absence of copy number alterations in these clusters (Fig. 4e) suggest that they are likely untransformed *tdTomato*⁺ cells. Notably, *Hic1*⁺ MSC cells are a mixture of cells at various states including primitive universal fibroblasts and also differentiated cells, e.g., fibroblasts and pericytes, which are abundant in the kidney at all stages but appear not transformable by *Dicer1*^{D1693N} mutation as evidenced by the prolonged tumor latency when *Dicer1*^{D1693N} mutations were activated late (Fig. 1d). Thus, it is possible that Fibro Perturb cells in HDT tumors may arise from activating *Dicer1*^{D1693N} in *Hic1*⁺ fibroblasts, whereas the SM Vasc may derive from activating *Dicer1*^{D1693N} in *Hic1*⁺ mural cells. However, future studies are needed to address whether these rare subsets, particularly the Fibro Perturb cluster, can contribute to sarcomagenesis upon accumulating additional genomic mutations.

We further demonstrate that the epithelial lining in our kidney tumors was negative for tdTomato, indicating that it did not belong to the neoplastic lineage. Nevertheless, these epithelial areas frequently displayed a botryoid or even cystic appearance. We propose that such cystic changes, commonly observed in tumors from the LGMT DICER1—such as PPB I and CN—are not a result of DICER1 disruption in epithelial cells, as previously suggested⁴. Instead, we hypothesize that these cystic transformations are likely a reactive phenomenon, driven by neoplastic mesenchymal proliferation, which may exert pressure on adjacent structures, such as kidney tubules or alveolar spaces, leading

to their compression and subsequent cystic dilation due to fluid or air entrapment. Notably, while previous studies suggested that cystic changes could reflect a non-autonomous mechanism where epithelial loss of *Dicer1* promoted the transformation of the underlying mesenchymal cells^{13,34}, their contribution to the development of LGMT DICER1, which is driven by RNase IIIb-mutant *DICER1* rather than loss of *DICER1*, remains to be addressed. It is possible that *DICER1* may be epigenetically silenced in the epithelium of a subset of these tumors, potentially influencing the growth of the subepithelial mesenchymal cells that are already transformed by RNase IIIb-mutant *DICER1* through this non-autonomous mechanism.

Notably, while human DRTP-associated mesenchymal neoplasms closely mimic RMS clinically and morphologically, our previous molecular analysis distinguished DRTP sarcomas from bona-fide RMS subtypes (fusion-negative, fusion-positive, and *MYOD1*-mutated RMS) based on DNA methylation profiles, suggesting distinct cellular origins for these tumors⁹. Conventional RMS subtypes are typically thought to originate from satellite-cell derived post-natal muscle and results from aberrant skeletal muscle differentiation^{35,36}. However, recent studies have challenged this notion, suggesting that RMS may originate from endothelial progenitors^{37,38} or mesenchymal stem cells^{33,39}. In our *Dicer1* GEMM, although our HDT tumors closely mimic conventional types of RMS morphologically and transcriptomically, we observed the expansion of the fibroblastic progenitor population that gave rise to divergent trajectories, including a step-wise continuum of rhabdomyoblastic differentiation, including *Pax7*⁺ and *Myf5*⁺ populations, supporting a fibroblastic origin of RMS in DRTP sarcomas. Accordingly, a recent study demonstrated that fibroblasts derived from the lateral plate mesoderm can trans-differentiate by activating myogenic programs during physiological development⁴⁰. Yaseen et al. showed that these fibroblasts not only acquire myogenic features and fuse with existing myofibers but also retain partial expression of their original fibroblastic transcriptional program. *Hic1*⁺ fibroblasts typically do not express *Pax7*, and do not contribute to the myogenic lineage in the developing limb⁴¹. However, these cells have a somitic origin^{41,42}. Notably, in our *Hic1*⁺ GEMM of synovial sarcoma, re-expression of *Pax7* has been linked to unmasking of their embryonic epigenome, suggesting a potential for lineage plasticity⁴³. In contrast, there was no expression of terminal myogenic transcription factors, such as *Myf5* or *MyoD1*, and no evidence of rhabdomyoblastic differentiation in the latter model. This suggests that our observations are specifically linked to the disruption of *Dicer1*, rather than being a broader feature of the *Hic1*⁺ lineage. In line with this, our findings provide clear evidence that these fibroblasts have the potential to undergo rhabdomyoblastic transformation upon *Dicer1* disruption, which may be linked to re-expression of embryonic programs and challenges the long-standing dogma that all RMS develop from a predetermined myogenic lineage.

Finally, although our spatial transcriptomic analysis is limited by the targeted panel, our data reveals a striking concordance between the cellular origins of murine and human DRTP sarcomas. We confirm that in the human disease, cancer cell states share transcriptomic

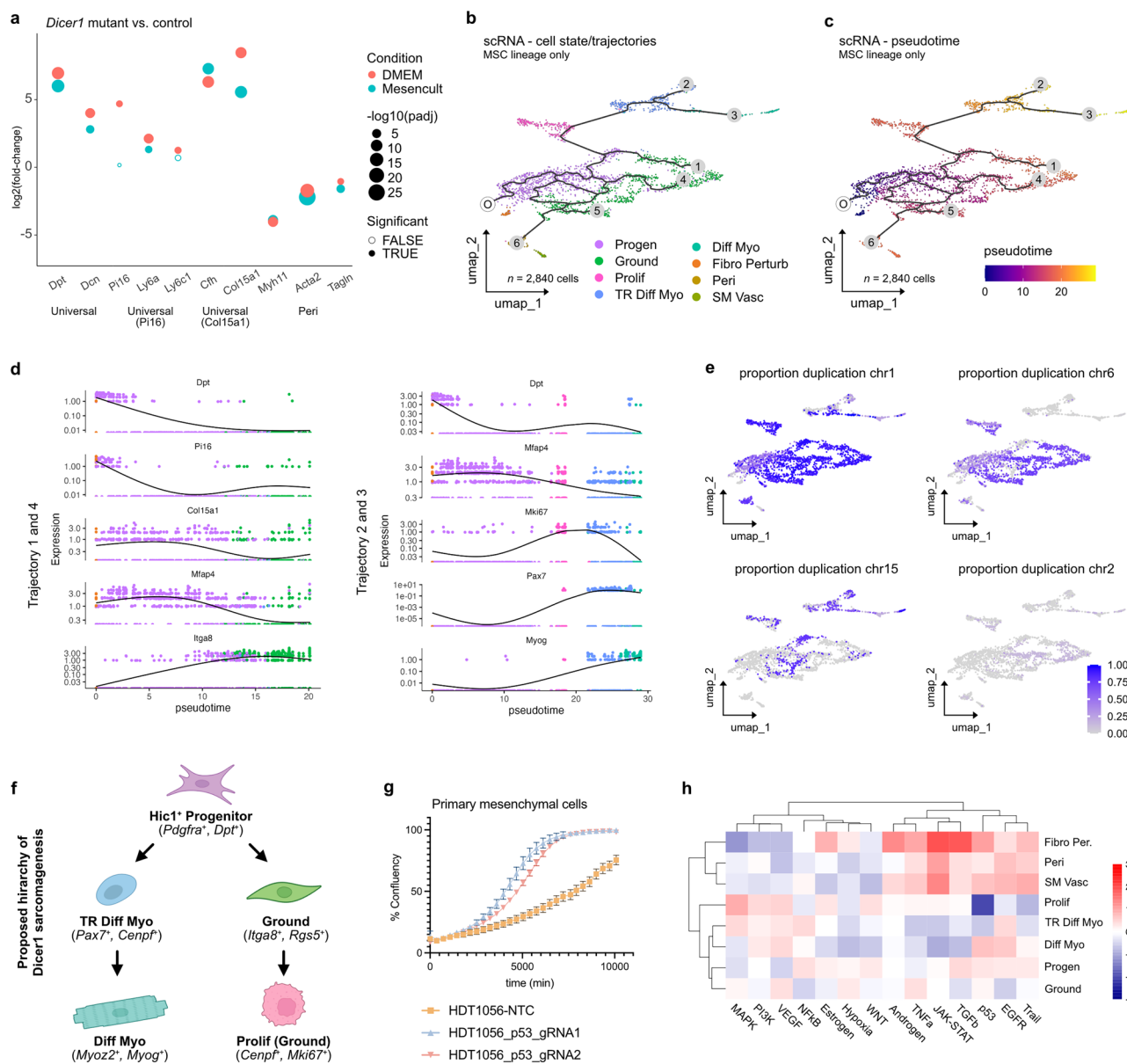


Fig. 4 | A *Hic1*⁺ universal fibroblast lineage as cell of origin for mouse

Dicer1 sarcoma. **a** Bulk RNA-seq data comparing specific gene expression in primary mesenchymal cells harvested from histologically normal murine kidney with *Dicer1* wild type and *Dicer1* mutant genotypes. Cells were cultured in either DMEM or Mesencult. *Dicer1* mutant cells display increased expression of universal fibroblast markers such as *Dpt*, *Dcn*, *Pli6*, *Ly6a*, *Ly6c1*, *Col15a1*, and reduced expression of lineage markers for differentiated fibroblasts (*Itga8*) or pericytes (*Myh11*, *Acta2*). Two-sided differential expression was assessed with the Wald test in DESeq2 and corrected for multiple comparisons with Benjamini–Hochberg.

b, c Trajectory and pseudotime analysis of scRNA-seq data. UMAP colored by cell type and pseudotime. O indicates origin, trajectories are numbered. **d** Gene trend plots of key cluster genes along the sarcomagenic trajectory (Ground and Diff Myo) defining the transformation process. Cells are colored according to cell type.

e UMAP of single-cell RNA-seq data showing the proportion of cells with chromosomal duplications (relating to Supplementary Fig. 9c, d). **f** Schematic of proposed hierarchy of sarcomagenesis in mouse *Dicer1* sarcoma. Created in BioRender.

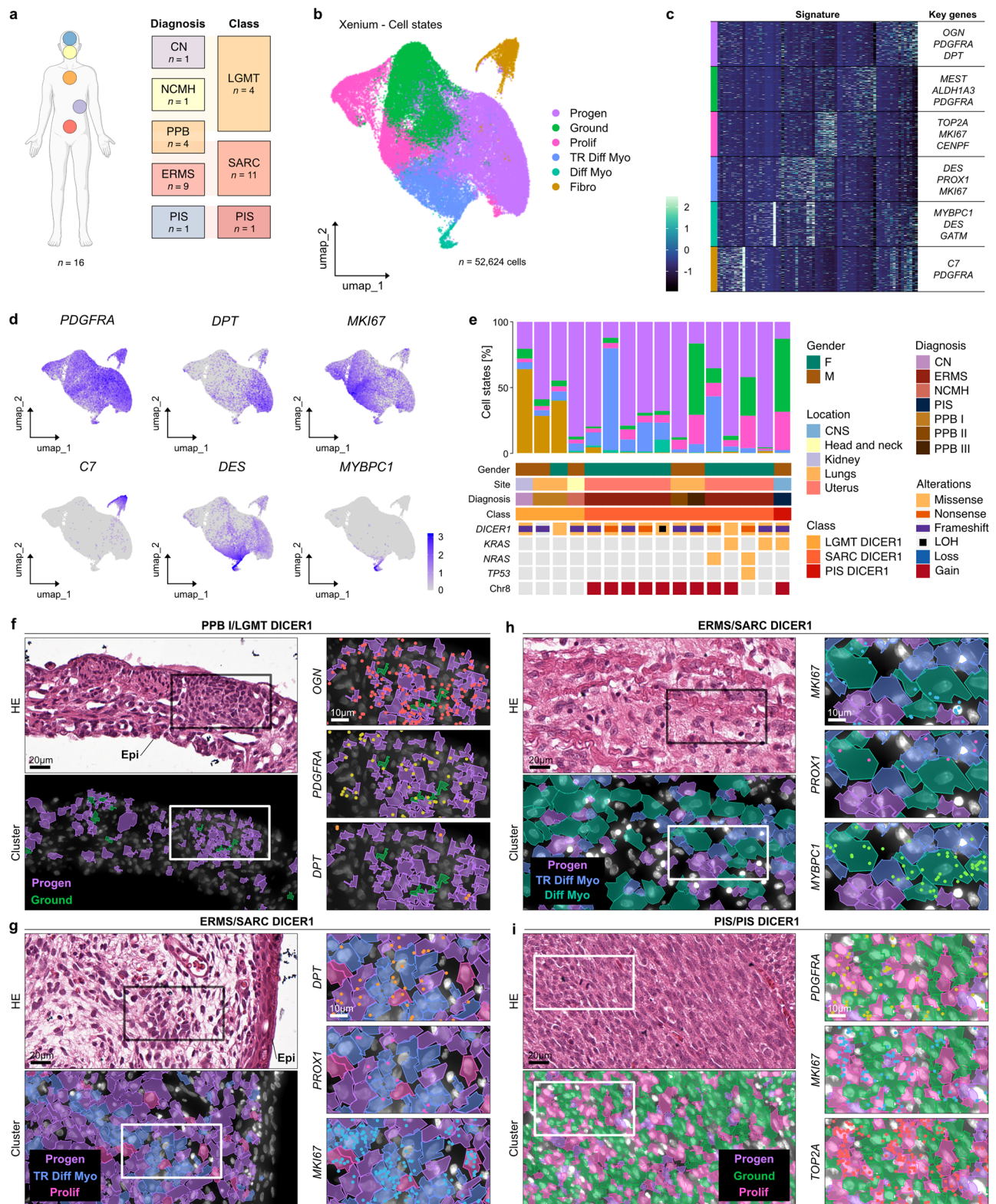
Kommos, F. (2024) <https://BioRender.com/m44x473>. **g** CRISPR/Cas9-mediated knockouts of p53 increases proliferation of *Dicer1*-mutant renal primary mesenchymal cells (HDT1056) in comparison to non-targeting control (NTC). Two independent guide RNAs (gRNA) were delivered via lentiviral transduction. *n* = 3 technically replicates for each condition. Data are presented as mean value ± SEM. **h** Gene set enrichment analysis along the tumor development continuum. The Prolif population exhibits upregulation of MAPK signaling and concurrent down-regulation of p53-associated pathways, highlighting key drivers of the proliferative cell state. Source data are provided as a Source Data file.

similarities with those observed in our murine model, including the presence of a progenitor population of universal fibroblastic lineage (*PDGFRA*⁺, *DPT*⁺). This striking similarity in cell states alongside the common frequent activation of the MAPK pathway identified in murine and human DRTP tumors further substantiates the relevance of our GEMM model for DRTP sarcoma and underscores its potential as a tool for advancing mechanistic understanding and developing novel therapeutic strategies.

Methods

Model development and cohort assembly

All animal experiments were conducted following ethical approval by the University of British Columbia animal care committee (A21-0023 and A21-0047). The conditional knockin mice expressing the floxed *Dicer1* p.D1693N (*Dicer1*^{+/f-D1693N}) was previously developed by our group¹⁰. The *Dicer1*^{R/m} mice⁴⁴ [RRID:MMRRC_015388-UNC] were obtained from the Mutant Mouse Resource & Research Centres



(MMRRC) depository. The Rosa26^{LSL-tdTomato} strain¹⁶ was obtained from The Jackson Laboratory (stock #: 007909). Hic1^{CreERT2} mice were previously generated as described elsewhere¹⁵. Mouse lines were maintained on a pure C57Bl/6J genetic background. We outcrossed these 4 strains together to develop the Hic1^{CreERT2/CreERT2}; Dicer1^{fl/fl-D1693N}; Rosa26^{LSL-tdTomato} and corresponding control strains. Genotyping was performed using primers in Supplementary Table 2 using ear punches obtained at 3 weeks of age. To induce CreERT2 nuclear translocation,

mice were injected at postnatal 10 (P10) days to 7 weeks old intraperitoneally with tamoxifen (dose: 2 mg every 24 h for 5 consecutive days for mice at 3–7 weeks old; 1 mg every 24 h for 3 consecutive days for mice at 2 weeks old; 1 mg for mice at P10).

For tumor development studies, we used 65 HDT mice (33 male, 32 female) and 55 control mice (26 male, 29 female). Mice were either found dead or sacrificed because they were morbid or reached the experimental end point (1.5 years). At the experimental endpoint, the

Fig. 5 | Murine cell states are largely recapitulated in human DRTP sarcoma.

a Patient samples ($n = 16$) encompassing three molecular classes of DICER1-associated mesenchymal neoplasms⁹. Created in BioRender. Kommos, F. (2024) <https://BioRender.com/m44x473>. **b** UMAP plot from targeted spatial transcriptomic (10× Genomics) data of mesenchymal cells colored by cell type. Cell types are annotated based on transcripts representative of lineage (relating to Supplementary Fig. 10d) and morpho-molecular correlation (relating to **f–i**): Progen progenitor, Ground ground, Prolif proliferative, TR Diff Myo transiting-differentiated myogenic cells, Diff Myo differentiated myogenic cells, Fibro fibroblasts. **c** Heatmap of top 15 signature defining transcripts, based on cell state annotation. Key genes of signatures are highlighted. **d** UMAP plots depicting expression of key fibroblast lineage markers and markers characteristic of identified cell types. **e** Summary of the cell type composition of patient samples with clinical and tumor genomic alteration information⁹. Class: low-grade mesenchymal tumor with DICER1 alteration, LGMT DICER1; sarcoma with DICER1 alteration, SARC DICER1; primary intracranial sarcoma with DICER1 alteration, PIS DICER1. Diagnosis: cystic nephroma, CN; embryonal rhabdomyosarcoma (DICER1-mutated), ERMS;

nasal chondroid mesenchymal hamartoma, NCMH; primary intracranial sarcoma, PIS; pleuropulmonary blastoma type I, II, and III, PPB I/II/III. LOH loss of heterozygosity. **f–i** Representative micrographs from independent tumors showing spatial mapping of cell clusters. Spatial transcriptomic cell segmentation shown was generated using multiple independent samples ($n = 16$). A PPB I (LGMT DICER1) shows cystic configuration with nodular and subepithelial growth of primitive tumors cells (Progen); expression of *OGN*, *PDGFRA*, and *DPT* is shown (**f**). DICER1-mutated ERMS (SARC DICER1) shows a subepithelial cambium layer and focal aggregation of more atypical tumor cells (TR Diff Myo); expression of *DPT*, *PROX1*, and *MKI67* is shown (**g**). A DICER1-mutated ERMS (SARC DICER1) with bona fide rhabdomyoblastic differentiation and colocalization of cells assigned to TR Diff Myo and Diff Myo. Expression of *MKI67*, *PROX1*, and *MYBPC1* is shown (**h**). A PIS DICER1 with fascicular and frank sarcomatous growth of atypical and highly proliferative tumor cells assigned to Ground and Prolif clusters; expression of *PDGFRA*, *MKI67*, and *TOP2A* is shown (**i**). HE Hematoxylin and Eosin. Source data are provided as a Source Data file.

tumor burden in each mouse did not reach the approved humane endpoint limit (10% of body weight). For penetrance analysis, the presence of macroscopic or microscopic tumors at death was used as the primary endpoint. For overall survival (OS) analysis, events were defined as either the death of the mice or their euthanasia due to morbid conditions. Five tumors from 5 HDT mice (4 males and 1 female) with age ranging PNW 36 to PNW 66 were used for scRNA-Seq analysis. Twenty-four tumors from 8 HDT males and 10 HDT females, age ranging from PNW 36 to PNW 81, and 15 non-neoplastic kidneys from 13 HDT mice (9 males, 4 females), age ranging from PNW 10 to PNW 61, were used for spatial transcriptomic analysis.

For scRNA-Seq analysis of normal mouse kidneys, 7 control mice (3 males, 4 females) were sacrificed at the age of PNW 4–36. For spatial transcriptomic analysis of normal mouse kidneys, 9 control mice (5 males and 4 females) were sacrificed at the age of PNW 10 to PNW 81.

Patient samples

A cohort of 16 patient samples of DICER1 mesenchymal neoplasms from 10 males and 6 females, with age ranging from 1 month to 69 years, was included in this study. They had all undergone histopathological evaluation by specialized pathologists, as well as molecular testing whenever applicable. Samples were collected at McGill University-affiliated hospitals in Montreal, Quebec, Canada, and the Institute of Pathology, University Hospital Heidelberg, mainly from collaborating institutions in accordance with ethics review board regulations. All eligible participants, regardless of sex, were included with signing an informed consent form. Clinical data, histological features, and/or sequencing results of tumors have previously been reported elsewhere⁹. Study of patient samples was approved by the institutional ethics committee (Heidelberg University, S660/2020) and performed in accordance with the Declaration of Helsinki. Due to the rareness of the disease, we do not have sufficient power to perform sex-specific analysis.

Tissue microarray construction

For murine tissue microarray (TMA, $n = 2$) construction, 1.0 mm cores of FFPE tissue from control and HDT mice were used, as described previously⁴⁵. For human samples, a previously constructed TMA was used, that included duplicates of 0.6 mm cores of FFPE tissues, including the 16 samples of DICER1 mesenchymal neoplasms used in this study.

Primary cell culture

The whole kidneys from $\text{Hic1}^{\text{CreERT2/CreERT2};\text{Rosa26}^{\text{LSL-tTomato}}}$ and $\text{Hic1}^{\text{CreERT2/CreERT2};\text{Dicer1}^{\text{fl/fl-D1693N}};\text{Rosa26}^{\text{LSL-tTomato}}}$ GEMM mice 4 weeks post tamoxifen treatment, which occurred when mice were 3-week-old, were minced with forceps to small chunks pipette-able by a 5 mL

serological pipette. Tumour chunks were then incubated in Gibco HBSS (no calcium, magnesium, phenol red, Catalog #14175103) and Stemcell 10× Collagenase/hyaluronidase in DMEM (Catalog # 07912), in a glass flask at 37 °C on a hot plate, until everything was dissolved, usually takes between 30 and 45 min. Enzymatic digestion was stopped with Advanced DMEM/F-12 (Catalog #12634010). Cells and tumour chunks were spun at 1000 rpm for 5 min. Pellets were resuspended in Gibco 0.25% Trypsin-EDTA (Catalog #25200072) for 5 min at 37 °C, followed by filtering through a Falcon 70 µm cell strainer (Catalog #08-771-2). Cells were pelleted and resuspended in Stemcell Mouse MesenCult™ Expansion media (Catalog #05513). tdTomato positive cells were enriched by FACS and subsequently cultured in Mesencult media or DMEM/F12 media supplemented with 10%FBS.

Histo- and Immunohistochemistry

Four micrometers-sections were cut and mounted on Permaflex Plus Microscopic Slides (Leica Biosystems). Sections were stained using standard protocols for Hematoxylin and Eosin (H&E). For immunohistochemistry of murine samples slides were subjected to heat-induced antigen retrieval in a high-pH buffer. Staining was performed using the Ventana Benchmark Ultra automated immunostainer (Ventana Medical Systems, Tucson, AZ, USA). Immunohistochemical staining of murine samples was performed with antibodies to tdTomato (Rockland, anti-RFP, 600-401-379, Dilution 1:400), PDGFRA (Cell Signaling, D1E1E, Dilution 1:100), p-ERK (Cell Signaling, 4376S, clone 20G11), and DESMIN (Abcam, Y66, ready-to-use).

Targeted panel sequencing

Genomic DNA was extracted with QIAamp DNA FFPE Advanced UNG Kit (Cat. # 56704). Mutations in *Kras* (exons 2 and 3), *Nras* (exons 2 and 3), *Hras* (exons 2 and 3), *Trp53* (exons 2-11), and *Dicer1* (exon 24) were sequenced on a NovaSeq 6000 machine using in house designed primers, and libraries were prepped with in house. Targeted panel data was processed using a modified version of the Nextflow⁴⁶ and Sarek⁴⁷ pipeline, version 3.4.2. Raw reads were trimmed with FastP and aligned with bwa-mem2. Overlapping primer intervals were merged with bedtools merge, version 2.30.0, and 100 bp of padding was added to either side of each interval. No duplicate marking was carried out and the “wes” parameter was set to true. Somatic variants were called on the primer intervals with GATK Mutect2 with “ignore_soft_clipped_bases” set to true, and annotated with SnpEff⁴⁸ and Ensembl VEP⁴⁹. For the variant calling step, the pipeline was modified to disable downsampling of the reads. Variants of interest were selected by filtering for those with variant allele fraction > 0.1, and excluding those annotated as “synonymous_variant” or “intron_variant” by SnpEff.

Kras qPCR and FISH

To evaluate *Kras* copy number status, we used a qPCR assay with SYBR Green detection and a custom primer for mouse *Kras* (Chr.6:145220129) and *Tfrc* (Chr.16:32626732). Samples were normalized to *Tfrc* and evaluated for a relative gain in *Kras*. HDT361 showed a high count for *Kras* upon qPCR, and gene amplification was identified using a fluorescence in-situ hybridization (FISH) assay. For this assay, we designed custom probes for *Kras* on 6qA1 (RP23-300B22, green; Empire genomics) and 6qG3 (RP23-118J6, red; Empire genomics).

SWGS and data analysis

DNA was extracted from cored FFPE HDT tumor or non-neoplastic kidney samples and sheared for library preparation and sequencing as we described previously⁵⁰. Single-end reads were processed with a previously validated Nextflow pipeline. Read quality was assessed via FastQC 0.11.9 and reads were trimmed with Trimmomatic 0.39.29. Illumina adapters and leading/trailing bases with quality scores below 3 were removed, as were 4 base sections with average quality scores below 15. Reads shorter than 40 base pairs were dropped. Reads were aligned to mm10 with bwa-mem2 2.1.30. Following alignment, reads were processed with QDNaseq 1.34.32 to determine segmented relative copy-number and call gains and losses using various bin sizes: 30, 50, 100 kbp. The 50 kbp bin size was selected as the best balance of resolution and variance. One sample was excluded due to a highly variable and fragmented copy-number profile, consistent with poor sample quality. For statistical comparisons of copy number alterations, a two-sided Fisher's exact test was used as implemented in the exact2Tomato positive cells were *enric2* R package, which implements improved selection of the matching confidence interval for such two-sided tests.

Bulk RNA-seq

RNA from frozen cell pellets was extracted using the Zymo Quick-RNA Miniprep Kit (Zymo cat# R1054). Total RNA integrity is assessed with the TapeStation 4200 (Agilent Technologies; RNA Assay Cat No.5067-5576) and quantified with the Qubit RNA HS assay (ThermoFisher; Cat No. Q32852). Total RNA is used as input to enrich for mRNA with the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB; Cat No. E7490). Strand-specific libraries are prepared using the DNBSEQ Fast RNA Library Prep kit (Complete Genomics; Cat No. 940-001522-00) and sequenced at Vancouver Prostate Centre. Library quality and size are checked with the TapeStation 4200 (Agilent Technologies; D1000 Assay Cat No. 5067-5582) and then quantified using the Qubit dsDNA HS Assay (ThermoFisher; Cat No. Q32854). The libraries are sequenced using a Complete Genomics DNBSEQ-G400RS High-throughput Sequencer to generate 30 million 100-bp paired-end reads per library.

Bulk RNA-Seq data analysis

Paired-end reads were processed with the nf-core rna-seq pipeline, version 3.14.0 and aligned to mm10. The trimmer was parameter was set to "trimgalore" and the aligner to "star-salmon", with extra salmon quant arguments supplied as "-seqBias -gcBias". Following alignment and quantification, the counts were imported with tximport version 1.26.1 and differential expression between the *Dicer1* mutant and control samples was assessed within each media type using DESeq2 version 1.38.3. Genes with fewer than 10 counts in at least 3 samples were excluded from analysis.

Cloning, lentivirus packaging, and cell proliferation

The gRNA sequences for knocking out Trp53 in murine cells was *aggagctcctgacactcgga*⁵¹ and gRNA2 sequence was *accctgtcaccgagacccc*⁵². They were cloned into the lentiCRISPRv2GFP (Plasmid #82416) vector following the Zhang lab protocol⁵³. For viral packaging, early passage HEK293T17 cells, cultured in DMEM + 10% FBS, authenticated by short tandem repeat analysis and negative for mycoplasma contamination, were seeded in 10 cm dishes and then

transfected with 5 µg of lentiCRISPRv2GFP-p53gRNA plasmids, 4 µg of 2 and 1 µg of VSV.G plasmids using 20 µL of Mirus Bio™ TransIT™-LT1 Transfection Reagent (Catalog #MIR2300). Virus was harvested in DMEM + 10% FBS + 1.1 g/100 ml BSA (Catalog # A2153-10G) + 0.01 M HEPES (Catalog #15630080). All the collected media was filtered with a 0.45 µm filter (Catalog #09-761-35). Beckman Optima XE Ultracentrifuge was prechilled to 4 °C. 36–40 mL of filtered harvest media was aliquoted into a 50 ml tube (#NC9530407) and spun down at 25,000 rpm at 4 °C for 90 min. Reconstituted viral pellet was stored at –80 °C. Fifty microliters of concentrated virus was used per 50,000 cells. Plates were spun at 2000 rpm at 37 °C for 2 h. Selection with eGFP sorting was carried out one week after transduction.

Growth curve generation

Primary cells were seeded in 96-well plates in triplicates (500 cells/well) and were allowed to attach for 24 h to achieve an initial confluency of 10%. Live-image for 7 days was performed with CELLCYTE X™ from Cytena. The mean confluency of cells was plot over time for illustrating their growth behavior.

scRNA-seq

Sample preparation. Tumour and normal kidney tissues were minced to 1 mm³ with forceps and incubated at 37 °C with Stemcell 10× Collagenase/hyaluronidase in DMEM (10× Genomics, C#07912) in HBSS for 30 min. The minced tissues were subsequently incubated in Trypsin-EDTA (0.25%) for 5 min at 37 °C. Cells were then filtered through a 70 µm cell strainer. Live cells were enriched with Miltenyi Biotec Dead Cell Removal Kit. Mesenchymal cell populations were then isolated by FACS sorting with tdTomato positive enrichment and antibody depletion of Epcam, CD45, CD31, CD11b, and F4/80. Cells were then prepared according to the 10× Genomics Chromium Single Cell Universal 3' Gene Expression kit protocol.

Alignment and quality control. Raw reads were aligned with Cell Ranger count 7.1.0 to a combined reference genome, composed of the mm10 reference with the addition of the tdTomato reporter sequence¹⁵. The combined reference genome was generated with Cell Ranger mkref 7.1.0. Following alignment, per-sample quality control was conducted. Low quality cells with > 25% of reads originating from mitochondrial genes, <1000 UMIs detected, or <500 genes detected were discarded.

Integration, clustering, cell-type annotation, and pathway activity.

Samples originating from HDT and/or control mice were integrated using Seurat 4.3.0.1⁵⁴. Counts were normalized per cell by library size followed by log1p transformation. The top 2000 highly variable genes per sample were selected using the VST method and the top 2000 genes across samples were used for integration via reciprocal principal components analysis (PCA). After integration and dimensionality reduction via PCA, shared nearest neighbors were identified in PCA space and clustered with the Louvain algorithm, using the first 30 components. The resulting clusters were manually annotated via the examination of differentially expressed upregulated genes, in combination with markers of kidney and *Hic1*⁺ MSC types obtained from the literature. To achieve a pure *Hic1*⁺ MSC lineage following examination of all cell types, the data was subset to the relevant clusters in the control (Uni Fibro, Fibro, Mural, Cycle) or HDT (Fibro Perturb, Fibro, Mural, Cycle, TR Diff Myo, Diff Myo) samples. If necessary, residual tdTomato- cells, or cells expressing epithelial (*Epcam*⁺, *n* = 102), immune (*Ptprc*⁺, *n* = 113), or transitional epithelial (*Krt19*⁺, *n* = 376) markers were excluded from the subset data. The resulting data, filtered for the *Hic1*⁺ MSC lineage, was reprocessed as described above for the unfiltered data; a reduced number of principal components were used for neighbor finding and clustering after selection via an elbow plot.

Trajectory inference. The *Hic1*⁺ MSC lineage obtained from the HDT samples was used as input for trajectory inference with Monocle 3, version 1.3.4⁵⁵. Based on the UMAP obtained during dimensionality reduction with Seurat, trajectory was inferred, and the cells were placed in a pseudotime ordering, setting the node in the graph surrounded by the greatest number of *Dpt*⁺ cells as the earliest point. Trajectories to Ground, Cycle, and Diff Myo were examined by selecting the relevant branches, holding the start point constant, and varying the end point to the most distant node in the ending cluster of interest. Selected genes of interest were visualized in pseudotime over the respective trajectories.

Copy number analysis. To infer copy-number from the scRNA-seq data, inferCNV 3.2.0 was used. The raw counts for all cells present in the control samples and the *Hic1*⁺ MSC lineage from the HDT samples were combined and used as input to inferCNV. All parameters were set to their default setting, except for “cutoff”, which was set to 0.1 as prescribed for 10× data in the inferCNV documentation.

Xenium spatial transcriptomics

Slide preparation and imaging. For spatial transcriptomic analyses (10× Genomics Xenium), four slides were prepared. Two slides each contained a section from one of the two murine TMAs. A third slide included full kidney sections from three control mice and one HDT mouse harboring a LGMT-like lesion. The fourth slide featured a section of the human TMA. Tissue sections were prepared as per the Xenium In Situ for FFPE—Tissue Preparation Guide (10× Genomics, CG000578 version Rev C) protocol. Five micrometer sections were cut and mounted onto the Sample Area of a Xenium slide. The slides were air dried, followed by a 3 h incubation at 42 °C. Subsequently, the slides were processed, over a two-day span, as per the Xenium In Situ for FFPE—Deparaffinization and Decrosslinking (10× Genomics, CG000580 version Rev C) and Xenium In Situ Gene Expression (10× Genomics, CG000582 version Rev C) protocols. On day one, slides were incubated at 60 °C for 2 h on a Xenium thermocycler adaptor plate (BioRad, C1000 Touch Thermocycler), followed by deparaffinized through a series of immersions in xylene, 100%, 96%, and 70% ethanol, and rehydration with nuclease-free water. Slides were assembled into Xenium cassettes for all downstream steps, starting with decrosslinking and permeabilization. Gene panel probes were hybridized overnight (minimum 18 h) at 50 °C. The pre-designed dataset for the murine slide included 379 genes (10× Genomics, Mouse Tissue Atlas Panel, PNI000627). The pre-designed dataset for the human slide included 379 genes (10× Genomics, Human Multi-Tissue and Cancer Panel, PNI000626). On day two, slides were subjected to a post-hybridization wash (37 °C for 2 h), ligation (37 °C for 2 h), amplification (30 °C for 2 h), autofluorescence quenching and nuclei staining steps. Slides were then stored in PBS-T, overnight, at 4 °C, in the dark. The following day, slides and consumables were loaded onto the Xenium Analyzer instrument according to the Xenium Analyzer (10× Genomics, CG000584, version Rev B) protocol. The whole slide area was selected for probe decoding and image acquisition, and the run was initiated. After run completion, slides were H&E stained as per the “Xenium in Situ Gene Expression—Post Xenium Analyzer H&E Staining” (10× Genomics, CG000613, version Rev A) protocol, and scanned using a slide scanner (AperioAT2, Leica).

Segmentation and Xenium data analysis. The Xenium bundles were segmented with Xenium Ranger 1.7.1.1, with the expansion distance around nuclei set to 5 μm. The bundles were then re-segmented with Proseg 1.1.4 using the “xenium” preset, converted to the Baysor format with the “proseg-to-baysor” utility bundled with Proseg, and finally re-imported with Xenium Ranger 1.7.1.1³⁶. Cell visualization and morpho-molecular correlation was performed in the 10× Xenium Explorer (Version 3.0.1). Resulting Xenium objects were analyzed

using Seurat 4.3.0.1⁵⁴. Quality control was conducted and low-quality cells with less than 3 counts and less than 3 features, were discarded. After normalization, integration, and dimensionality reduction via PCA, shared nearest neighbors were identified in PCA space and clustered with the Louvain algorithm. The resulting clusters were manually annotated via the examination of differentially expressed upregulated genes, in combination with markers of kidney and *Hic1*⁺ MSC types obtained from the literature, as well as morpho-molecular correlation. To achieve a pure tumor cell-associated lineage following examination of all cell types, the data was subset to the relevant clusters in the patient samples (Fibro, Progen, Ground, TR Diff Myo, Diff Myo, and Prolif). The resulting data was reprocessed as described above for the unfiltered data; a reduced number of principal components were used for neighbor finding and clustering after selection via an elbow plot. Trajectory inference of patient samples was performed with Monocle 3, version 1.3.4⁵⁵. On the basis of the UMAP obtained during dimensionality reduction with Seurat, trajectory was inferred, and the cells were placed in a pseudotime ordering, setting the node in the graph surrounded by the greatest number of *DPT*⁺ cells as the earliest point.

For comparison of human cell states to metaprograms of RMS tumors³³, RMS gene signatures were intersected with genes present on the targeted panel. Based on a limited representation of the RMS signatures, the AddModuleScore function of Seurat was used to infer metaprogram scores onto human cell states.

Statistics and reproducibility

Sample size determination was based on anticipated variability and the effect size that was observed in the investigator’s lab for similar experiments. No data were excluded from the analyses. Kaplan–Meier plots were generated using Prism 8. For pairwise comparisons, unpaired two-tailed *t* tests were used to calculate *P*-values. Both statistical methods were carried out using Prism (GraphPad Software v. 9). *p*-values < 0.05 were considered significant.

Reporting summary

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

Data availability

The raw and processed murine scRNA-seq data has been deposited into NCBI GEO under accession #: [GSE288990](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE288990). The raw and processed murine Xenium data has been deposited into NCBI GEO under accession #: [GSE289001](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE289001). The raw data for murine targeted panel data and sWGS is available in NCBI SRA under accession #: [PRJNA1221971](https://www.ncbi.nlm.nih.gov/sra/study/PRJNA1221971). Mouse primary cell culture bulk RNA-Seq raw data has been deposited into NCBI under accession #: [GSE309820](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE309820). For human spatial transcriptomic data, as participants did not agree to publicly share their raw transcriptomic data by deposition in a public repository, their raw data can be shared upon request for research-related, non-commercial purposes through the corresponding authors. All requests will be reviewed and processed within one month. There is no specific time limit for how long data will be available once the access has been granted for specified research use through the data access agreement. Source data are provided with this paper.

Code availability

All code to reproduce the analyses described in this study is available on GitHub: [dicer1-cell-of-origin](https://github.com/Huntsmanlab/dicer1-cell-of-origin) [<https://github.com/Huntsmanlab/dicer1-cell-of-origin>].

References

- de Kock, L., Wu, M. K. & Foulkes, W. D. Ten years of DICER1 mutations: provenance, distribution, and associated phenotypes. *Hum. Mutat.* **40**, 1939–1953 (2019).

2. Gonzalez, I. A. et al. DICER1 tumor predisposition syndrome: an evolving story initiated with the pleuropulmonary blastoma. *Mod. Pathol.* **35**, 4–22 (2022).
3. de Kock, L., Priest, J. R., Foulkes, W. D. & Alexandrescu, S. An update on the central nervous system manifestations of DICER1 syndrome. *Acta Neuropathol.* **139**, 689–701 (2020).
4. Hill, D. A. et al. DICER1 mutations in familial pleuropulmonary blastoma. *Science* **325**, 965 (2009).
5. Heravi-Moussavi, A. et al. Recurrent somatic DICER1 mutations in nonepithelial ovarian cancers. *N. Engl. J. Med.* **366**, 234–242 (2012).
6. Wang, Y. et al. The oncogenic roles of DICER1 RNase IIIb domain mutations in ovarian Sertoli-Leydig cell tumors. *Neoplasia* **17**, 650–660 (2015).
7. Anglesio, M. S. et al. Cancer-associated somatic DICER1 hotspot mutations cause defective miRNA processing and reverse-strand expression bias to predominantly mature 3p strands through loss of 5p strand cleavage. *J. Pathol.* **229**, 400–409 (2013).
8. McCluggage, W. G. & Foulkes, W. D. DICER1-sarcoma: an emerging entity. *Mod. Pathol.* **34**, 2096–2097 (2021).
9. Kommoss, F. K. F. et al. Genomic characterization of DICER1-associated neoplasms uncovers molecular classes. *Nat. Commun.* **14**, 1677 (2023).
10. Wang, Y. et al. Biallelic Dicer1 mutations in the gynecologic tract of mice drive lineage-specific development of DICER1 syndrome-associated cancer. *Cancer Res.* **83**, 3517–3528 (2023).
11. Bean, G. R., Anderson, J., Sangoi, A. R., Krings, G. & Garg, K. DICER1 mutations are frequent in Müllerian adenosarcomas and are independent of rhabdomyosarcomatous differentiation. *Mod. Pathol.* **32**, 280–289 (2019).
12. Oikawa, K. et al. Liver-specific DICER1 syndrome model mice develop cystic liver tumors with defective primary cilia. *J. Pathol.* **264**, 17–29 (2024).
13. Wagh, P. K. et al. Cell- and developmental stage-specific Dicer1 ablation in the lung epithelium models cystic pleuropulmonary blastoma. *J. Pathol.* **236**, 41–52 (2015).
14. Fraire, C. R. et al. An imbalance between proliferation and differentiation underlies the development of microRNA-defective pineoblastoma. *Genes Dev.* **39**, 727–750 (2025).
15. Scott, R. W., Arostegui, M., Schweitzer, R., Rossi, F. M. V. & Underhill, T. M. Hic1 defines quiescent mesenchymal progenitor subpopulations with distinct functions and fates in skeletal muscle regeneration. *Cell Stem Cell* **25**, 797–813 e9 (2019).
16. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
17. Antonescu, C. R., Reuter, V. E., Keohan, M. L., Hwang, S. & Chi, P. DICER1-associated anaplastic sarcoma of the kidney with coexisting activating PDGFRA D842V mutations and response to targeted kinase inhibitors in one patient. *JCO Precis Oncol.* **6**, e2100554 (2022).
18. Kommoss, F. K. F. et al. Clinicopathologic and molecular analysis of embryonal rhabdomyosarcoma of the genitourinary tract: evidence for a distinct DICER1-associated subgroup. *Mod. Pathol.* **34**, 1558–1569 (2021).
19. Diaz Coronado, R. Y. et al. Primary central nervous system sarcoma with DICER1 mutation-treatment results of a novel molecular entity in pediatric Peruvian patients. *Cancer* **128**, 697–707 (2022).
20. Koelsche, C. et al. Primary intracranial spindle cell sarcoma with rhabdomyosarcoma-like features share a highly distinct methylation profile and DICER1 mutations. *Acta Neuropathol.* **136**, 327–337 (2018).
21. Hiemcke-Jiwa, L. S. et al. Pleuropulmonary blastoma (PPB) and other DICER1-associated high-grade malignancies are morphologically, genetically and epigenetically related - A comparative study of 4 PPBs and 6 sarcomas. *Ann. Diagn. Pathol.* **60**, 152002 (2022).
22. Abbasi, S. et al. Distinct regulatory programs control the latent regenerative potential of dermal fibroblasts during wound healing. *Cell Stem Cell* **28**, 581–583 (2021).
23. Soliman, H. et al. Pathogenic potential of hTat: treatment of 1-expressing cardiac stromal progenitors. *Cell Stem Cell* **26**, 459–461 (2020).
24. Novella-Rausell, C., Grudniewska, M., Peters, D. J. M. & Mahfouz, A. A comprehensive mouse kidney atlas enables rare cell population characterization and robust marker discovery. *iScience* **26**, 106877 (2023).
25. Kragestein, B. K. et al. The transcriptional and regulatory identity of erythropoietin producing cells. *Nat. Med.* **29**, 1191–1200 (2023).
26. Muhl, L. et al. Single-cell analysis uncovers fibroblast heterogeneity and criteria for fibroblast and mural cell identification and discrimination. *Nat. Commun.* **11**, 3953 (2020).
27. Buechler, M. B. et al. Cross-tissue organization of the fibroblast lineage. *Nature* **593**, 575–579 (2021).
28. Dahlgren, M. W. & Molofsky, A. B. Adventitial cuffs: regional hubs for tissue immunity. *Trends Immunol.* **40**, 877–887 (2019).
29. Sitnik, K. M. et al. Context-dependent development of lymphoid stroma from adult CD34(+) adventitial progenitors. *Cell Rep.* **14**, 2375–2388 (2016).
30. Merrick, D. et al. Identification of a mesenchymal progenitor cell hierarchy in adipose tissue. *Science* **364**, eaav2501 (2019).
31. McKellar, D. W. et al. Large-scale integration of single-cell transcriptomic data captures transitional progenitor states in mouse skeletal muscle regeneration. *Commun. Biol.* **4**, 1280 (2021).
32. Tabula Sapiens, C. et al. The tabula sapiens: a multiple-organ, single-cell transcriptomic atlas of humans. *Science* **376**, eabl4896 (2022).
33. Danielli, S. G. et al. Single cell transcriptomic profiling identifies tumor-acquired and therapy-resistant cell states in pediatric rhabdomyosarcoma. *Nat. Commun.* **15**, 6307 (2024).
34. Yin, Y. et al. Fibroblast growth factor 9 regulation by MicroRNAs controls lung development and links DICER1 loss to the pathogenesis of pleuropulmonary blastoma. *PLoS Genet.* **11**, e1005242 (2015).
35. Hettmer, S. & Wagers, A. J. Muscling in: uncovering the origins of rhabdomyosarcoma. *Nat. Med.* **16**, 171–173 (2010).
36. Rubin, B. P. et al. Evidence for an unanticipated relationship between undifferentiated pleomorphic sarcoma and embryonal rhabdomyosarcoma. *Cancer Cell* **19**, 177–191 (2011).
37. Drummond, C. J. et al. Hedgehog pathway drives fusion-negative rhabdomyosarcoma initiated from non-myogenic endothelial progenitors. *Cancer Cell* **33**, 108–124 e5 (2018).
38. Searcy, M. B. et al. PAX3-FOXO1 dictates myogenic reprogramming and rhabdomyosarcoma identity in endothelial progenitors. *Nat. Commun.* **14**, 7291 (2023).
39. Charytonowicz, E., Cordon-Cardo, C., Matushansky, I. & Ziman, M. Alveolar rhabdomyosarcoma: is the cell of origin a mesenchymal stem cell? *Cancer Lett.* **279**, 126–136 (2009).
40. Yaseen, W. et al. Fibroblast fusion to the muscle fiber regulates myotendinous junction formation. *Nat. Commun.* **12**, 3852 (2021).
41. Arostegui, M., Scott, R. W., Bose, K. & Underhill, T. M. Cellular taxonomy of Hic1(+) mesenchymal progenitor derivatives in the limb: from embryo to adult. *Nat. Commun.* **13**, 4989 (2022).
42. Zhang, B. et al. A human embryonic limb cell atlas resolved in space and time. *Nature* **635**, 668–678 (2024).
43. Hill, L. A. et al. The fibroblast epigenome underlies SS18::SSX-mediated transformation in synovial sarcoma. *Nat. Commun.* **16**, 10830 (2025).
44. Yi, R. et al. Morphogenesis in skin is governed by discrete sets of differentially expressed microRNAs. *Nat. Genet.* **38**, 356–362 (2006).

45. Alkushi, A. et al. Identification of prognostically relevant and reproducible subsets of endometrial adenocarcinoma based on clustering analysis of immunostaining data. *Mod. Pathol.* **20**, 1156–1165 (2007).
46. Di Tommaso, P. et al. Nextflow enables reproducible computational workflows. *Nat. Biotechnol.* **35**, 316–319 (2017).
47. Hanssen, F. et al. Scalable and efficient DNA sequencing analysis on different compute infrastructures aiding variant discovery. *NAR Genom. Bioinform.* **6**, lqae031 (2024).
48. Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).
49. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
50. Jamieson, A. et al. Targeted and shallow whole-genome sequencing identifies therapeutic opportunities in p53abn endometrial cancers. *Clin. Cancer Res.* **30**, 2461–2474 (2024).
51. Bieging-Rolett, K. T. et al. Zmat3 is a key splicing regulator in the p53 tumor suppression program. *Mol. Cell* **80**, 452–469 e9 (2020).
52. Leibold, J. et al. Somatic tissue engineering in mouse models reveals an actionable role for WNT pathway alterations in prostate cancer metastasis. *Cancer Discov.* **10**, 1038–1057 (2020).
53. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
54. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 e29 (2021).
55. Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
56. Jones, D. C. et al. Cell simulation as cell segmentation. *Nat. Methods* **22**, 1331–1342 (2025).

Acknowledgements

We appreciate the technical support of Dionzie Ong, Rebecca Ho, Eunice Li, Emma Guo, Christine Chow, Yimei Qin, and all staff in the Animal Research Centre of British Columbia Cancer Research Institute. We would also like to express our gratitude to the collaborating clinicians, as well as to the patients who generously provided their samples for this research. This project was supported by research funds from a Terry Fox New Frontiers Program Project Grant (#1082 and #1155, D.G.H., U.M., and Y.W.), Canadian Institutes of Health Research (CIHR) grants (MOP-154290, D.G.H.; PJT#178179, Y.W.; PJT#195754, D.G.H., Y.W., and G.B.M.; FDN-148390, W.D.F.), and the philanthropic supports to British Columbia Ovarian Cancer Research Centre through the BC Cancer Foundation and the VGH/UBC foundation. F.K.F.K. is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 523898075. J.Y.H.Z. is a recipient of CIHR CGS-D. Y.W. is supported by an early career investigator award from Ovarian Cancer Research Alliance (ECIG-2025-3-2014). D.G.H. is supported by the Canadian Research Chair program in Molecular and Genomic Pathology.

Author contributions

F.K.F.K.: resource, methodology, investigation, formal analysis, visualization, writing; J.Y.H.Z.: methodology, investigation, formal analysis, visualization, writing; B.L.: formal analysis, data curation, visualization, writing; S.C. and J.S.: Methodology, investigation; Y.M., L.A.H., and J.B.: investigation; D.M.: formal analysis, data curation; W.S.: methodology; K.S.C.: discussion and manuscript editing; A.R. and A.v.D.: resource; W.D.F.: resource and manuscript editing; G.B.M.: discussion and funding acquisition; T.M.U.: resource, supervision, manuscript editing; Y.W.: conceptualization, methodology, investigation, writing, supervision, project management, funding acquisition. D.G.H.: conceptualization, supervision, funding acquisition. All authors have read and approved the manuscript.

Competing interests

T.M.U. is a co-founder of Mesintel Therapeutics Inc. and Mesentech Inc. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-70971-6>.

Correspondence and requests for materials should be addressed to Yemin Wang or David G. Huntsman.

Peer review information *Nature Communications* thanks Jason Chan and the other anonymous reviewers for their contribution to the peer review of this work. [A peer review file is available].

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026

¹Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada. ²Institute of Pathology, University Hospital Heidelberg, Heidelberg, Germany. ³Department of Basic and Translational Research, British Columbia Cancer Research Institute, Vancouver, BC, Canada. ⁴Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada. ⁵Department of Cellular and Physiological Sciences, University of British Columbia, Vancouver, BC, Canada. ⁶Biological Sciences Platform, Holland Bone and Joint Program, Sunnybrook Research Institute, Department of Laboratory Medicine & Pathobiology, University of Toronto, Toronto, ON, Canada. ⁷British Columbia Children's Hospital Research Institute, Vancouver, BC, Canada. ⁸Department of Pediatrics, Children's Medical Center Research Institute, University of Texas Southwestern Medical Center, Dallas, TX, USA. ⁹Department of Neuropathology, Heidelberg University Hospital, Heidelberg, Germany. ¹⁰Clinical Cooperation Unit Neuropathology, German Consortium for Translational Cancer Research (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany. ¹¹Department of Human Genetics, McGill University,

Montreal, QC, Canada. ¹²Cancer Axis, Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, QC, Canada. ¹³Cancer Research Program, Research Institute of the McGill University Health Centre, Montreal, QC, Canada. ¹⁴Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Research Institute, Vancouver, BC, Canada. ¹⁵Ovarian Cancer Research Centre and Vancouver Prostate Centre, Vancouver Coastal Health Research Institute, Vancouver, BC, Canada. ¹⁶Department of Obstetrics and Gynaecology, University of British Columbia, Vancouver, BC, Canada. ¹⁷These authors contributed equally: Felix K. F. Kommoss, Joyce Yu Han Zhang, Branden J. Lynch. ¹⁸These authors jointly supervised this work: Yemin Wang, David G. Huntsman. ✉ e-mail: yemin.wang@ubc.ca; dhuntsma@bccancer.bc.ca