Charting the Spatial Landscape of Cancer Hallmarks

Mustafa Sibai[1,2]*, Sergi Cervilla[1]*, Daniela Grases[1], Eva Musulen[1,3], Arola Fortian[4], Margarita Romeo[4], Adrià Bernat[4], Collin Tokheim[5,6], Manel Esteller[1,7], Jordi Barretina[4], Li Ding[8], Matthew Bailey[9^] and Eduard Porta-Pardo[1,2^]

1 – Josep Carreras Leukaemia Research Institute (IJC), Badalona, Spain
2 – Barcelona Supercomputing Center (BSC)
3 – Department of Pathology, Hospital Universitari General de Catalunya Grupo-QuirónSalud, Sant Cugat del Vallès, Spain.
4 – Institut de Recerca Germans Trias i Pujol (IGTP), Badalona, Spain
5 – Department of Data Science, Dana-Farber Cancer Institute, Boston, MA 02215, USA
6 – Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA
7 – Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
8 – Department of Medicine and Genetics, Siteman Cancer Center, Washington University in St. Louis, St. Louis, MO, USA
9 - Department of Biology and Simmons Center for Cancer Research, Brigham Young University, Provo, Utah 84602, USA

* These authors contributed equally
^ To whom correspondence should be addressed: eporta@carrerasresearch.org / matthew.bailey@byu.edu

Abstract

Cancer Hallmarks have been indispensable to rationalize the complex phenotypic properties acquired by cancer cells and their surrounding tumor microenvironment (TME) as they progress from normalcy to malignancy. While great insights have been gained from their use to understand bulk tissue transcriptomics, our knowledge remained incomplete, as it was not possible to determine the origin and relative localization of the activity of each Cancer Hallmark within tumor tissues. Thanks to recent advances in spatial transcriptomics, we can now explore the spatial contribution of each compartment within a tumor to each Cancer Hallmark at near single-cell resolution.

Here, we used spatial transcriptomics from fifteen different primary tumor samples from seven tissues to quantify the contribution of the main tumor compartments -cancer cells and the TME- to each of thirteen Cancer Hallmarks. Our results revealed a Pan-Cancer architecture: with the activity of some hallmarks consistently coming from the cancer compartment across all cancer types (Cancer-associated hallmarks) and the activity of others consistently coming from the TME (TME-associated hallmarks). Furthermore, we observed that even within each compartment (cancer or TME) there was a consistent internal architecture, as many of these hallmarks exhibited higher activity at the border between compartments. Our study provides the first spatial map of Cancer Hallmark activities in tumors at almost cellular resolution, paving the way for a better understanding of the underpinnings of hallmark acquisition by tumor cell types.

INTRODUCTION

Cancer is a complex disease whose phenotypes arise from diverse mechanisms happening throughout multiple molecular layers. The concept of Cancer Hallmarks provides an integrative approach to rationalize these mechanisms, as they distill the complex molecular machinery and the myriad of (epi)genetic and transcriptomic aberrations observed in a tumor into logical principles through which cells acquire the features needed in their transformation from normalcy to neoplastic growth states [1]. In fact, many of the currently available cancer drugs have been developed to target specific Cancer Hallmarks, such as inhibitors of cyclin-dependent kinases to prevent cancer cells from 'Sustaining proliferative signaling' or immune-activating anti-CTLA4 and anti-PD-L1 antibodies to prevent them from 'Avoiding immune destruction' [2,3].

In the current era of big data, Cancer Hallmarks have been applied to rationalize diverse types of omics information, such as those generated by The Cancer Genome Atlas, and to understand the genome-wide effects of individual genes on tumors [4]. These data were generated to identify cancer-driving mutations [5], gene amplification [6], genomic rearrangements [7], and quantify the transcriptome-wide gene expression changes, providing a proxy to the cellular state pertaining to the stage of the disease [8]. The typical process for identifying whether specific hallmarks are more or less active in the studied tumor type or patient sample is by associating them with specific genes, biological pathways, or functional properties [8]. However, until recently, most omics datasets have been generated from high-throughput profiling of bulk tumors, limiting our understanding of the interplay between the different cell types that form the two major compartments within a tumor: cancer cells and their surrounding tumor microenvironment (TME). Despite the difficulty in disentangling the cellular origin of the activities of Cancer Hallmarks, the importance of the TME in studying Cancer Hallmarks was already evident when Hanahan and Weinberg, more than a decade ago, expanded their original six Cancer Hallmarks so that, among others, this conceptual framework could be applied to "expand upon the functional roles and contributions made by recruited stromal cells to tumor biology" [3].

The advent of spatial transcriptomics (ST) in the last few years initiated a new era promising to untangle the complex ecosystem of tumors in their spatial context [9]. Among the different spatial transcriptomics options, the VISIUM platform is now becoming widely used due to its ability to profile the whole transcriptome of a relatively large tissue section (a square with a side of 6.5mm). Moreover, while it does not reach single-cell resolution, having in each 55µm spot about 5 to 20 cells, recent ST-tailored algorithms have been developed to counteract this limitation and computationally increase the resolution of each spot to almost single-cell [10]. Together, these advancements finally allow us to map the position and transcriptional landscape of different

compartments within a tumor and begin to unravel the interactions among them. Consequently, mapping the Cancer Hallmarks on ST data would allow us to finally determine the spatial contribution of the different tumor compartments to each Hallmark, paving the way to understand the complex cellular cross-talk that promotes malignancy.

In this study, we quantified the activity of Cancer Hallmarks in fifteen cancer samples from seven different cancer types (**Fig. 1A**). Furthermore, we compared the activity of each hallmark in the cancer and TME compartments, which were defined based on histopathological annotations within each tumor sample. Our findings demonstrate a spatially variable activity of the hallmarks, revealing their distinct categorization based on whether the activity originates from cancer cells or the TME. Finally, we show how there is spatial heterogeneity in the activity of the different Cancer Hallmarks even within each tumor compartment, revealing an internal structure within cancer compartments. This architecture shows how senescent cells with active gene expression programs to avoid immune destruction are in the periphery of the cancer compartment that interacts with the TME, and an extensive epigenetic reprogramming expression pattern in the core of the cancer compartment that is far from the TME.

RESULTS

Spatial Transcriptomics illustrates the spatial variability of Cancer Hallmarks in ovarian cancer

The response to PARP inhibitors in ovarian cancer patients with BRCA1/2 mutations is very heterogeneous. Given the interplay between cancer cells and the TME in patients with these mutations, we thought that spatial transcriptomics might help to begin to understand the interactions between the different cell types in such tumors. To that end, we collected two serial sections from an ovarian cancer sample of High-Grade Serous Ovarian Cancer (HGSOC) subtype obtained from a BRCA1 mutant patient with stage III disease in a primary surgery (prior to any systemic treatment). A pathological review of the hematoxylin and eosin (H&E) slides suggested broad cancer and TME compartments (**Fig. 1B**). Concurrently, we quantified the tumor's transcriptomic profile on a single capture area of VISIUM platform (**Methods**). We preprocessed the HGSOC sample and computationally enhanced the resolution of each spot with BayesSpace [10]. As a first control, we compared the expression of two important markers of cancer cells (*EPCAM* and *MYC*) and another marker for the TME (TGFB) in both histological (and consecutive) sections [8–10] (**Fig. 1C**). While EPCAM was measured by its normalized gene expression, MYC activity was measured by the expression of its transcriptional targets and TGFB pathway activity was measured by the expression of its responsive genes (**Methods**). Our results showed that both consecutive sections (10 µm apart) have very spatially-similar expression patterns, suggesting that the data generated is of good quality and robust. In both sections, EPCAM expression and MYC activity are largely limited to a compartment in the middle of the tumor, and TGFB activity to the area immediately surrounding the compartment with MYC and EPCAM expression, suggesting that these are the cancer and TME compartments, respectively.

We then used Nonnegative Matrix Factorization (NMF) [11] to systematically identify the different gene-expression compartments within the tumor (**Methods, Fig. S1**). To capture the general granularity in the HGSOC sample, we used five activity maps from NMF. Then, each of these areas was labeled based on their overlap with histopathological annotations. Based on the H&E image, the pathologist annotated the regions with carcinoma and TME components, and then labeled the factors based on their overlap with these annotations (**Fig. 1D**).

Next, we quantified the activity of each Cancer Hallmark in each spot of the spatial transcriptomics dataset. To that end, we compiled lists of genes associated with each of 13 hallmarks (all 14 hallmarks from [1] except microbiome) by a strategy similar to that used in [12], where each gene list was formed from a collection of hallmark-associated pathways from the latest "Pathway Commons" database. The number of associated hallmarks per gene followed a similar trend to that reported by *Iorio et. al.* [12], showing that the majority of genes were associated with

few hallmarks (**Fig. S2**). We then used the Seurat R package to score the activity of each hallmark, whereby the average gene expression levels of each hallmark were computed and subtracted by the aggregated expression of control gene sets. Genes from each hallmark-associated gene set are binned based on averaged expression such that the control features are randomly selected from each bin. The score of each hallmark was then centered and scaled and overlaid on the histology image, revealing the spatial landscape of the activity of each Cancer Hallmark (**Fig. 1E, Fig. S3**).

Afterward, we quantified the contribution of the cancer and TME compartments to each of the hallmarks (**Fig. 1F**), by classifying each spot according to which of the five compartments it belongs to and combining that with the histopathological annotations. The boxplot revealed a significantly higher contribution of the cancer compartment to five hallmarks: 'Evading growth suppressors', 'Enabling replicative immortality', 'Genome instability and mutation', 'Senescent cells', and 'Nonmutational epigenetic reprogramming', when compared to the four TME compartments. On the other hand, the activities of 'Avoiding immune destruction', 'Tumor-promoting inflammation', 'Sustaining proliferative signaling', 'Activating invasion and metastasis', 'Inducing angiogenesis', and 'Deregulating cellular energetics' were significantly higher in the TME compartments. These observations suggest compartment-specific Cancer Hallmark activation patterns with a distinct spatial distribution in this analyzed sample, and raise the hypothesis that this pattern might be present across cancer types.

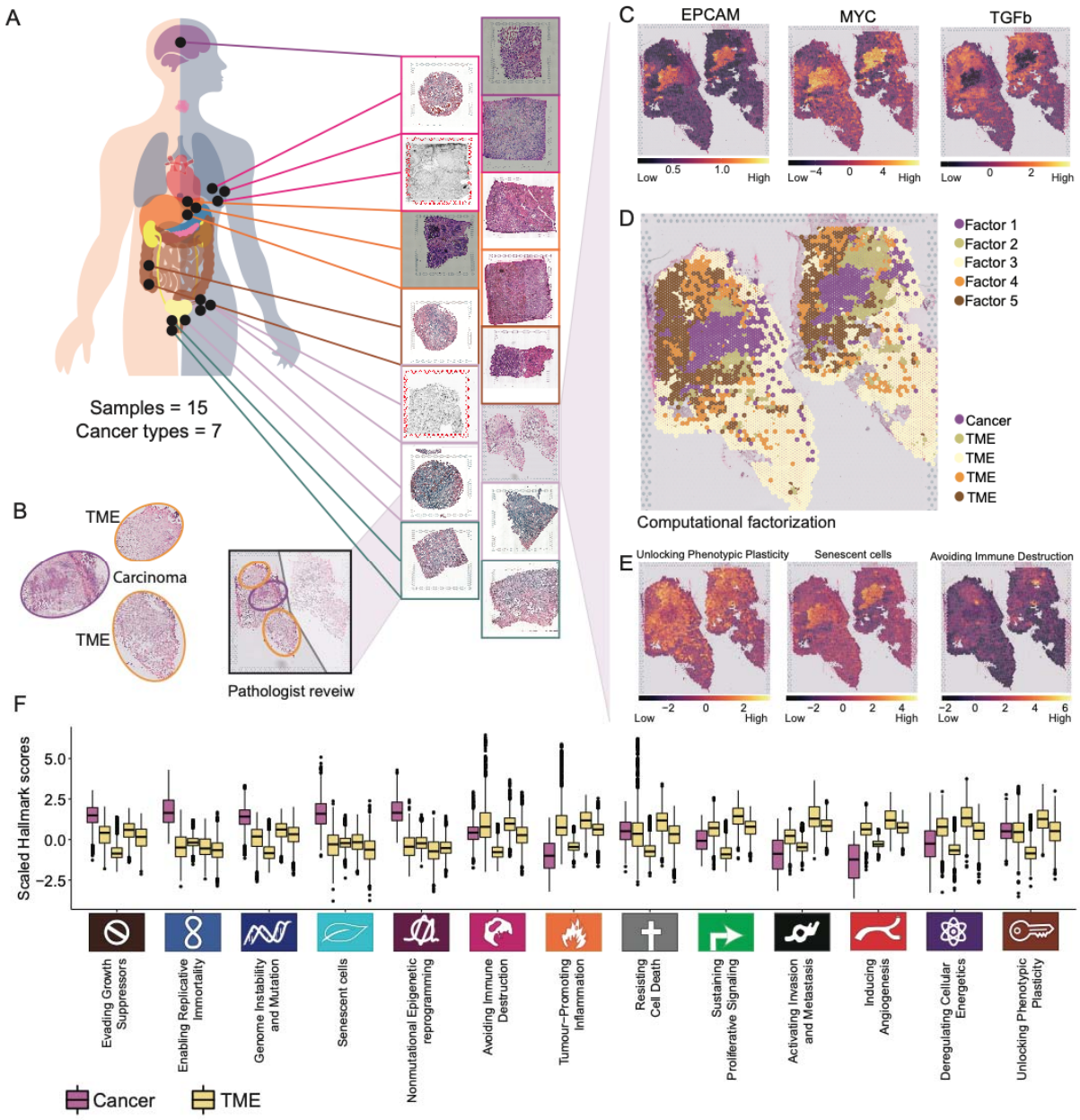

Figure 1 - Integrating spatial transcriptomics and histopathological annotations to identify the source of Cancer Hallmarks. A) An illustration of all the VISIUM datasets used in this study. The datasets were generated from 15 samples across 7 major cancer types, and were all collected from public repositories except one of the ovarian cancer sample datasets that we generated - the High-Grade Serous Ovarian Carcinoma (HGSOC). The squared sections on the right side refer to the H&E or immunofluorescence staining as the initial reference for histopathological annotations and the subsequent downstream analyses. **B)** Each one of the 15 sections was reviewed by a pathologist to broadly assign the regions with carcinoma or TME phenotype. The pathology review of the HGSOC sample is shown here. **C)** Expression levels of EPCAM and MYC transcription factor activity referring to the carcinoma region, and TGFB pathway activity referring to the TME, all of which overlapping with histopathological annotations of the HGSOC VISIUM sample with enhanced subspot-level resolution. **D)** Nonnegative Matrix Factorization (NMF) on the computationally enhanced HGSOC VISIUM sample identifying 5 major factors, which were classified by the pathologist into carcinoma or TME based on the H&E image review. **E)** The scaled scores of 3 Cancer Hallmarks overlaid on the background H&E histology image of the computationally enhanced HGSOC sample. **F)** Boxplots revealing the contribution of the two tumor compartments- cancer and TME- that were labeled by the pathologist based on the H&E image combined with computational factorization to each of the 13 Cancer Hallmarks in the computationally enhanced HGSOC sample.

Spatial Transcriptomics reveals a Pan-Cancer architecture of Cancer Hallmarks

We then applied the same analysis to fourteen different publicly available VISIUM datasets generated from tumor samples collected from seven different tissues. Importantly, nine of these datasets used the VISIUM protocol for fresh-frozen samples and six the protocol for FFPE samples, ensuring that we could identify any potential biases due to the spatial transcriptomics protocol. Moreover, these samples have been collected from seven different cancer types: glioblastoma (1), breast (3), liver (3), colorectal (2), ovarian (3, including the sample explained in **Fig. 1B-F**), prostate (2), and cervix (1) (**Fig. 1A**). All of these public datasets were collected from the 10x Genomics website (**Methods**) except the liver samples, which were collected from [13]. Main quality statistics, such as the final number of genes and spots for each processed VISIUM data, were considered sufficient for a reliable downstream analysis (**Table S1)**.

Following the approach described in the previous section, we first used NMF to identify the five major tumor compartments in each sample (**Fig. S1**). These compartments (n = 75, five compartments per sample) were then labeled either "Cancer" or "TME" according to their overlap with histopathological annotations (**Fig. S4**). Two of these compartments overlapped with clearly distinct histopathological annotations, so we label them as "unassigned". In two other factors, the histopathological annotations indicated the presence of both, stroma and cancer cells, within the same area, so we assigned a "Cancer/TME" label.

Next, we quantified the activity of each Cancer Hallmark in each spot of each dataset. Again, following the approximation described in **Fig. 1**, this allowed us to quantify the contribution of the different compartments (cancer or TME) to each hallmark in each sample. Our results showed a variable contribution of tumor compartments to all hallmarks within each dataset (**Fig. S4**) that were mostly consistent with those observed in the ovarian cancer sample from **Fig. 1,** with the activity of some hallmarks being higher either in the cancer or the TME compartments consistently across most cancer types.

To test whether this represented a Pan-Cancer architecture, we calculated the correlation between the embeddings of each of the five NMF factors of each sample with the activity of each Cancer Hallmark. Next, we used a 2D hierarchically clustered heatmap to plot the correlations between each NMF factor and the Cancer Hallmark activities and determine whether the patterns we observed in the ovarian cancer sample are conserved across tumor types (**Fig. 2A**). Clustering on the columns revealed three major dendrogram branches, one of which clusters all but two of the cancer-labeled compartments across all the different tumor samples together (Chi-Square $p < 1e-15$), while the other two branches cluster all but one of the TME-labeled compartments (Chi-square $p < 1e-15$). Notably, this clustering was independent of the VISIUM protocol (Chi-square $p > 0.45$)

or the cancer type (Chi-square $p > 0.98$), suggesting that the results are robust to the VISIUM protocol and potentially represent a true Pan-Cancer architecture.

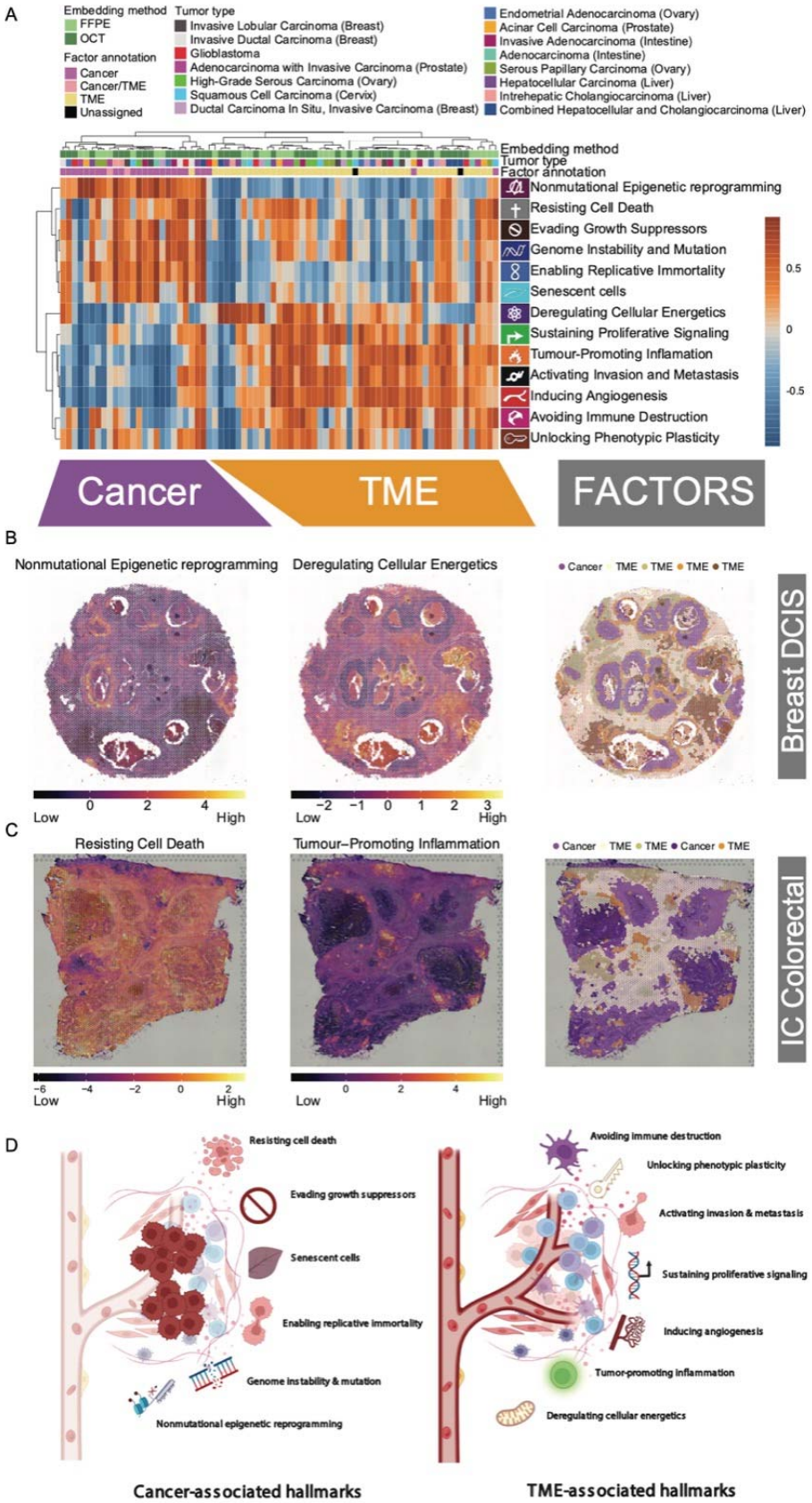

Figure 2 - Spatial transcriptomics clusters Cancer Hallmarks based on their cellular origin. A) 2D-hierarchically clustered heatmap of the Pearson correlations between the NMF factor embeddings and the scaled hallmark scores across all 15 computally enhanced VISIUM samples, labeled by all 75 factor annotations, tumor type, or tissue embedding method (fresh frozen or FFPE). **B)** The overlay of the scaled scores of Cancer Hallmarks exemplifying their tumor compartment-specific activity in Cancer (left) or TME (middle), based on the factor annotations (right) in the case of Ductal Carcinoma in Situ with Invasive Carcinoma Breast sample. **C)** In the case of Invasive Adenocarcinoma Colorectal sample, the overlaid scaled hallmark scores exemplify a scenario where a Cancer Hallmark (left)- which is labeled as cancer-associated on the Pan-Cancer clustered heatmap- may exhibit its activity in both compartments in some cancer types, while still having other hallmarks that are compartment-specific (middle), based on the factor annotations (right). **D)** A graphical illustration recapitulating the observation of having a Pan-Cancer architecture where some Cancer Hallmarks have the propensity to be associated with cancer compartments and others with TME compartments.

Clustering on the rows showed two major dendrogram branches. One branch grouped six hallmarks that appear to have positive correlations with cancer compartments and negative correlations with TME-labeled compartments. These Cancer Hallmarks include “Nonmutational Epigenetic Reprogramming”, “Resisting Cell Death”, “Genome Instability and Mutation”, “Enabling Replicative Immortality” and “Senescent Cells”. The other branch, on the other hand, showed a completely opposite pattern and grouped together Cancer Hallmarks with positive correlations with TME-labeled compartments and negative correlations with cancer-labeled compartments. These Cancer Hallmarks include “Deregulating Cellular Energetics”, “Sustaining Proliferative Signaling”, “Tumor-Promoting Inflammation”, “Activating Invasion and Metastasis”, “Inducing Angiogenesis”, “Avoiding Immune Destruction” and “Unlocking Phenotypic Plasticity”.

Notably, the overall activity of Cancer Hallmarks seems to clearly define a Pan-Cancer spatial architecture. Some Cancer Hallmarks are truly cancer- or TME-specific, such as “Nonmutational Epigenetic Reprogramming" or “Deregulating Cellular Energetics”, respectively. An example of both of these Hallmarks can be seen in the Ductal Carcinoma in Situ with Invasive Carcinoma sample (**Fig. 2B**). We noticed, though, that some Cancer Hallmarks generally associated with cancer-labeled compartments were also active in a few TME-labeled compartments and vice versa. For example, “Resisting Cell Death” activity has a positive correlation with the factors defining 17 out of 25 cancer-labeled compartments, but also with the factors defining 21 out of 46 TME-labeled compartments. One of the clearest examples of this phenomenon is in one of the colorectal cancer samples, where the activity of this Cancer Hallmark is high throughout the entire histological area (**Fig. 2C**). We similarly observed a compartment-unspecific activity of “Unlocking phenotypic plasticity” (**Fig. 2A**). This may suggest that the activation of this hallmark (and that of “Resisting cell death”) is potentially tumor type- or context-dependent.

<u>Cancer Hallmark activity varies within tumor compartments</u>

While our results showed that most Cancer Hallmarks are mostly active in either cancer cells or the TME (**Fig. 2D**), we noticed that the activity *within* these major tumor compartments was

also variable in some cases. For example, “Senescence” seems to be more active on the upper left border of the cancer compartment in the HGSOC sample, **Fig. 1E, F**).

To quantify this phenomenon and, therefore, explore the intra-compartment variability of the activity of Cancer Hallmarks, we split the spots of each cancer compartment into “TME-adjacent” and “TME-distant” subcompartments based on their minimum distance to the TME. Following the same reasoning, each spot in the TME compartment was split into “cancer-adjacent” and “cancer-distant” subcompartments. We then quantified the contribution of the defined tumor subcompartments to each hallmark in the different VISIUM datasets (**Fig. S5**).

Our results showed an internal architecture of the activity of Cancer Hallmarks within each main tumor compartment (**Fig. 3A**). In the case of the cancer-labeled compartments, most Cancer Hallmarks were usually more active in the area that is TME-adjacent. The only exception is “Nonmutational Epigenetic Reprogramming”, which is more active in the cancer subcompartment that is TME-distant. Similarly, most Cancer Hallmarks associated with TME-compartments had higher activity in the TME areas that are cancer-adjacent. These Cancer Hallmarks include “Avoiding Immune Destruction”, “Unlocking Phenotypic Plasticity”, “Activating Invasion and Metastasis” or “Sustaining Proliferative Signaling” among others. The only exception in this case was “Deregulating Cellular Energetics”, which, for most samples, is higher in the TME subcompartment that is cancer-distant.

While our Pan-Cancer results revealed general principles of the internal architecture of Cancer Hallmark activity in the cancer and TME compartments, there were some notable exceptions. For example, in the case of the HGSOC sample, we observed that “Enabling Replicative Immortality” had higher activity in the cancer subcompartment that is TME-adjacent (**Fig. 3B**). In another case, the Ductal Carcinoma In Situ with Invasive Carcinoma sample (**Fig. 3C**), we observed intriguing Cancer Hallmark patterns. The TME subcompartments that are in direct contact with cancer-associated subcompartments showed sustained high hallmark scores in both the TME- and Cancer-specific hallmarks–most notably “Evading Growth Suppressors”. We find this observation intriguing for this cancer type because of the established dependence of the endocrine system on breast cancer growth [14]. In a third example, the Invasive Adenocarcinoma Colorectal Sample (**Fig. 3D**), “Senescent cells” did not show any notable differential activity in all defined subcompartments, and whereas “Nonmutational Epigenetic Reprogramming” was found highly active in the cancer compartments, there was no notable differential activity between its subcompartments either. These examples of hallmark activities deviating from the general pattern could be a proxy for cancer stage-dependent activation patterns.

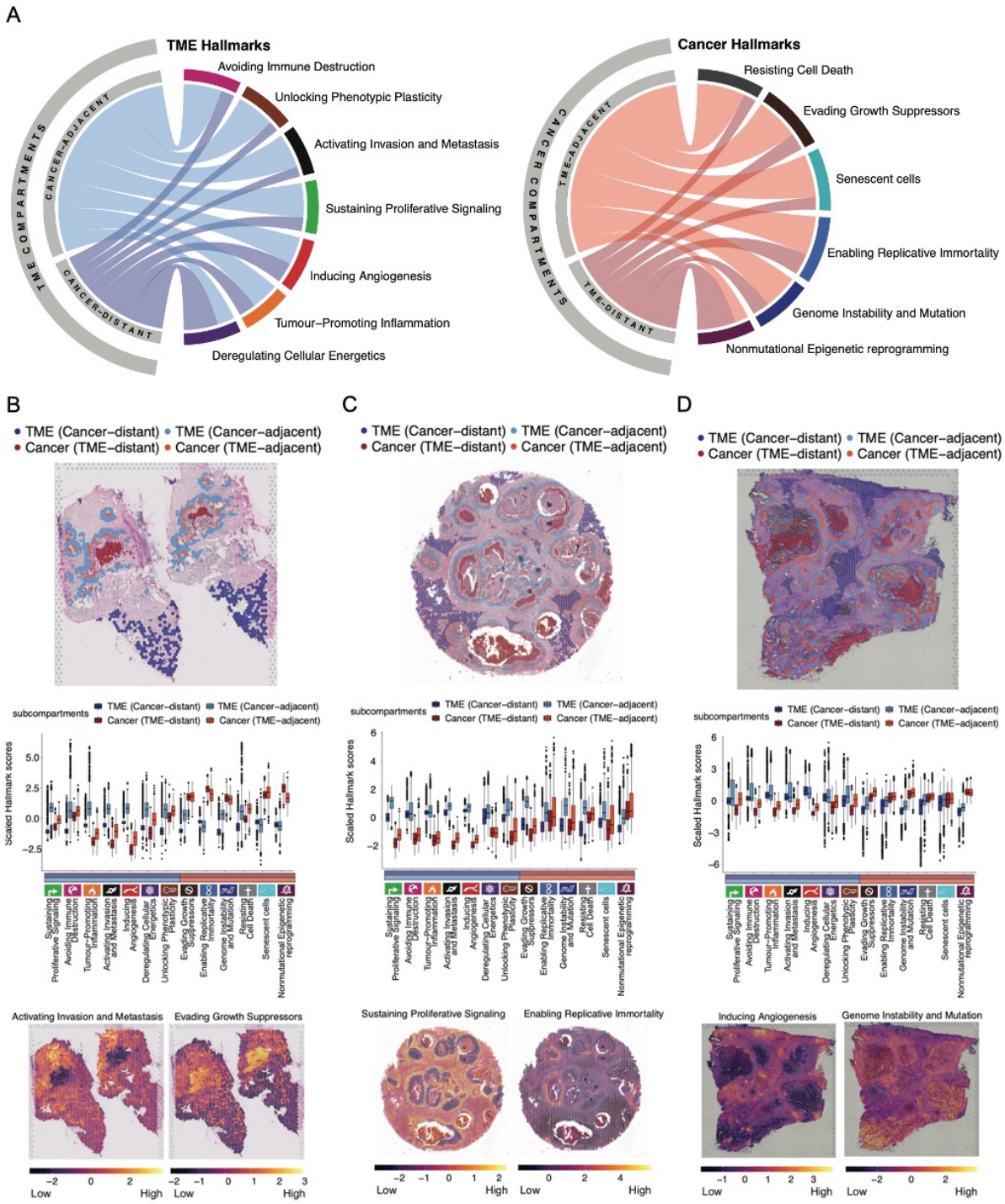

Figure 3 - Spatial transcriptomics reveals the internal architecture of tumor compartments. A) Circos plots demonstrating in how many instances a Cancer-associated Hallmark (Fig. 2D) is significantly more active in the cancer subcompartments that are either adjacent to or distant from TME compartments (left), and in how many instances a TME-associated Hallmark (Fig. 2D) is significantly more active in the TME subcompartments that are either adjacent to or distant from Cancer compartments (right). The subcompartments were defined based on the minimum distance of each spot of one compartment to those of the other compartment. **B-D)** The quantified contribution of the defined subcomparments to each of the 13 Cancer Hallmarks, showing the defined subcompartments on the histology image (Top), and their contribution to the Hallmarks (middle), with examples of compartment-specific activities of Cancer Hallmarks with their scaled scores overlaid on the histology image (Bottom), in the B) High-Grade Serous Ovarian Carcinoma, C) Ductal Carcinoma in Situ with Invasive Carcinoma of the Breast, and D) Invasive Adenocarcinoma Colorectal samples.

DISCUSSION

By distilling the complex molecular and cellular biology of cancer into fourteen clear features required by tumors to become malignant, Cancer Hallmarks have made it possible to study this disease from a more systematic perspective. However, until recently, the insights from analyses using Cancer Hallmarks were limited due to the lack of resolution of bulk sequencing technologies. By using data from spatial transcriptomics, our approach provides the first spatial landscape of Cancer Hallmarks at near single-cell resolution, revealing that their activity follows a Pan-Cancer architecture and that most Cancer Hallmarks can be clearly assigned to one of the two main compartments of tumors: cancer cells or their surrounding TME (**Fig. 2D**).

This architecture also highlights the extensive crosstalk between these two tumor compartments. For example, while the cancer-labeled compartments had the high activity of the "Evading Growth Suppressors" and "Enabling Replicative Immortality" hallmarks, the TME had higher activity of the "Sustaining Proliferative Signaling" hallmark, highlighting how both tumor compartments contribute to complementary Cancer Hallmarks so that cancer cells can grow. Similar examples are "Resisting Cell Death", coming mostly from the cancer-labeled compartments, and "Avoiding Immune Destruction", coming from the TME-labeled compartments. Together, these two hallmarks show how the inhibition of the immune system (Avoiding Immune Destruction) and its effects on cancer cells (Resisting Cell Death) are features contributed by both tumor compartments.

Our results also revealed a potential internal architecture within both, the TME and the cancer compartments, with some Cancer Hallmarks being more active in specific areas in relation to the distance to the opposite compartment (**Fig. 3A**). This was particularly interesting in the case of the cancer compartment, as it suggests that cancer cells have different transcriptional programs depending on their exposure to the TME. In particular, it seems that cancer cells that are in close proximity to the TME have higher activity of "Senescent cells", "Resisting Cell Death" and "Evading Growth Suppressors", three Cancer Hallmarks that are extremely relevant for cell-to-cell communication. For example, senescence is classically believed to act as an alternative cell-intrinsic fail-safe mechanism akin to apoptosis, keeping malignant progression at bay [15], until recent reports emerged to suggest its ability to promote tumor development in certain contexts [16]. Senescence-associated secretory phenotype (SASP) is the posited mechanism through which

senescent cells promote tumor phenotypes, whereby these cells convey in paracrine fashion various signaling molecules to neighboring cancer cells and other cells in the TME so as to propagate the acquisition of other hallmarks [16–18]. On the other hand, cancer cells further away from the TME usually have higher activity only in one of the Cancer Hallmarks: "Nonmutational Epigenetic Reprogramming". This Cancer Hallmark has been proposed to contribute to intratumoral heterogeneity and, therefore, be partly responsible for generating phenotypic diversity within cancer cells [19]. This cancer compartment structure suggests a model where an outer layer of cells with the necessary phenotypes to survive a direct interaction with the TME is protecting an inner core of cancer cells that are exploring the phenotypic landscape by means of epigenetic reprogramming.

Figure 4 - The Spatial Map of Cancer Hallmarks. Our approach allows from H&E images (left) a comprehensive analysis of Cancer Hallmark activity (center) that can be used to understand which Cancer Hallmarks are most relevant to each tumor region (right).

In conclusion, we hope that this study highlights how the concept of Cancer Hallmarks can be used to better understand spatial transcriptomics data (**Fig. 4**). We believe that this approach has certain advantages over single-cell analyses. For example, all the analyses that revealed the internal architecture of Cancer Hallmarks within the cancer compartment depending on the distance to the TME (and vice versa) would not have been possible with single-cell data, as they are based on the physical distance between cells. Furthermore, this approach will be even more powerful as newer versions of spatial transcriptomics platforms increase their resolution in the near future. Other future directions of this research should also include a more detailed characterization of the histopathological images. For example, while we were limited in most cases due to the low resolution of some images (and the need to be consistent in the labeling across all datasets), our pathologist identified several areas in some images as having strong immune infiltration, which were also consistent with the gene expression patterns (and Cancer Hallmarks) that we identified in such compartments. Together, these developments should provide a more refined and detailed

landscape of the spatial architecture of Cancer Hallmarks. Finally, we believe that as the number of available spatial transcriptomics datasets increases, we might be able to use their Cancer Hallmark maps (**Fig. 4**) to understand how the spatial distribution of the activity of Cancer Hallmarks correlates with clinical outcomes, potentially becoming "spatial biomarkers" [20]. This would be even more powerful for cases where we have matching spatial transcriptomics from different timepoints, such as before and after cancer treatment, allowing us to understand how Cancer Hallmarks vary across time and space. Overall, we believe that charting the spatial landscape of Cancer Hallmarks holds great promise to better understand this disease and improve the care of cancer patients.

METHODS

Sample collection of high-grade serous carcinoma of the ovary

Fresh tissue HGSOC sample was collected during the primary surgery, after obtaining written informed consent from the patient, as required by the IEC of Hospital Germans Trias I Pujol. After the surgery, the patient received platinum-based chemotherapy and currently remains in complete response after the last cycle.

Tissue processing and spatial transcriptomics

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 -80ºC. For sectioning, blocks were equilibrated to -20ºC in the cryochamber (LEICA CM1950). 10 µm-thick sections were placed onto the active areas (6mm x 6mm) 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 H&E 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 minutes and cDNA Synthesis and amplification were 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 x total spots on the Capture Area) x 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.

Sequencing, data processing and filtering

Raw fastq files of HGSOC were processed with the “spaceranger” command line tool (version 1.2.2, 10x Genomics) and mapped to the pre-built human reference genome (GRCh38). Fifteen

publicly available VISIUM tumor datasets from seven tissues (**Fig. 1A**) were downloaded from a recent publication [13] (three liver leading-edge tumor samples; HCC-1L, cHC-1L, ICC-1L) 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 HGSOC and each downloaded ST dataset, the output from SpaceRanger was loaded individually in the STutility environment [21]. Three steps for filtering were performed in each of them. First, genes with less than 5 UMI were excluded. Second, mitochondrial and ribosomal genes were removed from the analysis, alongside the non-protein-coding genes. The gene type tags that were kept in this step: 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. Finally, spots were filtered out depending on two conditions: if they are out of the tissue or if the number of features is lower than a threshold value established by the distribution of each sample. Then, the resulting filtered samples were normalized by variance stabilizing transformation on the expression data using SCTransform [22] (return.only.var.genes = FALSE, variable.features.n = NULL, variable.features.rv.th = 1.1).

Transcription factor and pathway activity assessment

MYC transcription factor activity was assessed using DoRothEA and Viper R packages after filtering for high confidence regulons [23,24]. TGFB pathway activity was assessed using PROGENy [25].

Spatial enhancement and gene expression imputation with BayesSpace

The normalized expression data was converted to a SingleCellExperiment object with the respective spatial coordinates and metadata. In the preprocessing step (spatialPreprocess) [10], normalization was skipped and 15 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, the clustering was computed by spatialCluster using 50000 iterations, 15 PC, 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. The previous results from the clustering were used as initialization of the algorithm using 200000 iterations and 0.3 of prior jitter. The model xgboost with a maximum number of boosting iterations of 100 (nrounds = 100) computed the expression of the complete set of genes for each subspot. The final object was converted back to Seurat and the expression matrix was scaled (ScaleData) to perform downstream analysis.

NMF-based clustering and annotations

To decompose each dataset into five main components, NMF was used through RunNMF (STutility) on the scaled data of the enhanced objects. To show mutually exclusive clusters from these factors, the embeddings were scaled and centered for each factor. Then, the label annotated for each subspot was selected by the maximum scaled factor activity. Factors were then labeled by a pathologist; whereby histopathological annotations of the H&E image were used to label the factors. Assigning the labels of factors for invasive ductal carcinoma of the breast was decided based on the anti-CD3 and DAPI immunofluorescence images and the published histopathological annotations [10]. Likewise, labels were assigned to factors of endometrial adenocarcinoma of the ovary based on the anti-Cytokeratin, anti-CD45 and DAPI immunofluorescence images [10].

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 in their corresponding gene sets were discarded, yielding a total of 37,522 gene sets with a maximum of 425 genes per pathway. To assign genes for each of the 13 hallmarks, specific keywords [12] that are known or postulated to be associated with each hallmark were searched in the names of the pathways (**Table S3**), and the union of the corresponding gene sets was considered the genes associated with that hallmark. After excluding entries with "CHEBI" symbol corresponding to small chemical compounds, the total unique number of genes associated with all hallmarks was 5,244 (**Table S4**). For each hallmark-associated gene list, the percentage of overlapping genes with the normalized SCT data of each sample was measured (**Table S5**).

Scoring the hallmarks

A module score was computed for each hallmark by using AddModuleScore from Seurat [22] on the enhanced objects. To compare the hallmark profiles across factors, Pearson correlation was calculated from the scaled embeddings and scaled scores of the hallmarks. A 2D heatmap was generated using the 'pheatmap' R package [26], where the whole correlations were plotted and clustered using "correlation" as the distance method.

Quantification of hallmark activities in tumor subcompartments

TME and Cancer compartments were defined by the annotated factors, whereby the spots with factor activity higher than one standard deviation from the mean were selected, and spots belonging to both TME and Cancer were discarded. To establish adjacent and distant subcompartments within each cluster, we computed the Euclidean distance between the two compartments and used the top and bottom 25% minimum distance of each subspot to label them as "distant" and "adjacent",

respectively. The Wilcoxon test was used to assess if the mean difference of the scaled hallmarks scores between adjacent and distant subcompartments was different from zero. P-values were adjusted using Bonferroni, whereby comparisons were deemed significant if they had p-values less than 0.05/13 = 0.0038. Dotcharts and boxplots were generated with ggpubr [27], and circos plot with “circlize” R package [28].

Graphical figures

Individual panels with statistical data were generated with the R packages “ggplot2” and “pheatmap”. The different panels were put together using Illustrator. Panel 2D was generated with BioRender.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the kind and generous patients who donated the samples that made this study possible, as well as the scientists that generated the datasets for sharing them.

FUNDING

M. S. and S. C. are funded by the Asociación Española Contra el Cáncer (AECC) project LABAE20038PORT. E.P-P. is supported by a Ramon y Cajal fellowship from the Spanish Ministry of Science (RYC2019-026415-I) and is a member of the 10x Clinical Translational Research Network. J. B. is supported by the Grupo Español de Investigación en Cáncer de Ovario (GEICO) Translational Research Award "Andrés Poveda" 2021.

REFERENCES

1. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. American Association for Cancer Research; 2022;12: 31–46. doi:10.1158/2159-8290.CD-21-1059

2. Bailón-Moscoso N, Romero-Benavides JC, Ostrosky-Wegman P. Development of anticancer drugs based on the hallmarks of tumor cells. Tumour Biol. Tumour Biol; 2014;35: 3981–3995. doi:10.1007/S13277-014-1649-Y

3. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell. Elsevier; 2011;144: 646–674. doi:10.1016/J.CELL.2011.02.013/ATTACHMENT/3F528E16-8B3C-4D8D-8DE5-43E0C98D8475/MMC1.PDF

4. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, et al. The cancer genome atlas pan-cancer analysis project. Nature Genetics. 2013. doi:10.1038/ng.2764

5. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell. Cell Press; 2018;173: 371-385.e18. doi:10.1016/J.CELL.2018.02.060/ATTACHMENT/B0ABFAFA-A1FF-4F7D-8842-B49BA8D32E08/MMC1.XLSX

6. Lamy PJ, Fina F, Bascoul-Mollevi C, Laberenne AC, Martin PM, Ouafik L, et al. Quantification and clinical relevance of gene amplification at chromosome 17q12-q21 in human epidermal growth factor receptor 2-amplified breast cancers. Breast Cancer Res. BioMed Central; 2011;13: 1–12. doi:10.1186/BCR2824/FIGURES/2

7. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell. Elsevier B.V.; 2011;144: 27–40. doi:10.1016/J.CELL.2010.11.055/ATTACHMENT/F05F620E-45EC-41D3-81B9-1AFB409EE44F/MMC6.PDF

8. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. Science (80-). American Association for the Advancement of Science; 2017;357. doi:10.1126/SCIENCE.AAN2507/SUPPL_FILE/SUPPLEMENTARY-TABLES.ZIP

9. Rao A, Barkley D, França GS, Yanai I. Exploring tissue architecture using spatial transcriptomics. Nat 2021 5967871. Nature Publishing Group; 2021;596: 211–220. doi:10.1038/s41586-021-03634-9

10. Zhao E, Stone MR, Ren X, Guenthoer J, Smythe KS, Pulliam T, et al. Spatial transcriptomics at subspot resolution with BayesSpace. Nat Biotechnol 2021. Nature Publishing Group; 2021; 1–10. doi:10.1038/s41587-021-00935-2

11. Lin X, Boutros PC. Optimization and expansion of non-negative matrix factorization. BMC Bioinformatics. BioMed Central Ltd.; 2020;21: 1–10. doi:10.1186/S12859-019-3312-5/FIGURES/4

12. Iorio F, Garcia-Alonso L, Brammeld JS, Martincorena I, Wille DR, McDermott U, et al. Pathway-based dissection of the genomic heterogeneity of cancer hallmarks' acquisition with SLAPenrich. Sci Reports 2018 81. Nature Publishing Group; 2018;8: 1–16. doi:10.1038/s41598-018-25076-6

13. Wu R, Guo W, Qiu X, Wang S, Sui C, Lian Q, et al. Comprehensive analysis of spatial architecture in primary liver cancer. Sci Adv. American Association for the Advancement of Science; 2021;7: 3750. doi:10.1126/SCIADV.ABG3750/SUPPL_FILE/SCIADV.ABG3750_TABLES_S1_TO_S5.ZIP

14. Hankinson SE, Colditz GA, Willett WC. The lifelong interplay of genes, lifestyle, and hormones. Breast Cancer Res. BioMed Central; 2004;6: 213–218. doi:10.1186/BCR921/TABLES/1

15. He S, Sharpless NE. Senescence in Health and Disease. Cell. Cell Press; 2017;169: 1000–1011. doi:10.1016/J.CELL.2017.05.015

16. Wang B, Kohli J, Demaria M. Senescent Cells in Cancer Therapy: Friends or Foes? Trends in Cancer. Cell Press; 2020;6: 838–857. doi:10.1016/J.TRECAN.2020.05.004

17. Faget D V., Ren Q, Stewart SA. Unmasking senescence: context-dependent effects of SASP in cancer. Nat Rev Cancer 2019 198. Nature Publishing Group; 2019;19: 439–453. doi:10.1038/s41568-019-0156-2

18. Lee S, Schmitt CA. The dynamic nature of senescence in cancer. Nat Cell Biol 2019 211. Nature Publishing Group; 2019;21: 94–101. doi:10.1038/s41556-018-0249-2

19. Torres CM, Biran A, Burney MJ, Patel H, Henser-Brownhill T, Cohen AHS, et al. The linker histone H1.0 generates epigenetic and functional intratumor heterogeneity. Science (80-). American Association for the Advancement of Science; 2016;353. doi:10.1126/SCIENCE.AAF1644/SUPPL_FILE/MORALES-TORRES-SM.PDF

20. Nirmal AJ, Maliga Z, Vallius T, Quattrochi B, Chen AA, Jacobson CA, et al. The spatial landscape of progression and immunoediting in primary melanoma at single cell resolutionAtlas of Primary Melanoma. Cancer Discov. 2022; doi:10.1158/2159-8290.CD-21-1357

21. Bergenstråhle J, Larsson L, Lundeberg J. Seamless integration of image and molecular analysis for spatial transcriptomics workflows. BMC Genomics. BioMed Central; 2020;21: 1–7. doi:10.1186/S12864-020-06832-3/FIGURES/2

22. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive Integration of Single-Cell Data. Cell. Cell Press; 2019;177: 1888-1902.e21. doi:10.1016/J.CELL.2019.05.031/ATTACHMENT/2F8B9EBE-54E6-43EB-9EF2-949B6BDA8BA2/MMC3.PDF

23. Garcia-Alonso L, Holland CH, Ibrahim MM, Turei D, Saez-Rodriguez J. Benchmark and integration of resources for the estimation of human transcription factor activities. Genome Res. 2019; doi:10.1101/gr.240663.118

24. Alvarez MJ, Shen Y, Giorgi FM, Lachmann A, Ding BB, Hilda Ye B, et al. Functional characterization of somatic mutations in cancer using network-based inference of protein activity. Nat Genet. 2016; doi:10.1038/ng.3593

25. Schubert M, Klinger B, Klünemann M, Sieber A, Uhlitz F, Sauer S, et al. Perturbation-response genes reveal signaling footprints in cancer gene expression. Nat Commun 2017 91. Nature Publishing Group; 2018;9: 1–11. doi:10.1038/s41467-017-02391-6

26. Kolde R. pheatmap□: Pretty Heatmaps. R Packag version 1012 https//CRANR-project.org/package=pheatmap. 2019;

27. Kassambara A. ggpubr: “ggplot2” Based Publication Ready Plots. R package version 0.4.0. 2020;

28. Z G, L G, R E, M S, B B. circlize Implements and enhances circular visualization in R. Bioinformatics. Bioinformatics; 2014;30: 2811–2812. doi:10.1093/BIOINFORMATICS/BTU393