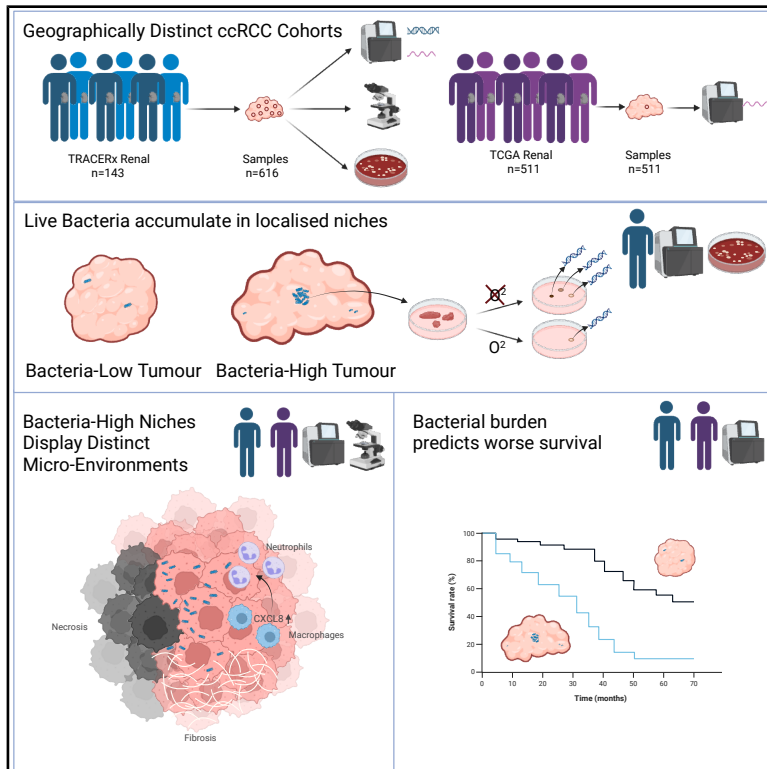


# Bacteria inside kidney tumors associate with myeloid inflammation and disease progression: A TRACERx Renal study

## Graphical abstract



## Authors

Alice C. Martin, Anne-Laure Cattin, Parise K. Lockwood, ..., Kevin Litchfield, Maximiliano G. Gutierrez, Samra Turajlić

## Correspondence

samra.turajlic@crick.ac.uk

## In brief

Live bacteria accumulate within localized niches in primary kidney cancer tumors and associate with worse survival. High bacterial abundance niches are characterized by neutrophil-rich, pro-tumor immune cell profiles, corroborating patterns in other solid tumors and warranting future work into their potential as novel therapeutic targets.

## Highlights

- Live bacteria accumulate in discrete niches within primary ccRCC tumors
- The dominant intratumoral genus, *Cutibacterium*, predicts worse survival
- *Cutibacterium* hotspots are characterized by neutrophils and myeloid inflammation
- *In vitro*, *Cutibacterium* induces neutrophil-recruiting chemokines in macrophages



## Translation to Patients

Martin et al., 2027, Med 8, 101335  
January 8, 2027 © 2026 The Author(s). Published by Elsevier Inc.  
<https://doi.org/10.1016/j.medj.2026.101335>

Article

# Bacteria inside kidney tumors associate with myeloid inflammation and disease progression: A TRACERx Renal study

Alice C. Martin,<sup>1,2,3</sup> Anne-Laure Cattin,<sup>1,2</sup> Parise K. Lockwood,<sup>4</sup> Martha A. Zepeda-Rivera,<sup>5,6,7</sup> Clara Becker,<sup>8</sup> Taja Barber,<sup>1,2</sup> Irene Lobon,<sup>1</sup> Krupa Thakkar,<sup>3</sup> Leo J. Bickley,<sup>1</sup> Zayd Tippu,<sup>1,2,9</sup> Laura Marandino,<sup>1,2,9</sup> Marija Miletić,<sup>1,2,9</sup> Lucy Meader,<sup>10</sup> Mary Green,<sup>10</sup> Donald Bell,<sup>11</sup> Brian Hanley,<sup>1,2,9,12</sup> Barbara Ibarzo Yus,<sup>1,2</sup> Fiona Byrne,<sup>1,2</sup> Charlotte Spencer,<sup>1,9</sup> Scott T.C. Shepherd,<sup>1,2,9,12</sup> Annika Fendler,<sup>1</sup> Lewis Au,<sup>1,9</sup> Daqi Deng,<sup>1,2</sup> Ángel Fernández Sanromán,<sup>1,13</sup> Sarah Rudman,<sup>14</sup> James Larkin,<sup>9</sup> Kevin Litchfield,<sup>3</sup> Maximiliano G. Gutierrez,<sup>4</sup> and Samra Turajlic<sup>1,2,9,15,16,17,\*</sup>

<sup>1</sup>Cancer Dynamics Laboratory, Francis Crick Institute, London, UK

<sup>2</sup>Cancer Dynamics Laboratory, Cancer Research UK Manchester Institute, The University of Manchester, Manchester, UK

<sup>3</sup>Tumour Immunogenomics and Immunosurveillance Laboratory, UCL Cancer Institute, London, UK

<sup>4</sup>Host-Pathogen Interactions in Tuberculosis Laboratory, Francis Crick Institute, London, UK

<sup>5</sup>Department of Genomic Medicine, MD Anderson Cancer Center, Houston, TX, USA

<sup>6</sup>Department of Immunology, James P. Allison Institute, Houston, TX, USA

<sup>7</sup>Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Seattle, WA, USA

<sup>8</sup>Human Biology Division, Fred Hutchinson Cancer Center, Seattle, WA, USA

<sup>9</sup>Renal and Skin Units, Royal Marsden NHS Foundation Trust, London, UK

<sup>10</sup>Experimental Histopathology Facility, Francis Crick Institute, London, UK

<sup>11</sup>Advanced Light Microscopy Facility, Francis Crick Institute, London, UK

<sup>12</sup>Melanoma and Kidney Cancer Group, Institute of Cancer Research, London, UK

<sup>13</sup>Spanish National Cancer Research Center (CNIO), Madrid, Spain

<sup>14</sup>Guy's and St Thomas' NHS Foundation Trust, London, UK

<sup>15</sup>Cancer Research UK Manchester Institute, Manchester, UK

<sup>16</sup>The Christie NHS Foundation Trust, Manchester, UK

<sup>17</sup>Lead contact

\*Correspondence: [samra.turajlic@crick.ac.uk](mailto:samra.turajlic@crick.ac.uk)

<https://doi.org/10.1016/j.medj.2026.101335>

**CONTEXT AND SIGNIFICANCE** Bacteria have increasingly been found to contribute to immune remodeling and disease progression across different solid tumor types. This study investigated whether intratumoral bacteria play a similar role in kidney cancer, where these same immune changes characterize aggressive disease and treatment resistance. Analyzing two large and geographically distinct patient cohorts, the researchers found that live bacteria accumulated in discrete niches within tumors and that higher bacterial abundance predicted worse patient survival. Bacterial hotspots were consistently characterized by the same neutrophil-driven immune profiles that mark aggressive disease, a pattern further confirmed in cell culture experiments showing that live bacteria directly sustained the production of immune-signaling molecules that attract neutrophils. These findings corroborate patterns observed in other solid tumors and establish intratumoral bacteria as a previously overlooked component of the kidney cancer immune landscape, laying the groundwork for exploring them as potential novel therapeutic targets.

## SUMMARY

**Background:** Intratumoral bacteria are increasingly recognized as active contributors to neutrophil-driven inflammation and disease progression in solid tumors. Yet, in clear cell renal cell carcinoma (ccRCC), where myeloid inflammation is a key feature of aggressive disease, live bacterial presence, their intratumoral distribution, and interaction with the tumor microenvironment remain unexplored.

**Methods:** We genetically characterized the microbiome of primary ccRCC tumors across two independent cohorts (TRACERx: 616 samples, 143 patients; The Cancer Genome Atlas [TCGA]: 511 samples and patients),

confirming live bacterial presence through culture. Multi-regional sampling enabled within-tumor analysis of bacterial distribution. Associations with the tumor microenvironment were explored at scale using bulk RNA sequencing (RNA-seq) across both cohorts and corroborated in depth through spatial transcriptomics and *in vitro* infection assays.

**Findings:** Across both cohorts, the commensal *Cutibacterium* emerged as the dominant genus, accumulating in discrete niches enriched toward the tumor center and associating with more advanced (TRACERx  $p = 0.024$ ; TCGA  $p = 0.0011$ ) tumors and shorter progression-free survival (TRACERx  $p = 0.0022$ ; TCGA  $p = 4.4 \times 10^{-4}$ ). Both across and within tumors, *Cutibacterium* accumulation associated with neutrophil recruitment and myeloid inflammation. *In vitro*, *Cutibacterium* induced sustained expression of neutrophil-attracting chemokines in macrophages, recapitulating patient tumor findings.

**Conclusions:** These findings establish intratumoral bacteria as novel prognostic components of the ccRCC tumor microenvironment and suggest an active role in shaping local myeloid inflammation, laying the groundwork for investigating their potential as therapeutic targets.

**Funding:** This work was supported by the Francis Crick Institute (Cancer Research UK, MRC, Wellcome Trust). TRACERx Renal is funded by Cancer Research UK (C50947/A18176).

## INTRODUCTION

Commensal bacteria are fundamental in shaping host immunity in health and disease.<sup>1</sup> In cancer, the role of distinct pro- and anti-tumor gut microbial communities in influencing immune activity, disease progression, and response to immunotherapy<sup>2–5</sup> has already inspired microbiome-targeting clinical trials,<sup>6–9</sup> including in clear cell renal cell carcinoma (ccRCC) (NCT07383441). More recently, work exploring bacterial communities within tumors<sup>10–13</sup> has highlighted opportunities for novel prognostic and therapeutic developments centered around these bacteria, which directly interact with tumor cells and the tumor microenvironment (TME).

Similar to cancer-related gut microbes, distinct tumor-resident bacterial populations have been described with either pro- or anti-tumor effects. In select cancers, peptides from intratumoral bacteria can be presented on tumor cells, inducing anti-tumor T cell activation.<sup>14,15</sup> However, the most commonly reported activity of intratumoral microbes is the induction of pro-tumorigenic chronic inflammation, characterized by neutrophil recruitment and myeloid inflammation.<sup>10,13,16</sup> In ccRCC, where myeloid inflammation is a key feature of advanced disease,<sup>17,18</sup> bacterial DNA has previously been detected in patient tumors.<sup>19–21</sup> However, the presence of intact, live bacteria and their interactions with tumor cells and the TME remain to be elucidated.

We therefore genetically profile the intratumoral microbiome of 143 multi-regionally sampled primary ccRCC tumors ( $n = 616$ ) from the TRACERx Renal (TxRenal) cohort,<sup>22</sup> confirming bacterial presence through *in situ* imaging and culture. We explore associations with clinical features including survival within TxRenal and validate findings in The Cancer Genome Atlas (TCGA) ( $n = 511$ ). Leveraging bulk RNA sequencing (RNA-seq), fluorescent *in situ* hybridization (FISH), and spatial transcriptomics, we explore phenotypic changes associated with bacterial accumulation in depth and at breadth. Finally, we recapitulate these phenotypic changes *in vitro* using infection assays, thereby providing evidence for an active involvement of bacteria in disease progression.

## RESULTS

### *Cutibacterium* dominates primary ccRCC tumors

To establish the microbial landscape of primary ccRCC tumors, we extracted bacterial reads from whole-genome sequencing (WGS) of 143 primary tumors, profiled across multiple regions (median = 4,  $n = 616$ ) to capture spatial diversity. The samples were taken from a central tumor slice to minimize surgical contamination (Figure 1A). Bacterial reads were identified using a custom pipeline and normalized by sequencing depth to enable cross-sample comparison (Figure S1A; STAR Methods; Table S1).

After stringent decontamination (Figure 1B), we identified nine bacterial genera enriched in tumors relative to matched peripheral blood. Interestingly, most genera were strict or facultative anaerobes, suggesting potential selective pressure for tumor colonization. *Cutibacterium*, a skin commensal implicated in chronic inflammatory conditions such as acne vulgaris and prosthetic joint infections,<sup>23,24</sup> was the most abundant bacterial genus, contributing a median of 86% (interquartile range [IQR], 77%–93%) of all bacterial reads per sample (Figures 1C and 1D) and driving 97.9% of variation in bacterial abundance across samples (Figure S1B).

Because *Cutibacterium* is a common skin commensal, we used multiple approaches to confirm that its detection in tumors was not driven by contamination (Tables S2 and S3). In our decontamination pipeline, we excluded genera found in water and DNA controls that underwent the same extraction protocol and shipping preparation as our tumor samples. We also removed genera for which 5% or more of their variance across the cohort was explained by hospital site or sequencing batch. To further confirm the lack of contamination during sample preparation and sequencing, 6 samples were re-sequenced after 3 months, yielding consistent *Cutibacterium* abundance but unreliable detection of less-abundant genera present at fewer than 25 counts (Figure 1E). *Cutibacterium* was therefore absent from our lab controls and not associated with hospital site or sequencing. Nonetheless, given the ubiquity of *Cutibacterium* on human skin, it is possible that a portion of *Cutibacterium*

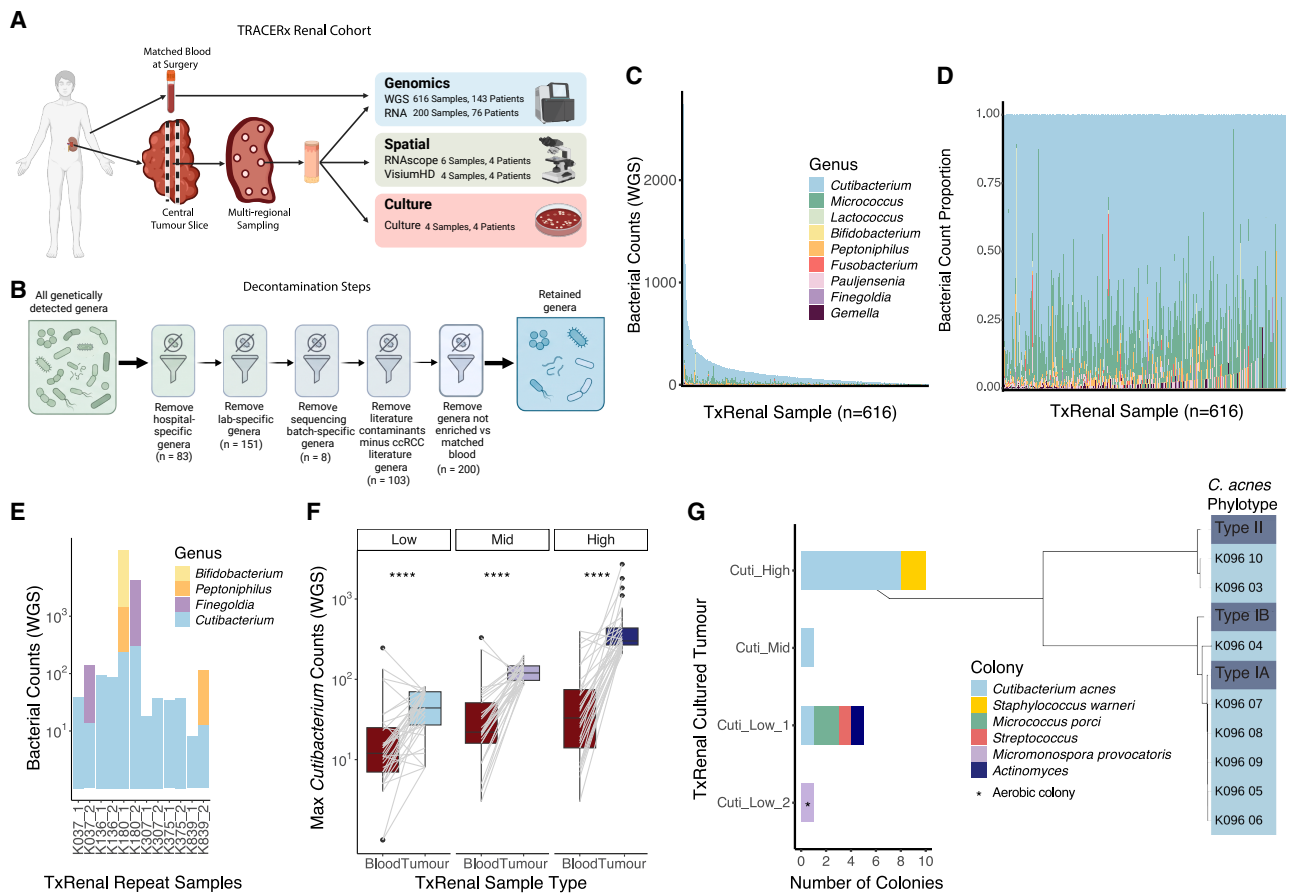


**Figure 1. Characterizing the intratumoral microbiome of ccRCC tumors**

(A) Sampling workflow and sample numbers of different modalities used for intratumoral bacteria characterization in the TxRenal cohort. Steps to minimize and recognize contaminants during sample acquisition highlighted.

(B) Contaminant filters implemented and number of genera screened out at each step.

(C and D) Per sample, total and relative abundance, respectively, of the bacterial genera in the TxRenal cohort. Bacterial reads are normalized by total sample read depth.

(E) Bacterial counts detected in 6 samples processed and sequenced in duplicate at a 3-month interval.

(F) Paired maximum *Cutibacterium* count difference in tumor and peripheral blood for high, mid, and low abundance tumors. Gray lines link samples from the same patient. Significance was assessed by paired Wilcoxon test.

(G) Bacterial colonies cultured from tumors classified as high, mid, or low based on WGS read count. Inset shows a cladogram of *Cutibacterium acnes* strains cultured from the high tumor, with reference strains from each of the three most common *Cutibacterium acnes* phylotypes.

(A and B) were created in BioRender. Martin, A. (2026) <https://BiorRender.com/be72tff>.

ns:  $p > 0.05$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$

reads stem from contamination. However, any background signal is negligible relative to tumor enrichment, as *Cutibacterium* abundance was significantly lower in matched blood samples collected during surgery and processed alongside the tumors (Figure 1F).

Next, to assess whether the bacterial reads detected through sequencing originated from live bacteria, we cultured snap-frozen tumor samples belonging to “high,” “mid,” or “low” *Cutibacterium* tertiles, as defined by WGS, under aerobic and anaerobic conditions. Colonies were shotgun sequenced for species identification (Table S4; Figure 1G), confirming *Cutibacterium acnes* as the most abundant species. Both the “Cuti-high” and “Cuti-mid” samples yielded *C. acnes*

colonies, with most colonies stemming from the Cuti-high sample: eight individual *C. acnes* colonies from three genetically distinct strains (Figure 1G). In contrast, only one of the two “Cuti-low” samples grew a *C. acnes* colony, while the other showed no *Cutibacterium* growth, matching the relative abundance determined through WGS (Figure 1G) and confirming that *Cutibacterium* is not a ubiquitous contaminant. Across patients, *C. acnes* colonies predominantly belonged to phylotype IA, with no phylotype IB colonies cultured. *Micrococcus*, the second most abundant genus in the WGS cohort, yielded two colonies, further supporting accurate detection of viable bacteria with our sequencing approach. The remaining cultured species were not retained in our final WGS data, reflecting the stringency of our

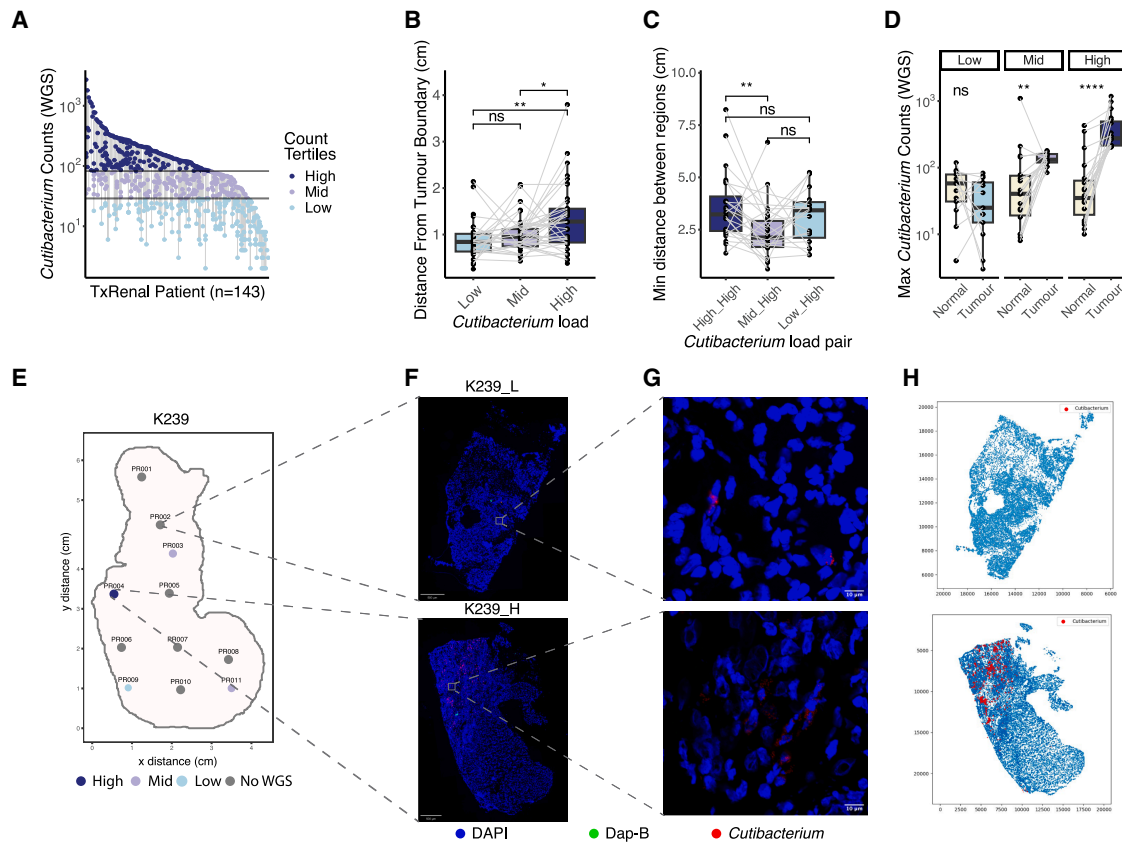


**Figure 2. Spatial distribution of *Cutibacterium* in ccRCC tumors**

(A) *Cutibacterium* abundance of each region within individual tumors, colored by tertiles of abundance.

(B) Median distance from tumor boundary of high, mid, and low regions for tumors that span at least two load groups for tumors >7 cm in size. Lines connect regions from the same tumor. Paired Wilcoxon test, Cuti-high vs. Cuti-mid  $n = 50$ ,  $p = 0.014$ ; Cuti-high vs. Cuti-low  $n = 33$ ,  $p = 0.0063$ ; Cuti-mid vs. Cuti-low  $n = 42$ ,  $p = 0.36$ .

(C) Minimum distance to a high *Cutibacterium* load region for regions in the same tumor in each of the abundance tertiles. Lines connect regions from the same tumor. Paired Wilcoxon test, high-high vs. high-mid  $n = 20$ ,  $p = 0.0056$ ; high-high vs. high-low  $n = 13$ ,  $p = 0.34$ ; high-mid vs. high-low  $n = 15$ ,  $p = 0.095$ .

(D) Maximum *Cutibacterium* abundance in primary tumor relative to its matched tumor-adjacent normal tissue.

(E) WGS-based *Cutibacterium* abundance tumor map. Dot color represents absolute *Cutibacterium* abundance by tertile. Gray regions were not profiled by WGS.

(F) RNAseq images of high and low regions from the same tumor, *Cutibacterium* 16S probe in red, Dap-B negative control in green, and DAPI in blue.

(G) Confocal image of bacteria-positive cells from each region.

(H) Plot representations of the images for clarity, with inset signal quantification.

ns:  $p > 0.05$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ .

decontamination pipeline, which excludes genera where true and contaminant strains cannot currently be distinguished. Together, these results demonstrate the presence of viable bacteria in ccRCC primary tumors and validate relative *Cutibacterium* abundance inferred from WGS.

### Intratumoral heterogeneity in *Cutibacterium* distribution

Having established the presence of *Cutibacterium* in primary tumors, we next investigated its distribution within tumors. We found considerable heterogeneity in bacterial abundance across tumor regions, with 80% of tumors containing a Cuti-high region, also harboring regions with mid or low abundance (Figure 2A). This suggests that *Cutibacterium* does not proliferate uniformly but instead accumulates in discrete niches.

Exploring the localization of these Cuti-high regions, we found that, in large tumors (>7 cm), Cuti-high regions were significantly further from the tumor edge than Cuti-mid or Cuti-low regions within the same tumor ( $n = 71$  tumors with  $\geq 2$  abundance groups; Figure 2B). This supports a true biological signal origin, rather than surgical contamination, which would be expected to accumulate at the tumor boundary. Measuring the distance from each region to the closest Cuti-high region revealed that Cuti-mid regions were significantly closer to Cuti-high regions than other Cuti-high or Cuti-low regions were ( $n = 30$  tumors with  $\geq 2$  Cuti-high-Cuti-high/mid/low pairs; Figure 2C), suggesting that Cuti-mid regions represent the edges of Cuti-high proliferation niches. Conversely, the longer distance between Cuti-high regions indicates that spatially distinct colonization or proliferation events can occur in the same tumor.

To explore whether *Cutibacterium* presence extends beyond the tumor in the same way as has been reported for tumor-related immune cells,<sup>25,26</sup> we assessed its presence in the tumor-adjacent renal tissue from 50 patients. We classified tumors into Cuti-high/mid/low tertiles based on each tumor's maximum-abundance region and found detectable, but significantly lower, abundance of *Cutibacterium* in adjacent renal tissue for all but Cuti-low tumors (Figure 2D). Adjacent to Cuti-high tumors, normal tissue harbored increasing *Cutibacterium* abundance the closer the sample was taken to the tumor (Figure S1C). Thus, *Cutibacterium* preferentially accumulates within tumors and, at the tumor/normal boundary, toward tumors.

Given the variation in bacterial abundance across tumor regions and beyond, we next investigated how *Cutibacterium* is organized within regions using a *Cutibacterium*-specific RNA-scope probe and a *Bacillus subtilis* (Dap-B) probe as a control (Figure S2A). Four late-stage (III/IV) tumors were analyzed: one Cuti-low (K988) and three Cuti-high (K239, K096, and K153). For two of the Cuti-high tumors (K239 and K153), we examined one WGS-defined Cuti-high region and a second, spatially distant region for which, because WGS-profiled tissue was unavailable, *Cutibacterium* status was defined by RNAscope signal abundance.

Imaging confirmed heterogenous *Cutibacterium* abundance across and within tumors. In both Cuti-high tumors for which a second region was imaged, the distal region showed similar bacterial signal abundance to the Cuti-low tumor ( $13\text{--}29$  spots;  $1.6 \times 10^{-7}\text{--}1.8 \times 10^{-7}$  spots/area, Figures 2E–2H; Figures S2B–S2E). In contrast, all Cuti-high regions contained markedly more bacteria ( $94\text{--}639$  spots;  $5.8 \times 10^{-7}\text{--}4.9 \times 10^{-6}$  spots/area), either distributed evenly throughout the tumor section (K153\_H) or clustered in aggregates (K239\_H and K096\_H).

To better visualize the distribution of *Cutibacterium* signal within these hotspots, we imaged them using high-resolution confocal microscopy. Bacterial signal was closely associated with cell nuclei along the z axis (Figure 2G; Videos S1 and S2), consistent with *Cutibacterium* adhering to or residing within cells. Given that *Cutibacterium* is non-motile, this association with cells in the tumor is likely critical for colonization and spread.

### ***Cutibacterium* burden predicts survival**

We next investigated whether intratumoral accumulation of *Cutibacterium* correlates with clinical outcomes (Table S5). As outcomes are measured per patient, we represented multi-regionally sampled tumors by each tumor's maximum *Cutibacterium* abundance region. This revealed a significant enrichment of *Cutibacterium* in late-stage tumors (III/IV) (Figure 3A). Patients with Cuti-high and Cuti-mid tumors also had significantly worse progression-free survival (PFS) than those with Cuti-low tumors, with little difference between the Cuti-high and Cuti-mid groups (Figure 3B; Figures S3A and S3B). This is consistent with Cuti-mid regions representing the periphery of Cuti-high niches, such that limited sampling may miss the highest abundance areas. Given the enrichment of *Cutibacterium* in late-stage tumors, we repeated the survival analysis stratified by stage. *Cutibacterium* abundance remained significantly prognostic of PFS in late- but not early-stage disease (Figure 3C). Similar

trends were seen for overall survival (OS), although they did not reach significance (Figures S3C and S3D).

We validated these findings in 511 treatment-naïve, primary ccRCC tumors from TCGA, processed with our custom pipeline (Table S6). Consistent with TxRenal, *Cutibacterium* was the dominant genus detected across tumors (median 82%, IQR 70%–90%). It was also enriched in late-stage tumors (Figure 3D) and was associated with poorer PFS and, in this larger cohort additionally OS, in patients with Cuti-high and Cuti-mid tumors compared to those with Cuti-low tumors (Figure 3E; Figures S3E–S3G). Strikingly, in TCGA, *Cutibacterium* abundance additionally was associated with worse PFS and OS even in stage I and II tumors, which are typically considered low risk (Figure 3F; Figure 1E).

Beyond stage, several clinical factors are routinely used for risk stratification in localized ccRCC.<sup>27</sup> To investigate the association of *Cutibacterium* with these clinical features, we ran a linear mixed-effects model in TxRenal, including stage as a random effect. *Cutibacterium* abundance was significantly associated with several features indicative of tumor growth (tumor size, tumor thrombus presence, and perinephric fat invasion) as well as with necrosis, across stages (Figure 3G). *Cutibacterium* accumulation may therefore promote or exploit these aggressive tumor features.

As we also identified other, less abundant, bacteria in these tumors, we explored whether any of them similarly associated with survival (PFS) by performing an elastic net Cox regression on all genera found across both cohorts (Figures 3H and 3I). One other genus (*Finegoldia*) was associated with PFS in TxRenal but did not reach significance independently ( $p = 0.21$ ), while one additional genus significantly associated with PFS in TCGA (*Fusobacterium*  $p = 0.0007$ ). Only *Cutibacterium* consistently predicted worse PFS across both cohorts (Figures 3H and 3I). Additionally, our duplicate runs showed that low-abundance genera were inconsistently detected by non-targeted sequencing (Figure 1E). Therefore, although other bacteria may contribute to the survival signal when present, *Cutibacterium* serves as a more robust biomarker due to its ubiquity across geographies and sufficient abundance for reliable detection.

### ***Cutibacterium*-high regions are enriched in neutrophil recruitment and response genes**

To better understand why *Cutibacterium* accumulation associates with disease progression, we analyzed transcriptional changes linked to bacterial abundance in a subset of TxRenal (200 samples, 76 patients) and all of TCGA (511 samples).

In both cohorts, tumors with higher *Cutibacterium* abundance were enriched for bacterial sensing pathways, suggesting active recognition of and response to *Cutibacterium* (Figures 4A and 4B). Additional significantly enriched pathways involved immune cell migration and activation, particularly of granulocytes and neutrophils. Indeed, in TxRenal, the four most upregulated genes with higher *Cutibacterium* abundance included potent neutrophil chemoattractants (*CXCL8/IL-8* and *PPBP/CXCL7*) and neutrophil-specific receptors (*GPR97* and *FCGR3B/CD16*) (Figure 4C). Similarly, in TCGA, significantly enriched immune response pathways were driven by neutrophil- and myeloid-related genes (e.g., antimicrobial humoral response: *CXCL1*, *CXCL8*, *SLPI*, *PLA2G1B*, and *ROMO1*). Consistently, *ARG1*,

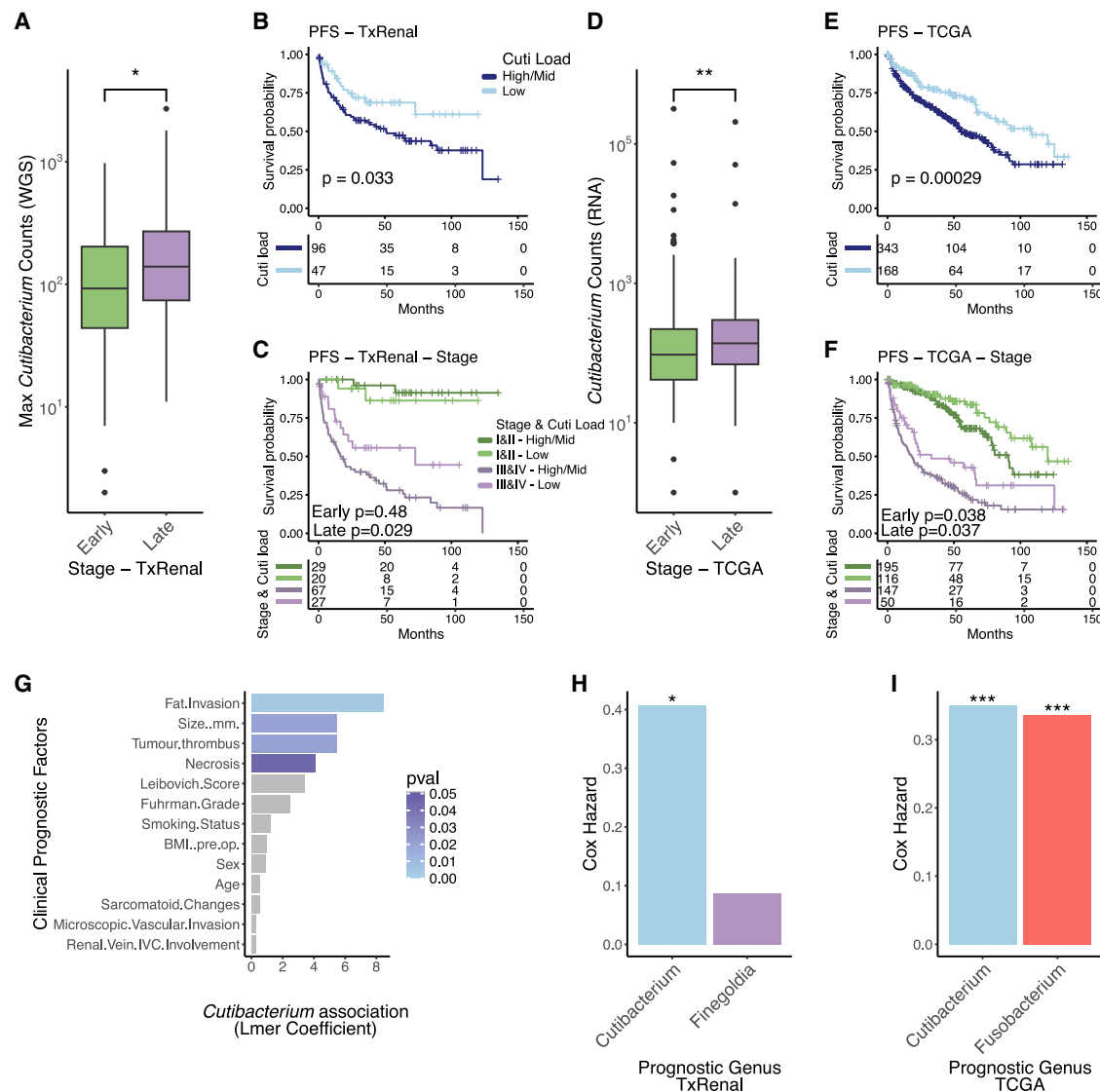


**Figure 3. Associations of intratumoral *Cutibacterium* abundance with clinical prognosis**

(A) Maximum *Cutibacterium* abundance in tumors split by early (I and II) vs. late (III and IV) stage in boxplots, for the TxRenal cohort.

(B and C) Kaplan-Meier survival curves for progression-free survival for the TxRenal cohort, grouped by *Cutibacterium* load or both stage and load, respectively. Patients are split into high/mid and low *Cutibacterium* load groups based on tertiles. The region with maximum *Cutibacterium* abundance was selected to represent the tumor.

(D) *Cutibacterium* abundance in tumors split by early (I and II) vs. late (III and IV) stage in boxplots, for the TCGA cohort.

(E and F) Kaplan-Meier survival curves for progression-free survival for the TCGA cohort, grouped by *Cutibacterium* load or both stage and load, respectively. Patients are split into high/mid and low *Cutibacterium* load groups based on tertiles.

(G) Linear mixed effect regression coefficient of clinical factor predicting *Cutibacterium* abundance (log10) with stage as random effect, colored by significance.

(H and I) Bacterial genera identified through elastic net Cox regression as being prognostic of progression-free survival in TxRenal and TCGA cohorts, respectively. Univariate significance indicated by stars.

\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ .

primarily expressed by neutrophils in humans,<sup>28,29</sup> was among the most upregulated genes as *Cutibacterium* abundance increased (Figure 4D). Together, both datasets show a strong link between *Cutibacterium* burden and the recruitment and response of myeloid cells, especially neutrophils.

We also explored the evolution of these signatures as *Cutibacterium* burden goes from low to mid/high within the same tumor

(Figure 4E). Aside from ATP metabolic pathways, all pathways upregulated with *Cutibacterium* abundance across patients were also upregulated with abundance within tumors, with considerably more pathways reaching significance than expected by chance ( $pval < 0.05$ : 17/33, BH  $p_{adj} < 0.05$ : 9/33), supporting a *Cutibacterium*-associated effect rather than inter-patient variance.

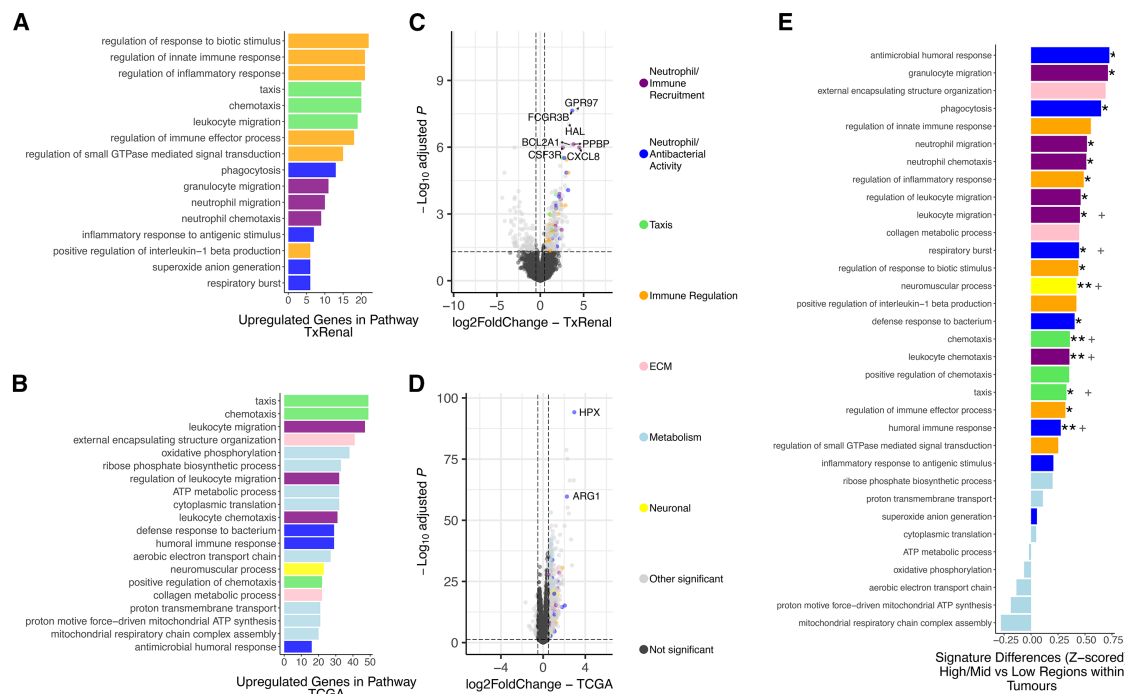


**Figure 4. Transcriptomic characterization of *Cutibacterium* accumulation within and across tumors**

(A and B) Pathway enrichment of human genes significantly upregulated with *Cutibacterium* abundance (log10) across tumors in the TxRenal and TCGA cohorts, respectively. Colored by pathway category.

(C and D) Volcano plots of genes up- and downregulated with logged *Cutibacterium* abundance across tumors in the TxRenal and TCGA cohorts, respectively. Genes are colored by the pathway category they contribute to.

(E) Median difference in pathway enrichment between high/mid and low regions within tumors (Z scored across samples and signatures). Significance was calculated using a paired Wilcoxon test (stars) and false discovery rate (FDR) corrected (crosses).

\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ .

### ***Cutibacterium* accumulates in distinct tumor topologies**

To refine the observations from bulk sequencing to spatially resolved cellular neighborhoods, we performed spatial transcriptomics (Visium HD) on serial sections of the three Cuti-high regions and the Cuti-low tumor previously profiled with RNAscope, using the *Cutibacterium* signal from the serial RNAscope to map bacterial localization (Figure 5A; Figure 2G; Figures S2A–S2D and S4–S7). As expected from bulk transcriptomics, genes upregulated in the Cuti-high vs. Cuti-low samples were significantly enriched for bacterial sensing and neutrophil or granulocyte recruitment pathways (Figure 5B), confirming that these regions are representative of more common bacterial dynamics.

In two of the three Cuti-high regions, *Cutibacterium* displayed distinct topological patterns consistent with the transcriptomic signals in TxRenal and TCGA (Figure 5C–F). Firstly, in K096\_H (Figure 5C), bacterial aggregates localized around areas of necrosis, defined by pathologist review of the H&E. This accumulation is consistent with the enrichment of *Cutibacterium* in necrosis-positive tumors in TxRenal (Figure 3G). In addition, a published gene signature characteristic of peri-necrotic tumor tissue<sup>30</sup> was enriched in Cuti-high/mid tumor regions, both across (TCGA) and within tumors (TxRenal) (Figure 5D), indicating a consistent link between *Cutibacterium* and necrosis.

Secondly, in K239\_H, *Cutibacterium* and immune cells were largely excluded from the tumor and instead localized to the

stroma, which contained collagen-rich, fibrotic areas as assessed on the H&E (Figure 5E). This spatial pattern recapitulates the enrichment of extracellular matrix (ECM) remodeling pathways observed in Cuti-high/mid tumors across TCGA (Figure 4B), suggesting that this bacterial localization within stromal regions undergoing matrix remodeling is likely a recurrent feature across patients.

In the third Cuti-high region (K153\_H), bacteria were evenly distributed throughout the tumor section (Figure 5F) and did not associate with notable topological features. Together, these data reveal three distinct modes of *Cutibacterium* distribution in the TME. Two modes, stromal and peri-necrotic, align with pervasive patterns across independent RNA-seq cohorts and display bacterial clustering, whereas the third, more dispersed bacterial distribution does not associate with distinctive niche structures.

### **Proximity to *Cutibacterium* is marked by neutrophil accumulation and macrophage reprogramming**

We next explored the phenotype of *Cutibacterium*-host cell interactions. Because the *Cutibacterium* signal was obtained from a serial section, we could not resolve direct bacteria-human cell-cell contacts; instead, we quantified cellular neighborhoods at increasing distances from the mapped *Cutibacterium* signal (Figure 5A; Figure S8A).

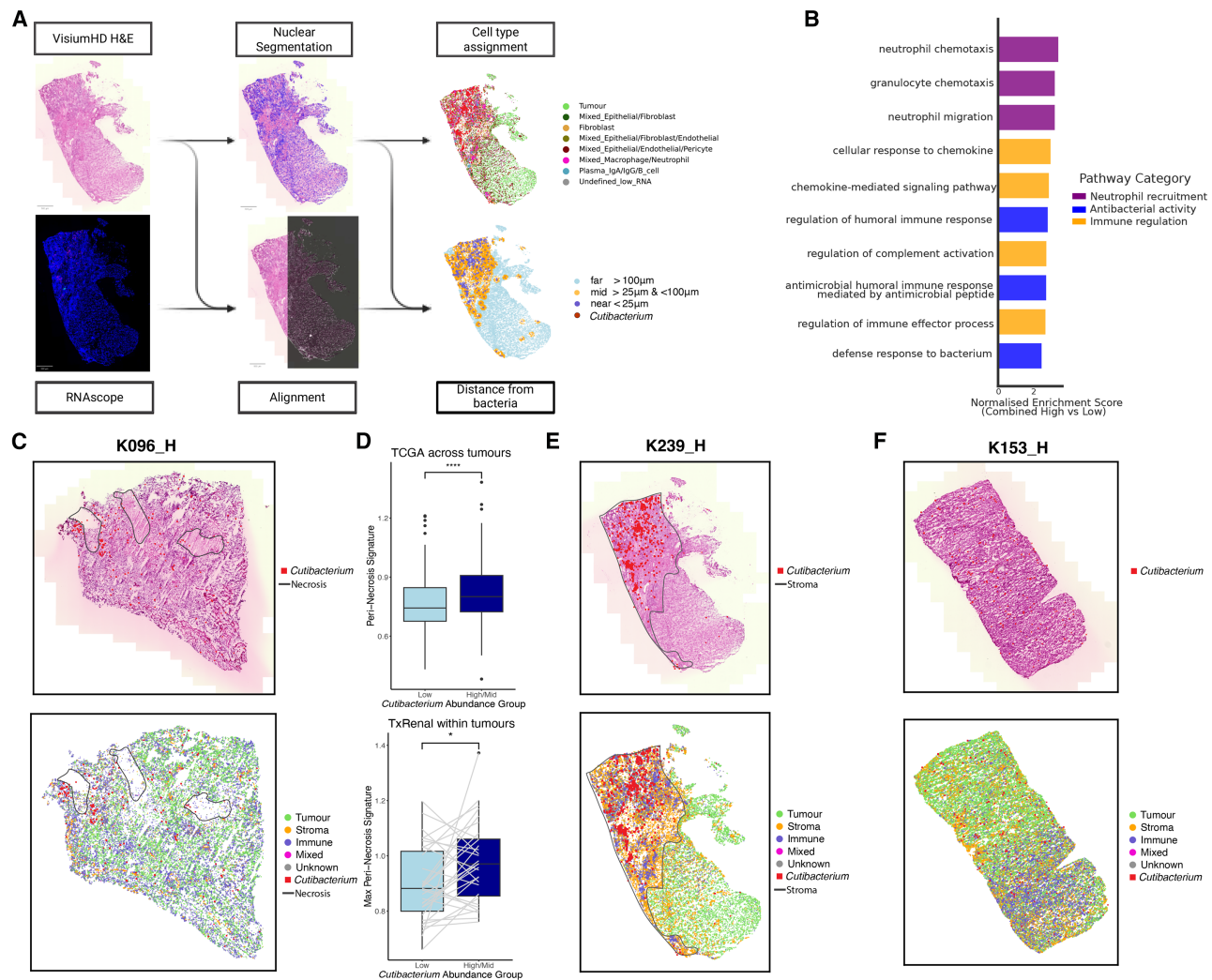


**Figure 5. Visualization of *Cutibacterium*-associated topological niches through spatial transcriptomics**

(A) Spatial transcriptomics workflow to identify *Cutibacterium*-proximal cells. Images used are from “Cuti-high” load sample K239\_H.

(B) Pathway enrichment of differentially expressed genes in all high vs. low *Cutibacterium* samples.

(C) *Cutibacterium* distribution from serial RNAscope mapped onto Visium HD H&E and broad cell type classifications for sample K096\_H. Areas of necrosis are indicated on both images.

(D) Peri-necrosis signature enrichment in high/mid vs. low regions across tumors in the TCGA cohort and within tumors in the TxRenal cohort. Significance assessed with unpaired and paired Wilcoxon tests, respectively.

(E and F) Same as (C) for samples K239\_H and K153\_H. Distinct stroma is outlined in K239\_H.

\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ .

Consistent with neutrophil recruitment and chemokine signaling enriched in the bulk cohorts, both neutrophil presence and macrophage expression of neutrophil-recruiting chemokines (*CXCL8*, *CXCL6*, *CXCL3*, *CXCL2*, and *CXCL1*) were enriched near *Cutibacterium* across Cuti-high samples (Figures 6A–6C). All neutrophil-containing clusters in Cuti-high samples increased in frequency closer to *Cutibacterium*, reaching significantly higher frequency “near” bacteria than “far” (Figure 6A). Neutrophil lineage markers (e.g., *CSF3R* and *SELL*) were found in clusters alongside either macrophage (e.g., *CD163* and *MS4A6A*) or monocyte (e.g., *CD14* and *VCAN*) markers (Figure S8B), due to VisiumHD probe bins capturing transcripts from closely adjacent cells, thereby re-

flecting the recurrent co-localization of these cell populations. Within these mixed clusters, the ratio of neutrophil (*CSF3R* and *FCGR3B*) to monocyte/macrophage (*CD68* and *CD14*) marker-expressing cells increased with bacterial proximity (Figure 6B), supporting consistent neutrophil accumulation near bacteria across Cuti-high samples. By contrast, the Cuti-low sample (K988\_L) contained fewer myeloid cells and showed no enrichment near *Cutibacterium* (Figures 6A and 6B), supporting that neutrophil accumulation associates with higher bacterial burden.

All macrophages expressed genes characteristic of late-stage ccRCC macrophages<sup>31</sup> across both Cuti-high and Cuti-low samples (e.g., immunosuppressive markers: *CD163*, *MRC1/CD206*,

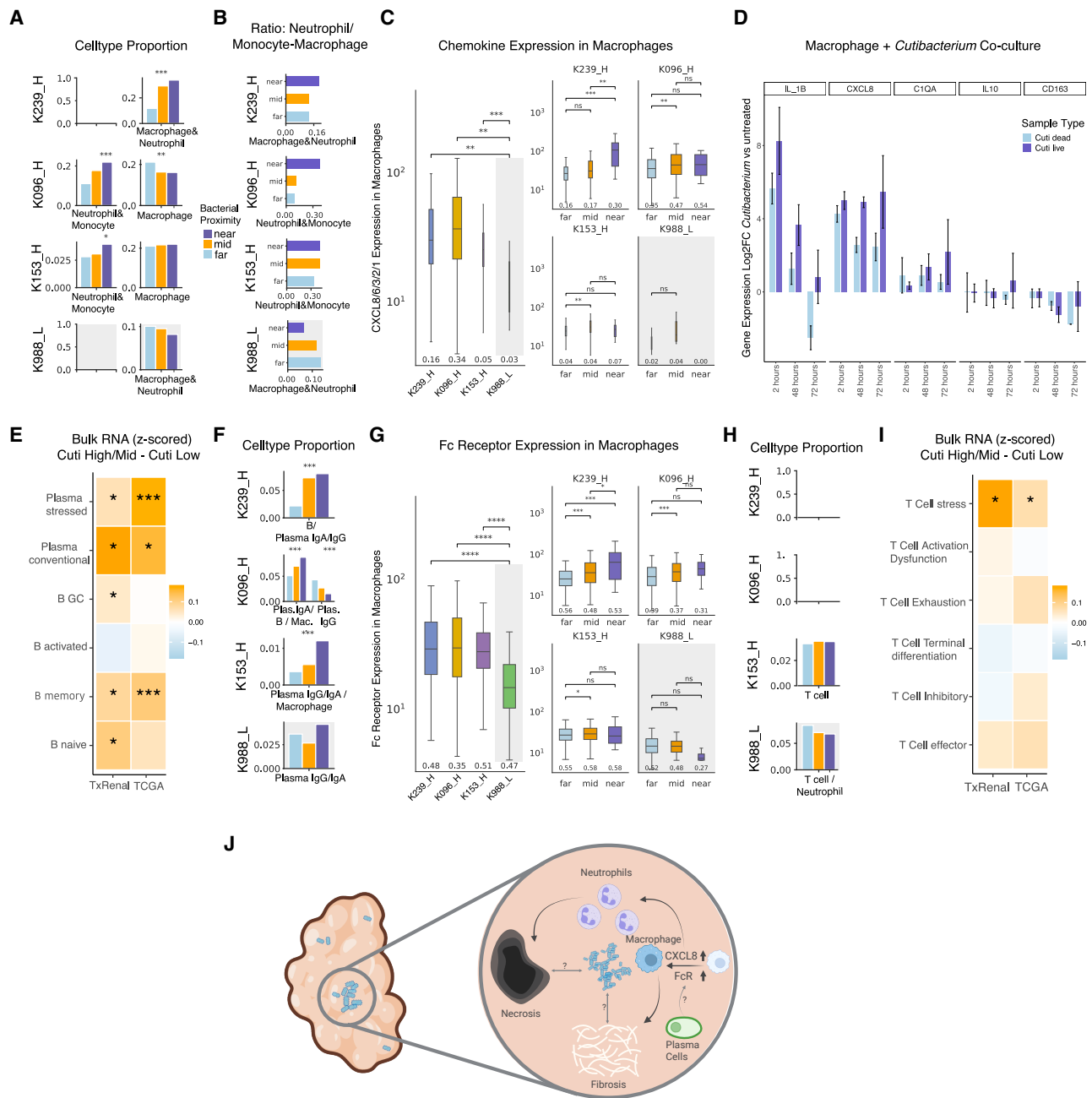


**Figure 6. Characterization of cell type dynamics near *Cutibacterium* in tumor tissue**

Across the figure, the Cuti-low sample K988\_L has been highlighted with a gray background to easily differentiate it from the three Cuti-high samples.

(A) Proportions of cells near, mid, and far from bacteria that each cell type makes up per sample. Near <25  $\mu$ m, mid >25 and < 100  $\mu$ m, far >100  $\mu$ m. Two-sided Fisher's test to assess significance, FDR correction across all cell types and samples. Myeloid clusters are shown.

(B) Ratio of cells expressing neutrophil lineage markers (*CSF3R* or *FCGR3B*) over cells expressing monocyte/macrophage lineage markers (*CD14* or *CD68*) in neutrophil-containing clusters, near, mid, and far from bacteria.

(C) Difference in expression of neutrophil-recruiting chemokines (*CXCL8*, *CXCL6*, *CXCL3*, *CXCL2*, or *CXCL1*) by macrophage clusters across samples and bacterial proximity compartments. Box widths show the ratio of chemokine-expressing cells over myeloid marker expression (*CD68* or *CD14*), capped at 1, and are indicated numerically below the box. Permutation test for significance, FDR corrected across samples and bacterial proximity compartments.

(D) Log2 fold change of gene expression by immunosuppressive macrophages with addition of dead or live *Cutibacterium* (MOI = 10), relative to untreated macrophages at the same time point, as assessed by qPCR.

(legend continued on next page)

*MS4A6A*, and *FOLR2*; complement cascade: *C1QA*, *C1QB*, and *C1QC*; cysteine cathepsins: *CTSS*; lipid transport: *APOE*) (Figure S8C). However, macrophages in Cuti-high samples expressed significantly higher levels of neutrophil-recruiting chemokines than in the Cuti-low sample, where expression was low or absent (Figure 6C). Macrophage chemokine expression additionally increased with bacterial proximity, including in K096\_H where macrophage frequency was higher “far” from bacteria (Figure 6A), consistent with a spatial gradient of macrophage reprogramming, rather than a compositional shift. Notably, the proportion of chemokine-expressing macrophages was lower in K153\_H than in the other two Cuti-high samples, yet chemokine expression was significantly higher closer to bacteria and relative to the Cuti-low sample. This suggests less efficient but still proximity-dependent macrophage reprogramming in this sample, potentially reflecting its diffuse bacterial distribution. Together, despite variation across the three Cuti-high samples, these spatial findings are consistent in directionality across distinct topological contexts and corroborate the bulk RNA results, supporting an association between *Cutibacterium* accumulation, neutrophil enrichment, and neutrophil-recruiting chemokine upregulation in immunosuppressive macrophages.

#### Live *Cutibacterium* induces *CXCL8* expression by immunosuppressive macrophages in culture

To test whether *Cutibacterium* actively drives the macrophage changes observed in patient tumors, we infected human monocyte-derived macrophages with live or heat-killed *C. acnes*. Monocytes were first differentiated with M-CSF to model the immunosuppressive macrophages found in ccRCC<sup>31</sup> and then infected for up to 3 days, with samples collected at 2, 48, and 72 h. Uninfected macrophages served as controls.

Both live and heat-killed *Cutibacterium* rapidly induced *CXCL8* at 2 h (Figure 6D). However, only live bacteria sustained similar *CXCL8* expression at 48 and 72 h, while the response to heat-killed bacteria waned over time. Inflammatory *IL-1B* expression similarly spiked at 2 h but declined progressively under both conditions, showing only a transient acute inflammatory response to *Cutibacterium*. Conversely, expression of immunosuppressive markers *IL10* and *CD163* did not show consistent changes (Figure 6D). Live *C. acnes* therefore induced a more lasting and

stronger *CXCL8* response in immunosuppressive macrophages than dead bacteria did, without altering the macrophages’ immunosuppressive expression profile, mirroring the phenotype observed in patient tumors.

#### Bacterial proximity is marked by B-lineage enrichment and limited T cell involvement

In addition to direct myeloid cell engagement by *Cutibacterium*, enrichment of antimicrobial humoral response pathways with increasing bacterial abundance in the bulk RNA cohorts (Figure 4A–C) suggested potential involvement of antibody-mediated myeloid inflammation. Consequently, we scored published tumor-associated B-lineage signatures<sup>32</sup> in both bulk RNA cohorts, revealing enrichment of memory B cell, conventional plasma cell, and stressed plasma cell (characterized by heat shock protein expression) signatures with higher *Cutibacterium* both within (TxRenal) and across (TCGA) tumors (Figure 6E).

Consistent with the bulk findings, all three Cuti-high spatial samples contained plasma-associated clusters that were significantly enriched in proximity to bacteria (Figure 6F), comprising predominantly plasma cell markers (*IgA/IgG*, *JCHAIN*, and *MZB1*) mixed with antigen-experienced B-lineage genes (*BLK*, *FCRL5*, and *POU2AF1*) (Figure S8D). In two samples (K096\_H and K153\_H), clusters enriched in proximity to bacteria displayed mixed macrophage-plasma transcriptional features, reflecting close spatial overlap between the two populations. Consistently, expression of antibody receptors (*FCGR1A*, *FCGR2A*, *FCGR2B*, *FCGR3A*, and *FCAR*) in macrophages increased with bacterial proximity and was significantly higher in Cuti-high compared to Cuti-low samples (Figure 6G), suggesting enhanced capacity for Fc-mediated sensing within bacteria-rich niches.

In contrast, T cell associations with bacterial abundance were minimal. No T cell-related pathways were enriched in the bulk differential analysis (Figures 4A and 4B), and in the single spatial Cuti-high sample with a T cell cluster (K153\_H), it showed no significant positional bias relative to bacteria (Figure 6H). Bulk scoring of ccRCC-specific T cell signatures<sup>31,33</sup> similarly revealed no enrichment with bacterial abundance aside from a stress-associated signature (Figure 6I). Collectively, across two independent bulk RNA cohorts and mirrored at spatial resolution, these findings indicate that bacteria-associated immune

(E) Enrichment of published cancer-related B-lineage signatures in high/mid vs. low regions across tumors in the TCGA cohort and within tumors in the TxRenal cohort, respectively. Significance assessed with unpaired and paired Wilcoxon tests, respectively, and FDR corrected. Signature scores were Z scored across all signatures and samples to enable comparison. The median difference between high/mid and low samples per signature is shown.

(F) Cell type proportions per bacterial proximity group as for (A), B lineages shown.

(G) Difference in expression of Fc receptors (*FCGR1A*, *FCGR2A*, *FCGR2B*, *FCGR3A*, or *FCAR*) by macrophage clusters across samples and bacterial proximity compartments. Box widths show the ratio of chemokine-expressing cells over myeloid marker expression (CD68 or CD14), capped at 1, and are indicated numerically below the box. Permutation test for significance, FDR corrected across samples and bacterial proximity compartments.

(H) Cell type proportions per bacterial proximity group as for (A), T cells shown.

(I) Enrichment of published ccRCC-related T-lineage signatures in high/mid vs. low regions across tumors in the TCGA cohort and within tumors in the TxRenal cohort, respectively. Significance assessed with unpaired and paired Wilcoxon test, respectively, and FDR corrected. Signature scores were Z scored across all signatures and samples to enable comparison. The median difference between high/mid and low samples per signature is shown.

(J) Proposed model of *Cutibacterium* interactions with the TME in ccRCC primary tumors: proximity to bacteria is marked by neutrophil enrichment and expression of neutrophil-attracting chemokines in macrophages. Plasma cell co-localization and upregulation of Fc receptors may indicate antibody-dependent macrophage stimulation, further supporting the chronic inflammatory state. This chronic myeloid inflammation may drive the observed co-localization with necrosis and fibrosis, which may be favorable to further bacterial proliferation.

This figure (J) was created in BioRender. Martin, A. (2026) <https://BioRender.com/yzjmc8k>.

\**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001; \*\*\*\**p* < 0.0001.

landscapes are characterized by myeloid and humoral enrichment, with limited evidence of T cell involvement.

## DISCUSSION

In this study, we demonstrate that prognostically relevant bacteria, predominantly *Cutibacterium*, locally colonize ccRCC primary tumors. This genus has repeatedly been detected within tumors across geographically distinct ccRCC cohorts<sup>19–21</sup> and is enriched in the urine of ccRCC patients relative to healthy individuals,<sup>34</sup> highlighting its potential as a clinically implementable prognostic biomarker. While some previous ccRCC studies have chosen to disregard *Cutibacterium* solely due to its status as a skin commensal,<sup>19</sup> our careful decontamination approach covering hospital sampling through to sequencing enabled confident detection of true *Cutibacterium* enrichment in ccRCC. This result was independently confirmed in a study by Ghaddar and colleagues,<sup>35</sup> who developed an approach to distinguish truly present bacterial reads from contaminant reads of common skin commensals, validating their model on culture-confirmed *Cutibacterium* infections and reporting high-confidence truly present *Cutibacterium* reads in over 88% of TCGA WGS ccRCC samples.

Across cohorts and modalities in our study, bacterial abundance consistently coincided with neutrophil recruitment and myeloid inflammation. This aligns with previous findings of pro-tumor intratumoral bacteria recruiting neutrophils in other cancer types,<sup>10,36,37</sup> as well as the established role of *Cutibacterium* in driving neutrophil recruitment via CXCL8 induction in acne vulgaris, prosthetic joint infections, and sarcoidosis.<sup>38–44</sup> Similar to acne vulgaris, we predominantly isolated *C. acnes* colonies of phylotype IA from ccRCC tumors and no phylotype IB colonies, which are predominantly found on healthy skin.<sup>45</sup> In culture, phylotype IA induces stronger macrophage expression of pro-inflammatory cytokines (CXCL8 and IL-6) than phylotype IB,<sup>46</sup> suggesting selective tumor colonization by more inflammatory strains. CXCL8, together with related chemokines, characterizes myeloid inflammation, an established feature of advanced disease and immunotherapy resistance in ccRCC.<sup>17,18,31,33</sup> Consistent with this, we show live but not heat-killed *Cutibacterium* inducing sustained CXCL8 expression in immunosuppressive macrophages *in vitro*, supporting an active role for bacteria in maintaining chronic myeloid inflammation. These findings align with experimental work demonstrating that intratumoral bacteria drive tumor growth through neutrophil recruitment in mice and organoids,<sup>10,36</sup> suggesting a similar mechanism may be responsible for the observed associations in ccRCC.

Alongside these myeloid changes, plasma and antigen-experienced B cell gene expression was enriched near bacteria across bulk and spatial transcriptomic analyses, with concordant increases in macrophage Fc receptor expression in bacteria-rich regions. This raises the possibility that locally produced antibodies engage Fc receptor signaling, potentially amplifying inflammatory responses beyond direct bacterial recognition. Stress-associated lymphoid signatures also associated with bacterial abundance in bulk transcriptomics. Characterized by heat shock protein expression and enriched in late-stage

ccRCC,<sup>31</sup> these signatures have been linked to poor prognosis and immunotherapy resistance in pan-cancer analyses.<sup>47,48</sup> As heat shock protein induction is a common host response to bacterial exposure,<sup>49</sup> whether lymphoid stress represents an additional axis of bacterially mediated immune dysfunction warrants further investigation.

These immune changes occur within distinct spatial tumor contexts, with bacteria recurrently associated with matrix remodeling and necrosis. ECM remodeling results from bacteria-mediated chronic inflammation across diseases<sup>50</sup> and can promote tumor growth by altering tissue architecture and immune infiltration.<sup>51</sup> Necrosis is similarly linked to aggressive tumor behavior<sup>27,52</sup> and increases in mouse tumors injected with bacteria relative to vehicle.<sup>36</sup> Neutrophils may contribute to this process through NETosis,<sup>30</sup> neutrophils' natural antibacterial response,<sup>53</sup> which *C. acnes* induces *in vitro*.<sup>54</sup> Both necrosis and matrix remodeling may therefore arise downstream of bacteria-driven myeloid inflammation, although local hypoxia within necrotic or stromal niches may also favor persistence of *Cutibacterium* as an aerotolerant anaerobe. In one sample, bacteria were diffusely distributed without clear necrotic or stromal niches, yet neutrophil enrichment and proximity-dependent macrophage reprogramming were still observed, albeit more weakly, favoring the hypothesis that bacterial aggregation and niche formation follow initial colonization and myeloid inflammation rather than preceding it.

Based on these findings, we propose a model (Figure 6J) in which *Cutibacterium* locally colonizes tumors and is recognized by myeloid cells, initiating a chronic inflammatory response characterized by CXCL8 production in immunosuppressive macrophages, which amplifies neutrophil recruitment and may be further reinforced by plasma cell co-localization and Fc-mediated macrophage activation. Over time, this bacteria-associated inflammatory niche may drive tissue remodeling processes such as necrosis and ECM remodeling, reshaping the TME in ways that promote tumor progression.

By establishing the prognostic relevance of intratumoral bacteria in ccRCC and linking them to TME changes associated with disease progression and immunotherapy resistance, this study identifies tumor-associated bacteria as candidate diagnostic targets and marks them as promising avenues for therapeutic investigation. Selective depletion of pro-tumorigenic bacteria in colorectal cancer models has restored anti-tumor T cell immunity,<sup>55</sup> raising the possibility that analogous strategies could be developed in ccRCC. Work in ccRCC is likely to be relevant in other solid tumors, namely prostate cancer, where similar *Cutibacterium* dynamics have been described.<sup>56–61</sup> Our study therefore establishes intratumoral bacteria as novel actors in the TME of ccRCC and lays the groundwork for future work to determine their potential as therapeutic targets, the elimination or modulation of which could extend survival for patients with advanced ccRCC and beyond.

## Limitations of the study

Our study has certain limitations. Untargeted sequencing, while enabling normalization relative to human DNA content, limited bacterial read depth, restricting confident classification to genus level and resulting in variable detection of low-abundance

genera. Our stringent decontamination pipeline ensured true tumor enrichment of reported genera but may have excluded biological strains where contaminant strains of the same genus were more abundant. Furthermore, while designed to provide cellular resolution of signal detected in the broader bulk cohorts rather than serving as primary evidence in isolation, spatial investigations were restricted to four samples as a result of limited availability of appropriate samples, a common limitation of human studies. Therefore, beyond our co-cultures with macrophages, future studies should refine cell-type-specific interactions, including through proteomic techniques, and address mechanisms through which bacteria may actively promote the TME changes observed *in situ*. Finally, with a single snapshot at the time of surgery, we could not assess the timing or route of tumor colonization. Further work is needed to determine how bacteria arise in ccRCC tumors and persist despite immune recognition.

## RESOURCE AVAILABILITY

### Lead contact

Requests for further information, resources, and reagents should be directed to and will be fulfilled by the lead contact, Samra Turajlic ([samra.turajlic@crick.ac.uk](mailto:samra.turajlic@crick.ac.uk)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

All code used in analyses and figure generation is available on FigShare (<https://doi.org/10.25418/crick.30865199>). All data needed to reproduce analyses, including bacterial read counts, associated clinical and metadata, and spatially resolved VisiumHD expression matrices, can also be found on FigShare (<https://doi.org/10.25418/crick.30865199>). Raw images and spatial transcriptomics are available upon request to the corresponding author. The raw sequencing data for the TxRenal WGS cohort can be accessed by becoming a member of the Genomics England Research Network (<https://www.genomicsengland.co.uk/research/academic>) or, for industry partners, the Discovery Forum (<https://www.genomicsengland.co.uk/research/research-environment>). Data from the TCGA and TxRenal RNA datasets used in this study can be obtained from their source publications and repositories as detailed in the [STAR Methods](#). Any additional information required to reanalyze the data reported in this work paper is available from the [lead contact](#) upon request.

## ACKNOWLEDGMENTS

We thank Prof. Susan Bullman and Dr. Evangelos Giampazolias for their guidance and expertise in planning bacterial culture and imaging work, the clinical trial teams at our recruiting hospital sites for enabling this research, and members of the Turajlic and Litchfield labs for their support throughout this project. This work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK, the UK Medical Research Council, and the Wellcome Trust. This work was funded by Cancer Research UK (CRUK) (C50947/A18176). This research was made possible through access to data in the National Genomic Research Library, which is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care). The National Genomic Research Library holds data provided by patients and collected by the NHS as part of their care and data collected as part of their participation in research. The National Genomic Research Library is funded by the National Institute for Health Research and NHS England. The Wellcome Trust, Cancer Research UK, and the Medical Research Council have also funded research infrastructure. L.M. was supported by the Royal Marsden Cancer

Charity. Graphical summaries and workflow diagrams were created with [BioRender.com](#).

## AUTHOR CONTRIBUTIONS

A.C.M.: conceptualization, validation, investigation, visualization, methodology, writing—original draft, and writing—review and editing. A.-L.C.: conceptualization, validation, investigation, and methodology. P.K.L.: validation, investigation, and methodology. M.A.Z.-R.: validation, methodology, and writing—review. C.B.: validation and methodology. T.B.: validation, methodology, and project administration. I.L.: conceptualization, methodology, and writing—review. K.T.: methodology. L.J.B.: methodology. Z.T.: data curation. L. Marandino: data curation. M.M.: data curation and writing—review. L. Meader: methodology. M.G.: methodology. D.B.: methodology. B.H.: methodology. B.I.Y.: methodology and writing—review. F.B.: project administration. C.S.: methodology. S.T.C.S.: methodology. A.F.: methodology. L.A.: methodology. D.D.: methodology. Á.F.S.: methodology, data curation, and writing—review. S.R.: project administration. J.L.: project administration. K.L.: conceptualization and supervision. M.G.G.: supervision, methodology, validation, and writing—review. S.T.: supervision, funding acquisition, conceptualization, writing—original draft, project administration, and writing—review and editing.

## DECLARATION OF INTERESTS

S.T. has received speaking fees from Roche, AstraZeneca, Novartis, and Ipsen. L.M. has received travel and accommodation grants from Immunocore and research funding from AstraZeneca.

## DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE MANUSCRIPT PREPARATION PROCESS

During the preparation of this work, the authors used ChatGPT in order to proofread and refine the text. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
  - TRACERx Renal (TxRenal) cohort
  - TCGA renal cohort
  - Sample collection – TxRenal cohort
- **METHOD DETAILS**
  - Whole genome sequencing
  - RNA-sequencing
  - 16S-sequencing
  - Bacterial culture from tumor samples
  - Bacterial and TME imaging by FISH (1) imaging protocol
  - Bacterial and TME imaging by FISH (2) probe testing
  - VisiumHD protocol
  - Cutibacterium-macrophage co-cultures
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Bacterial identification pipeline (0)
  - Bacterial identification pipeline (1) bacterial read identification from WGS and RNAseq
  - Bacterial identification pipeline (2) removing technical noise
  - Bacterial identification pipeline (3) bacterial read identification from 16S sequencing
  - Bacterial identification pipeline (4) removing contaminants from bacterial data
  - Bacterial identification pipeline (5) aligning to T2T reference genome
  - Bacterial identification pipeline (6) validating repeatability

- Tumour-level spatial analyses
- Survival analyses
- Clinical risk factor analyses
- Differential gene expression and enrichment analyses
- Taxonomy identification of bacterial cultures
- RNAscope image processing
- Macroscopic H&E sample annotations
- Cell type classification from VisiumHD
- Bacterial proximity analyses from VisiumHD data
- Differential gene expression analyses from VisiumHD data
- qPCR cytokine profiles from co-cultures
- Summary of evidence supporting true *Cutibacterium* enrichment in ccRCC primary tumors

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.medj.2026.101335>.

Received: April 20, 2026

Revised: August 28, 2026

Accepted: September 28, 2026

## REFERENCES

1. Zheng, D., Liwinski, T., and Elinav, E. (2020). Interaction between microbiota and immunity in health and disease. *Cell. Res.* 30, 492–506. <https://doi.org/10.1038/s41422-020-0332-7>.
2. Zitvogel, L., Ma, Y., Raoult, D., Kroemer, G., and Gajewski, T.F. (2018). The microbiome in cancer immunotherapy: Diagnostic tools and therapeutic strategies. *Science* 359, 1366–1370. <https://doi.org/10.1126/science.aar6918>.
3. Bullman, S., Eggermont, A., Johnston, C.D., and Zitvogel, L. (2021). Harnessing the microbiome to restore immunotherapy response. *Nat. Cancer* 2, 1301–1304. <https://doi.org/10.1038/s43018-021-00300-x>.
4. Xie, M., Li, X., Lau, H.C.-H., and Yu, J. (2025). The gut microbiota in cancer immunity and immunotherapy. *Cell. Mol. Immunol.* 22, 1012–1031. <https://doi.org/10.1038/s41423-025-01326-2>.
5. Helmink, B.A., Khan, M.A.W., Hermann, A., Gopalakrishnan, V., and Wargo, J.A. (2019). The microbiome, cancer, and cancer therapy. *Nat. Med.* 25, 377–388. <https://doi.org/10.1038/s41591-019-0377-7>.
6. Dizman, N., Meza, L., Bergerot, P., Alcantara, M., Dorff, T., Lyou, Y., Frankel, P., Cui, Y., Mira, V., Llamas, M., et al. (2022). Nivolumab plus ipilimumab with or without live bacterial supplementation in metastatic renal cell carcinoma: a randomized phase 1 trial. *Nat. Med.* 28, 704–712. <https://doi.org/10.1038/s41591-022-01694-6>.
7. Baruch, E.N., Youngster, I., Ben-Betzalel, G., Ortenberg, R., Lahat, A., Katz, L., Adler, K., Dick-Necula, D., Raskin, S., Bloch, N., et al. (2021). Fecal microbiota transplant promotes response in immunotherapy-refractory melanoma patients. *Science* 371, 602–609. <https://doi.org/10.1126/science.abb5920>.
8. Routy, B., Lenehan, J.G., Miller, W.H., Jamal, R., Messaoudene, M., Daisley, B.A., Hes, C., Al, K.F., Martinez-Gili, L., Punčochář, M., et al. (2023). Fecal microbiota transplantation plus anti-PD-1 immunotherapy in advanced melanoma: a phase I trial. *Nat. Med.* 29, 2121–2132. <https://doi.org/10.1038/s41591-023-02453-x>.
9. Ebrahimi, H., Dizman, N., Meza, L., Malhotra, J., Li, X., Dorff, T., Frankel, P., Llamas-Quitiquit, M., Hsu, J., Zengin, Z.B., et al. (2024). Cabozantinib and nivolumab with or without live bacterial supplementation in metastatic renal cell carcinoma: a randomized phase 1 trial. *Nat. Med.* 30, 2576–2585. <https://doi.org/10.1038/s41591-024-03086-4>.
10. Galeano Niño, J.L., Wu, H., LaCourse, K.D., Kempchinsky, A.G., Bar-yames, A., Barber, B., Futran, N., Houlton, N., Sather, C., Scinska, E., et al. (2022). Effect of the intratumoral microbiota on spatial and cellular heterogeneity in cancer. *Nature* 611, 810–817. <https://doi.org/10.1038/s41586-022-05435-0>.
11. Lu, Y., Xu, L., Chen, W., Liu, W., Zhang, Y., Zhou, Q., Wang, N., Zhang, Y., Bai, H., Xu, S., et al. (2025). Intrahepatic Microbial Heterogeneity in Multifocal Hepatocellular Carcinoma and Its Association with Host Genomic and Transcriptomic Alterations. *Cancer Discov.* 15, 1630–1648. <https://doi.org/10.1158/2159-8290.CD-24-1259>.
12. Fu, A., Yao, B., Dong, T., Chen, Y., Yao, J., Liu, Y., Li, H., Bai, H., Liu, X., Zhang, Y., et al. (2022). Tumor-resident intracellular microbiota promotes metastatic colonization in breast cancer. *Cell* 185, 1356–1372.e26. <https://doi.org/10.1016/j.cell.2022.02.027>.
13. Battaglia, T.W., Mimpfen, I.L., Traets, J.J.H., van Hoeck, A., Zevenijn, L.J., Geurts, B.S., de Wit, G.F., Noë, M., Hofland, I., Vos, J.L., et al. (2024). A pan-cancer analysis of the microbiome in metastatic cancer. *Cell* 187, 2324–2335.e19. <https://doi.org/10.1016/j.cell.2024.03.021>.
14. Kalaora, S., Nagler, A., Nejman, D., Alon, M., Barbolin, C., Barnea, E., Keltalaars, S.L.C., Cheng, K., Vervier, K., Shental, N., et al. (2021). Identification of bacteria-derived HLA-bound peptides in melanoma. *Nature* 592, 138–143. <https://doi.org/10.1038/s41586-021-03368-8>.
15. Naghavian, R., Faigle, W., Oldrati, P., Wang, J., Toussaint, N.C., Qiu, Y., Medici, G., Wacker, M., Freudenmann, L.K., Bonté, P.E., et al. (2023). Microbial peptides activate tumour-infiltrating lymphocytes in glioblastoma. *Nature* 617, 807–817. <https://doi.org/10.1038/s41586-023-06081-w>.
16. Pushalkar, S., Hundeyin, M., Daley, D., Zambirinis, C.P., Kurz, E., Mishra, A., Mohan, N., Aykut, B., Usyk, M., Torres, L.E., et al. (2018). The pancreatic cancer microbiome promotes oncogenesis by induction of innate and adaptive immune suppression. *Cancer Discov.* 8, 403–416. <https://doi.org/10.1158/2159-8290.CD-17-1134>.
17. Rappold, P.M., Vuong, L., Leibold, J., Chakiryan, N.H., Curry, M., Kuo, F., Sabio, E., Jiang, H., Nixon, B.G., Liu, M., et al. (2022). A Targetable Myeloid Inflammatory State Governs Disease Recurrence in Clear-Cell Renal Cell Carcinoma. *Cancer Discov.* 12, 2308–2329. <https://doi.org/10.1158/2159-8290.CD-21-0925>.
18. McDermott, D.F., Huseni, M.A., Atkins, M.B., Motzer, R.J., Rini, B.I., Escudier, B., Fong, L., Joseph, R.W., Pal, S.K., Reeves, J.A., et al. (2018). Clinical activity and molecular correlates of response to atezolizumab alone or in combination with bevacizumab versus sunitinib in renal cell carcinoma. *Nat. Med.* 24, 749–757. <https://doi.org/10.1038/s41591-018-0053-3>.
19. Liss, M.A., Chen, Y., Rodriguez, R., Pruthi, D., Johnson-Pais, T., Wang, H., Mansour, A., White, J.R., and Kaushik, D. (2020). 2020, 1, 6, Microbiome within Primary Tumor Tissue from Renal Cell Carcinoma May Be Associated with PD-L1 Expression of the Venous Tumor Thrombus. *Adv. Urol.* 2020. <https://doi.org/10.1155/2020/9068068>.
20. Jingushi, K., Kawashima, A., Tanikawa, S., Saito, T., Yamamoto, A., Uemura, T., Sassi, N., Ishizuya, Y., Yamamoto, Y., Kato, T., et al. (2024). *Cutibacterium* acnes-derived extracellular vesicles promote tumor growth in renal cell carcinoma. *Cancer Sci.* 115, 2578–2587. <https://doi.org/10.1111/cas.16202>.
21. Kovaleva, O.V., Podlesnaya, P., Sorokin, M., Mochalnikova, V., Kataev, V., Khlopko, Y.A., Plotnikov, A.O., Stilidi, I.S., Kushlinskii, N.E., and Gratchev, A. (2022). Macrophage Phenotype in Combination with Tumor Microbiome Composition Predicts RCC Patients' Survival: A Pilot Study. *Biomedicines* 10, 1516. <https://doi.org/10.3390/biomedicines10071516>.
22. Turajlic, S., Xu, H., Litchfield, K., Rowan, A., Horswell, S., Chambers, T., O'Brien, T., Lopez, J.I., Watkins, T.B.K., Nicol, D., et al. (2018). Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. *Cell* 173, 595–610.e11. <https://doi.org/10.1016/j.cell.2018.03.043>.
23. Dréno, B., Pécastaings, S., Corvec, S., Veraldi, S., Khammari, A., and Roques, C. (2018). *Cutibacterium acnes* (Propionibacterium acnes) and acne vulgaris: a brief look at the latest updates. Preprint at Blackwell Publishing Ltd. <https://doi.org/10.1111/jdv.15043>.

24. Thoraval, L., Varin-Simon, J., Ohl, X., Velard, F., Refruveille, F., and Tang-Fichaux, M. (2025). Cutibacterium acnes and its complex host interaction in prosthetic joint infection: Current insights and future directions. *Res. Microbiol.* 176, 104265. <https://doi.org/10.1016/J.RES-MIC.2024.104265>.
25. Krishna, C., DiNatale, R.G., Kuo, F., Srivastava, R.M., Vuong, L., Chowell, D., Gupta, S., Vanderbilt, C., Purohit, T.A., Liu, M., et al. (2021). 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.e6. <https://doi.org/10.1016/j.ccell.2021.03.007>.
26. Xu, W., Lu, J., Liu, W.R., Anwaier, A., Wu, Y., Tian, X., Su, J.Q., Qu, Y.Y., Yang, J., Zhang, H., and Ye, D. (2023). Heterogeneity in tertiary lymphoid structures predicts distinct prognosis and immune microenvironment characterizations of clear cell renal cell carcinoma. *J. Immunother. Cancer* 11, e006667. <https://doi.org/10.1136/jitc-2023-006667>.
27. Leibovich, B.C., Blute, M.L., Cheville, J.C., Lohse, C.M., Frank, I., Kwon, E.D., Weaver, A.L., Parker, A.S., and Zincke, H. (2003). Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma: A stratification tool for prospective clinical trials. *Cancer* 97, 1663–1671. <https://doi.org/10.1002/cncr.11234>.
28. Zhang, H., Zhu, X., Friesen, T.J., Kwak, J.W., Pisarenko, T., Mekvanich, S., Velasco, M.A., Randolph, T.W., Kargl, J., and Houghton, A.M. (2022). Annexin A2/TLR2/MYD88 pathway induces arginase 1 expression in tumor-associated neutrophils. *J. Clin. Invest.* 132, e153643. <https://doi.org/10.1172/JCI153643>.
29. Uhlen, M., Karlsson, M.J., Zhong, W., Tebani, A., Pou, C., Mikes, J., Lakshmikanth, T., Forsström, B., Edfors, F., Odeberg, J., et al. (2019). A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. *Science* 366, eaax9198. <https://doi.org/10.1126/science.aax9198>.
30. Adrover, J.M., Han, X., Sun, L., Fujii, T., Sivetz, N., Daßler-Plenker, J., Evans, C., Peters, J., He, X.-Y., Cannon, C.D., et al. (2025). Neutrophils drive vascular occlusion, tumour necrosis and metastasis. *Nature* 645, 484–495. <https://doi.org/10.1038/s41586-025-09278-3>.
31. Braun, D.A., Street, K., Burke, K.P., Cookmeyer, D.L., Denize, T., Pedersen, C.B., Gohil, S.H., Schindler, N., Pomerance, L., Hirsch, L., et al. (2021). Progressive immune dysfunction with advancing disease stage in renal cell carcinoma. *Cancer Cell* 39, 632–648.e8. <https://doi.org/10.1016/j.ccell.2021.02.013>.
32. Fitzsimons, E., Qian, D., Enica, A., Thakkar, K., Augustine, M., Gamble, S., Reading, J.L., and Litchfield, K. (2024). A pan-cancer single-cell RNA-seq atlas of intratumoral B cells. *Cancer Cell* 42, 1784–1797.e4. <https://doi.org/10.1016/j.ccell.2024.09.011>.
33. Motzer, R.J., Banchereau, R., Hamidi, H., Powles, T., McDermott, D., Atkins, M.B., Escudier, B., Liu, L.F., Leng, N., Abbas, A.R., et al. (2020). Molecular Subsets in Renal Cancer Determine Outcome to Checkpoint and Angiogenesis Blockade. *Cancer Cell* 38, 803–817.e4. <https://doi.org/10.1016/j.ccell.2020.10.011>.
34. Ahn, H.K., Kim, K., Park, J., and Kim, K.H. (2022). Urinary microbiome profile in men with genitourinary malignancies. *Investig. Clin. Urol.* 63, 569–576. <https://doi.org/10.4111/icu.20220124>.
35. Ghaddar, B., Blaser, M.J., and De, S. (2026). Reliable detection of Host-Microbe Signatures in cancer using PRISM. *Cancer Cell* 44, 879–890.e3. <https://doi.org/10.1016/j.ccell.2026.01.007>.
36. Garcia Castillo, J., Fernandez, S., Campbell, T., Gonzalez Ventura, D., Williams, J., Ybarra, J., Flores Hernandez, N., Wells, E., Portnoy, D.A., and DuPage, M. (2025). Cellular mechanisms underlying beneficial versus detrimental effects of bacterial antitumor immunotherapy. *Immunity* 58, 2002–2018.e5. <https://doi.org/10.1016/j.immuni.2025.05.001>.
37. Silver, N.L., Dai, J., Kerr, T.D., Altemus, J., Garg, R., Simmons, H., Alban, T., Noel-Romas, L., Makarov, V., Shih, D.J.H., et al. (2026). Intratumoral bacteria are immunosuppressive and promote immunotherapy resistance in head and neck squamous cell carcinoma. *Nat. Cancer* 7, 80–97. <https://doi.org/10.1038/s43018-025-01067-1>.
38. Webster, G.F., and Leyden, J.J. (1980). Characterization of serum-independent polymorphonuclear leukocyte chemotactic factors produced by *Propionibacterium*. *Acnes. Inflammation* 4, 261–269. <https://doi.org/10.1007/BF00915027>.
39. Madli Puhvel, S., and Sakamoto, M. (1980). Cytotoxin Production by Comedonal Bacteria (*Propionibacterium* Acnes, *Propionibacterium Granulosum* and *Staphylococcus Epidermidis*). *J. Invest. Dermatol.* 74, 36–39. <https://doi.org/10.1111/1523-1747.ep12514592>.
40. Dessinioti, C., and Katsambas, A.D. (2010). The role of *Propionibacterium* acnes in acne pathogenesis: facts and controversies. *Clin. Dermatol.* 28, 2–7. <https://doi.org/10.1016/j.clinidermatol.2009.03.012>.
41. Kim, J., Ochoa, M.-T., Krutzik, S.R., Takeuchi, O., Uematsu, S., Legaspi, A.J., Brightbill, H.D., Holland, D., Cunliffe, W.J., Akira, S., et al. (2002). Activation of Toll-like Receptor 2 in Acne Triggers Inflammatory Cytokine Responses1. *J. Immunol.* 169, 1535–1541. <https://doi.org/10.4049/jimmunol.169.3.1535>.
42. Nagy, I.N., Pivarcsi, A., Koreck, A., Má Rta Szé, Á., and Kemé, L. (2005). Distinct Strains of *Propionibacterium* Acnes Induce Selective Human B-Defensin-2 and Interleukin-8 Expression in Human Keratinocytes through Toll-like Receptors.J. Invest. Dermatol. 124 (5), 931–938. <https://doi.org/10.1111/j.0022-202X.2005.23705.x>.
43. Jugeau, S., Tenaud, I., Knol, A.C., Jarrousse, V., Quereux, G., Khammari, A., and Dreno, B. (2005). Induction of toll-like receptors by *Propionibacterium* acnes. *Br. J. Dermatol.* 153, 1105–1113. <https://doi.org/10.1111/j.1365-2133.2005.06933.x>.
44. Dubus, M., Varin, J., Papa, S., Rammal, H., Chevrier, J., Maisonneuve, E., Mauprivez, C., Mongaret, C., Gangloff, S.C., Refruveille, F., and Kerdjoudj, H. (2020). Interaction of Cutibacterium acnes with human bone marrow derived mesenchymal stem cells: a step toward understanding bone implant-associated infection development. *Acta Biomater.* 104, 124–134. <https://doi.org/10.1016/j.actbio.2019.12.028>.
45. Hamann, T., Brüggemann, H., Feidenhansl, C., Ruci, E., Gallinger, J., Gallinat, S., and Hüpeden, J. (2025). Distinct Intraspecies Variation of Cutibacterium acnes and Staphylococcus epidermidis in Acne Vulgaris and Healthy Skin. *Microorganisms* 13, 299. <https://doi.org/10.3390/microorganisms13020299>.
46. Cros, M.P., Mir-Pedrol, J., Toloza, L., Knödseder, N., Maruotti, J., Zouboulis, C.C., Güell, M., and Fábrega, M.-J. (2023). New insights into the role of Cutibacterium acnes-derived extracellular vesicles in inflammatory skin disorders. *Sci. Rep.* 13, 16058. <https://doi.org/10.1038/s41598-023-43354-w>.
47. Chu, Y., Dai, E., Li, Y., Han, G., Pei, G., Ingram, D.R., Thakkar, K., Qin, J.J., Dang, M., Le, X., et al. (2023). Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance. *Nat. Med.* 29, 1550–1562. <https://doi.org/10.1038/s41591-023-02371-y>.
48. Yang, Y., Chen, X., Pan, J., Ning, H., Zhang, Y., Bo, Y., Ren, X., Li, J., Qin, S., Wang, D., et al. (2024). Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes. *Cell* 187, 4790–4811.e22. <https://doi.org/10.1016/j.cell.2024.06.038>.
49. Stewart, G.R., and Young, D.B. (2004). Heat-shock proteins and the host-pathogen interaction during bacterial infection. *Curr. Opin. Immunol.* 16, 506–510. <https://doi.org/10.1016/j.coi.2004.05.007>.
50. Wynn, T.A., and Vannella, K.M. (2016). Macrophages in Tissue Repair, Regeneration, and Fibrosis. Preprint at Cell Press. <https://doi.org/10.1016/j.immuni.2016.02.015>.
51. Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K.J., and Werb, Z. (2020). Concepts of extracellular matrix remodelling in tumour progression and metastasis. Preprint at Nature Research. <https://doi.org/10.1038/s41467-020-18794-x>.
52. Oza, B., Eisen, T., Frangou, E., Stewart, G.D., Bex, A., Ritchie, A.W.S., Kaplan, R., Smith, B., Davis, I.D., Stockler, M.R., et al. (2022). External Validation of the 2003 Leibovich Prognostic Score in Patients Randomly Assigned to SORCE, an International Phase III Trial of Adjuvant Sorafenib

- in Renal Cell Cancer. *J. Clin. Oncol.* **40**, 1772–1782. <https://doi.org/10.1200/JCO.21>.
53. Wang, H., Kim, S.J., Lei, Y., Wang, S., Wang, H., Huang, H., Zhang, H., and Tsung, A. (2024). Neutrophil extracellular traps in homeostasis and disease. Preprint at Springer Nature. <https://doi.org/10.1038/s41392-024-01933-x>.
54. Kim, H.J., Lee, Y.S., Lee, B.S., Han, C.H., Kim, S.G., and Kim, C.H. (2024). NLRP3 inflammasome activation and NETosis positively regulate each other and exacerbate proinflammatory responses: implications of NETosis inhibition for acne skin inflammation treatment. *Cell. Mol. Immunol.* **21**, 466–478. <https://doi.org/10.1038/s41423-024-01137-x>.
55. Wang, M., Rousseau, B., Qiu, K., Huang, G., Zhang, Y., Su, H., Le Bihan-Benjamin, C., Khatri, I., Artz, O., Foote, M.B., et al. (2024). Killing tumor-associated bacteria with a liposomal antibiotic generates neoantigens that induce anti-tumor immune responses. *Nat. Biotechnol.* **42**, 1263–1274. <https://doi.org/10.1038/s41587-023-01957-8>.
56. Davidsson, S., Mölling, P., Rider, J.R., Unemo, M., Karlsson, M.G., Carlsson, J., Andersson, S.O., Elgh, F., Söderquis, B., and Andrén, O. (2016). Frequency and typing of *Propionibacterium acnes* in prostate tissue obtained from men with and without prostate cancer. *Infect. Agents Cancer* **11**, 26. <https://doi.org/10.1186/s13027-016-0074-9>.
57. Brüggemann, H., and Al-Zeer, M.A. (2020). Bacterial signatures and their inflammatory potentials associated with prostate cancer. Preprint at Blackwell Munksgaard. <https://doi.org/10.1111/apm.13021>.
58. Fassi Fehri, L., Mak, T.N., Laube, B., Brinkmann, V., Ogilvie, L.A., Mollenkopf, H., Lein, M., Schmidt, T., Meyer, T.F., and Brüggemann, H. (2011). Prevalence of *Propionibacterium acnes* in diseased prostates and its inflammatory and transforming activity on prostate epithelial cells. *Int. J. Med. Microbiol.* **301**, 69–78. <https://doi.org/10.1016/j.ijmm.2010.08.014>.
59. RONALD, J., and COHEN, B.A.S.J.E.M.T.S. (2005). PROPIONIBACTERIUM ACNES ASSOCIATED WITH INFLAMMATION IN RADICAL PROSTATECTOMY SPECIMENS: A POSSIBLE LINK TO CANCER EVOLUTION? *J. Urol.* **173**, 1969–1974.
60. Alexeyev, O.A., Marklund, I., Shannon, B., Golovleva, I., Olsson, J., Andersson, C., Eriksson, I., Cohen, R., and Elgh, F. (2007). Direct visualization of *Propionibacterium acnes* in prostate tissue by multicolor fluorescent in situ hybridization assay. *J. Clin. Microbiol.* **45**, 3721–3728. <https://doi.org/10.1128/JCM.01543-07>.
61. Ashida, S., Kawada, C., Tanaka, H., Kurabayashi, A., Yagyu, K.I., Sakamoto, S., Maejima, K., Miyano, S., Daibata, M., Nakagawa, H., and Inoue, K. (2024). Cutibacterium acnes invades prostate epithelial cells to induce BRCAness as a possible pathogen of prostate cancer. *Prostate* **84**, 1056–1066. <https://doi.org/10.1002/pros.24723>.
62. Liu, J., Lichtenberg, T., Hoadley, K.A., Poisson, L.M., Lazar, A.J., Cherniack, A.D., Kovatich, A.J., Benz, C.C., Levine, D.A., Lee, A.V., et al. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* **173**, 400–416.e11. <https://doi.org/10.1016/j.cell.2018.02.052>.
63. Fernández-Sanromán, Á., Fendler, A., Tan, B.J.Y., Cattin, A.L., Spencer, C., Thompson, R., Au, L., Lobon, I., Pallikonda, H.A., Martin, A., et al. (2025). Tracking Nongenetic Evolution from Primary to Metastatic ccRCC: TRACERx Renal. *Cancer Discov.* **15**, OF1–OF23. <https://doi.org/10.1158/2159-8290.CD-24-0499>.
64. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
65. Wood, D.E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257. <https://doi.org/10.1186/S13059-019-1891-0/FIGURES/2>.
66. Hillmann, B., Al-Ghalith, G.A., Shields-Cutler, R.R., Zhu, Q., Knight, R., and Knights, D. (2020). SHOGUN: a modular, accurate and scalable framework for microbiome quantification. *Bioinformatics* **36**, 4088–4090. <https://doi.org/10.1093/BIOINFORMATICS/BTAA277>.
67. Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Preprint at Nature Publishing Group. <https://doi.org/10.1038/s41587-019-0209-9>.
68. Hoffman, G.E., and Schadt, E.E. (2016). variancePartition: Interpreting drivers of variation in complex gene expression studies. *BMC Bioinf.* **17**, 483. <https://doi.org/10.1186/s12859-016-1323-z>.
69. Kassambara, A., Kosinski, M., and Biecek, P. (2026). survminer: Drawing Survival Curves using “ggplot2.”. *Datanovia*.
70. Simon, N., Friedman, J., Hastie, T., and Tibshirani, R. (2011). Regularization paths for Cox’s proportional hazards model via coordinate descent. *J. Stat. Software* **39**, 1–13. <https://doi.org/10.18637/jss.v039.i05>.
71. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Software* **67**, 1–48. <https://doi.org/10.18637/jss.v067.i01>.
72. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
73. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). ClusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287. <https://doi.org/10.1089/omi.2011.0118>.
74. Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSEA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinformatics* **14**, 7. <https://doi.org/10.1186/1471-2105-14-7>.
75. Hall, B.G., and Nisbet, J. (2023). Building Phylogenetic Trees From Genome Sequences With kSNP4. *Mol. Biol. Evol.* **40**, msad235. <https://doi.org/10.1093/molbev/msad235>.
76. Schmidt, U., Weigert, M., Broaddus, C., and Myers, G. (2018). Cell detection with star-convex polygons. In *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Springer Verlag), pp. 265–273. [https://doi.org/10.1007/978-3-030-00934-2\\_30](https://doi.org/10.1007/978-3-030-00934-2_30).
77. Bankhead, P., Loughrey, M.B., Fernández, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., et al. (2017). QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878. <https://doi.org/10.1038/s41598-017-17204-5>.
78. Blampey, Q., Mulder, K., Gardet, M., Christodoulidis, S., Dutertre, C.A., André, F., Ginhoux, F., and Cournède, P.H. (2024). Sopa: a technology-invariant pipeline for analyses of image-based spatial omics. *Nat. Commun.* **15**, 4981. <https://doi.org/10.1038/s41467-024-48981-z>.
79. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>.
80. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15. <https://doi.org/10.1186/s13059-017-1382-0>.
81. Fang, Z., Liu, X., and Peltz, G. (2023). GSEAPy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics* **39**, btac757. <https://doi.org/10.1093/bioinformatics/btac757>.
82. The National Genomic Research Library v5.1, Genomics England.
83. (2015). Sample Preparation and Pretreatment Guide for Fresh Frozen Tissue using RNAscope® Fluorescent Multiplex Assay PART 1 Document Number 320513 Compatible with the RNAscope® Fluorescent Multiplex Detection Kit User Manual Part 2, see Document Number 320293. *ACD Bio*. [https://acdbio.com/system/files\\_force/320513\\_FreshFrozen%20SamplePrep%20Part1\\_UM\\_1\\_1052015.pdf](https://acdbio.com/system/files_force/320513_FreshFrozen%20SamplePrep%20Part1_UM_1_1052015.pdf).
84. Contents Handbook Overview and Navigation Overview 4 Visium HD Spatial Gene Expression Reagent Kits.

85. Santucci, P., Greenwood, D.J., Fearn, A., Chen, K., Jiang, H., and Gutierrez, M.G. (2021). Intracellular localisation of *Mycobacterium tuberculosis* affects efficacy of the antibiotic pyrazinamide. *Nat. Commun.* **12**, 3816. <https://doi.org/10.1038/s41467-021-24127-3>.
86. SuperScript™ IV First-Strand Synthesis System Protocol Outline.
87. O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., et al. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745. <https://doi.org/10.1093/nar/gkv1189>.
88. Richardson, L., Allen, B., Baldi, G., Beracochea, M., Bileschi, M.L., Burdett, T., Burgin, J., Caballero-Pérez, J., Cochrane, G., Colwell, L.J., et al. (2023). MGnify: the microbiome sequence data analysis resource in 2023. *Nucleic Acids Res.* **51**, D753–D759. <https://doi.org/10.1093/nar/gkac1080>.
89. Eisenhofer, R., Minich, J.J., Marotz, C., Cooper, A., Knight, R., and Weyrich, L.S. (2019). Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. Preprint at Elsevier. <https://doi.org/10.1016/j.tim.2018.11.003>.
90. Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., and Walker, A.W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87. <https://doi.org/10.1186/s12915-014-0087-z>.
91. Routy, B., Le Chatelier, E., Derosa, L., Duong, C.P.M., Alou, M.T., Dailière, R., Fluckiger, A., Messaoudene, M., Rauber, C., Roberti, M.P., et al. (2018). Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91–97. <https://doi.org/10.1126/science.aan3706>.
92. Poore, G.D., Kopylova, E., Zhu, Q., Carpenter, C., Fraraccio, S., Wandro, S., Kosciolk, T., Janssen, S., Metcalf, J., Song, S.J., et al. (2020). Microbiome analyses of blood and tissues suggest cancer diagnostic approach. *Nature* **579**, 567–574. <https://doi.org/10.1038/s41586-020-2095-1>.
93. Mao, Y., and Zhang, G. (2022). A complete, telomere-to-telomere human genome sequence presents new opportunities for evolutionary genomics. Preprint at Nature Research. <https://doi.org/10.1038/s41592-022-01512-4>.
94. Gene Ontology Consortium; Aleksander, S.A., Balhoff, J., Carbon, S., Cherry, J.M., Drabkin, H.J., Ebert, D., Feuerhann, M., Gaudet, P., Harris, N.L., et al. (2023). The Gene Ontology knowledgebase in 2023. *Genetics* **224**, iyad031. <https://doi.org/10.1093/genetics/iyad031>.
95. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst.* **1**, 417–425. <https://doi.org/10.1016/j.cels.2015.12.004>.
96. Manresa, M.C., Godson, C., and Taylor, C.T. (2014). Hypoxia-sensitive pathways in inflammation-driven fibrosis. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **307**, R1369–R1380. <https://doi.org/10.1152/ajpregu.00349.2014>.
97. Syrtveit, A.A., Ingebriktsen, L.M., Tegnander, A.F., Akslen, L.A., Wik, E., and Hovik, E.A. (2025). Tumor necrosis associates with aggressive breast cancer features, increased hypoxia signaling and reduced patient survival. *Sci. Rep.* **15**, 45518. <https://doi.org/10.1038/s41598-025-29905-3>.
98. Hiraoka, N., Ino, Y., Sekine, S., Tsuda, H., Shimada, K., Kosuge, T., Zavada, J., Yoshida, M., Yamada, K., Koyama, T., et al. (2010). Tumour necrosis is a postoperative prognostic marker for pancreatic cancer patients with a high interobserver reproducibility in histological evaluation. *Br. J. Cancer* **103**, 1057–1065. <https://doi.org/10.1038/sj.bjc.6605854>.
99. Meier-Kolthoff, J.P., and Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* **10**, 2182. <https://doi.org/10.1038/s41467-019-10210-3>.
100. Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114. <https://doi.org/10.1038/s41467-018-07641-9>.
101. Liu, Z., Chen, L., Cheng, H., Ao, J., Xiong, J., Liu, X., Chen, Y., Mao, Y., and Ji, M. (2024). Virtual formalin-fixed and paraffin-embedded staining of fresh brain tissue via stimulated Raman CycleGAN model. *Sci. Adv.* **10**, eadn3426.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE                                                        | SOURCE                                  | IDENTIFIER                                                                                                                                |
|----------------------------------------------------------------------------|-----------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Bacterial and virus strains</b>                                         |                                         |                                                                                                                                           |
| <i>Cutibacterium acnes</i>                                                 | ATCC                                    | ATCC 6919 = DSM1897                                                                                                                       |
| <i>Escherichia coli</i>                                                    | NEB                                     | DH5alpha                                                                                                                                  |
| <b>Biological samples</b>                                                  |                                         |                                                                                                                                           |
| Samples from 143 primary ccRCC patient tumors                              | TRACERx Renal Study                     | EGAS00001002793                                                                                                                           |
| Leukocyte cones                                                            | NHS Blood and Transplant Service        | NC24                                                                                                                                      |
| <b>Chemicals, peptides, and recombinant proteins</b>                       |                                         |                                                                                                                                           |
| M-CSF                                                                      | Miltenyi                                | 130–096–491                                                                                                                               |
| <b>Critical commercial assays</b>                                          |                                         |                                                                                                                                           |
| AllPrep DNA/RNA mini kit                                                   | Qiagen                                  | Cat# 80004                                                                                                                                |
| RNAscope™ Multiplex Fluorescent v2 Assay                                   | Bio-Techne                              | Cat# 323100                                                                                                                               |
| Visium HD, Human Transcriptome, 6.5 mm                                     | 10× Genomics                            | PN-1000675                                                                                                                                |
| RNA Mini Kit                                                               | Invitrogen                              | Cat # 12-183-018A                                                                                                                         |
| SuperScript™ IV First-Strand Synthesis System                              | Invitrogen                              | Cat # 18091050                                                                                                                            |
| <b>Deposited data</b>                                                      |                                         |                                                                                                                                           |
| TCGA KIRC RNA                                                              | EGA                                     | dbGaP - phs000178                                                                                                                         |
| TCGA KIRC PFS data                                                         | Lui et al. <sup>62</sup>                | <a href="https://doi.org/10.1016/j.cell.2018.02.052">https://doi.org/10.1016/j.cell.2018.02.052</a>                                       |
| TRACERx Renal RNAseq data                                                  | Fernández-Sanromán et al. <sup>63</sup> | –                                                                                                                                         |
| Visium HD                                                                  | This paper                              | <a href="https://doi.org/10.25418/crick.30865199">https://doi.org/10.25418/crick.30865199</a>                                             |
| <b>Oligonucleotides</b>                                                    |                                         |                                                                                                                                           |
| dapB RNAscope Negative Control Probe                                       | Bio-Techne                              | Cat# 312038-C2                                                                                                                            |
| <i>Cutibacterium</i> 16S RNAscope Probe – targeting 597–638 of NR_113028.1 | Bio-Techne                              | NA                                                                                                                                        |
| 16S probes - V3-V4, 341F-806R                                              | Novogene                                | NA                                                                                                                                        |
| <b>Software and algorithms</b>                                             |                                         |                                                                                                                                           |
| Trimmomatic v0.39                                                          | Bolger et al. <sup>64</sup>             | <a href="https://github.com/usadellab/trimmomatic">https://github.com/usadellab/trimmomatic</a>                                           |
| Kraken2 v2.1.3                                                             | Wood et al. <sup>65</sup>               | <a href="https://github.com/DerrickWood/kraken2">https://github.com/DerrickWood/kraken2</a>                                               |
| SHOGUN                                                                     | Hillmann et al. <sup>66</sup>           | <a href="https://github.com/knights-lab/SHOGUN">https://github.com/knights-lab/SHOGUN</a>                                                 |
| Qiime2 v2023                                                               | Boyle et al. <sup>67</sup>              | <a href="https://qiime2.org/">https://qiime2.org/</a>                                                                                     |
| VariancePartition R Package                                                | Hoffman et al. <sup>68</sup>            | <a href="https://github.com/GabrielHoffman/variancePartition">https://github.com/GabrielHoffman/variancePartition</a>                     |
| Survminer R Package                                                        | Kassambara et al. <sup>69</sup>         | <a href="https://github.com/kassambara/survminer">https://github.com/kassambara/survminer</a>                                             |
| Glmnet R Package                                                           | Simon et al. <sup>70</sup>              | <a href="https://github.com/cran/glmnet">https://github.com/cran/glmnet</a>                                                               |
| Lme4 R Package                                                             | Bates et al. <sup>71</sup>              | <a href="https://github.com/lme4/lme4/">https://github.com/lme4/lme4/</a>                                                                 |
| DESeq2 R Package                                                           | Love et al. <sup>72</sup>               | <a href="https://github.com/thelovelab/DESeq2">https://github.com/thelovelab/DESeq2</a>                                                   |
| clusterProfiler R Package                                                  | Yu et al. <sup>73</sup>                 | <a href="https://github.com/YuLab-SMU/clusterProfiler">https://github.com/YuLab-SMU/clusterProfiler</a>                                   |
| Gsva R Package                                                             | Hänzelmann et al. <sup>74</sup>         | <a href="https://github.com/rcastelo/GSVA">https://github.com/rcastelo/GSVA</a>                                                           |
| kSNP4                                                                      | Hall et al. <sup>75</sup>               | <a href="https://github.com/kissake/kSNP4">https://github.com/kissake/kSNP4</a>                                                           |
| Space Ranger v3.0.0                                                        | 10× Genomics                            | <a href="https://www.10xgenomics.com/support/software/space-ranger/3.0">https://www.10xgenomics.com/support/software/space-ranger/3.0</a> |
| StarDist v0.9.1                                                            | Schmidt et al. <sup>76</sup>            | <a href="https://github.com/stardist/stardist">https://github.com/stardist/stardist</a>                                                   |
| QuPath v0.5.1                                                              | Bankhead et al. <sup>77</sup>           | <a href="https://github.com/qupath">https://github.com/qupath</a>                                                                         |
| Sopa v2.0                                                                  | Blampey et al. <sup>78</sup>            | <a href="https://github.com/prism-oncology/sopa">https://github.com/prism-oncology/sopa</a>                                               |
| SciPy v1.15                                                                | Virtanen et al. <sup>79</sup>           | <a href="https://github.com/scipy/scipy">https://github.com/scipy/scipy</a>                                                               |

(Continued on next page)

### Continued

| REAGENT or RESOURCE       | SOURCE                    | IDENTIFIER                                                                        |
|---------------------------|---------------------------|-----------------------------------------------------------------------------------|
| ScanPy v.1.11             | Wolf et al. <sup>80</sup> | <a href="https://github.com/scverse/scanpy">https://github.com/scverse/scanpy</a> |
| GSEAPy v.1.1.5            | Fang et al. <sup>81</sup> | <a href="https://github.com/zqfang/GSEAPy">https://github.com/zqfang/GSEAPy</a>   |
| <b>Other</b>              |                           |                                                                                   |
| Fluorescent slide scanner | AKOYA                     | PhenolImager™                                                                     |
| Confocal microscope       | Olympus/Evident           | Olympus FV-4000                                                                   |

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

### TRACERx Renal (TxRenal) cohort

The TRACERx Renal prospective cohort study has been continuously recruiting patients since its initiation and ethical approval in 2012 (National Health Service Research Ethics Committee approval 11/LO/1996). The study was conducted in accordance with recognized ethical guidelines. The study is sponsored by the Royal Marsden NHS Foundation Trust and recruits from four hospital sites across London and Edinburgh: Royal Marsden Hospital NHS Foundation Trust, Guy's and St Thomas' Hospital NHS Foundation Trust, Royal Free Hospital NHS Foundation Trust and Western General Hospital (NHS Lothian).

The following criteria were applied when recruiting patients: i) aged over 18; ii) confirmed or suspected renal cell carcinoma leading to nephrectomy or metasectomy; iii) disease managed and treated following national or local standard of care; iv) written informed consent to the use of fresh and archived tissue and blood and anonymized clinical information for research. Patients are excluded based on the following criteria: i) concomitant physical or mental problems deemed to prevent completion of treatment or follow-up; ii) lack of adequate tissue.

In addition to the general cohort requirements, this study further excluded patients if: i) they had received neo-adjuvant immunotherapy; ii) they had documented germline renal cell carcinoma disposition. Tumors were only included if they were: i) primary tumors; ii) histologically confirmed as being ccRCC.

Clinical and demographic variation in the cohort is representative of the patient population eligible for nephrectomy at the time of recruitment. Demographic data available was self reported and includes: Sex, Age, Ethnicity, Pre-operative BMI and smoking status. Information on gender, race and socioeconomic status was not collected. Histological data available was reported by pathologists and includes: Tumor size at nephrectomy, TNM stage, highest Fuhrman grade, Necrosis, peri-nephric fat invasion, microscopic vascular invasion, renal vein IVC involvement, presence of a tumor thrombus, presence of sarcomatoid changes. Clinical data available was reported by an oncologist and includes: PFS from nephrectomy, OS from nephrectomy, neo-adjuvant targeted therapy, and first line therapy and best response. Progression Free Survival is defined as time from curative surgery to relapse for stage 1–3 and from surgical resection to first progression for stage 4. Overall survival is defined as cancer-specific death (Sup. Table S4). While progression free survival is stages 1–3 would more accurately be framed as recurrence free survival, this nomenclature is the one adopted across TRACERx Renal publications and therefore retained for consistency.

### TCGA renal cohort

Raw fastq files were obtained for the bulk RNA-sequenced TCGA ccRCC cohort (KIRC) (Controlled Access under dbGaP - phs000178). Clinical data was obtained from the GDC data portal and PFS data from the publication by Liu and colleagues.<sup>62</sup>

Samples were only included if they respected the following criteria: i) primary tumor sample; ii) no prior therapy iii) available PFS and OS data.

### Sample collection – TxRenal cohort

All tumors collected under the TRACERx Renal study were processed following the same protocol as previously described.<sup>22</sup> Briefly, following macroscopic pathological review, tumors were dissected along the longest axis and the resulting tumor slice extensively sampled using 6-mm sterile punch biopsy needles. To avoid cross-contamination between samples from the same tumor, a new sterile punch needle was used for each sample. Samples are named with prefix R for primary tumor region, THR for primary tumor thrombus, N for tumor adjacent tissue and B/BC for blood.

Patient blood was collected at the time of surgery and centrifuged to obtain plasma and buffy coat fractions. Buffy coat was extracted for sequencing.

## METHOD DETAILS

### Whole genome sequencing

143 primary tumors spanning all 4 tumor stages was selected for whole genome sequencing (WGS) from the TRACERx Renal cohort. From these tumors, a median of 4 regions per tumor underwent DNA and RNA co-extraction from the stored, snap-frozen punch

biopsies as previously described.<sup>22</sup> Extracted DNA was sent to Genomics England<sup>82</sup> for sequencing using Illumina's paired-end WGS, with a target read-depth of 100×.

### RNA-sequencing

We used RNA data from a published subset of 76 patient tumors from the TxRenal cohort which underwent multi-regional RNA sequencing.<sup>63</sup> Briefly, RNA was co-extracted with DNA as described above. For samples with sufficient RNA quality (RNA integrity score  $\geq 5$ ), 100ng of RNA was rRNA depleted and paired-end sequenced with a target read-depth of 100×.

### 16S-sequencing

Contaminants introduced during DNA extraction, sample preparation for shipping and in commercial sequencing facilities, were identified by sequencing 8 controls with 16S rDNA sequencing (V3-V4, 341F-806R). Three water controls were taken through the same extraction protocol as our tumor samples and three more were added when preparing samples for shipping to sequencing facilities. In a final two samples, known amounts of DNA from eight bacterial genera were added. The 8 controls were sequenced with 16S rDNA sequencing (V3-V4, 341F-806R).

### Bacterial culture from tumor samples

Tumor regions were selected for culturing based on their WGS-estimated *Cutibacterium* load, or the load of adjacent regions when insufficient tissue from profiled regions was available.  $\sim 1\text{mm}^3$  of snap frozen tumor tissue was homogenised and spread on two fastidious anaerobe agar plates supplemented with 10% defibrinated horse blood. Both plates were incubated at 37°C for 14 days. One plate was kept in aerobic conditions and the other in anaerobic conditions. DNA was extracted from all resulting colonies and shotgun sequenced for strain identification. As  $\sim 0.5\text{mm}^3$  of snap-frozen tumor tissue is needed for successful culture per condition, several adjacent tumor regions were combined to reach the minimum volume. In cases where the target  $1\text{cm}^3$  of total tumor tissue could not be reached due to tissue availability, no colony growth was observed. Due to variability in input material, only cases with growth of at least one colony were considered for further analysis; these consistently ranged from  $1\text{cm}^3$  to  $1.2\text{cm}^3$  of total input tissue. Absence of bacterial growth in these lower volume cases regardless of genomically-defined *Cutibacterium* status, is consistent with bacterial growth depending on tissue-associated bacterial load rather than contamination introduced during processing, which would be expected to be independent of tissue volume.

### Bacterial and TME imaging by FISH (1) imaging protocol

Tumor regions were selected for imaging based on their WGS-estimated *Cutibacterium* load, or distance from "Cuti-High" regions when insufficient tissue from profiled regions was available. Snap-frozen tumor tissue was embedded in OCT and 10-15 $\mu\text{m}$  cryosections placed on charged microscopy slides. Slices were fixed in 4% PFA for 1 h before commencement of the standard RNAscope protocol.<sup>83</sup> Pre-treatment was carried out without epitope retrieval and with a 10-min incubation of Pre-Treat Pro at 40°C. Signal amplification and detection was carried out using the automated Leica Bond system.

The following probes were used: i) DAPI ii) Dap-B (312038-C2 - Bio-Techne) - Opal 570 - Channel 1 iii) *Cutibacterium* 16S - Opal 690 - Channel 2. Slides were scanned using the Phenolmager™ (AKOYA) at 40× and both probes exposed for 220ms.

For high-resolution images, regions of interest were imaged with a confocal microscope (Olympus FV-4000). Images were taken at 40× with a 3× zoom. For each region of interest, 8 positions in the z stack were imaged at 0.64 $\mu\text{m}$  intervals.

### Bacterial and TME imaging by FISH (2) probe testing

The *Cutibacterium* 16S probe was designed with the provider (Bio-Techne), targeting 597–638 of NR\_113028.1. To test the probe's accuracy, we constructed a *Cutibacterium*-tissue sandwich, where tumor adjacent normal kidney tissue was sliced in half and pelleted *Cutibacterium acnes* (ATCC 6919) was streaked between the two slices before fixing the sandwich in PFA. We also tested probe specificity using a PFA-fixed pellet of *Escherichia Coli* (DH5alpha). Both controls were marked with three RNAscope probes coupled with their respective Opal fluorophores: Pan-bacterium 16S probe - Opal 690 - Channel 1, *Cutibacterium* 16S probe - Opal 780 - Channel 2, Dap-B probe - Opal 570 - Channel 3. The Dap-B probe targets the dihydrodipicolinate reductase gene of *Bacillus subtilis*, a common soil bacterium which is not expected in human tissue and is therefore used to detect non-specific binding. The slides were scanned at 40× with 200ms exposure for each of the three probes using the Phenolmager™. The pan-bacterium marking was visible across all three controls, with close overlap with the *Cutibacterium* marking in the positive controls. No *Cutibacterium* signal was discernible in the *E. coli* pellet, confirming the specificity of the probe (Figures S9A and S9B).

### VisiumHD protocol

Four tumor regions were selected based on their WGS-estimated *Cutibacterium* load. As described above, 2 serial cryosections were taken and placed on charged slides. For each of the four samples, one section underwent the RNAscope protocol for *Cutibacterium* identification, the second followed the standard fresh frozen VisiumHD protocol following manufacturer's instructions,<sup>84</sup> starting with H&E staining. The only deviation from the protocol consisted in skipping the addition of the Ribonuclease Vanadyl Complex working solution prior to H&E imaging as its addition washed the eosin out of the tissue. cDNA was

amplified for 11 cycles and sequenced on Illumina NovaSeq X+ to a total of 900M reads across the four samples, pooled by percent of tissue in the capture area which ranged from 14% to 33%.

### Cutibacterium-macrophage co-cultures

Undifferentiated monocytes were obtained from healthy donor PBMCs as previously described.<sup>85</sup> Monocytes were differentiated through co-culture with M-CSF (50ng/mL) for 1 week 15000 differentiated macrophage cells were co-cultured under aerobic conditions with i) nothing ii) live *Cutibacterium acnes* (ATCC 6919) at a multiplicity of infection of 10 iii) heat-inactivated *Cutibacterium acnes* (ATCC 6919) at a multiplicity of infection of 10, for 2, 48 and 72 h respectively. This was repeated for 3 separate donors. Due to limited cells obtained from one donor, the 72 h time point could not be performed for that donor. At the appropriate time point, the cells were spun down and supernatant removed to remove dead cells. 150μL of trizol was then added before freezing at −80°C. RNA was extracted using the Invitrogen RNA Mini Kit following manufacturer's instructions. 100ng of resulting RNA was converted to cDNA for RT-qPCR. Standard RT-qPCR protocol was followed<sup>86</sup> with the following primers (Merck Life Sciences):

- CXCL8 - Forward: GAGAGTATTGAGAGTGGACCAC, Reverse: CACAACCCTCTGCACCCAGTTT;
- IL1B - Forward: CCACAGACCTTCCAGGAGAATG, Reverse: GTGCAGTTCAGTGATCGTACAGG;
- C1QA - Forward: CGAGCACCAGACGGGAAGAAAG, Reverse: AGGTTCCCCCTGGTCTCCTTTA;
- IL10 - Forward: TCTCCGAGATGCCTTCAGCAGA, Reverse: TCAGACAAGGCTTGGCAACCCA;
- CD163 - Forward: GCCGAAAAGTCCAGTGTGAAGG, Reverse: GCCATTCAAGTCGGACATCTG;
- beta-actin - Forward: CACCATTGGCAATGAGCGGTC, Reverse: AGGTCTTTCGGATGTCCACGT.

Each primer-condition pair was run in duplicate.

Supernatant and cell lysate from the 2 h and 72 h co-cultures were grown under anaerobic conditions and colony forming units quantified to ensure live bacterial growth throughout the co-cultures. At both timepoints, more CFUs were obtained from cell lysate than supernatant and none from the heat-killed culture conditions, indicating continued survival both inside and out of macrophages (Figures S10A–S10E).

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Bacterial identification pipeline (0)

To reliably identify bacterial reads from WGS or bulk RNA-seq, a pipeline was designed combining rigorous alignment criteria, technical denoising and algorithmic and experimental decontamination. The final outputs are validated technically and experimentally (Sup.Figure S1A).

For better comparison between the TxRenal and TCGA datasets, only genera identified in both were kept for further analysis.

### Bacterial identification pipeline (1) bacterial read identification from WGS and RNAseq

For WGS, only reads that do not map to the human hg38 reference genome are considered and filtered for high quality reads (trimomatic<sup>64</sup>), sliding window-phred>15, length>35bp). All RNA-seq reads are considered and similarly filtered for high quality reads. Reads were first mapped to the NCBI's RefSeq-82 database<sup>87</sup> of prokaryotic and human (hg38) genomes using a fast k-mer approach (KRAKEN2<sup>65</sup>) to filter out human reads. The remaining reads were aligned to the MGnify's gut microbiome ref.<sup>88</sup> (as of April 2023) using bowtie2's direct genomic alignment in SHOGUN.<sup>66</sup> This ensures only microbes known to be found in humans were considered.

To express bacterial counts relative to human DNA content, thereby making them comparable across samples, they were normalised by division by the sample's total read depth (including human reads) and multiplication by the rounded median read count across the cohort (TxRenal WGS: 3,000,000,000, TxRenal RNA, TCGA RNA: 100,000,000).

### Bacterial identification pipeline (2) removing technical noise

To remove any low-quality genomes, each reference genome was assessed for mapping bias. Any genome that was assigned more than 100 reads across the TxRenal WGS cohort was removed if more than 30% of those reads overlapped in the same position, as this is indicative of human genome contamination in the bacterial reference genome. Nine reference genomes were identified in this way in the TxRenal WGS cohort and removed from consideration across the WGS and RNA-seq cohorts. The same approach could not be applied to the RNA-seq cohorts as a strong bias toward mapping to highly expressed 16S genes is expected for RNA data.

### Bacterial identification pipeline (3) bacterial read identification from 16S sequencing

Qiime2<sup>67</sup> was used to classify bacterial genera from 16S sequencing data. A probe-specific classifier was built from the silva 138 database. Raw reads were then denoised using dada2 and classified at the genus-level using qiime2's sklearn-based feature classifier.

#### Bacterial identification pipeline (4) removing contaminants from bacterial data

We implemented safeguards in our sampling procedure to reduce the risk of bacterial contamination. Samples were taken from a central tumor section which would not have been exposed to the environment during surgery. Additionally, we used a new, sterile punch needle for each sample to minimise the risk of cross-contamination between samples from the same tumor. Beyond these mitigations, we identified contaminants that may have been introduced at any stage from surgery to sequencing experimentally where possible and algorithmically when experimental identification was not possible.

Contaminants introduced during DNA extraction, sample preparation for shipping and in commercial sequencing facilities, were identified by sequencing 8 controls with 16S rDNA sequencing (V3-V4, 341F-806R). Three water controls were taken through the same extraction protocol as our tumor samples and three more were added when preparing samples for shipping to sequencing facilities. In a final two samples, known amounts of DNA from eight bacterial genera were added. The 8 controls were sequenced with 16S rDNA sequencing (V3-V4, 341F-806R). Any genera that were identified in any of the water controls or in addition to the eight expected genera in the positive control were removed from our tumor datasets (Sup. Table S2).

Hospital and sequencing-batch-specific contaminants, for which we couldn't retrospectively sequence blanks, were identified using the variancePartition R package.<sup>68</sup> 45 and 58 genera were removed from the TxRenal and TCGA datasets respectively because >5% of their variance was explained by either Hospital site or the sequencing batch they were in (sequencing batch available for TxRenal but not TCGA) (Sup. Table S2).

Genera reported as common contaminants in low-biomass studies<sup>89,90</sup> were removed from consideration unless also previously identified in relation to kidney cancer.<sup>19–21,34,91</sup> (Sup. Table S3).

Finally, in the TxRenal cohort, peripheral blood samples taken at the time of surgery were extracted and sequenced alongside the tumor samples. As enrichment in the tumor through proliferation is expected for biologically relevant bacteria, we only kept genera that showed significant enrichment in the tumor. This was assessed using a paired Wilcoxon test of the max abundance region in the tumor relative to the peripheral blood. Compared to tumour-adjacent normal renal tissue the microenvironment of which is likely to be affected by the tumor, peripheral blood was deemed as a cleaner enrichment control and the bacterial presence in normal tissue was instead explored biologically. As no matched blood control was available for TCGA, only genera that passed these stringent filters in TxRenal were considered in TCGA (Sup. Table S2).

Throughout decontamination, we opted to completely remove all reads assigned to genera which were identified as potential contaminants. This approach was chosen due to concerns of introducing noise when regressing batch effects in sparse matrices such as bacterial read counts, one of the major concerns leading to the retraction of Poore and colleagues TCGA microbiome paper.<sup>92</sup>

This stringent decontamination approach yields genera with high confidence biological enrichment in tumors, but may eliminate some true biological signal, especially for genera for which both biological and contaminant strains may exist.

Secondary analyses on bacteria localisation within the tumor (Figure 1E) further support true bacterial presence by identifying significant enrichment in the center of the tumor, rather than the tumor boundary where external contamination would be expected to accumulate.

#### Bacterial identification pipeline (5) aligning to T2T reference genome

To verify that reads mapping to *Cutibacterium*, the most abundant genus, are not human reads from regions badly resolved in the hg38 reference genome, all reads assigned to *Cutibacterium* were mapped back to the telomere to telomere (T2T) human reference genome<sup>93</sup> using bwa. Only 1 of 141,748 *Cutibacterium* reads aligned to the T2T genome. As only 30 bases of this read aligned to T2T vs. 150 to the *Cutibacterium* genome, it was kept as a true *Cutibacterium* assignment.

#### Bacterial identification pipeline (6) validating repeatability

To confirm reliable recall of bacterial genera, six samples were WGS twice at a 3 months interval (Figure 1E). Genera with more than 25 assigned counts had equivalent counts in each repeat, with a median difference of only 9 reads and a maximum difference of 63 reads (=10% of counts), further confirming lack of contamination at the sequencing step. However, genera with fewer than 25 assigned counts could fail to be detected in one of the repeats, indicating that while WGS is a reliable approach to detect high abundance genera, it is likely to misrepresent the diversity of less abundant genera.

#### Tumour-level spatial analyses

Annotations of sample distances relative to one another and the tumor boundary were obtained from annotated pictures of the tumor slices following punch-biopsy. “Low”, “Mid” and “Cuti-High” regions are defined by tertiles of *Cutibacterium* abundance. 7 cm was chosen as the threshold for small tumors following clinical guidelines in ccRCC. All statistical tests comparing “Low”, “Mid” and “Cuti-High” within tumors are paired Wilcoxon tests to account for shared patient background. Continuous correlations are assessed by spearman's test.

#### Survival analyses

For survival analyses, cox regressions were run using the Survminer R package.<sup>69</sup> In the TxRenal cohort, tumors were defined by their maximum abundance region. Survival was initially assessed with logged *Cutibacterium* as a continuous to avoid loss of information from categorisation (TxRenal: PFS  $p = 0.0022$ , OS  $p = 0.21$ ; TCGA: PFS  $p = 0.00044$ , OS  $p = 3.99 \times 10^{-5}$ ). Patients were then split into

tertiles based on *Cutibacterium* abundance for visualisation and group-based comparisons. We selected tertiles rather than an arbitrary absolute threshold because no biologically established cut-off currently exists for defining bacteria-high ccRCC tumor. Survival was represented graphically using Kaplan Meier curves. The “Cuti-High” and “Mid” tertiles were grouped due to comparable survival profiles. Survival benefit was assessed within pathological stage groupings as described in the text.

To identify other genera potentially contributing to progression free survival, we used an elastic net-based cox regression with 5-fold cross validation using the glmnet R package<sup>70</sup> with family “cox”. Significance of resulting associations was determined with a univariate cox regression.

All TCGA patients were treatment-naïve by inclusion criteria. Within TxRenal, 7 patients received neo-adjuvant VEGFi prior to surgery; *Cutibacterium* abundance in these patients was not significantly different from stage-matched treatment-naïve patients, suggesting neo-adjuvant therapy did not substantially alter bacterial load. Post-operatively, 11 patients received VEGFi and 4 received checkpoint inhibitors (CPI); no association was observed between *Cutibacterium* abundance and VEGFi treatment response. Given the small number of CPI-treated patients ( $n = 4$ ) we do not draw conclusions from this group, though we note a trend consistent with the broader prognostic association, where treatment responders showed lower *Cutibacterium* abundance pre-treatment than non-responders.

### Clinical risk factor analyses

Pathological stages were grouped into early (stage I & II) and late (stage III & IV) when comparing *Cutibacterium* abundance, (Wilcoxon test for significance). Association with other clinical risk factors was assessed using a linear mixed effect model predicting logged *Cutibacterium* abundance with grouped pathological stage as a random effect using the R package lme4.<sup>71</sup> Significance of the clinical factor predicting *Cutibacterium* abundance was determined by running an anova test against a null model predicting *Cutibacterium* abundance with only the random effect of pathological stage.

### Differential gene expression and enrichment analyses

To identify human genes associated in expression with *Cutibacterium* abundance across tumors, a differential gene expression analysis was run using the DESeq2 R package<sup>72</sup> and *Cutibacterium* abundance (log10) as a continuous variable. Significantly upregulated genes with a log2 fold change greater than 0.5 were tested for enrichment of “Biological Process” pathways using enrichGO from the clusterProfiler R package.<sup>73</sup> The top 20 significantly enriched pathways are shown, colored by thematic groupings.

To test for pathway enrichment within tumors, gene signatures were scored for each sample using the gsva R function<sup>74</sup> with ssGSEA method. The maximum pathway enrichment score amongst each tumor’s Cuti-High/Mid regions was compared to the maximum pathway enrichment score in each tumor’s Cuti-Low regions using a paired Wilcoxon test for significance. When comparing pathway enrichments by tertile across tumors in TCGA, the same method was applied but significance assessed with an unpaired Wilcoxon test. Where several gene signatures were assessed at once, tests were FDR corrected. Additionally, when several signatures were displayed in the same plot, scores were first normalised globally across all signatures and samples (Z-scored) before plotting the median difference in z-scores across all samples for a given signature, to enable more meaningful comparison of effect magnitude across pathways. Gene signatures assessed for enrichment in this study are summarised below:

peri\_necrosis\_adover<sup>30</sup>: *S100A9*, *LGALS3*, *FTH1*, *LCN2*, *JUND*, *PNRC1*, *NDRG1*, *SLC2A1*, *S100A8*, *CEBPB*, *FTL1*, *GPI1*, *RACK1*, *GM112114*, *CXCL2*, *MMP12*, *ALDOA*, *IL1RN*, *DDIT3*, *HSPA1B*

motzer\_t\_effector<sup>33</sup>: *CD8A*, *EOMES*, *PRF1*, *IFNG*, *CD274*

braun\_Tcell\_Inhibitory<sup>31</sup>: *BATF*, *BCL6*, *BTLA*, *CD160*, *CD244*, *CTLA4*, *GATA3*, *HAVCR2*, *HIF1A*, *ID2*, *ID3*, *IRF4*, *KLRG1*, *LAG3*, *NFATC1*, *PDCD1*, *PRDM1*, *TIGIT*, *TNFRSF14*, *TOX*

braun\_Tcell\_Terminal\_differentiation<sup>31</sup>: *TIGIT*, *PDCD1*, *CD274*, *CTLA4*, *LAG3*, *HAVCR2*, *CD244*, *CD160*

braun\_Tcell\_Exhaustion<sup>31</sup>: *CD3E*, *PDCD1*, *IL2RA*, *CD69*, *ICOS*, *CD24*, *TNFRSF4*, *TNFRSF9*, *CD28*, *CD27*

braun\_Tcell\_Activation\_Dysfunction<sup>31</sup>: *ALDOA*, *ANKRD52*, *B3GNT2*, *BCL3*, *BCL6*, *BRAP*, *C6orf89*, *CABLES1*, *CACNB3*, *CAMK2D*, *CCNG2*, *CCR7*, *CD19*, *CD5*, *CD74*, *CERK*, *CLN6*, *CRTC3*, *CSNK1D*, *DTWD1*, *EIF3F*, *EML4*, *FAM168B*, *FAS*, *FASN*, *FBXL20*, *FOXO1*, *FRRS1*, *GATA3*, *GNA13*, *GPR132*, *GPR35*, *GRASP*, *IER5*, *IFI30*, *IFT80*, *IKZF2*, *IL2RG*, *INSYN2B*, *ISCU*, *ITGAE*, *JAK2*, *KDM4B*, *KDM5B*, *KLF3*, *LATS2*, *LGMN*, *MAP2K7*, *MKRN1*, *MTMR3*, *MYB*, *NDUFV1*, *NOTCH2*, *NPEPPS*, *NR4A1*, *NSMCE1*, *OSBPL11*, *PARP3*, *PCGF5*, *PHC2*, *PIAS2*, *PIGH*, *PIM3*, *PINK1*, *PIP4P1*, *PNRC1*, *POU2AF1*, *QRICH1*, *RAB2A*, *RARA*, *RASGEF1B*, *RASSF2*, *RELB*, *RNF149*, *SCIN*, *SEC24B*, *SGPP1*, *SH2B3*, *SLAMF6*, *SLAMF7*, *SLC25A33*, *SLC28A2*, *SLC2A6*, *SNTB1*, *SNX18*, *SPI1*, *SPRY1*, *SWAP70*, *TBC1D22B*, *TMEM243*, *TNIP1*, *TRAT1*, *TUBB2B*, *UTP15*, *WTAP*, *YPEL3*, *ZFP62*

braun\_Tcell\_Cell\_stress<sup>31</sup>: *BAG3*, *BLOC1S5-TXNDC5*, *CALU*, *DNAJB1*, *DUSP1*, *EGR1*, *FOS*, *FOSB*, *HIF1A*, *HSP90AA1*, *HSP90AB1*, *HSP90AB2P*, *HSP90AB3P*, *HSP90B1*, *HSPA1A*, *HSPA1B*, *HSPA6*, *HSPB1*, *HSPH1*, *IER2*, *JUN*, *JUNB*, *NFKBIA*, *NFKBIZ*, *RGS2*, *SLC2A3*, *SOC3*, *UBC*, *ZFAND2A*, *ZFP36*, *ZFP36L1*.

Ev\_B\_naive<sup>32</sup>: *CD19*, *MS4A1*, *FCER2*, *TCL1A*, *FCMR*.

Ev\_B\_mem<sup>32</sup>: *CD19*, *MS4A1*, *BLK*, *LY86*, *LTB*, *CRIP1*.

Ev\_B\_activated<sup>32</sup>: *CD19*, *MS4A1*, *CD69*, *CD83*, *NR4A2*.

Ev\_B\_GC<sup>32</sup>: *CD19*, *MS4A1*, *SUGCT*, *BCL6*, *SSBP2*, *STAG3*, *RGS13*, *LCK*, *MYO1E*, *IRAG2*, *LAT2*, *RRAS2*, *MARCKSL1*, *SERPINA9*, *PRAPSAP2*, *LMO2*, *CD22*, *IL4R*.

Ev\_Plasma\_conventional<sup>32</sup>: *JCHAIN*, *IGKC*, *IGHG1*, *IGHG3*, *SDC1*, *DERL3*, *XBP1*, *MZB1*.

Ev\_Plasma\_stressed<sup>32</sup>: JCHAIN, IGKC, SDC1, DERL3, XBP1, MZB1, HSPA1B, HSPA1A, DNAJB1.

Hypoxia\_response (GO: 0001666)<sup>94</sup>: ABAT, ACVRL1, ADA, ADM, ADORA1, ADSL, AGER, ALAD, ALAS2, ANG, ANGPT2, APAF1, BIRC2, ARNT, ASCL2, ATM, BMP2, BNIP3, CA9, CAPN2, CASP3, CASP9, CAT, CAV1, CAV3, CBFA2T3, CD38, CHRNA4, CHRNA7, CHRN2, COMT, CLDN3, CPT1A, CREBBP, CRYAB, CYP1A1, ACE, DIO3, DPP4, DRD2, EDNRA, EGR1, EP300, EPAS1, EPO, ERCC2, ERCC3, F7, PTK2B, VEGFD, FOSL2, HIF1A, HK2, HMOX2, HSD11B2, HSPG2, IL1A, ITGA2, ITPR1, ITPR2, KCNA5, KCNJ8, KCNJ11, KCNMA1, LEP, LIF, LOXL2, LTA, SMAD3, SMAD4, MB, MECP2, MMP2, MMP14, MT3, COX1, COX2, CYTB, MTHFR, ND1, ND2, ND4, ND5, MYB, NF1, NOS1, NOS2, NPPC, NR4A2, P2RX3, PENK, PGF, PIN1, PKLR, PLAT, PLA2, PLOD1, PLOD2, PML, PPARA, PRKAA1, PSEN2, RAD21, REST, RYR1, RYR2, CXCL12, SLC2A1, SLC6A4, SOD2, SOD3, SOX4, SRF, ADAM17, TEK, TFAM, TFRC, TGFB2, TGFB3, TGFB3, TH, THBS1, TLR2, TNF, HSP90B1, UCP2, UCP3, USF1, VCAM1, VEGFA, VEGFB, VEGFC, XRCC1, CXCR4, FZD4, BECN1, LIMD1, PDLIM1, LONP1, ADIPOQ, TM9SF4, ARNT2, CITED2, POSTN, SCAP, P2RX2, SCFD1, TXN2, SLC2A8, ANGPTL4, DDIT4, EGLN1, ALKBH5, AGTRAP, NGB, TRPV4, HIF3A, CIAO3, PLEKHN1, WDR83, AJUBA, MYOCD, EGLN2, EGLN3, CYGB, WTIP, FUND1, MIR223, CD24, MTCO2P12.

Hypoxia\_Hallmark<sup>95</sup>: ACKR3, ADM, ADORA2B, AK4, AKAP12, ALDOA, ALDOB, ALDOC, AMPD3, ANGPTL4, ANKZF1, ANXA2, ATP3, ATP7A, B3GALT6, B4GALNT2, BCAN, BCL2, BGN, BHLHE40, BNIP3L, BRS3, BTG1, CA12, CASP6, CAV1, CAVIN1, CAVIN3, CCN1, CCN2, CCN5, CCNG2, CDKN1A, CDKN1B, CDKN1C, CHST2, CHST3, CITED2, COL5A1, CP, CSRP2, CXCR4, DCN, DDIT3, DDIT4, DPYSL4, DTNA, DUSP1, EDN2, EFNA1, EFNA3, EGFR, ENO1, ENO2, ENO3, ERO1A, ERRF1, ETS1, EXT1, F3, FAM162A, FBP1, FOS, FOSL2, FOXO3, GAA, GALK1, GAPDH, GAPDH, GBE1, GCK, GCNT2, GLRX, GPC1, GPC3, GPC4, GPI, GRHRP, GYS1, HAS1, HDLBP, HEXA, HK1, HK2, HMOX1, HOXB9, HS3ST1, HSPA5, IDS, IER3, IGFBP1, IGFBP3, IL6, ILVBL, INHA, IRS2, ISG20, JMJD6, JUN, KDEL3, KDM3A, KIF5A, KLF6, KLF7, KLHL24, LALBA, LARGE1, LDHA, LDHC, LOX, LXN, MAFF, MAP3K1, MIF, MT1E, MT2A, MXI1, MYH9, NAGK, NCAN, NDRG1, NDST1, NDST2, NEDD4L, NFIL3, NOCT, NR3C1, P4HA1, P4HA2, PAM, PCK1, PDGFB, PDK1, PDK3, PFKFB3, PFKL, PFKP, PGAM2, PGF, PGK1, PGM1, PGM2, PHKG1, PIM1, PKLR, PKP1, PLAC8, PLAUR, PLIN2, PNRC1, PPARGC1A, PPPIA4, PPP1R15A, PPP1R3C, PRDX5, PRKCA, PYGM, RBPJ, RORA, RRAGD, S100A4, SAP30, SCARB1, SDC2, SDC3, SDC4, SELENBP1, SERPINE1, SIAH2, SLC25A1, SLC2A1, SLC2A3, SLC2A5, SLC37A4, SLC6A6, SRPX, STBD1, STC1, STC2, SULT2B1, TES, TGFB3, TGFB1, TGM2, TIPARP, TKTL1, TMEM45A, TNFAIP3, TPBG, TPD52, TPI1, TPST2, UGP2, VEGFA, VHL, VLDLR, WSB1, XPNPEP1, ZFP36, ZNF292.

As both fibrosis and necrosis are linked to hypoxia<sup>96–98</sup> which may favor bacterial colonisation, we explored the association between bacterial abundance and hypoxia in the bulk cohorts. Within tumors, higher bacterial abundance associated with enriched hypoxia signatures<sup>94,95</sup> (TxRenal) (Sup.Figures S11A and S11B), consistent with local hypoxia within bacteria-associated necrotic and stromal niches. Across tumors however (TCGA), hypoxia signatures showed no consistent enrichment in “Cuti-High/Mid” relative to “Cuti-Low” tumors, with different signatures showing opposing directional trends (Sup.Figure S11A and S11B), arguing against global tumor hypoxia as a primary driver of colonisation. However, given the pseudo-hypoxic state of ccRCC tumors induced by the foundation loss of *VHL*, these transcriptomic signatures may not uniquely capture truly oxygen deficient regions, limiting inferences in the association between bacterial presence and hypoxia.

### Taxonomy identification of bacterial cultures

The taxonomy of each whole genome sequenced colony was determined using taxonomic assignments from the Type Strain Genome Server (TYGS).<sup>99</sup> Taxonomic assignments of *Cutibacterium acnes* genomes were in agreement with average nucleotide identity,<sup>100</sup> comparisons to the *C. acnes* type strain genome (NCBI CP025934.1) (ATCC 6919 = DSM1897), genome size and GC content. Genomically identical *Cutibacterium acnes* strains were identified based on an average nucleotide identity (ANI) greater than 99.9. Phylotype was assigned based on groupings observed in a reference-free whole-genome phylogeny generated using kSNP4<sup>75</sup> (kmer = 17, FCK = 0.653) with reference genomes from each common phylotype (Type IA: DSM1897 - Taxonomy ID: 1122995; Type IB: HL059PA1 - Taxonomy ID: 765091; Type II: HL001PA1 - Taxonomy ID: 765105).

### RNAscope image processing

*Cutibacterium* presence in tumor FISH images was determined based on *Cutibacterium* 16S and Dap-B probe intensity (15/255). We found that most positions where the two probes overlapped, coincided with very high *Cutibacterium* signal (>200) but low Dap-B signal (~15–30), stemming from autofluorescence in the 520nm channel where Dap-B was positioned (Sup.Figures S9C and S9D). On average across the samples, less than 3% of *Cutibacterium*-positive pixels (a pixel is ~0.1µm) overlapped with a Dap-B pixel that had intensities greater than two-thirds that of the *Cutibacterium* pixel. These were all removed from further analysis. In these areas of high density *Cutibacterium* probe signal, at whole-slide scanning resolution, individual RNA probe spots are not optically resolved and therefore appear as a confluent signal as seen in Sup.Figures S9C and S9D). To further ascertain true bacterial staining by the *Cutibacterium* probe in these regions, we increased the DAPI intensity in our confocal images of the fluorescent stains. This revealed presence of DNA outside of the nucleus overlapping with *Cutibacterium* signal as expected for intact bacteria. This extra-nuclear DNA signal associated with *Cutibacterium* through the z stack (SupFig.9.E-F), further supporting true bacterial presence in these samples.

To quantify bacterial abundance across a sample, bacterial “spots” were defined as any cluster of at least 50 pixels (~5 µm) that are within 5 pixels of one another.

### Macroscopic H&E sample annotations

The areas marked as stroma and necrosis on VisiumHD samples were annotated on the H&E image by a pathologist.

### Cell type classification from VisiumHD

Alignment, tissue and fiducial detection, and barcode/UMI counting were performed with space ranger using standard parameters (10× Genomics Space Ranger v3.0.0). Nuclei were segmented from the H&Es using the stardist model<sup>76</sup> in QuPath.<sup>77</sup> Reads from VisiumHD's 2 μm × 2 μm bins are assigned into the nuclear mask using Sopa<sup>78</sup> with a nuclear expansion of 0.25–0.75× the nuclear radius to achieve comparable median cell area for each sample. The resulting pseudo-single cell data was filtered to remove any cells with fewer than 10 counts and any genes assigned to fewer than 10 cells. To minimise misclustering driven by highly expressed but non-cell type-specific genes, which can be exacerbated by VisiumHD's low per-cell transcript capture, we excluded mitochondrial genes and *MALAT1* from analysis. Read counts were normalised by each cell's total counts and log transformed for differential analyses.

For each sample, the pseudo-single cells were clustered with the leiden algorithm, followed by inspection of the top 25 differentially expressed genes and expression of lineage marker genes in each cluster for cell type assignment. Clusters with overlapping marker gene expression were re-clustered to improve resolution where possible (Sup. Figures S1–S4). Of note, cell clusters with low total read count per cell, consequently showed frequent lineage marker gene dropout but could still be classified based on differential gene expression. Low RNA clusters which overlapped or were adjacent to the main cluster of the same cell type on the umap and differentially expressed the same lineage markers were combined with the main cluster. Clusters where no distinctive lineage markers could be identified were marked as “Undefined\_low\_RNA”.

One limitation of the VisiumHD method in cell type assignment is that, as transcripts are captured by fixed probe bins, transcripts from closely adjacent cells can be captured by the same probe bin. Additionally, as expected from fresh frozen input tissue, nuclear borders in the H&E were less sharp than they would be for fixed samples reducing our ability to disambiguate some closely adjacent nuclei during segmentation.<sup>101</sup> Consequently, we observed some clusters with cells expressing lineage markers of distinct cell types which we labeled as “mixed”. For K239\_H and K988\_L, where neutrophil markers did not appear in the top 25 differentially expressed genes of any cluster, all significantly differentially expressed genes with log fold change greater than 1 were considered to identify the clusters harboring neutrophils.

### Bacterial proximity analyses from VisiumHD data

Bacterial coordinates were mapped onto the VisiumHD data by aligning the serial RNAscope image with the H&E and applying the corresponding affine transform to the RNAscope coordinates. For K153\_H, the tissue had adopted a slightly distorted shape due to a tear. The alignment was therefore done in three sections, individually aligning and transforming the top, middle, and bottom of the RNAscope image.

For each cell, distance from the closest bacterium was then calculated using the cKDTree function from the SciPy python package.<sup>79</sup> Cells were categorised as “near” bacteria if within 25 μm of the closest bacterium, “mid” from bacteria if within 100 μm of the closest bacterium and “far” if beyond 100 μm.

For each cell type, the proportion of near, mid and far cells the cell type makes up in each sample was then compared to identify conserved trends relative to bacterial proximity. Significant difference in near vs. far proportions was assessed using a two-sided Fisher's exact test with FDR correction across cell types and samples.

In Neutrophil-containing mixed clusters, the ratio of cells expressing neutrophil markers (*CSF3R* or *FCGR3B*) over cells expressing monocyte and macrophage markers (*CD68* or *CD14*) was calculated far, mid and near from bacteria to confirm neutrophil enrichment near bacteria within the mixed clusters.

### Differential gene expression analyses from VisiumHD data

To confirm the chosen samples recapitulated gene expression dynamics seen in the bulk cohort, gene expression across all cells of the three “Cuti-High” samples and of the “Cuti-Low” sample were compared. Differentially expressed genes were identified using ScanPy's rank\_genes\_groups\_df function. Only genes present in all four samples were considered in this analysis. Pathway enrichment of the differentially expressed genes was determined using gseapy's<sup>81</sup> prerank function. The top 10 significantly enriched pathways with more than 5 genes contributing to the enrichment were shown.

Due to the high sparsity of VisiumHD data at cell-level resolution, arising from the small capture area per bin relative to whole-cell transcriptome coverage, when assessing expression of a small set of specific genes (Neutrophil-recruiting chemokines: *CXCL1*, *CXCL2*, *CXCL3*, *CXCL6*, *CXCL8*; Fc receptors: *FCGR1A*, *FCGR2A*, *FCGR3A*, *FCAR*, *FCGR2B*) expression was considered only in cells expressing at least one gene in the set (sum of non-log-transformed, read-depth normalised counts). Box widths in the boxplots represent the proportion of cells in a given sample or proximity compartment expressing at least one marker gene, normalised by the proportion expressing at least one lineage marker (*CD68* or *CD14* for macrophages), to correct for differential transcriptome capture efficiency and cell type identity across samples or proximity compartments. Differences in expression scores between groups were assessed using a two-sided permutation test (10,000 permutations, test statistic: difference in medians, random seed 0). The *p*-value was computed as the proportion of permuted statistics with absolute value greater than or equal to the observed statistic, with a +1 correction applied to numerator and denominator, and FDR corrected across samples

and proximity comparisons. A permutation test was used as a distribution-free alternative to parametric tests, given that expression score distributions are skewed and non-normal.

To assess whether bacteria-proximal chemokine upregulation in the K239\_H macrophage cluster was attributable to neutrophil doublet contamination, a one-sided binomial test was performed on the co-detection of chemokine and neutrophil lineage marker genes within the same cell. The null dropout rate for neutrophil marker genes was estimated using a stably expressed housekeeping gene (*EEF2*) detection rate as a proxy for general transcriptome capture efficiency in this cluster. The observed proportion of chemokine-positive, neutrophil-marker-negative cells (149/155) significantly exceeded the null expectation (expected 136.2;  $p = 0.0003$ ). Therefore, the increase in chemokine expression with proximity to bacteria can confidently be attributed to macrophages, rather than stemming solely from neutrophil doublets.

### qPCR cytokine profiles from co-cultures

For each condition, the mean qPCR cycle number of the technical repeats for each of the cytokines was normalised by the human reference (beta-actin) and the log2 fold change relative to the time-matched, untreated macrophage control calculated.

### Summary of evidence supporting true *Cutibacterium* enrichment in ccRCC primary tumors

Given that sample collection was initiated prior to the start of this microbiome study, matched environmental controls per patient at the time of surgery were not collected. Nonetheless, we have taken a rigorous decontamination approach which addresses the two main areas of contamination risk.

- (1) Surgical environment: any bacterial genera for which more than 5% of variance across the cohort was explained by hospital site were fully removed from consideration, thereby removing bacteria enriched at specific hospitals. Moreover, samples were obtained from whole tumor resections rather than needle biopsies, where contamination from patient skin would be considerably more likely. From these resections, samples were taken from a central tumor slice which would not have been exposed to the surgical environment, mitigating the risk of bacteria introduced at the time of surgery. Critically, we observe significant enrichment of *Cutibacterium* in the center of these central slices (Figure 2B), providing orthogonal evidence that these bacteria were present in the tumor prior to surgery as surgical contaminants would be expected to accumulate at the tumor surface exposed during surgery, which is the opposite of what we observe. We also draw attention to one of the two Cuti-Low samples cultured where no *Cutibacterium* colonies were grown, demonstrating that the genus is not ubiquitously present across samples as may be expected if added from host skin contamination.
- (2) Laboratory process: contaminants present throughout the laboratory process were characterised using 16S sequencing of negative controls (kit reagents, extraction blanks), and all genera identified were removed from tumor WGS data. Additionally, any bacterial genera for which more than 5% of variance across the cohort was explained by sequencing batch were fully removed from consideration, and commonly reported laboratory contaminants identified in the literature were removed unless previously identified as present in ccRCC. We also confirmed reproducible *Cutibacterium* abundance across 6 samples which were re-sequenced 3 months after initial sequencing (Figure 1E) confirming lack of bacterial DNA contamination in the processing, shipping and sequencing of samples.
- (3) Across processes: we removed all reads associated with any genus flagged as a potential contaminant. This is a conservative approach that may remove some true signal where genuine reads do not considerably outweigh contaminant reads. However, we deliberately opted for this conservative approach to minimise the risk of retaining contaminants. Further, we only considered genera with significant enrichment in tumor relative to matched patient blood obtained at the time of surgery, ensuring that only genera which have proliferated in the tumor are considered.

Finally, our results are corroborated by a recent independent study<sup>35</sup> designed to differentiate true from contaminant *Cutibacterium* in tumor samples, which found high-confidence truly present *Cutibacterium* reads in over 88% of TCGA WGS ccRCC samples.