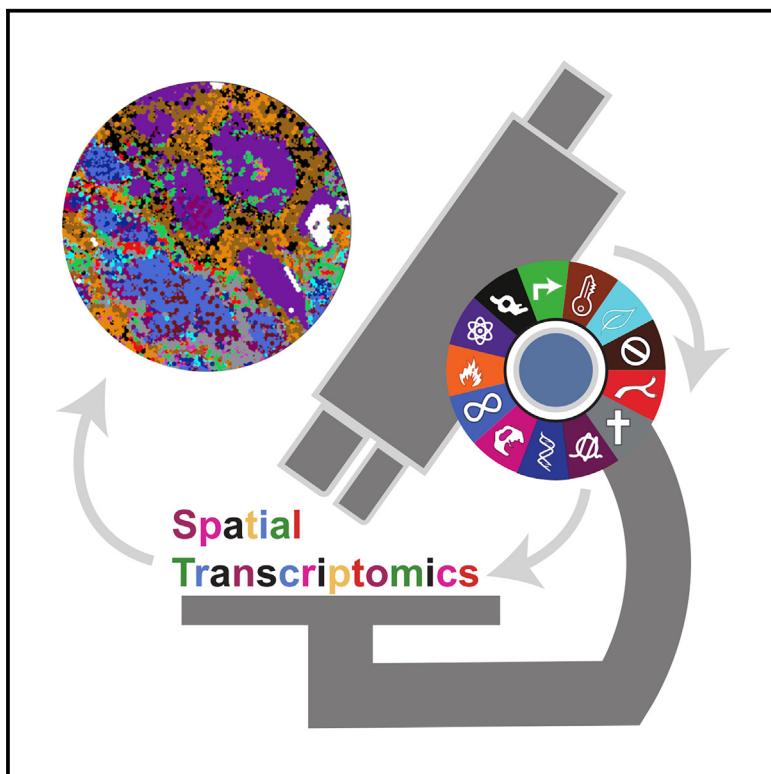


## The spatial landscape of cancer hallmarks reveals patterns of tumor ecological dynamics and drug sensitivity

### Graphical abstract



### Authors

Mustafa Sibai, Sergi Cervilla,  
Daniela Grases, ..., Manel Esteller,  
Matthew H. Bailey, Eduard Porta-Pardo

### Correspondence

eporta@carrerasresearch.org

### In brief

Sibai et al. investigate the spatial organization of cancer hallmarks in primary untreated tumors, revealing compartment-specific contributions of cancer cells and the tumor microenvironment to hallmark activity. This approach highlights how genomic and spatial factors influence hallmark specialization and interdependencies, providing insights into tumor ecology and drug sensitivity that could inform therapeutic strategies.

### Highlights

- Cancer cells drive seven out of thirteen cancer hallmarks, while the TME governs the rest
- Genomic distance correlates with hallmark activity, leading to clone specialization
- Hallmarks interrelate at TME-cancer junctions, influencing drug response
- Spatial hallmark patterns offer biomarkers for neoadjuvant treatment efficacy



## Article

# The spatial landscape of cancer hallmarks reveals patterns of tumor ecological dynamics and drug sensitivity

Mustafa Sibai,<sup>1,2,15</sup> Sergi Cervilla,<sup>1,15</sup> Daniela Grases,<sup>1,15</sup> Eva Musulen,<sup>1,3</sup> Rossana Lazcano,<sup>4</sup> Chia-Kuei Mo,<sup>5</sup> Veronica Dávalos,<sup>1</sup> Arola Fortian,<sup>6</sup> Adrià Bernat,<sup>6</sup> Margarita Romeo,<sup>6</sup> Collin Tokheim,<sup>7</sup> Jordi Barretina,<sup>6</sup> Alexander J. Lazar,<sup>4</sup> Li Ding,<sup>5</sup> the DUTRENEO Study Investigators, Enrique Grande,<sup>8</sup> Francisco X. Real,<sup>9,10,11</sup> Manel Esteller,<sup>1,10,12,13,16</sup> Matthew H. Bailey,<sup>14,16</sup> and Eduard Porta-Pardo<sup>1,2,16,17,\*</sup>

<sup>1</sup> Josep Carreras Leukaemia Research Institute (IJC), Badalona, Spain

<sup>2</sup> Barcelona Supercomputing Center (BSC), Barcelona, Spain

<sup>3</sup> Department of Pathology, Hospital Universitari General de Catalunya Grupo-QuirónSalud, Sant Cugat del Vallès, Spain

<sup>4</sup> The University of Texas MD Anderson Cancer Center, Houston, TX, United States

<sup>5</sup> Department of Medicine and Genetics, Siteman Cancer Center, Washington University in St. Louis, St. Louis, MO, USA

<sup>6</sup> Institut de Recerca Germans Trias i Pujol (IGTP), Badalona, Spain

<sup>7</sup> Department of Data Science, Dana-Farber Cancer Institute, Boston, MA, USA

<sup>8</sup> Medical Oncology Department, MD Anderson Cancer Center Madrid, Madrid, Spain

<sup>9</sup> Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid, Spain

<sup>10</sup> Centro de Investigación Biomédica en Red Cancer (CIBERONC), Madrid, Spain

<sup>11</sup> Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Barcelona, Spain

<sup>12</sup> Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

<sup>13</sup> Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona, Barcelona, Catalonia, Spain

<sup>14</sup> Department of Biology and Simmons Center for Cancer Research, Brigham Young University, Provo, UT, USA

<sup>15</sup> These authors contributed equally

<sup>16</sup> Senior author

<sup>17</sup> Lead contact

\*Correspondence: [eporta@carrerasresearch.org](mailto:eporta@carrerasresearch.org)

<https://doi.org/10.1016/j.celrep.2024.115229>

## SUMMARY

Tumors are complex ecosystems of interacting cell types. The concept of cancer hallmarks distills this complexity into underlying principles that govern tumor growth. Here, we explore the spatial distribution of cancer hallmarks across 63 primary untreated tumors from 10 cancer types using spatial transcriptomics. We show that hallmark activity is spatially organized, with the cancer compartment contributing to the activity of seven out of 13 hallmarks, while the tumor microenvironment (TME) contributes to the activity of the rest. Additionally, we discover that genomic distance between tumor subclones correlates with differences in hallmark activity, even leading to clone-hallmark specialization. Finally, we demonstrate interdependent relationships between hallmarks at the junctions of TME and cancer compartments and how they relate to sensitivity to different neoadjuvant treatments in 33 bladder cancer patients from the DUTRENEO trial. In conclusion, our findings may improve our understanding of tumor ecology and help identify new drug biomarkers.

## INTRODUCTION

Over the last decade, we have witnessed an increasing array of discoveries describing the molecular origins of oncogenesis. Among others, we now have catalogs of cancer driver genes and mutations,<sup>1</sup> their interactions with the immune system,<sup>2</sup> or new clinically relevant gene expression profiles.<sup>3</sup> However, the specific molecular changes driving oncogenesis depend on a myriad of variables, such as the specific cancer type,<sup>4</sup> the biological sex of the patient,<sup>5</sup> or the patient's germline background.<sup>6</sup> In the context of this endless complexity,<sup>7</sup> cancer hallmarks emerge as a powerful integrative approach to rationalize these mechanisms, distilling the legion of molecular aberrations

observed in a tumor into logical principles describing the necessary features for a tumor to become malignant.<sup>8–10</sup>

Cancer hallmarks have been applied to rationalize diverse types of omics datasets.<sup>11,12</sup> The typical process to identify which hallmarks are active in each tumor starts by associating them with specific gene expression signatures.<sup>13,14</sup> However, until recently, most omics datasets were generated from the profiling of bulk tumors, limiting our understanding of the interplay between the two major tumor compartments: cancer cells and their surrounding tumor microenvironment (TME). Despite the difficulty in disentangling the cellular origin of cancer hallmarks, the importance of the TME in studying them has been evident for years.<sup>9</sup>



The analysis of single-cell transcriptomics from human tumors revealed an unforeseen richness of cell types and states in the TME, including different immune, stromal, or fibroblast cells.<sup>15</sup> Cancer cells themselves are diverse and complex. In most tumors, there are many types of cancer cells, each showing different gene activity patterns. These cells may vary in how developed they are or in their reactions to conditions such as low oxygen (hypoxia) or immune signals (such as interferon).<sup>16</sup> It is worth noting that, just as different cell types oftentimes only do a subset of all functions carried out by the organ where they are located, this transcriptional heterogeneity opens the possibility that the phenotypes described by cancer hallmarks are not active in all individual cancer cells but emerge in the whole tumor through the cooperation of multiple cell populations, each one of them specialized in certain phenotypes.<sup>16</sup>

While single-cell experiments can identify and quantify this cellular heterogeneity, they destroy the tissue structure. Spatial transcriptomics, on the other hand, allows us to explore this transcriptional heterogeneity while keeping the physiological context of the cells.<sup>17</sup> Among others, spatial transcriptomics has revealed that the area between the tumor and healthy tissue contains a specialized subset of cancer cells,<sup>18</sup> the location of niches of cancer cells with specific functions,<sup>19</sup> that the tumor edges of renal cell carcinomas are enriched in cancer cells undergoing epithelial-mesenchymal transition,<sup>18,20</sup> or the different spatial structures (or lack thereof) of glioblastoma cells and their surrounding TME.<sup>21</sup>

Here, we quantified how the activity of cancer hallmarks in 63 tumor samples from 10 cancer types is distributed through space using spatial transcriptomics. Our findings demonstrate that the distribution of cancer hallmark activities is not random but instead follows a spatial pattern. Cell identity plays a crucial role in determining the activity of these hallmarks, with some showing consistently higher activity in cancer cells while others are more active in the TME. Somatic copy number alterations have a limited impact on the variability of cancer hallmark activities in most cancers but, given sufficient genomic heterogeneity, some clones specialize in specific cancer hallmarks. We also show how the architecture of cancer hallmarks across the entire tumor is largely shaped by the spatial interdependence between the phenotypes of cancer cells and those of the TME, leading to a conserved pancancer tumor architecture and revealing complex ecosystem dynamics. Finally, differences in this spatial interdependence between cancer hallmarks correlate with outcomes in 33 muscle-invasive bladder cancer patients from the DUTRENEO trial,<sup>22</sup> showing how the spatial organization of the tumor ecosystem dictates its sensitivity to different treatments.

## RESULTS

### Cancer hallmarks are spatially organized

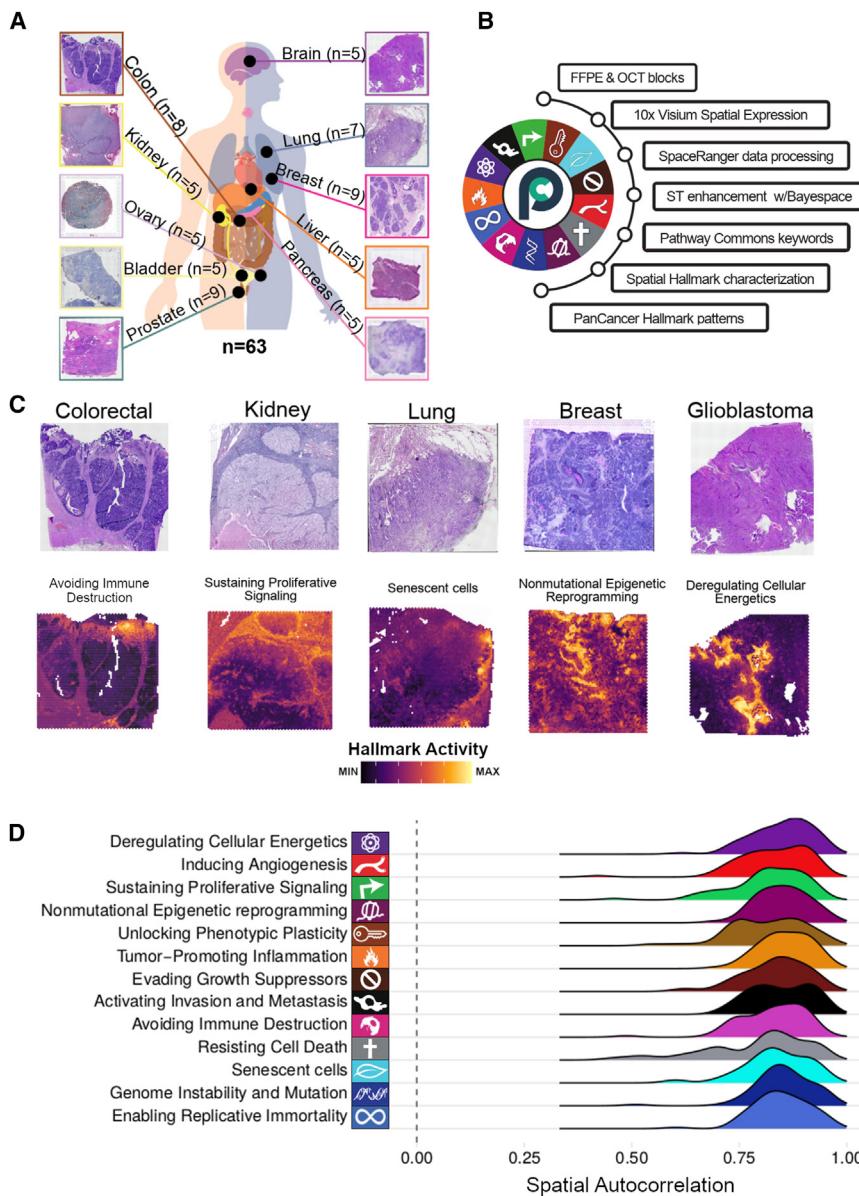
We have assembled a pancancer spatial transcriptomics (VISIUM<sup>23</sup>) cohort, with 63 primary untreated tumor samples from 10 cancer types: breast ( $n = 9$ ), prostate ( $n = 9$ ), lung ( $n = 7$ ), brain ( $n = 5$ ), colorectal ( $n = 8$ ), ovary ( $n = 5$ ), bladder ( $n = 5$ ), liver ( $n = 5$ ), pancreas ( $n = 5$ ), and kidney ( $n = 5$ ). Of these, 32 samples were collected from public repositories, while we generated the data for the remaining 31 samples (Figure 1A; Table S1, Section A). The spatial transcriptomics data from all

samples had sufficient high-quality data (Table S1, Section A). Overall, our cohort is representative of the most common cancer types.

To quantify the activity of different cancer hallmarks, we created a transcriptional signature describing each hallmark (Figures 1B and 1C). We manually curated pathways from the Pathway Commons database to generate gene sets associated with 13 cancer hallmarks (all except “Polymorphic Microbiomes”; STAR Methods; Figure S1A). This ensured that the association of pathways to each cancer hallmark was based on the collective knowledge of the cancer research community. We established these associations taking into account a trade-off between the biological intertwining that cancer hallmarks have with one another and their individual roles. For example, to quantify the activity of the “Nonmutational epigenetic reprogramming” hallmark, we created a gene set with pathways such as “HDACs deacetylate histones,” “Chromatin modifying enzymes,” or “HATs acetylate histones.” Individually, each of these pathways only contributes to the hallmark activity in a subset of the samples (Figure S1B), but, when integrated in a single signature, we can quantify the hallmark across all samples from all cancer types. The exact pathways and genes included in each cancer hallmark signature are described in Table S1, Section B.

The number of pathways associated with each hallmark ranged between four and 15 (Figure S1C), and the number of genes per hallmark ranged between 150 and 600 (Figure S1C). In total, the signatures of all 13 hallmarks have 2,699 unique genes, and 62% of them are associated with a single hallmark (Figure S1D). This balances the specificity of each gene set while accounting for the biological intertwining across hallmarks. Finally, we scored the activity of each hallmark’s gene set in the different samples. We also computationally enhanced our spatial transcriptomics data with BayesSpace,<sup>24</sup> which splits each VISIUM spot (which has a ~55-μm diameter) into six subspots (each ~20 μm in diameter). After scaling and centering these scores within each sample, the final score is what we refer to as the “hallmark activity” for the rest of the manuscript.

To make sure the gene lists for each cancer hallmark are reliable and measurable in all samples, we checked how many genes from each hallmark were included in the VISIUM data. On average, over 80% of the genes for each hallmark were found across all samples (Figure S1E). We tested how stable the cancer hallmark signatures are under two conditions: missing key genes (to see the impact of excluding important genes) and adding similar but possibly unnecessary genes (from related pathways). For the first test, we compared the original cancer hallmark scores with scores from modified gene sets, where some genes were randomly left out from each hallmark pathway (STAR Methods; Figure S2A). The number of excluded genes ranged between 25% and 50% per pathway and, for each fraction that we tested, we repeated the experiment five times. For the second control, we correlated the original hallmark scores with a number of control gene sets that were each composed of the original gene signature plus a number of unique genes from potentially important pathways for the tested hallmark (STAR Methods). Both controls showed the robustness of our signatures (Figure S2B).



Confident in our method for measuring cancer hallmark activity, we analyzed it in the spatial transcriptomics data. We found that the activity of all 13 cancer hallmarks is not random but is concentrated in specific tumor areas. To quantify this phenomenon, we calculated the spatial autocorrelation of each hallmark by computing the Pearson correlation between the original hallmark activity vector and its spatial lag vector.<sup>25</sup> This statistic ranges between  $-1$  and  $1$ , with features repelling each other in space tending to  $-1$ , those with random distribution across space tending to  $0$ , and those with clustering in space tending to  $1$ . The average spatial autocorrelation of all hallmarks was around  $0.8$ , confirming that cancer hallmarks are spatially clustered (Figure 1D). The spatial autocorrelation of each cancer hallmark was also higher than that of random gene sets across all

**Figure 1. Quantifying the activity of cancer hallmarks across space**

(A) Overview of the spatial transcriptomics cohort, showing the number of samples for each tumor type studied.

(B) Workflow to build cancer hallmark signatures.

(C) Comparison between H&E images (top) and the hallmark's activity (bottom) in five of the studied cancer types.

(D) Ridge plot showing the spatial autocorrelation distribution for each of the 13 cancer hallmarks across 63 samples. Higher values indicate non-random clustering, while lower values suggest random distribution within the tissue.

samples, both at the pancancer (Figure S3A) as well as the individual cancer type levels (Figure S3B).

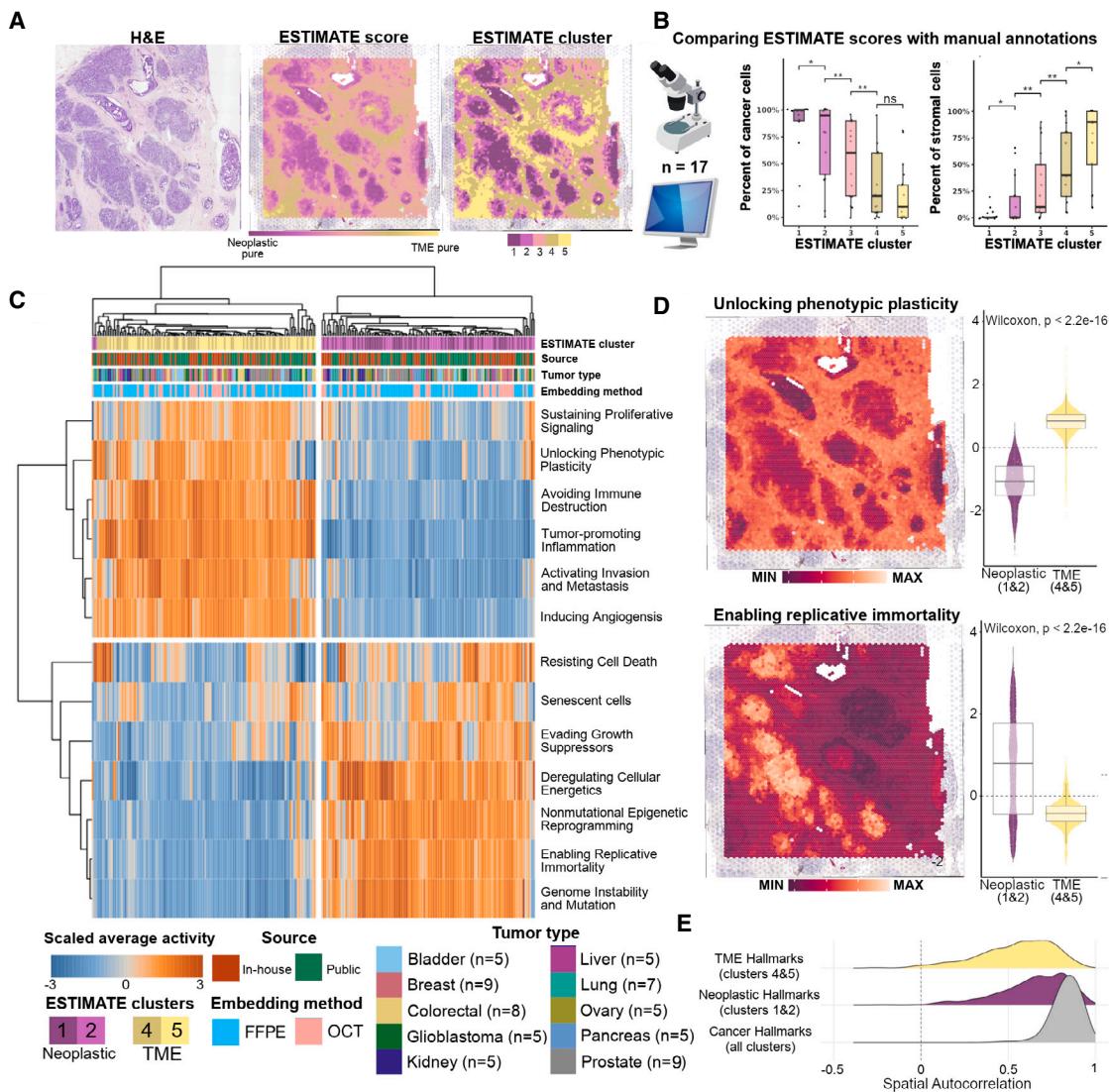
### The activity of cancer hallmarks follows a pancancer tumor architecture

One potential factor contributing to the spatial organization of cancer hallmarks is cell identity: cancer cells and the TME can potentially contribute differently to each hallmark. Most studies on cancer hallmarks have focused on cancer cells, but there is increasing recognition that cells in the TME may also contribute to some hallmarks.<sup>8</sup>

To explore this in detail, we classified each subspot in each tumor sample into one of the two major tumor compartments: cancer cells and the TME. We used ESTIMATE<sup>26</sup> to measure the number of cancer cells in each subspot of every sample (STAR Methods). Based on these scores, we grouped the subspots into five clusters: cluster 1 had a high number of cancer cells, while cluster 5 was mostly made up of TME cells (Figure 2A). We confirmed these groupings

with histopathological annotations from certified pathologists, who analyzed high-quality hematoxylin and eosin (H&E) images from 17 samples across five cancer types (Figures 2B and S4A; Table S1, Section C). The results show a strong match between ESTIMATE-based clusters and the percentage of cancer cells identified by pathologists. Cluster 1 spots contained over 90% cancer cells and less than 5% stromal cells on average, while cluster 5 spots had less than 10% cancer cells and over 85% stromal cells.

To further confirm ESTIMATE's accuracy, we used VISIUM CytAssist and Xenium data from serial sections of colorectal, lung adenocarcinoma, and ovarian adenocarcinoma samples. This allowed us to pinpoint cell types at single-cell resolution and compare them with VISIUM-based annotations. The findings show that ESTIMATE scores align well with epithelial cell



**Figure 2. The pancancer architecture of cancer hallmarks**

(A) H&E of a breast-invasive carcinoma sample (left) with the computed ESTIMATE scores for each individual subspot (center), ranging from pure neoplastic cell content in purple to pure TME content in yellow. The score is categorized into five clusters by applying k-means (right).

(B) Benchmark of ESTIMATE scores on 17 samples from five cancer types annotated by human pathologists to quantify the percentage of stromal and cancer cells within each cluster.

(C) Heatmap showing the scaled average cancer hallmark scores (rows) for each sample cluster (columns). Annotations include (1) ESTIMATE clusters (1 and 2 for neoplastic compartments, 4 and 5 for TME compartments), (2) sample source (public or in-house), (3) tumor type, and (4) tissue-embedding method.

(D) Distribution of hallmark scores in neoplastic (clusters 1 and 2) and TME (clusters 4 and 5) compartments, along with their spatial patterns in the breast-invasive carcinoma sample, for “Unlocking phenotypic plasticity” (top) and “Enabling replicative immortality” (bottom).

(E) Spatial autocorrelation distribution for all hallmarks across each sample, showing the seven hallmarks with higher activity in neoplastic compartments (clusters 1 and 2) and the six hallmarks with higher activity in TME compartments (clusters 4 and 5).

locations identified in the Xenium data, supporting our approach to mapping cancer cells (Figures S4B–S4D). Finally, we analyzed the average cancer hallmark activity in different ESTIMATE clusters across all tumor samples and types. The results reveal a clear pancancer pattern, with hallmark activities grouping by tumor compartments (neoplastic or TME,  $p < 0.0001$ , Fisher’s exact test) rather than cancer type ( $p > 0.9$ ), tissue preparation method ( $p > 0.9$ ), or data source

( $p > 0.8$ ). This highlights the key role of cell identity in shaping cancer hallmark spatial patterns.

Seven hallmarks showed higher activity in ESTIMATE clusters 1 and 2 (cancer cell-dominant regions; Figure 2C): “Evading growth suppressors,” “Enabling replicative immortality,” “Deregulating cellular energetics,” “Senescent cells,” “Nonmutational epigenetic reprogramming,” “Genomic instability and mutations,” and “resisting cell death.” Occasionally, hallmarks such as

"Senescence" and "Resisting cell death" were also active in TME areas, indicating senescent immune or stromal cells with survival mechanisms that may support cancer progression.<sup>9</sup> For instance, "Senescent cells" activity was particularly high in the TME of kidney, lung, and prostate adenocarcinomas, influenced by various immune cells (Figures S5A and S5B). The remaining six hallmarks were more active in ESTIMATE clusters 4 and 5, which represent the TME (Figure 2C). These include "Inducing angiogenesis," "Activating invasion and metastasis," and "Unlocking phenotypic plasticity," which help cancer cells modify the extracellular matrix and blood vessels to enable their spread and acquire stem-like properties through processes such as epithelial-to-mesenchymal transition (EMT). Two other hallmarks, "Avoiding immune destruction" and "Tumor-promoting inflammation," highlight the role of immune cells and cancer-secreted molecules, such as the senescence-associated secretory phenotype (SASP), in driving these activities within the TME. Lastly, "Sustaining proliferative signaling" was also more active in the TME, indicating that cancer cells rely on external signals such as mitogens from their surroundings to maintain their growth. Interestingly, a similar phenomenon has been observed where the TME promotes the proliferation of normal epithelial cells, eventually leading to malignant tumor formation.<sup>27</sup>

Unlike the tumor compartment, which mainly consists of cancer cells, the TME includes a mix of various cell types. These cell types and their states can differ between cancer types and even among patients with the same cancer. To determine which cell types contribute to the cancer hallmarks more active in the TME, we measured the presence of over 30 cell types in each sample by analyzing the expression of multiple well-known markers for each type (STAR Methods).

We first used the three samples with matching VISIUM and Xenium data to test the accuracy of this approach, showing that it provided relatively accurate estimates of cell type composition (Figures S5C and S5D). Importantly, there is little overlap between the genes used to quantify the activity of the different hallmarks and those used to identify different cell types (Figure S6A). We analyzed the correlation between the activity of each of the six TME hallmarks and the quantified signatures of different cell types to determine which cells contribute to each hallmark (Figure S6B).

Our analysis revealed that the activity of "Avoiding immune destruction" is strongly linked to immune cells such as T cells, T helper cells, B cells, and monocytes. "Tumor-promoting inflammation" is associated not only with immune cells but also with endothelial cells, fibroblasts, and pericytes. Similarly, "Activating invasion and metastasis" and "Unlocking phenotypic plasticity" are tied to fibroblasts and pericytes. "Inducing angiogenesis" correlates with endothelial cells, as expected, but also with immune cells, pericytes, and fibroblasts. In contrast, "Sustaining proliferative signaling" shows weak correlations with individual cell types, suggesting that multiple cell types work together to drive this activity. These relationships vary across tumor types, indicating that the same cell types may influence a hallmark in some tumors but not others (Figure S6C).

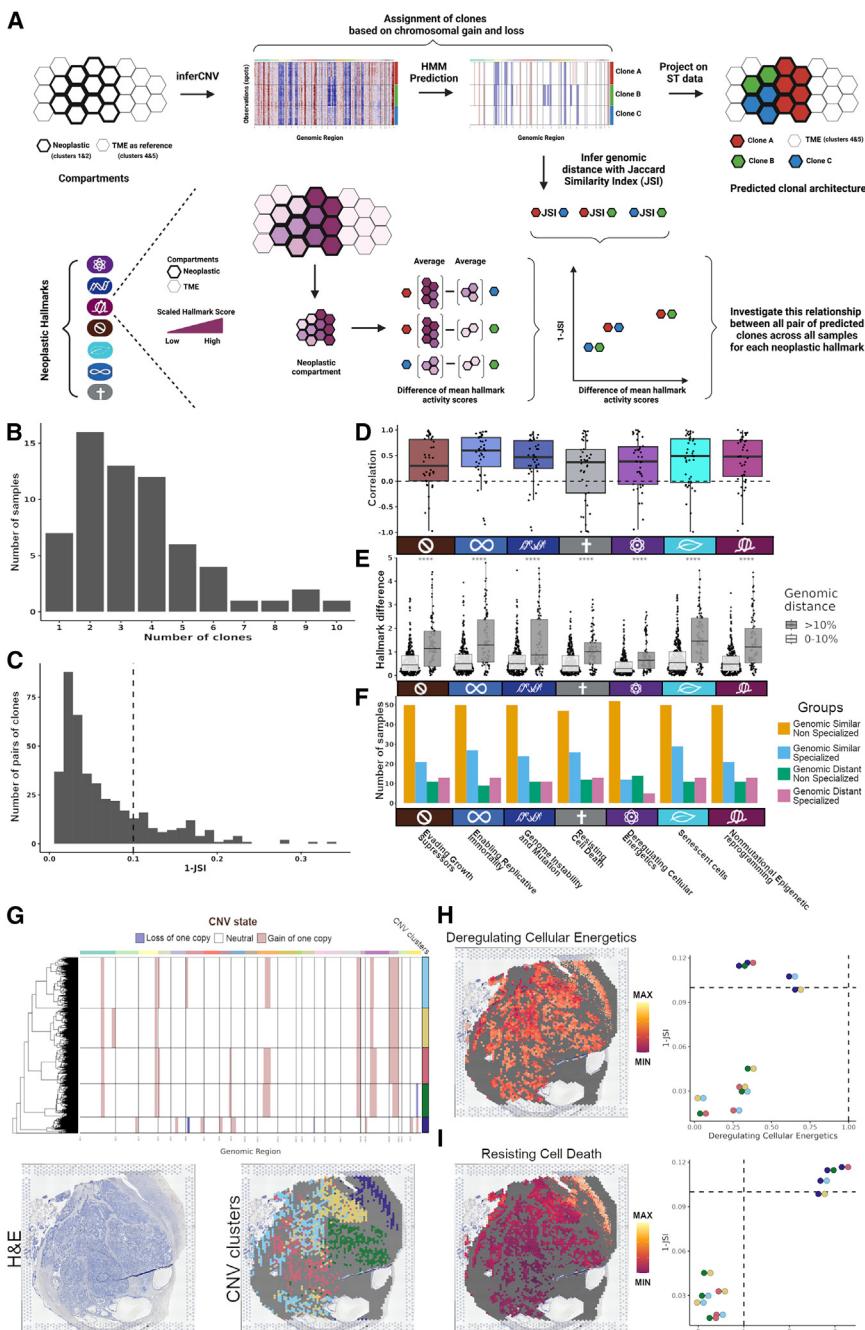
These findings suggest that many TME hallmarks arise from the combined actions of multiple cell types within specific tu-

mor regions rather than from individual cell types. Additionally, a single cell type can contribute to multiple cancer hallmarks, highlighting the complex, many-to-many relationships between cell types in the TME and cancer hallmarks. This approach demonstrates how the cancer hallmark framework can uncover pancancer patterns that might otherwise be hidden due to variations in cell type composition across different tumor types.

We investigated how cancer hallmark activities are spatially organized within the tumor's two compartments. Although cancer hallmark activity differs significantly between compartments (Figure S7A), there is notable variability among individual spots within the same compartment. For example, in a colorectal cancer sample, "Unlocking phenotypic plasticity" was highly active in 25% of TME spots, which were not randomly distributed but concentrated in specific tumor regions (Figure 2D, top). Similarly, while "Enabling replicative immortality" is generally more active in the neoplastic compartment, some neoplastic spots showed lower activity than certain TME spots (Figure 2D, bottom). A detailed analysis of all hallmarks' spatial distribution within each compartment revealed that their activity is strongly clustered rather than random. Across 63 samples, the average spatial autocorrelation was 0.64 for neoplastic hallmarks within neoplastic compartments and 0.55 for TME hallmarks within TME compartments (Figures 2E and S7B). Additionally, the spatial autocorrelation for all 13 hallmarks was significantly higher than for random gene signatures (Figure S7C). This suggests that, while cell identity (cancer or TME) strongly influences hallmark activity and spatial patterns, other factors also shape their internal spatial distribution within the TME and neoplastic compartments.

### The interplay between genetic and phenotypic variability throughout space

As cancer cells evolve, they can acquire somatic mutations that cause transcriptional and phenotypic differences between clones, including variations in cancer hallmark activity.<sup>28</sup> We used inferCNV<sup>29</sup> (STAR Methods; Figure S8A) to find tumor areas with different CNV states suggesting that they could be different genetic clones. Overall, we found 223 clones across the 63 samples, with most of the samples containing between two and four clones. Next, we calculated the genetic distance of each pair of clones in each sample. We did so based on the fraction of the genome that differs in CNV state between both clones (Figures 3A and 3B). The vast majority of the clone pairs have similar CNV states across the genome, with only 21% of them having a genomic distance greater than 10% (Figure 3C). To examine the link between genetic and phenotypic variation, we compared genomic distance with differences in cancer hallmark activity between clone pairs. There was a positive correlation, with Pearson coefficients ranging from 0.3 to 0.4 (Figure 3D). Grouping genomic distances into two categories (0%–10% and >10%) confirmed these findings, showing that genetically distant pairs had significantly larger differences in neoplastic hallmark activity on average (Figure 3E). Overall, these results show that genomic variation often correlates with changes in the gene expression patterns and the overall activity of cancer hallmarks in cancer cells.



**Figure 3. Relationship between cancer hallmark activity and genomic alterations**

(A) Diagram illustrating the process of using InferCNV to infer clone CNV status in a spatial transcriptomics sample, calculate genomic distances between clone pairs, and compare these to differences in neoplastic hallmark activity between clone pairs across all samples.

(B) Barplot showing the number clones per sample.

(C) Histogram of the distribution of the genomic distances between all pairs of clones in two bins: those having 0%-10% and >10% genomic distance.

(D) Boxplots showing the Pearson correlation coefficients from correlating the average difference in a neoplastic hallmark's activity and the genomic distance between pairs of clones within each sample.

(E) Boxplots comparing the difference in activities of neoplastic hallmarks between assigned categories of genomic distances.

(F) Barplots showing the number of samples in which the different genomic-phenotypic groups occur for each neoplastic hallmark.

(G) Top: Hidden Markov Model predictions of genomic states reported by inferCNV for an ovarian cancer sample whose H&E image is shown on the bottom left and the projected clones on the histology images on the bottom right.

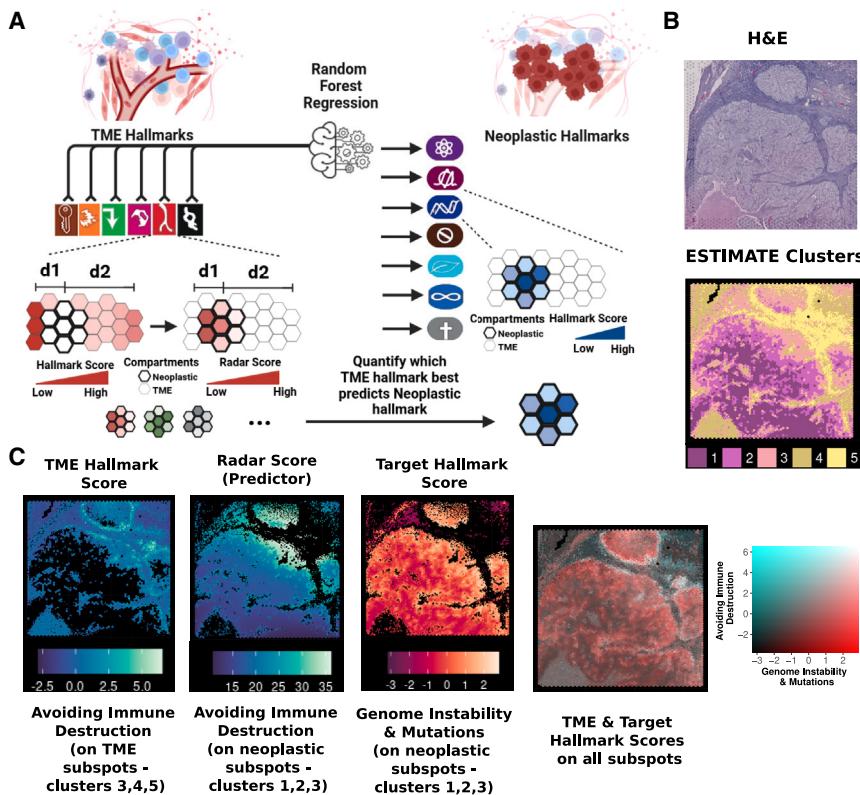
(H) Left: the scaled activity of "Deregulating cellular energetics" is overlaid on the neoplastic compartment in the ovarian cancer sample. Right: a scatterplot showing the difference of this hallmark's scaled activity and the genomic distance between the different pairs of clones.

(I) Same as (H) but for "Resisting cell death" hallmark.

We classified each pair of clones depending on their difference at the genomic and hallmark levels into one of four possible scenarios (Figure S8B). The first, and most common, one is where two clones with low genomic distance also have low average differences in their hallmark activities. We observed this in 265 pairs of clones across all samples (64% of the total pairs) and in at least one pair of clones in 89% of samples. The second scenario is that of high genomic distance and high average difference in hallmark activity, which could be considered genetic divergence leading to phenotypic specialization. Overall, this pattern was

rare across pairs of clones, as we only observed it in 51 pairs across all samples (12% of the total pairs). However, at the sample level, it was relatively common, as these 51 pairs of clones were spread across 23% of all samples.

The other two scenarios include the cases where the difference between a pair of clones in one of the variables is low but in the other variable is high. In this case, it is likely that genomics plays a less important role in determining the phenotype of the cancer cell. Instead, other factors such as cancer plasticity or the interplay between cancer cells and the TME through non-genetic cues are likely more relevant. For example, the third scenario includes pairs of clones that have differences in their hallmark activity, despite the fact that they are genetically similar ( $\leq 10\%$  of genomic distance), which happened in 35 pairs of clones (8%) across all samples, and in at least one pair of clones in 20% of samples. Finally, some clones have similar hallmark activities despite having high ( $\geq 30\%$ ) genomic distance (low-high). This scenario suggests phenotypic convergence, which we



**Figure 4. Describing the spatial interdependence between neoplastic and TME hallmarks**

(A) Illustration of using random forest regression to predict the spatial distribution of each neoplastic hallmark within the neoplastic compartment using the combination of the TME hallmarks from the TME compartment through their conversion into radar scores within the neoplastic compartment where the targets are located.

(B) An H&E image of a kidney cancer sample shown on the top, and its tumor compartments as defined by ESTIMATE's clusters are shown at the bottom.

(C) Example from a kidney tumor model where "Genome instability and mutation" is the target and "Avoiding immune destruction" is a predictor. From left to right: the scaled activity of "Avoiding immune destruction" in TME subspots (clusters 3–5); its radar score in the neoplastic compartment (clusters 1–3); the scaled activity of the target hallmark in neoplastic subspots (clusters 1–3); and the combined scaled activities of both hallmarks across the tissue section. Colocalized hallmark activities appear in white, while red shows the target hallmark without the TME hallmark, and cyan indicates the reverse.

observed in 62 pairs of clones across all samples (15% of the total pairs) and in at least one pair of clones in 38% of samples.

These scenarios can be further explored individually within each hallmark. Overall, when counting the number of samples having at least one pair in these scenarios, we find similar patterns across hallmarks except for "Deregulating cellular energetics" (Figure 3F). In this hallmark specifically, we find less presence of samples that have at least one clone pair that is specialized (with high hallmark difference) regardless of their genomic distance. One of the ovarian samples in which we inferred five different spatial constrained clones (Figure 3G) exhibited different scenarios depending on hallmark. For instance, "Deregulating cellular energetics" had a homogeneous pattern across the space without specialization to one or more clones (Figure 3H). However, for "Resisting cell death," we observe a clear specialization of this hallmark in one of the clones, which notably coincides with the most distinct CNV profile (Figure 3I). Scenarios that differ from the last one suggest that, while genomic evolution may play a role in shaping the spatial architecture of cancer hallmarks, by segregating cancer clones with specific functions (i.e., hallmarks) into certain areas of the tumor, there is likely a significant involvement of cell-extrinsic events that influence the spatial landscape of the neoplastic hallmarks, potentially through interactions with signaling cells from the TME.<sup>30</sup>

#### Quantifying the spatial interdependence of cancer hallmarks

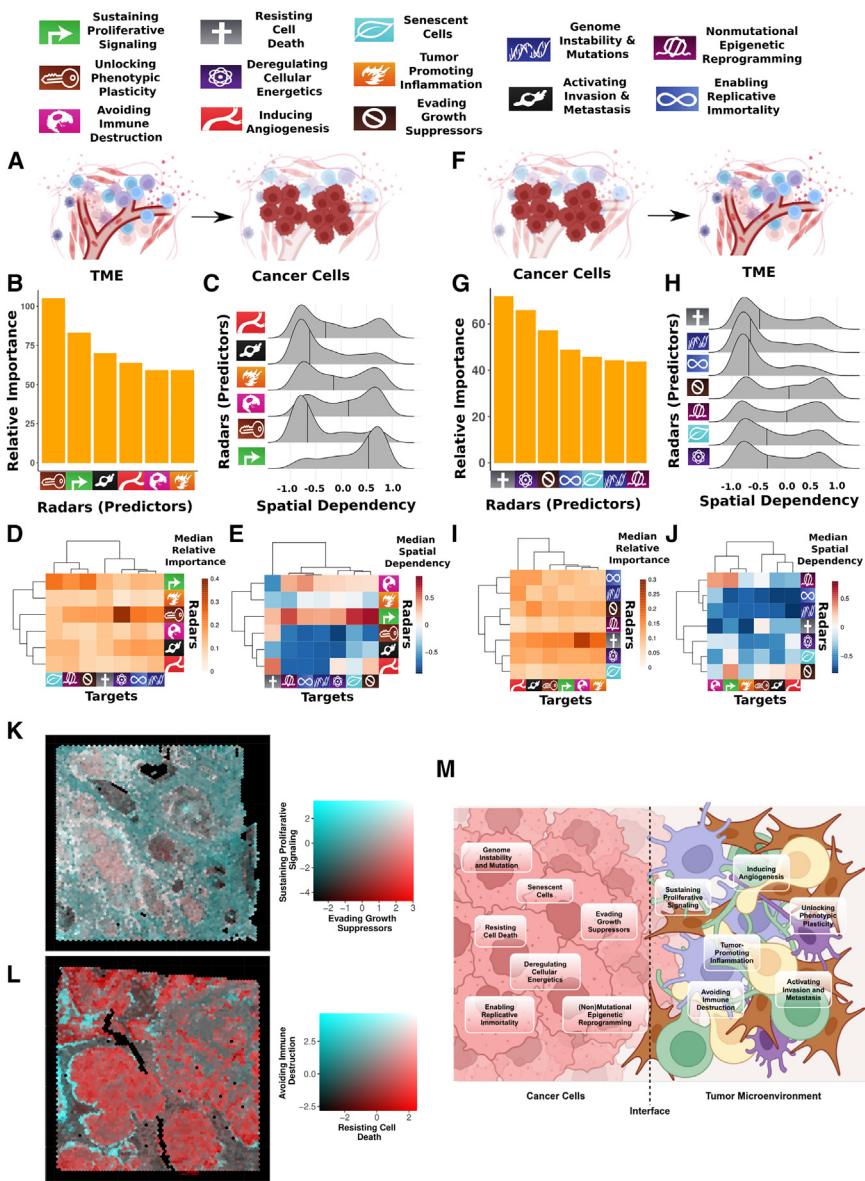
Another factor influencing the spatial organization of cancer hallmarks could be their interdependence—where the activity of one

hallmark is affected by the location of others. By using the phenotypic concept of cancer hallmarks, we can analyze this interdependence through an ecological

dynamics perspective.<sup>31</sup> In this view, the location of cells with specific phenotypes is shaped by the presence or absence of cells with other phenotypes. This kind of functional segregation is common in normal tissues, where different functions are assigned to specific cells in defined areas to optimize overall tissue performance. Similarly, recent studies show that cancer cells adopt different transcriptional programs and phenotypes depending on their position relative to the tumor's invasive front.<sup>32</sup>

To study the spatial interdependence of cancer hallmarks, we developed a random forest regression to model how the spatial patterns of TME hallmarks predict neoplastic hallmarks (Figure 4A) and vice versa. To incorporate the geometric distribution of hallmark activity, we transformed the activity of the predictor compartment into radar scores within the target compartment (STAR Methods). These radar scores represent how close each subspot in the target compartment is to subspots with high or low hallmark activity in the predictor compartment, accounting for spatial distance. For example, when modeling how TME hallmark activities influence those in cancer cells, we converted TME activity into radar scores for the cancer cell compartment (Figure 4A). After calculating radar scores for all hallmarks, we trained the model on 80% of subspots and used the remaining 20% to predict the target hallmark activity (response). This method is illustrated with "Avoiding immune destruction," where TME activity is used to generate radar scores in the neoplastic compartment to predict the location of the target hallmark "Genome instability and mutation" (Figures 4B and 4C).

Finally, we use Shapley additive explanations (SHAP) values to understand the model's predictions, focusing on



In the color legend, colocalized hallmarks at maximum activity are shown in white. Red indicates the predictor is active without the target, and black represents low activity in both hallmarks.

(L) The concept shown in (K) is shown here for the negative spatial dependency between the predictor ("Avoiding Immune Destruction") and its target ("Resisting Cell Death") in a hepatocellular carcinoma sample, with a coloring scheme similar to that in (K).

(M) Schema showing the median spatial location of cells highly active in each of the different hallmarks according to our pancancer data.

how relationships between radar features and the target hallmark were prioritized. This approach allows us to assess two critical aspects: feature importance and spatial dependency. Feature importance provides insights into which predictor hallmarks within one compartment are most influential in explaining the spatial distribution of a target hallmark in the other compartment. Spatial dependency, on the other hand, reveals the nature of these relationships. A positive spatial dependency indicates that the predictor and target hallmarks are frequently active in close proximity along their shared interface, while a negative spatial dependency sug-

gests they are generally active farther apart. Additionally, non-linear spatial dependencies point to patterns where predictor and target hallmarks are aligned in specific proximal or distal regions of the tumor. For instance, when modeling "Avoiding immune destruction" as a predictor for "Genome instability and mutation," we observed that these hallmarks tended to be active near each other, reflecting a positive spatial dependency. This approach can provide valuable insights into the spatial interplay between hallmarks across compartments and their ecological dynamics (Figure 4C).

**Figure 5. The spatial interdependence between cancer hallmarks**

(A) Icons summarizing the concept of the random forest regression experiment, where we are predicting the distribution of the hallmarks in cancer cells using the information of the location of the hallmarks from the TME.

(B) Barplot showing the sum of relative importances (y axis) of each predicting TME hallmark (x axis).

(C) Ridge plots of the Pearson correlation coefficients for the spatial dependency (calculated by correlating the predictor and its SHAP values as a proxy to its target) for all targets across all samples, where positive coefficients indicate a positive spatial dependency and negative coefficients indicate a negative spatial dependency.

(D) Heatmap clustering the predictor hallmarks (rows) according to their relative importance when predicting the target hallmarks (columns).

(E) Same as (D) but clustering according to the spatial dependency.

(F) Icons summarizing the concept of the random forest regression experiment, where we are predicting the distribution of the hallmarks in the TME using the information of the location of the hallmarks from the neoplastic compartment.

(G) Barplot showing the sum of relative importances (y axis) of each predicting neoplastic hallmark (x axis).

(H) Ridge plots of the Pearson correlation coefficients for the spatial dependency (calculated by correlating the predictor and its SHAP values as a proxy to its target) for all targets across all samples, where positive coefficients indicate a positive spatial dependency and negative coefficients indicate a negative spatial dependency.

(I) Heatmap clustering the predictor hallmarks (rows) according to their relative importance when predicting the target hallmarks (columns).

(J) Same as (I) but clustering according to the spatial dependency.

(K) Example from an invasive breast carcinoma showing a positive spatial dependency between the predictor ("Sustaining proliferative signaling") and the target ("Evading growth suppressors").

### The coordinated ecological dynamics of cancer hallmarks

We used our random forest modeling strategy to quantify the spatial interdependence between cancer hallmarks from both compartments across the entire cohort (Figure 5). The models from both experiments (predicting the spatial location of the neoplastic hallmarks using the location of the TME hallmarks, and vice versa) are all very accurate (mean  $R^2$  of 0.82 across all models; Figure S9A), suggesting that this approach can, indeed, capture the non-genetic cues driving the spatial organization of the cancer hallmarks of each compartment.

In the first experiment, we predicted the spatial distribution of neoplastic hallmarks using the TME activity (Figure 5A). The two main TME hallmarks influencing the spatial location of neoplastic hallmarks were "Unlocking phenotypic plasticity" and "Sustaining proliferative signaling" (Figure 5B). The contributions of other TME hallmarks were relatively similar, indicating that the spatial distribution of cancer traits typically arises from the combined influence of multiple TME hallmarks. TME hallmarks showed variability in how their activity depended on spatial proximity. For example, "Sustaining proliferative signaling" had a strong positive spatial dependency, meaning its activity predicted targets by proximity to high activity. On the other hand, "Unlocking phenotypic plasticity" and "Activating invasion and metastasis" exhibited strong negative spatial dependency, meaning that their activity in TME cells is lower near cancerous areas with high activity in neoplastic hallmarks (Figure 5C). Other predictors, such as "Avoiding immune destruction," "Tumor-promoting inflammation," and "Inducing angiogenesis," displayed no consistent pattern and influenced targets based on either proximity or distance (Figure 5C).

Examining the detailed relationships between specific hallmark pairs revealed additional insights into the ecological dynamics of tumor cells with different phenotypes, both in relative importance (Figure 5D) and spatial dependency (Figure 5E). For example, "Sustaining proliferative signaling" was important in predicting the spatial location of "Evading growth suppressors," "Senescent cells," and "Nonmutational epigenetic reprogramming" (Figure 5D), with a strong positive spatial dependency (Figure 5E), highlighting the positive selection of co-localization of cells with these functions from both compartments. Finally, we analyzed the random forest models for specific tumor types. We found no significant differences in model accuracy across cancer types for any target neoplastic hallmark (Figure S9B). However, the importance of some TME hallmarks varied in predicting their targets across different tumor types (Figure S9C). For instance, "Sustaining proliferative signaling" was the least important predictor in glioblastoma but the most important in breast cancer (Figures S9C–S9E).

In the second experiment, we predicted the distribution of TME hallmarks using the spatial distribution of neoplastic hallmarks (Figure 5F). In this case, we observed a clear gradient in terms of relative importance, showing that "Resisting cell death" is the most crucial neoplastic hallmark to predict the spatial distribution of the TME hallmarks (Figure 5G). This is followed by "Deregulating cellular energetics," "Evading growth suppressors," "Enabling replicative immortality," "Senescent cells," "Genomic instability and mutations," and "Nonmutational epige-

netic reprogramming." Additionally, the predictors varied in their spatial dependency with their targets (Figure 5H). Overall, "Resisting cell death," "Genomic instability and mutations," and "Enabling replicative immortality" showed a strong negative spatial dependency, suggesting a distant spatial relationship with their targets.

Our analysis of specific hallmark pairs showed intriguing patterns (Figures 5I and 5J). For example, "Resisting cell death" was highly important in predicting the spatial distribution of "Avoiding immune destruction" and "Tumor-promoting inflammation." A positive spatial dependency between "Nonmutational epigenetic reprogramming" and "Avoiding immune destruction" suggests that cancer cells undergoing epigenetic changes often interact with immune cells in a weakened state. Across different tumor types, the accuracy of the model did not vary significantly for predicting any TME hallmark (Figure S10A). However, "Deregulating cellular energetics" stood out as a key predictor of TME hallmarks with a negative spatial dependency, especially in liver adenocarcinomas (Figure S10B). A similar trend was seen for "Resisting cell death" in lung adenocarcinomas. Neoplastic hallmarks in other tumor types showed more variation in importance and spatial dependencies, aligning with patterns seen at the pancancer level (Figures S10B–S10D).

We show an example of these ecological dynamics in a breast-invasive carcinoma sample (Figure 5K). Here, cancer cells "Sustaining proliferative signaling" are colocalized with those "Evading growth suppressors" at the tumor-TME interface, marked by white-colored subspots. In contrast, cells deeper in the tumor core show reduced proliferative activity, indicated by red-colored subspots. This pattern suggests that cancer cells with impaired growth suppression mechanisms rely on interactions with TME accessory cells for proliferation signals or that cells at the invasive margin inherently exhibit higher proliferation rates. Another striking example is the relationship between "Resisting cell death" and "Avoiding immune destruction," which displays a strong negative spatial dependency (Figure 5J). This suggests that cancer cells may not need to resist cell death when the immune response is inactive. In a hepatocellular carcinoma sample (Figure 5L), cancer cells actively "Resisting cell death" (red-colored subspots) are predominantly located away from regions with active "Avoiding immune destruction" (cyan-colored subspots), illustrating this spatial dynamic.

Finally, we integrated the results from both experiments to create a comprehensive pancancer model of tumor architecture. This model ranked hallmarks in each compartment based on their median spatial dependency when predicting hallmarks in the other compartment (Figure 5M). The analysis reveals distinct cancer cell phenotypes organized by their proximity to the TME.

At the core of the neoplastic compartment, farthest from the TME, cancer cells predominantly exhibit "Resisting cell death," "Enabling replicative immortality," and "Genome instability and mutation." Surrounding this core is a middle layer of cells involved in "Deregulating cellular energetics" and "Senescence." The outermost layer, at the TME interface, features cancer cells active in "Evading growth suppressors" and "Nonmutational epigenetic reprogramming."

From the TME perspective, hallmarks are spatially distributed relative to their proximity to cancer cells. At the interface with cancer cells, TME cells are active in "Sustaining proliferative signaling," "Tumor-promoting inflammation," and "Avoiding immune destruction." These cells likely play a dual role, initially resisting cancer proliferation but eventually being co-opted to provide growth signals and suppress immune responses. In contrast, the core of the TME compartment, farther from the cancer mass, is enriched with cells exhibiting "Activating invasion and metastasis," "Inducing angiogenesis," and "Unlocking phenotypic plasticity." This layered organization underscores the dynamic interplay between tumor and TME compartments in shaping the tumor's ecological architecture.

### The ecological dynamics between cancer hallmarks correlate with clinical outcomes

While we identified pancancer relationships between cancer hallmarks in both tumor compartments, there was a notable inter-patient variability in the patterns we observed. These variations could indicate that certain cellular phenotypes are more relevant in some patients than others, potentially influencing tumor sensitivity to different drugs. To investigate this hypothesis, we generated spatial transcriptomics data using VISIUM for 33 muscle-invasive bladder cancer patients enrolled in the DUTRENEO clinical trial<sup>22</sup> (Table S1, Section A).

Pre-treatment samples were collected via transurethral bladder tumor resection (TURBT). The TURBT samples were used to evaluate inflammatory activity, allowing for the classification of tumors as immune hot or immune cold. Patients with immune-cold tumors received neoadjuvant chemotherapy, while those with immune-hot tumors were randomized to receive either standard neoadjuvant chemotherapy or neoadjuvant immunotherapy with checkpoint inhibitors targeting PD-L1 and CTLA4 (durvalumab and tremelimumab, respectively). Of the 73 patients ultimately included in the trial, we generated spatial transcriptomics data of pre-treated tumors from 33 patients, aiming to represent a range of patients with complete, partial, and no pathological response across all treatment groups.

We evaluated whether the pancancer spatial architecture of cancer hallmarks was conserved in this tissue-specific cohort. Our findings reveal that the distribution of cancer hallmarks between the neoplastic and TME compartments (Figure 6A) closely mirrors the pattern observed in the pancancer cohort (Figure 2A). This architecture is consistent regardless of the tumor's immune classification (immune cold or immune hot) or the patient's response to different neoadjuvant treatments. Next, we applied our random forest approach to model the relationships between hallmarks within each compartment for individual patients (Table S1, Section D; Figure S11). While the overall patterns were similar to the pancancer model, distinct differences emerged that appeared to be influenced by the tumor's immune status (Figures 6B and 6C).

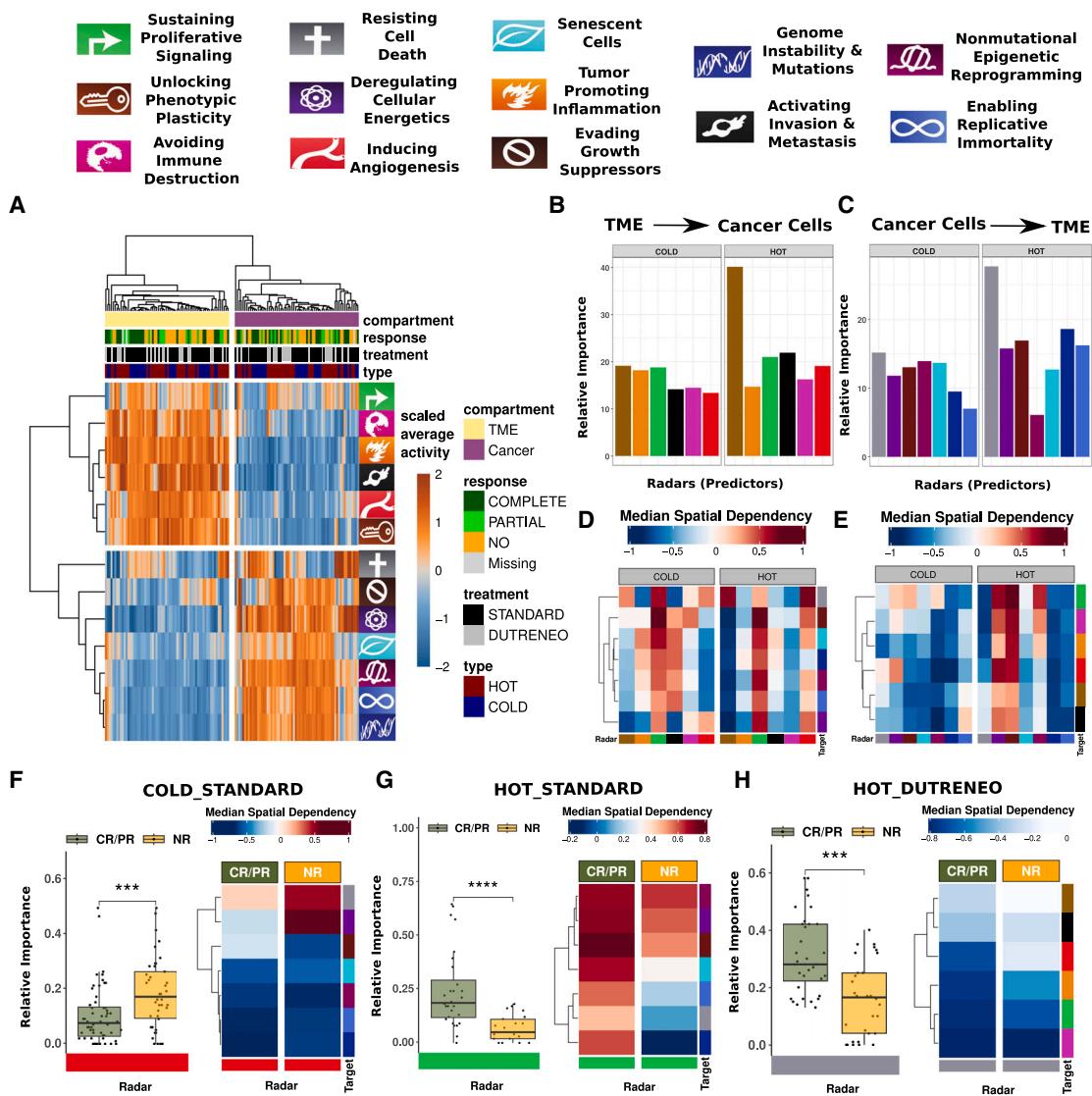
For instance, "Unlocking phenotypic plasticity" (Figure 6B) and "Resisting cell death" (Figure 6C) were the most important predictors in the TME and neoplastic compartments respectively, similar to the pancancer results. However, both hallmarks had significantly greater importance in immune-hot tumors than in immune-cold ones. Additionally, the spatial dependencies be-

tween hallmarks in both compartments (Figures 6D and 6E) varied between immune-hot and immune-cold tumors. This was especially evident for "Inducing angiogenesis," which predicts the distribution of neoplastic compartment hallmarks by proximity in immune-hot tumors but primarily by distance in immune-cold tumors (Figure 6D). In contrast, "Sustaining proliferative signaling" predicts the distribution of neoplastic hallmarks by proximity in both immune-hot and immune-cold tumors (Figure 6D), highlighting the need for TME accessory cells to sustain cancer cell proliferation across most tumor types. Overall, these findings suggest a model where the spatial relationship between different cancer hallmarks, while maintaining a similar architecture to the pancancer cohort (Figure 5M), can vary based on the tumor's underlying biology.

Given the relevance of spatial relationships in drug sensitivity, particularly in immune-based therapies,<sup>33</sup> we analyzed whether spatial interactions between hallmarks differed between patients who responded to various neoadjuvant therapies and those who did not. Our results show that the relative importance of specific TME or neoplastic hallmarks was significantly linked to sensitivity or resistance to these treatments. In the immune-cold group, which included only patients receiving neoadjuvant chemotherapy, "Inducing angiogenesis" was a stronger predictor of cancer cell distribution in non-responders than in responders (Figure 6F). Specifically, this hallmark predicted the proximity (positive spatial dependency) of cancer cells exhibiting "Resisting cell death" and "Deregulated cellular energetics," suggesting that these phenotypes might impede chemotherapy delivery or efficacy in non-responders. Interestingly, "Inducing angiogenesis" did not influence chemotherapy sensitivity in immune-hot tumors. Instead, in immune-hot tumors, "Sustaining proliferative signaling" emerged as a key differentiator between responders and non-responders (Figure 6G). This hallmark was more significant in predicting the location of active cancer cells in responders. Given its positive spatial dependency with all neoplastic hallmarks in responding patients (but not in non-responding patients), it is plausible that chemotherapy disrupts the TME in a way that critically affects tumor sustenance in immune-hot patients. Finally, among patients receiving neoadjuvant immunotherapy, the neoplastic hallmark "Resisting cell death" was significantly more important in responders compared to non-responders (Figure 6H). The negative spatial dependency of this hallmark suggests that reactivating immune cells is insufficient if the tumor ecosystem is structured around a core of cancer cells resistant to apoptosis.

### DISCUSSION

While the first cancer genomics projects were very successful in identifying recurrently mutated cancer driver genes,<sup>1,34</sup> fueling the first wave of personalized treatments, driver mutations interact with each other,<sup>35</sup> the tissue of origin,<sup>36</sup> the TME,<sup>2,36</sup> the germline genome,<sup>37</sup> and many other factors. Advances in single-cell genomics have shown how human tumors consist of an ever-increasing number of different cell types.<sup>38–41</sup> These cell types, together, shape the phenotype and evolution of the tumor. In this context, cancer hallmarks become a powerful tool to



**Figure 6. The spatial interdependence between cancer hallmarks depicts therapeutic vulnerabilities in muscle-invasive bladder cancer**

(A) Heatmap showing scaled average cancer hallmark scores (rows) for ESTIMATE clusters across subspots in 33 samples from the muscle-invasive bladder cancer (MIBC) cohort (columns). Annotations include (1) ESTIMATE clusters (1 and 2 for neoplastic compartments, 4 and 5 for TME compartments), (2) pathological response to neoadjuvant treatment, (3) treatment type (standard chemotherapy or DUTRENEO immunotherapy), and (4) immune infiltration type (immune cold or immune hot based on TIS classification).

(B) Relative importance of random forest models using TME hallmark radars as features to predict target neoplastic hallmarks, split by immune-cold and immune-hot samples from the MIBC cohort.

(C) Relative importance of random forest models in the reverse experiment: using neoplastic hallmark radars to predict target TME hallmarks, also split by immune-cold and immune-hot samples.

(D and E) Heatmaps showing median spatial dependencies for each predictor-target pair, corresponding to the experiments in (B) and (C), respectively, categorized by immune status.

(F–H) Boxplots of the relative importance of radars predicting specific hallmarks, divided by pathological response to neoadjuvant treatment (CR/PR, complete/partial responders; NR, non-responders). (F) "Inducing angiogenesis" predicting neoplastic hallmarks in immune-cold samples treated with standard chemotherapy. (G) "Sustaining proliferative signaling" predicting neoplastic hallmarks in immune-hot samples treated with standard chemotherapy. (H) "Resisting cell death" predicting TME hallmarks in immune-hot samples treated with DUTRENEO immunotherapy. Boxplots are paired with heatmaps showing median spatial dependencies between predictors (radars) and target hallmarks by pathological response. Colors for predictors and targets match the background colors of their icons at the top of the figure.

simplify this complexity to a reduced set of phenotypes required for every tumor.

We introduced a new approach to quantify the spatial heterogeneity of cancer hallmarks and identified key factors shaping their distribution. The first major factor is the division into two tumor compartments: cancer cells and the TME. Most hallmarks are predominantly active in one compartment or the other, but rarely in both. This near-binary classification, suspected previously, is now confirmed across 10 cancer types and shown to follow a structured spatial pattern. Some results challenge conventional assumptions. For instance, "Sustaining proliferative signaling" is often associated with cancer cells, yet we found it is more active in the TME, especially at the border with the neoplastic compartment. This aligns with studies showing that pre-cancerous TME can trigger proliferation in adjacent epithelial cells,<sup>27</sup> illustrating how the TME sustains proliferative signaling. Additionally, we observed many-to-many relationships between TME cell types and hallmarks, where a single cell type can contribute to multiple hallmarks, and several cell types influence a single hallmark. This finding underscores the need to study the spatial interactions between phenotypes<sup>42</sup> rather than focusing solely on cell type distribution, a shift from the current norm in spatial biology.

Our experiments revealed an intriguing pattern: as cancer cells diverge genomically, they often show corresponding divergence in the activity levels of different cancer hallmarks. This divergence can lead to the specialization of certain hallmarks within specific clones. Notably, this functional specialization may enhance the overall fitness of the tumor.<sup>42</sup> For instance, if the high activity of one hallmark physically limits another within the same cell, the tumor could optimize both by assigning these functions to different clones.<sup>43</sup> In such cases, genetic alterations that stabilize this segregation would likely undergo positive selection.<sup>44</sup> Interestingly, we also observed phenotypic convergence among genetically distant clones, as well as phenotypic divergence that occurred without genomic alterations. While these scenarios were less frequent overall, they appeared in most tumors, highlighting the significant role of non-genetic factors in shaping the gene expression patterns of cancer cells.<sup>45,46</sup>

We hypothesized that part of these non-genetic cues shaping the spatial positioning of the activity of cancer hallmarks could be reflected in their relative position to the cancer hallmarks from the other compartment. Our machine-learning models showed a conserved pancancer tumor architecture. Cancer cells at the tumor core predominantly exhibit hallmarks such as "Resisting cell death," "Enabling replicative immortality," and "Genome instability and mutation." These features likely reflect the need for sustained survival and proliferative capacity in regions less influenced by external factors. In contrast, the TME interface is characterized by hallmarks such as "Evasive growth suppressors" and "Nonmutational epigenetic reprogramming," suggesting that cancer cells at the invasive front are highly adaptive and interact dynamically with the TME. From the perspective of the TME, hallmarks such as "Sustaining proliferative signaling," "Tumor-promoting inflammation," and "Avoiding immune destruction" are concentrated at the interface with cancer cells, potentially serving dual roles: initially resisting tumor progression but eventually being co-opted to support cancer

growth. Meanwhile, deeper TME regions show enrichment for hallmarks such as "Activating invasion and metastasis" and "Inducing angiogenesis," indicating that spatially distinct regions of the TME play specialized roles in tumor progression. The spatial interdependence of cancer hallmarks between compartments highlights the ecological dynamics within tumors. This organization is reminiscent of how tissues optimize the location of cell types performing different functions to maximize the overall functionality of the entire tissue.<sup>43,47</sup> This perspective sheds light on the non-random nature of hallmark activity, suggesting that the spatial coordination of cancer and TME functions is a critical feature of tumor progression.

Finally, we validated our results in an additional spatial transcriptomics dataset from muscle-invasive bladder cancer patients from the DUTRENEO clinical trial.<sup>22</sup> The overall distribution of hallmark activities across the two major tumor compartments is conserved in this cancer-specific cohort, highlighting the robustness of this architecture. The inclusion of the DUTRENEO cohort also allowed us to explore whether variations in the spatial relationships between cancer hallmarks correlate with differences in clinical outcomes. This not only highlights new important aspects of the determinants of drug sensitivity but also provides a potential roadmap toward exploiting the spatial tumor architecture as a drug biomarker for personalized medicine.

In conclusion, the spatial architecture of cancer hallmarks provides a new lens to understand tumor biology, moving beyond bulk or single-cell analyses to incorporate the spatial and ecological context of tumors. This approach highlights the importance of tailored therapeutic strategies based on the spatial distribution and interdependence of hallmarks within the tumor ecosystem. For example, targeting hallmark-specific spatial niches could improve drug delivery and efficacy in immune-cold tumors, while reactivating apoptotic pathways may enhance the success of immunotherapy in tumors with resistant cores.

### Limitations of the study

While our study represents one of the largest spatial transcriptomics datasets to date, it is not comprehensive, as many cancer types remain unexamined. Additionally, the lack of information on driver mutations in these patients limits our ability to generalize findings across other molecular layers. Although we used InferCNV to predict the copy number states of tumor regions, this approach is not as precise as whole-genome sequencing, which would allow for accurate identification of somatic alterations. Another limitation is that VISIUM lacks single-cell resolution. To address this, we enhanced VISIUM's resolution using the BayesSpace method whenever possible. Future studies employing newer spatial transcriptomics platforms with single-cell resolution could provide greater detail and further refine the insights presented here.

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Eduard Porta-Pardo ([eporta@carrerasresearch.org](mailto:eporta@carrerasresearch.org))

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- Our spatial data have been deposited at Zenodo as <https://doi.org/10.5281/zenodo.1404496> and are publicly available as of the date of publication.
- All original code has been deposited at Zenodo and is publicly available <https://doi.org/10.5281/zenodo.14046805> as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

### ACKNOWLEDGMENTS

M.S. and S.C. are funded by the AECC project LABAE20038PORT. M.S. is also supported by a la Caixa Foundation Fellowship (LCF/BQ/DR22/11950021). E.P.-P. is funded by the Spanish Ministry of Science (RYC2019-026415-I) and Fundación Fero (BFERO2002.6). J.B. receives support from GEICO (GEICO 118-T). Josep Carreras Leukaemia Research Institute is supported by the CERCA Programme. This research was funded by MCIN/AEI/10.13039/501100011033 and the ERDF through PID2021-125282OB-I00, Generalitat de Catalunya (2021 SGR 01494) and the Sarah Jennifer Knott Foundation. The authors would like to thank Abel Gonzalez-Perez, Nuria Lopez-Bigas, and Eliezer Van Allen for their valuable feedback and the kind and generous patients who donated the samples that made this study possible. This work has been done as part of the PhD Programme in "Biochemistry, molecular biology and biomedicine" from Universitat Autònoma de Barcelona.

### AUTHOR CONTRIBUTIONS

Conceptualization, E.P.-P., M.S., and S.C.; methodology, M.S., M.H.B., S.C., and D.G.; formal analysis, M.S., S.C., D.G., E.M., R.L., C.-K.M., and A.J.L.; resources, E.P.-P., M.E., F.X.R., E.G., and L.D.; writing of the manuscript, all authors.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

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

During the preparation of this work, the authors used ChatGPT to reduce the character count. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

### 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
- METHOD DETAILS
  - Generation of spatial transcriptomics data from fresh frozen samples
  - Generation of spatial transcriptomics data from formalin-fixed paraffin-embedded (FFPE) samples
  - Generation of spatial transcriptomics data using Xenium
  - Sequencing, data processing and filtering
  - Spatial enhancement and gene expression imputation with Bayes-Space
  - Assigning genes to hallmarks
  - Assigning the neoplastic and TME compartments
  - Scoring the hallmarks

- Building gene exclusion controls
- Building gene addition controls
- Correction of the distances between subspots
- Spatial autocorrelation of hallmark activities
- Building spatial distribution controls
- Data preprocessing of image-based Xenium technology
- Transferring Xenium cell type assignments to deconvolve Visium data
- Detection of cancer clones by quantifying copy number alterations
- Modeling the spatial relationship between cancer hallmarks of the neoplastic and those of the TME compartments with random forests
- TME cell type composition predictions

### ● QUANTIFICATION AND STATISTICAL ANALYSIS

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2024.115229>.

Received: October 21, 2023

Revised: August 15, 2024

Accepted: December 31, 2024

Published: January 24, 2025

### REFERENCES

- Bailey, M.H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., Colaprico, A., Wendl, M.C., Kim, J., Reardon, B., et al. (2018). Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* 173, 371–385.e18.
- Thorsson, V., Gibbs, D.L., Brown, S.D., Wolf, D., Bortone, D.S., Ou Yang, T.-H., Porta-Pardo, E., Gao, G.F., Plaisier, C.L., Eddy, J.A., et al. (2018). The Immune Landscape of Cancer. *Immunity* 48, 812–830.e14.
- Prat, A., Pineda, E., Adamo, B., Galván, P., Fernández, A., Gaba, L., Díez, M., Viladot, M., Arancé, A., and Muñoz, M. (2015). Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast* 24, S26–S35.
- Hoadley, K.A., Yau, C., Hinoue, T., Wolf, D.M., Lazar, A.J., Drill, E., Shen, R., Taylor, A.M., Cherniack, A.D., Thorsson, V., et al. (2018). Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* 173, 291–304.e6.
- Rubin, J.B., Lagas, J.S., Broestl, L., Sponagel, J., Rockwell, N., Rhee, G., Rosen, S.F., Chen, S., Klein, R.S., Imoukhuede, P., and Luo, J. (2020). Sex differences in cancer mechanisms. *Biol. Sex Differ.* 11, 17.
- Huang, K.-L., Mashl, R.J., Wu, Y., Ritter, D.I., Wang, J., Oh, C., Paczkowska, M., Reynolds, S., Wyczalkowski, M.A., Oak, N., et al. (2018). Pathogenic Germline Variants in 10,389 Adult Cancers. *Cell* 173, 355–370.e14.
- Weinberg, R.A. (2014). Coming full circle—from endless complexity to simplicity and back again. *Cell* 157, 267–271.
- Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. *Cell* 100, 57–70.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674.
- Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 12, 31–46.
- Cancer Genome Atlas Research Network; Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R.M., Ozenberger, B.A., Elliott, K., Shmulevich, I., Sander, C., and Stuart, J.M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* 45, 1113–1120.
- Li, Y., Porta-Pardo, E., Tokheim, C., Bailey, M.H., Yaron, T.M., Stathias, V., Geffen, Y., Imbach, K.J., Cao, S., Anand, S., et al. (2023). Pan-cancer proteogenomics connects oncogenic drivers to functional states. *Cell* 186, 3921–3944.e25.

13. Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhor, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., et al. (2017). A pathology atlas of the human cancer transcriptome. *Science* 357, eaan2507. <https://doi.org/10.1126/science.aan2507>.
14. Iorio, F., Garcia-Alonso, L., Brammell, J.S., Martincorena, I., Wille, D.R., McDermott, U., and Saez-Rodriguez, J. (2018). Pathway-based dissection of the genomic heterogeneity of cancer hallmarks' acquisition with SLA-Penrich. *Sci. Rep.* 8, 6713.
15. de Visser, K.E., and Joyce, J.A. (2023). The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. *Cancer Cell* 41, 374–403.
16. Barkley, D., Moncada, R., Pour, M., Liberman, D.A., Dryg, I., Werba, G., Wang, W., Baron, M., Rao, A., Xia, B., et al. (2022). Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. *Nat. Genet.* 54, 1192–1201.
17. Rao, A., Barkley, D., França, G.S., and Yanai, I. (2021). Exploring tissue architecture using spatial transcriptomics. *Nature* 596, 211–220.
18. Hunter, M.V., Moncada, R., Weiss, J.M., Yanai, I., and White, R.M. (2021). Spatially resolved transcriptomics reveals the architecture of the tumor-microenvironment interface. *Nat. Commun.* 12, 6278.
19. Ravi, V.M., Will, P., Kueckelhaus, J., Sun, N., Joseph, K., Salie, H., Vollmer, L., Kuliesiute, U., von Ehr, J., Benotmane, J.K., et al. (2022). Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell* 40, 639–655.e13.
20. Li, R., Ferdinand, J.R., Loudon, K.W., Bowyer, G.S., Laidlaw, S., Muyas, F., Mamanova, L., Neves, J.B., Bolt, L., Fasouli, E.S., et al. (2022). Mapping single-cell transcriptomes in the intra-tumoral and associated territories of kidney cancer. *Cancer Cell* 40, 1583–1599.e10.
21. Greenwald, A.C., Darnell, N.G., Hoefflin, R., Simkin, D., Mount, C.W., Gonzalez Castro, L.N., Harnik, Y., Dumont, S., Hirsch, D., Nomura, M., et al. (2024). Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell* 187, 2485–2501.e26.
22. Wei, L.Y., Zhang, X.J., Wang, L., Hu, L.N., Zhang, X.D., Li, L., and Gao, J.N. (2020). DUTRENEO Trial: A randomized phase II trial of DUrvatumab and TRELimumab versus chemotherapy as a NEOadjuvant approach to muscle-invasive urothelial bladder cancer (MIBC) patients (pts) prospectively selected by an interferon (INF)-gamma immune signature. *OncoTargets Ther.* 38, 6497–6509. [https://doi.org/10.1200/JCO.2020.38.15\\_suppl.5012](https://doi.org/10.1200/JCO.2020.38.15_suppl.5012).
23. Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 353, 78–82.
24. Zhao, E., Stone, M.R., Ren, X., Guenthoer, J., Smythe, K.S., Pulliam, T., Williams, S.R., Uytingco, C.R., Taylor, S.E.B., Nghiem, P., et al. (2021). Spatial transcriptomics at subspot resolution with BayesSpace. *Nat. Biotechnol.* 39, 1375–1384.
25. Bergenstråhl, J., Larsson, L., and Lundeberg, J. (2020). Seamless integration of image and molecular analysis for spatial transcriptomics workflows. *BMC Genom.* 21, 482.
26. Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P.W., Levine, D.A., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612.
27. Skrupskelyte, G., Rojo Arias, J.E., Dang, Y., Han, S., Bejar, M.T., Colom, B., Fowler, J.C., Jones, P.H., Rulands, S., Simons, B.D., and Alcolea, M.P. (2024). Pre-cancerous Niche Remodelling Dictates Nascent Tumour Survival. Preprint at bioRxiv 2024, 07.04.602022. <https://doi.org/10.1101/2024.07.04.602022>.
28. Marusyk, A., Tabassum, D.P., Altrock, P.M., Almendro, V., Michor, F., and Polyak, K. (2014). Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature* 514, 54–58.
29. inferCNV of the Trinity CTAT Project. <https://github.com/broadinstitute/inferCNV.Preprint>.
30. Househam, J., Heide, T., Cresswell, G.D., Spiteri, I., Kimberley, C., Zapata, L., Lynn, C., James, C., Mossner, M., Fernandez-Mateos, J., et al. (2022). Phenotypic plasticity and genetic control in colorectal cancer evolution. *Nature* 611, 744–753.
31. Zahir, N., Sun, R., Gallahan, D., Gatenby, R.A., and Curtis, C. (2020). Characterizing the ecological and evolutionary dynamics of cancer. *Nat. Genet.* 52, 759–767.
32. Lin, J.-R., Wang, S., Coy, S., Chen, Y.-A., Yapp, C., Tyler, M., Nariya, M.K., Heiser, C.N., Lau, K.S., Santagata, S., and Sorger, P.K. (2023). Multiplexed 3D atlas of state transitions and immune interaction in colorectal cancer. *Cell* 186, 363–381.e19.
33. Williams, H.L., Frei, A.L., Koessler, T., Berger, M.D., Dawson, H., Michelin, O., and Zlobec, I. (2024). The current landscape of spatial biomarkers for prediction of response to immune checkpoint inhibition. *npj Precis. Oncol.* 8, 178.
34. ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium (2020). Pan-cancer analysis of whole genomes. *Nature* 578, 82–93.
35. Park, S., and Lehner, B. (2015). Cancer type-dependent genetic interactions between cancer driver alterations indicate plasticity of epistasis across cell types. *Mol. Syst. Biol.* 11, 824.
36. Cohen, R.L., and Settleman, J. (2014). From cancer genomics to precision oncology—tissue's still an issue. *Cell* 157, 1509–1514.
37. Sayaman, R.W., Saad, M., Thorsson, V., Hu, D., Hendrickx, W., Roelandts, J., Porta-Pardo, E., Mokrab, Y., Farshidfar, F., Kirchhoff, T., et al. (2021). Germline genetic contribution to the immune landscape of cancer. *Immunity* 54, 367–386.e8.
38. Kinker, G.S., Greenwald, A.C., Tal, R., Orlova, Z., Cuoco, M.S., McFarland, J.M., Warren, A., Rodman, C., Roth, J.A., Bender, S.A., et al. (2020). Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat. Genet.* 52, 1208–1218.
39. Tang, F., Li, J., Qi, L., Liu, D., Bo, Y., Qin, S., Miao, Y., Yu, K., Hou, W., Li, J., et al. (2023). A pan-cancer single-cell panorama of human natural killer cells. *Cell* 186, 4235–4251.e20.
40. Zheng, L., Qin, S., Si, W., Wang, A., Xing, B., Gao, R., Ren, X., Wang, L., Wu, X., Zhang, J., et al. (2021). Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* 374, abe6474.
41. Luo, H., Xia, X., Huang, L.-B., An, H., Cao, M., Kim, G.D., Chen, H.-N., Zhang, W.-H., Shu, Y., Kong, X., et al. (2022). Pan-cancer single-cell analysis reveals the heterogeneity and plasticity of cancer-associated fibroblasts in the tumor microenvironment. *Nat. Commun.* 13, 6619.
42. Thomas, F., DeGregori, J., Marusyk, A., Dujon, A.M., Ujvari, B., Capp, J.P., Gatenby, R., and Nedelcu, A.M. (2024). A new perspective on tumor progression: Evolution via selection for function. *Evolution, Medicine, and Public Health* 12, 172–177.
43. Haussler, J., Szekely, P., Bar, N., Zimmer, A., Sheftel, H., Caldas, C., and Alon, U. (2019). Tumor diversity and the trade-off between universal cancer tasks. *Nat. Commun.* 10, 5423.
44. Seteven, F.A., and Yanai, I. (2024). The origin of novel traits in cancer. *Trends in Cancer* 10, 880–892.
45. Whiting, F.J.H., Househam, J., Baker, A.-M., Sottoriva, A., and Graham, T.A. (2024). Phenotypic noise and plasticity in cancer evolution. *Trends Cell Biol.* 34, 451–464.
46. Laplane, L., and Maley, C.C. (2024). The evolutionary theory of cancer: challenges and potential solutions. *Nature Reviews Cancer* 24, 718–733.
47. Adler, M., Moriel, N., Goeva, A., Avraham-David, I., Mages, S., Adams, T.S., Kaminski, N., Macosko, E.Z., Regev, A., Medzhitov, R., and Nitzan, M. (2023). Emergence of division of labor in tissues through cell interactions and spatial cues. *Cell Rep.* 42, 112412.
48. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalex, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888–1902.e21.

49. Manukyan, A., Bahry, E., Wyler, E., Becher, E., Pascual-Reguant, A., Plumbom, I., Dikmen, H.O., Elezkurtaj, S., Conrad, T., Altmüller, J., et al. (2023). VoltRon: A Spatial Omics Analysis Platform for Multi-Resolution and Multi-omics Integration using Image Registration. Preprint at bioRxiv 12, 571667. <https://doi.org/10.1101/2023.12.15.571667>.
50. Wright, M.N., and Ziegler, A. (2015). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Preprint at arXiv. <https://doi.org/10.18637/jss.v077.i01>.
51. Komisarczyk, K., Kozminski, P., MakSYMiuK, S., and Biecek, P. (2023). treeshap: fast SHAP values computation for tree ensemble models.
52. Wu, S.Z., Al-Eryani, G., Roden, D.L., Junankar, S., Harvey, K., Andersson, A., Thennavan, A., Wang, C., Torpy, J.R., Bartonicek, N., et al. (2021). A single-cell and spatially resolved atlas of human breast cancers. *Nat. Genet.* 53, 1334–1347.
53. Wu, Y., Yang, S., Ma, J., Chen, Z., Song, G., Rao, D., Cheng, Y., Huang, S., Liu, Y., Jiang, S., et al. (2022). Spatiotemporal Immune Landscape of Colorectal Cancer Liver Metastasis at Single-Cell Level. *Cancer Discov.* 12, 134–153.
54. Meylan, M., Petitprez, F., Becht, E., Bougoün, A., Pupier, G., Calvez, A., Giglioli, I., Verkarre, V., Lacroix, G., Verneau, J., et al. (2022). Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. *Immunity* 55, 527–541.e5.
55. Wu, R., Guo, W., Qiu, X., Wang, S., Sui, C., Lian, Q., Wu, J., Shan, Y., Yang, Z., Yang, S., et al. (2021). Comprehensive analysis of spatial architecture in primary liver cancer. *Sci. Adv.* 7, eabg3750.
56. Gracia Villacampa, E., Larsson, L., Mirzazadeh, R., Kvastad, L., Andersson, A., Mollbrink, A., Kokaraki, G., Monteil, V., Schultz, N., Appelberg, K.S., et al. (2021). Genome-wide spatial expression profiling in formalin-fixed tissues. *Cell Genom.* 1, 100065.
57. Cui Zhou, D., Jayasinghe, R.G., Chen, S., Herndon, J.M., Iglesia, M.D., Navale, P., Wendl, M.C., Caravan, W., Sato, K., Storrs, E., et al. (2022). Spatially restricted drivers and transitional cell populations cooperate with the microenvironment in untreated and chemo-resistant pancreatic cancer. *Nat. Genet.* 54, 1390–1405.
58. Erickson, A., He, M., Berglund, E., Marklund, M., Mirzazadeh, R., Schultz, N., Kvastad, L., Andersson, A., Bergensträhle, L., Bergensträhle, J., et al. (2022). Spatially resolved clonal copy number alterations in benign and malignant tissue. *Nature* 608, 360–367.
59. Franzén, O., Gan, L.-M., and Björkegren, J.L.M. (2019). PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database* 2019, baz046. <https://doi.org/10.1093/database/baz046>.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Chemicals, peptides, and recombinant proteins</b>		
Hydrochloric Acid Solution, 0.1N	Fisher Chemical	SA54
Potassium Hydroxide Solution, 8M	Millipore Sigma	P4494
SSC Buffer 20X Concentrate	Millipore Sigma	S66391L
Tris (1M) pH 8	Invitrogen	AM9855G
TE Buffer, TRIS-EDTA, 1X Solution, pH 8.0	Thermo Fisher Scientific	BP24731
Urea Solution, 8M	Millipore Sigma	51457
Dimethyl sulfoxide (molecular biology grade)	Millipore Sigma	41639-100 ML
Tween 20 Surfact-Amps Detergent Solution (10% solution)	Thermo Fisher Scientific	28320
KCl (2 M), RNase-free	Invitrogen	AM9640G
<b>Critical commercial assays</b>		
Rneasy Mini Kit	Quiagen	Cat# 74104
Rneasy FFPE Kit	Quiagen	Cat# 73504
Visium Accessory Kit	10x Genomics	Cat# PN-1000194
Visium Spatial Tissue Optimization Slide & Reagent Kit	10x Genomics	Cat# PN-1000193
Visium Spatial Gene Expression Slide & Reagent Kit, 16 rxns	10x Genomics	Cat# PN-1000184
Dual Index Plate TT Set A, 96 rxns	10x Genomics	Cat# PN-1000215
Visium Spatial for FFPE Gene Expression Kit, Human Transcriptome, 16 rxns	10x Genomics	Cat# PN-1000336
Dual Index Plate TS Set A, 96 rxns	10x Genomics	Cat# PN-1000251
KAPA SYBR FAST qPCR Master Mix (2X)	KAPA Biosystems	KK4600
SPRIselect Reagent Kit	Beckman Coulter	B23318
Xenium Decoding Consumables (1 run, 2 slides)	10x Genomics	Cat# PN-1000487
Xenium Decoding Reagents (1 run, 2 slides)	10x Genomics	Cat# PN-1000461
Xenium Slides & Sample Prep Reagents (2 slides, 2 rxns)	10x Genomics	Cat# PN-1000460
Kit, Xenium Human Multi-Tissue and Cancer Panel (2 rxns)	10x Genomics	Cat# PN-1000626
<b>Deposited data</b>		
The spatial landscape of Cancer Hallmarks reveals patterns of tumor ecology and drug sensitivity	This paper	<a href="https://doi.org/10.5281/zenodo.14044964">https://doi.org/10.5281/zenodo.14044964</a>
<b>Software and algorithms</b>		
Space Ranger 1.3.1	10x Genomics	<a href="https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/installation">https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/installation</a>
Loupe Browser 5.0	10x Genomics	<a href="https://support.10xgenomics.com/spatial-gene-expression/software/visualization/latest/installation">https://support.10xgenomics.com/spatial-gene-expression/software/visualization/latest/installation</a>
STUtility 1.1.1	Bergenstråhle et al. <sup>25</sup>	<a href="https://ludvigla.github.io/STUtility_web_site/">https://ludvigla.github.io/STUtility_web_site/</a>
Seurat 4 & 5	Stuard et al. <sup>48</sup>	<a href="https://satijalab.org/seurat/">https://satijalab.org/seurat/</a>
BayesSpace 1.8.2	Zhao et al. <sup>24</sup>	<a href="https://github.com/edward130603/BayesSpace">https://github.com/edward130603/BayesSpace</a>
ChatGPT 3.5	N/A	<a href="https://chatgpt.com/">https://chatgpt.com/</a>

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ESTIMATE 1.0.1	Yoshihara et al. <sup>26</sup>	<a href="https://bioinformatics.mdanderson.org/public-software/estimate/">https://bioinformatics.mdanderson.org/public-software/estimate/</a>
Voltron 1.0.0	Manukyan et al. <sup>49</sup>	<a href="https://github.com/BIMSBioinfo/VoltRon">https://github.com/BIMSBioinfo/VoltRon</a>
inferCNV 1.21.0	inferCNV of the Trinity CTAT Project <sup>29</sup>	<a href="https://github.com/broadinstitute/infercnv">https://github.com/broadinstitute/infercnv</a>
ranger 0.16.0	Wright et al. <sup>50</sup>	<a href="https://github.com/imbs-hl/ranger">https://github.com/imbs-hl/ranger</a>
Treeshap 0.3.1	Komisarczyk et al. <sup>51</sup>	<a href="https://github.com/ModelOriented/treeshap">https://github.com/ModelOriented/treeshap</a>
Deposited Code	This paper	<a href="https://doi.org/10.5281/zenodo.14046805">https://doi.org/10.5281/zenodo.14046805</a>

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All tumor samples of the PanCancer cohort for which we generated spatial transcriptomics data ( $n = 31$ ) were obtained under the ethics committee approval 2022/78-APA-HUGC. Samples from the DUTRENEO trial ( $n = 33$ ) were obtained during the trial itself (NCT03472274). Details about the DUTRENEO design and patient allocation can be found in the original publication<sup>22</sup>.

## METHOD DETAILS

### Generation of spatial transcriptomics data from fresh frozen samples

Frozen Ovarian Cancer samples (HGSOC) were embedded in cryomolds using OCT (Tissue Tek Sakura) using a bath of isopentane and liquid nitrogen (Tissue preparation guide CG000240, 10x Genomics). Following freezing in OCT, blocks were stored at  $-80^{\circ}\text{C}$ . RNA integrity was assessed by calculating RNA Integrity Number (RIN) of freshly collected tissue sections with Qiagen protocol (RNeasy Mini Kit 74104) and analyzed by TapeStation. Samples with Rin above 7 were selected for the experiments. For sectioning, blocks were equilibrated to  $-20^{\circ}\text{C}$  in the cryochamber (LEICA CM1950). 10  $\mu\text{m}$ -thick sections were placed onto the active areas (6 mm  $\times$  6 mm) of chilled 10x genomics Visium slides. For Tissue Optimization, 7 sections were placed on a Tissue Optimization Slide (3000394, 10X Genomics) then fixed in chilled methanol and stained according to the Visium Spatial Tissue Optimization User Guide (CG000238, 10X Genomics) to determine optimal permeabilization time for HGSOC. For tissue optimization experiments, fluorescent images were taken with a TRITC filter using a 10X objective and 900 ms exposure time. For Gene Expression samples, 2 consecutive sections were placed on a chilled Visium Spatial Gene Expression Slide (2000233, 10X Genomics), and adhered by warming the back of the slide. Tissue sections were then fixed in chilled methanol and Hematoxylin & Eosin staining was performed according to the Demonstrated Protocol (CG000160, 10x genomics). Brightfield histology images were taken using a 10X objective (Plan APO) on a Nikon Eclipse Ti2, images were stitched together using NIS-Elements software (Nikon) and exported as tiff files. Following imaging, HGSOC samples were permeabilized for 30 min and cDNA Synthesis and amplification was performed following Visium Spatial Gene Expression User Guide (CG000239, 10X Genomics).

Libraries were prepared according to the Visium Spatial Gene Expression User Guide (CG000239, 10X Genomics) and sent for sequencing Using HiseqX 150PE (2x 150bp) applying 1% Phix. Sequencing depth was calculated with the formula (Coverage Area  $\times$  total spots on the Capture Area)  $\times$  50,000 read pairs/spot. Sequencing was performed using the following read protocol: read 1: 28 cycles; i7 index read: 10 cycles; i5 index read: 10 cycles; read 2: 90 cycles.

### Generation of spatial transcriptomics data from formalin-fixed paraffin-embedded (FFPE) samples

RNA integrity was assessed by calculating DV200 of RNA extracted from freshly collected tissue sections. Briefly, Tissue blocks were placed in the microtome and cut to expose the tissue. 4 sections 10  $\mu\text{m}$  thick were placed in a chilled Eppendorf tube and RNA extraction protocol from Qiagen was performed (Rneasy FFPE Kit 73504) and RNA was analyzed by TapeStation. 21 Samples with DV200  $\geq 25\%$  were selected for experiments. FFPE samples were placed in the microtome and sectioned 7  $\mu\text{m}$  thick, after floating on a water bath at  $42^{\circ}\text{C}$ , sections were placed on Visium Spatial Gene Expression slides (2000233, 10X Genomics). After sectioning, the slides were dried at  $42^{\circ}\text{C}$  for 3 h. The slides were then placed inside a slide mailer, sealed with parafilm, and left overnight at Room temperature. The slides were deparaffinized by successive immersions in xylene and ethanol followed by H&E staining according to Demonstrated Protocol (CG000409, 10X Genomics).

Brightfield images were taken using a 10X objective (Plan APO) on a Nikon Eclipse Ti2, images were stitched together using NIS-Elements software (Nikon) and exported as tiff files. After imaging, the glycerol and cover glass were carefully removed from the Visium slides by holding the slides in an 800 mL water beaker and letting the glycerol diffuse until the cover glass detached and density changes were no longer visible in the water. The slides were then dried at  $37^{\circ}\text{C}$ .

Libraries were prepared according to the Visium Spatial Gene Expression for FFPE User Guide (CG000407, 10X Genomics) and sent for sequencing Using HiseqX 150PE (2x 150bp) applying 1% Phix. Sequencing was performed using the specific for FFPE following read protocol: read 1: 28 cycles; i7 index read: 10 cycles; i5 index read: 10 cycles; read 2: 50 cycles.

### Generation of spatial transcriptomics data using Xenium

The Xenium workflow began by sectioning 7 $\mu$ m FFPE tissue sections onto a Xenium slide, according to the "Xenium *In Situ* for FFPE-Tissue Preparation Guide" (CG000578 Rev C, 10X Genomics). Briefly, the sections were cut, floated in an RNase-free water bath at 42°, and carefully placed onto the capture area of a Xenium slide (PN- 1000465). We strategically placed three 7  $\times$  7mm sections in each slide. After sectioning, the slides were incubated at 42° per 3 h, and kept in a sealed bag with desiccators at 4° overnight. The next day, the slides were shipped to the Dresden Genome Center Facility, where the experiments were conducted.

At the Dresden Genome Center Facility, the Xenium slides were processed following the "Xenium *In Situ* for FFPE- Deparaffinization and Decrosslinking" protocol (CG000580 Rev C, 10X Genomics). The slides were equilibrated to room temperature and the protocol of deparaffinization was performed.

Subsequently, the Xenium slides were assembled into Xenium cassettes (PN-1000566, 10X Genomics), which allow for the incubation of slides on the Xenium Thermocycler Adapter inside a Thermocycler machine with a closed lid for optimal temperature control. The slides were processed using the "Xenium Slides and Sample Prep Reagents" kit (PN-1000460, 10X Genomics), starting with incubation in a decrosslinking and permeabilization solution at 80°C for 30 min, followed by washes with PBS-T.

The Xenium slides were then processed according to the "Xenium *In Situ* Gene Expression" user guide (CG000582 Rev D, 10X Genomics). The slides were incubated at 50°C overnight for approximately 19 h with the gene expression panel ("Xenium Human Multi-Tissue and Cancer Panel", PN-1000626, 10X Genomics, which targets 377 human genes). This was followed by a series of washes and steps, including a post-hybridization wash at 37°C for 30 min, a ligation at 37°C for 2 h, and an amplification step at 30°C for 2 h. After additional washing steps, the slides were treated with an autofluorescence quencher and a nuclei staining step.

Finally, at the end of the second day of the protocol, the two slides in the cassettes were loaded into the Xenium Analyzer. The first step in the instrument consists in a sample scan, where images of the fluorescent nuclei in each section are given, and these images allow the user to select and determine the regions to be included in the analysis. For the 6 samples, we selected all the tissue to be analyzed.

The run in the Xenium Analyzer lasted around 50 h. After that, the instrument was emptied of consumables and the Xenium slides carefully removed. PBS-T was added to the slides and a post-run H&E staining was performed.

### Sequencing, data processing and filtering

Raw fastq files of two ovarian cancer, four colorectal cancer, four breast cancer, five bladder cancer, five prostate cancer, seven lung cancer, and four glioblastoma samples were processed with the "spaceranger" command line tool (version 1.3.1, 10x Genomics) and mapped to the pre-built human reference genome (GRCh38). Thirty-six publicly available VISIUM tumor datasets from seven tissues were downloaded from recent publications<sup>19,52–58</sup> and the 'spatial gene expression' panel of 10x Genomics (<https://www.10xgenomics.com/resources/datasets>). The samples were obtained from the following sections: 'Visium Spatial Targeted Demonstration Data (v1)', 'Visium Spatial Fluorescent Demonstration (v1)', 'Visium Spatial for FFPE Demonstration (v1)'. For each downloaded ST dataset, the output from SpaceRanger was loaded individually in the STUtility environment.<sup>25</sup> Each sample was filtered following three steps. First, genes with less than 5 UMIs were excluded. Second, mitochondrial and ribosomal genes, and non protein-coding genes were removed from the analysis. Genes that were kept were tagged as protein\_coding, TR\_V\_gene, TR\_D\_gene, TR\_J\_gene, TR\_C\_gene, IG\_LV\_gene, IG\_V\_gene, IG\_J\_gene, IG\_C\_gene and IG\_D\_gene. Lastly, spots were filtered out depending on two conditions: if they were isolated from the majority of the tissue and/or if the number of features was lower than a threshold value established by the distribution in each sample. Each filtered sample was then individually normalized by variance stabilizing transformation on the expression data using SCTtransform (return.only.var.genes = FALSE, variable.features.n = NULL, variable.features.rv.th = 1.1).<sup>48</sup>

### Spatial enhancement and gene expression imputation with BayesSpace

As each of the VISIUM spots may contain a number of cells (1–10), we sought to conduct our downstream analysis in a way that better approximates single-cell resolution. We therefore used the novel BayesSpace R package that leverages a fully Bayesian statistical framework by using information from spatial neighborhoods to computationally enhance the resolution of spatial transcriptomics data.<sup>24</sup> To that end, the normalized gene expression data was converted to a SingleCellExperiment object with the respective spatial coordinates and metadata. In the preprocessing step (spatialPreprocess), normalization was skipped and 15 principal components (PCs) were computed using the top 2000 highly variable genes in the SCT assay. The number of clusters was chosen by the elbow of the pseudo-log-likelihood plot (qTune). Then, spots were clustered by spatialCluster using 50000 iterations, 15 PCs, gamma = 3 (recommended for Visium), and the initialization method k-means. To increase the resolution of the data, each spot was divided into 6 subspots using spatialEnhance from BayesSpace. For those samples that were processed using the "-reorient-images" flag in SpaceRanger, their "imagerow" and "imagecol" columns were switched to prevent inaccurate enhanced clustering. The resulting clusters were used as the initialization of the algorithm using 200000 iterations and 0.3 of prior jitter. Finally, xgboost was used to impute the expression of the complete set of genes for each subspot with a maximum number of boosting iterations of 100.

### Assigning genes to hallmarks

The full list of pathways was downloaded from the Pathway Commons data portal (v12, 2019) (<http://www.pathwaycommons.org/archives/PC2/v12/>). From this collection, pathways with less than 4 genes or more than 475 genes in their corresponding gene sets

were discarded, as were those with “CHEBI” symbol corresponding to small chemical compounds, yielding a total of 3,250 pathways. To assign genes for each Hallmark, a two-stage pathway filtering process was undertaken. First, specific keywords that can be associated with specific Hallmarks were searched in the names of the pathways following a similar previously reported approach,<sup>14</sup> and the resulting pathways were assigned to their associated Hallmark. Second, for Hallmarks that were assigned an excessive number of pathways, ChatGPT (v. 3.5) was used to narrow-down the number of pathways in those Hallmarks by prompting it to retain pathways that are minimally involved across Hallmarks. Following this pathway-to-Hallmark assignment, each list was manually inspected to ensure the credibility of these assignments. This approach was applied for all 13 Hallmarks except “Resisting cell death” as it required a more careful assignment of pathways. Since pathways that are directly related to cell death mechanisms such as apoptosis may contain pro- as well as anti-apoptotic genes, this could mask the downstream signature interpretation. We therefore manually assigned a number of pathways that could directly be related to survival mechanisms, which would indirectly imply resisting cell death.

### Assigning the neoplastic and TME compartments

“ESTIMATE” R package<sup>26</sup> was used to quantify the neoplastic cell purity within each subspot in the enhanced VISIUM dataset. This package has been extensively and successfully used to quantify the presence of cancer, immune or stromal cells in RNAseq profiles from bulk tumor samples. As such, we assumed that, since each (sub)spot in our dataset can contain multiple cells (often between 1 and 10), it could be treated as a mini-bulk RNAseq and, therefore, deconvoluted with the same approach. To perform the quantification, the gene expression of each subspot was first extracted and transformed into GCT format. Then, the “estimate” function was used to compute the stroma, immune, and ESTIMATE scores. Finally, ESTIMATE scores were added as metadata, and they were used to create five clusters within each sample using the k-means algorithm. The cluster with the lowest ESTIMATE values corresponds to the maximum purity of neoplastic cells, whereas the cluster with the highest ESTIMATE values corresponded to the minimum purity of neoplastic cells (TME). An expert pathologist validated these scores, who determined the neoplastic and stroma cell content in each cluster through H&E images from seventeen samples.

### Scoring the hallmarks

To compute the activity of a Hallmark, the enhanced object was first converted back to Seurat, and the expression matrix was scaled (ScaleData) to perform downstream analysis. A module score was computed for each Hallmark’s gene set using AddModuleScore from Seurat<sup>48</sup> on the enhanced objects, followed by scaling and centering the obtained scores within each sample, yielding what we refer to as “Hallmark activity”. The resulting Hallmark activities were then averaged per ESTIMATE cluster depending on the neoplastic cell content. Rows and columns were clustered using “average” as the clustering method and “correlation” as the distance measure after scaling the columns. The same Seurat function was used to score the activities of each individual pathway from which the genes formed the signature of “Nonmutational epigenetic reprogramming”.

### Building gene exclusion controls

Two levels of controls were built in order to ensure the reliability of the genes in the selected pathways to sufficiently represent the quantified Hallmark scores. First, 25% of the genes from each Hallmark-associated pathway were randomly excluded and the remaining genes were combined for each Hallmark to form a control gene signature. The second level was similarly applied with 50% exclusion. Each level was repeated 5 times to obtain a total of 10 randomized Hallmark control signatures. The controls were scored using the same parameters with AddModuleScore from Seurat.

### Building gene addition controls

In order to investigate whether adding new genes to each of the 13 Hallmark gene sets could influence the signature scores, the percentage of overlap between the genes of each pathway (excluding the final 109 pathways) from the initially filtered list (3,250 pathways) and the genes of each of the 13 Hallmark gene-sets was computed. Pathways with an overlap of at least 60% and at most 90% were subsequently assigned as an accessory pathway list for the associated Hallmark based on an assumption that those pathways with such a high overlap and which were not included initially may also be important for the associated Hallmark. The union of the unique set of genes from these accessory pathway lists associated with one Hallmark were stored as a list of accessory genes that could be assigned to that Hallmark. From this accessory list for each Hallmark, a number of genes corresponding to 5%, 10%, 25%, and 50% of the size of the original Hallmark-associated gene set were selected and added to the original gene set. This process was repeated 5 times for each of the added percentages. The resulting Hallmark gene set controls were scored using the same parameters with AddModuleScore from Seurat.

### Correction of the distances between subspots

The distance between the center of two adjacent spots is 100  $\mu\text{m}$  in the VISIUM technology, which has a spot diameter of 55  $\mu\text{m}$ . These distances must be considered after enhancing, as two subspots coming from the same spot will have a shorter distance than two adjacent subspots from two different spots. Therefore, the spot location grid created by BayesSpace at the spot level was multiplied by 100  $\mu\text{m}$  to preserve the distance between spots. Then, each subspot was shifted from the center following the

same hexagonal distribution as BayesSpace enhancing, but instead of applying 1/3 of the distance between spots, a third of the radius (27.5/3) was used to determine the center of each subspot.

### Spatial autocorrelation of hallmark activities

To determine the spatial autocorrelation of Hallmark activities in the tissue, we modified the function CorSpatialGenes() from STUtility to apply it to our enhanced VISIUM samples. Each neighborhood was computed using 10 neighbors (5 subspots coming from the same spot and 5 others from adjacent spots) and weighted by the corrected subspot distances. Spatial autocorrelation was computed for the 7 TME Hallmarks within ESTIMATE clusters 4&5 and the 6 neoplastic Hallmarks within ESTIMATE clusters 1&2, and for all 13 Hallmarks within the whole tissue of each sample.

### Building spatial distribution controls

Control signatures were built to compare their spatial autocorrelation with those of Hallmark signatures. These controls were generated by randomly selecting genes of various sizes: 50, 100, 200, 350, and 500. For each size, 25 sets of genes were randomly sampled from the intersection of all genes captured across all samples. The controls were scored using the same parameters with AddModuleScore from Seurat. Subsequently, spatial autocorrelation was calculated for each gene set, and the average value was determined for each sample and size.

### Data preprocessing of image-based Xenium technology

Gene expression profiles matrices from Xenium Analyzer (v1.5.0.3) of each sample were imported using Seurat (v5). After removing cells with no expression, the expression data of each sample was normalized by variance stabilizing transformation (SCTransform). Principal component analysis (PCA) was done on each sample dataset with 30 dimensions. Cells were clustered using the Louvain algorithm implemented in the FindClusters function using a resolution of 0.5. Cluster identities were defined by the top over-expressed genes detected in each cluster through FindAllMarkers (logfc.threshold = 0.75, min.pct = 0.1, only.pos = T). Annotation of cell types was carried out using a set of markers provided by 10x Genomics for the selected multi-tissue and cancer gene panel of 377 genes.

### Transferring Xenium cell type assignments to deconvolve Visium data

First, we applied image registration using the software Voltron<sup>49</sup> by defining common landmarks between Xenium (DAPI) and Visium (H&E) images to align Visium and Xenium samples from consecutive sections. After obtaining the spot coordinates of Visium in the Xenium reference space, we assigned each spot the corresponding set of Xenium cells whose centroids are within 27.5 µm of the spot centroids. Finally, relative abundance of each cell type was computed at the spot level.

### Detection of cancer clones by quantifying copy number alterations

inferCNV method<sup>29</sup> was used to detect copy number alterations (CNA) in each spot based on gene expression. Initially, the compartment labels from ESTIMATE clusters were assigned to the spot resolution using the max-voting rule. In case of a tie, the label "Buffer" was assigned. Subsequently, raw gene expression counts were used to infer CNA in neoplastic spots (clusters 1 & 2) with TME spots (clusters 4 & 5) serving as a reference. The "Buffer" spots were excluded from this analysis. The inferCNV function was executed with the following parameters: run(object\_infCNV, num\_threads = 1, cutoff = 0.1, cluster\_by\_groups = F, plot\_steps = T, denoise = T, no\_prelim\_plot = F, k\_obs\_groups = 1, HMM = T, leiden\_resolution = 0.005, BayesMaxPNormal = 0.2).

The output results were obtained from step 19, which included spot groups from the dendrogram and the CNA state of each gene after filtering low-probability events. We then transferred back to the subspot resolution, excluding spot groups with fewer than 10 spots (60 subspots) as they might be outliers in the dendrogram. Next, we calculated the genomic distance between each pair of groups within a sample by determining the number of genes with the same CNA state divided by the total number of genes used to infer CNA states (JSI score). For those clones with a similarity higher than 99%, they were merged to avoid overclustering of clones. CNA events of these merged clones were defined as the most frequent event in a given gene. Then, JSI scores were recomputed by including the merged clones. To quantify the Hallmark difference, we averaged the Hallmark score within each group and computed the absolute difference. Finally, we correlated, for each sample and Hallmark, the genomic distance (1 -JSI) and the absolute Hallmark difference using Pearson correlation.

### Modeling the spatial relationship between cancer hallmarks of the neoplastic and those of the TME compartments with random forests

Before running the machine learning experiments, the scaled Hallmark scores were shifted to positive values in order to minimize the effect of negative values from the originally scaled Hallmark scores on downstream interpretation. The shifted scores were calculated by transforming the minimum scaled Hallmark scores into zero and shifting the rest of the values accordingly. A small constant is then added for each value to prevent dividing by zero in downstream analysis. These shifted values are then used in the input of the following machine learning experiments.

The spatial relationship between the Hallmarks of the neoplastic (ESTIMATE clusters 1&2&3) and TME (ESTIMATE clusters 3&4&5) compartments (including the intermediate cluster #3) was modeled on the subspot level by running two machine learning experiments.

- 1) predicting the spatial distribution of each one of the 7 neoplastic Hallmarks as individual targets using the 6 TME Hallmarks altogether as the features (predictors), yielding 7 targets \* 63 samples = 441 models in total.
- 2) vice-versa: predicting the spatial distribution of each one of the 6 TME Hallmarks as individual targets using the 7 neoplastic Hallmarks altogether as the features (predictors), yielding 6 targets \* 63 samples = 378 models in total.

In either of the experiments, as the predictors must be located in the same subspots (compartment) where the target is located, we therefore translated the activity of predictors from their original compartment into an activity within the target compartment that serves much like a “radar” (Figures 4A and S4). These “radar” scores are constructed in each of the 63 samples that ultimately serve as the predictors in each model. In the case of the first experiment, the radar scores for each of the 6 TME Hallmarks in each neoplastic subspot are calculated following 2 major steps.

- 1) Each “radar” in each neoplastic subspot first captures the shifted Hallmark score from each one of the TME subspots weighted by the inverse of the corrected distance between each TME subspot and the neoplastic subspot where the radar is located.
- 2) These individual spatially weighted scores that are coming from all TME subspots in each neoplastic subspot are then summed up into a single score for each neoplastic subspot, which we finally call the “radar” score.

The same process is applied reversely when running the second experiment. Radar scores as predictors of their target Hallmarks from either experiments are then used to train the model on 80% of randomly-selected subspots with the ranger package,<sup>50</sup> followed by predicting the target Hallmark activity (response) using the rest 20% of the subspots with the computation of the shapley additive values (SHAP).<sup>51</sup> We used the SHAP values to identify how the model prioritizes the relationships between the features (radars) and the target Hallmark using two metrics. The first is feature importance, which is the mean of absolute values of SHAP values. After scaling these values across features from 0 to 100, the fraction of each feature is used to infer the relative importance of that feature in predicting its target Hallmark. The second metric is feature dependency, which profiles the relationship between the feature and its SHAP value as a proxy to its target. Pearson correlation is calculated between them to infer the directionality of these relationships. Finally, as the predictors are weighted by distances we, therefore, call the latter metric in this context as “spatial dependency” for a more intuitive understanding.

### TME cell type composition predictions

Cell type markers from the *PanglaoDB* database<sup>59</sup> were used to quantify the spatial distribution of certain cell types pertinent to the TME. Cell types from three main organ categories were selected from the database to represent the TME; the immune system, vasculature, connective tissue, and nervous system. For each cell type, only canonical markers were obtained, followed by retaining cell types with at least 10 canonical marker genes, yielding a total of 45 cell types.

Graphical figures Individual panels with statistical data were generated with the R packages “ggplot2”, “pheatmap”, “Complex-Heatmap”. Figures 3A, 4A, and 5M were generated using BioRender, and all individual panels were put together using GIMP or Inkscape.

### QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details of each experiment can be found in the figure legend or the methods section.