communications biology

Article

https://doi.org/10.1038/s42003-024-06478-x

Single-cell transcriptional profiling of clear cell renal cell carcinoma reveals a tumor-associated endothelial tip cell phenotype

Check for updates

Justina Zvirblyte[1], Juozas Nainys[1,5], Simonas Juzenas[1], Karolis Goda[1], Raimonda Kubiliute[2], Darius Dasevicius[3], Marius Kincius[4], Albertas Ulys[4], Sonata Jarmalaite[2,4] ✉ & Linas Mazutis[1] ✉

Clear cell renal cell carcinoma (ccRCC) is the most prevalent form of renal cancer, accounting for over 75% of cases. The asymptomatic nature of the disease contributes to late-stage diagnoses and poor survival. Highly vascularized and immune infiltrated microenvironment are prominent features of ccRCC, yet the interplay between vasculature and immune cells, disease progression and response to therapy remains poorly understood. Using droplet-based single-cell RNA sequencing we profile 50,236 transcriptomes from paired tumor and healthy adjacent kidney tissues. Our analysis reveals significant heterogeneity and inter-patient variability of the tumor microenvironment. Notably, we discover a previously uncharacterized vasculature subpopulation associated with epithelial-mesenchymal transition. The cell-cell communication analysis reveals multiple modes of immunosuppressive interactions within the tumor microenvironment, including clinically relevant interactions between tumor vasculature and stromal cells with immune cells. The upregulation of the genes involved in these interactions is associated with worse survival in the TCGA KIRC cohort. Our findings demonstrate the role of tumor vasculature and stromal cell populations in shaping the ccRCC microenvironment and uncover a subpopulation of cells within the tumor vasculature that is associated with an angiogenic phenotype.

The asymptomatic nature of clear cell renal cell carcinoma (ccRCC), the most common renal cancer, often leads to diagnosis in late III or IV stage with survival probability of 59% and 20%, respectively, also, ~30% of cases metastasize[1]. Previous efforts aimed at characterizing ccRCC tumors have provided valuable insights into the genomic[2], transcriptomic and epigenetic[3,4] landscape of both the tumor and the tumor microenvironment (TME). It is now well-established that the most abundant genomic alterations in ccRCC involve the loss of regions in 3p chromosome (occurring in >90% of cases) and von Hippel–Lindau gene mutations (>50% of cases). These alterations lead to impaired degradation and abnormal accumulation of hypoxia-inducible factors[2,3], resulting in a highly vascularized tumor appearance. Moreover, ccRCC tumors exhibit a high degree of immune infiltration[5,6]. Consequently, the most common first-line treatment options for the localized disease involve surgical removal of the tumor, while advanced disease may be treated with VEGF pathway inhibitors, standalone or in combination with immune checkpoint blockade therapies[2,7,8]. However, owing to a high degree of intra- and inter-tumor heterogeneity, these treatments benefit only a fraction of patients, and often result in acquired resistance and further disease progression[2,9].

Recent advancements in microfluidics and molecular barcoding have enabled high-throughput transcriptional, epigenomic and even multi-omic tissue profiling at the single cell resolution, yielding important biological insights. For instance, using single-cell RNA sequencing (scRNA-seq) a plethora of single-cell resolution healthy and cancerous tissue atlases have been constructed, revealing the phenotypic complexity and plasticity of the tumor microenvironment[10–13]. In the context of ccRCC, single-cell techniques have shed light on the cell of origin of ccRCC[14,15], malignancy-related transcriptional programs of the tumor[16] and the heterogeneous

[1]Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius 10257, Lithuania. [2]Institute of Biosciences, Life Sciences Center, Vilnius University, Vilnius 10257, Lithuania. [3]National Center of Pathology, Affiliate of Vilnius University Hospital Santaros Klinikos, Vilnius 08406, Lithuania. [4]National Cancer Institute, Vilnius 08660, Lithuania. [5]Present address: Droplet Genomics, Vilnius 10257, Lithuania. ✉e-mail: sonata.jarmalaite@nvi.lt; linas.mazutis@bti.vu.lt

tumor-associated immune cell infiltrate[17–20]. Furthermore, the phenotypical changes of immune cell populations along advancing disease stage[21] and immunotherapy treatment[18,22] have been characterized in detail.

Upon the widespread adoption of the single cell profiling techniques there was a noticeable paradigm shift in the field of cancer research—a systemic view of the tumor as a highly orchestrated ecosystem took over the tumor cell-centric point of view. This shift has highlighted the crucial role of other players in the TME, including various subpopulations of stromal and endothelial cells that have been discovered to have an impact on disease progression, response to therapy and patient survival[23,24]. While considerable efforts have been made to characterize the ccRCC tumor microenvironment at the single-cell level, most of the previous studies focused on tumor or immune cells, leaving the role of other cell types within the ccRCC TME poorly understood. In this study, we aimed to address this gap by profiling fresh ccRCC tumor and matched healthy adjacent tissue samples using droplet-based scRNA-seq, omitting cell sorting and enrichment steps in order to capture the diverse phenotypes present in the TME, including the stromal cell populations. As a result, we captured all major specialized epithelial and endothelial cell populations in healthy adjacent kidney tissue, including a progenitor-like epithelial cell phenotype resembling the cell of origin for ccRCC. Furthermore, we described five tumor endothelium subpopulations and discovered a previously uncharacterized tip-like cell phenotype. Within the TME, we identified well-described immunosuppressive tumor associated macrophage (TAM) populations and exhausted infiltrating T cells[21]. Through cell-cell communication analysis, we inferred the interactions between various cell types within the TME, revealing tumor vasculature and stromal cell involvement in maintaining an immunosuppressive niche. Expression of genes involved in these interactions was associated with worse overall survival in the TCGA KIRC cohort. Overall, our results complement ongoing ccRCC TME characterization efforts by introducing a tumor-associated endothelial phenotype and highlighting the importance as well as potential therapeutic relevance of stromal and endothelial cells in the TME.

Results

Single-cell profiling of healthy and tumor tissues reveals inter-patient variability and epithelial ccRCC progenitor-like population in healthy tissue

To dissect the transcriptional landscape of the human ccRCC tumor microenvironment (TME), we profiled fresh tumor ($n = 8$) and healthy adjacent ($n = 9$) kidney tissue samples (histology slides provided in Supplementary Fig. S1) using a droplet-based scRNA-seq platform (Fig. 1a). To capture the diverse range of cell types constituting the TME, our experimental strategy involved rapid isolation of dissociated cells in microfluidic droplets, without any enrichment or sorting steps (see Methods). Following quality control, batch correction and doublet removal (see Methods), we obtained a total of 50,236 single-cell transcriptomes that were then clustered using graph-based spectral clustering. The cell types belonging to each cluster were identified manually based on differentially expressed top 25 marker genes (adjusted p value < 0.05; cluster vs the rest of cells, Mann-Whitney U test with Benjamini-Hochberg correction), validated by an extensive literature review (Fig. 1b, f and Supplementary Information Table 1).

Healthy-adjacent samples displayed all major epithelial and endothelial cell populations characteristic of a healthy kidney (Fig. 1b)[25–27]. By omitting the cell enrichment step, we could successfully capture diverse cell types that are known to be highly sensitive to handling and extended workflow procedures[28]. For example, we captured both, ascending (*DNASE1L3*) and descending (*AQP1*, *SLC14A1*) parts of the vasa recta, as well as glomerular endothelium marked by *IGFBP5* and *SOST* expression. The epithelial compartment encompassed cells from various specialized nephron segments, including rare populations such as intercalating cells of type A and B (expressing marker genes *ATP6V1G3* and *SLC26A4*, respectively), as well as podocytes (*NPHS2*, *PODXL*). Interestingly, in contrast to tumor, all healthy tissue samples comprised a population of epithelial progenitor-like cells, similar to that described by ref. 14 (Fig. 1e) and ccRCC "cell of origin" PT-B phenotype delineated by Zhang et al., (Supplementary Fig. S2a). This population expressed genes associated with de-differentiated injured kidney epithelium, such as *PROM1* and *ITGB8*[29], as well as *CD24* and *SOX4*, which have been implicated in kidney development and mark proximal tubule and distal nephron response to acute kidney injury[30] (Fig. 1e). Therefore, the epithelial progenitor-like cell population in our dataset likely represents a de-differentiated phenotype, similar to a potential cell of origin for ccRCC disease (PT-B, Supplementary Fig. S2a).

The tumor samples encompassed localized and locally advanced pT1a and pT3a pathologic stages of ccRCC (Fig. 1c, Supplementary Data S1). These samples exhibited high immune cell infiltration, including several populations of tumor-associated macrophages and T cells (Fig. 1b). The stromal cells separated into myofibroblast (type I, IV and VI collagens, *FN1*, *TIMP2*, *ACTA2*), vascular smooth muscle cell (*TAGLN*, *ACTA2*, *SNCG*) and mesangial/vSMC (*BGN*, *PDGFRB*, *TAGLN*) clusters. Tumor endothelium completely separated from healthy-adjacent endothelial populations (Fig. 1b) and included ascending vasa recta-like cells (*ACKR1*, *DNASE1L3*) as well as heterogeneous vasculature subpopulations expressing tumor-associated endothelial markers *PLVAP*, *VWF*, *SPARC*, *INSR*, *ANGPT2*, and others (Supplementary Data S2, S3). Tumor vasculature exhibited distinct expression patterns as compared to healthy endothelium (Fig. 1f). While four out of five vasculature subpopulations identified in our data have been described previously[14–16], one tumor vasculature subpopulation (Tumor vasculature 3 comprising 151 cells) appeared to be novel in the context of ccRCC and featured upregulation of *LY6H*, *PGF*, *LOX*, *CHST1*, and type IV collagen (Fig. 1f), consistent with a tip-cell phenotype[31].

The tumor cells in all samples expressed canonical markers *CA9*, *NDUFA4L2*, *VEGFA* and segregated into three subpopulations, out of which one (Tumor cells 1) was patient-specific (126 cells in population, 120 of them specific to patient P9, Supplementary Fig. S2b). Notably, these cells exhibited elevated expression of progenitor-like phenotype marker *SLC17A3*, which was not highly expressed in the healthy-adjacent epithelial progenitor cells (Fig. 1e, Supplementary Fig. S2c). Furthermore, Tumor cells 1 population was the most distinct from other tumor cells based on unsupervised hierarchical clustering (Fig. 1f, Supplementary Fig. S2c). These cells over-expressed genes such as vitamin D binding protein *GC* and *HLA-G*, the latter being involved in immunosuppressive interactions (Fig. 2c), as well as *FABP7*, crucial for lipid uptake and storage in hypoxic conditions when de novo lipid synthesis is repressed[32]. Additionally, these cells were marked by high expression of pan-cancer marker *MDK*[33], along with *IFI27* and *SOD2* (Supplementary Fig. S2c), both of which play a role in interferon response[22]. Consistently, Tumor cells 1 was the only tumor cell population not enriched for hypoxia, but instead enriched for oxidative phosphorylation and adipogenesis. Considering the elevated expression of *VCAM1* and *SLC17A3*, it is possible to envision that this small patient-specific population could represent an intermediate progenitor-tumor cell phenotype.

The cellular composition of tumor tissues, as expected, displayed noticeable variability across the patients as compared to their matched pair of healthy-adjacent tissues (Fig. 1d, Supplementary Data S4). A common theme to all tumor samples was a high number of immune cells infiltrating the TME, accompanied by almost complete loss of specialized kidney-specific epithelial and endothelial cell populations (Fig. 1c, d and Supplementary Fig. S2b). Except for Tumor cells 1, no other cell phenotype was patient-specific; cell population composition analysis by patient ID confirmed adequate representation of cells of different origins (Supplementary Fig. S2b). To quantitively assess tumor sample heterogeneity, we calculated Shannon entropy for each broad cell category[11]. Low entropy values for a cell phenotype indicate that it is rarely shared between samples, meaning that the level of heterogeneity within samples is high. In tumor samples, the heterogeneity was highest for stromal, endothelial and tumor cells, whereas healthy adjacent tissue samples exhibited comparatively lower heterogeneity (Supplementary Fig. S2d, e). Such diverse TME snapshots among different patients in our and other ccRCC studies[15,34] suggest that patient stratification may rely on the abundance of specific cellular phenotypes

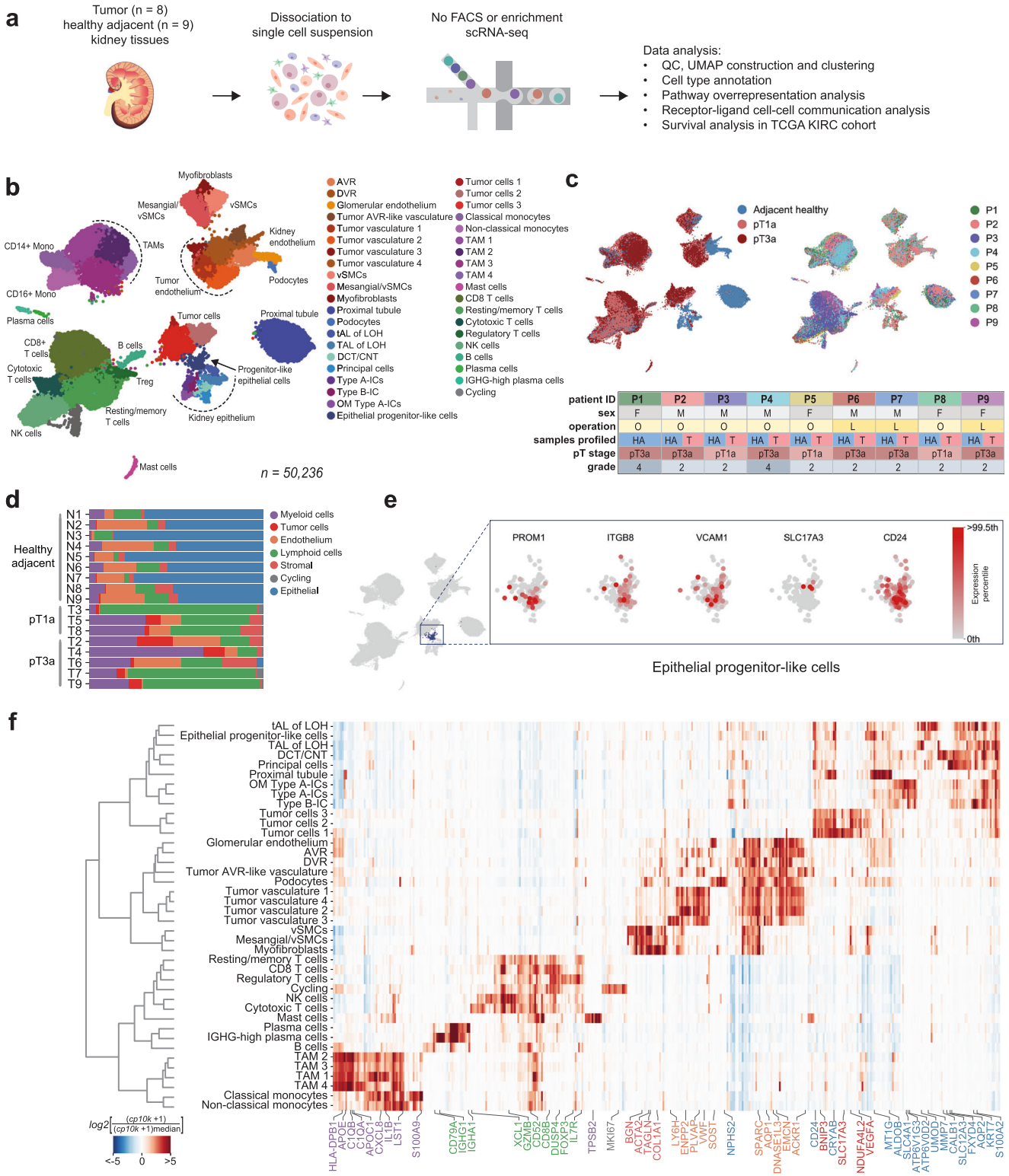

Fig. 1 | Profiling the ccRCC microenvironment. a Experimental design. **b** Global single cell transcriptional map of ccRCC. **c** Clinical information of collected samples and corresponding UMAPs of cells annotated by disease stage (adjacent healthy, pT1a and pT3a) and patient ID (P1–P9). Healthy adjacent samples (blue) almost completely separate from the tumor (light and dark red). **d** Sample composition by major cell type. Notably, healthy adjacent samples are enriched with specialized kidney epithelial and endothelial cells, while tumor samples are enriched for immune cells. **e** Expression of ccRCC cell of origin markers in epithelial progenitor-like cell population. **f** Global heatmap for population-specific markers. Only genes with Benjamini-Hochberg adjusted *p* value < 0.05 are shown. Color of the gene name indicates major cell type. AVR ascending vasa recta, DVR descending vasa recta, vSMCs vascular smooth muscle cells, LOH loop of Henle, tAL thin ascending limb, TAL thick ascending limb, DCT/CNT distal convoluted/connecting tubule, ICs intercalated cells, OM outer medullary, TAM tumor associated macrophages. All graphic elements in the figure were created by the first author.

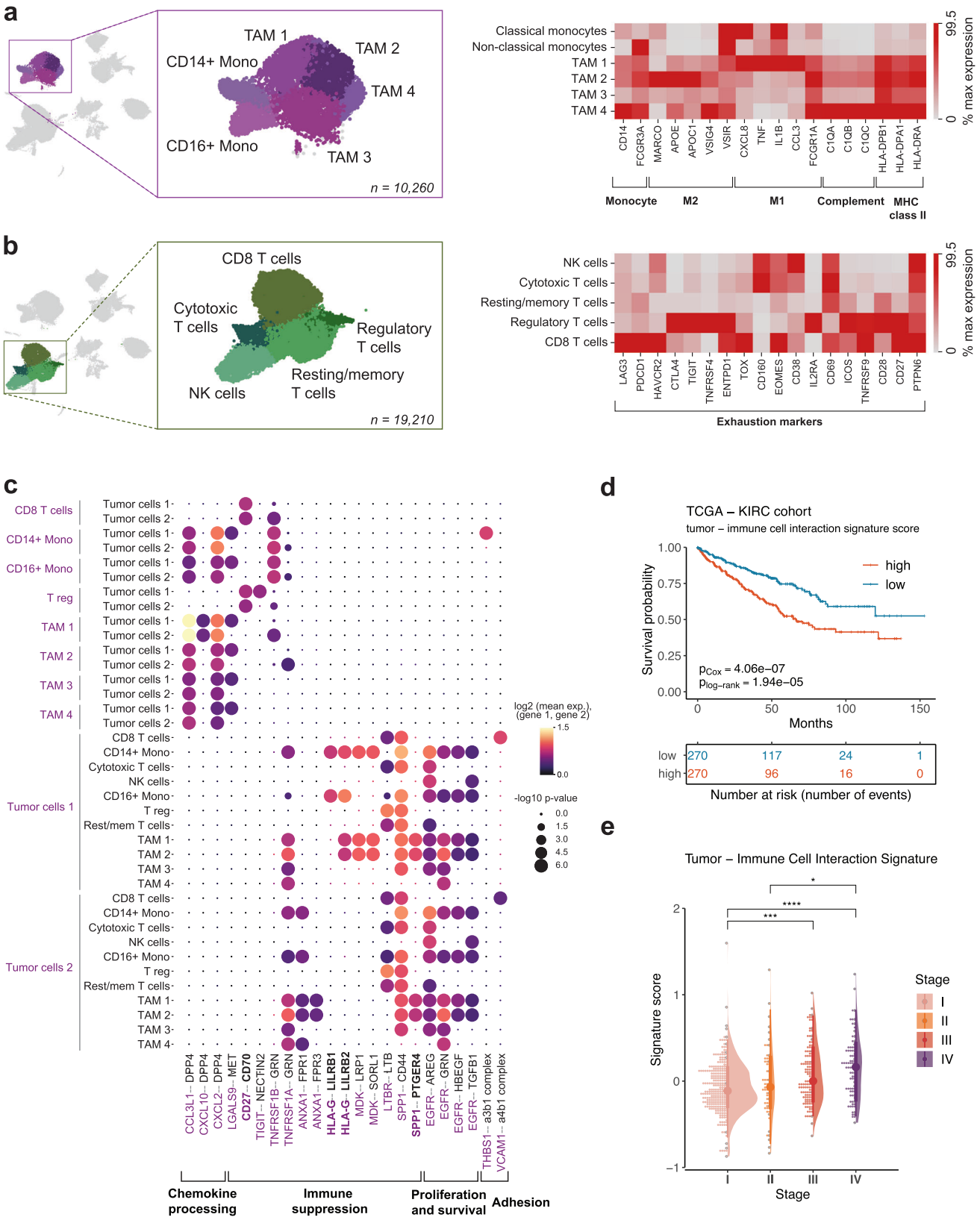

Fig. 2 | Characterization of immune cell populations found in ccRCC. a Myeloid cell compartment consists of CD14+ and CD16+ monocytes and four populations of tumor associated macrophages diverse in expression of polarization markers. **b** Lymphoid cells in ccRCC display heterogeneous exhaustion profile. **c** Immunosuppressive interactions of clinical importance revealed by cell-cell communication analysis between immune and tumor cells using CellPhoneDB. **d** Tumor-immune cell interaction signature expression in TCGA KIRC cohort is associated with a worse overall survival. **e** Tumor-immune cell interaction signature increases along the progression of the ccRCC disease.

within the TME, rather than patient-specific phenotypes. This underscores the importance of revisiting strategies for biomarker selection to aid personalized treatment options in ccRCC.

Tumor-associated macrophages exhibit phenotypic heterogeneity and immunosuppressive tumor–immune interaction signature is associated with poor survival

ccRCC is recognized as highly immune infiltrated tumor with a dynamic microenvironment. The compositional changes that occur along tumor stage progression[21] and in response to immunotherapy treatment[22,35] have a profound impact on patient survival. Therefore, the phenotypic states of immune populations represent potentially druggable targets for advanced and metastatic ccRCC treatments.

Within the immune compartment, we identified all major lymphoid and myeloid cell populations including plasma cells (*IGKC*, *IGHG1*), B cells (*CD79A*, *MS4A1*), mast cells (*TPSB2*), NK cells (*GZMB*, *NKG7*), classical (*CD14*) and non-classical (*FCGR3A*) monocytes and two major groups of T cells and macrophages (Fig. 1b), in concordance with previous ccRCC studies[18,19,21]. As expected, the tumor samples were enriched in TAMs that clustered into four transcriptionally distinct subpopulations (Fig. 2a). The TAM 1 and TAM 2 cells expressed genes hinting towards M1 and M2 polarization, respectively (Fig. 2a), thus encompassing a traditional view of TAM dichotomy. However, TAM 3 and TAM 4 subpopulations did not follow a clear activation pattern, despite their marker genes seemed to reflect an alternatively activated macrophage phenotype (Fig. 1f, Supplementary Information Table 1). For example, while the expression of certain immunosuppressive genes, such as *MARCO*, were clearly diminished in TAM 3/4 cells, other immune-response modulating genes such as *VSIG4*[36] or *VSIR* were highly expressed in TAM 4 population. In addition, among all TAM populations, TAM 4 demonstrated the highest expression of complement system C1Q genes (Fig. 2a), products of which are known to promote tumor progression in ccRCC by interacting with tumor-produced complement system molecules[37]. Interestingly, some complement components were not only specific to the tumor cells but also present in the stromal compartment, suggesting potential stromal cell involvement in tumor progression (Supplementary Fig. S3a). These findings support the notion that ccRCC TME is enriched in suppressive macrophages that adapt to the microenvironment-derived signals influencing disease progression[6,10,21].

The lymphoid compartment predominantly consisted of CD8 T cells (*CD8B*, *DUSP4*), CD4 regulatory T cells (*FOXP3*, *TNFRSF4*), resting/memory T cells (*IL7R*, *CD52*), cytotoxic T cells (*XCL1*, *KLRB1*) and natural killer cells (*GZMB*, *NKG7*). These subpopulations expressed multiple exhaustion markers (Fig. 2b), with classic immune-checkpoint molecule *PDCD1* expressed abundantly in CD8 T cell cluster and *CTLA4* enriched in regulatory T cells. The cytotoxic T cell population shared the exhaustion pattern with NK cells characterized by high expression of *CD160*, *EOMES*, *CD38* and *CD69*. As expected, resting/memory T cells displayed the least exhausted phenotype compared to other lymphoid cell populations (Fig. 2b). Given the established exhaustion profile of lymphoid cells and immunosuppressive phenotype of myeloid cells[18,21,38], we evaluated the crosstalk of these immune cell populations and tumor cells.

Receptor-ligand analysis (see Methods) revealed multiple interactions involved in chemokine processing, immune suppression and sustained survival of tumor cells (Fig. 2c, Supplementary Data S5, S6). For example, tumor cells were predicted to communicate with monocytes and TAMs through the immune checkpoint *HLA-G–LILRB1/2* axis, which is involved in promoting the immunosuppressive M2 phenotype and immune escape of the tumor[39]. Interestingly, both pro-inflammatory (M1) and anti-inflammatory (M2) TAMs received signals from tumor cells via *SPP1–PTGER4* interaction, known to promote macrophage polarization towards tumor-supporting phenotype in hepatocellular carcinoma[40]. Another important interaction observed in the TME involved T-cell co-stimulatory *CD27–CD70* axis, targeted at CD8 T cells and CD4 regulatory T cells. Recent studies have shown that this cell-cell interaction is associated with a pro-tumoral effect, primarily driven by chronic stimulation of T cells leading to exhaustion, enhanced survival of regulatory T cells, and recruitment of TAMs[41]. Furthermore, the expression of interaction signature (gene set of both receptors and ligands, Supplementary Data S7) was associated with significantly lower overall survival (Fig. 2d, Supplementary Data S8) and steadily increased along the progression of the disease in the TCGA KIRC dataset (Fig. 2e). Therefore, our analysis of the ccRCC TME reveals the extensive network of immune and cancer cell interactions that are involved in establishing an immune-suppressive TME for sustained tumor survival and growth.

Tumor endothelial cells are diverse and play a role in re-shaping the tumor microenvironment, associated with worse overall survival

The highly vascularized appearance of ccRCC tumors is often attributed to the abnormal accumulation of hypoxia-inducible factors[2,3] that create pseudohypoxic conditions and subsequently increase production of angiogenic factors. To this day, the heterogeneity and possible regulatory role of the tumor vasculature in ccRCC remains poorly described. Focusing on ccRCC endothelium in our scRNA-seq dataset we identified five tumor vasculature (TV) subpopulations (Fig. 3a) that were markedly distinct from healthy kidney endothelium (Fig. 3b) and featured upregulation of genes important in vascularization, angiogenesis and disease progression. For instance, among the multiple overexpressed genes (Supplementary Data S9), the TV cells displayed elevated levels of the fenestration marker *PLVAP*, which is recognized as a therapeutic target in hepatocellular carcinoma[42]; *ANGPT2*, which stimulates angiogenesis in autocrine manner and is involved in recruitment of immunosuppressive TAMS[43]; *IGFBP7*, which is clinically used acute kidney injury urinary biomarker[44]. Moreover, endothelial migration stimulating insulin receptor (*INSR*) was overexpressed in tumor endothelium and is known to be associated with poor overall survival in bladder cancer, which, similarly to ccRCC, can become resistant to VEGF pathway targeted therapy[45]. These findings highlight the abnormal, fenestrated nature of tumor endothelial cells and might provide future guidance for tumor-specific vasculature identification in ccRCC.

Within the tumor vasculature we found an ascending vasa recta-like population that was transcriptionally closer to the healthy endothelium cells than to other tumor vasculature cells (Fig. 3c), as noted in previous work[15]. Intriguingly, our ccRCC atlas also unveiled an uncharacterized population of tumor vasculature (referred to as TV 3) that appeared as the most distinct from the rest of TV cells (Fig. 3c). This population was marked by high expression of tip cell markers *LOX*, *PXDN*, *LY6H* and *PGF*[31,46] (Supplementary Fig. S4a, Supplementary Data S10), characteristic of a tip cell phenotype. Furthermore, TV 3, along with TV 1 and TV 4, displayed elevated expression of extracellular matrix constituents, including pro-angiogenic and potentially pro-metastatic collagen type IV and perlecan (*HSPG2*) (Fig. 3c)[47–49]. Meanwhile, TV 2 overexpressed multiple genes implicated in tumor progression, such as VEGF receptor *FLT1*, *ESM1*, *ANGPT2*, *KCNE3*, coagulation factor VIII (*F8*) (Fig. 3c), which are involved in tumor-associated angiogenesis[49,50]. In addition, TV 2 was marked by high expression of autotaxin (*ENPP2*), a potent stimulator of tumor development and invasion, which has been associated with acquiring resistance to the antiangiogenic drug sunitinib in ccRCC[51] (Fig. 3c). Interestingly, a fraction of cells from all tumor vasculature sub-populations expressed *INHBB* and *SCGB3A1* (Supplementary Fig. S4a), which, in concert with perivascular *TNC* (in our dataset expressed by myofibroblasts, Fig. 5a, b), have recently been demonstrated to orchestrate the pro-metastatic niche in lung metastasis models in mice[52]. Thus, the tumor vasculature in ccRCC appears to be highly heterogeneous and expresses a variety of angiogenesis-related and tumor-promoting factors.

Subsequently, we investigated the potential interactions between tumor vasculature and other cell types within the TME. Cell-cell communication analysis using CellPhoneDB[53] revealed crosstalk between vascular and immune cells involved in angiogenesis, immune suppression and adhesion (Fig. 3d, Supplementary Fig. S3b). Unexpectedly, our analysis revealed that tumor vasculature delivers immunosuppressive signals

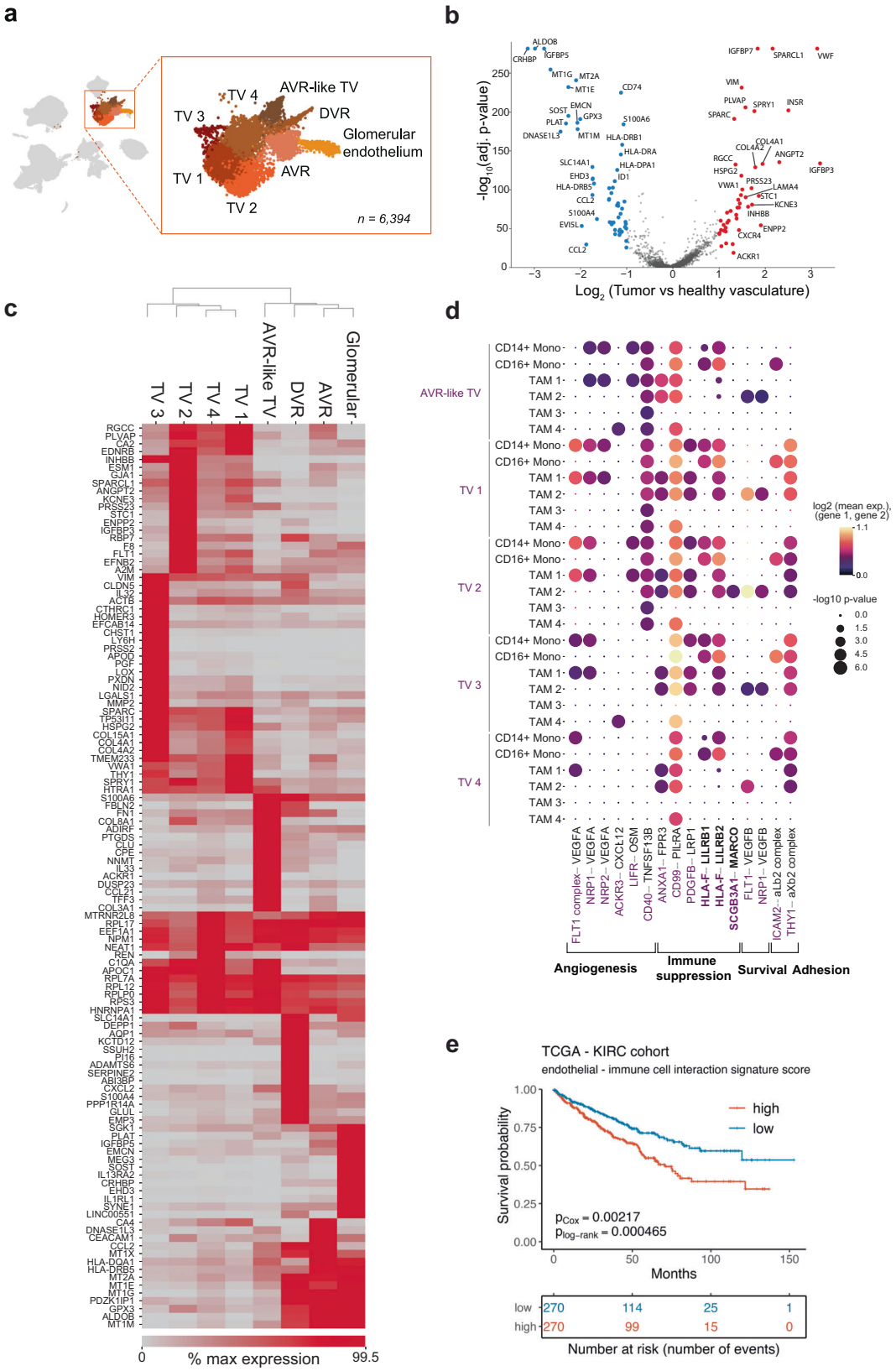

Fig. 3 | Assessing the heterogeneity of tumor vasculature of ccRCC. a A close-up of endothelial cell subpopulations. **b** Tumor and healthy vasculature comparison shows upregulation of angiogenesis related genes in tumor vasculature. **c** Differential gene expression between vasculature subpopulations. Only genes with Benjamini-Hochberg adjusted *p* value < 0.05 are shown. **d** Tumor endothelium and myeloid cells demonstrate abundant cell-cell interactions. **e** Collective tumor vasculature–immune cell communication signature expression is associated with a worse overall survival in TCGA KIRC dataset. AVR ascending vasa recta, DVR descending vasa recta, TV tumor vasculature.

previously thought to be confined to the tumor cells, such as the interactions between *TIGIT* and *NECTIN2* (Supplementary Fig. S3b) or *HLA-F* and *LILRB1/2* (Fig. 3d). Also, we observed several known interactions mediated by myeloid cell produced TNF-α with tumor endothelium i.e *TNF* – *NOTCH1* (Supplementary Fig. S3b), which induces *JAG1* expression and enhances migration and proliferation of endothelial cells upon subsequent VEGF exposure[54]. Importantly, a higher degree of cell-cell communication between tumor vasculature and immune cells, as evaluated by higher expression of receptor and ligand pairs, was found to result in a significantly lower overall survival in TCGA KIRC cohort (Fig. 3e).

These findings suggest notable tumor vasculature participation in tumor progression and tumor microenvironment shaping through the expression of angiogenesis-related genes, tumor-promoting extracellular matrix molecules, and active immunosuppressive communication with immune cells.

A subpopulation of tumor endothelium expresses genes involved in epithelial-mesenchymal transition associated with worse patient survival

The tip cell-like tumor vasculature population (TV 3 in Fig. 3a) expressed *LOX*, *PXDN*, *LY6H* and *PGF*, which are not only denoted as tip cell markers, but have also been implicated in tumor growth promotion within the TME. For example, placental growth factor (*PGF*), a member of VEGF family, can directly interact with VEGF receptors and increase vascular permeability while promoting M2 macrophage polarization[55]. In *PGF*-deficient mice, tumor-associated M1 macrophage polarization is largely restored while tumor vasculature appears normalized[56]. Lysil oxidase *LOX* and peroxidase *PXDN* are involved in cross-linking of the collagen type IV rich extracellular matrix and basement membrane, which is essential for growth factor induced endothelial cell proliferation and survival[57]. Inhibition of ECM cross-linking through lysil oxidase knockdown has been shown to impair vessel sprouting[31]. The transcriptional profile of tumor vasculature 3 population corresponded to angiogenic tip-cell phenotype extensively characterized by ref. 31 (Supplementary Fig. S4b) and could potentially be involved in promoting tumor progression.

Molecular Signatures Database Hallmark gene set over-representation analysis in tumor, tumor vasculature and stromal cell populations (top 100 marker genes) revealed, as expected, hypoxia and glycolysis terms in tumor cells (Fig. 4a, Supplementary Data S11). However, this analysis also uncovered an enrichment of epithelial-mesenchymal transition (EMT) associated genes in all tumor vasculature and stromal cell subpopulations. Interestingly, the overexpression of EMT pathway overlapping genes for AVR-like tumor vasculature (Fig. 4b) and TV 3 population (Fig. 4c) was associated with a significantly worse overall survival in the TCGA KIRC cohort. In this context, it is important to note that the specific genes overlapping with the EMT differed between these subpopulations (Supplementary Data S12). Also, even though other cell populations, such as stromal cells and the rest of tumor vasculature had a significant overlap with the EMT pathway (Supplementary Fig. S5a), no effect on patient survival in the TCGA KIRC cohort was observed (Supplementary Figs. S5b–g). Overall, our findings highlight the presence of a tip cell-like tumor endothelium subpopulation associated with an aggressive phenotype, potentially influencing ccRCC disease progression and survival.

Stromal cells remodel the ECM and potentially contribute to immunosuppression of TAM populations

Finally, we investigated the putative roles of stromal cells in the ccRCC tumor microenvironment. While stromal cells have been recognized as important components of the TME[34], their specific contribution in ccRCC have received much less attention compared to immune or tumor cells. Graph-based clustering of our dataset revealed three cell populations within the stromal cells: vascular smooth muscle cells (vSMCs), myofibroblasts and mesangial/vSMCs (Fig. 5a, b, Supplementary Data S13). The vSMCs expressed markers *TAGLN*, *ACTA2,* and *MYH11*, while myofibroblasts were enriched for ECM constituents (Collagen types I, III, IV, VI and fibronectin) including markers *TIMP1* and *ACTA2* (Fig. 5b). The precise annotation of the third stromal cell population was challenging due to simultaneous upregulation of mesangial marker *PDGFRB* and vSMC genes (Supplementary Information Table 1). Interestingly, this population featured substantial transcriptional differences between tumor and healthy tissue (Supplementary Fig. S6, Supplementary Data S14). In tumor samples, the mesangial/vSMC population overexpressed tumor marker *NDUFA4L2* as well as some stress-related genes, such as *CD36*, which is upregulated in chronic kidney disease and associated with poor prognosis in ccRCC[58,59], and renin (*REN*), which is expressed by mesangial cells under disturbed homeostasis[60] (Supplementary Fig. S6). Thus, it appears that the mesangial/vSMC population is reactive to the disruptive microenvironmental changes exerted by the tumor.

Cell-cell interaction analysis between stromal and immune cells revealed putative interactions related to stromal cell proliferation and survival, as well as immune cell suppression and adhesion. Majority of immunosuppressive signals originating from the stromal cells were directed at TAM 1 and TAM 2 subpopulations (Fig. 5c). For instance, we identified *ANXA1–FPR1* interaction, which is involved in anti-inflammatory macrophage polarization and tumor progression in various cancers[61,62]. Furthermore, we found an indication of myofibroblast and mesangial/vSMC communication with cytotoxic T cells via *HLA-E–KLRC1*, which has recently been proposed as a new targetable path of T cell exhaustion in bladder cancer[63]. Treatment of *HLA-E* positive tumors with anti-KLRC1 antibodies has shown a strong effect in restoring the anti-tumor immunity[64]. Interestingly, our analysis shows that this communication signature is associated with worse overall survival in the TCGA KIRC dataset (Fig. 5d), and the expression of genes involved in the stromal-immune cell communication increased with advancing stage of the disease (Fig. 5e). Collectively, our results suggest that stromal cells could be actively involved in modulating the tumor microenvironment in ccRCC through therapeutically relevant paths.

Discussion

The single-cell transcriptomic studies have provided valuable insights about the origin of ccRCC[14,15], malignancy programs of the tumor[16], immune cell population phenotypical changes during tumorigenesis[21] and immunotherapy treatment[18,22] among other. Complementing these ongoing efforts to better characterize ccRCC tumor microenvironment we profiled single-cell transcriptomes of human ccRCC tumor samples along with healthy adjacent tissues. In contrast to previous studies that used cell enrichment prior to scRNA-seq, our strategy relied on a rapid isolation of cells from ccRCC specimens, without involving any type of sorting or cell enrichment. As a result, we could capture a rich diversity of cells constituting heterogeneous TME that were either significantly depleted or absent in previous studies. Given that immune compartment in our dataset largely recapitulated previous findings[17–22], we mainly focused on the phenotypic heterogeneity and cellular interactions of the often overlooked and underappreciated endothelial and stromal cell populations.

Endothelial cells are very important in ccRCC tumorigenesis and to this day remain the main targets of therapeutics in advanced and metastatic disease[2]. The tumor endothelial cells identified in our study include a previously uncharacterized tip cell phenotype, enriched for epithelial-mesenchymal transition pathway genes that are associated with poor patient overall survival. Indeed, the previous single-cell studies in ccRCC have also captured endothelial cells, however, these were most often represented by two major phenotypic subpopulations that are also found in our ccRCC atlas. For instance, Zhang et al. reported *ACKR1*+ and *EDNRB*+ endothelium, while Long et al. reported *VCAM1*+ and *VCAM1*- vasculature populations. Consistently, in our dataset we find a population co-expressing ascending vasa recta marker *ACKR1* and *VCAM1* (tumor AVR-like vasculature), however, *EDNRB* is expressed by tumor vasculature 1, 2, and 4 populations, but not by tumor vasculature 3 (Supplementary Fig. S4a), further supporting that this endothelial (*PECAM1*+) phenotype has not been characterized in ccRCC.

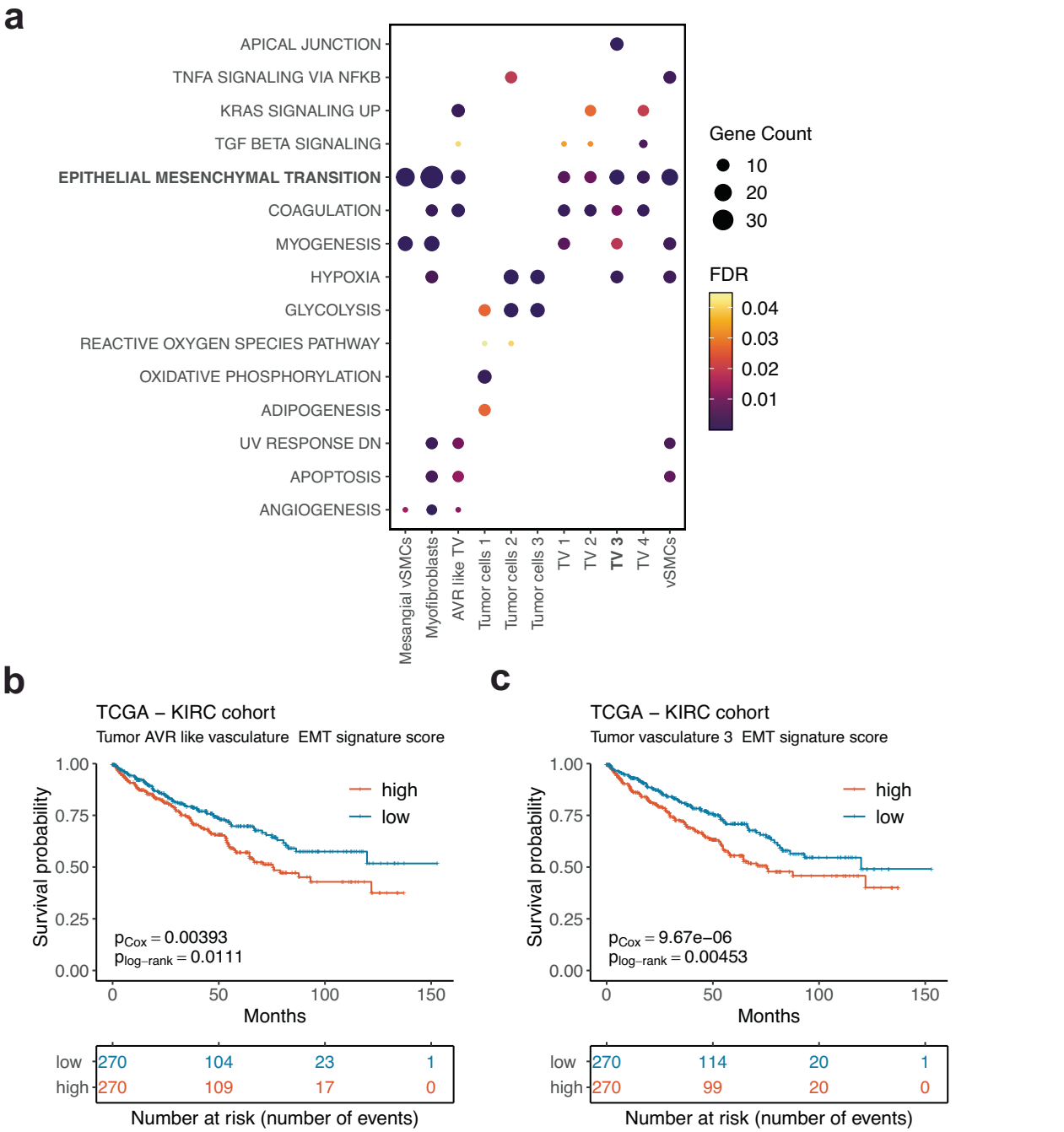

Fig. 4 | MSigDB Hallmark pathway overrepresentation analysis. a Tumor vasculature and stromal cell populations are enriched in epithelial-mesenchymal transition (EMT) signature. **b** Tumor AVR-like vasculature and **c** tip-like tumor vasculature 3 signature genes overlapping with EMT pathway associate with worse overall survival in the TCGA KIRC cohort.

The tip cell population (TV 3) in our dataset is very similar to a tip cell population observed in lung cancer (*LOX*, *PXDN*, *PGF*, *LXN*, collagen type IV enriched, Supplementary Fig. S4a, b) where it was shown to correlate with worse patient survival[31]. The authors have found this phenotype the most congruent across several species and tumor types, including kidney cancer (as determined by bulk proteomics), which raises a question about why previous single-cell studies of ccRCC did not capture this rare population. Furthermore, the authors demonstrated that tip cell marker *LOX* knock-down impaired vessel sprouting, suggesting that the reported population in ccRCC might be of interest for future research as a potential therapeutic target.

In line with our findings, Long et al. showed that *VCAM1*+ population (labeled as AVR-like tumor vasculature in our dataset) is enriched for EMT signature[16], yet our pathway over-representation analysis indicates similar association with EMT for all tumor vasculature and stromal cell populations, not just the AVR-like population (Fig. 4a). On another hand, the worse overall survival in association with EMT was pronounced only for AVR-like and the tumor vasculature 3 populations, further emphasizing the diversity of tumor endothelial cells and potential importance of the reported tip cell phenotype. Alchahin et al. also reported association with EMT for endothelial and stromal cells, but did not discriminate healthy kidney and tumor endothelial cells. On the contrary to our findings, they report lower endothelial cell abundance in tumor samples as compared to healthy tissues[20]. Such discrepancies between different studies can be related to technical aspects, for instance, processing of the samples, and further

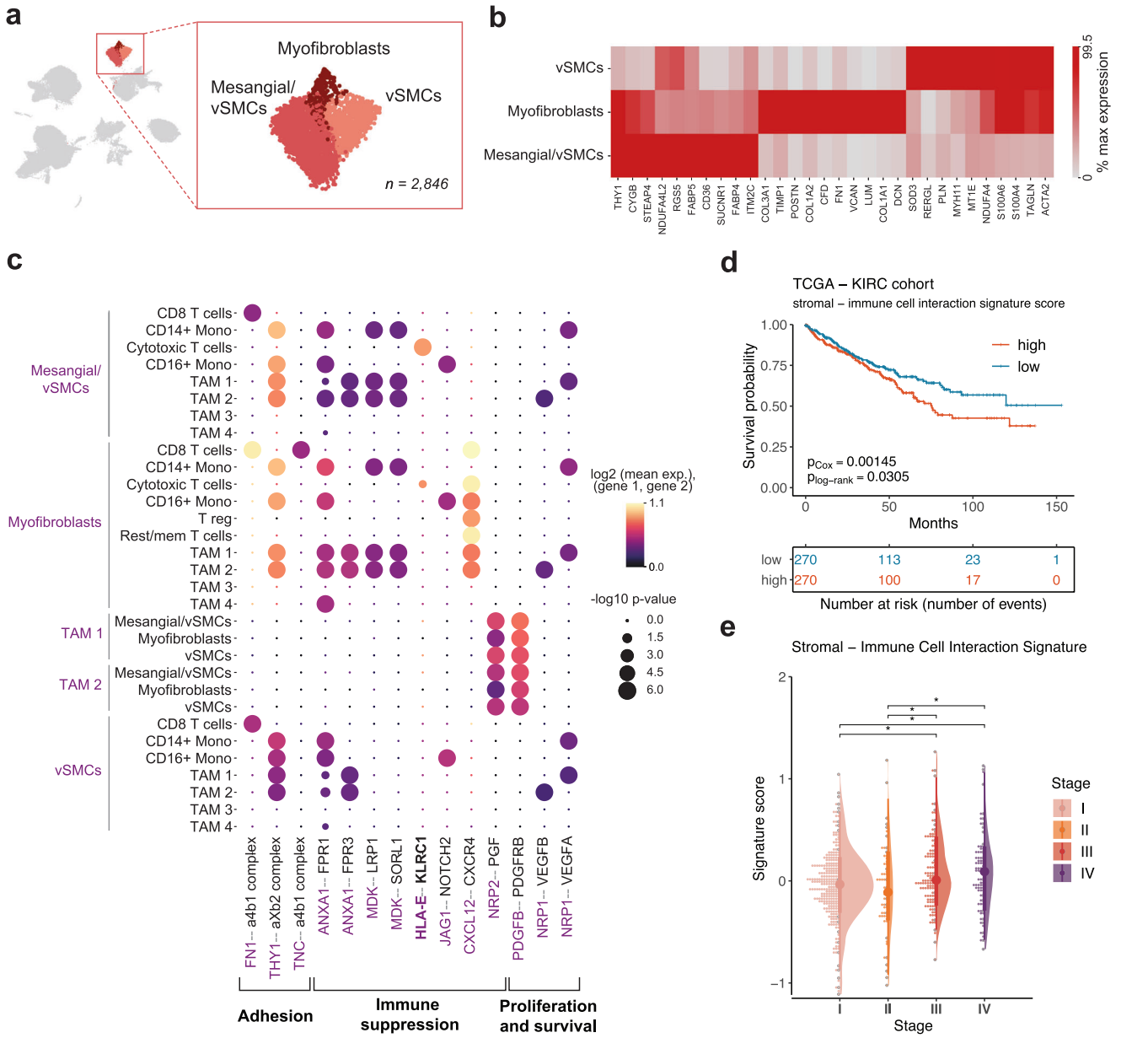

Fig. 5 | Assessing the heterogeneity of stromal cells in the TME. a Stromal cell populations consisting of vSMCs, myofibroblasts and mesangial/vSMCs. **b** Differential gene expression between stromal cell subpopulations. Only genes with Benjamini-Hochberg adjusted *p* value < 0.05 are shown. **c** Stromal and immune cells exhibit immunosuppressive interactions mediated by stromal cells. **d** Expression of collective stromal-immune cell interaction signature gene set associates with worse overall survival in the TCGA KIRC cohort. **e** Stromal-immune cell interaction signature expression increases along the progression of the ccRCC disease. vSMCs vascular smooth muscle cells.

underline the importance for accurate phenotypic characterization of the tumor vasculature cells in ccRCC.

Our findings suggest two major modes of action of the tumor vasculature cells in the TME. First, remodeling of the ECM by active deposition of various ECM constituents and expression of their modifying agents related to EMT (i.e., *LOX*, *PXDN* in tumor vasculature 3) and second, active engagement in cellular communication in the tumor microenvironment, mostly involved in immune suppression and angiogenesis maintenance. Interestingly, spatial transcriptomic profiling of ccRCC by *Li et al.*, showed that collagen producing endothelial cells localize at the tumor–normal interface enriched in EMT-high tumor cells and *IL1B*+ macrophages[17]. These findings are also corroborated by our results suggesting that tumor endothelial cells might indeed contribute to EMT in ccRCC and interact with TAMs. The cell-cell communication analysis uncovered diverse interactions of clinical relevance enriched in the tumor vasculature and stromal cell communication with immune cells (Figs. 3d, 5c). For instance, in 2021, a phase I–II clinical trial (ID NCT04913337) began for LILRB1 and LILRB2 inhibitor as a monotherapy or in combination with Pembrolizumab (anti PD-1) for advanced or metastatic solid tumors, including ccRCC. Inhibition of LILRB2 reprograms myeloid cells to a stimulatory (pro-inflammatory) state, while inhibition of LILRB1 stimulates the reprogramming of both myeloid and lymphoid cells. Our analysis suggests that *LILRB1/2*+ immune cells interact not only with tumor cells, but also with endothelial cells. Similarly, endothelial cell-expressed *NECTIN2* associated with *TIGIT* expressed by regulatory T cells, an interaction that has gained increased attention over the last few years and is currently exploited in a multitude of clinical trials[65]. Another intriguing interaction observed between TV 2 and TAM 2 populations was *SCGB3A1–MARCO*. As demonstrated recently, *SCGB3A1*, a secreted secretoglobin family member produced by endothelial cells, is a crucial component of a pro-metastatic

niche and induces stem cell properties in cancer cells, while macrophages are also required for the niche maintenance[52]. However, *SCGB3A1–MARCO* interaction in ccRCC, to our knowledge, has not been described.

It is worth emphasizing that stromal cells in our dataset were involved in communication with immune cells in a suppressive manner, suggesting their participation in maintaining a pro-tumorigenic niche, especially considering the difference of mesangial/vSMCs population expression in tumor vs healthy adjacent tissue. Moreover, the communication signature expression associated with worse overall survival and increased along the progression of the disease in the TCGA KIRC dataset. On a side note, increase of stromal cells has recently been shown in recurrent RCC as compared to primary disease, furthermore, stromal cell-produced Galectin-1 inhibitor significantly reduced tumor mass and improved anti-PD-1 immunotherapy efficacy in murine models[66]. Another report showed that co-targeting stromal cells expressing PDGFRs and endothelial cells expressing VEGFRs delays tumor vascularization and has clinical efficacy in pancreatic neuroendocrine tumors[43]. Therefore, there is a need for in-depth characterization of ccRCC stromal cells and further validation of their pro-tumorigenic properties. Understanding the role of stromal cells in the TME could provide valuable insights for the development of targeted therapies.

Overall, our study introduces a tumor-associated endothelial tip cell phenotype and provides new insights into the characterization of the TME in ccRCC. We propose that tumor endothelial cells favor tumor progression and potentially metastatic dissemination through the expression of metastasis promoting factors, specific extracellular matrix components and indirectly via targetable interactions with immune cells in the TME. Undoubtedly, future functional studies are needed to elucidate the exact roles of the described diverse tumor endothelial cells and explore their potential as therapeutic targets in ccRCC.

Study limitations

Like any other, this study is not without limitations. Single-cell RNA sequencing results generally suffer from data sparsity and tissue dissociation biases. The latter is particularly relevant to adhesive cells, such as epithelial or tumor cells, that are more challenging to dissociate into single cell suspension as opposed to infiltrating immune cells[28]. Therefore, even though immune cell infiltration is a common characteristic of ccRCC, the exact cellular composition of tumors in our as well as other scRNA-seq datasets[16,18,20] is likely to be affected by the dissociation protocols and other experimental variables, inflating the immune compartment at the expense of the tumor cell capture. We aimed at minimizing these biases by reducing the sample handling time in order to extend the viability of cells, and deliberately avoided the FACS that is known to cause damage to the fragile cells. Whilst our efforts led to a recovery of rich cell phenotypes, including ccRCC endothelial tip-cell population, future studies will be required to validate our findings. Moreover, functional in-vitro and in-vivo characterization will be necessary to elucidate the role of tip-cell population in the disease progression or response to therapy, as such experiments were out of scope of this work. Finally, another compromise taken due to the selected study design involves sacrifices to scRNA-seq data quality. The data sparsity did not permit us applying imputation, pseudotime or RNA velocity algorithms that could provide further insights into tumor biology. Nonetheless, despite the existing limitations, our study reveals previously under-characterized cell populations and their putative interactions thereby not only complementing ccRCC characterization, but also suggesting new directions for future research.

Methods

Sample acquisition

Fresh ccRCC tumor ($n = 8$) and healthy-adjacent ($n = 9$) paired kidney tissues were obtained from the National Cancer Institute (Vilnius, Lithuania) with informed patient consent and a Vilnius Regional Bioethics Committee approval No.2019/2-1074-586. All ethical regulations relevant to human research participants were followed. No patient had received prior systemic therapy for their cancer. Samples were collected during an open or laparoscopic, partial or radical nephrectomy surgery, placed on ice and rapidly (<1 h) transferred to the laboratory for dissociation. Sample T1 (tumor from patient P1) was highly necrotic, thus excluded from analysis. Clinical characteristics of all samples profiled are provided in Supplementary Data S1.

Sample processing

Sample preparation was performed according to the scRNA-seq protocol[67], yet without FACS-based enrichment. Briefly, patient-derived tumor tissues were dissociated using Tumor Dissociation Kit (Miltenyi Biotec, cat. no.130-095-929) in an automated instrument gentleMACS Octo Dissociator with Heaters (Miltenyi Biotec) as per manufacturer's instructions. Healthy-adjacent tissues were dissociated using Tissue Dissociation Kit I (Miltenyi Biotec, cat. no. 130-110-201). After dissociation, red blood cells were removed from the samples using RBC lysis reagent (Miltenyi Biotec, cat. no.130-094-183). After RBC lysis, cells were washed three times in ice-cold 1X DPBS (Gibco, cat. no. 14080-048) at 500 *g* for 5 min. Cell viability and count were assessed using Trypan Blue dye (Gibco, cat. no. 15250061) on a hemocytometer. No further enrichment or selection of cells was performed. Cell suspension was immediately loaded onto inDrops platform[68] for cell barcoding experiment.

Single-cell barcoding, library preparation and sequencing

Dissociated cells were isolated in 1 nl droplets and their transcriptomes barcoded using a modified version of inDrops protocol[69]. Specifically, instead of linear cDNA amplification by in vitro transcription we used template switching and PCR amplification. For that purpose, we isolated the cells at occupancy 0.1 alongside barcoding beads (Atrandi Biosciences, cat. no. DG-BHB-C) and reverse transcription/lysis mix, the latter supplemented with a template-switching oligonucleotide, TSO (see Table 1 for composition). We used cell barcoding chip (Atrandi Biosciences, cat.no. MCN-05) to inject the cells, DNA barcoding beads, and RT/lysis mix at flow rates of 250, 60, 250 µl/h, respectively. The droplet stabilization oil (Atrandi Biosciences, cat. no. MON-DSO2) was set at 700 µl/h. The emulsion was collected off-chip on ice rack and briefly exposed to UV light (5 min at 6.5 J/cm^2 of 350 nm, Atrandi Biosciences, cat.no. MHT-LAS2) to release the photo-cleavable RT primers from the barcoding hydrogel beads. The RT

Table 1 | Lysis/RT reaction mix for single-cell mRNA barcoding

Reagent	Amount, µl	Concentration in droplet
Nuclease-free water	21	---
5X RT buffer (Thermo Scientific, Cat. No EP0751)	60	1X
TSO primer (0.5 mM)	15	25 µM
dNTP (10 mM each, Thermo Scientific, Cat. No. R0192)	15	0.5 mM
10% (v/v) Igepal CA-630 (Sigma Aldrich, Cat. No. 18896-50 mL)	9	0.3%
RiboLock RNAse Inhibitor (Thermo Scientific, Cat. No. EO0382)	15	1 U/ul
Maxima H Minus Reverse Transcriptase (Thermo Scientific, Cat. No. EP0751)	15	10 U/ul
Total volume	150	---

reaction was performed at 42 °C for 60 min followed by 5 min at 85 °C. The post-RT emulsion was burst with 10% emulsion breaker (Atrandi Biosciences, cat.no. MON-EB1) and pooled material was used for subsequent library construction.

Library construction

The barcoded-cDNA was purified twice with 0.8X AMPure XP reagent (BeckMan Coulter, cat. co. A63881) as per manufacturer's instructions. Next, cDNA was PCR amplified with KAPA HiFi Hot Start Ready Mix (Roche, cat.no. KK2601) using cDNA FWD primer and cDNA REV primers (see Table 2). Amplified DNA was fragmented and ligated to adapter using instruction and reagents provided by NEBNext® Ultra™ II FS DNA Library Prep (NEB, cat.no. E7805S). Finally, the libraries were amplified by 12 rounds of indexing PCR (2X KAPA HiFi Hot Start Ready Mix, Roche, cat.no. KK2601). Library quality was assessed using Bioanalyzer DNA High Sensitivity chip (Agilent, cat. no. 50674626). The libraries were sequenced on Illumina NextSeq 550 platform in multiple batches using either NextSeq 500/550 High Output Kit v2.5 (75 Cycles) (Illumina, cat. no. 20024906) or NextSeq 500/550 High Output Kit v2.5 (150 Cycles) (Illumina, cat. no. 20024907).

Raw sequencing data processing

The STARsolo pipeline (https://github.com/jsimonas/solo-in-drops) was used to process the data and to obtain expression matrices. STAR (version 2.7.6a) was run with the following parameters: --soloMultiMappers Uniform, -- soloType CB_UMI_Simple, -- soloUMIfiltering MultiGeneUMI, and --soloCBmatchWLtype 1MM. Homo sapiens (human) genome assembly GRCh38 (hg38) and Ensembl v93 annotations were used as the reference.

Data analysis: quality control, doublet, and RBC removal

Starting with cell x gene matrices, analysis was performed in Python using scanpy toolkit (Table 3). All notebooks are provided at https://github.com/zvirblyte/2023_ccRCC. Briefly, the raw count matrices were uploaded into an AnnData object and filtered by total transcript count and mitochondrial count fraction. The threshold for mitochondrial counts for all libraries was 20%. The total transcript count threshold was determined by evaluating the total count distribution and was selected permissive at minimum 400 UMIs per cell (300 UMIs for libraries T3.1, T9.1, N3.3, N4.3, N2.3). Doublets were removed using Scrublet[70] (v0.2.3) in the same PCA space used for initial UMAP construction. Scrublet was applied on each emulsion separately. Briefly, the procedure for doublet removal consisted of 1) Calculating doublet scores for each cell in each emulsion using Scrublet; 2) high-resolution graph-based clustering using Scanpy's Louvain algorithm implementation (resolution = 60); 3) evaluation of mean doublet score and fraction of predicted doublets per cluster; 4) manual inspection of doublet-rich clusters in the interactive SPRING application[71], 5) removal of clusters with high mean doublet score and doublet fraction and no cluster-specific gene expression. This procedure, starting from UMAP construction at step 2) was repeated a total of two times and 913 cells (<2% of the total cell population) were removed. Transcriptomes with >1% of total raw counts originating from hemoglobin genes (HBB, HBA1, HBA2, HBD) were considered as red blood cells (RBCs) and 47 such transcriptomes were removed from further analysis.

UMAP construction, clustering, and annotation

After filtering and QC steps we retained 50,236 single cells that were used to construct a graph and UMAP representation (Fig. 1b). The procedure consisted of 1) normalization to 10,000 total counts, log-transformation and scaling; 2) selection of highly variable genes; 3) PCA; 4) batch correction using Harmony[72]; 5) graph construction and 6) UMAP representation. After normalization, genes with 15 CPTT (counts per ten thousand) in not less than 25 cells were considered abundant and retained, furthermore, mitochondrial and ribosomal genes were excluded and top 2000 abundant and highly variable genes, based on Fano factor (as in ref. 68), were used for PCA. To remove batch effects due to different batches of barcoding beads the dataset integration was performed using function scanpy.external.pp.harmony_integrate() with the batch variable 'beads'. Then, adjacency graph was constructed using sc.pp.neighbors() with n_neighbors = 30 and UMAP representation was built using sc.tl.umap() with min_dist = 0.4. The resulting representation was used for exploration in interactive SPRING application. Graph-based spectral clustering with varying number of clusters (k) was performed using sklearn.cluster.SpectralClustering() function, the clustering results were explored in the interactive SPRING environment, and $k = 43$ was selected for annotation. Differential gene expression analysis (Mann Whitney U test with Bonferoni-Hochberg correction) was performed and top 25 marker genes for each cluster (adjusted p value < 0.05) were used for in-depth literature analysis and manual cell type annotation (Supplementary Information Table 1, Supplementary Data S2).

Sample heterogeneity quantification

To quantify sample heterogeneity, Shannon entropy of samples was calculated for each broad cell category as described in ref. 11. Briefly, entropy

Table 2 | List of DNA oligonucleotides

Name	Sequence
Template-switching primer	
TSO	5′-AAGCAGTGGTATCAACGCAGAGTACATrGrGrG
cDNA amplification primers	
cDNA REV primer	5′-AAGCAGTGGTATCAACGCAGAGT
cDNA FWD primer	5′-CTACACGACGCTCTTCCGATCT
Ligation adapter	
Ligation FWD primer	5′-GATCGGAAGAGCACACGTCTGAACTCCAGTCAC
Ligation REV primer	5′-GCTCTTCCGATCT
Indexing PCR primers	
Forward PCR index primer	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
PE2-ind1	CAAGCAGAAGACGGCATACGAGAT**CGTGAT**GTGACTGGAGTTCAGACGTGT
PE2-ind2	CAAGCAGAAGACGGCATACGAGAT**ACATCG**GTGACTGGAGTTCAGACGTGT
PE2-ind3	CAAGCAGAAGACGGCATACGAGAT**GCCTAA**GTGACTGGAGTTCAGACGTGT
PE2-ind4	CAAGCAGAAGACGGCATACGAGAT**TGGTCA**GTGACTGGAGTTCAGACGTGT
PE2-ind5	CAAGCAGAAGACGGCATACGAGAT**CACTGT**GTGACTGGAGTTCAGACGTGT
PE2-ind6	CAAGCAGAAGACGGCATACGAGAT**ATTGGC**GTGACTGGAGTTCAGACGTGT

Table 3 | Software and algorithms

Software	Version	Reference
solo-in-drops	v1.0	https://github.com/jsimonas/solo-in-drops
STAR	2.7.6a	https://github.com/alexdobin/STAR, https://doi.org/10.1101/2021.05.05.442755
scanpy	v1.8.0	[78], https://scanpy.readthedocs.io/en/stable
harmonypy	v0.0.5	[72], https://github.com/slowkow/harmonypy
scrublet	v0.2.3	[70], https://github.com/swolock/scrublet
SPRING viewer	N/A	[71], https://github.com/AllonKleinLab/SPRING_dev
scikit-learn	v1.0.2	https://scikit-learn.org/stable
statsmodels	v0.12.2	https://www.statsmodels.org/v0.12.2
scipy	v1.6.2	[79], https://scipy.org
anndata	v0.7.6	https://doi.org/10.1101/2021.12.16.473007, https://anndata.readthedocs.io/en/latest
numpy	v1.20.1	https://numpy.org/doc/1.20/index.html
pandas	v1.2.4	https://pandas.pydata.org
louvain	v0.7.1	https://github.com/vtraag/louvain-igraph
umap	v0.5.1	https://umap-learn.readthedocs.io/en/latest
matplotlib	v3.2.2	https://matplotlib.org/stable/index.html
seaborn	v0.11.0	https://seaborn.pydata.org
jupyterlab	v2.2.6	https://jupyter.org
CellPhoneDB	v2.0	[53], https://cellphonedb.readthedocs.io/en/latest/index.html
R	v4.2.1	https://www.r-project.org/
tidyverse	v1.3.2	https://www.tidyverse.org/
biomaRt	v2.52.0	https://bioconductor.org/packages/biomaRt/
clusterProfiler	v4.4.4	https://bioconductor.org/packages/clusterProfiler/
TCGAbiolinks	v2.24.3	https://bioconductor.org/packages/TCGAbiolinks/
survival	v3.3-1	https://CRAN.R-project.org/package=survival
survminer	v0.4.9	https://cran.r-project.org/package=survminer
CellTypist	V1.6.2	https://www.celltypist.org/

values were calculated for sample frequency in each cell group (stromal, endothelial, tumor, lymphoid, myeloid, epithelial and cycling). To account for differences in the number of cells per group, we subsampled 100 cells from each group 100 times with replacement and calculated the Shannon entropy using function scipy.stats.entropy(). Cells from cluster "Tumor cells 1" were excluded, as they were sample specific.

Receptor–ligand interaction analysis

Log-normalized expression values for all cell types, excluding healthy epithelial cell populations and cycling cells were used to infer cell-cell interactions using CellphoneDB v.2.0.0[53] with method "statistical_analysis" and default parameters. Significant (p value < 0.05) cell-cell interactions were explored and selected interactions are shown in Figs. 2c, 3d, 5c and Supplementary Fig. S3b. Cell-cell interaction signatures for subsequent survival analysis (as in Fig. 2d) were constructed by taking both the receptor and ligand genes in the set (provided in Supplementary Data S7). Cell-cell interaction analysis results are provided in Supplementary Data S5 and S6.

CellTypist label transfer

To examine the similarity of ccRCC tumor endothelial cell types to the ones described by ref. 31, a CellTypist[73] model was trained for label transfer according to a tutorial available at https://www.celltypist.org/. Briefly, Goveia et al., endothelial cell scRNA-seq matrix and metadata was obtained from https://endotheliomics.shinyapps.io/lung_ectax/, the matrix was log-normalized and filtered to exclude nontumor endothelial cells and patient #5 specific phenotype. Then, the model was trained on the dataset without gene filtering and applied for label transfer to our endothelial cell log-normalized matrix with parameter majority_voting=True. Similarly, a model was trained on ref. 15 dataset obtained from GEO (at GSE159115). The dataset was filtered to epithelial cells only without gene filtering and applied for label transfer to our epithelial cell log-normalized matrix with parameter majority_voting=True. The results are presented in Supplementary Figs. S2a and S4b.

Gene set over-representation analysis

Gene set over-representation analysis was employed to evaluate the potential functional significance of a given gene signature. The analysis utilized gene sets obtained from the Hallmark Pathways of the MSigDB database v7.5.1[74]. Gene signatures were then submitted to a hypergeometric test implemented in the enrichGO() function of the clusterProfiler R package[75] using genes that were detected (nonzero UMI counts) in kidney tissue samples as a universe (background reference). The pathways having FDR (Benjamini-Hochberg) values below 0.05 were considered as significantly over-represented.

Survival analysis

TCGA KIRC cohort bulk RNA-seq (upper quartile FPKM normalized) and clinical data were downloaded from the NCI GDC Data Portal[76] using the TCGAbiolinks R package[77]. Cell type signature scoring of the TCGA bulk RNA-seq samples was performed by calculating an arithmetic mean of the z-score transformed expression values for all genes in a given signature. The used gene-wise z-score transformation equalized differences in the gene expression abundances, so that lowly and highly expressed genes would have the same scale and, thus equal weight in the score. The association between signature score and overall survival time was assessed by Kaplan-Meier and multivariate Cox regression analyses. Log-rank tests and Wald tests, respectively, were used to evaluate statistical significance (at level of 0.05) of the performed survival analyses. For the Kaplan-Meier analysis, stratified signature (high—greater or equal than the median signature score; low—lower than the median signature score) was used, while for the multivariate Cox regression analysis, the continuous signature score values were used with patient age and sex as covariates. The survival analyses were conducted using the survival and the survminer R packages.

Statistics and reproducibility

Single-cell RNA-seq datasets for paired healthy-adjacent kidney ($n = 9$) and ccRCC tumor ($n = 8$) samples were generated in this study. One sample (patient P1 tumor) was excluded from analysis due to high necrosis level in tissue. Detailed descriptions of the statistical analyses in this study are provided in the respective methods section. Significance threshold for p values and adjusted p values was <0.05.

Reporting summary

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

Data availability

The data generated in this study are available in Gene Expression Omnibus (GEO) at GSE242299. Publicly available datasets used were downloaded from GEO (at GSE159115) and https://endotheliomics.shinyapps.io/lung_ectax/. TCGA KIRC cohort bulk RNA-seq (upper quartile FPKM normalized) and clinical data were downloaded from the NCI GDC Data Portal using the TCGAbiolinks R package.

Code availability

All Jupyter notebooks for scRNA-seq and R scripts for other analyses presented in this manuscript are publicly available at https://github.com/zvirblyte/2023_ccRCC. Software versions used are provided in Table 3.

Received: 25 September 2023; Accepted: 21 June 2024;
Published online: 28 June 2024

References

1. Hsieh, J. J. et al. Renal cell carcinoma. *Nat. Rev. Dis. Prim.* **3**, 17009 (2017).
2. Dizman, N., Philip, E. J. & Pal, S. K. Genomic profiling in renal cell carcinoma. *Nat. Rev. Nephrol.* **16**, 435–451 (2020).
3. Cancer Genome Atlas Research, N. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499**, 43–49 (2013).
4. Sato, Y. et al. Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat. Genet.* **45**, 860–867 (2013).
5. Senbabaoglu, Y. et al. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* **17**, 231 (2016).
6. Chevrier, S. et al. An immune atlas of clear cell renal cell carcinoma. *Cell* **169**, 736–749 e718 (2017).
7. Motzer, R. J. et al. Avelumab plus Axitinib versus sunitinib for advanced renal-cell carcinoma. *N. Engl. J. Med.* **380**, 1103–1115 (2019).
8. Rini, B. I. et al. Pembrolizumab plus Axitinib versus sunitinib for advanced renal-cell carcinoma. *N. Engl. J. Med.* **380**, 1116–1127 (2019).
9. Braun, D. A. et al. Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nat. Med.* **26**, 909–918 (2020).
10. Azizi, E. et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* **174**, 1293–1308 e1236 (2018).
11. Chan, J. M. et al. Signatures of plasticity, metastasis, and immunosuppression in an atlas of human small cell lung cancer. *Cancer Cell* **39**, 1479–1496 e1418 (2021).
12. Zilionis, R. et al. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* **50**, 1317–1334 e1310 (2019).
13. Massalha, H. et al. A single cell atlas of the human liver tumor microenvironment. *Mol. Syst. Biol.* **16**, e9682 (2020).
14. Young, M. D. et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* **361**, 594–599 (2018).
15. Zhang, Y. et al. Single-cell analyses of renal cell cancers reveal insights into tumor microenvironment, cell of origin, and therapy response. *Proc. Natl Acad. Sci. USA* **118**, https://doi.org/10.1073/pnas.2103240118 (2021).
16. Long, Z. et al. Single-cell multiomics analysis reveals regulatory programs in clear cell renal cell carcinoma. *Cell Discov.* **8**, 68 (2022).
17. Li, R. et al. Mapping single-cell transcriptomes in the intra-tumoral and associated territories of kidney cancer. *Cancer Cell* **40**, 1583–1599 e1510 (2022).
18. Krishna, C. et al. Single-cell sequencing links multiregional immune landscapes and tissue-resident T cells in ccRCC to tumor topology and therapy efficacy. *Cancer Cell* **39**, 662–677 e666 (2021).
19. Borcherding, N. et al. Mapping the immune environment in clear cell renal carcinoma by single-cell genomics. *Commun. Biol.* **4**, 122 (2021).
20. Alchahin, A. M. et al. A transcriptional metastatic signature predicts survival in clear cell renal cell carcinoma. *Nat. Commun.* **13**, 5747 (2022).
21. Braun, D. A. et al. Progressive immune dysfunction with advancing disease stage in renal cell carcinoma. *Cancer Cell* **39**, 632–648 e638 (2021).
22. Bi, K. et al. Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. *Cancer Cell* **39**, 649–661 e645 (2021).
23. Chen, Y., McAndrews, K. M. & Kalluri, R. Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat. Rev. Clin. Oncol.* **18**, 792–804 (2021).
24. Nagl, L., Horvath, L., Pircher, A. & Wolf, D. Tumor endothelial cells (TECs) as potential immune directors of the tumor microenvironment —new findings and future perspectives. *Front. Cell Dev. Biol.* **8**, 766 (2020).
25. Stewart, B. J. et al. Spatiotemporal immune zonation of the human kidney. *Science* **365**, 1461–1466 (2019).
26. Lake, B. B. et al. A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. *Nat. Commun.* **10**, 2832 (2019).
27. Balzer, M. S., Rohacs, T. & Susztak, K. How many cell types are in the kidney and what do they do? *Annu. Rev. Physiol.* **84**, 507–531 (2022).
28. Denisenko, E. et al. Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biol.* **21**, 130 (2020).
29. Schreibing, F. & Kramann, R. Mapping the human kidney using single-cell genomics. *Nat. Rev. Nephrol.* **18**, 347–360 (2022).
30. Rudman-Melnick, V. et al. Single-cell profiling of AKI in a murine model reveals novel transcriptional signatures, profibrotic phenotype, and epithelial-to-stromal crosstalk. *J. Am. Soc. Nephrol.* **31**, 2793–2814 (2020).
31. Goveia, J. et al. An integrated gene expression landscape profiling approach to identify lung tumor endothelial cell heterogeneity and angiogenic candidates. *Cancer Cell* **37**, 21–36 e13 (2020).
32. Bensaad, K. et al. Fatty acid uptake and lipid storage induced by HIF-1alpha contribute to cell growth and survival after hypoxia-reoxygenation. *Cell Rep.* **9**, 349–365 (2014).
33. Filippou, P. S., Karagiannis, G. S. & Constantinidou, A. Midkine (MDK) growth factor: a key player in cancer progression and a promising therapeutic target. *Oncogene* **39**, 2040–2054 (2020).
34. Shi, Y. et al. Decoding the multicellular ecosystem of vena caval tumor thrombus in clear cell renal cell carcinoma by single-cell RNA sequencing. *Genome Biol.* **23**, 87 (2022).
35. Au, L. et al. Determinants of anti-PD-1 response and resistance in clear cell renal cell carcinoma. *Cancer Cell* **39**, 1497–1518 e1411 (2021).
36. Vogt, L. et al. VSIG4, a B7 family-related protein, is a negative regulator of T cell activation. *J. Clin. Investig.* **116**, 2817–2826 (2006).
37. Roumenina, L. T. et al. Tumor cells hijack macrophage-produced complement C1q to promote tumor growth. *Cancer Immunol. Res.* **7**, 1091–1105 (2019).
38. Pritykin, Y. et al. A unified atlas of CD8 T cell dysfunctional states in cancer and infection. *Mol. Cell* **81**, 2477–2493 e2410 (2021).
39. Carosella, E. D., Gregori, S. & Tronik-Le Roux, D. HLA-G/LILRBs: a cancer immunotherapy challenge. *Trends Cancer* **7**, 389–392 (2021).
40. Liu, L. et al. Construction of TME and Identification of crosstalk between malignant cells and macrophages by SPP1 in hepatocellular carcinoma. *Cancer Immunol. Immunother.* **71**, 121–136 (2022).
41. Flieswasser, T. et al. The CD70-CD27 axis in oncology: the new kids on the block. *J. Exp. Clin. Cancer Res.* **41**, 12 (2022).
42. Wang, Y. H. et al. Plasmalemmal vesicle associated protein (PLVAP) as a therapeutic target for treatment of hepatocellular carcinoma. *BMC Cancer* **14**, 815 (2014).
43. De Palma, M., Biziato, D. & Petrova, T. V. Microenvironmental regulation of tumour angiogenesis. *Nat. Rev. Cancer* **17**, 457–474 (2017).
44. Guzzi, L. M. et al. Clinical use of [TIMP-2]*[IGFBP7] biomarker testing to assess risk of acute kidney injury in critical care: guidance from an expert panel. *Crit. Care* **23**, 225 (2019).
45. Roudnicky, F. et al. High expression of insulin receptor on tumour-associated blood vessels in invasive bladder cancer predicts poor overall and progression-free survival. *J. Pathol.* **242**, 193–205 (2017).
46. Abe, Y. et al. A single-cell atlas of non-haematopoietic cells in human lymph nodes and lymphoma reveals a landscape of stromal remodelling. *Nat. Cell Biol.* **24**, 565–578 (2022).
47. Elgundi, Z. et al. Cancer metastasis: the role of the extracellular matrix and the heparan sulfate proteoglycan perlecan. *Front Oncol.* **9**, 1482 (2019).
48. Lindgren, M. et al. Type IV collagen as a potential biomarker of metastatic breast cancer. *Clin. Exp. Metastasis* **38**, 175–185 (2021).

49. Dumas, S. J. et al. Phenotypic diversity and metabolic specialization of renal endothelial cells. *Nat. Rev. Nephrol.* **17**, 441–464 (2021).
50. Samuelson Bannow, B. et al. Factor VIII: long-established role in haemophilia A and emerging evidence beyond haemostasis. *Blood Rev.* **35**, 43–50 (2019).
51. Su, S. C. et al. Autotaxin-lysophosphatidic acid signaling axis mediates tumorigenesis and development of acquired resistance to sunitinib in renal cell carcinoma. *Clin. Cancer Res.* **19**, 6461–6472 (2013).
52. Hongu, T. et al. Perivascular tenascin C triggers sequential activation of macrophages and endothelial cells to generate a pro-metastatic vascular niche in the lungs. *Nat. Cancer* **3**, 486–504 (2022).
53. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
54. Sainson, R. C. et al. TNF primes endothelial cells for angiogenic sprouting by inducing a tip cell phenotype. *Blood* **111**, 4997–5007 (2008).
55. Lee, W. S., Yang, H., Chon, H. J. & Kim, C. Combination of anti-angiogenic therapy and immune checkpoint blockade normalizes vascular-immune crosstalk to potentiate cancer immunity. *Exp. Mol. Med.* **52**, 1475–1485 (2020).
56. Rolny, C. et al. HRG inhibits tumor growth and metastasis by inducing macrophage polarization and vessel normalization through downregulation of PlGF. *Cancer Cell* **19**, 31–44 (2011).
57. Lee, S. W. et al. Peroxidasin is essential for endothelial cell survival and growth signaling by sulfilimine crosslink-dependent matrix assembly. *FASEB J.* **34**, 10228–10241 (2020).
58. Yang, X. et al. CD36 in chronic kidney disease: novel insights and therapeutic opportunities. *Nat. Rev. Nephrol.* **13**, 769–781 (2017).
59. Xu, W. H. et al. Elevated CD36 expression correlates with increased visceral adipose tissue and predicts poor prognosis in ccRCC patients. *J. Cancer* **10**, 4522–4531 (2019).
60. Guessoum, O., de Goes Martini, A., Sequeira-Lopez, M. L. S. & Gomez, R. A. Deciphering the identity of renin cells in health and disease. *Trends Mol. Med.* **27**, 280–292 (2021).
61. Araujo, T. G. et al. Annexin A1 as a regulator of immune response in cancer. *Cells* **10**, https://doi.org/10.3390/cells10092245 (2021).
62. Moraes, L. A. et al. Annexin-A1 enhances breast cancer growth and migration by promoting alternative macrophage polarization in the tumour microenvironment. *Sci. Rep.* **7**, 17925 (2017).
63. Salomé, B. et al. NKG2A and HLA-E define an alternative immune checkpoint axis in bladder cancer. *Cancer Cell* **40**, 1027–1043 e1029 (2022).
64. Andre, P. et al. Anti-NKG2A mAb is a checkpoint inhibitor that promotes anti-tumor immunity by unleashing both t and NK cells. *Cell* **175**, 1731–1743 e1713 (2018).
65. Chauvin, J. M. & Zarour, H. M. TIGIT in cancer immunotherapy. *J. Immunother. Cancer* **8**, https://doi.org/10.1136/jitc-2020-000957 (2020).
66. Peng, Y. L. et al. Single-cell transcriptomics reveals a low CD8(+) T cell infiltrating state mediated by fibroblasts in recurrent renal cell carcinoma. *J. Immunother. Cancer* **10**, https://doi.org/10.1136/jitc-2021-004206 (2022).
67. Quintanal-Villalonga, Á. et al. Protocol to dissociate, process, and analyze the human lung tissue using single-cell RNA-seq. *STAR Protoc.* **3**, https://doi.org/10.1016/j.xpro.2022.101776 (2022).
68. Klein, A. M. et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
69. Zilionis, R. et al. Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* **12**, https://doi.org/10.1038/nprot.2016.154 (2017).
70. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291 e289 (2019).
71. Weinreb, C., Wolock, S. & Klein, A. M. SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. *Bioinformatics* **34**, 1246–1248 (2018).
72. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
73. Dominguez Conde, C. et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science* **376**, eabl5197 (2022).
74. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
75. Wu, T. et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovations* **2**, 100141 (2021).
76. Grossman, R. L. et al. Toward a shared vision for cancer genomic data. *N. Engl. J. Med.* **375**, 1109–1112 (2016).
77. Colaprico, A. et al. TCGAbiolinks: an R/bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **44**, e71 (2016).
78. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
79. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

Acknowledgements

We are especially grateful to the patients at the National Cancer Institute, Vilnius, Lithuania for participating in this study. This work received funding from European Regional Development Fund [01.2.2-LMT-K-718-04-0002] under grant agreement with the Research Council of Lithuania. The work in S. Ja. group was funded by grant no. S-MIP-17-54. S. Ju. was supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 101030265. We are grateful to Rapolas Zilionis for valuable discussions and input on data analysis, and to the members of the Oncourology Department at the National Cancer Institute (Lithuania) for their valuable support and kind assistance.

Author contributions

J.Z., J.N., K.G. single-cell RNA-seq experiments, library preparation and sequencing; J.N., R.K. biospecimen logistics and processing; M.K., A.U., patient consent, biospecimen curation, acquisition and logistics; D.D. histology; J.Z. data analysis and interpretation, initial manuscript draft; S.Ju. data management and analysis; J.N., S.Ju., S.Ja., L.M. proofreading; J.Z., L.M. manuscript revision and preparation; S.Ja., and L.M. study design and funding acquisition; L.M. supervision. All authors have read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-024-06478-x.

Correspondence and requests for materials should be addressed to Sonata Jarmalaite or Linas Mazutis.

Peer review information *Communications Biology* thanks Brandon J. Manley and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Simona Chera and Johannes Stortz.