Spatial Profiling of the Immune Microenvironment by Multiplex Imaging Identifies Barriers to Immunotherapy in Colorectal Liver Metastases

Ranchu Cheng

Wolfson College

Department of Oncology

University of Oxford

This thesis is submitted for the degree of

MSc by Research

August 2025

Table of Contents

Abstract

Colorectal liver metastasis (CRLM) is the most frequent site of distant metastasis in colorectal cancer (CRC), affecting around 50% of CRC patients. Only 10–30% of CRLM patients were eligible for curative resection, with a low 5-year survival rate (35–55%) and a high recurrence rate (70%). Notably, most CRLMs are resistant to immune checkpoint blockade, prompting a need to map how immune failure is organised in space. In this study, we profiled the CRLM tumour microenvironment across tumour core, peritumour region, and healthy background liver using two highly multiplexed, single-cell datasets: Cell DIVE^{TM} spatial proteomics (13 patients; 1,617 500µm × 500µm regions of interest (ROIs); 2.9 M cells) and CosMx^{TM} 6k spatial transcriptomics (4 patients; 1,233 ROIs; 1.34 M cells). Our multiscale pipelines quantified cell abundances, defined neighbourhood clusters from ten-nearest neighbours, tested short-range interactions with permutation-controlled proximity statistics, resolved distance-dependent associations with cross-pair correlation functions (cross-PCF), and captured higher-order geometry with persistent homology-based topological data analysis. Inference was performed at the slide level with multiplicity control. Across cohorts, the peritumour rim concentrated adaptive and antigen-presenting programmes. Short-range coupling of CD8^{+} T-cells with conventional dendritic cells, together with CD4–dendritic, CD4–B, and CD4–CD8 proximities, demonstrated the formation of tertiary lymphoid structure (TLS)-like hubs at the tumour border. CosMx identified a CCL19/CCL21-expressing fibroblast subset, consistent with fibroblastic reticular cell-like stromal support, while Cell DIVE confirmed higher frequencies of both activated and exhausted T cells at the rim, indicating active yet regulated immunity. The tumour core exhibited layered suppression. Cell DIVE revealed adjacency between T-cells and aSMA^{+}/FAP^{+} cancer-associated fibroblasts (CAF). CosMx resolved extracellular matrix remodelling fibroblasts, enriched regulatory T-cells, and abundant myeloid programmes, which together tracked with exclusion of cytotoxic effectors. Cross-PCFs showed reductions in TLS couplings within the tumour core and reinforcements of CAF to immune cell proximities. Topological data analysis separated regions using component and loop statistics, highlighting potential fragmented T-cell geometry interleaved with stromal interfaces at the peritumour region and interface-dominated topology with granulocytic admixture in the tumour core. The healthy liver displayed simple parenchymal vascular segregation consistent with baseline tolerance. These spatial maps nominate promising therapeutic strategies to mitigate immunotherapy resistance. Peritumour-focused strategies include expanding and licensing cross-priming dendritic cells with CD40 agonism and increasing dendritic cell supply at the border. While tumour core-focused approaches include reducing immunosuppressive activities/secretions from CAF and T-regs, and targeting immunosuppressive myeloid cells.

Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed cancer and the second leading cause of cancer-related mortality worldwide, with over 1.9 million new cases and 935,000 deaths annually (Sung *et al.,* 2021). The liver is the most common site of distant spread, affecting up to 50% of CRC patients (Tsilimigras *et al.,* 2021). Approximately 15-25% of patients present with synchronous liver metastases at the time of initial diagnosis, while another 18–25% of patients will develop metachronous metastases within 5 years of diagnosis (Wang *et al.,* 2023a; van der Geest *et al.,* 2015). This predilection is largely due to the liver's unique anatomy, blood flow, and intrinsic immune-

tolerant environment. As the first organ to filter blood draining from the colon via the portal vein, it acts as the primary recipient for circulating tumour cells (Vidal-Vanaclocha, 2008). This "first-pass" effect is compounded by the liver's internal structure, where exceptionally slow blood flow and highly permeable sinusoidal vessels result in a prolonged transit time, maximising the opportunity for cancer cells to interact with the sinusoidal endothelium and thereby promoting their extravasation into the liver parenchyma (Mielgo and Schmid, 2020; Liu *et al.,* 2023).

In fact, successful metastasis is an inefficient, multi-step process known as the metastatic cascade, and arrest in the liver is only one part of the journey (Tsilimigras *et al.,* 2021). This cascade begins with intravasation, where cancer cells invade local blood vessels at the primary tumour site. After surviving transit through the circulatory system, these cells must then undergo extravasation to exit the bloodstream at the liver, which represents a fertile "soil" with rich blood supply for cancer "seeds" (Paget, 1989). To grow beyond microscopic clusters, these micro-metastases must induce the formation of new blood vessels, also known as angiogenesis. The final and rate-limiting step is colonisation, where disseminated cancer cells must adapt to the foreign liver microenvironment to proliferate and form clinically detectable macro-metastasis (Fares *et al.*, 2020; Shasha *et al.,* 2022).

Clinically, the presence of colorectal liver metastases (CRLM) is the primary determinant of prognosis for patients with advanced CRC. The current standard of care for CRLM is a multimodal strategy centred on surgical resection, which remains the only potentially curative option. However, only around 10–20% of CRLM patients are eligible for upfront liver resection due to the number, size, or location of their tumours (De Greef *et al.,* 2016). For most patients with initially unresectable disease, systemic chemotherapy, often combined with targeted agents, has become the initial treatment option (Bolhuis *et al.,* 2020; De Greef *et al.,* 2016). Standard-of-care chemotherapy typically consists of 5-Fluorouracil and Leucovorin, plus either Irinotecan (FOLFIRI) or Oxaliplatin (FOLFOX), to control disease and, in some cases, as neoadjuvant treatment to downsize tumours for potential conversion surgery, allowing up to 30–40% of CRLM patients to achieve complete liver resection (Van Cutsem et al., 2010; Moretto *et al.,* 2024; De Greef *et al.,* 2016). However, a phase III trial, which compared perioperative FOLFOX4 chemotherapy to surgery alone in patients with resectable CRLM, showed that although perioperative chemotherapy significantly improved progression-free survival, it did not translate into a long-term overall survival benefit. Additionally, the risk of postoperative complications was notably higher in the chemotherapy group, particularly with prolonged perioperative treatment (Nordlinger *et al.,* 2013). The addition of biologic agents targeting vascular endothelial growth factor (VEGF), such as Bevacizumab, or epidermal growth factor receptor (EGFR) inhibitors in RAS wild-type tumours, such as Cetuximab and Panitumumab, has also modestly increased response rates and enabled a subset of patients with unresectable disease to become candidates for potentially curative resection (Van Cutsem *et al.,* 2010; De Greef *et al.,* 2016). However, the overall impact of these targeted agents on conversion rates is constrained by significant limitations and inconsistent trial results. Due to potential wound healing risks and complications, the use of perioperative Bevacizumab (anti-VEGF) may be limited (Tamiya *et al.,* 2009; Gordon *et al.,* 2009). While some small studies have shown promising conversion rates, a key phase III trial found that the addition of Bevacizumab to Oxaliplatin-based chemotherapy resulted in only a marginal,

non-statistically significant increase in complete (R0) resections compared to chemotherapy alone (6.3% vs 4.9%; P=0.24) (Okines *et al.,* 2009). The efficacy of anti-EGFR agents is also limited, being restricted to *RAS wild-type* patients and showing highly inconsistent results (De Greef *et al.,* 2016). While trials like OPUS and CRYSTAL reported a doubling of R0 resections with Cetuximab, the absolute rates remained low (<8%) (Bokemeyer *et al.,* 2009; Van Cutsem *et al.,* 2011). This contrasts sharply with the approximately 40% resection rate reported in studies such as CELIM, a difference likely attributable to specialist surgical assessment rather than drug efficacy alone (Folprecht *et al.,* 2010; De Greef *et al.,* 2016). Moreover, the entire strategy was questioned by large trials like COIN and NORDIC VII, which demonstrated no benefit from adding Cetuximab to Oxaliplatin-based chemotherapy, regardless of KRAS status (Maughan *et al.,* 2011; Tveit *et al.,* 2012).

Hence, despite significant advances in multimodality treatment, the five-year overall survival rate for patients with CRLM remains modest, ranging from 30% to 50% after curative-intent liver resection (Tsilimigras *et al.,* 2021). Furthermore, recurrence is common in CRLM, with nearly 70% of patients relapsing within two years of surgery (Buisman *et al.,* 2020; Vadisetti *et al.,* 2024). The poor prognosis of CRLM underscores a critical need to expand therapeutic strategies beyond conventional chemotherapy and mutation-targeted treatments to improve long-term outcomes for this large patient population.

Immunotherapy has gained prominence as an effective therapeutic strategy in recent years. The advent of immune checkpoint inhibitors (ICIs), particularly antibodies targeting PD-1/PD-L1 and CTLA-4, has revolutionised the treatment of numerous malignancies by restoring anti-tumour T-cell activities (Esfahani *et al.,* 2020). However, their success in metastatic CRC is largely confined to a small subset of metastatic CRC cases (around 5%), characterised by mismatch repair deficiency or high microsatellite instability tumours (Chen *et al.,* 2023a). The vast majority of metastatic CRCs are mismatch repair proficient or microsatellite stable, and these subtypes are known to be resistant to ICI, whether administered as monotherapy or in combination regimens (Chen *et al.,* 2020; Eng *et al.,* 2019). More importantly, emerging evidence suggests the liver's unique immune environment promotes CRLM's resistance to ICIs. For instance, one phase II trial highlighted a significant contrast in response rates between patients without liver metastases (21.7%) and those with liver metastases (0%) (Fakih *et al.,* 2021). Further supporting this, a secondary analysis of the CCTG CO.26 study found that in patients receiving durvalumab (anti-PD-L1) plus Tremelimumab (anti-CTLA-4), the presence of liver metastases was associated with significantly worse outcomes, including a lower disease control rate (14% vs 49%) and shorter progression-free survival (1.82 vs 2.04 months) (Chen *et al.,* 2020).

The failure of immunotherapy in CRLM is largely due to the complex cellular ecosystem of the tumour microenvironment (TME), which actively orchestrates immune evasion. The TME is a dynamic landscape of non-malignant cells, including a diverse group of immune cells, stromal cells like cancer-associated fibroblasts (CAFs), and endothelial cells, all embedded within an extracellular matrix (ECM) (Xiao *et al.,* 2021). These components engage in a complex crosstalk that can either foster anti-tumour immunity or, more commonly in CRLM, create a profoundly immunosuppressive barrier that renders therapies like ICIs ineffective (Binnewies *et al.,* 2018; Hinshaw and Shevde, 2019). For

instance, an abundance of regulatory T-cells (T-regs) and M2-polarised macrophages can directly inhibit the function of cytotoxic $CD8^+$ T-cells, while dense stromal networks built by CAFs can physically exclude immune cells from the tumour core and further suppress them through secreted factors (Mariathasan *et al.,* 2018; Kalluri, 2016). Meanwhile, the barrier is reinforced by the liver's natural disposition as an immune-tolerant organ, a feature necessary to process the constant stream of antigens it receives from the gut. The local tolerance is actively maintained by specialised antigen-presenting cells (APCs) that regulate immune responses and maintain homeostasis (Thomson and Knolle, 2010; Dou *et al.,* 2018; Heymann *et al.,* 2015). Understanding the complex cellular interaction pattern is therefore paramount to deciphering the mechanisms of resistance and designing effective new immunotherapies for CRLM.

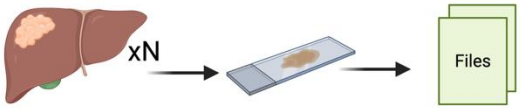
Initial efforts to characterise key players within the CRLM TME relied on techniques such as bulk sequencing, flow cytometry, and single-cell RNA sequencing (Qiao *et al.,* 2023; Che *et al.,* 2021; Geng *et al.,* 2022). However, a fundamental limitation of these dissociation-based methods is that they destroy the native spatial context of the tissue. Without spatial information, potential architectural motifs (e.g., local cell-cell interaction patterns) that underpin immune exclusion, suppressive niche formation, and therapeutic failure could be masked. To overcome the limitation, spatial biology has emerged as a transformative approach, enabling the profiling of cells *in situ* while preserving their coordinates (Schürch *et al.,* 2020; Pelka *et al.,* 2021; Keren *et al.,* 2018; Angelo *et al.,* 2014; Gerdes *et al.,* 2013; Stachtea *et al.,* 2022; Xiao *et al.,* 2024; Nirmal *et al.,* 2022; Salachan *et al.,* 2023). Early spatial studies in primary colorectal cancer have demonstrated the power of this approach. Using CODEX, Schürch et al. (2020) identified multiple recurring "cellular neighbourhoods" (i.e., multicellular communities that serve as the building blocks of the tissue) within CRC and showed that the degree of spatial organisation within or between "cellular neighbourhoods" at the invasive margin is a powerful predictor of patient survival. Similarly, by integrating single-cell and spatial transcriptomics in primary CRC, Xiao et al. (2024) were able to reveal intensive crosstalk between adjacent tumour and stromal regions, allowing for the inference of key molecular interactions, such as the C5AR1-RPS19 ligand-receptor pair. While these works provided a foundational blueprint for CRC TME spatial biology, the specific spatial rules governing the unique TME of liver metastases remain less defined. Our study hence aimed to build upon prior spatial analysis pipelines to specifically dissect the spatial logic of immune failure in the liver.

This study introduces a comprehensive analytical pipeline that integrates multiple bioinformatic tools to dissect the spatial arrangements of CRLM TME components in different tissue regions: the tumour core, the peritumoural margin, and the adjacent healthy liver. Meanwhile, by analysing the TME at a whole slide/tissue level and separating the tissue into defined spatial compartments (tumour core, peritumour, and background liver), we can make inferences about how the tumour has overcome the liver's immune response. We hypothesised that TME spatial metrics would trace a directional transition from healthy liver to peritumour to tumour, with the peritumour region showing intense immune response and immune–stromal crosstalk. Beyond conventional cell count calculations and cellular neighbourhood identifications, we also quantified spatial organisation across a continuum of biological scales. This includes using permutation-based proximity tests and cross-pair correlation

functions (cross-PCF) to statistically define significant short- and medium-range interactions, revealing the distance-dependent "attraction" or "exclusion" between every cell type (Bull *et al.,* 2025; Nirmal *et al.,* 2024). These proximity metrics also identify scale-specific couplings that sometimes may be obscured in general neighbourhood summaries, which are calculated by averaging across all regions. Moreover, we introduced topological data analysis (TDA) as a novel approach to capture higher-order, multicellular spatial patterns, allowing us to mathematically describe the "shape" of cellular communities (Bull *et al.,* 2024). This multi-scale approach provided a rich, quantitative language to describe tissue structure, enabling us to identify key architectural motifs that define the tumour, peritumour, and healthy liver compartments. To ensure our findings are robust and comprehensive, we applied the entire pipeline to two distinct, highly multiplexed spatial datasets. The first is a 13-patient spatial proteomics cohort (Cell DIVE™), which provided protein-level data on cell identity and functional state (e.g., proliferation, exhaustion) with the statistical power of a larger cohort (Gerdes *et al.,* 2013). The second is a 4-patient spatial transcriptomics cohort (NanoString CosMx™), which offered better phenotypic depth with its 6,000-gene panel, allowing for the fine-grained resolution of cell subtypes (He *et al.,* 2022a). By integrating diverse bioinformatics, biomathematical, and statistical approaches, we aimed to exploit spatial data to characterise interactions among key components of the CRLM TME, particularly immune and fibroblastic cells, across distinct tissue regions, and to build a spatial atlas of architectural barriers to immunity to inform rational therapeutic targeting and sensitise resistant tumours to immunotherapy.

Methodology

Sample Information

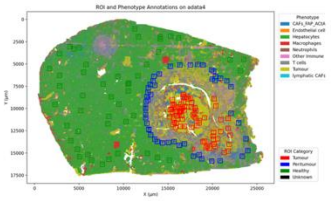


- **Cell DIVE** spatial proteomics cohort: 13 FFPE whole-slide CRLM samples.
- **CosMx** spatial transcriptomics cohort: 4 FFPE whole-slide CRLM samples.

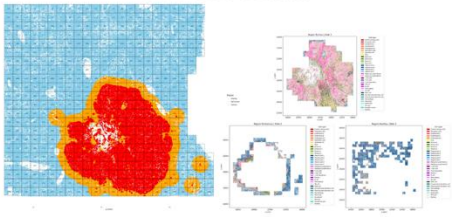
1. Region Annotations, ROI Segmentation, ROI Selection and Quality Control

(500 µm × 500 µm) ROIs from Tumour, Peritumour, Healthy regions.

Cell DIVE cohort:
1617 ROIs in total.



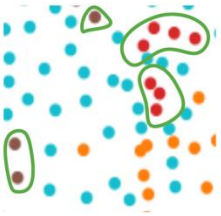
CosMx cohort:
1233 ROIs in total.



2. Data Analysis (Python Based Pipelines)

2.1 Normalised Cell Count

2.1.1 Normalised cell count per ROI based on all cells in a ROI

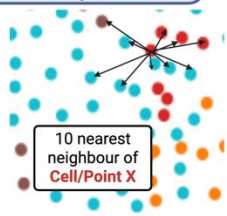


- *Normalised Fraction of cell type **x** = No. of cells of phenotype **x** in the ROI / No. of **ALL cells** in the ROI.*
- For each region, ROI-level metrics were averaged per slide to generate slide-level metrics.
- Region-wise comparisons: Benjamini–Hochberg FDR-corrected paired Wilcoxon signed-rank tests.

2.1.2 Normalised cell count of 'child' cell types per ROI based on cells belonging to the corresponding 'parent' broad cell type in a ROI

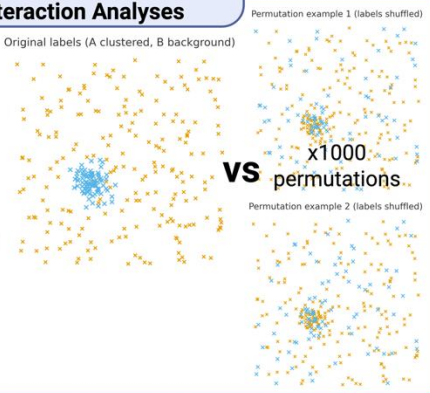
- *Normalised Fraction of cell type **x** = No. of cells of phenotype **x** in the ROI / No. of all cells with **broad type B** in the ROI.*
- Slide-level metrics: restricted to ROIs with sufficient infiltration of a given broad type (i.e., > healthy-tissue–derived cellularity baseline).
- Region-wise comparisons: Benjamini–Hochberg FDR-corrected paired Wilcoxon signed-rank tests.

2.2 Cellular Neighbourhood Cluster Identification and Comparison



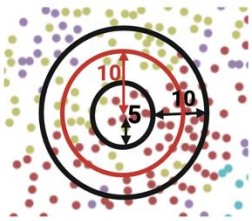
- K-nearest neighbour/KNN (k=10).
- K-means clustering.
- *Proportion of NC x = No. of cells labelled as NC x in the ROI / No. of all cells in the ROI.*
- For each region, ROI-level metrics were averaged per slide to generate slide-level metrics.
- Region-wise comparisons: Benjamini–Hochberg FDR-corrected paired Wilcoxon signed-rank tests.

2.3 Permutation-Controlled, Spatial Proximity-based Cell-Cell Spatial Interaction Analyses



- For each cell type pair (**cell-type *i*, neighbour-type *j***): compare the observed 10-nearest neighbour frequency to a null distribution obtained from 1 000 random permutations of cell-type labels, yielding a Z-score and an exact permutation p-value for every slide.
- CellDIVE dataset: interactions between cell types were only considered significant when the p-value < 0.05 in at least half of the slides from that region (n_slides = 7).
- CosMx dataset: p-value < 0.05 in all slides from that region (n_slides = 4).

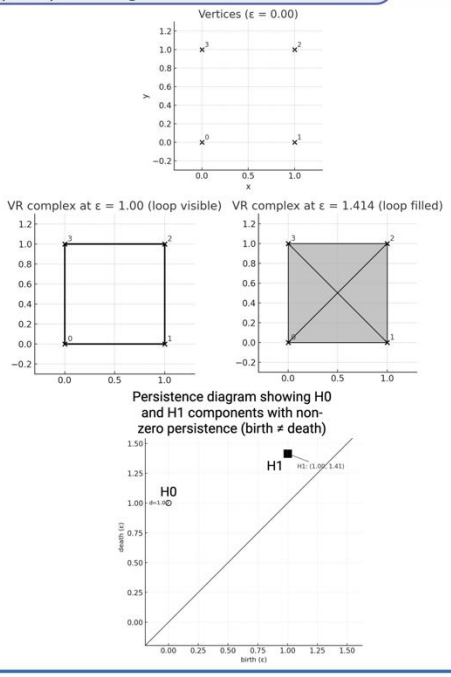
2.4 Cross-Pair Correlation Functions-based Cell-Cell Spatial Interaction Analyses



Bins to define point j 'neighbours' of point x: 0-10µm, 5-15µm , 10-20µm...

- *Cross-PCF between **cell-type i and j**, $g_{ij}(r)$: a ratio describing whether the observed No. of **i-j pairs of points** separated by distance R is higher or lower than would be expected under a statistical null model.*
- For each region, ROI-level metrics were averaged per slide to generate slide-level metrics.
- Max_R=250 µm, Annulus_Step=5 µm, Annulus_Width=10 µm.

2.5 Topological Data Analysis (TDA) of Neighbourhood Clusters



- Measured spatial organisation of neighbourhood clusters using Vietoris–Rips filtration based persistent homology on ROI point clouds.
- Two modes for point clouds build-up per ROI
  a. Single-cluster: centroids of all cells in one neighbourhood cluster.
  b. Pairwise-cluster: centroids from an unordered cluster pair combined into one set.
- ***H0 features: connected components, H1 features: holes in the surface.***
- Each feature barcode was vectorised into metrics. Extracted only "significant bars" by thresholding at ≥50 µm for H_0 and ≥10 µm for H_1.
- Region-wise discriminations: Random forests (1,000 trees) with ROI-level metrics as inputs + MDI importance ranked features.
- For each region, ROI-level metrics were averaged per slide to generate slide-level metrics.
- Region-wise comparisons: Benjamini–Hochberg FDR-corrected paired Wilcoxon signed-rank tests.

Figure 1. Overview of the methods used in this study. Text: simplified method summary. Figures: simplified (mathematical) model concepts/visualisation.

1. Data Collection and Source Description

The Cell DIVE spatial proteomics dataset comprised thirteen single-cell–resolved datasets, each generated from a formalin-fixed, paraffin-embedded (FFPE) whole-slide CRLM sample from a distinct patient who had received neoadjuvant chemotherapy followed by liver resection. All tissue samples were obtained from the pathology archive with appropriate ethical approval (REC reference: 21/YH/0206, IRAS project ID: 301652). Patient demographic data are provided in Table 1. Raw multiplexed spatial proteomic images were processed and annotated by a DPhil student and a junior clinician. Cell phenotypes were assigned based on expression across 19 protein marker channels, yielding 23 distinct phenotypes. Using each cell's spatial coordinates and the corresponding scatter plot of the whole-slide image, multiple 500 µm × 500 µm regions of interest (ROI) were manually extracted from each slide, and a region label (Tumour, Peritumour, or Healthy) was assigned to each ROI by the junior clinician. In total, the curated Cell DIVE dataset included 1,617 ROIs profiling 2,923,140 cells. Although extraction and annotation were manual, ROI selection was guided by biological intent: all 572 Tumour ROIs were chosen to reflect tumour cores; the 513 Peritumour ROIs were focused on invasive margins and stromal interfaces, defined as regions approximately 200–500 µm from tumour boundaries; and the 532 Healthy ROIs were selected to represent the spatial organisation of background liver. An example of ROI extraction was shown in Fig.2, and the final per-sample ROI counts by region are summarized in Table.2A. Each row of a ROI-level dataset represented one single cell, and the dataset contained 7 informative metadata: the x and y coordinates of the cell (in µm), cell type (more granular phenotype), broad type (more generalised phenotype), region label of the ROI, ROI identity number (ID), and slide ID.

Age	Gender	Date of resection	number of metastases	Size of Largest metastasis (mm)	Margin (mm)	Histology	Neoadjuvant chemotherapy	Regimen 1	Biological agent	KRAS mutation status	Synchronous colon cancer	Primary colorectal cancer site	Primary Nodal Stage	Dissemination	Liver site	Recurrence site	Death	Date Last Follow-up
72	M	25/04/2012	1	100	15	CRLM	0	N/A	none	N/A	synchronous	left	x	liver only	multiple segments same side	liver and lung	1	21/08/2012
64	M	10/10/2012	4	15	1	CRLM	0	N/A	none	mutant	synchronous	left	1	liver only	bilobar disease	liver and lung	1	05/06/2013
67	F	26/04/2013	1	45	4	CRLM	0	none	none	N/A	metachronous	left	2	liver only	single segment	none	0	19/01/2021
69	F	17/10/2012	3	22	1	CRLM	0	N/A	none	N/A	metachronous	left	0	liver only	bilobar disease	lung	0	21/10/2020
44	M	12/06/2012	3	30	1	CRLM	0	N/A	none	N/A	metachronous	left	2	liver only	multiple segments same side	none	0	15/12/2020
63	F	01/06/2012	1	64	5	CRLM	0	N/A	none	mutant	synchronous	Left	0	liver only	single segment	none	0	05/10/2019
70	F	29/04/2014	1	40	2	CRLM	1	FOLFOX	none	N/A	metachronous	left	1	liver only	single segment	liver and bone	1	12/03/2015
66	M	12/06/2013	7	18	0	CRLM	1	FOLFOX	cetuximab	wild-type	synchronous	right	1	liver only	bilobar disease	none	0	01/03/2021
58	M	11/06/2013	1	35	40	CRLM	1	CAPOX	none	N/A	synchronous	left	2	liver only	bilobar disease	none	0	18/02/2021
54	F	23/05/2013	1	30	4	CRLM	0	none	none	mutant	synchronous	left	2	liver only	bilobar disease	none	0	06/02/2018
72	M	20/03/2013	2	15	0	CRLM	1	FOLFOX	none	N/A	synchronous	left	1	liver and lung	bilobar disease	liver	1	25/04/2018
63	M	06/11/2013	1	75	18	CRLM	1	FOLFOX	none	N/A	synchronous	left	0	liver only	multiple segments same side	liver	1	01/09/2014
76	M	06/11/2013	4	45	3	CRLM	1	FOLFIRI	cetuximab	wild-type	metachronous	left	1	liver only	bilobar disease	Liver and lung and bone	1	03/04/2014

Table 1. Patient (Slide) demographic data for the Cell DIVE dataset.

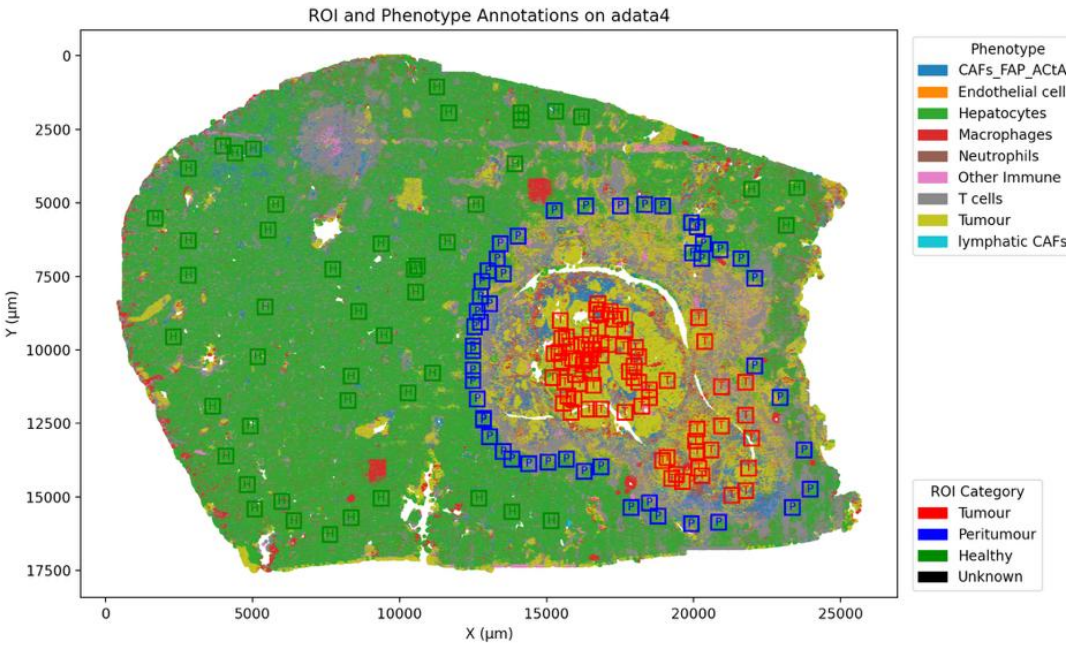


Figure 2. Cell DIVE dataset ROI generation and annotation. *This figure is an example of manual ROI extraction and annotation from a slide of the Cell DIVE dataset (slide 4). The size of all squared ROIs was 500 µm × 500 µm. Blue ROIs were annotated as peritumour ROIs, Red ROIs were annotated as tumour core ROIs, and Green ROIs were annotated as healthy liver ROIs. Note: this image was generated by other members of the lab.*

The CosMx 6K-panel (Nanostring) spatial transcriptomics dataset comprised four single-cell-resolved datasets, each derived from a whole-slide CRLM sample from a different patient. All samples were FFPE tissues obtained from the pathology archive, with the same ethical approval (REC reference: 21/YH/0206, IRAS project ID: 301652) as the Cell DIVE dataset. Patient demographic data was not available here. CosMx™ Human 6k spatial transcriptomics molecular imaging data were processed and annotated by a trained postdoctoral researcher using a workflow designed for multiplexed in situ transcriptomic profiling. Transcript-level flat files generated by the CosMx pipeline were processed to produce a gene expression matrix. Quality control involved removing cells with low gene counts and segmentation errors, and removing outlier ROIs with distorted expression profiles. Gene counts were normalised using total counts normalisation. Dimensionality reduction was performed with principal component analysis, followed by Harmony for integration across slides. Cell types were annotated using supervised clustering, guided by an integrated reference matrix derived from CRC and liver tissue. Cluster assignments were refined by validating canonical marker gene expression. After cell typing, spatial regions were defined by classifying cells according to their Euclidean distance from the nearest malignant epithelial cell (≤50 µm = Tumour; 50–500 µm = Peritumour; >500 µm = Healthy) (Fig.3A). 35 distinct phenotypes were identified and each whole-slide-level dataset was subdivided into 500 µm × 500 µm ROI-level datasets (Fig.3B). Because tumour borders are often irregular and some tumour areas are highly fibrotic, the initial strategy of retaining only square-shaped ROIs in which all cells carried a single region label yielded a small subset of ROIs and introduced substantial selection bias by underrepresenting heterogeneous border and fibrotic zones. To mitigate this, we implemented a refined ROI-level region assignment: ROIs in which more than 60% of cells were originally labelled as Tumour were designated as Tumour ROIs, and all cells within them were relabelled as Tumour. For the Peritumour region, we included (i) ROIs containing only Tumour

and Peritumour-labelled cells in which ≥40% of cells were Peritumour, (ii) ROIs containing only Peritumour and Healthy-labelled cells in which >50% of cells were Peritumour, (iii) ROIs containing only Peritumour cells. In addition, 38 ROIs that initially contained cells from all three regions-but had only minor contributions from Tumour and Healthy labels-were manually annotated as Peritumour. To avoid overrepresentation of the Healthy region and to better balance ROI counts across regions, we randomly subsampled 50% of the ROIs in which all cells were originally labelled as Healthy. This distance-based approach captured the spatial organisation of the tissue, distinguishing the tumour core, its peripheral microenvironment, and the more distant, potentially unaffected tissue. In total, the curated dataset included 1,233 ROIs profiling 1,339,397 cells: 464 Tumour ROIs, 327 Peritumour ROIs, and 442 Healthy ROIs. Figure 2B showed how each region dataset looked for one example whole-slide, and the final per-sample ROI counts by region are shown in Table.2B.

A. Cell DIVE ROI Counts

Sample (slide ID)	Tumour	Peritumour	Healthy	Total
0	35	48	39	122
1	41	67	26	134
3	10	28	18	56
4	72	49	48	169
5	19	29	54	102
8	93	37	62	192
9	41	41	56	138
10	69	63	44	176
11	37	42	37	116
12	52	30	27	109
13	33	33	57	123
14	16	24	44	84
16	54	22	20	96

B. CosMx ROI Counts

Sample (slide ID)	Tumour	Peritumour	Healthy	Total
1	144	84	83	311
2	127	67	91	285
3	109	49	119	277
4	62	127	171	360

Table 2. ROI count summaries. A. Summary of per-sample ROI counts by region of the Cell DIVE cohort, 13 samples in total. B. Summary of per-sample ROI counts by region of the CosMx cohort, 4 samples in total.

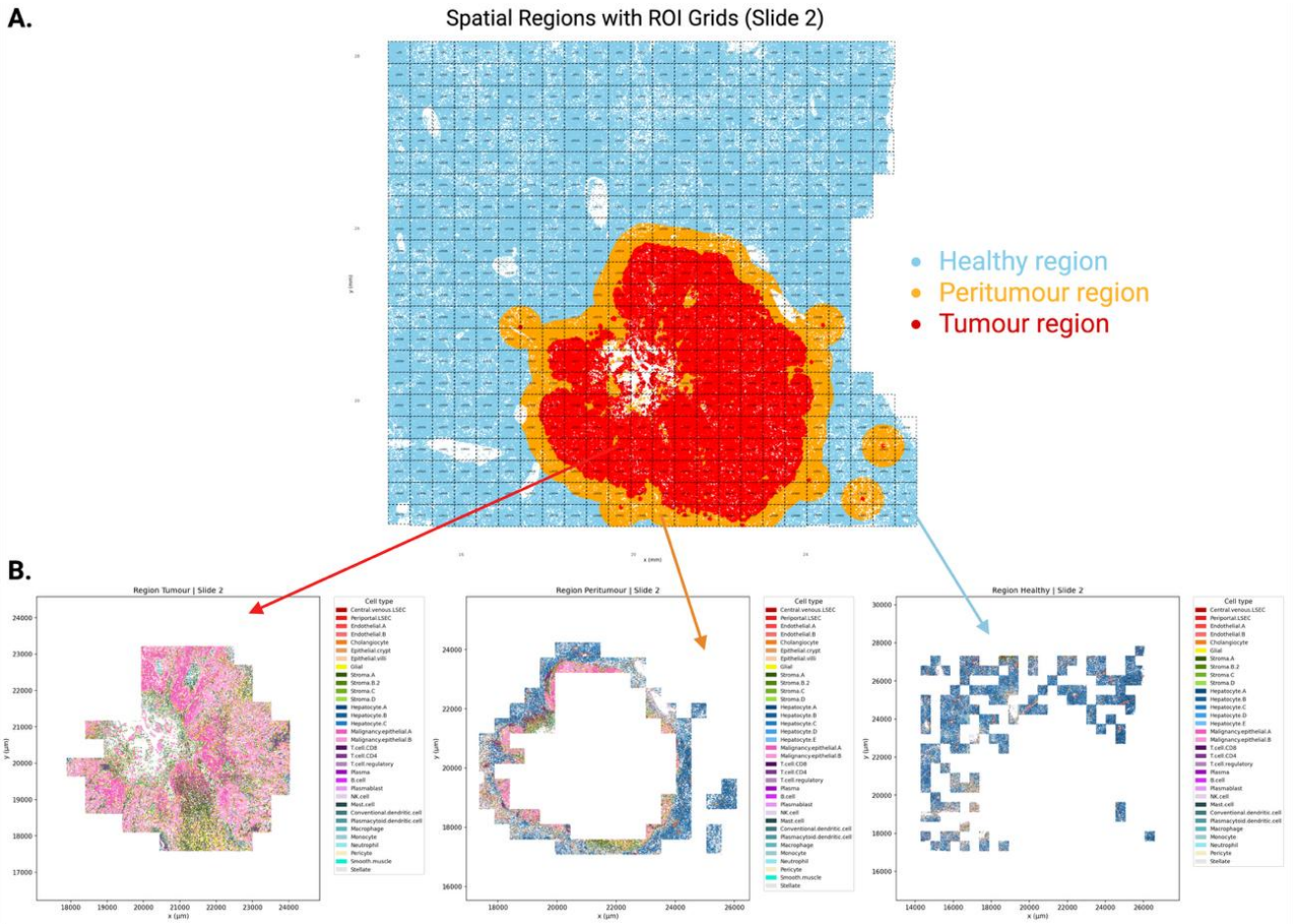


Figure 3. CosMx dataset ROI generation, annotation, and separation. A. An example of ROI generation and annotation from a slide of the CosMx dataset (slide 2). Each whole-slide image was subdivided into multiple 500 µm × 500 µm ROIs (black grids). Red: tumour core region, Orange: peritumour region, Light blue: healthy liver region. B. Scattered plots showing three different regional datasets for slide 2. Each regional dataset contained multiple ROIs, and each ROI contained multiple cells and their spatial coordinates.

2. Measurement of Spatial Metrics and Statistical Analysis

Most spatial analyses used individual ROIs as the primary unit, with regional datasets derived by aggregating ROI-level metrics. The only exception was the permutation-based inference of cell–cell interactions using Scimap (v2.3.4), which takes all ROIs from the same region on a given slide as input and automatically computes slide-level averages. From the spatially resolved point clouds defined by the coordinates of cells within each ROI, we compared the three tissue regions using five complementary approaches: normalised cell counts, cellular neighbourhood proportions, cell–cell correlations, cross-pair correlation functions, and persistent homology based on Vietoris–Rips filtrations.

The same analytical pipeline was applied to both the Cell DIVE and CosMx datasets independently, enabling us to validate between the two datasets. Minor modifications are detailed explicitly in the corresponding sections below.

2.1 Normalised Cell Count

2.1.1 Normalise Cell Count per ROI based on All Cells in a ROI

To enable quantitative comparison of cell phenotype abundances across heterogeneous tissue microenvironments, we normalised cell counts for each cell type within every ROI to the total number of cells in that ROI. This per-ROI normalisation yielded unitless fractions that were robust to variation in cellularity and ROI size.

Normalised Fraction of cell type X = the number of cells of phenotype X in the ROI / the number of all cells in the ROI.

For each region (Tumour, Peritumour, Healthy), all per-ROI CSVs were processed using a custom Python pipeline, and ROIs were parsed one at a time to compute normalised fractions. Metadata (region, slide, ROI ID) and all per-phenotype normalised counts were written to individual CSV files for traceability. All ROI-level normalised cell count calculations were performed using the in-house muspan Python package (v1.1.1) (Bull *et al.,* 2025), which enabled domain construction, cell label assignment, and summary statistics extraction. Then, for each region, all per-ROI normalised count CSVs were aggregated, and missing cell type columns were filled with zeros to accommodate rare phenotypes.

We performed statistical comparisons between region pairs for the normalised cell count of targeted cell types. A conservative ROI inclusion criterion was applied in 2.1.1, where only ROIs containing at least one cell from any of the targeted phenotypes were included in downstream analyses. This approach would potentially prevent inflation of variance and potential selection bias in region-wise comparisons. For region-wise comparisons, non-parametric paired Wilcoxon signed-rank tests were performed across slides (patients) for each targeted cell type and region pair (Healthy vs Peritumour, Healthy vs Tumour, Peritumour vs Tumour) (Hollander *et al.,* 2013). Multiple testing was addressed using the Benjamini–Hochberg procedure, with a false discovery rate (FDR) threshold of 0.05 (Benjamini *et al.,* 1995). All statistical analyses were conducted at the slide level (N=13 for Cell DIVE, and N=4 for CosMx), with boxplots and volcano plots generated to visualise phenotype distributions and differential abundance between regions, including adjusted p-values and effect sizes.

To visualise global patterns in cell type composition and assess whether tissue regions cluster distinctly based on their normalised cell count profiles, we applied Uniform Manifold Approximation and Projection (UMAP) to the normalised cell count matrix. Specifically, for each ROI, the vector of normalised fractions for all cell types was used as input to UMAP, enabling two-dimensional visualisation of region-wise similarities and differences. UMAP was performed using the umap-learn package (v0.5.7), with parameters set to (n_neighbors=50, min_dist=0.5, spread=2.0, random_state=42) (McInnes *et al.,* 2018; Becht *et al.,* 2019). Cell type fractions were transformed using the centred log-ratio (CLR) transformation to mitigate the effects of the constant-sum constraint before multivariate analysis (Aitchison, 1982; Yerke *et al.,* 2024), batch effects were corrected with the ComBat algorithm (scanpy v1.11.3) to remove slide-level effects while preserving region-level differences (Johnson *et al.,* 2006; Wolf *et al.,* 2018), and feature matrices were

standardised (z-score normalisation) using the StandardScaler function (scikit-learn v1.6.1) before UMAP (Pedregosa *et al.,* 2011). UMAPs were generated with points coloured by both region and slide, to visualise the degree of separation between tissue regions and potential slide-specific effects. To compare different biological focuses, UMAPs were generated using either all detected cell types or only the subset of targeted cell types (N=18 for Cell DIVE, and N=21 for CosMx). CLR was only applied to the full-set cell type fractions.

2.1.2 Normalised Cell Count of 'Child' Cell Types per ROI based on Cells Belonging to the Corresponding 'Parent' Broad Cell Type in a ROI

This analysis was restricted to ROIs with sufficient infiltration of the relevant broad cell type (immune cells or CAFs), determined using a healthy liver tissue baseline. Specifically, for each broad type, we calculated the median number of cells per ROI in the healthy region. ROIs from all regions (Tumour, Peritumour, Healthy) were excluded from downstream analysis for a given broad type if their cell count fell below the healthy-region median. This ROI exclusion was applied separately and independently to immune cells and CAFs, resulting in two curated sets of ROIs representing immune-rich and CAF-rich niches, respectively.

Within each retained ROI, we calculated the proportion of each specific subtype (which we termed the 'child' cell type) relative to its overarching category (the 'parent' broad type). That is, for each phenotype X within broad type B in a given ROI, the normalised fraction was defined as:

Normalised Fraction of cell type X = the number of cells of phenotype X in the ROI / the number of all cells with broad type B in the ROI.

For each targeted child cell type, the mean of these normalised fractions was then calculated across all ROIs belonging to each region for every slide (patient), yielding one region-specific value per slide per phenotype. Region-wise paired comparisons were conducted using the Wilcoxon signed-rank test as in previous sections. Multiple testing correction was applied using the Benjamini–Hochberg procedure. Visualisations of regional differences in niche composition were generated using box plots, volcano plots, and bar plots.

To visualise patterns of compositional heterogeneity among immune or CAF subtypes within their respective "cellular niches", we performed UMAP embedding separately for each broad type. Only ROIs meeting the healthy-region-derived cellularity threshold for the relevant broad type were included in each UMAP. For each ROI, the vector of normalised fractions of all constituent child subtypes served as the UMAP input. Before embedding, as in 2.1.1, all feature matrices were batch-corrected and standardised. UMAP (umap-learn v0.5.7) was performed with n_neighbors=50, min_dist=0.5, spread=1.0, and random_state=42. Embeddings were visualised with points coloured by tissue region and by slide to highlight both biological and potential technical clustering.

Having established cell-type quantifications, we next interrogated the spatial relationships among cell types at increasing levels of complexity, from local neighbourhood clustering (Section 2.2) to

scale-dependent pairwise interactions (Sections 2.3–2.4) and higher-order topological features (Section 2.5).

2.2 Cellular Neighbourhood Cluster Identification and Comparison

To systematically define and compare cellular neighbourhoods within each tissue region, we adapted the established analytic strategies from Schürch et al. (2020).

The first step was single-cell neighbourhood construction, which was applied independently to the two datasets. For every cell in a dataset, we computed a cell-type composition vector based on its 10 nearest spatial neighbours (10-NN) using the NearestNeighbors algorithm in Python (scikit-learn v1.6.1). Our custom Neighborhoods class (adapted from Schürch et al.) grouped cells by region or ROI, found the 10 nearest neighbours for each cell, and counted the number of each cell type present in those local neighbourhoods. This yielded a per-cell matrix capturing the immediate neighbourhood cell-type composition. The resulting feature matrix was batch-corrected for slide effects using the ComBat algorithm (scanpy v1.11.3), while preserving regional differences. Batch-corrected neighbourhood matrices were then clustered using K-means clustering (KMeans, scikit-learn), with the optimal number of clusters (k) selected via an elbow-plot of within-cluster inertia (sum of squared distances) for k = 1–15. We chose k=9 clusters for Cell DIVE downstream analysis (k=10 for CosMx), corresponding to the “elbow” point. Each cell was assigned a neighbourhood cluster label, and both the per-cell neighbourhood matrix and original cell metadata were updated accordingly. For each cluster, the mean frequency (z-scored across clusters) of each cell type among member neighbourhoods was visualised in a heatmap.

At the ROI level, we calculated the proportion of cells assigned to each cluster. Region-level differences in neighbourhood cluster composition were computed by averaging across all ROIs within each region and visualised with stacked bar plots. To determine whether the prevalence of specific neighbourhood clusters differed across regions, we focused on targeted clusters (Clusters 0, 4, 5, 7 for Cell DIVE; Clusters 0, 2, 7 for CosMx) and compared their proportions using paired Wilcoxon signed-rank tests at the slide level. For each region, the fraction of cells belonging to each cluster was averaged across ROIs within each slide to obtain slide-level cluster abundances. Pairwise comparisons between regions (Healthy, Peritumour, Tumour) were performed for each targeted cluster, and p-values were corrected for multiple testing across clusters and region pairs using the Benjamini-Hochberg FDR procedure. Results were visualised as box plots of slide-level cluster abundances per region, with significant differences (FDR-adjusted $p < 0.05$) annotated on the plots.

Lastly, to visualise global patterns in neighbourhood composition, UMAP embedding was applied using the proportions of targeted clusters in each ROI as inputs. Batch correction with ComBat (scanpy v1.11.3) and standardised with StandardScaler (scikit-learn v1.6.1) were applied to the cluster proportions before UMAP embedding. UMAP was performed using umap-learn (v0.5.7) with n_neighbors=50, min_dist=0.5, spread=2.0, and random_state=42, resulting in two-dimensional UMAPs that were visualised with ROIs coloured by tissue region and by slide, allowing assessment of region separation and batch effects.

2.3 Permutation-Controlled, Spatial Proximity-based Cell-Cell Spatial Association/Interaction Analyses

To complement the cellular neighbourhood analysis and quantify short-range spatial interactions/correlation between cell types, we quantified short-range (10-NN) cell–cell interactions with scimap v2.3.4 (Nirmal *et al.,* 2024). For every tissue region, we first merged all ROIs from the same slide, yielding one point cloud per slide and region. Cell-level metadata (x/y coordinates, cell-type annotation, and slide ID) were loaded into an AnnData object and passed to sm.tl.spatial_interaction with method="knn", knn=10, permutation=1000, pval_method="zscore". For each ordered cell type pair (cell-type *i*, neighbour-type *j*), the function compared the observed 10-NN frequency to a null distribution obtained from 1000 random permutations of cell-type labels, yielding a Z-score (interaction enrichment if $Z > 0$, avoidance if $Z < 0$) and an exact permutation p-value for every slide.

For each region, the resulting slide-level directional Z-scores and permutation *p*-values were concatenated to form a matrix of all possible ordered cell type pairs. For the Cell DIVE dataset, to prioritise reproducible signals while allowing inter-patient variability, interactions were only shown when the permutation-derived p-value was less than 0.05 in at least half of the slides from that region (n_slides = 7). While for the CosMx dataset, to identify robust interactions, we considered only those pairs for which the permutation-derived p-value was less than 0.05 in all slides from that region (n_slides = 4).

When visualising these results, we focused on a targeted panel of cell types relevant to the tumour microenvironment, including CAFs, immune cells, endothelial cells, and tumour cells (replaced by hepatocytes for the healthy region). The directional interaction matrices were visualised as heatmaps, with each pairwise direction (e.g., CAF→CD8 T and CD8 T→CAF) was displayed separately.

2.4 Cross-Pair Correlation Functions (cross-PCF)-based Cell-Cell Spatial Association/Interaction Analyses

To assess broader spatial relationships between cell types beyond immediate neighbours, we applied cross-PCFs to compute distance-resolved spatial associations. This method, implemented via the muspan spatial-statistics Python package (v1.1.1), computed the function $g_{ij}(r)$, representing the density of j-type cells at distance r from i-type cells, normalised by a random spatial distribution baseline.

For both Cell DIVE dataset and CosMx dataset, cross-PCF were computed for all ordered pairs of cell types within each ROI. For each ROI, the file containing per-cell coordinates and cell-type annotations was loaded. Using muspan's domain object, cell coordinates were added as spatial points and cell-type labels were assigned accordingly. The physical boundaries of each ROI were estimated using an alpha-shape algorithm ($\alpha = 95$ μm), which robustly captures ROI shape for edge correction during cross-PCF calculation. To enable exhaustive spatial analysis, all possible ordered cell-type pairs, including both self-self and directional pairs, were enumerated. For each such pair, the $g_{ij}(r)$ value

was computed using the function cross_pair_correlation_function, with a maximum search radius of 250 µm, radial bins defined by annulus_step = 5 µm and annulus_width = 10 µm, and a boundary exclusion of 10 µm to reduce edge artefacts. The function was run with exclude_zero=True to ignore zero-distance artefacts. Any failed computations (e.g., due to absent cell types) were recorded as NaN. For each ROI, all computed cross-PCF curves (with distance as the index) were saved for downstream aggregation and statistical comparison. To ensure computational efficiency, this workflow was parallelised at the ROI level using the Python multiprocessing module. The computational pipeline and parameter choices were discussed with the muspan developers and analyses were conducted following the recommended procedures described in the muspan tutorial documentation.

For each tissue region (Tumour, Peritumour, Healthy), we computed region-level cross-PCFs by averaging cross-PCFs from all available ROIs. For every ROI, all ordered cell-type pairwise cross-PCFs were first computed as described above. The resulting ROI-level PCF matrices were aggregated region-wise, where each ROI-level cross-PCF file reported $g_{ij}(r)$ values for every cell-type pair across a range of spatial distances r. To ensure robust summary statistics, for cell-type pair in each region, all available ROI-level cross-PCF curves were aligned, and the mean and standard deviation (s.d.) were calculated at each distance bin. In cases where a particular ROI lacked a given cell-type pair (e.g., if one or both cell types were absent), a value of zero was imputed, and all ROIs contributed equally to the regional mean and s.d. for that pair. The resulting region-level matrices of mean and s.d. $g_{ij}(r)$ values were saved as summary files for visualisation. For direct visual comparison between region pairs, mean cross-PCF curves for each cell-type pair were plotted with 95% confidence intervals, based on the region-level s.d. and the number of ROIs. By including all available ROIs, this approach captures the full diversity and frequency of $g_{ij}(r)$ values present across entire tissue sections. As a result, the analysis provides a comprehensive overview of spatially enriched (clustering) or depleted (avoidance) cell-type interactions at multiple length scales, allowing for robust region-wise comparisons of cellular spatial architecture.

2.5 Topological Data Analysis (TDA) of Neighbourhood Clusters

To quantify the higher-order spatial organisation of cellular neighbourhood clusters in CRLM tissue, we applied persistent homology-based TDA to the distribution of cells in each ROI.

For each ROI, spatial point clouds were generated to represent the distribution of neighbourhood clusters in two modes: (1) by selecting the centroids of all cells belonging to a single neighbourhood cluster (single-cluster analysis), and (2) by selecting the centroids of all cells belonging to either member of a specific unordered pair of clusters (pairwise-cluster analysis). In the pairwise case, all points from both clusters were combined into one population for topological analysis. These point clouds served as input for subsequent TDA, capturing both within- and between-cluster spatial structures. Vietoris–Rips filtrations were computed for each set using the muspan Python package (v1.1.1), with a maximum distance parameter (max_distance = 707 µm for both Cell DIVE and CosMx, corresponding to the diagonal of a 500 µm×500 µm ROI). Persistent homology was calculated up to dimension 1, yielding H_0 (connected components) and H_1 (loops) barcodes (Otter *et al.,* 2017).

Analyses were restricted to clusters or cluster pairs with at least three cells per ROI, a threshold reflecting both mathematical and statistical reliability. TDA computations were parallelised via Python's multiprocessing to efficiently process large datasets. For each ROI, persistence diagrams for all clusters and pairs were saved in serialised (pickle) format and as summary plots.

To enable robust region-wise statistical analysis, each persistence diagram was vectorised into an interpretable set of summary features (Ali *et al.,* 2023). For every single-cluster and pairwise-cluster diagram, we extracted a standard suite of features (mean, s.d., and selected percentiles of bar births, deaths, lifespans, and midpoints, along with persistence entropy) using the muspan package. In addition to these original full-set statistics, we computed a new set of statistics based on thresholded ("significant") features when restricting analysis to bars with persistence above biologically informed thresholds (≥50 µm for H_0, ≥10 µm for H_1). This design maximised biological interpretability while minimising the influence of noise (i.e., small, artefactual topological structures). For downstream analyses, we focused on biologically relevant neighbourhood clusters and their pairs, as defined in Section 2.2 for both Cell DIVE and CosMx datasets independently. To reduce noise from irrelevant regions, ROIs were retained only if they contained at least one significant H_0 component for any targeted cluster or cluster pair. The final selected TDA feature set for each ROI included significant bar counts (for H_0 and H_1), the mean and s.d. of significant lifespans (for H_0 and H_1), and the mean and s.d. of significant H_1 birth times. To ensure that H_0 significant bar counts corresponded to the number of connected components present just before connections form at the 50 µm threshold, we incremented each count by one. This adjustment was required because each H_0 bar above the threshold represented a merging event and the true number of components is one greater than the number of bars observed.

For each ROI, extracted persistence features were aggregated at the slide level by computing the mean of each feature across all ROIs within each slide-region combination. To explore potential redundancy and intercorrelations between selected TDA features, we calculated pairwise Spearman rank correlation coefficient (ρ) and squared correlations (R^2) between all slide-level features (N = 80 for Cell DIVE, and N = 48 for CosMx) across all slides and regions with SciPy v1.15.2 (Virtanen *et al.,* 2020). For each unique feature pair, the ρ and corresponding FDR-adjusted p-value were computed. A feature pair was classified as "strongly correlated" if $R^2 \geq 0.7$, and the proportion of such strongly correlated pairs was reported. The full feature-feature correlation matrix was visualised as a lower-triangle heatmap. For the Cell DIVE dataset, only 4.7% of all correlation tests showed strong correlations (9.3% for CosMx), indicating low between-feature redundancy and supporting the use of the full selected feature panel as input for classification (Supplementary Fig.1, A/B) (Gregorutti *et al.,* 2017).

To assess the region-discriminatory capacity of our selected topological feature set, we trained random forest (RF) classifiers to separate each region pair (Tumour vs Peritumour, Tumour vs Healthy, Peritumour vs Healthy) (Breiman, 2001). For each binary comparison, the input feature matrix consisted of all selected TDA features calculated at the ROI level (N = 80 features per ROI for Cell DIVE, and N = 48 for CosMx), with each row representing a single ROI and the region label as the prediction target. Because ROIs from the same slide are not statistically independent, we implemented 5-fold

grouped cross-validation using slide identity as the grouping variable, ensuring all ROIs from a given slide were confined to a single fold. Random forest models were trained with 1,000 estimators, and all other hyperparameters were left at their default values (RandomForestClassifier, scikit-learn v1.6.1). Model accuracy was reported as mean ± s.d. across cross-validation folds. After training, mean decrease in impurity (MDI) importances were calculated for each feature, with fold-wise means and s.d. used to rank features and generate importance bar plots.

To discover TDA features that differed significantly between regions, we performed pairwise region comparisons for all selected features using paired Wilcoxon signed-rank tests at the slide level. All tests were performed across all slides with measurements in both regions. Multiple testing was controlled using the Benjamini–Hochberg FDR method (FDR-adjusted $p < 0.05$), and results were visualised using two approaches. For each feature, box plots of slide-level means were created to compare distributions across regions, with significant p-values annotated for each region pair. Additionally, volcano plots were generated to display the magnitude of feature differences between tissue regions.

To visualise global patterns and assess region-wise separation based on TDA features, we applied UMAP to the ROI-level TDA feature matrix. Before dimensionality reduction, all features were standardised, and batch effects were corrected as in sections 2.1 and 2.2. UMAP (umap-learn v0.5.7) was performed using n_neighbors=50, min_dist=0.8, and spread=2, with a random seed =42 for reproducibility. For the full dataset, UMAP embeddings were generated using all selected TDA features, with points coloured by both tissue region and slide to visualise clustering and batch effects. In addition, for each region pair, UMAP was repeated using the top N most informative features identified by random forest importance (N=30 for Cell DIVE, N=20 for CosMx, so around 40% of all selected features for each dataset), enabling comparison of tissue region discrimination based on the most discriminatory TDA descriptors. All embeddings were visualised as two-dimensional scatter plots.

3. Software and Reproducibility

All computational analyses were performed in Python 3.12 using a suite of open-source scientific libraries, including pandas (v2.2.3), numpy (v2.2.5), matplotlib (v3.10.1), and seaborn (v0.13.2) for data processing, visualisation, and figure generation. Core spatial statistics and topological analyses were conducted using the muspan package (v1.1.1), with parallelisation enabled via Python's multiprocessing module to accelerate ROI-level analyses. Statistical comparisons, including paired Wilcoxon signed-rank tests and false discovery rate correction, were performed using scipy.stats (v1.15.2) and statsmodels (v0.14.4). Dimensionality reduction and batch correction relied on umap-learn (v0.5.7), scikit-learn (v1.6.1), and scanpy (v1.11.3), while spatial interaction analysis was carried out with scimap (v2.3.4). All codes for data pre-processing, spatial analysis, and downstream figure production are provided at https://github.com/rach-crc/MRes-Oncology-Thesis-Code. Analysis was performed on a MacOS M2 system.

Results

1. Differences in Normalised Cell Count between Regions

We first sought to understand how the proportions of different cell types varied across distinct tissue microenvironments. This initial step confirmed the accuracy of our regional definition strategy for both Cell DIVE and CosMx datasets, as tumour cells were primarily found in tumour ROIs and hepatocytes were concentrated in healthy liver ROIs (Fig.4). However, since our research interest focused on immune cells and CAFs, when quantifying the prevalence of cell phenotypes across regions, we limited the analysis to only cell types of biological interest and excluded tumour and hepatocyte populations from region-wise comparisons.

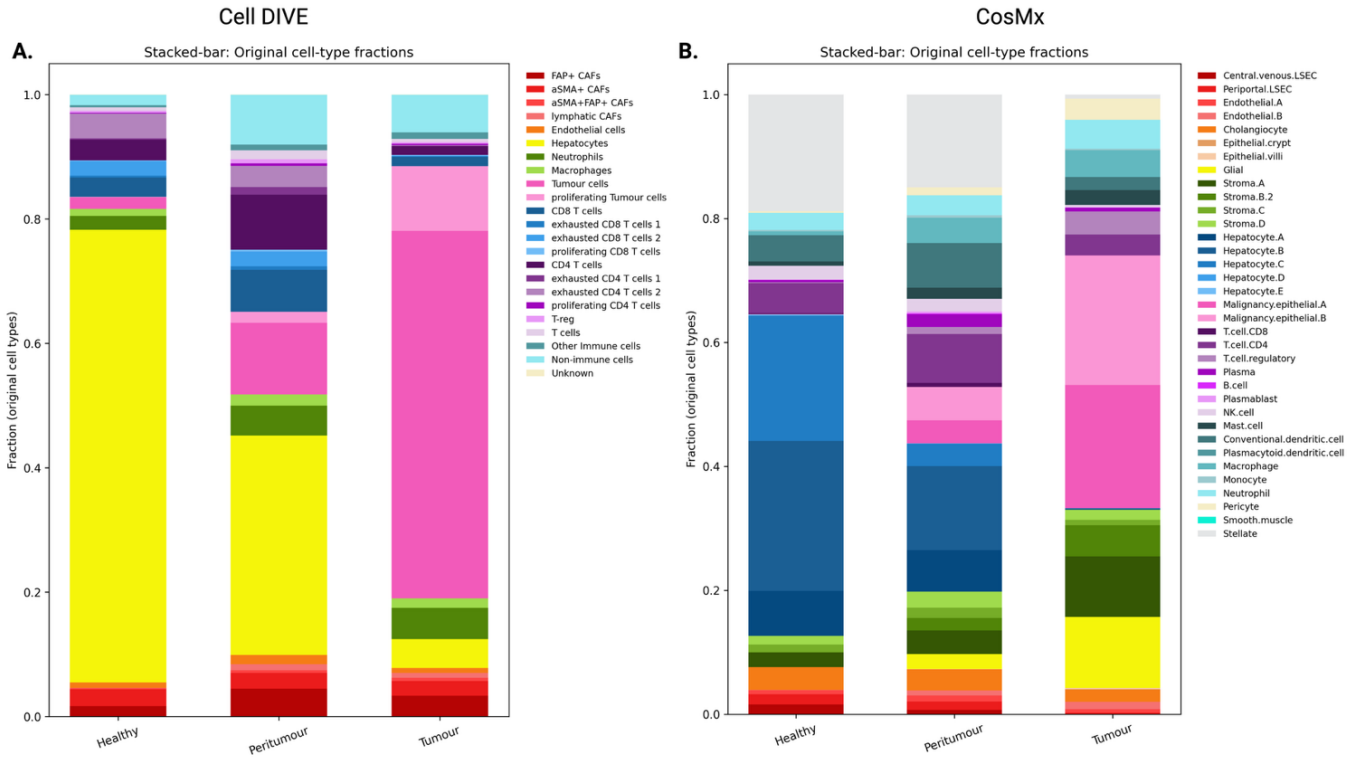


Figure 4. Regional variation in normalised cell counts for all cell type across CRLM tissue compartments in Cell DIVE and CosMx datasets. A. Cell DIVE dataset: stacked bar plots display the mean fraction of each cell type (as a proportion of all detected cells in a ROI) within healthy, peritumour, and tumour ROIs. Each bar represents the average composition across all ROIs assigned to that region, coloured by cell type as indicated in the legend at right. Hepatocytes (yellow) predominate in healthy-region ROIs, while tumour cells (pink) are largely confined to tumour-region ROIs, confirming accurate region annotation. Immune and fibroblast populations are variably distributed, with a notable enrichment of adaptive lymphocytes (blue and purple) in the peritumour compartment. B. CosMx dataset: stacked bar plots show the mean fraction of each cell type across healthy, peritumour, and tumour ROIs, coloured according to the cell-type legend. As in the proteomics data, hepatocyte subclasses (blue) are dominant in healthy-region ROIs, while malignant epithelial subsets (pink) are highly enriched in tumour ROIs. Peritumour regions display a broad admixture of immune, fibroblast, and endothelial populations. CAF, cancer-associated fibroblast; LSEC, liver sinusoidal endothelial cell; aSMA, alpha–smooth muscle actin; FAP, fibroblast activation protein; exhausted T cell 1, PD-1 high; exhausted T cell 2, CTLA-1 high; Stroma.x indicated different fibroblasts.

We applied identical analytical frameworks to the Cell DIVE spatial proteomics cohort (18 targeted cell types and 2 targeted broad types) and the CosMx spatial transcriptomics cohort (21 targeted cell

types and 2 targeted broad types) (Table.4), enabling direct cross-platform comparison that leveraged Cell DIVE's larger sample size and functional status annotations, alongside CosMx's finer phenotypic resolution. Figure 5 presents a representative raw multiplexed image of CRLM tissue from the Cell DIVE dataset, with selected markers indicated. While figure 6 shows the UMAP and flightpath plots for the CosMx spatial dataset, along with a heatmap displaying the top 30 most differentially expressed genes for each cell type.

Dataset	Broad_Type	Cell_type
CellDIVE	Immune Cells	"Macrophages", "Neutrophils", "CD8 T cells", "exhausted CD8 T cells 1", "exhausted CD8 T cells 2"," proliferating CD8 T cells","CD4 T cells", "T-reg", "exhausted CD4 T cells 1", "exhausted CD4 T cells 2", "proliferating CD4 T cells", "T cells", "Other Immune cells "
CellDIVE	CAFs	"FAP+ CAFs", "aSMA+ CAFs", "aSMA+ FAP+ CAFs", "lymphatic CAFs"
CellDIVE	Not targeted (Endothelial)	"Endothelial cells"
CosMx	Immune Cells	"Neutrophil", "Macrophage", "Monocyte", "Mast.cell", "Conventional.dendritic.cell", "Plasmacytoid.dendritic.cell", "Plasma", "Plasmablast", "T.cell.CD4", "B.cell", "T.cell.CD8", "T.cell.regulatory", and "NK.cell"
CosMx	CAFs	"Stroma.A", "Stroma.B.2", "Stroma.C", "Stroma.D"
CosMx	Not targeted (Endothelial)	"Endothelial.A", "Endothelial.B", "Central.venous.LSEC", "Periportal.LSEC"

Table 4. Targeted cell types and broad cell-type categories involved for conducting paired Wilcoxon signed-rank test.

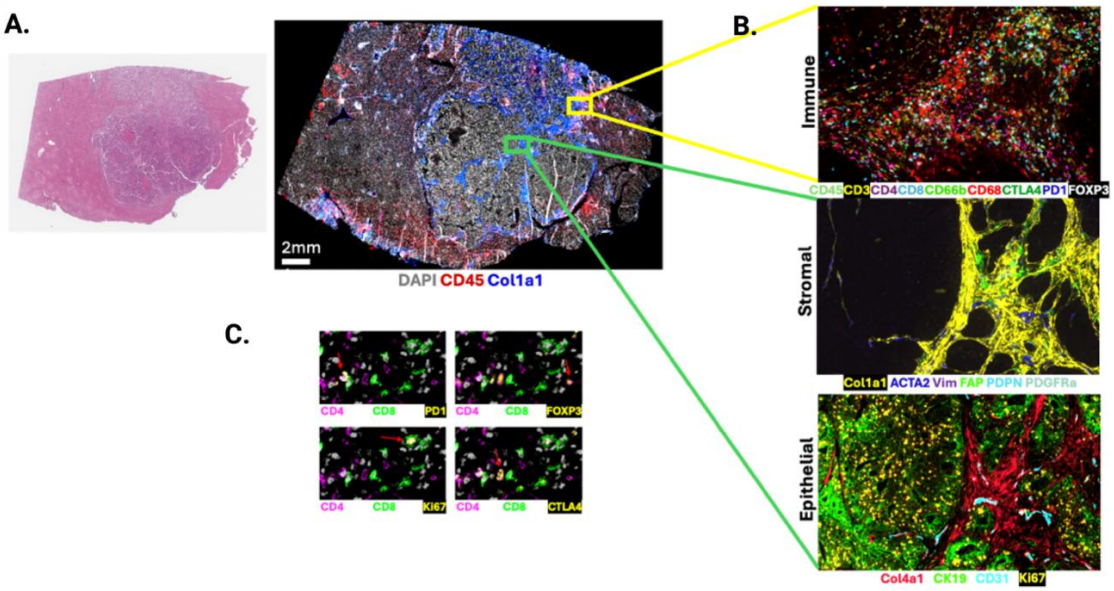


Figure 5. Example raw multiplexed image from the Cell DIVE cohort. A. Left: H&E staining of one representative patient slide. Right: whole-slide CRLM image with surrounding uninvolved liver visualising the indicated markers. B. High magnification of antibody panel for different cell types. C. Identification of exhausted ($PD1^+/CTLA4^+$), proliferating ($Ki67^+$) and T-reg subpopulations. Red arrows indicating double positive cells. *Note: these images were generated and pre-processed by other members of the lab.*

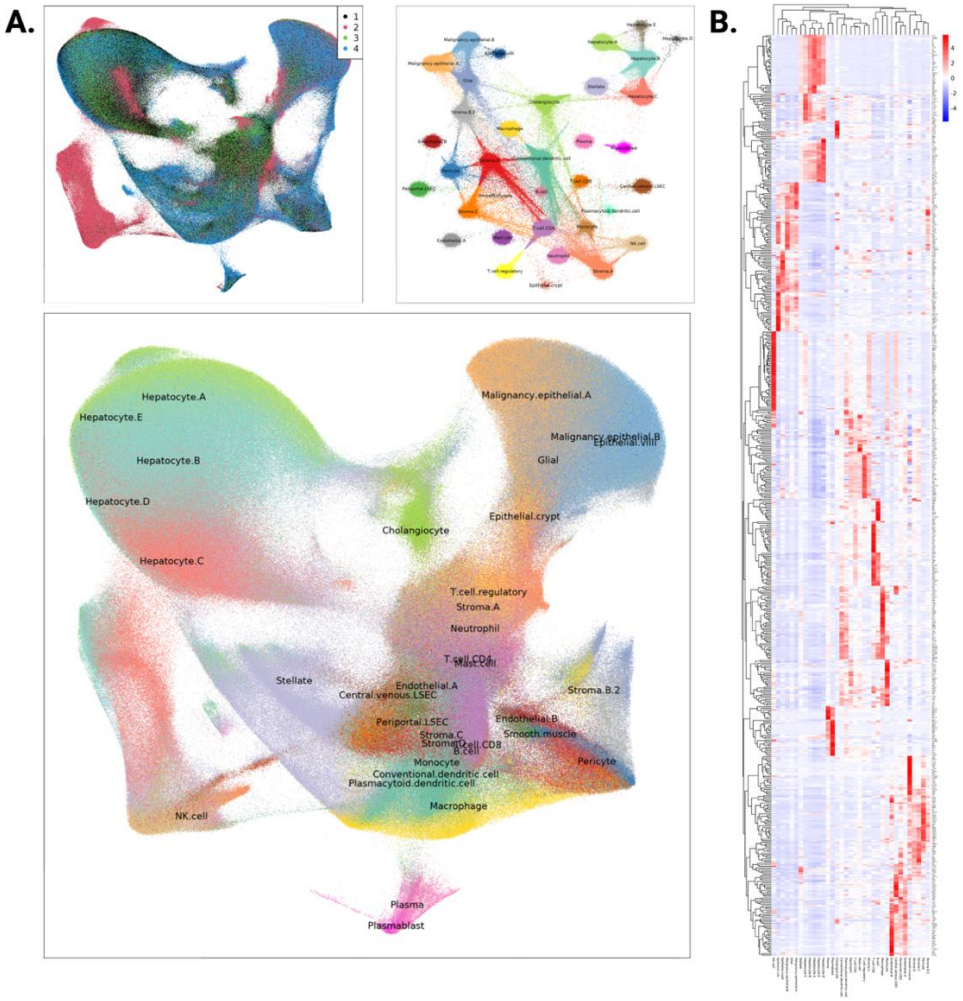


Figure 6. Overview of the cell types contained in the CosMx spatial transcriptomic dataset. A. UMAP and flightpath plots for the CosMx spatial dataset. B. Heatmap of top 30 most differentially expressed genes for each cell type identified within the dataset. *Note: these images were generated and pre-processed by other members of the lab.*

1.1 Spatial Distribution of Targeted Cell Types across Tissue Region

To characterise how major immune, fibroblast, and endothelial populations were distributed across tumour, peritumour, and healthy tissue regions in CRLM, we applied a unified cell-type quantification workflow (Method 2.1.1) to both the Cell DIVE spatial proteomics and CosMx spatial transcriptomics datasets. Although we only included ROIs containing at least one cell from any targeted phenotypes for analyses, in practice, all ROIs from either the Cell DIVE or CosMx dataset were reserved. Guided by the approach, we asked: Among ROIs containing any of our targeted phenotypes, does the slide-level normalised cell count for each phenotype differ significantly between any two regions?

1.1.1 Cell DIVE Cohort

Using the ROI-level cell-type fractions derived from multiplexed proteomic imaging (13 slides, 1,617 ROIs, 2,923,140 cells), we observed that the peritumour microenvironment of CRLM exhibited a pronounced enrichment of both lymphoid and endothelial cell populations compared to tumour core and healthy liver regions (Fig.4, A). These trends were confirmed in paired Wilcoxon signed-rank tests. Within the immune compartment, the peritumour region had significantly higher proportions of $CD4^+$ T-cells (including the proliferating subset, exhausted subsets (PD-1-high, CTLA4-high), and the

activated subset), $CD8^+$ T-cells (including an exhausted subset (CTLA4-high) and the activated subset), T-regs, and unclassified T-cells than the tumour core (adjusted $p < 0.05$) (Fig. 7; Supplementary Fig.2, A). The increased abundance of both conventional and dysfunctional T-cells in the peritumour region reflects an active yet dysregulated immune response at the invasive margin. This pattern is characteristic of an infiltrated–excluded tumour immune microenvironment (TIME), a TME phenotype frequently observed in solid tumours, where multiple immune cells accumulate at the tumour border, but cytotoxic lymphocytes (CTLs) are largely excluded from the tumour core. (Binnewies *et al.,* 2018; Tauriello *et al.,* 2018; Galon *et al.,* 2006). Broad endothelial cells were also more represented at the peritumour margin than at the tumour core, likely reflecting liver's native vascular architecture, such as the portal triads and active tumour-associated lymph-angiogenesis, a process strongly correlated with lymphatic invasion and poor prognosis in CRLM. (Ülger *et al.,* 2018; Muralidharan *et al.,* 2014). In contrast, myeloid populations (neutrophils, macrophages), CAFs, and "other immune cells" showed no significant differences between peritumour and tumour, indicating that their distribution may be less spatially restricted. Comparison with healthy background liver revealed further enrichment of lymphoid and specific fibroblast subsets in the peritumour region (Fig.7; Supplementary Fig.2, B), while tumour regions were relatively depleted of immune populations and exhibited a modest increase in CAFs (Fig.7; Supplementary Fig.2, C).

UMAP embedding of ROI-level targeted cell-type fractions demonstrated clear regional clustering, with peritumour ROIs forming a distinct intermediate group between tumour and healthy regions, consistent with a transitional microenvironment at the invasive margin (Fig.8). The modest overlaps at the regional boundaries underscored gradual shifts in regional microenvironment patterns of CRLM.

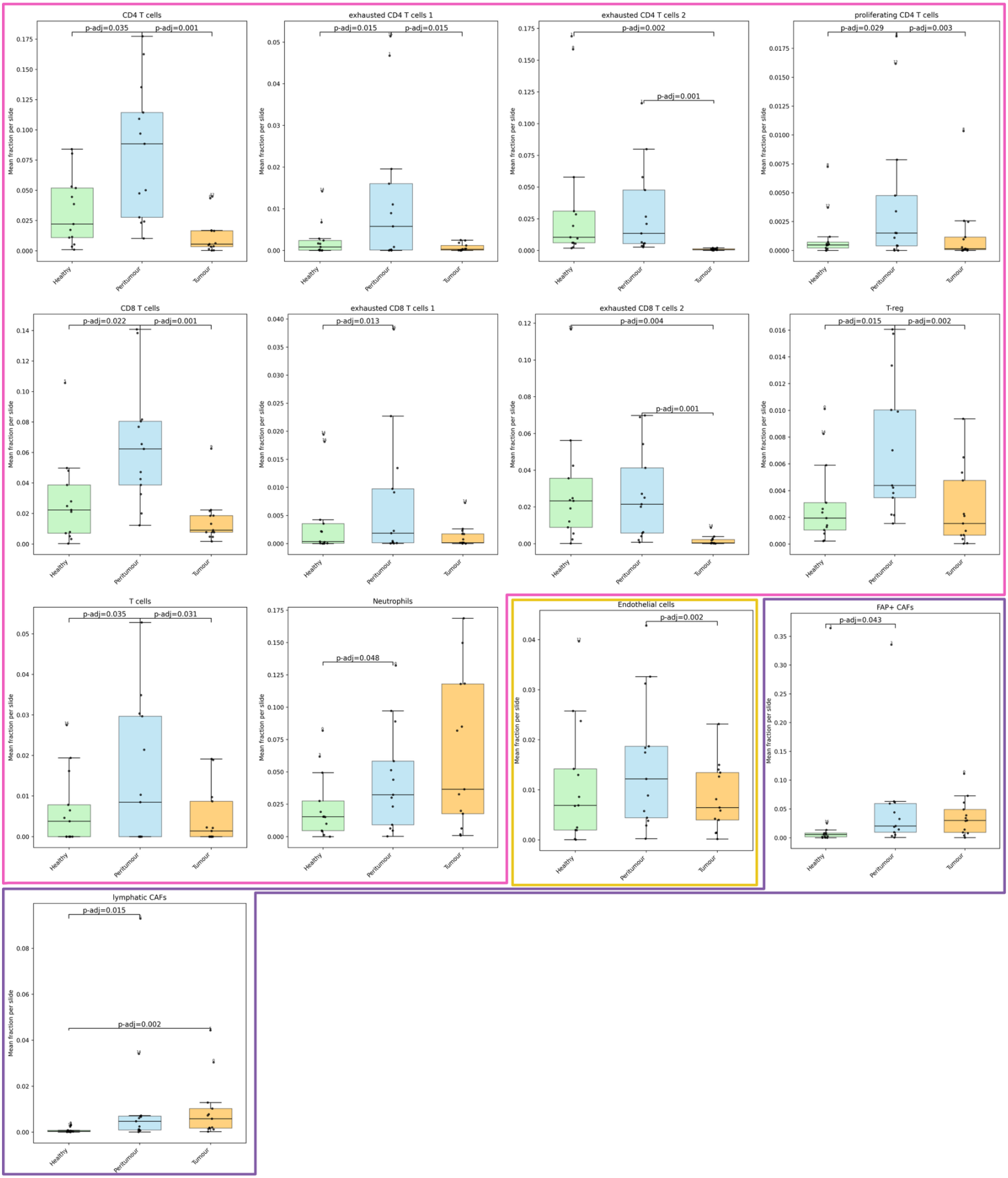


Figure 7. Region-specific abundance of targeted immune, fibroblastic, and endothelial cell types in CRLM: Cell DIVE cohort. Box plots display the distribution of slide-level mean normalised cell counts for each targeted cell type (y-axis) across three tissue regions (x-axis). Each panel presents a distinct cell type. Immune cell panels are outlined in pink,

endothelial cell in yellow, and fibroblastic cell panels in purple. Each box summarises values from all slides (patients) within the corresponding region (n = 13 per region). Individual black dots represent individual slide (patient) and denote slide-level means. Outliers were labelled by slide number. Box colours: light green for Healthy, sky blue for Peritumour, and orange for Tumour. Horizontal bars and numeric labels above the boxes indicate pairwise statistical comparisons between regions. “p-adj” values represent FDR–adjusted p-values from paired Wilcoxon signed-rank tests, performed for each cell type and region pair. Only comparisons with FDR-adjusted p < 0.05 are shown. Box plot elements show interquartile ranges (boxes), medians (horizontal lines), and whiskers extending to 1.5 × IQR.

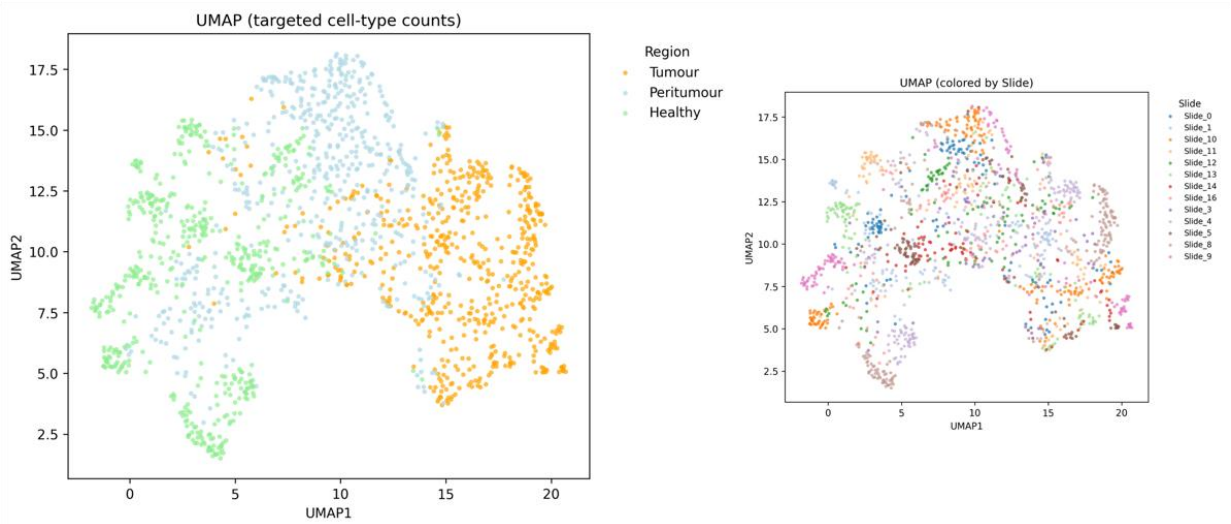


Figure 8. UMAP embedding of ROI-level targeted cell-type fractions reveals region- and slide-associated clustering in the Cell DIVE cohort. Left: UMAP of ROIs, using vectors of normalised fractions for all targeted cell types, coloured by tissue region: Tumour (orange), Peritumour (light blue), and Healthy (light green). Each point represents one ROI. Region-wise separation is apparent, with peritumour ROIs occupying an intermediate position between tumour and healthy regions. Right: UMAP embedding of the same ROIs, coloured by slide. Points from different slides were mixed, revealing a relatively low degree of slide-level (batch) effects after batch-correction.

1.1.2 CosMx Cohort

To validate and further dissect these spatial trends at higher phenotypic resolution, we analysed the CosMx dataset (4 slides, 1,233 ROIs, 1,339,397 cells). Regionally, by observation, peritumour tissue displayed an expected mixture of cell types and was relatively more enriched in lymphoid cells compared to other regions. The tumour core appeared to contain more CAFs and myeloid cells, while the healthy liver was dominated by hepatocytes and contained smaller proportions of other cell populations (Fig.4, B). To confirm these trends, we applied paired Wilcoxon signed-rank tests. Although the small sample size limited statistical power (i.e., none of the differences described below reached statistical significance), qualitative patterns closely mirrored those observed in the Cell DIVE cohort. The peritumour region showed relative enrichment in a range of immune and fibroblast cell types compared to the tumour core, including conventional dendritic cells (cDC), multiple T-cell and B-cell subsets, natural killer (NK) cells, generic fibroblasts (Stroma.C), and a fibroblast subtype characterised by CCL19/CCL21 and antigen-presentation genes (Stroma.D) (Fig.9; Supplementary Fig.3, A). The endothelial compartment followed the same trend. Periportal and central-venous liver-sinusoidal endothelial cells (LSEC) were both higher in proportions at the peritumour region, consistent with the canonical portal-to-central zonation of LSECs described by single-cell and spatial-transcriptomic liver atlases (Su *et al.,* 2021). Alongside LSECs, an uncharacterised "endothelial A"

population was also more abundant at peritumour, hinting at a dense lymphatic network in the peritumour rim, however, its precise phenotype remains inferential and needs further validation. Clinically, high peritumour lymphatic-vessel density has previously been linked to earlier recurrence and poorer disease-free survival in CRLM (Muralidharan *et al.,* 2014).

As in Cell DIVE, the tumour core of the CosMx cohort showed an “immune-cold” profile. More importantly, the higher resolution of CosMx revealed that this microenvironment is more than a passive state of immune exclusion, as it also maintains abundant immunosuppressive activities. The tumour core demonstrated enrichments of T-regs, which are known to directly inhibit effector T-cell function and create a tolerogenic milieu (Sojka *et al.,* 2008). This immunosuppressive landscape is potentially reinforced by diverse myeloid populations and a fibroblast subset (Stroma.B.2) expressing genes for both ECM-remodelling and neutrophil chemo-attraction. The ECM-remodelling signature is consistent with the formation of a dense, desmoplastic barrier that physically impedes T-cell infiltration, a hallmark of immune exclusion orchestrated by CAFs (Sahai *et al.,* 2020; Xiao *et al.,* 2023; Yuan *et al.,* 2023). Simultaneously, some tumour-associated neutrophils (TANs) and tumour-associated macrophages (TAMs) were found to be immunosuppressive and tumour-promoting (Zhou *et al.,* 2025; Pan *et al.,* 2020). Furthermore, the enrichment of an "Endothelial.B" population, while phenotypically undefined, strongly implies active tumour neoangiogenesis, a process essential for sustaining the growth of the metastatic lesion. These newly formed tumour vessels are often structurally abnormal and leaky, contributing to a hypoxic and acidic microenvironment that further suppresses CTL function and promotes immune escape (Chen *et al.,* 2023b). Lastly, paralleling the Cell DIVE findings, comparison with healthy liver showed LSEC predominating in healthy ROIs, while immune cells, fibroblasts, and two endothelial subtypes (Endothelial.A/B) were more prominent in peritumour regions (Fig.9; Supplementary Fig.3, B). Comparison of healthy liver with tumour regions (Fig.9; Supplementary Fig.3, C) also revealed enrichment of LSECs in healthy ROIs, along with cDCs, B cells, CD8 T-cells, and plasmacytoid dendritic cells (pDC), whereas fibroblasts and other immune cell types were relatively increased within tumours.

UMAP embedding of ROI-level targeted cell-type fractions further confirmed robust region-specific clustering and the intermediate positioning of peritumour ROIs, although slide 4 exhibited a distinct cluster pattern, possibly reflecting patient-specific biology or technical effects (Fig. 10).


T.cell.CD4
T.cell.CD8
NK.cell
B.cell
Macrophage
Neutrophil
Monocyte
Mast.cell
Conventional.dendritic.cell
Plasmacytoid.dendritic.cell
Plasmablast
Plasma
T.cell.regulatory
Central.venous.LSEC
Periportal.LSEC
Endothelial.A
Endothelial.B
Stroma.B.2
Stroma.C
Stroma.D
Mean fraction per slide
Healthy
Peritumour
Tumour

Figure 9. Region-specific abundance of targeted immune, fibroblastic, and endothelial cell types in CRLM: CosMx cohort. Box plots display the distribution of slide-level mean normalised cell counts for each targeted cell type (y-axis) across three tissue regions (x-axis). Each panel presents a distinct cell type. Immune cell panels are outlined in pink, endothelial cell in yellow, and fibroblastic cell panels in purple. Each box summarises values from all slides (patients) within the corresponding region (n = 4 per region). Individual black dots represent individual slide (patient) and denote slide-level means. Outliers were labelled by slide number. Box colours: light green for Healthy, sky blue for Peritumour, and orange for Tumour. None of the region-wise comparisons were significant (paired Wilcoxon signed-rank tests with Benjamini–Hochberg FDR correction), hence no "p-adj" annotation above individual box plots.

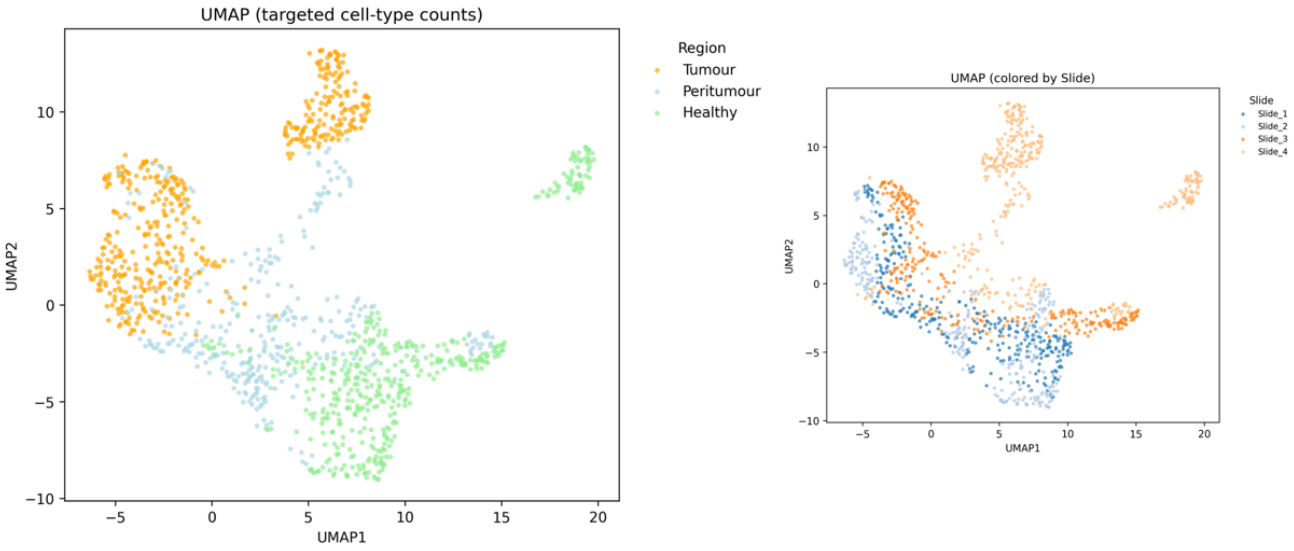


Figure 10. UMAP embedding of ROI-level targeted cell-type fractions reveals region- and slide-associated clustering in the CosMx cohort. Left: UMAP of ROIs, using vectors of normalised fractions for all targeted cell types, coloured by tissue region: Tumour (orange), Peritumour (light blue), and Healthy (light green). Each point represents one ROI. Region-wise separation is apparent, with peritumour ROIs occupying an intermediate position between tumour and healthy regions. Right: UMAP embedding of the same ROIs, coloured by slide. Points from different slides were mixed, revealing a relatively low degree of slide-level (batch) effects after batch correction.

1.2 Composition of Immune and Fibroblastic "Niches" Across Tissue Regions

To further characterise the composition of immune- and fibroblast-rich microenvironments in CRLM, we quantified the relative abundance of specific cell subtypes, which for convenience we termed 'child' cell types, within their overarching broad cell-type categories (namely, total immune cells and CAFs). This analysis was performed on a per-ROI basis, allowing us to determine the proportional makeup of immune-/CAF- rich microenvironments (e.g., the proportion of $CD8^+$ T-cells within the total immune infiltrate).

In contrast to Section 1.1, where slide-level means were calculated across all ROIs within each region, this analysis was restricted to ROIs with sufficient infiltration of a given broad type (immune cell or CAFs). Specifically, we applied a healthy-tissue–derived cellularity threshold, using the median cell count of a given broad cell type in healthy-region ROIs as the baseline (see Method 2.1.2). This exclusion criterion ensured that only ROIs with at least a typical level of immune or fibroblastic cell representation were retained for region-wise comparisons, thereby improving biological

interpretability, mitigating the influence of biologically sparse or technically compromised ROIs, and increasing the robustness of statistical inference.

Within each high-infiltration ROI, normalising each child cell type to its parent broad type (i.e., the proportion of $CD8^+$ T-cells in all immune cells in a ROI), rather than to total cells like in Section 1.1, enabled a context-specific assessment of niche composition and allowed us to detect shifts in the internal makeup of immune or fibroblastic compartments that may be masked by whole-sample averaging. As a complementary perspective to the global distribution analyses presented in Section 1.1, this targeted approach focused explicitly on microenvironments with sufficient cellular content and addressed the following question: Among ROIs with sufficient infiltration of a given broad type (immune cells or CAFs), does the slide-level fraction of each targeted child cell type (relative to its parent broad type) differ significantly between any two tissue regions?

1.2.1 Cell DIVE Cohort

For the Cell DIVE dataset, the median immune cell count in the healthy region was 177.50 cells per ROI, and the median CAF count in the healthy region was 12.00 cells per ROI. Restricting the analysis to ROIs that met the immune-cell infiltration threshold and normalising each targeted subtype to the total immune compartment revealed distinct internal make-ups across regions. Composition plots showed that the tumour immune niche was skewed toward neutrophils, whereas peritumour niches were dominated by adaptive lymphocytes ($CD8^+$ and $CD4^+$ T-cells), and healthy liver displayed a more balanced mixture of $CD4^+$, $CD8^+$ and exhausted (CTLA4-high) T-cells (Fig.11). Paired Wilcoxon signed-rank tests across slides further confirmed these patterns (Fig.12). Within the immune compartment, the fraction of activated $CD4^+$ and $CD8^+$ T-cells was significantly higher in peritumour than tumour ROIs (adjusted $p < 0.05$). T-cell dysfunction was also regionally patterned. Exhausted $CD4^+$ T-cells 1 (PD-1-high) were significantly enriched in peritumour immune-cell niches versus healthy immune-cell niches. Although the differences did not reach statistical significance, tumour immune-cell niches appeared to have the lowest levels of PD-1-high exhausted T-cells among all regions. Likewise, exhausted $CD4^+$ T-cells 2 (CTLA4-high) and exhausted $CD8^+$ T-cells 2 (CTLA4-high) were significantly depleted in the tumour relative to both healthy and peritumour compartments, and the immune-cell niche of the healthy liver had the highest level of CTLA4-high exhausted T-cells (adjusted $p < 0.05$). In contrast, neutrophils occupied a significantly greater proportion of the tumour immune niche than either healthy or peritumour niches.

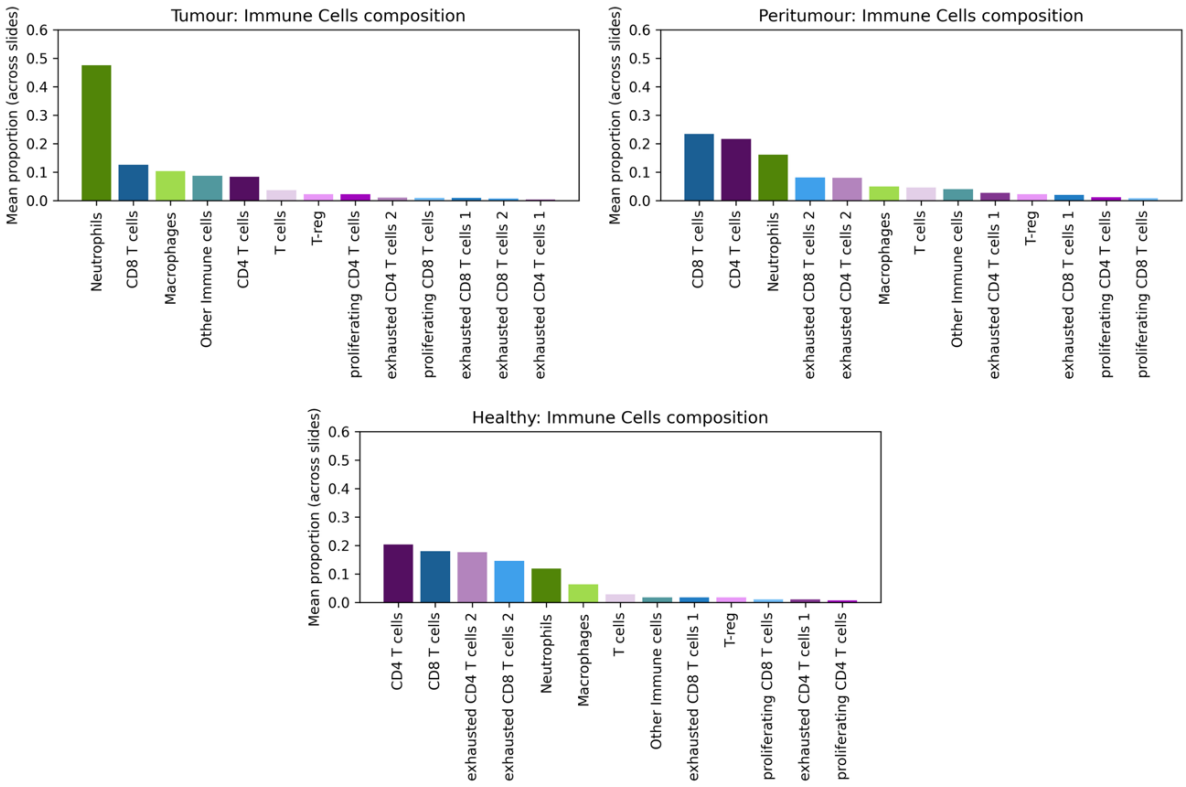


Figure 11. Cell DIVE cohort: Regional composition of immune compartments in CRLM. Immune cell composition in immune-rich niches. Bar plots display the mean fraction of each immune cell subtype within the immune compartment, for tumour, peritumour, and healthy regions. Only ROIs with immune cell counts exceeding the healthy-region median were included. Bars represent the proportion of each targeted immune cell type normalised to the total immune cell count per ROI, then averaged across all ROIs and slides for each region.

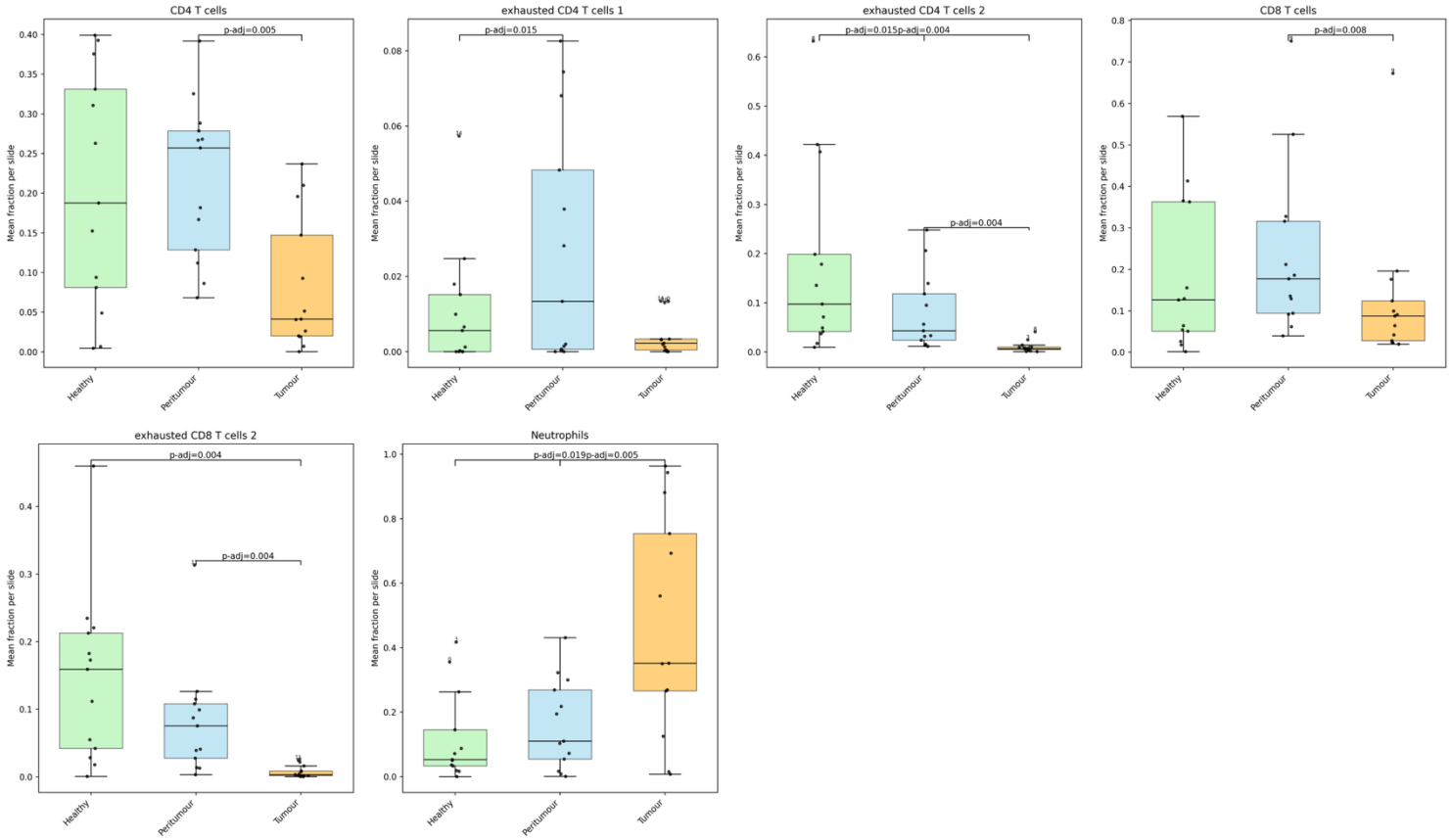


Figure 12. Cell DIVE cohort: Regional variation in immune subtype proportions within broad immune-rich compartments across slides. Box plots show slide-level mean normalised fractions of each immune cell subtype within the immune-rich niche, stratified by region (healthy, peritumour, tumour). Individual black dots represent individual slide (patient) and denote slide-level means. Outliers were labelled by slide number. Box colours: light green for Healthy, sky blue for Peritumour, and orange for Tumour. Horizontal bars and numeric labels above the boxes indicate pairwise statistical comparisons between regions. "p-adj" values represent FDR–adjusted p-values from paired Wilcoxon signed-rank tests, performed for each cell type and region pair. Only comparisons with FDR-adjusted $p < 0.05$ are shown.

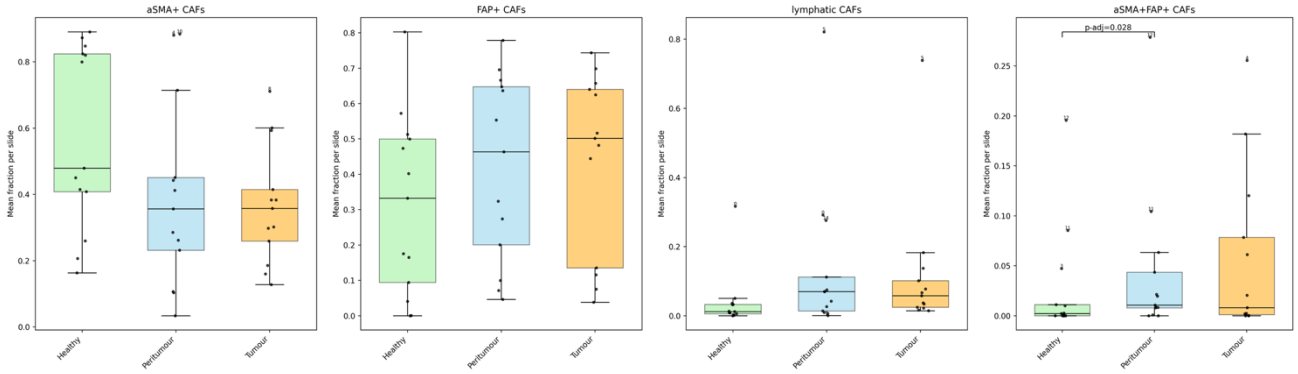


Figure 13. Cell DIVE cohort: Regional variation in CAF subtype proportions within broad CAF-rich compartments across slides. Box plots show slide-level mean normalised fractions of each CAF subtype within the CAF-rich niche, stratified by region. Individual black dots represent individual slide (patient) and denote slide-level means. Outliers were labelled by slide number. Box colours: light green for Healthy, sky blue for Peritumour, and orange for Tumour. Horizontal bars and numeric labels above the boxes indicate pairwise statistical comparisons between regions. "p-adj" values represent FDR–adjusted p-values from paired Wilcoxon signed-rank tests, performed for each cell type and region pair. Only comparisons with FDR-adjusted $p < 0.05$ are labelled.

Nevertheless, when the analysis was restricted to ROIs that met the CAF infiltration threshold, regional differences in CAF composition were far less pronounced. Composition plots revealed that CAF niches in all three regions were similarly dominated by alpha–smooth muscle actin (aSMA)+ CAFs and fibroblast activation protein (FAP)+ CAFs, with no substantial shifts in the relative contribution of individual CAF subsets. This was corroborated by paired Wilcoxon signed-rank tests across slides, which did not detect any significant regional differences in the internal composition of CAF compartments (Fig.13). These findings suggested that, unlike the immune compartment, CAF niches might be more conserved across tumour, peritumour, and healthy regions in CRLM.

UMAP embedding of ROI-level compositions within the immune cell compartment revealed partially distinct clustering by region, with healthy and tumour ROIs occupying divergent territories and peritumour ROIs diffusely interspersed between the two (Fig.14, A). By contrast, the UMAP embedding of CAF compartment compositions did not show clear region-specific segregation (Fig.14, B). Instead, ROIs from different regions appeared largely intermixed, supporting that the relative composition of CAF subtypes was broadly similar across regions.

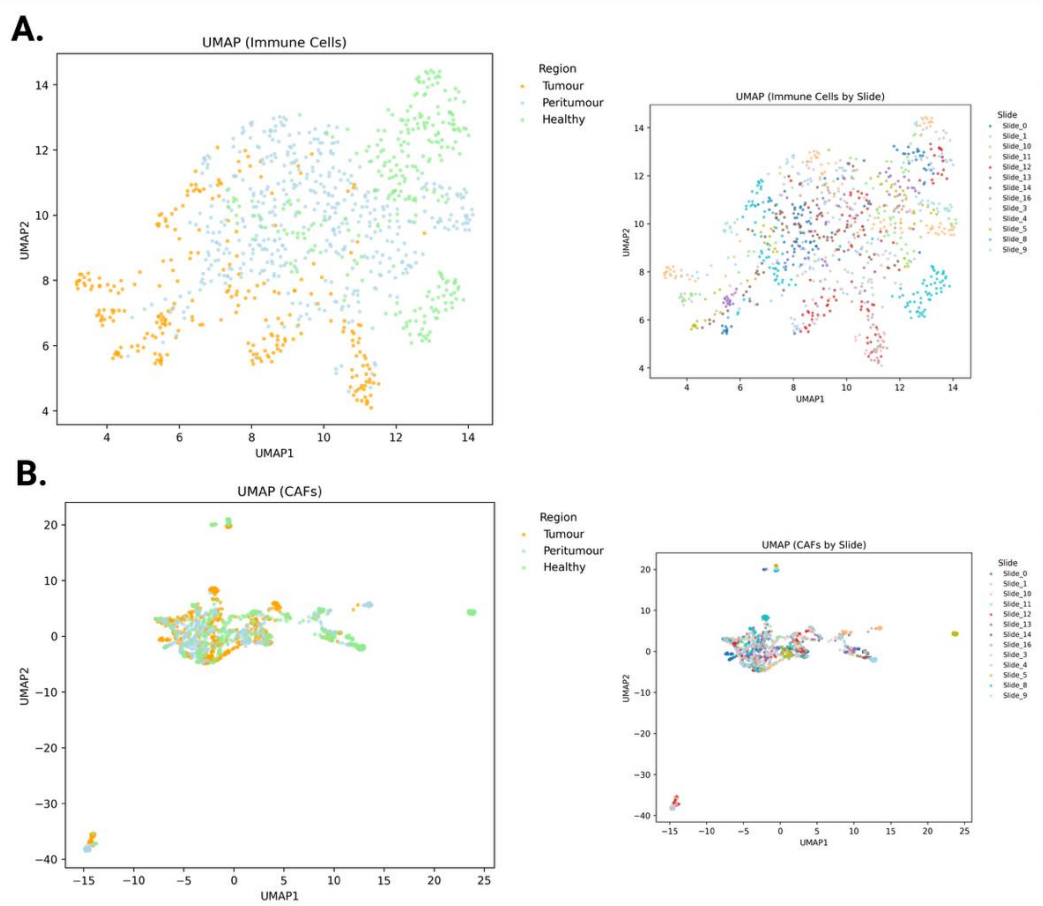


Figure 14. Cell DIVE cohort: UMAP embedding of immune and CAF compartment compositions in CRLM ROIs. A. UMAP of immune-rich ROIs. UMAP embedding was performed on the vector of normalised immune subtype fractions for each immune-rich ROI. Points represent individual ROIs, coloured by tissue region (left panel) or by slide (right panel). Tumour and healthy ROIs occupy largely distinct areas of the embedding, with peritumour ROIs occupying an intermediate space, reflecting the graded shift in immune composition across regions. B. UMAP of CAF-rich ROIs. UMAP embedding on the normalised CAF subtype fractions for each CAF-rich ROI. Points are coloured by region (left) or slide (right). ROIs from all regions showed substantial overlap in the embedding, supporting the notion of broadly conserved CAF niche composition across different regions.

1.2.2 CosMx Cohort

Using the same method, we restricted CosMx ROIs to immune-rich and CAF-rich niches and normalised each targeted subtype to its parent compartment. For the CosMx dataset, the median immune cell count in the healthy region was 127.00 cells per ROI, and the median CAF count in the healthy region was 33.00 cells per ROI. With only four slides, the patterns below are descriptive (adjusted $p > 0.05$).

Within immune-rich niches, peritumour ROIs were relatively enriched for adaptive lymphocytes and APC compared with tumour ROIs. Composition plots showed peritumour niches skewed toward $CD4^+$ T-cells and cDCs, with additional contributions from macrophages, neutrophils, NK cells, $CD8^+$ T-cells, etc. Whereas tumour immune-cell niches were dominated by macrophages and neutrophils, with a noticeable increase in T-regs. $CD8^+$ T-cells contributed only a small fraction across regions and were lowest in the tumour (Fig.15, A). Healthy immune niches also showed high proportions of $CD4^+$ T-

cells and cDCs, and comparatively lower T-regs. The box and whisker plots for paired Wilcoxon signed-rank test results reflected similar trends (Fig.16). Hence, the CosMx data mirrored the Cell DIVE pattern, where tumour immune niches are myeloid-skewed and relatively lymphoid-poor ("immune-cold"), the peritumour interface is comparatively lymphoid/APC enriched, and healthy liver is characterised by high cDC and $CD4^+$ T-cell proportions. CAF-rich niches showed more structure in CosMx than in Cell DIVE. Excluding the undefined Stroma A subtype, Stroma D and Stroma C dominated peritumour CAF niches. Stroma B.2 was relatively prominent in tumour CAF niches and nearly absent in healthy niches (Fig.15, B; Fig.17). These trends suggested a progressive shift in CAF niche composition across regions, characterised by a marked increase in Stroma B.2 abundance from healthy tissue to tumour, a pattern that was not apparent at the protein panel resolution of Cell DIVE.

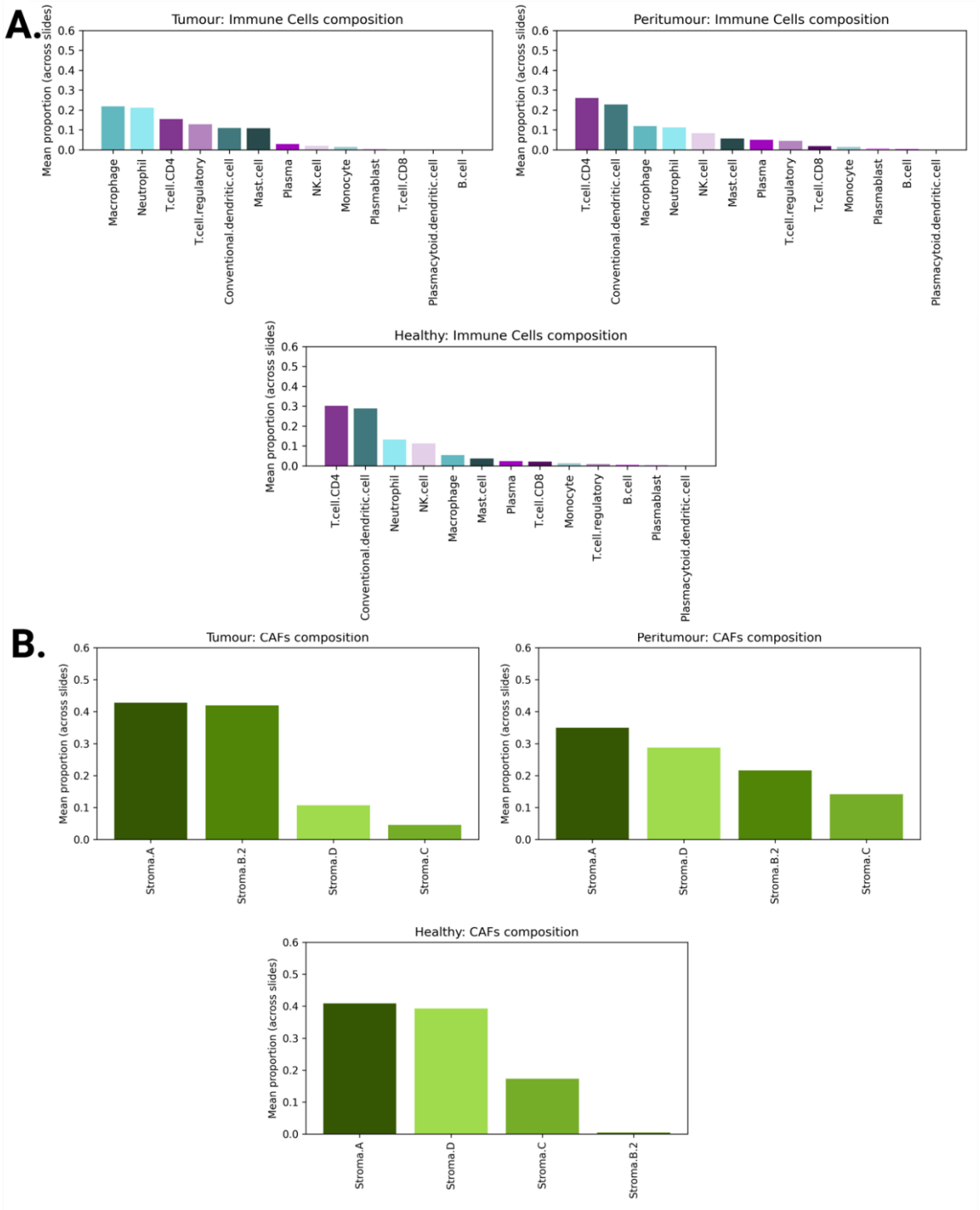


Figure 15. CosMx cohort: Regional composition of immune and cancer-associated fibroblast (CAF) compartments in CRLM. A. Immune cell composition in immune-rich niches. Bar plots display the mean fraction of each immune cell subtype within the immune compartment for each region. Only ROIs with immune cell counts exceeding the healthy-region median were included. Bars represent the proportion of each targeted immune cell type normalised to the total immune cell count per ROI, then averaged across all ROIs and slides for each region. B. CAF subtype composition in CAF-rich niches.

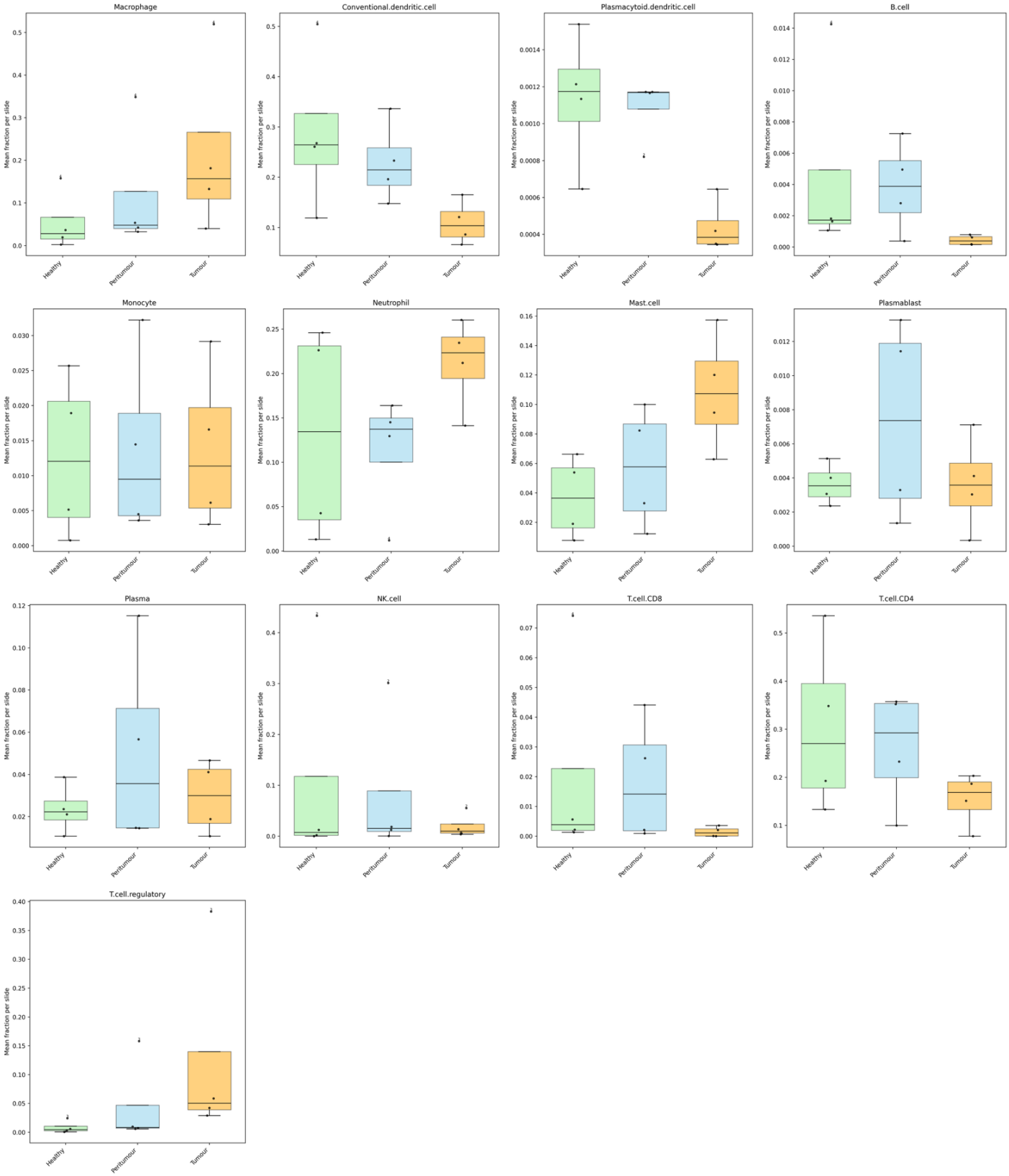


Figure 16. CosMx cohort: Regional variation in immune subtype proportions within broad immune-rich compartments across slides. Box plots show slide-level mean normalised fractions of each immune cell subtype within the immune-rich niche, stratified by region. Individual black dots represent individual slide (patient) and denote slide-level means. Outliers were labelled by slide number. Box colours: light green for Healthy, sky blue for Peritumour, and orange for Tumour. None of the region-wise comparisons were significant (paired Wilcoxon signed-rank tests with Benjamini–Hochberg FDR correction), hence no "p-adj" annotation above individual box plots.

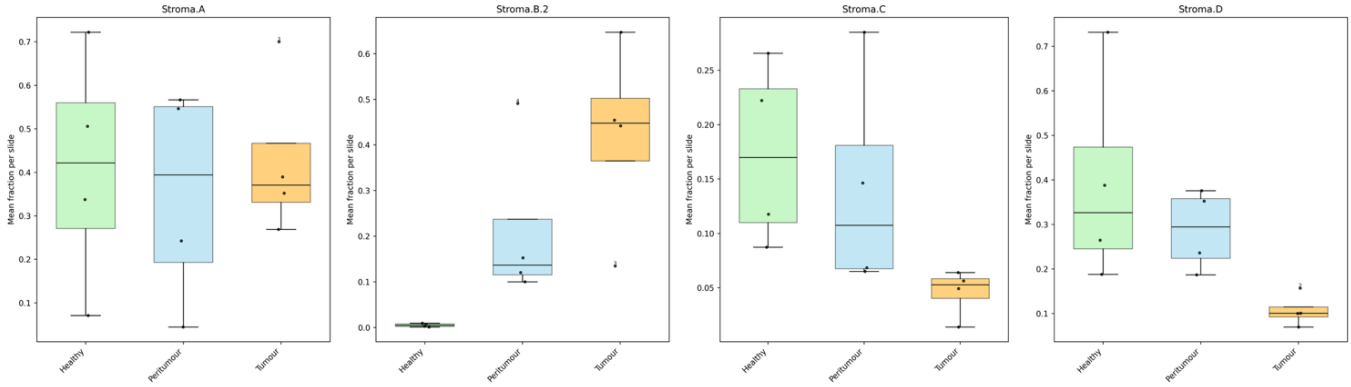


Figure 17. CosMx cohort: Regional variation in CAF subtype proportions within broad CAF-rich compartments across slides. Box plots show slide-level mean normalised fractions of each CAF subtype within the CAF-rich niche, stratified by region. Individual black dots represent individual slide (patient) and denote slide-level means. Outliers were labelled by slide number. Box colours: light green for Healthy, sky blue for Peritumour, and orange for Tumour. None of the region-wise comparisons were significant (paired Wilcoxon signed-rank tests with Benjamini–Hochberg FDR correction), hence no "p-adj" annotation above individual box plots.

UMAPs of niche compositions supported a graded regional structure. Immune-niche embeddings showed a smooth separation with tumour and healthy at opposite ends and peritumour ROIs distributed between them (Fig.18, A). CAF-niche embeddings displayed region-specific segregation between tumour and healthy liver regions, but substantial overlaps of peritumour CAF-niche compositions with both tumour and healthy regions were observed. This pattern suggested that peritumour CAF niches possess a hybrid phenotype, sharing compositional features with both adjacent tissue states rather than existing as a distinct, isolated compartment (Fig.18, B).

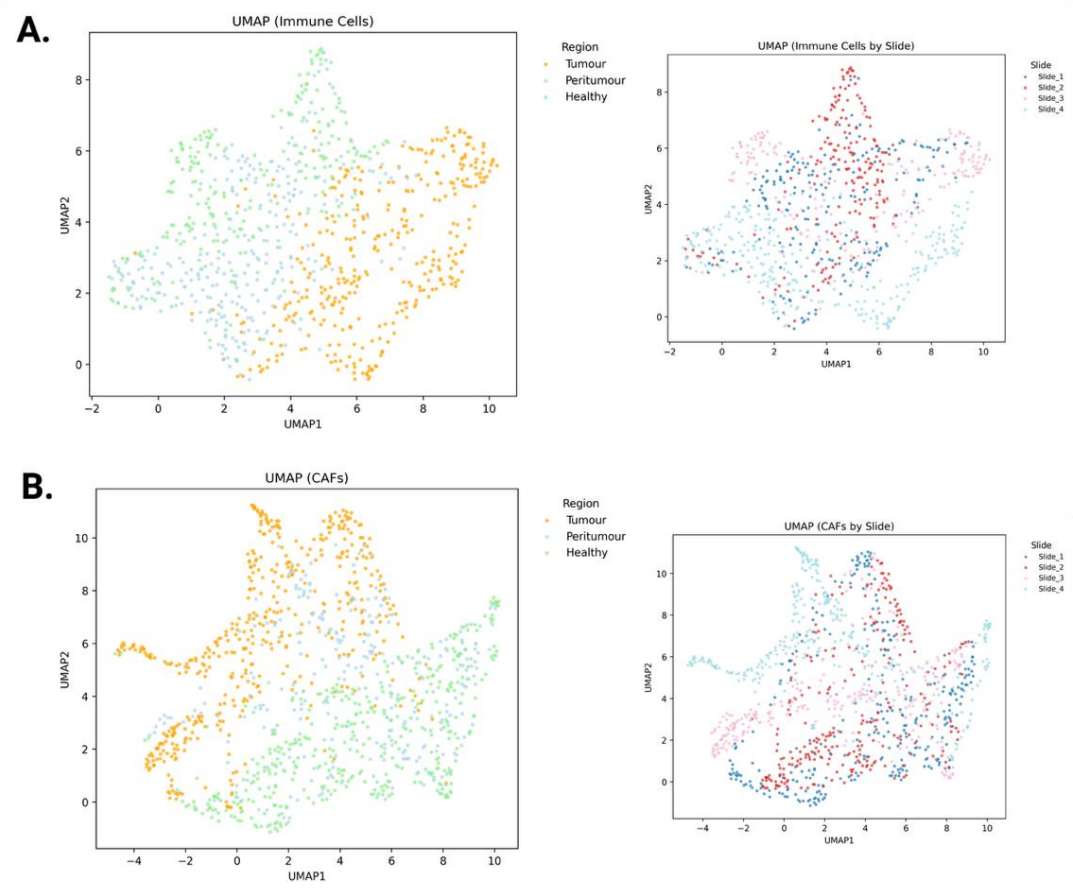


Figure 18. CosMx cohort: UMAP embedding of immune and CAF compartment compositions in CRLM ROIs. A. UMAP of immune-rich ROIs. Points represent individual ROIs, coloured by tissue region (left panel) or by slide (right panel). Tumour and healthy ROIs occupy largely distinct areas of the embedding, with peritumour ROIs occupying an intermediate space, reflecting the graded shift in immune composition across regions. B. UMAP of CAF-rich ROIs. Points are coloured by region (left) or slide (right). ROIs from Tumour and Healthy regions were separated but Peritumour ROIs showed substantial overlap with other two regions in the embedding.

2. Neighbourhood Clusters (NC) in CRLM and Differences in Targeted NC proportions between Three Regions

To move from millions of single cells to a tractable description of tissue ecology, we modelled each cell by the composition of its immediate microenvironment ("neighbourhood") defined by its 10 nearest neighbours (10-NN). This approach generated, for each cell, a vector of local cell-type composition that captures the immediate cellular context rather than identity alone. Clustering these neighbourhood vectors (after batch correction) identifies a small set of recurring spatial neighbourhood types, or "neighbourhood clusters," as described by Schürch et al. (2020). NCs simplify the data by aggregating noisy single-cell variation into stable microenvironmental states, enabling direct comparisons of how often specific spatial niches occur across tumour, peritumour, and healthy liver.

2.1 Cell DIVE cohort

K-means clustering of 10-NN neighbourhood composition vectors (after batch correction) resolved nine recurrent cellular NCs in the Cell DIVE cohort (elbow at k=9; Supplementary Fig.4, A). The z-scored heatmap of average neighbour counts per cluster showed clear, interpretable signatures (Fig.19, A): two hepatocyte-centred neighbourhoods (NC2/6); two tumour-cell–dominated neighbourhoods (NC1/8); a neutrophil-predominant neighbourhood (NC4); two T-cell–rich neighbourhoods (NC0 with a mixed proliferation/exhaustion profile (exhausted and proliferating $CD4^+$/$CD8^+$ T-cells together with Tregs and $aSMA^+FAP^+$ CAFs) and NC5 with an activation-skewed profile (activated $CD4^+$/$CD8^+$ T-cells, T-regs, and PD-1-high exhausted $CD4^+$ T-cells)); and a stromal–immune cell interface neighbourhood (NC7) comprising multiple CAF subsets (FAP^+, $aSMA^+$, $aSMA^+FAP^+$, lymphatic CAFs) plus endothelial cells, macrophages and $CD8^+$ T-cell subsets. Regionally, peritumour tissue displayed the broadest mixture of NCs**,** healthy liver was dominated by the hepatocyte NCs (NC2/6), and the tumour core by tumour-cell NCs (NC1/8) (Fig.19, B).

Focusing on immune/CAF-relevant NCs (NC0, NC4, NC5, NC7), paired Wilcoxon signed-rank tests at the slide level revealed robust regional differences (Fig.19, C). NC0 was significantly more abundant in peritumour than tumour and likewise higher in healthy than tumour (adjusted $p < 0.05$), with no significant difference between peritumour and healthy. This suggested that T-cell activity–enriched niches (proliferation and exhaustion) were maintained outside the tumour core but depleted within it. NC4 showed a progressive increase across regions, being significantly higher in tumour and peritumour than in healthy, while the difference between peritumour and tumour was not significant. This pattern reflected a regional gradient in neutrophil-dense niches peaking within the tumour microenvironment. NC5 was significantly enriched in peritumour relative to both healthy and tumour, pointing to accumulation of T-cell activation at the invasive margin. NC7 was significantly more abundant in both peritumour and tumour regions relative to healthy tissue, with a non-significant trend towards greater abundance in peritumour than tumour. Collectively, these findings reveal a spatial transition in niche composition across the tumour border and indicate that CRLM does mount an immune response. However, this response is maintained predominantly in the peritumour region rather than the tumour core, as evidenced by the enrichment of T-cell activation/exhaustion neighbourhoods and the expansion of the stromal–immune interface. This pattern mirrors the compartmentalised immune landscapes previously described in CRC (Schürch *et al.,* 2020; Pelka *et al.,* 2021).

UMAPs of ROI-level targeted-NC proportions did not form sharply separated regional clusters, rather, ROIs from different regions were largely intermixed (Fig.19, D). This lack of distinct regional clustering suggests that targeted NC proportions were insufficient to spatially segregate regions, likely because most neighbourhood types were shared across all regions and differed primarily in relative frequency rather than in their categorical presence.

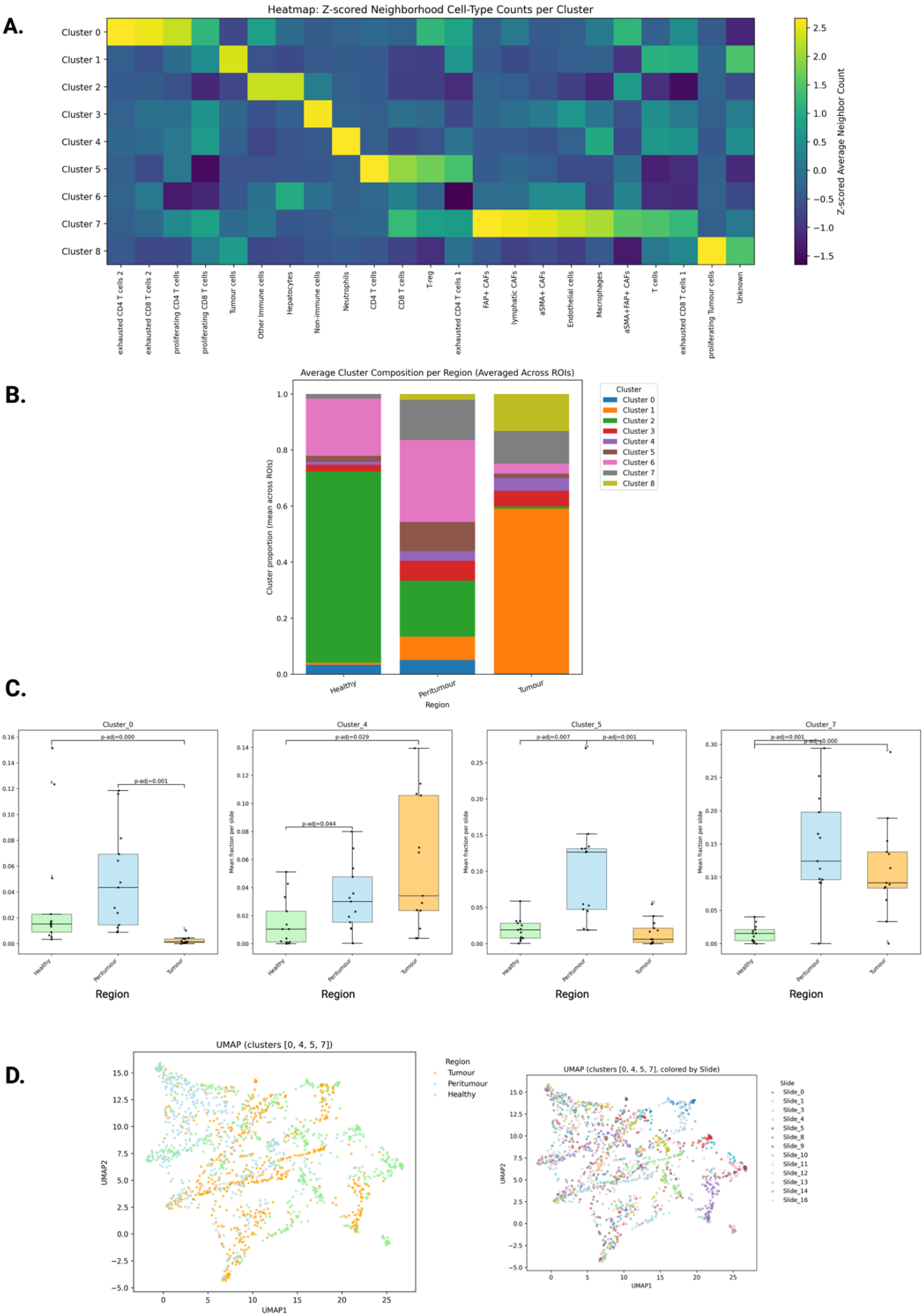


Figure 19. Identification and regional distribution of cellular neighbourhood clusters (NC) in CRLM by Cell DIVE. A.

Heatmap of neighbourhood cluster cell-type composition. Each row represents one K-means–defined NC, and each column a targeted cell type. Values indicate the z-scored mean number of each cell type per neighbourhood within that cluster. B. Average cluster composition by region. Stacked bar plots show the average proportion of each NC across all ROIs within each tissue region (healthy, peritumour, tumour). C. Regional comparison of targeted cluster abundances. Box plots display the slide-level proportions of four immune/CAF-relevant neighbourhood clusters (NC0, NC4, NC5, NC7), stratified by region. Each point represents the average proportion for a slide. Horizontal bars and numeric labels above the boxes indicate pairwise statistical comparisons between regions. "p-adj" values represent FDR–adjusted p-values from paired Wilcoxon signed-rank tests. Only comparisons with FDR-adjusted $p < 0.05$ are labelled. D. UMAP embedding of ROI-level targeted NC proportions. UMAP was applied to ROI-level vectors of targeted NC proportions, after batch correction and standardisation. Points represent individual ROIs, coloured by tissue region (left) or by slide (right).

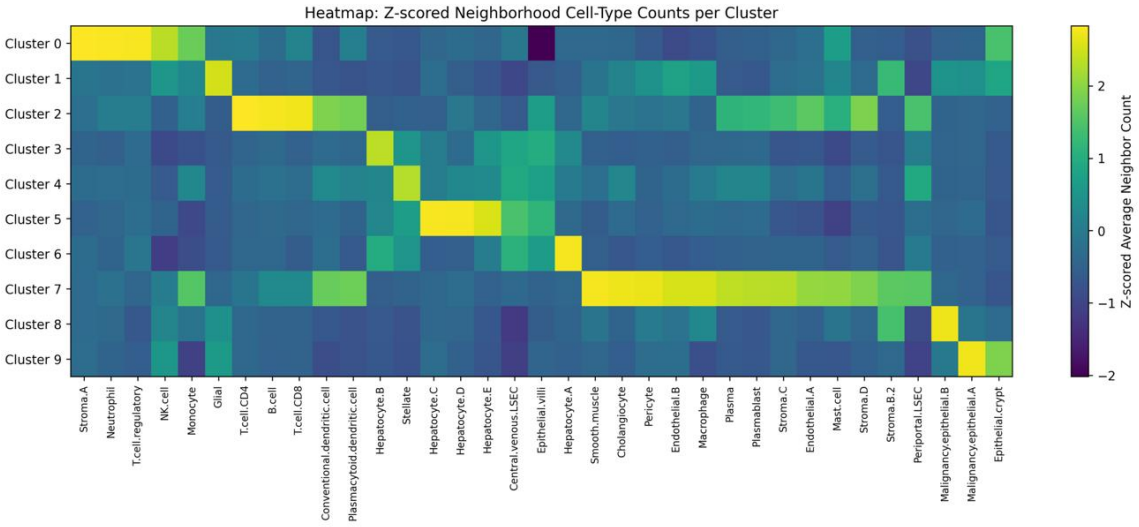
2.2 CosMx cohort

Applying the same analytical workflow to the CosMx dataset, we identified ten recurrent NCs based on K-means clustering (elbow at k=10, Supplementary Fig.4, B). The resulting z-scored heatmap of average neighbour counts per cluster also revealed a diverse and interpretable set of neighbourhood compositions (Fig.20, A), including hepatocyte-centred neighbourhoods (NC3-6) and malignant epithelial neighbourhoods (NC1/8/9), and several immune- and stroma-enriched NCs (NC0/2/7). Specifically, NC0 consisted of neutrophils, T-regs, NK cells, and monocytes, consistent with an immunosuppressive myeloid–lymphoid niche. NC2 reflected an adaptive and antigen-presenting neighbourhood, characterised primarily by $CD4^+$/$CD8^+$ T-cells, B-cells, cDCs and pDCs, along with smaller proportions of plasma cells, plasmablasts, endothelial cells, and two fibroblast subtypes (Stroma C and D). NC7 captured a stromal–immune interface niche, enriched for CAF populations, macrophages, monocytes, mast cells, plasma cells, plasmablasts, and multiple endothelial features. Notably, within NC2 (and to a lesser extent, NC7), the dense colocalisation of cDCs, B-cells, T-cells, and specialised CAFs seemed to reflect the cellular composition hallmarks of tertiary lymphoid structures (TLS), which are ectopic lymphoid aggregates that function as local sites of immune priming and have been associated with enhanced immunosurveillance and improved response to immunotherapy in multiple cancers (Sautès-Fridman *et al.,* 2019; Cabrita *et al.,* 2020; Helmink *et al.,* 2020).

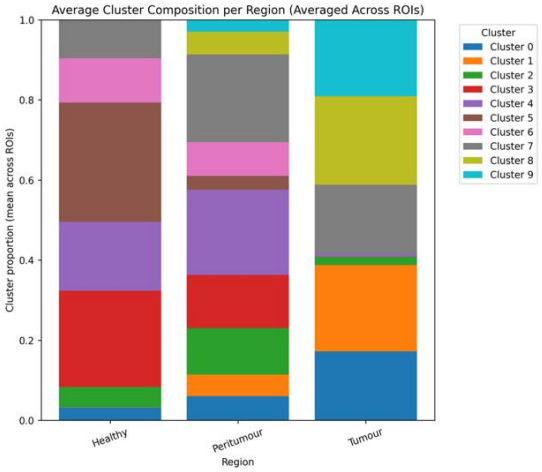
Region-level summaries showed healthy liver dominated by hepatocyte and sinusoidal endothelial NCs, peritumour tissue containing the broadest mixture with strong representation of NC2 and NC7, and tumour regions enriched for malignant epithelial NCs, along with elevated contributions from NC0 and NC7 (Fig.20, B). Although limited to four slides, the qualitative trends observed in slide-level comparisons aligned with the bar-plot-based interpretations (Fig.20, C). NC2 tended to peak in peritumour tissue and decline in tumour, supporting the concentration of adaptive immune hubs with antigen-presenting capability, including potential TLS-like structure formation, at the invasive margin. NC0 tended to increase from healthy to peritumour and was highest in tumour, reflecting a potential gradient of intratumoural immune suppression mediated by myeloid and T-reg programmes. NC7 was elevated in tumour-proximal regions, with peritumour generally exceeding tumour, and both were higher than healthy, indicating expansion of fibroblastic and vascular niches in and around the lesion.

Taken together, CosMx recapitulated the immune-relevant organisation seen with Cell DIVE but with a stronger suggestion of intratumoural immune suppression. The enrichment of NCs composed of myeloid, T-regs, and fibroblastic populations in the tumour core, alongside the restriction of adaptive and antigen-presenting programmes to the peritumour region, supported prior spatial profiling studies in CRC that highlight immune compartmentalisation and suppressive microenvironments (Schürch *et al.,* 2020; Pelka *et al.,* 2021). Lastly, CosMx's UMAPs of ROI-level proportions of targeted NCs (NC0/2/7) also failed to separate regions into discrete groups (Fig.20, D), which was expected when neighbourhood types are shared across regions.

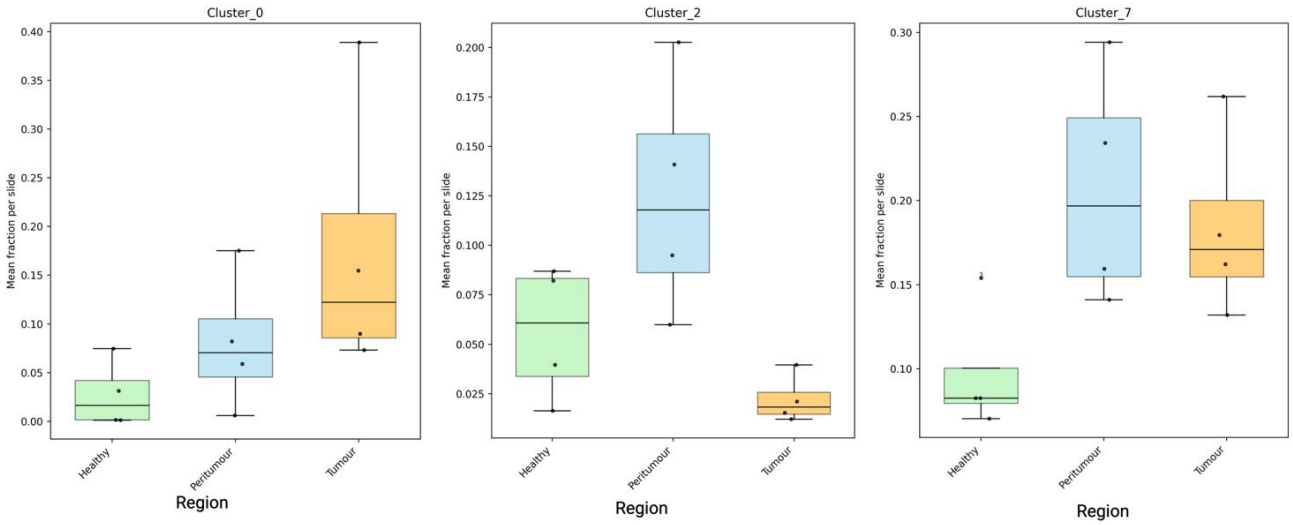
A.


Heatmap: Z-scored Neighborhood Cell-Type Counts per Cluster
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Stroma.A
Neutrophil
T.cell.regulatory
NK.cell
Monocyte
Glial
T.cell.CD4
B.cell
T.cell.CD8
Conventional.dendritic.cell
Plasmacytoid.dendritic.cell
Hepatocyte.B
Stellate
Hepatocyte.C
Hepatocyte.D
Hepatocyte.E
Central.venous.LSEC
Epithelial.villi
Hepatocyte.A
Smooth.muscle
Cholangiocyte
Pericyte
Endothelial.B
Macrophage
Plasma
Plasmablast
Stroma.C
Endothelial.A
Mast.cell
Stroma.D
Stroma.B.2
Periportal.LSEC
Malignancy.epithelial.B
Malignancy.epithelial.A
Epithelial.crypt
Z-scored Average Neighbor Count

B.


Average Cluster Composition per Region (Averaged Across ROIs)
Cluster proportion (mean across ROIs)
Region
Healthy
Peritumour
Tumour
Cluster
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9

C.


Cluster_0
Cluster_2
Cluster_7
Mean fraction per slide
Healthy
Peritumour
Tumour
Region

D.

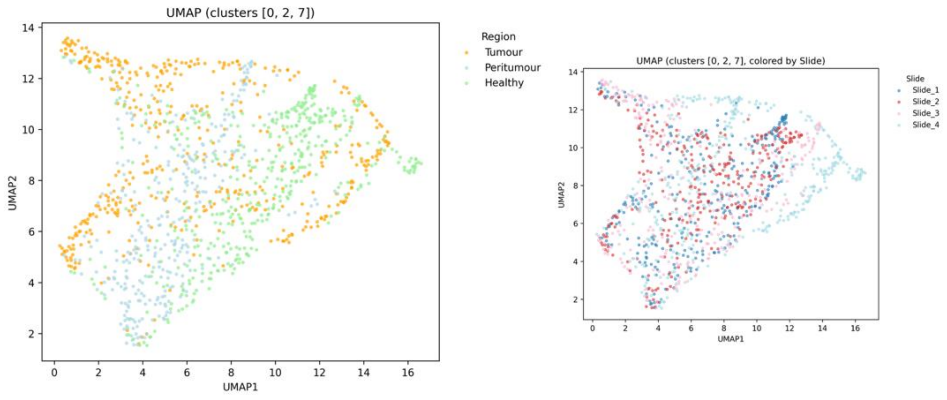

UMAP (clusters [0, 2, 7])
UMAP1
UMAP2
Region
Tumour
Peritumour
Healthy
UMAP (clusters [0, 2, 7], colored by Slide)
Slide
Slide_1
Slide_2
Slide_3
Slide_4

Figure 20. Identification and regional distribution of cellular neighbourhood clusters (NC) in CRLM by CosMx. A. Heatmap of neighbourhood cluster cell-type composition. Each row represents one K-means–defined NC, and each column a targeted cell type. Values indicate the z-scored mean number of each cell type per neighbourhood within that cluster. B. Average cluster composition by region. Stacked bar plots show the average proportion of each NC across all ROIs within each tissue region. C. Regional comparison of targeted cluster abundances. Box plots display the slide-level proportions of four immune/CAF-relevant neighbourhood clusters (NC0, NC2, NC7), stratified by region. Each point represents the average proportion for a slide. None of the region-wise comparisons were significant (paired Wilcoxon signed-rank tests with Benjamini–Hochberg FDR correction), hence no "p-adj" annotation above individual box plots. D. UMAP embedding of ROI-level targeted NC proportions. UMAP was applied to ROI-level vectors of targeted NC proportions, after batch correction and standardisation. Points represent individual ROIs, coloured by tissue region (left) or by slide (right).

3. Short-Range Spatial Interactions Between Cell Types in CRLM

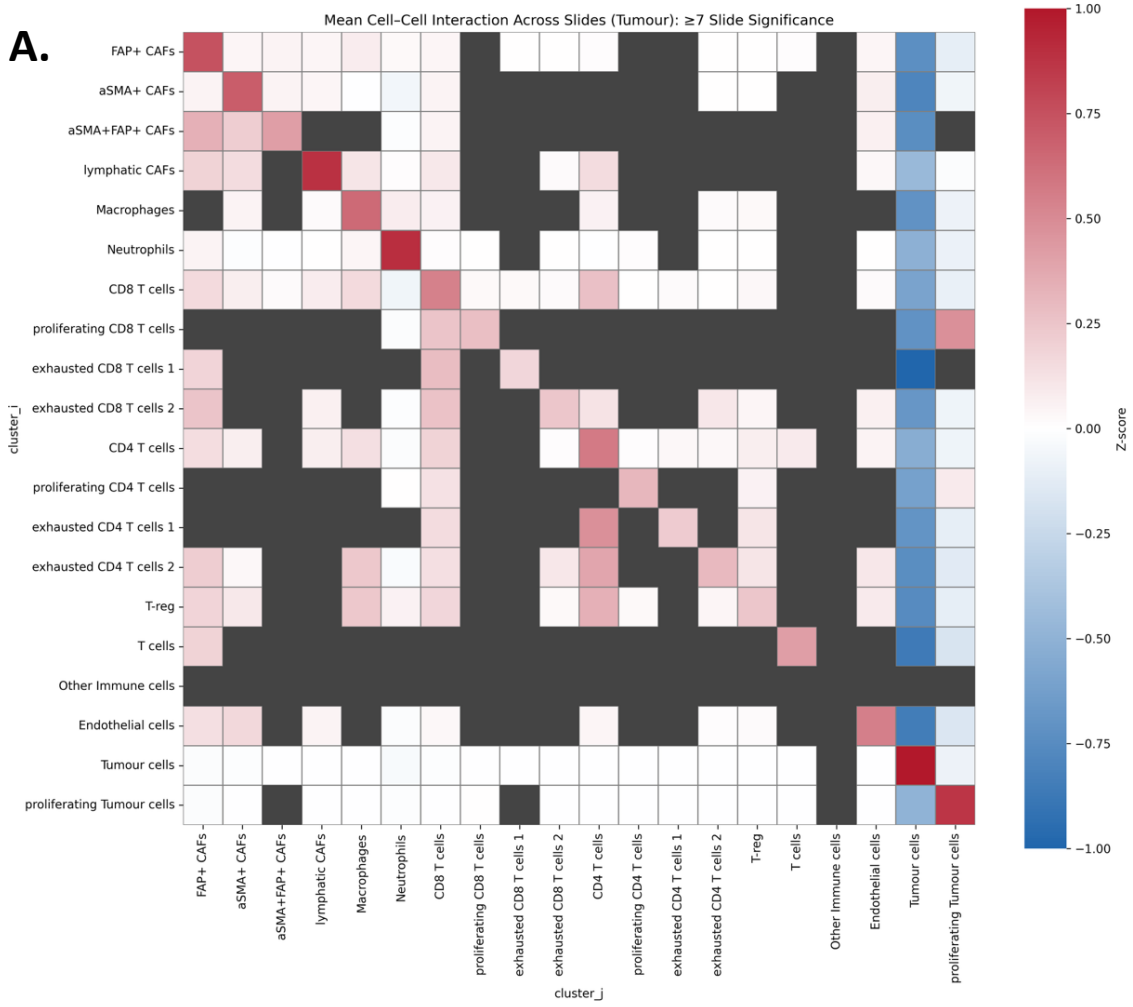
To characterise the spatial organisation of cell types within the CRLM microenvironment, we assessed short-range, directional cell–cell associations using a permutation-controlled KNN proximity test (k=10; see Methods 2.3). For each slide and region, we computed directional interaction scores by comparing observed cell-type i → neighbour-type j neighbour counts to a null distribution derived from 1,000 label permutations. Importantly, interactions were treated as directional (i → j) rather than symmetric to capture potential biological asymmetries in local spatial structuring, such as effector targeting of specific cell types.

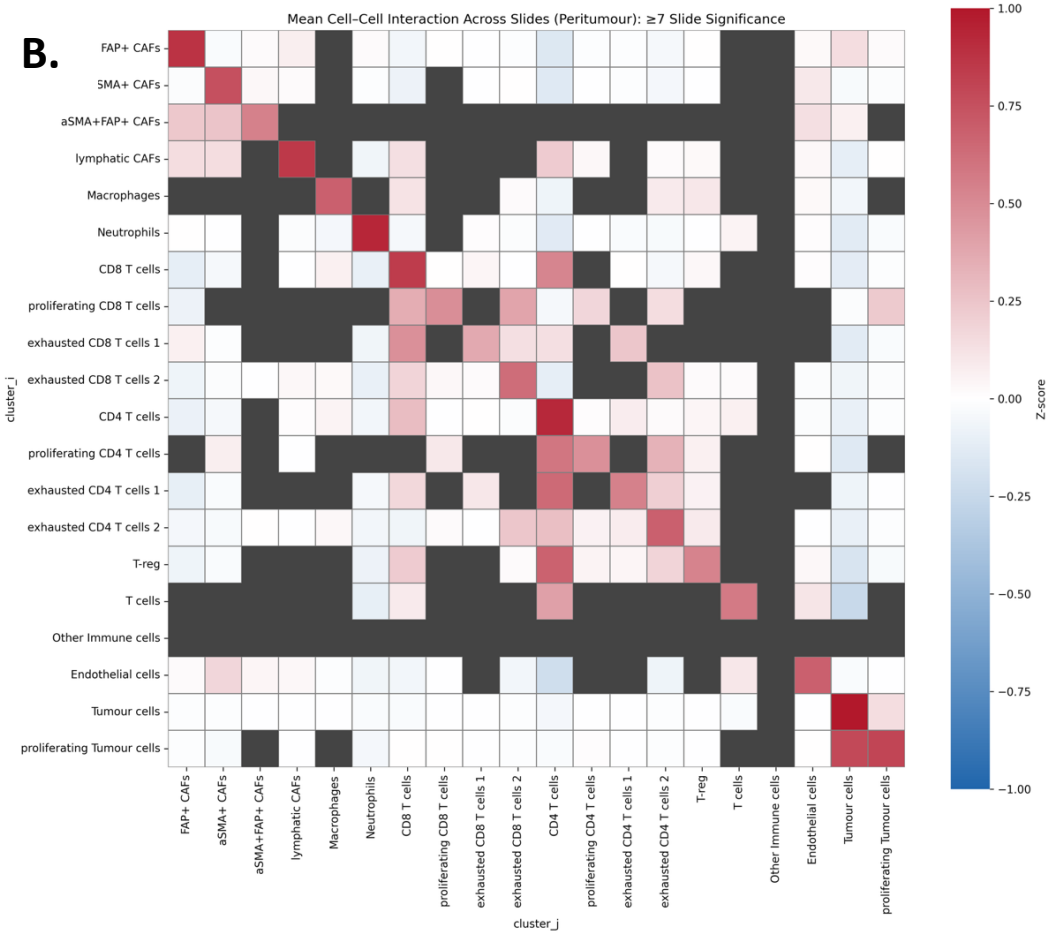
A Z-score reflecting the magnitude and direction of association, along with an empirical permutation-derived p-value, was computed for each ordered cell-type pair. The resulting "interaction" heatmaps visualise enrichment (positive Z-scores) in red, avoidance (negative Z-scores) in blue, and mask non-significant or non-reproducible interactions in grey. These directional proximity maps offer a high-resolution representation of local spatial structuring across regions, complementing compositional and neighbourhood-level analyses.

3.1 Cell DIVE cohort

The tumour core of CRLM exhibited four predominant short-range interaction patterns (Fig.21, A). First, there was widespread avoidance between malignant epithelial cells and multiple T-cell subsets (proliferating, activated, and exhausted $CD4^+$/$CD8^+$ T-cell subsets). Second, relative to other regions, tumour microenvironment showed more pronounced CAF–immune co-localisations, with (FAP^+/$aSMA^+$/$aSMA^+FAP^+$) CAFs interacting at short range with activated $CD8^+$ T-cells and macrophages, alongside enriched CAF–CAF cross-associations. Third, the tumour core demonstrated elevated local interactions between T-regs and other lymphoid subsets, suggesting spatial proximity between immunosuppressive and effector compartments within the immune-excluded tumour core. Fourth, unlike the other two regions which suggested spatial avoidance, FAP^+/$aSMA^+$ CAFs were closely clustered with T-regs in the tumour core. This likely reflects an active immunosuppressive axis in which α-SMA^+ and FAP^+ CAFs promote T-reg recruitment and differentiation via secretion of TGF-β, CXCL12, VEGF, and CCL5 (Liu *et al.,* 2019). Together, these features reinforced the spatially immune-excluded tumour architecture observed in previous analyses and further revealed certain immunosuppressive activities within the tumour.

In the peritumour region (Fig.21, B), tumour–T-cell avoidance was markedly reduced. Instead, the dominant interaction signatures reflected dense intra-lymphoid co-localisation, particularly among $CD4^+$ and $CD8^+$ T-cells across functional states. T-reg interactions with other lymphoid populations persisted, suggesting continued local immunomodulation. In parallel, we observed moderate avoidance between activated $CD4^+$/$CD8^+$ T-cell subsets and CAF subsets (FAP^+, $aSMA^+$, $aSMA^+FAP^+$) and a reduction in intra-CAF associations compared to the tumour core. Collectively, these patterns suggested a more immune-active, yet spatially regulated, microenvironment at the border zone. In the healthy liver region (Fig.21, C), the interaction map showed strong hepatocyte self-clustering with coherent avoidance of immune and endothelial phenotypes, and moderate co-localisations among lymphoid populations, reflecting the expected partitioning of immune and vascular elements in non-malignant tissue.




B.
Mean Cell–Cell Interaction Across Slides (Peritumour): ≥7 Slide Significance
FAP+ CAFs
SMA+ CAFs
aSMA+FAP+ CAFs
lymphatic CAFs
Macrophages
Neutrophils
CD8 T cells
proliferating CD8 T cells
exhausted CD8 T cells 1
exhausted CD8 T cells 2
CD4 T cells
proliferating CD4 T cells
exhausted CD4 T cells 1
exhausted CD4 T cells 2
T-reg
T cells
Other Immune cells
Endothelial cells
Tumour cells
proliferating Tumour cells
cluster_i
cluster_j
Z-score
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

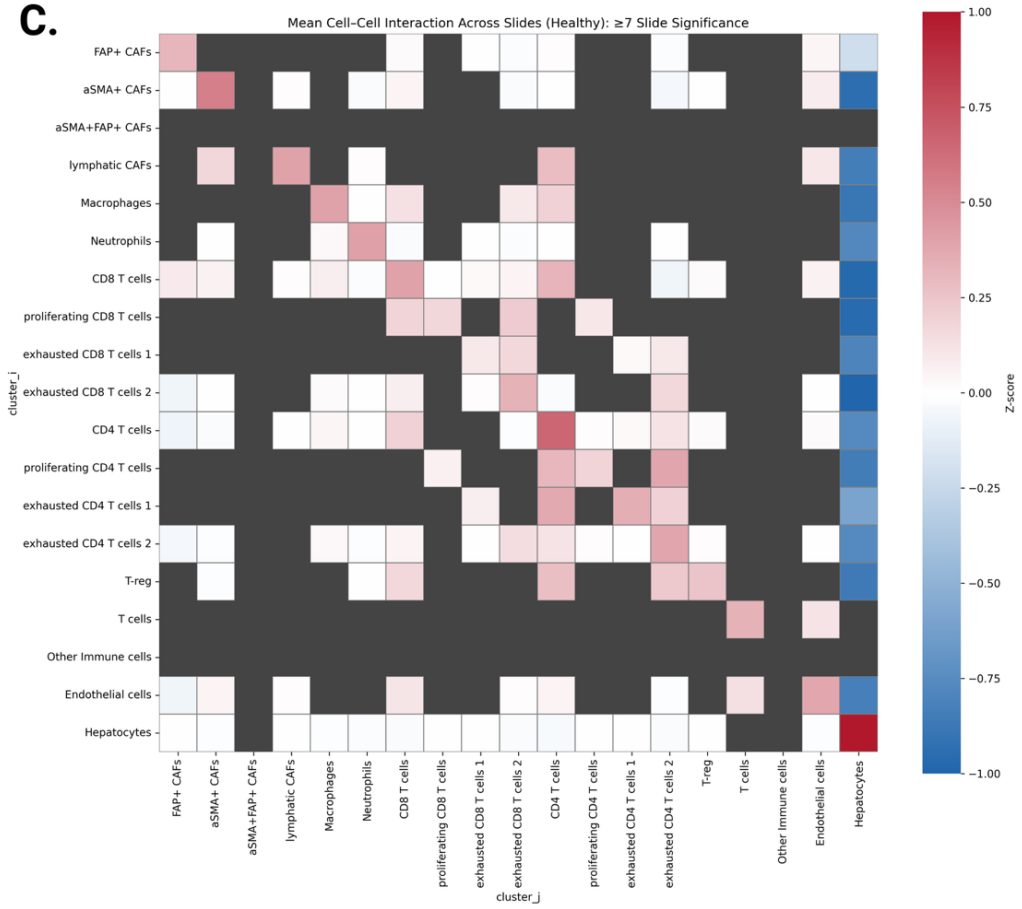
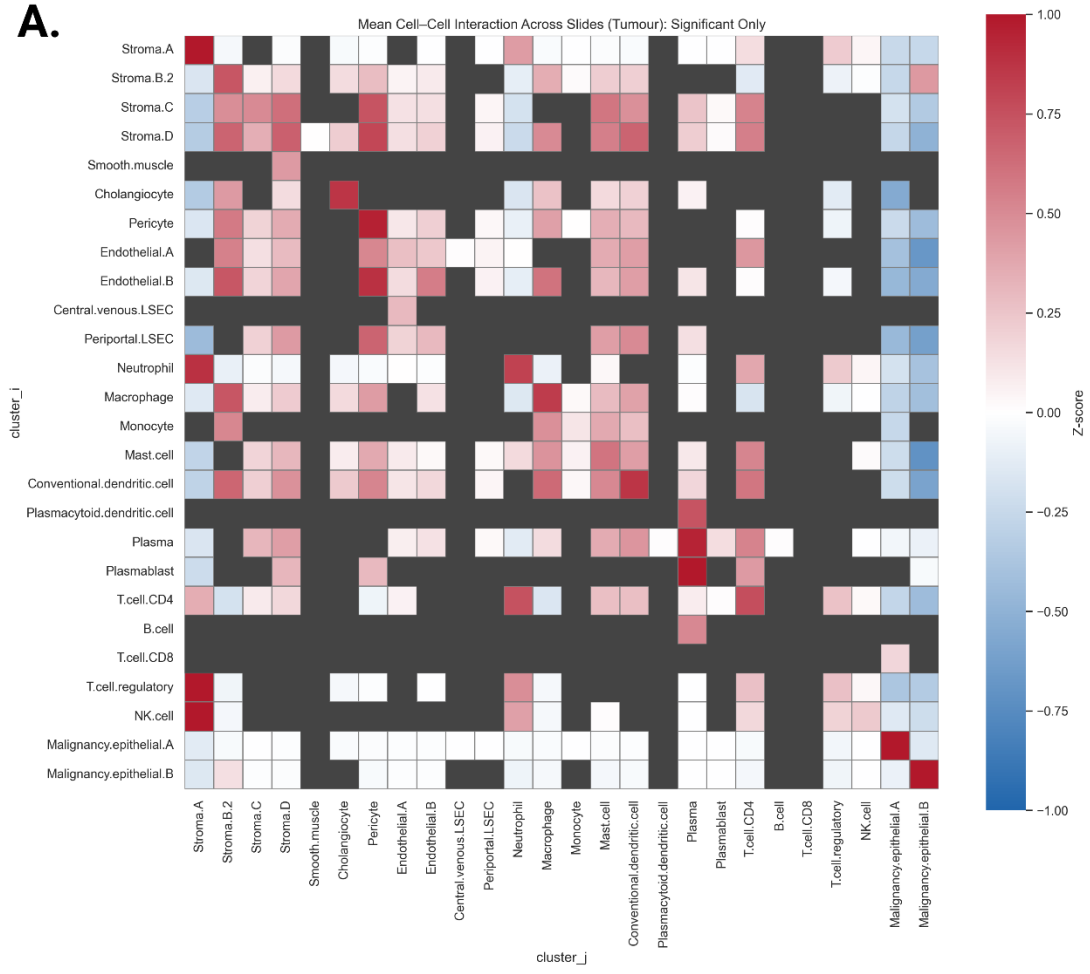
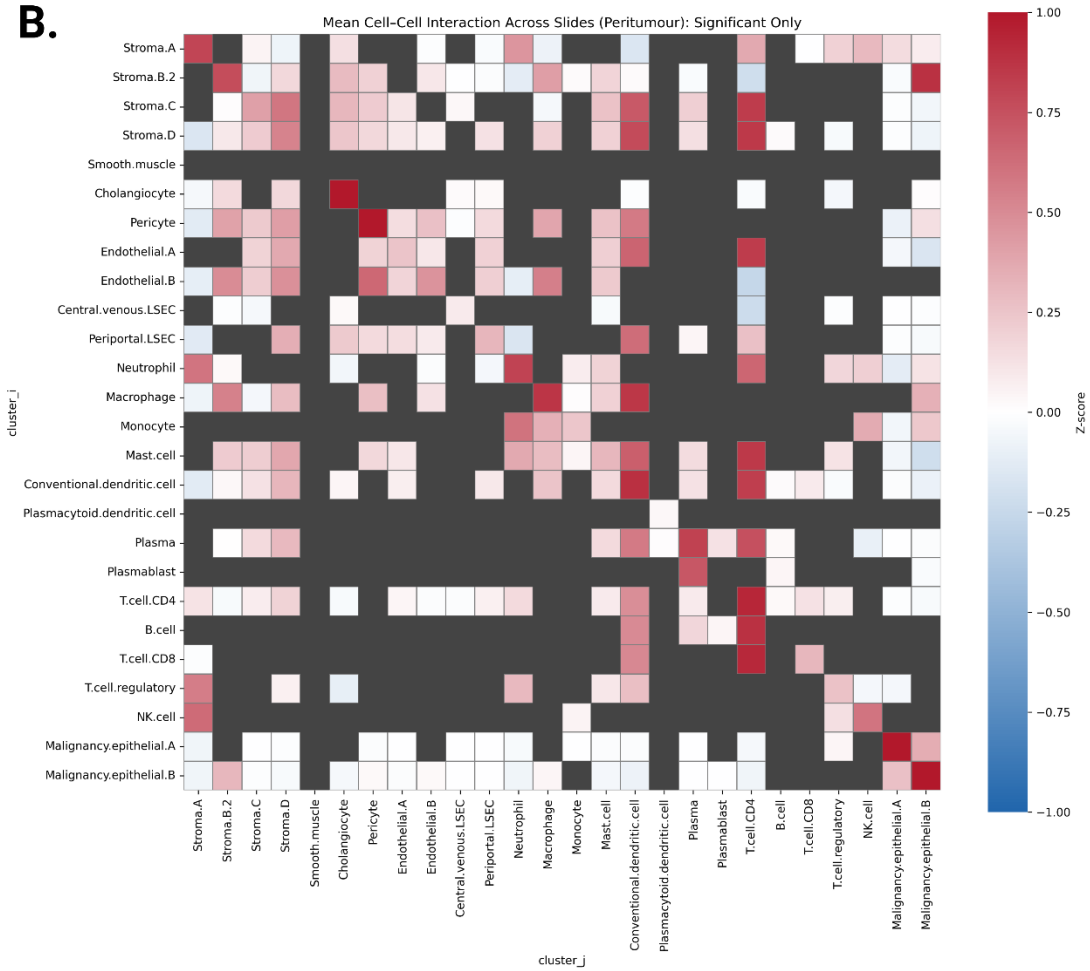

C.
Mean Cell–Cell Interaction Across Slides (Healthy): ≥7 Slide Significance
FAP+ CAFs
aSMA+ CAFs
aSMA+FAP+ CAFs
lymphatic CAFs
Macrophages
Neutrophils
CD8 T cells
proliferating CD8 T cells
exhausted CD8 T cells 1
exhausted CD8 T cells 2
CD4 T cells
proliferating CD4 T cells
exhausted CD4 T cells 1
exhausted CD4 T cells 2
T-reg
T cells
Other Immune cells
Endothelial cells
Hepatocytes
cluster_i
cluster_j
Z-score
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Figure 21. Cell DIVE short-range spatial interaction matrices for (A) tumour, (B) peritumour, and (C) healthy regions. Heatmaps display the mean slide-level Z-scores for all possible ordered cell-type pairs from the targeted cell-type panel, quantifying the enrichment (red) or avoidance (blue) of each cell type as a 10-nearest neighbour of another. Only interactions that reached permutation-derived significance ($p < 0.05$) in at least half of the slides for that region ($n = 7$) are shown. Non-significant cell pairs are masked in grey. The targeted cell-type panel includes CAFs, immune cells, endothelial cells, tumour cells (tumour and peritumour), and hepatocytes (healthy). Each interaction is directional. Positive Z-scores indicate that cell type j is enriched in the 10-NN of cell type i relative to a spatially permuted null.

3.2 CosMx cohort

The higher phenotypic resolution of CosMx resolved a more detailed interaction landscape. In the tumour core (Fig.22, A), malignant epithelial subsets showed widespread short-range avoidance with most immune lineages, consistent with the immune exclusion pattern seen in the Cell DIVE cohort. In contrast, the fibroblastic–myeloid axis was densely interlinked. Stroma B.2/C/D displayed recurrent co-localisation with macrophages, monocytes, mast cells, and DCs, and formed frequent contacts with endothelial subsets and mural cells (e.g., smooth muscle cells and pericytes). CAF–CAF cross-associations were common across B.2/C/D. T-regs were proximally engaged with $CD4^+$ T-cells and NK cells, reinforcing an immunoregulatory niche embedded within a fibroblast-consolidated core. At the peritumour margin (Fig.22, B), malignant-lymphoid avoidance was attenuated, and the dominant signature shifted to lymphoid engagements. Especially, robust short-range interactions were detected between $CD4^+$ T-cells and multiple partners, including cDCs, B-cells, $CD8^+$ T-cells, neutrophils, Stroma C, and Stroma D. B-cell couplings (B-cell↔ cDC/$CD4^+$ T-cells/plasma cells) were also more frequent at the peritumour region. Notably, T-reg engagements and CAF–CAF couplings persisted but with reduced extent relative to tumour, indicating an immune-active yet less fibroblast-consolidated interface. In the healthy liver (Fig.22, C), the heatmap again resolved the expected parenchymal–vascular partitioning. Hepatocyte subsets showed strong self-clustering and coherent avoidance of immune and endothelial phenotypes, whereas cholangiocytes, endothelial subsets, and pericytes formed a tightly interconnected vascular network. Adaptive immune co-localisations and immune cell accumulation around LSEC and vasculatures were also observed, with the correlations being slightly weaker than those seen in the peritumour region, mirroring physiological compartmentalisation.


A.
Mean Cell–Cell Interaction Across Slides (Tumour): Significant Only
cluster_i
cluster_j
Z-score
Stroma.A
Stroma.B.2
Stroma.C
Stroma.D
Smooth.muscle
Cholangiocyte
Pericyte
Endothelial.A
Endothelial.B
Central.venous.LSEC
Periportal.LSEC
Neutrophil
Macrophage
Monocyte
Mast.cell
Conventional.dendritic.cell
Plasmacytoid.dendritic.cell
Plasma
Plasmablast
T.cell.CD4
B.cell
T.cell.CD8
T.cell.regulatory
NK.cell
Malignancy.epithelial.A
Malignancy.epithelial.B
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00


B.
Mean Cell–Cell Interaction Across Slides (Peritumour): Significant Only
cluster_i
cluster_j
Z-score
Stroma.A
Stroma.B.2
Stroma.C
Stroma.D
Smooth.muscle
Cholangiocyte
Pericyte
Endothelial.A
Endothelial.B
Central.venous.LSEC
Periportal.LSEC
Neutrophil
Macrophage
Monocyte
Mast.cell
Conventional.dendritic.cell
Plasmacytoid.dendritic.cell
Plasma
Plasmablast
T.cell.CD4
B.cell
T.cell.CD8
T.cell.regulatory
NK.cell
Malignancy.epithelial.A
Malignancy.epithelial.B
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

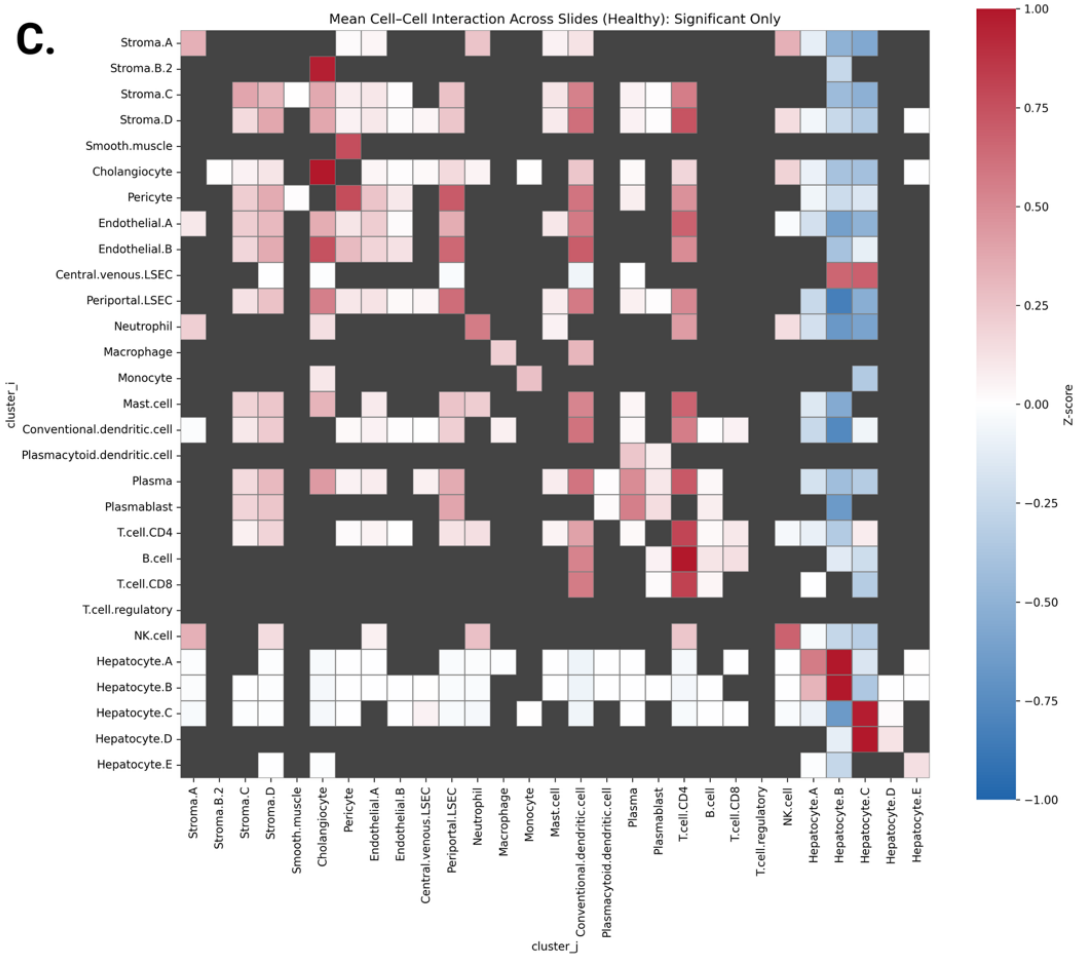


Figure 22. CosMx short-range spatial interaction matrices for (A) tumour, (B) peritumour, and (C) healthy regions. Heatmaps as above, for CosMx spatial transcriptomics data, showing only interactions for which all slides in a region (n = 4) reached permutation-derived significance ($p < 0.05$). The cell-type panel is expanded to include fine-grained immune, stromal, endothelial, malignant, and hepatocyte subpopulations. Directional Z-scores are displayed for each ordered cell-type pair, with significant interactions colour-coded and non-significant pairs masked. This stringent criterion prioritises highly robust, reproducible interactions.

Overall, CosMx confirmed the key spatial themes observed with Cell DIVE, an infiltrated-excluded TIME (Binnewies *et al.,* 2018), including general immune exclusion with concurrent fibroblast–myeloid cell co-localisation in the tumour core, increased lymphoid engagement at the peritumour margin, and a compartmentalised architecture in healthy liver.

4. Distance-Resolved Short-Range Interactions

To add spatial scale to the slide-level KNN heatmaps in Section 3, we computed cross-PCFs for selected, biologically informative pairs. Cross-PCFs report how the local density of a neighbour type varies with distance from an index cell (attraction/clustering when the curve is above the complete-spatial-randomness baseline ($g_{ij}(r)>1$); separation/exclusion when $g_{ij}(r)<1$), thereby revealing the distance over which interactions operate. We show focused panels rather than all pairs to emphasise mechanisms that matter for tumour control, including cytotoxic access, priming/coordination, regulatory proximity, and stromal scaffolding. Only contrasts between tumour core and peritumour

was emphasised in this section because healthy liver forms its own compartment (parenchyma–vascular segregation with modest immune co-localisation), offering limited insight into tumour–immune dynamics (Sections 1–3).

4.1 Cell DIVE cohort

In the Cell DIVE panel, we summarised three sets of distance-resolved interactions that parallel the themes from Section 3 and add the spatial scales at which they operate. The first set highlights immune exclusion and regulatory proximity (Fig.23). $CD8^+$ T-cells showed clear short-range clustering with tumour cells at the peritumour border, most evident within roughly 0–50 µm, whereas the tumour-core profile was flat and remained below the clustering baseline across distances. This indicated restricted CD8 access in the tumour centre and preferential effector engagement at the margin. Proliferating tumour cells were consistently separated from $CD8^+$ T-cells in both regions, indicating poor immunosurveillance and recognition of dividing cancer-cell clones. Tumour–exhausted $CD8^+$ T-cell pairs behaved similarly, where peritumour curves sat above tumour still showed sustained separation, consistent with limited contact even for dysfunctional CD8 states. Regulatory niches were prominent in both regions. T-regs maintained close-range proximity with $CD4^+$ and $CD8^+$ T-cell subsets, including their proliferating counterparts, with peaks within about 10–50 µm. Notably, the peritumour curve remained above the clustering baseline over a longer span than the tumour-core curves, indicating prolonged regulatory adjacency at the border. All these patterns supported restricted effector access in the tumour core and sustained local immunomodulation that persists even where lymphoid activity increases at the peritumour region (Wu *et al.,* 2022a; Binnewies *et al.,* 2018; Mariathasan *et al.,* 2018). Contacts between T-cells and macrophages peaked at small distances of about 10-30 µm in both regions, indicating comparable local coupling to tumour-associated macrophages (TAM). Clustering of $CD4^+$ and $CD8^+$ T-cells with neutrophils also happened around 0-20 µm at the peritumour interface, while the tumour curves suggested avoidance, which might be due to limited infiltration of T-cells. Notably, the specific subtypes of macrophages and neutrophils were not defined in our dataset. Hence, they could be either pro-tumour or anti-tumour (pro-inflammatory) (Mantovani *et al.,* 2017; Yan *et al.,* 2022).

The second set captures engaged immunity (Fig.24). $CD4^+$ and $CD8^+$ T-cells co-clustered, with a stronger and longer-ranged signal in peritumour that extended to roughly 0–140 µm, consistent with coordinated lymphoid activity at the tumour border. The third set focuses on fibroblastic/stromal context (Fig.25). Short-range associations between T cells and CAFs were a defining feature of the tumour core. $CD4^+$ and $CD8^+$ T-cells clustered with FAP^+ and $aSMA^+$ CAFs within approximately 0–40 µm in tumour, whereas the corresponding peritumour curves showed smaller peaks or suggested exclusion throughout. By contrast, T-cell interactions with $aSMA^+FAP^+$ CAFs tended to show exclusion in both regions. Moreover, CAF cross-associations were stronger in the tumour. Pairings such as $aSMA^+$ with FAP^+ CAFs showed pronounced sub-50 µm clustering in the tumour core, while this $aSMA^+ \leftrightarrow FAP^+$ pairing was sparser at the tumour border. This pattern may suggest a CAF-driven exclusion barrier that sequesters T-cells within a fibroblastic meshwork and limits their engagements with tumour cells. This characteristic of TME has been linked to T-cell exclusion and reduced response to ICIs in some previous studies (Bagaev *et al.,* 2021; Feig *et al.,* 2013; Mariathasan *et al.,* 2018).

Taken together, the cross-PCF profiles reinforce the spatial dichotomy described earlier. Peritumour tissue concentrated engaged lymphoid programmes, including CD8⁺ T-cell proximities to malignant cells, T-cell co-clustering, and T-cell coupling to macrophages, all operating over short length scales. The tumour core, in contrast, was marked by fibroblast consolidation, persistent immune↔CAF contacts, and sustained regulatory proximity, a configuration that aligned with immune exclusion and local immunosuppression.

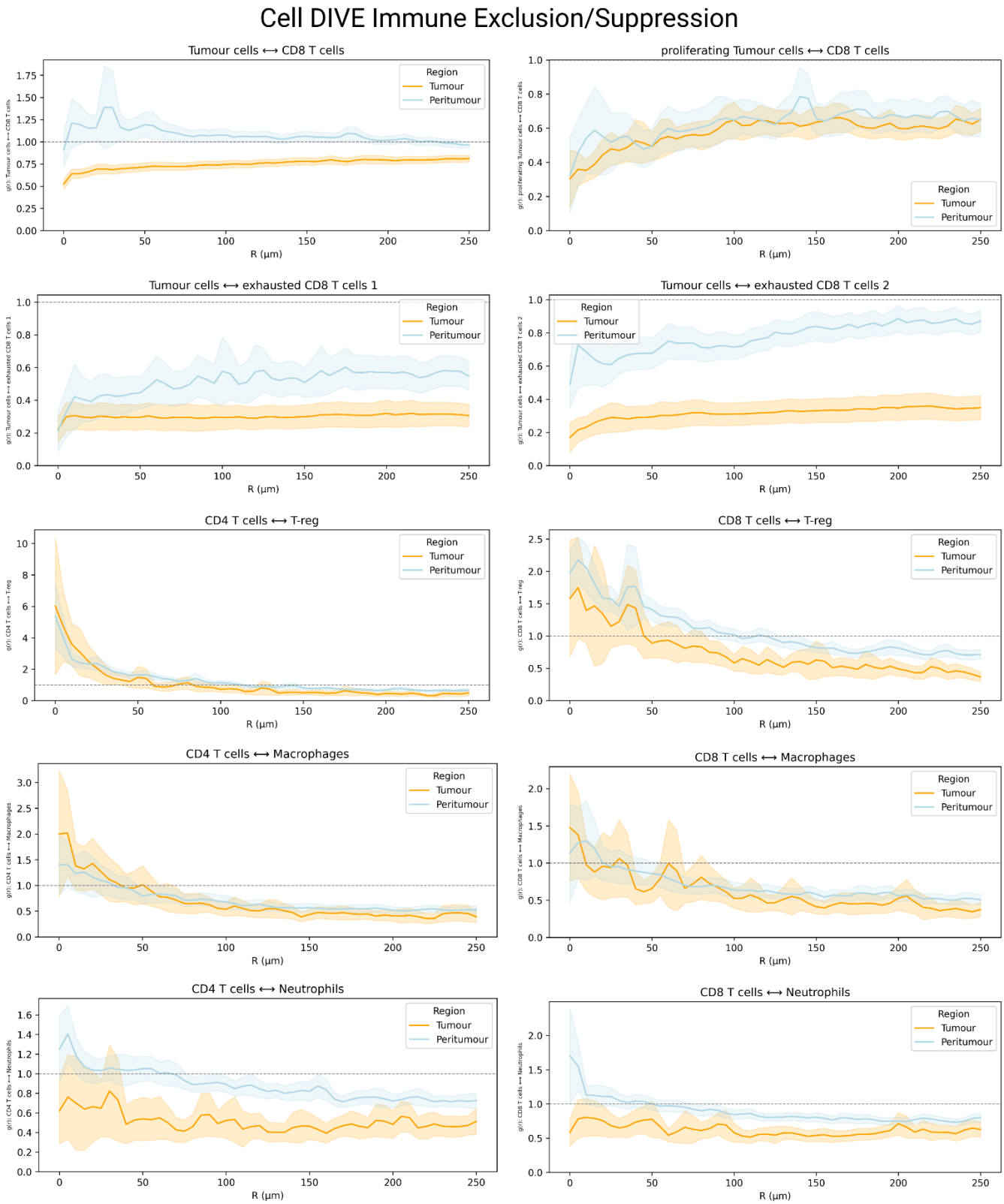


Figure 23. cross-PCF analysis of distance-resolved cell–cell interactions in CRLM (Cell DIVE cohort Panel 1: Immune Exclusion and Suppression). Each subplot displays the distance-dependent association between a pair of cell types, with only single-direction plots shown here (as the corresponding reversed plots exhibited the same patterns). The plots include region-level mean curves (coloured lines) and 95% confidence intervals (the coloured band around each line). The x-axis shows the radial distance R (µm) between cell pairs, the maximum R is 250 µm. The y-axis represents the cross-PCF ($g_{ij}(r)$) function, where values above 1 indicate spatial clustering of the two cell types at distance R, values below 1 indicate spatial separation, and a value of 1 corresponds to a random, uncorrelated spatial distribution. The horizontal dashed line marks the threshold for spatial randomness ($g_{ij}(r) = 1$). Orange: Tumour, Sky blue: Peritumour.

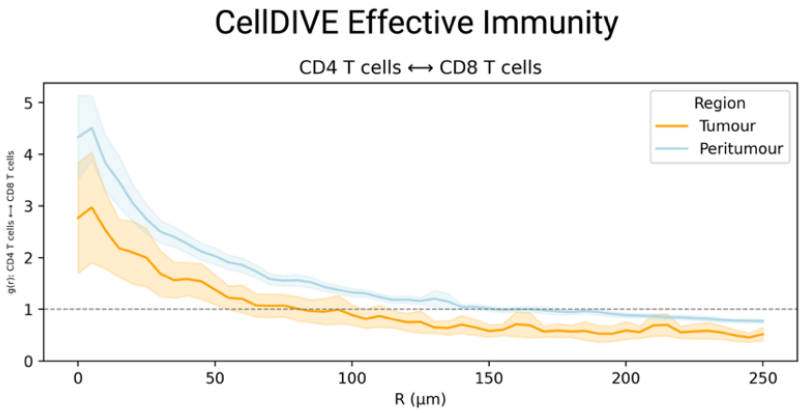


Figure 24. cross-PCF analysis of distance-resolved cell–cell interactions in CRLM (Cell DIVE cohort Panel 2: Effective Immunity).

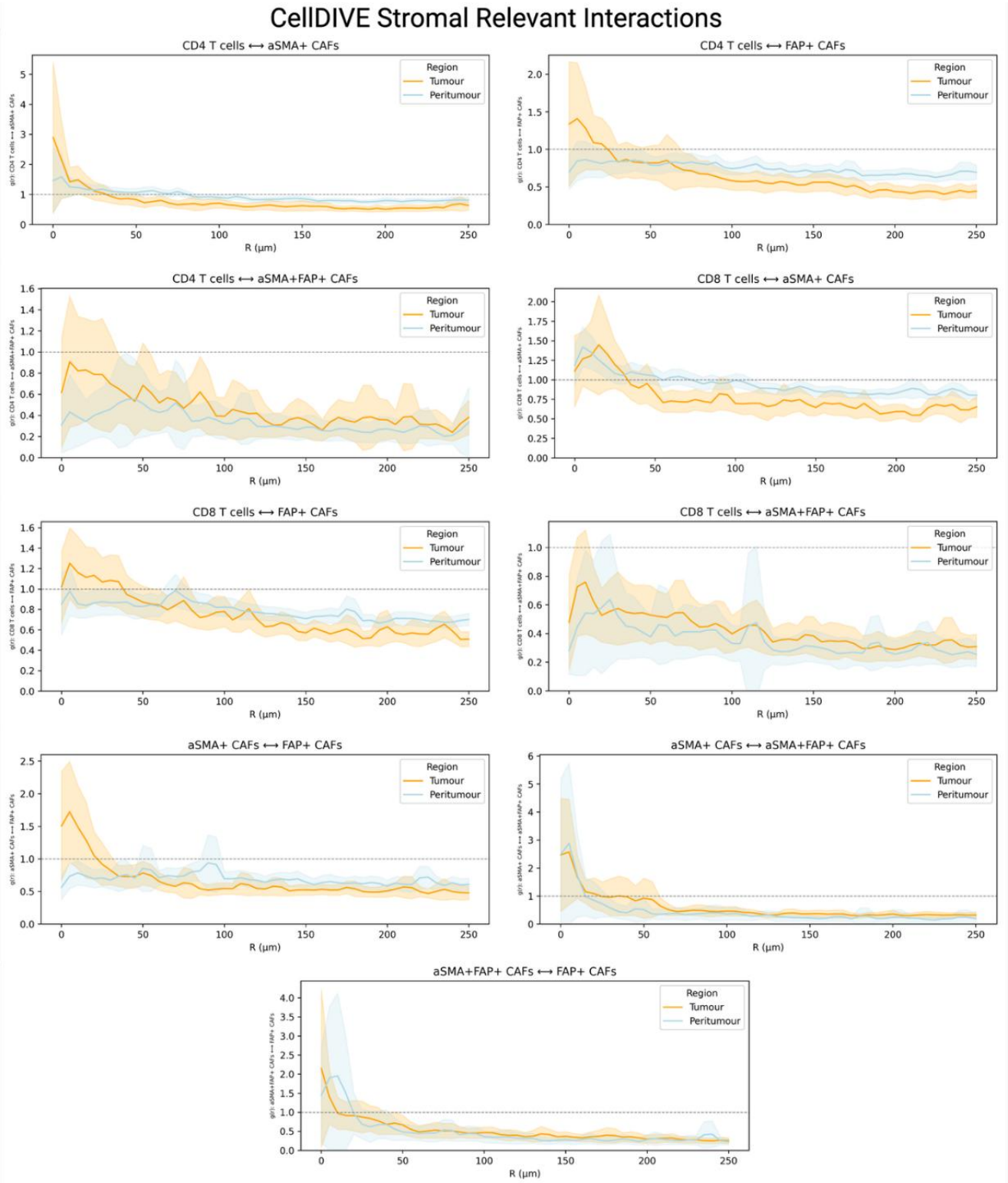


Figure 25. cross-PCF analysis of distance-resolved cell–cell interactions in CRLM (Cell DIVE cohort Panel 3: Stromal-Relevant Interactions).

4.2 CosMx cohort

In the CosMx panel, cross-PCFs again revealed the distance of enrichment or exclusion. The immune suppression and exclusion panel is shown in Figure.26. Direct contact between malignant epithelium and cytotoxic effectors was limited in both tumour and peritumour. $CD8^+$ T-cells exhibited a sustained exclusion pattern from malignant epithelial subsets, with peritumour curves consistently higher than

tumour but still indicative of spatial separation. NK↔Malignancy profiles followed a similar trend, with tumour curves slightly above peritumour yet still showing net avoidance. These patterns indicated a shared constraint on tumour–cytotoxic contact across tissue regions. Meanwhile, regulatory proximity was persistent. The $CD4^+$↔T-reg pair exhibited strong co-clustering with similar amplitudes in both regions, lasting over 0–90 µm. $CD8^+$ T-cell↔T-reg proximity was also evident peritumourally but shifted toward separation in the core. These results confirmed that regulatory niches persisted alongside increased lymphoid engagement in the peritumour region. They suggested that limited $CD8^+$ T-cell infiltration at the tumour core might be due to the suppression of $CD8^+$ T-cells at the peritumour area. Interestingly, NK↔T-reg proximity was comparatively higher at short range in tumour than peritumour, consistent with NK suppression by TGF-β–rich microenvironments (Viel *et al.,* 2016). Lymphoid-Myeloid contacts were heterogeneous. $CD4^+$ T-cell↔Macrophage clustering hovered just above neutral in both regions, while $CD8^+$ T-cell/NK↔Macrophage profiles indicated broad exclusion. $CD4^+$ T-cell–monocyte interactions showed sharp short-range clustering in both regions, with a slight tumour advantage. $CD8^+$ T-cell/NK↔Monocyte proximities remained largely non-associative, but a slightly higher clustering within very close distances at the tumour core could still be observed. $CD4^+$ T-cell/NK↔Neutrophil pairs exhibited clear immediate-distance clustering that relaxed by around 100–150 µm, with slightly stronger tumour signals, while $CD8^+$ T-cell↔Neutrophil curves reflected sustained avoidance. As in Cell DIVE, we did not know the specific subtypes of macrophages and neutrophils, and we could not confirm whether these myeloid cells were pro-tumour or anti-tumour.

Within the effective and engaged immunity axis (Fig.27), lymphoid–APC and lymphoid–lymphoid couplings were predominantly peritumour-biased at short distances. $CD8^+$ T-cells co-clustered with cDCs within 10-50 µm at the peritumour region, with curves consistently above those from the tumour across the full scale, indicating enhanced antigen-presenting engagement. $CD4^+$ T-cell↔cDC coupling was strong in both regions, with a slight peritumour advantage that persisted beyond 50 µm. $CD4^+$ T-cell↔B-cell and $CD4^+$ T-cell↔$CD8^+$ T-cell pairs exhibited pronounced peritumour co-clustering within the first 50 µm, while tumour profiles suggested spatial separation, consistent with localised TLS-like coordination at invasive fronts, which involves active antigen presentation and $CD8^+$ T-cell priming (Sautès-Fridman *et al.,* 2019; Helmink *et al.,* 2020). Whereas the attenuation of couplings between cDCs, $CD4^+$ T-cells, $CD8^+$ T-cells, and B-cells in the tumour core may indicate loss of or contain fewer TLS-like niches. Although great variances were seen in the proximity between NK cells and cDCs across all ROIs, NK↔cDC couplings appeared to be slightly more significant in the tumour region, suggesting potential attractions of cDCs to the tumour core by NK cells (Böttcher *et al.,* 2018a).

Fibroblast associations also displayed marked regional and subtype specificity (Fig.28). Stroma D and Stroma C showed strong peritumour clustering with $CD4^+$ T-cells at short distances, with Stroma D also forming a steep short-range peak with $CD8^+$ T-cells within around 0–20 µm at the border, suggestive of localised antigen-presenting activities involving cytotoxic lymphocytes and supported that specific fibroblast subsets mediate the formation of TLS. Most NK cell↔ fibroblast interactions were broadly avoidant in both regions, indicating minimal NK integration into stromal contexts. $CD4^+$/$CD8^+$ T-cell interactions with Stroma B.2 were largely neutral or avoidant across both regions,

offering limited evidence for directional engagement. Fibroblast cross-associations, in contrast, were consistently stronger in tumour than peritumour, supporting the presence of stromal frameworks that reinforce immune exclusion in the tumour core (Bagaev *et al.,* 2021; Mariathasan *et al.,* 2018; Binnewies *et al.,* 2018).

Overall, the CosMx dataset revealed that lymphoid priming and coordination, marked by enriched T-cell↔DC, T-cell↔T-cell, T-cell↔B-cell, and T-cell↔fibroblast interactions, were concentrated at the peritumour border. While the tumour core retained features of immune exclusion and suppression, including CTL/NK cell separation from malignant cells, localised regulatory niches, and fibroblast consolidation. Together with Cell DIVE, these patterns argued that spatial barriers to immunity in CRLM are maintained by both suppressive cellular proximities and a compartmentalised tissue architecture.

CosMx Immune Exclusion/Suppression

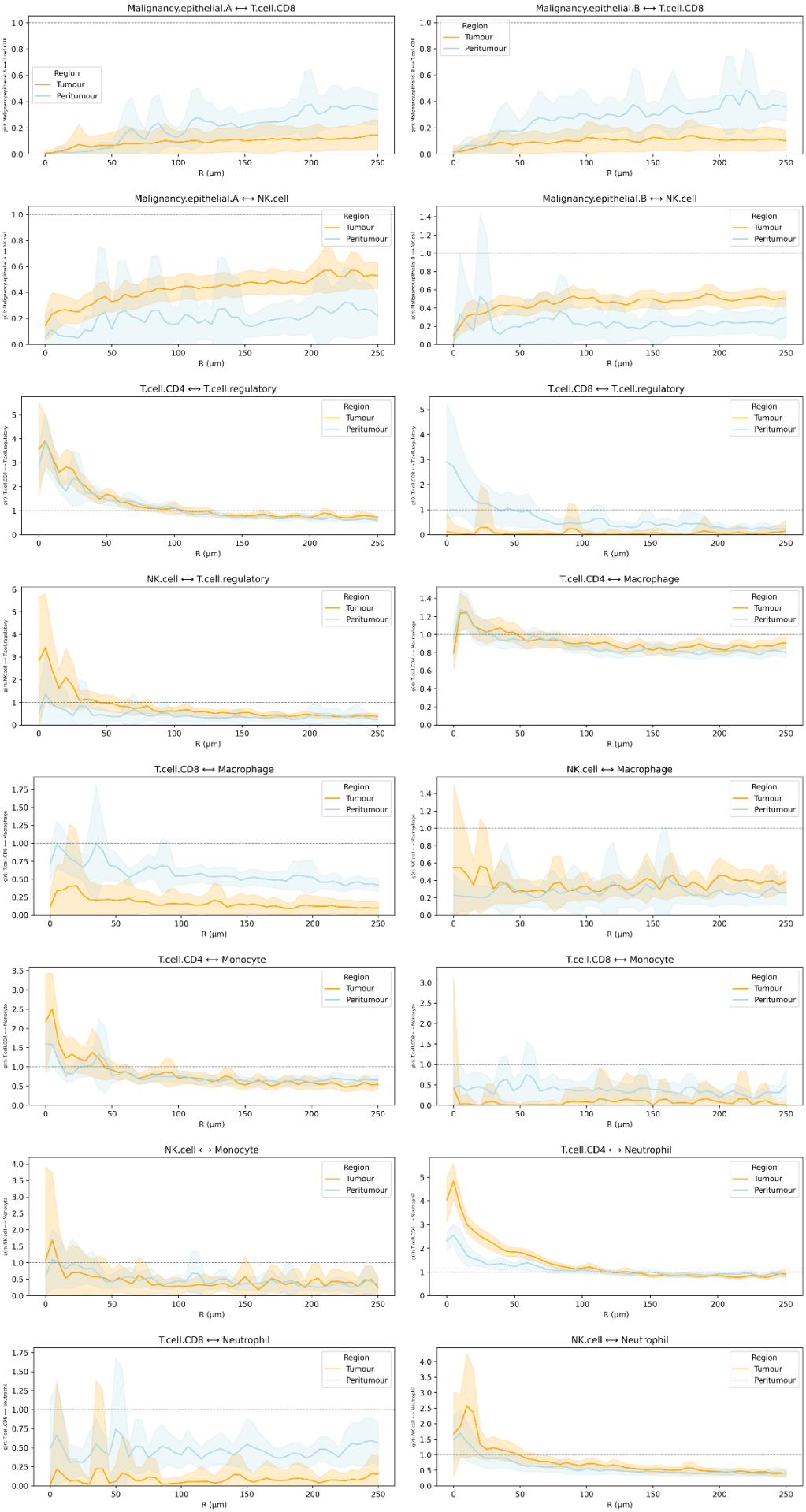


Figure 26. cross-PCF analysis of distance-resolved cell–cell interactions in CRLM (CosMx cohort Panel 1: Immune Exclusion and Suppression). Each subplot displays the distance-dependent association between a pair of cell types, with only single-direction plots shown here (as the corresponding reversed plots exhibited the same patterns). The plots include region-level mean curves (coloured lines) and 95% confidence intervals (the coloured band around each line). The x-axis shows the radial distance R (µm) between cell pairs, the maximum R is 250 µm. The y-axis represents the cross-PCF ($g_{ij}(r)$) function, where values above 1 indicate spatial clustering of the two cell types at distance R, values below 1 indicate spatial separation, and a value of 1 corresponds to a random, uncorrelated spatial distribution. The horizontal dashed line marks the threshold for spatial randomness ($g_{ij}(r) = 1$). Orange: Tumour, Sky blue: Peritumour.

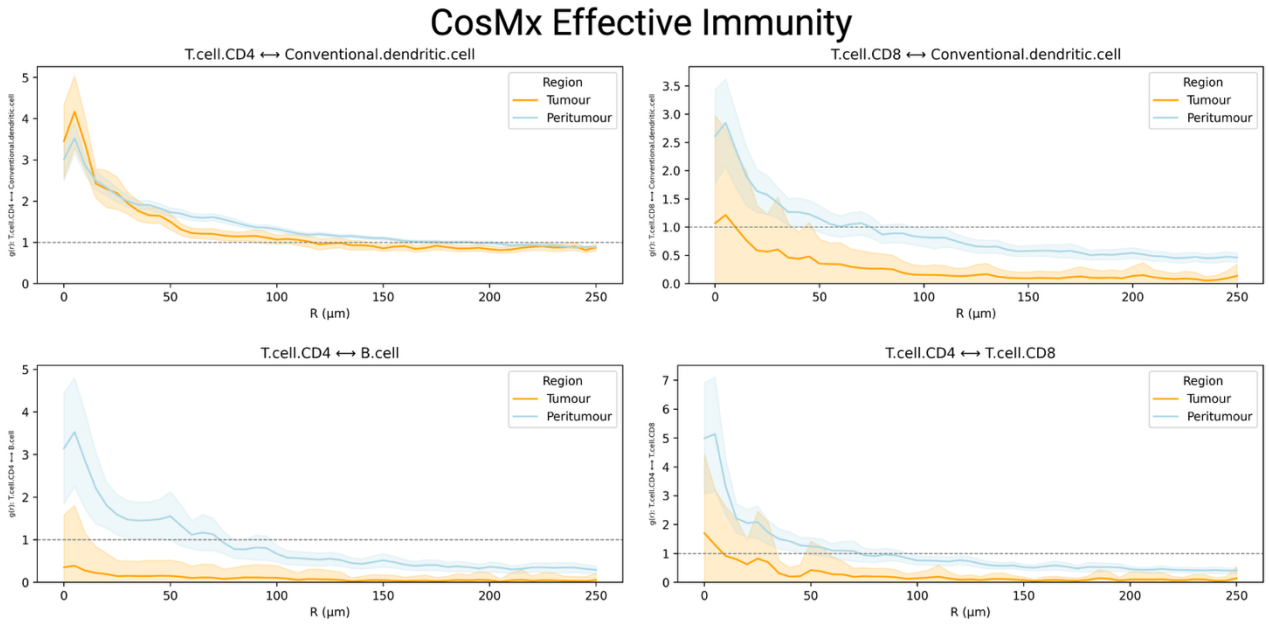


Figure 27. cross-PCF analysis of distance-resolved cell–cell interactions in CRLM (CosMx cohort Panel 2: Effective Immunity).

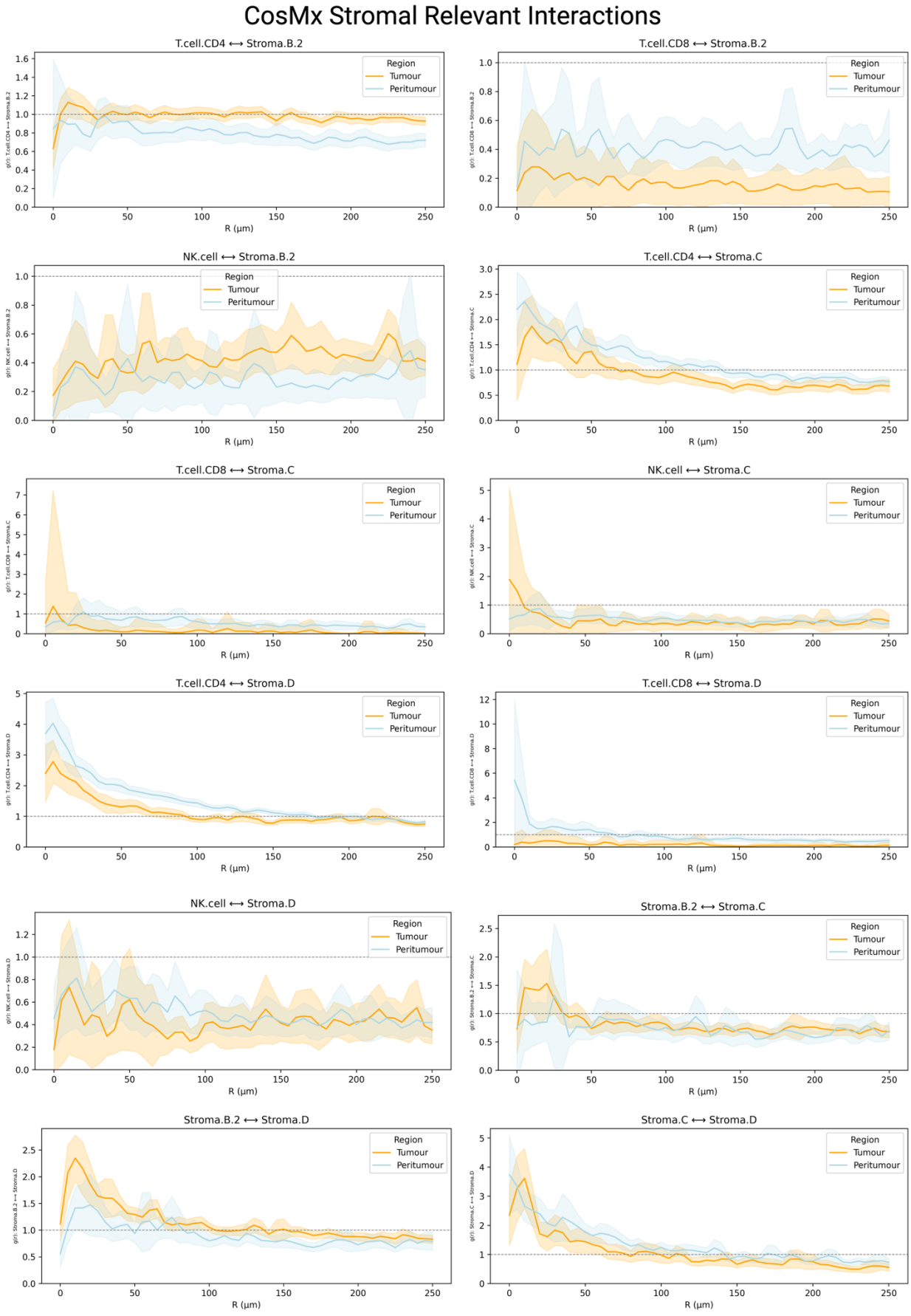


Figure 28. cross-PCF analysis of distance-resolved cell–cell interactions in CRLM (CosMx cohort Panel 3: Stromal-Relevant Interactions).

5. Topology of Neighbourhood Organisation

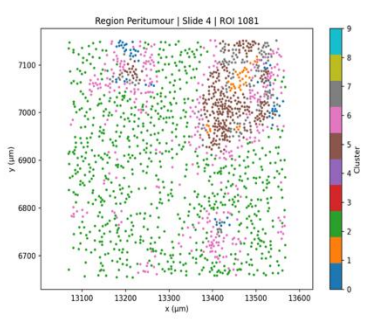
While cell type and neighbourhood proportions, KNN interactions, and cross-PCF analyses reveal which cell types are present, their proximity, and the distances at which they interact, these approaches do not capture how local neighbourhoods are organised into larger geometric patterns. For instance, they cannot distinguish whether T-cell neighbourhoods exist as many disconnected islands or as a mostly continuous sheet interrupted by small voids, nor can they assess the stability of these patterns across spatial scales. TDA addresses this limitation by summarising higher-order spatial organisation in a way that is robust to rotation, translation, and modest sampling noise, thus providing complementary insights beyond those available from pairwise proximity or simple compositional analyses (Ali *et al.,* 2023).

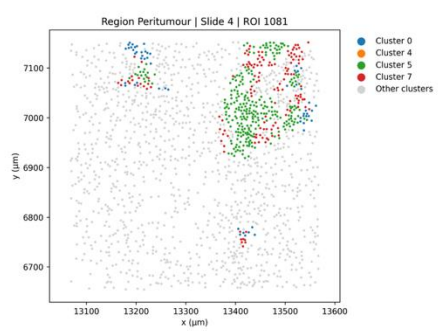
We used persistent homology-based TDA to characterise the geometry of NCs identified in result section 2. Starting from the centroids of all cells in a chosen NC (or a two-cluster combination), we grew a Vietoris–Rips filtration by increasing the connection distance r. As r expands, points that fall within the threshold link, and the persistence diagrams record two kinds of structure: H_0 bars track how many disconnected "islands" exist and how long they survive before merging; H_1 bars capture the appearance and disappearance of loops, i.e., ring-like voids. To keep the readout biologically interpretable, we treated only "significant" bars as those with persistence ≥50 µm for H_0 components and ≥10 µm for H_1 loops, roughly corresponding to the distance needed to bridge gaps of four to five lymphocyte diameters and to the scale of small architectural voids. To avoid redundancy and ensure interpretability, from each diagram we then extracted a compact feature set, including counts of significant H_0 and H_1 bars, the mean and s.d. of their lifespans, and for H_1 also the mean and s.d. of birth distance.

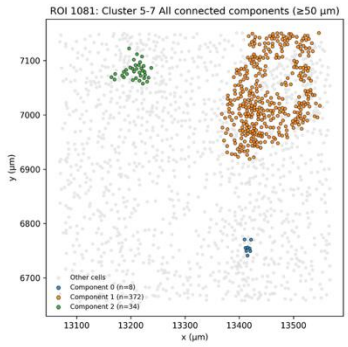
Figure 29 illustrates these ideas for one representative cluster pair (clusters 5↔7) across four peritumour ROIs, revealing distinct geometries. ROI 1081 showed three dense, well-separated islands that do not connect until r approaches around 150 µm. This yields three significant H_0 bars with long lifespans (mean ≈ 149 µm; narrow s.d.) and only two modest H_1 loops (birth mean ≈ 18 µm; lifespan mean ≈ 18 µm), an example of compartmentalisation. ROI 1095 comprised several mid-sized islands spaced more closely than in ROI 1081. Components merge earlier (significant H_0 lifespan mean ≈ 71 µm) and multiple loops arise at intermediate distances (four H_1 bars; birth mean ≈ 53 µm with large s.d.; lifespan mean ≈ 14.9 µm), consistent with a more interwoven mosaic punctuated by channel-like voids. ROI 1097 showed one large, mostly continuous area with a few smaller nearby clusters. Islands are nearer than in ROI 1081 (significant H_0 lifespan mean ≈ 71 µm), and loops are moderate in number and size. ROI 1100 was essentially continuous at 50 µm (no long-lived H_0 components) yet contains many small voids (16 H_1 bars; birth mean ≈ 14.9 µm; lifespan mean ≈ 16.9 µm), the geometry expected for a densely interdigitating sheet with fine perforations.

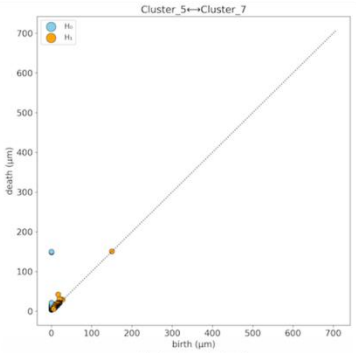
The table in Figure 29 reports these summary features for every ROI and targeted cluster (or cluster pair). In this framework, a high count of significant H_0 features implied the presence of many disconnected micro-regions, while a large mean H_0 lifespan suggested that these regions remained relatively far apart. Numerous H_1 features with short birth distance indicated a continuous structure

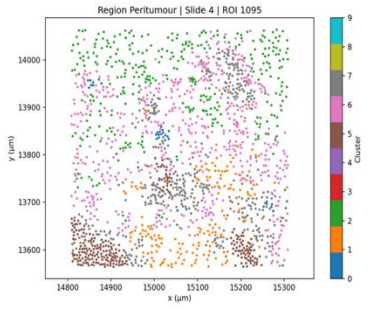
perforated by small loops, whereas higher H_1 birth distance pointed to voids that only emerged when larger-scale structure was engaged. Longer H_1 lifespans marked stable, ring-like gaps. With this proposed link between barcodes and tissue geometry established, we next compare targeted TDA feature panels between regions and identify the topological motifs that best distinguish tumour, peritumour, and healthy liver in CRLM.

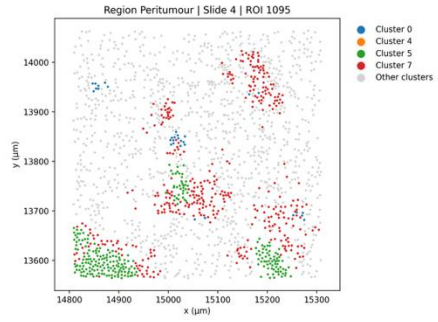


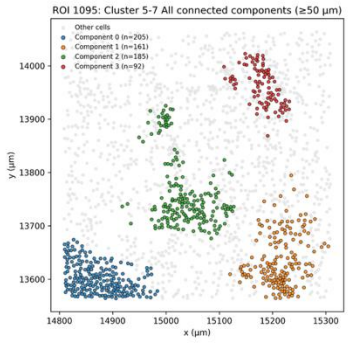


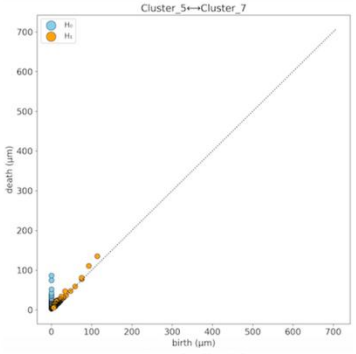


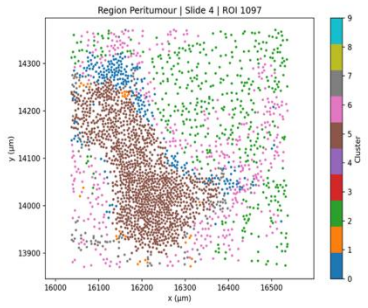


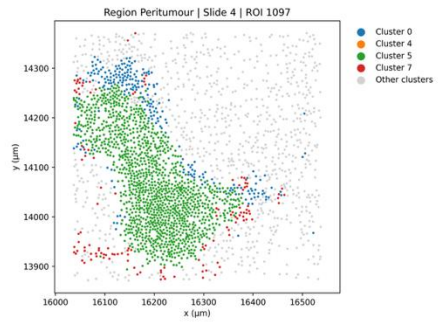


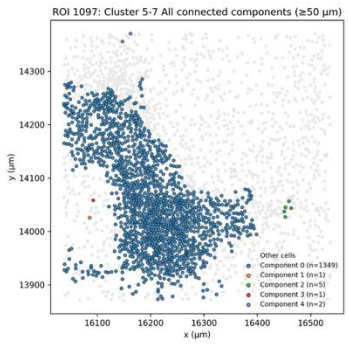


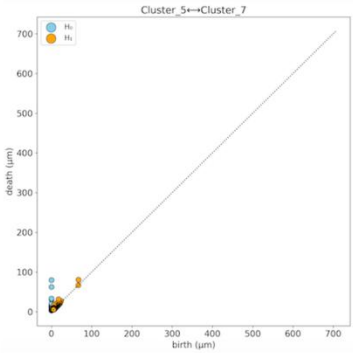


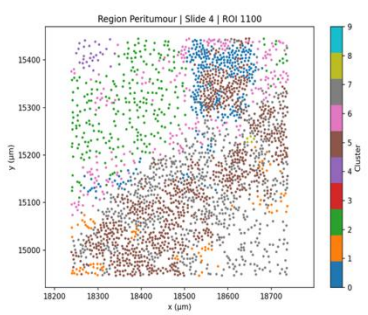


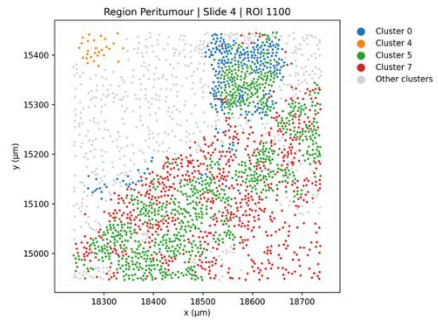


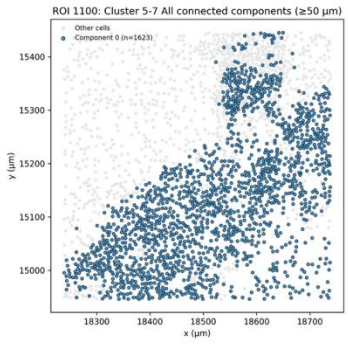


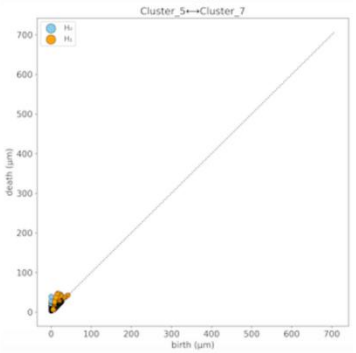












ROI	Cluster_5↔Cluster_7_H0_significant_lifespan_mean	Cluster_5↔Cluster_7_H0_significant_lifespan_std	Cluster_5↔Cluster_7_H0_significant_nBars	Cluster_5↔Cluster_7_H1_significant_birth_mean	Cluster_5↔Cluster_7_H1_significant_birth_std	Cluster_5↔Cluster_7_H1_significant_lifespan_mean	Cluster_5↔Cluster_7_H1_significant_lifespan_std	Cluster_5↔Cluster_7_H1_significant_nBars
ROI_1081	149.2215576171880	0.9327171753914230	3.0	18.344411849975600	2.20437323736127	18.12918281555180	10.189233278112400	2.0
ROI_1095	71.00935745239260	17.57371315770960	4.0	53.17813301086430	47.312919467697900	14.851780319213900	4.638507489876650	5.0
ROI_1097	71.1195068359375	12.297967561182300	3.0	24.75639319419860	28.599910052210700	12.563258886337300	1.769577823943020	4.0
ROI_1100	0.0	0.0	1.0	14.857345938682600	5.373802902332910	16.922575712204000	6.972626412366690	16.0

Figure 29. Persistent homology reveals geometric patterns of neighbourhood organisation in CRLM peritumour tissue. Top: Each row corresponds to one peritumour ROI (ROI_1081, ROI_1095, ROI_1097, ROI_1100), shown as a representative set illustrating the geometric diversity of spatial neighbourhoods formed by cells (points) from cluster 5 and cluster 7. First column: Full ROI spatial scatter plot with all neighbourhood cluster assignments. Second column: Same ROI spatial scatter plot highlighting only cells of clusters 5 and 7, coloured accordingly. Third column: Significant connected components (islands) formed by points from cluster 5 and cluster 7 as the distance threshold increases (here, r = 50 μm). Different connected groups are coloured separately. Fourth column: Persistence diagram (barcode plot) for H_0 (island/connected-component birth and death) and H_1 (loop birth and death) for points from cluster 5 and cluster 7; light blue dots: H_0 components, orange dots: H_1 loops, x-axis: birth distance, y-axis: death distance. Bottom: A summary table reporting quantitative TDA features for each ROI, including Significant H_0 bars: number, mean and s.d. of their lifespans (measuring size and persistence of disconnected "islands" of points from clusters 5 and 7). Significant H_1 bars: number, mean and s.d. of their birth distance (how far apart points must be for loops to appear) and their lifespans (size and persistence of ring-like voids).

5.1 Cell DIVE: TDA-based discrimination of tissue regions

To investigate the geometric organisation of functionally relevant microenvironments in CRLM, we extracted 80 topological descriptors from immune- and stroma-associated neighbourhoods across ROIs in the Cell DIVE cohort. These features were derived from persistent-homology calculations applied to targeted neighbourhood clusters NC0, NC4, NC5 and NC7 and their six pairwise combinations. Random forest classifiers trained on the full set of topological features achieved distinct region separations with high accuracy. Cross-validated accuracy for peritumour versus tumour was 72.7 ± 7.4 percent, while healthy versus peritumour and healthy versus tumour both exceeded 84 percent (84.6 ± 6.5 percent and 84.3 ± 3.3 percent, respectively). Thus, the topology of targeted neighbourhoods and their pairings, without using abundance or marker intensities, was sufficient to distinguish CRLM compartments. UMAP embeddings of the topological data further supported the discriminative abilities of the top RF features. UMAP embeddings were consistent with this. Using all batch-corrected TDA features produced a partial regional ordering with expected overlap (Fig.30). Restricting to the top 30 RF features for each comparison tightened separation: healthy versus tumour formed largely distinct manifolds (Fig.31, A); healthy versus peritumour showed a clear shift with modest mixing (Fig.31, B); and peritumour versus tumour separated into two clouds with a central overlap zone (Fig.31, C). To ground these effects in the underlying slide-level distributions, we summarised the top-30 features for each region-wise comparison with paired Wilcoxon signed-rank tests and used volcano plots to plot all highlighted features.

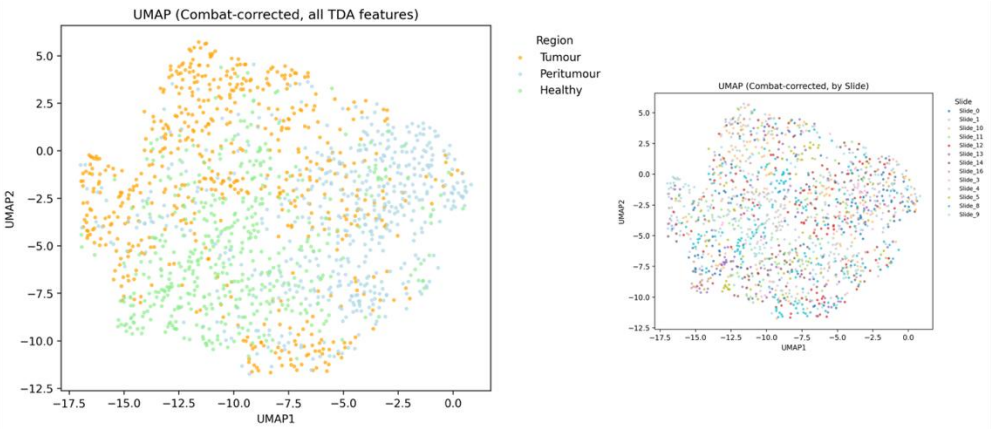


Figure 30. UMAP embedding of region-level topological data analysis (TDA) features across CRLM tissue regions (Cell DIVE). Left: Two-dimensional UMAP embedding of all batch-corrected, standardised TDA features extracted from targeted neighbourhood clusters and their pairwise combinations in Cell DIVE ROIs. Each point represents a single ROI, coloured by tissue region (tumour, peritumour, healthy). Right: The same embedding, with points coloured by slide.

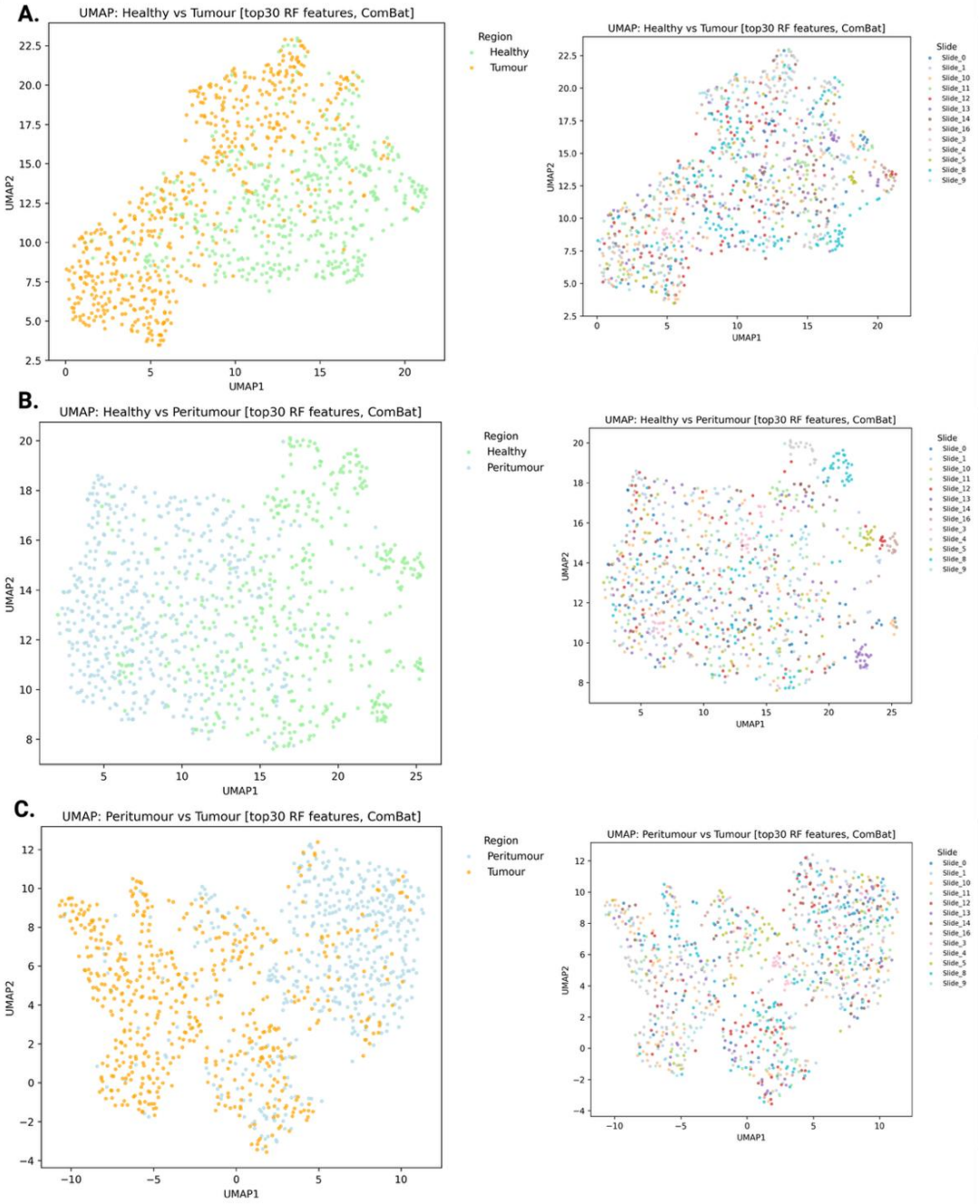


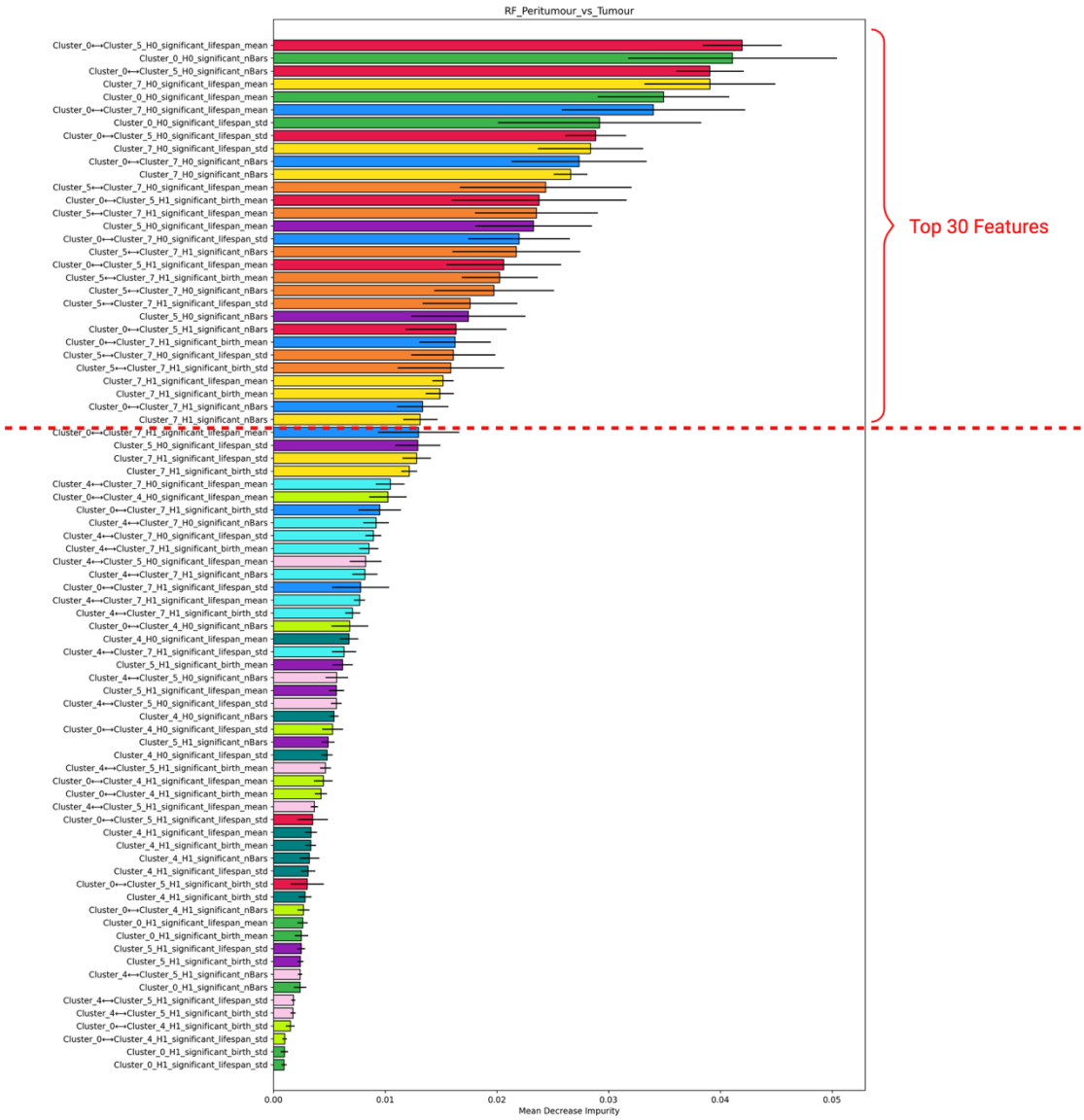
Figure 31. UMAP embedding of the top 30 most discriminatory TDA features for each region pair in CRLM (Cell DIVE). For each pairwise region comparison, UMAP was performed on the top 30 most important TDA features as ranked by random forest classifiers distinguishing those two regions. Left-side UMAP embeddings were coloured by region, right-side UMAP embedding were coloured by slide. A. Healthy (light green) vs Tumour (orange). B. Healthy vs Peritumour (sky blue). C. Peritumour vs Tumour.

Peritumour versus Tumour

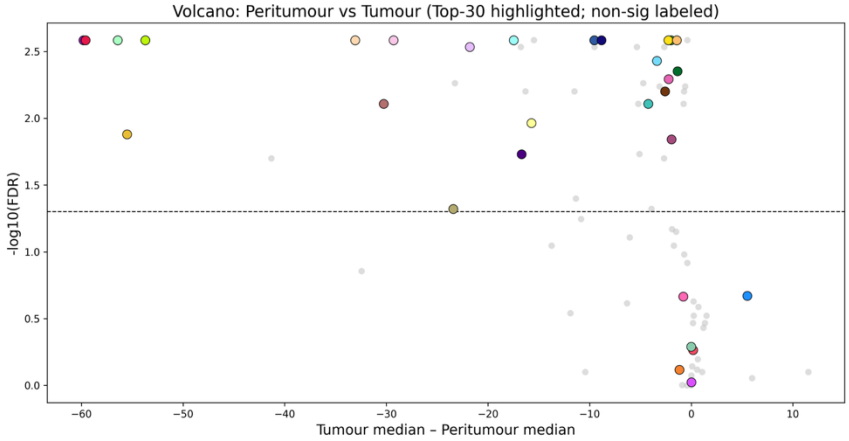
The top-30 features were dominated by H0 descriptors from T-cell neighbourhoods NC0 and NC5 and their pairings with each other and with the stromal–immune cell interface NC7 (Fig.32, A). High-ranking entries included the mean and variability of significant H_0 lifespans for NC0↔NC5 and NC0 alone, together with the corresponding counts of significant H_0 components. H_0 features from NC0↔NC7 and NC5↔NC7 appeared next, indicating that differences in how T-cell islands sat alongside the stromal-immune cell interface also differed between regions. H_1 loop metrics were present but secondary, mainly for NC0↔NC5 and NC5↔NC7. In the volcano plot, most of these top features lay to the left of zero on the x-axis, defined as the slide-level median difference tumour minus peritumour, and exceeded the FDR threshold (Fig.32, B), showing higher peritumour medians.

Geometrically, the significantly higher NC0, NC5, and NC0↔NC5 H_0/H_1 statistics observed in the peritumour region might indicate that the border zone contains more spatially disconnected T-cell micro-regions, which persisted as separate entities over greater distances. In contrast, the tumour core was characterised by fewer components with shorter lifespans, suggesting a collapse of T-cell spatial fragmentation within the tumour mass. Since NC0↔NC7/NC5↔NC7 H_0/H_1 statistics were also significantly higher in peritumour, they reinforced the idea that peritumour tissue preserved more long-lasting, discrete stromal–immune substructures, whereas these elements become homogenised in the tumour interior. A small number of tumour-biased points on the right involved NC7 H_1 birth distance or H_0 counts but did not pass FDR, reinforcing that the dominant, significant signal is the peritumour excess in T-cell fragmentation. These findings geometrically depicted an active tumour margin where multiple T-cell neighbourhoods coexist and interleave with a structured stromal-immune interface, consistent with the formation and maintenance of TLS (Sautès-Fridman *et al.,* 2019; Rodriguez *et al.,* 2021). Conversely, the homogenised intratumoural organisation suggested a breakdown of this complex architecture, pointing to the disruption or failed assembly of TLS within the tumour core.

A.



B.



Peritumour

Most Neg

- Cluster_0_H0_significant_lifespan_mean
- Cluster_0⟷Cluster_5_H0_significant_lifespan_mean
- Cluster_5⟷Cluster_7_H0_significant_lifespan_mean
- Cluster_5_H0_significant_lifespan_mean
- Cluster_0⟷Cluster_7_H0_significant_lifespan_mean
- Cluster_0⟷Cluster_5_H1_significant_birth_mean
- Cluster_0⟷Cluster_7_H1_significant_birth_mean
- Cluster_0_H0_significant_lifespan_std
- Cluster_5⟷Cluster_7_H1_significant_birth_mean
- Cluster_0⟷Cluster_5_H0_significant_lifespan_std
- Cluster_0⟷Cluster_7_H0_significant_lifespan_std
- Cluster_5⟷Cluster_7_H0_significant_lifespan_std
- Cluster_5⟷Cluster_7_H1_significant_birth_std
- Cluster_0⟷Cluster_5_H1_significant_lifespan_mean
- Cluster_5⟷Cluster_7_H1_significant_lifespan_mean
- Cluster_5⟷Cluster_7_H1_significant_lifespan_std
- Cluster_5⟷Cluster_7_H1_significant_nBars
- Cluster_0⟷Cluster_7_H0_significant_nBars
- Cluster_0⟷Cluster_5_H0_significant_nBars
- Cluster_0⟷Cluster_7_H1_significant_nBars
- Cluster_0_H0_significant_nBars
- Cluster_5⟷Cluster_7_H0_significant_nBars
- Cluster_0⟷Cluster_5_H1_significant_nBars
- Cluster_5_H0_significant_nBars
- Cluster_7_H0_significant_lifespan_std (non-significant)
- Cluster_7_H1_significant_lifespan_mean (non-significant)
- Cluster_7_H1_significant_nBars (non-significant)

Least Pos

Most Pos

Tumour

- Cluster_7_H1_significant_birth_mean (non-significant)
- Cluster_7_H0_significant_nBars (non-significant)
- Cluster_7_H0_significant_lifespan_mean (non-significant)

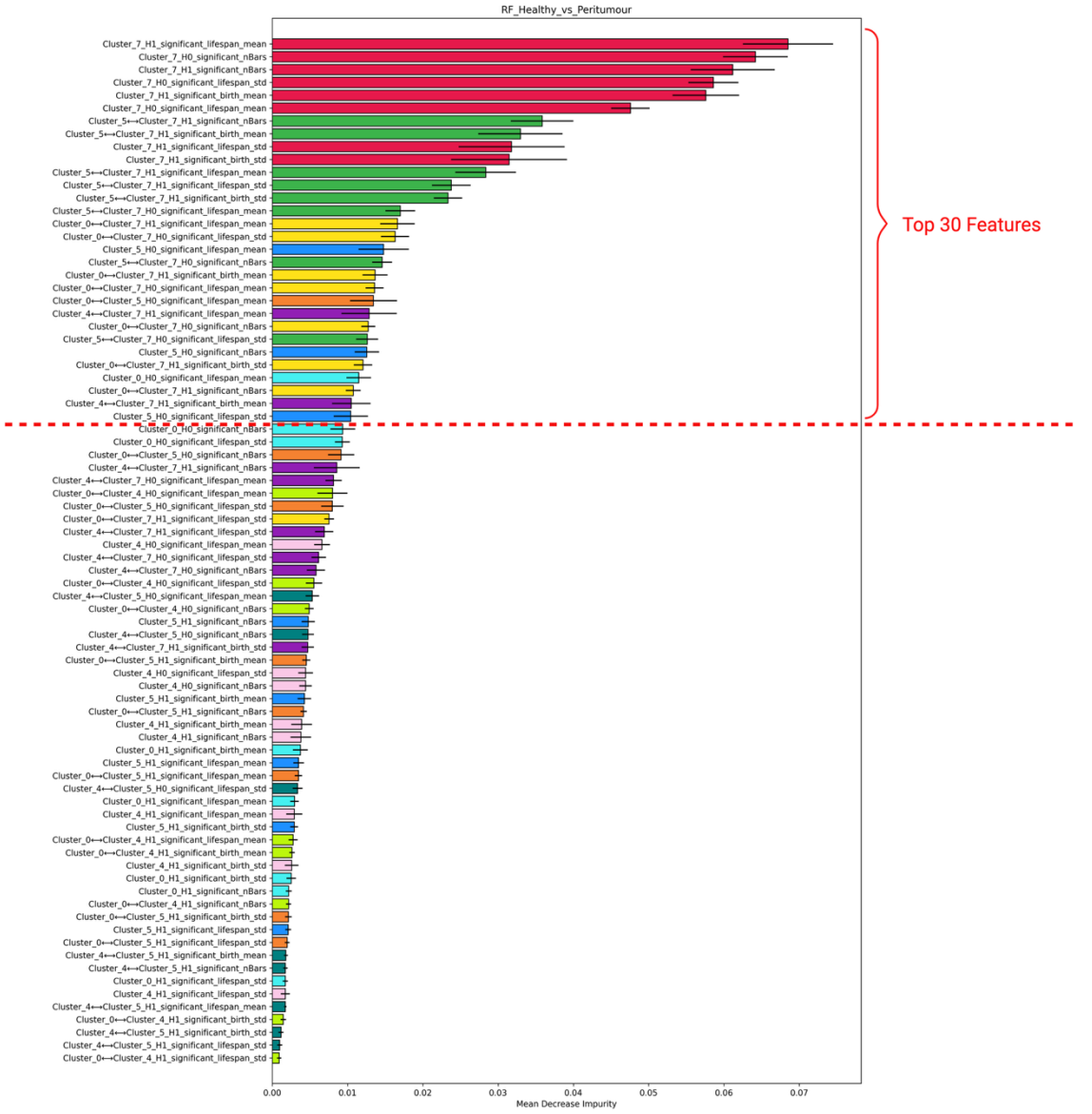
Least Neg

Figure 32. Topological features distinguishing peritumour and tumour regions in CRLM: feature importance and statistical significance (Cell DIVE). A. A ranked bar plot of the TDA features used to discriminate between peritumour and tumour regions, as identified by random forest classifier feature importance (mean decrease in impurity, MDI). Each bar corresponds to a single TDA feature, derived from persistent homology calculations applied to single and pairwise neighbourhood clusters (e.g., NC0, NC4, NC5, NC7, and their pairings), and represents its mean importance across cross-validation folds. The top 30 features, as ranked by average MDI, are indicated above a dashed red line. B. A volcano plot comparing the distribution of the same TDA features between regions. For each feature, the x-axis gives the difference in slide-level medians (tumour minus peritumour), and the y-axis shows the negative log10 of the FDR-adjusted p-value from a paired Wilcoxon signed-rank test. Each point represents one TDA feature, with the top 30 features highlighted and labelled, and those not passing FDR significance are noted as non-significant in the legend. Circle colours correspond to feature identities. The volcano plot enables assessment of both the effect size and statistical significance of region differences for each topological feature.

Healthy versus Peritumour

Here, the top-30 ranking shifted decisively toward geometric features within NC7 (Fig.33, A). The leading features were the number of significant H_0/H_1 features, their mean lifespans, and H_1 mean birth distance in NC7. Geometric descriptors from the NC5↔NC7 pair and NC0↔NC7 pair also appeared among the top contributors, but with smaller importances. The volcano plot again showed a clear directional pattern (Fig.33, B), where most top features fell to the right of zero effect size (peritumour minus healthy) and exceeded the FDR line, indicating that the peritumour region harboured more numerous and longer-lived stromal–immune interface micro-regions. Healthy liver, by contrast, sat lower for these metrics, indicating fewer discrete islands with less perforated architectures. Together, the RF and paired Wilcoxon signed-rank test results suggested that, compared to the healthy liver, the peritumour region exhibited a pronounced increase in spatial fragmentation and topological complexity within the stromal–immune interface (NC7), as well as in its interactions with T-cell neighbourhoods (NC5↔NC7, NC0↔NC7). The emergence of more numerous and persistent topological loops and islands highlights potential architectural reorganisation of the tumour border, consistent with the assembly of complex, immune-active, and stroma-rich microenvironments that are largely absent from healthy tissue.

A.


RF_Healthy_vs_Peritumour
Cluster_7_H1_significant_lifespan_mean
Cluster_7_H0_significant_nBars
Cluster_7_H1_significant_nBars
Cluster_7_H0_significant_lifespan_std
Cluster_7_H1_significant_birth_mean
Cluster_7_H0_significant_lifespan_mean
Cluster_5↔Cluster_7_H1_significant_nBars
Cluster_5↔Cluster_7_H1_significant_birth_mean
Cluster_7_H1_significant_lifespan_std
Cluster_7_H1_significant_birth_std
Cluster_5↔Cluster_7_H1_significant_lifespan_mean
Cluster_5↔Cluster_7_H1_significant_lifespan_std
Cluster_5↔Cluster_7_H1_significant_birth_std
Cluster_5↔Cluster_7_H0_significant_lifespan_mean
Cluster_0↔Cluster_7_H1_significant_lifespan_mean
Cluster_0↔Cluster_7_H0_significant_lifespan_std
Cluster_5_H0_significant_lifespan_mean
Cluster_5↔Cluster_7_H0_significant_nBars
Cluster_0↔Cluster_7_H1_significant_birth_mean
Cluster_0↔Cluster_7_H0_significant_lifespan_mean
Cluster_0↔Cluster_5_H0_significant_lifespan_mean
Cluster_4↔Cluster_7_H1_significant_lifespan_mean
Cluster_0↔Cluster_7_H0_significant_nBars
Cluster_5↔Cluster_7_H0_significant_lifespan_std
Cluster_5_H0_significant_nBars
Cluster_0↔Cluster_7_H1_significant_birth_std
Cluster_0_H0_significant_lifespan_mean
Cluster_0↔Cluster_7_H1_significant_nBars
Cluster_4↔Cluster_7_H1_significant_birth_mean
Cluster_5_H0_significant_lifespan_std
Top 30 Features
Cluster_0_H0_significant_nBars
Cluster_0_H0_significant_lifespan_std
Cluster_0↔Cluster_5_H0_significant_nBars
Cluster_4↔Cluster_7_H1_significant_nBars
Cluster_4↔Cluster_7_H0_significant_lifespan_mean
Cluster_0↔Cluster_4_H0_significant_lifespan_mean
Cluster_0↔Cluster_5_H0_significant_lifespan_std
Cluster_0↔Cluster_7_H1_significant_lifespan_std
Cluster_4↔Cluster_7_H1_significant_lifespan_std
Cluster_4_H0_significant_lifespan_mean
Cluster_4↔Cluster_7_H0_significant_lifespan_std
Cluster_4↔Cluster_7_H0_significant_nBars
Cluster_0↔Cluster_4_H0_significant_lifespan_std
Cluster_4↔Cluster_5_H0_significant_lifespan_mean
Cluster_0↔Cluster_4_H0_significant_nBars
Cluster_5_H1_significant_nBars
Cluster_4↔Cluster_5_H0_significant_nBars
Cluster_4↔Cluster_7_H1_significant_birth_std
Cluster_0↔Cluster_5_H1_significant_birth_mean
Cluster_4_H0_significant_lifespan_std
Cluster_4_H0_significant_nBars
Cluster_5_H1_significant_birth_mean
Cluster_0↔Cluster_5_H1_significant_nBars
Cluster_4_H1_significant_birth_mean
Cluster_4_H1_significant_nBars
Cluster_0_H1_significant_birth_mean
Cluster_5_H1_significant_lifespan_mean
Cluster_0↔Cluster_5_H1_significant_lifespan_mean
Cluster_4↔Cluster_5_H0_significant_lifespan_std
Cluster_0_H1_significant_lifespan_mean
Cluster_4_H1_significant_lifespan_mean
Cluster_5_H1_significant_birth_std
Cluster_0↔Cluster_4_H1_significant_lifespan_mean
Cluster_0↔Cluster_4_H1_significant_birth_mean
Cluster_4_H1_significant_birth_std
Cluster_0_H1_significant_birth_std
Cluster_0_H1_significant_nBars
Cluster_0↔Cluster_4_H1_significant_nBars
Cluster_0↔Cluster_5_H1_significant_birth_std
Cluster_5_H1_significant_lifespan_std
Cluster_0↔Cluster_5_H1_significant_lifespan_std
Cluster_4↔Cluster_5_H1_significant_birth_mean
Cluster_4↔Cluster_5_H1_significant_nBars
Cluster_0_H1_significant_lifespan_std
Cluster_4_H1_significant_lifespan_std
Cluster_4↔Cluster_5_H1_significant_lifespan_mean
Cluster_0↔Cluster_4_H1_significant_birth_std
Cluster_4↔Cluster_5_H1_significant_birth_std
Cluster_4↔Cluster_5_H1_significant_lifespan_std
Cluster_0↔Cluster_4_H1_significant_lifespan_std
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
Mean Decrease Impurity

B.

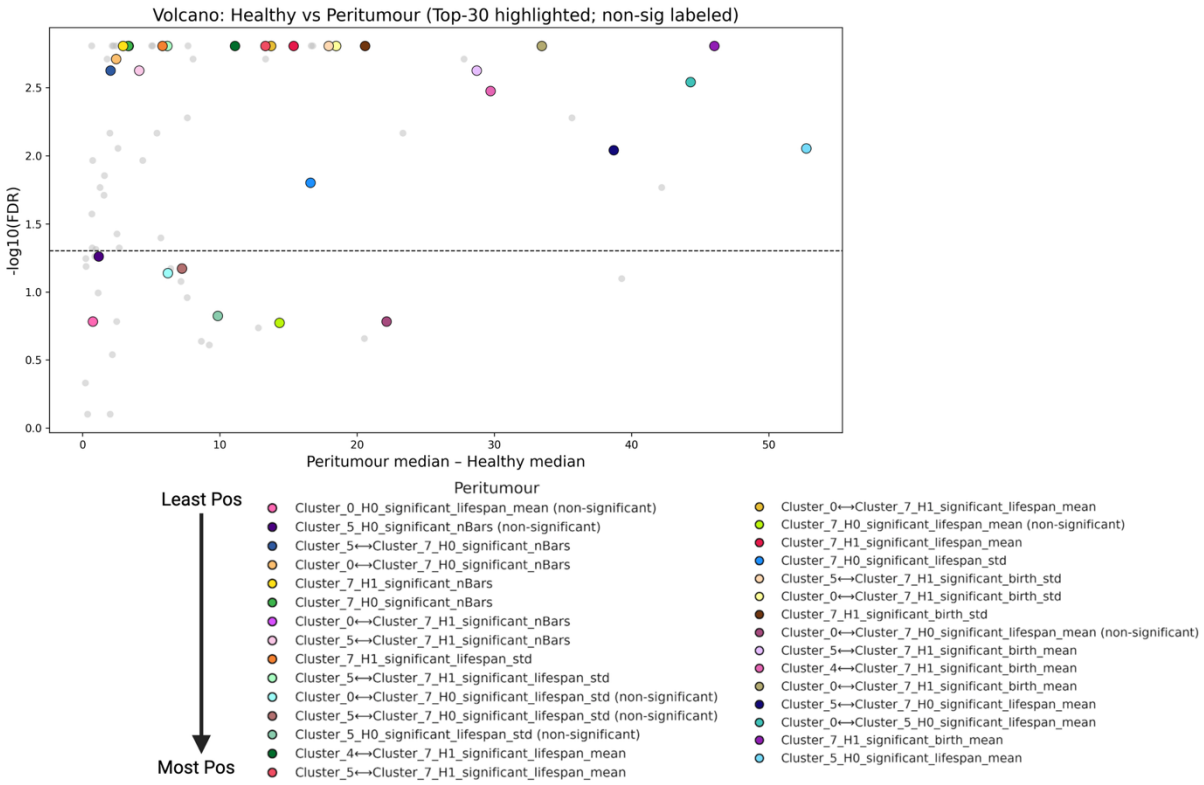

Volcano: Healthy vs Peritumour (Top-30 highlighted; non-sig labeled)
-log10(FDR)
0.0
0.5
1.0
1.5
2.0
2.5
0
10
20
30
40
50
Peritumour median – Healthy median
Least Pos
Most Pos
Peritumour
Cluster_0_H0_significant_lifespan_mean (non-significant)
Cluster_5_H0_significant_nBars (non-significant)
Cluster_5↔Cluster_7_H0_significant_nBars
Cluster_0↔Cluster_7_H0_significant_nBars
Cluster_7_H1_significant_nBars
Cluster_7_H0_significant_nBars
Cluster_0↔Cluster_7_H1_significant_nBars
Cluster_5↔Cluster_7_H1_significant_nBars
Cluster_7_H1_significant_lifespan_std
Cluster_5↔Cluster_7_H1_significant_lifespan_std
Cluster_0↔Cluster_7_H0_significant_lifespan_std (non-significant)
Cluster_5↔Cluster_7_H0_significant_lifespan_std (non-significant)
Cluster_5_H0_significant_lifespan_std (non-significant)
Cluster_4↔Cluster_7_H1_significant_lifespan_mean
Cluster_5↔Cluster_7_H1_significant_lifespan_mean
Cluster_0↔Cluster_7_H1_significant_lifespan_mean
Cluster_7_H0_significant_lifespan_mean (non-significant)
Cluster_7_H1_significant_lifespan_mean
Cluster_7_H0_significant_lifespan_std
Cluster_5↔Cluster_7_H1_significant_birth_std
Cluster_0↔Cluster_7_H1_significant_birth_std
Cluster_7_H1_significant_birth_std
Cluster_0↔Cluster_7_H0_significant_lifespan_mean (non-significant)
Cluster_5↔Cluster_7_H1_significant_birth_mean
Cluster_4↔Cluster_7_H1_significant_birth_mean
Cluster_0↔Cluster_7_H1_significant_birth_mean
Cluster_5↔Cluster_7_H0_significant_lifespan_mean
Cluster_0↔Cluster_5_H0_significant_lifespan_mean
Cluster_7_H1_significant_birth_mean
Cluster_5_H0_significant_lifespan_mean

Figure 33. Topological features distinguishing peritumour and healthy regions in CRLM: feature importance and statistical significance (Cell DIVE). A. A ranked bar plot of the TDA features used to discriminate between peritumour and healthy regions, as identified by random forest classifier feature importance (mean decrease in impurity, MDI). The top 30 features, as ranked by average MDI, are indicated above a dashed red line. B. A volcano plot comparing the distribution of the same TDA features between regions. Each point represents one TDA feature, with the top 30 features highlighted and labelled, and those not passing FDR significance are noted as non-significant in the legend.

Healthy versus Tumour

The dominant discriminants were H_0 descriptors from NC7, with both the number of significant NC7 components and their mean lifespans ranking at the very top of the feature list (Fig.34, A). NC7 H_1 metrics and NC0 H_0 measures followed. Correspondingly, the volcano plot placed key NC7 features to the right of zero for tumour minus healthy and above the FDR line (Fig.34, B), indicating that tumour cores contain more stromal–immune interface components that persist as separate entities over long distances than in healthy liver. Beyond single-cluster signals, pairwise features involving NC4 were prominent and showed the same tumour-shift. NC4↔NC7 H_1 birth and lifespan statistics and NC4↔NC7 H_0 lifespan were all higher in tumour than healthy, indicating neutrophil-rich neighbourhoods interleaving with the stromal–immune interface in the tumour core. H_0 features of NC0 showed significantly higher values in the healthy liver region, indicating that compared to the tumour core, the healthy liver baseline preserved more long-lasting, discrete T-cell-rich neighbourhoods. Nevertheless, the healthy liver lacked both the complex interface and the granulocytic admixture seen in the CRLM tumour core.

Taken together, the three comparisons outlined a coherent spatial progression that aligned with earlier neighbourhood analyses. The peritumour border was characterised by fragmented, T-cell-dominated geometry interleaved with loop-rich stromal-immune interfaces, reflecting active but compartmentalised immunity. Moving inward, the tumour core exhibited fewer and less fragmented T-cell regions, but its fragmented fibroblastic–immune components were often interwoven with neutrophil neighbourhoods. The healthy liver showed the simplest topology with intermediate (baseline) T-cell infiltration geometry, minimal interface modularity, and little granulocytic admixture.

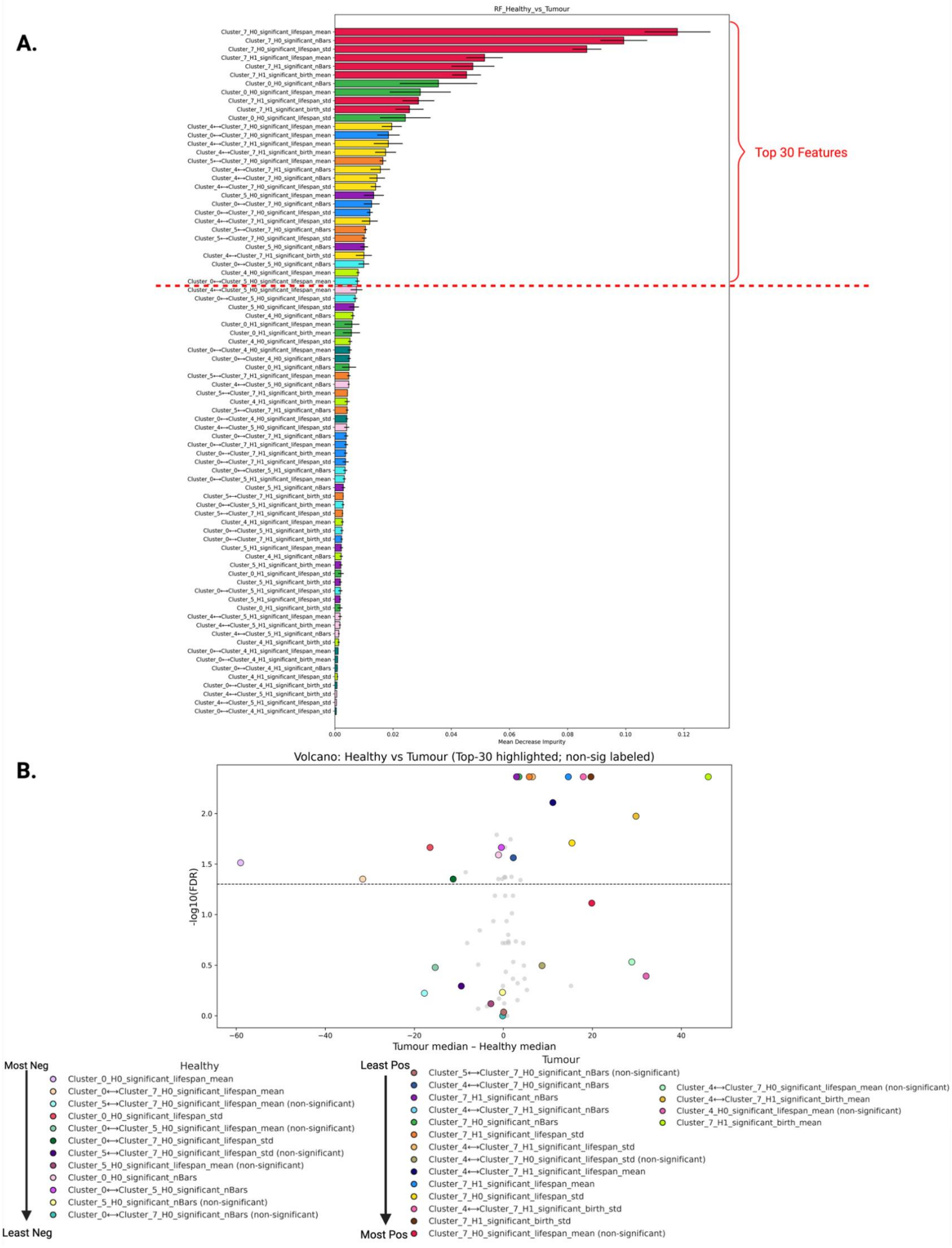

A.
RF_Healthy_vs_Tumour
Top 30 Features
Mean Decrease Impurity
B.
Volcano: Healthy vs Tumour (Top-30 highlighted; non-sig labeled)
-log10(FDR)
Tumour median – Healthy median
Most Neg
Least Neg
Healthy
Cluster_0_H0_significant_lifespan_mean
Cluster_0↔Cluster_7_H0_significant_lifespan_mean
Cluster_5↔Cluster_7_H0_significant_lifespan_mean (non-significant)
Cluster_0_H0_significant_lifespan_std
Cluster_0↔Cluster_5_H0_significant_lifespan_mean (non-significant)
Cluster_0↔Cluster_7_H0_significant_lifespan_std
Cluster_5↔Cluster_7_H0_significant_lifespan_std (non-significant)
Cluster_5_H0_significant_lifespan_mean (non-significant)
Cluster_0_H0_significant_nBars
Cluster_0↔Cluster_5_H0_significant_nBars
Cluster_5_H0_significant_nBars (non-significant)
Cluster_0↔Cluster_7_H0_significant_nBars (non-significant)
Least Pos
Most Pos
Tumour
Cluster_5↔Cluster_7_H0_significant_nBars (non-significant)
Cluster_4↔Cluster_7_H0_significant_nBars
Cluster_7_H1_significant_nBars
Cluster_4↔Cluster_7_H1_significant_nBars
Cluster_7_H0_significant_nBars
Cluster_7_H1_significant_lifespan_std
Cluster_4↔Cluster_7_H1_significant_lifespan_std
Cluster_4↔Cluster_7_H0_significant_lifespan_std (non-significant)
Cluster_4↔Cluster_7_H1_significant_lifespan_mean
Cluster_7_H1_significant_lifespan_mean
Cluster_7_H0_significant_lifespan_std
Cluster_4↔Cluster_7_H1_significant_birth_std
Cluster_7_H1_significant_birth_std
Cluster_7_H0_significant_lifespan_mean (non-significant)
Cluster_4↔Cluster_7_H0_significant_lifespan_mean (non-significant)
Cluster_4↔Cluster_7_H1_significant_birth_mean
Cluster_4_H0_significant_lifespan_mean (non-significant)
Cluster_7_H1_significant_birth_mean

Figure 34. Topological features distinguishing healthy and tumour regions in CRLM: feature importance and statistical significance (Cell DIVE). A. A ranked bar plot of the TDA features used to discriminate between healthy and tumour regions, as identified by random forest classifier feature importance (mean decrease in impurity, MDI). The top 30 features, as ranked by average MDI, are indicated above a dashed red line. B. A volcano plot comparing the distribution of the same TDA features between regions. Each point represents one TDA feature, with the top 30 features highlighted and labelled, and those not passing FDR significance are noted as non-significant in the legend.

5.2 CosMx: TDA-based discrimination of tissue regions

We applied the same topological workflow to CosMx ROIs, extracting 48 persistent-homology descriptors from immune- and stroma-associated neighbourhoods (NC0, NC2, NC7) and their pairwise combinations. Random-forest classifiers trained on these features achieved above-chance region separation despite the smaller cohort: peritumour versus tumour reached 71.2 ± 9.2 percent accuracy, healthy versus tumour 69.8 ± 8.2 percent, and healthy versus peritumour 68.2 ± 9.4 percent. As in Cell DIVE, these results indicated that topology features captured region-specific organisation, but to a lesser extent than in the Cell DIVE data. UMAP embeddings were consistent with this picture. Using all TDA features, ROIs from the three regions intermingled with only mild ordering by compartment (Fig.35). Restricting to the top 20 RF features per comparison tightened the structure. Healthy versus tumour showed the clearest separation (Fig.36, A), healthy versus peritumour resolved into partially distinct clusters with broad overlap (Fig.36, B), whereas peritumour versus tumour largely overlapped with a subtle shift (Fig.36, C). As in Cell DIVE, we summarised the top-20 features for each region-wise slide-level comparison with paired Wilcoxon signed-rank tests and created volcano plots for visualisation.

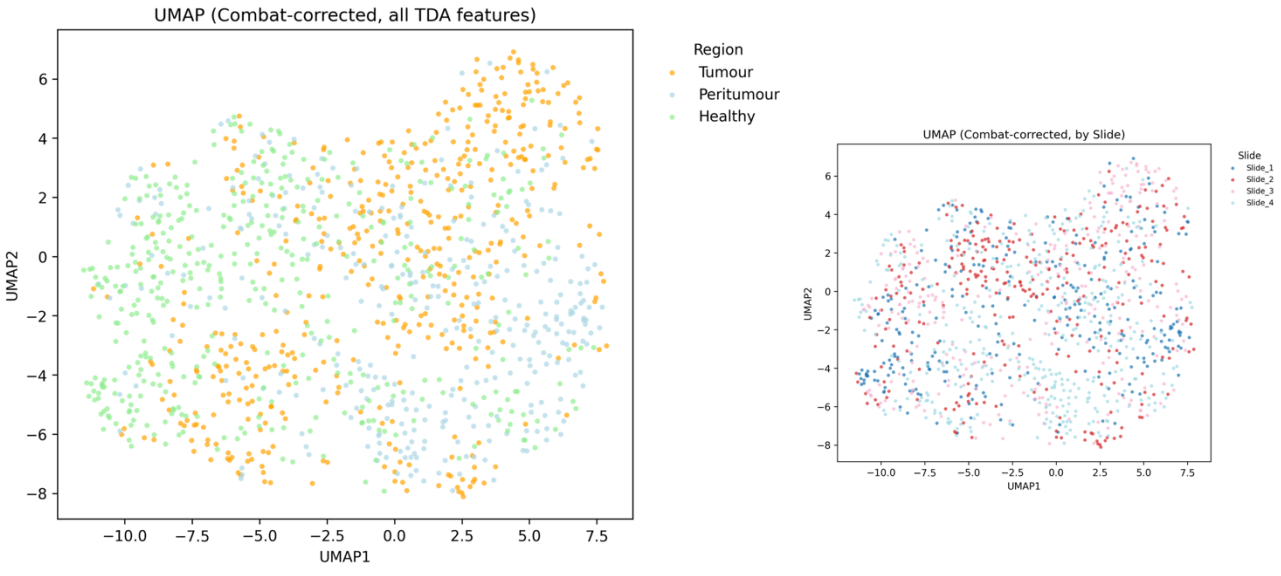


Figure 35. UMAP embedding of region-level topological data analysis (TDA) features across CRLM tissue regions (CosMx). Left: Two-dimensional UMAP embedding of all batch-corrected, standardised TDA features extracted from targeted neighbourhood clusters and their pairwise combinations in CosMx ROIs. Each point represents a single ROI, coloured by tissue region. Right: The same embedding, with points coloured by slide.

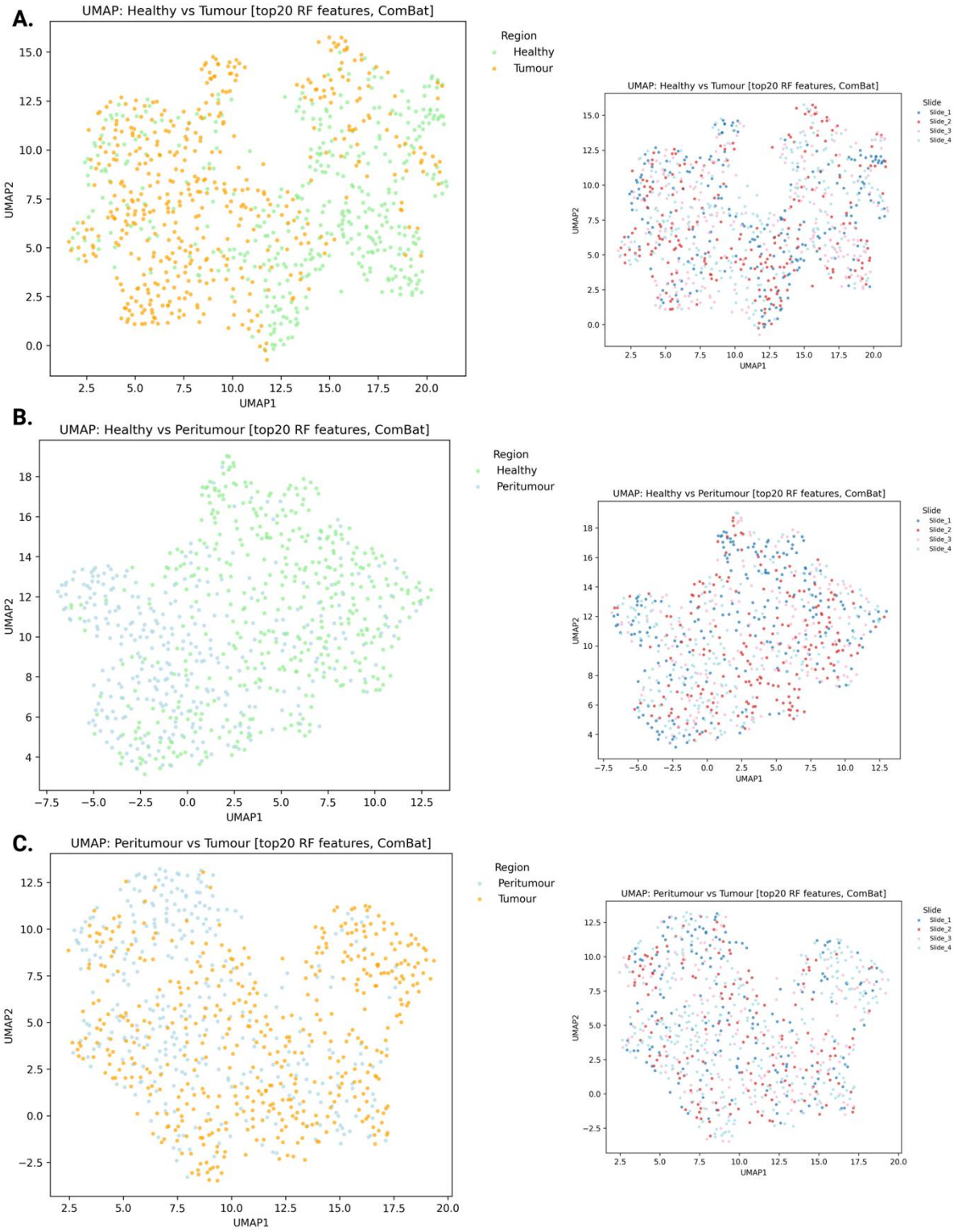
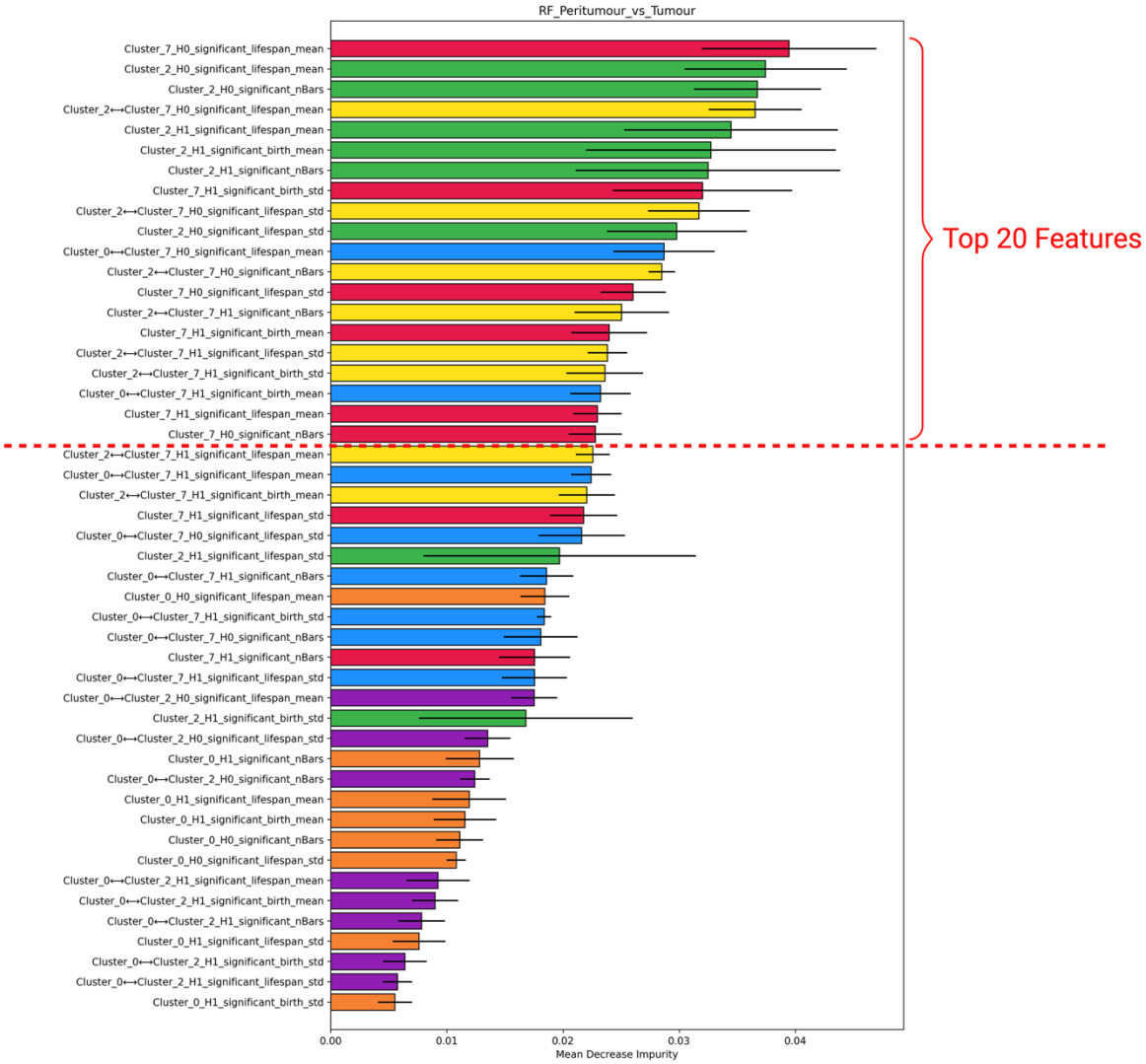


Figure 36. UMAP embedding of the top 20 most discriminatory TDA features for each region pair in CRLM (CosMx). For each pairwise region comparison, UMAP was performed on the top 20 most important TDA features as ranked by random forest classifiers distinguishing those two regions. Left-side UMAP embeddings were coloured by region, right-side UMAP embedding were coloured by slide. A. Healthy (light green) vs Tumour (orange). B. Healthy vs Peritumour (sky blue). C. Peritumour vs Tumour.

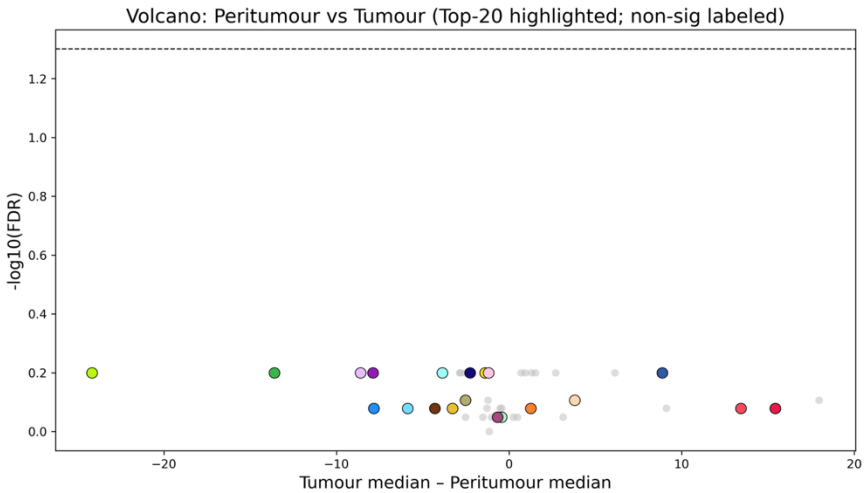
Peritumour versus Tumour

The highest-ranked features centred on NC7 and NC2. NC7 H_0 lifespan mean sat at the top of the RF list, followed closely by NC2 H_0 lifespan mean and NC2 H_0 counts, with NC2↔NC7 H_0 features (lifespan mean and s.d.) and several NC2 H_1 descriptors (lifespan mean and birth mean) also prominent (Fig.37, A). The volcano plot used tumour minus peritumour on the x-axis and showed a split of the top 20 features in different directions (Fig.37, B). Most NC2-driven points, including NC2 H_0 lifespan mean, NC2 H_0 counts, and NC2 H_1 birth and lifespan means, fell to the left of zero, indicating more numerous and longer-lived adaptive and antigen-presenting islands formed at the invasive margin. By contrast, NC7-centric H_0 quantities trended to the right or sat near the origin. For example, the NC7 H_0 lifespan mean and its s.d., and the pairwise NC0↔NC7 H_0 lifespan mean shifted toward positive values on the x-axis, consistent with more widely separated fibroblastic–immune interface components appearing inside the tumour core. Geometrically, the peritumour region favoured many NC2 islands and moderate NC2↔NC7 interleaving, whereas the tumour core maintained more persistent and collapsed NC7 compartments, same for NC0↔NC7 structures. Biologically, this pointed to an enrichment of adaptive/APC neighbourhoods at the border and a relative attenuation of those motifs within the tumour. The tumour core contained fewer, more widely separated micro-regions, characterised either by fibroblast–immune cell interfaces or by admixture with myeloid and T-regulatory programmes. Notably, since our 2.2 result section also revealed that the tumour core exhibited a higher proportion of NC0 and a comparably high proportion of NC7, similar to the peritumour region, this may suggest that, rather than being uniformly sparse, the tumour core may harbour a small number of large micro-regions densely populated with cells (all cells/points are connected in 50 μm), with each major cluster separated from the others. With only four slides, these were qualitative trends, but overall, they matched the region-level summaries of NC2 peaking at the peritumour in Section 2.2.

A.



B.



Most Neg

Least Neg

Peritumour

- Cluster_2_H1_significant_birth_mean (non-significant)
- Cluster_2_H0_significant_lifespan_mean (non-significant)
- Cluster_7_H1_significant_birth_std (non-significant)
- Cluster_2_H1_significant_lifespan_mean (non-significant)
- Cluster_2↔Cluster_7_H0_significant_lifespan_mean (non-significant)
- Cluster_2↔Cluster_7_H1_significant_birth_std (non-significant)
- Cluster_2_H0_significant_lifespan_std (non-significant)
- Cluster_2↔Cluster_7_H1_significant_lifespan_std (non-significant)
- Cluster_7_H1_significant_birth_mean (non-significant)
- Cluster_7_H1_significant_lifespan_mean (non-significant)
- Cluster_2↔Cluster_7_H1_significant_nBars (non-significant)
- Cluster_2_H0_significant_nBars (non-significant)
- Cluster_2_H1_significant_nBars (non-significant)
- Cluster_7_H0_significant_nBars (non-significant)
- Cluster_2↔Cluster_7_H0_significant_nBars (non-significant)

Least Pos

Most Pos

Tumour

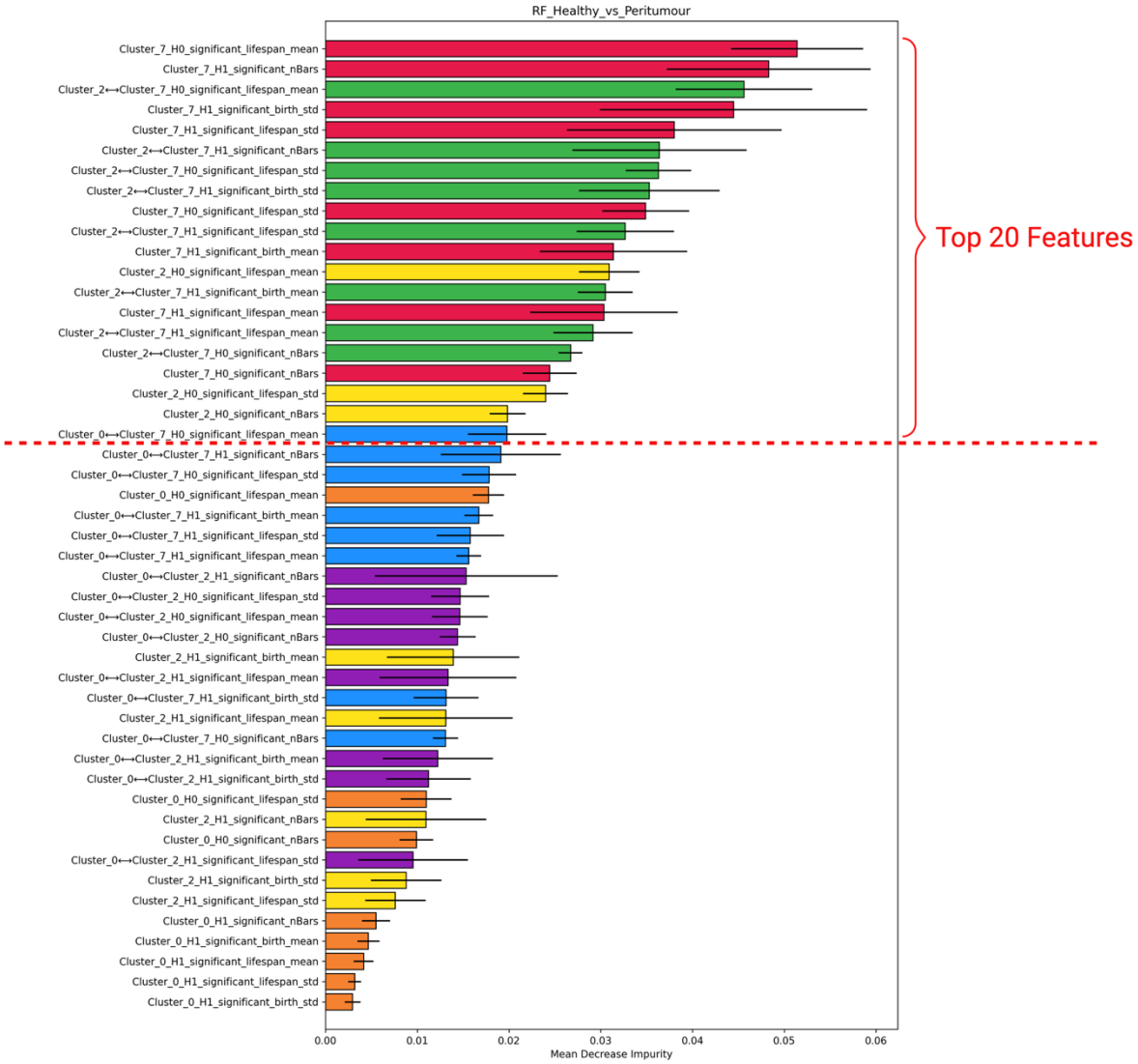
- Cluster_2↔Cluster_7_H0_significant_lifespan_std (non-significant)
- Cluster_7_H0_significant_lifespan_std (non-significant)
- Cluster_0↔Cluster_7_H1_significant_birth_mean (non-significant)
- Cluster_0↔Cluster_7_H0_significant_lifespan_mean (non-significant)
- Cluster_7_H0_significant_lifespan_mean (non-significant)

Figure 37. Topological features distinguishing peritumour and tumour regions in CRLM: feature importance and statistical significance (CosMx). A. A ranked bar plot of the TDA features used to discriminate between peritumour and tumour regions, as identified by random forest classifier feature importance (mean decrease in impurity, MDI). Each bar corresponds to a single TDA feature, derived from persistent homology calculations applied to single and pairwise neighbourhood clusters (e.g., NC0, NC2, NC7, and their pairings), and represents its mean importance across cross-validation folds. The top 20 features, as ranked by average MDI, are indicated above a dashed red line. B. A volcano plot comparing the distribution of the same TDA features between regions. For each feature, the x-axis gives the difference in slide-level medians (tumour minus peritumour), and the y-axis shows the negative log10 of the FDR-adjusted p-value from a paired Wilcoxon signed-rank test. Each point represents one TDA feature, with the top 20 features highlighted and labelled, and those not passing FDR significance are noted as non-significant in the legend. Circle colours correspond to feature identities. The volcano plot enables assessment of both the effect size and statistical significance of region differences for each topological feature.

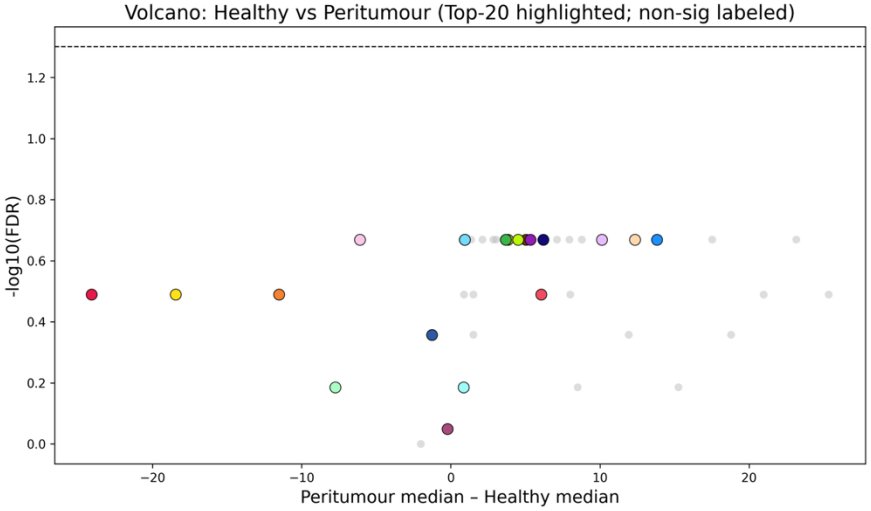
Healthy versus Peritumour

For this comparison, feature importance shifted decisively toward NC7 geometry (Fig.38, A). Top entries included NC7 H_0 lifespan mean and s.d., the number of NC7 H1 loops, and multiple NC2↔NC7 metrics spanning both H_0 and H_1. The volcano plot was defined as peritumour minus healthy (Fig.38, B). Here, the counts clearly favoured the peritumour margin: NC7 H_0 nBars, NC2↔NC7 H_0 nBars and NC2 H_0 nBars lie to the right of zero, indicating more discrete components in peritumour. In contrast, the separation scale was larger in healthy tissue: NC7 H_0 lifespan mean, NC7 H_0 lifespan s.d. and their NC2↔NC7 counterparts lie to the left, showing longer-lived components and broader spacing in healthy liver. The geometric picture was that peritumour accumulated a greater number of NC7 and NC2↔NC7 compartments and loop features, but these features merged/disappeared at smaller distances, whereas the healthy parenchyma contained fewer interfaces that were separated by greater distances. This again fitted a border that is busy and compartmentalised versus a simpler healthy-region topology.

A.



B.



Most Neg

Healthy

- Cluster_7_H0_significant_lifespan_mean (non-significant)
- Cluster_2↔Cluster_7_H0_significant_lifespan_mean (non-significant)
- Cluster_7_H0_significant_lifespan_std (non-significant)
- Cluster_2_H0_significant_lifespan_mean (non-significant)
- Cluster_2↔Cluster_7_H0_significant_lifespan_std (non-significant)
- Cluster_2_H0_significant_lifespan_std (non-significant)
- Cluster_0↔Cluster_7_H0_significant_lifespan_mean (non-significant)

Least Neg

Least Pos

Peritumour

- Cluster_2↔Cluster_7_H0_significant_nBars (non-significant)
- Cluster_2_H0_significant_nBars (non-significant)
- Cluster_7_H0_significant_nBars (non-significant)
- Cluster_7_H1_significant_nBars (non-significant)
- Cluster_2↔Cluster_7_H1_significant_lifespan_mean (non-significant)
- Cluster_2↔Cluster_7_H1_significant_nBars (non-significant)
- Cluster_2↔Cluster_7_H1_significant_lifespan_std (non-significant)
- Cluster_7_H1_significant_lifespan_std (non-significant)
- Cluster_7_H1_significant_birth_mean (non-significant)
- Cluster_7_H1_significant_lifespan_mean (non-significant)
- Cluster_2↔Cluster_7_H1_significant_birth_std (non-significant)
- Cluster_2↔Cluster_7_H1_significant_birth_mean (non-significant)
- Cluster_7_H1_significant_birth_std (non-significant)

Most Pos

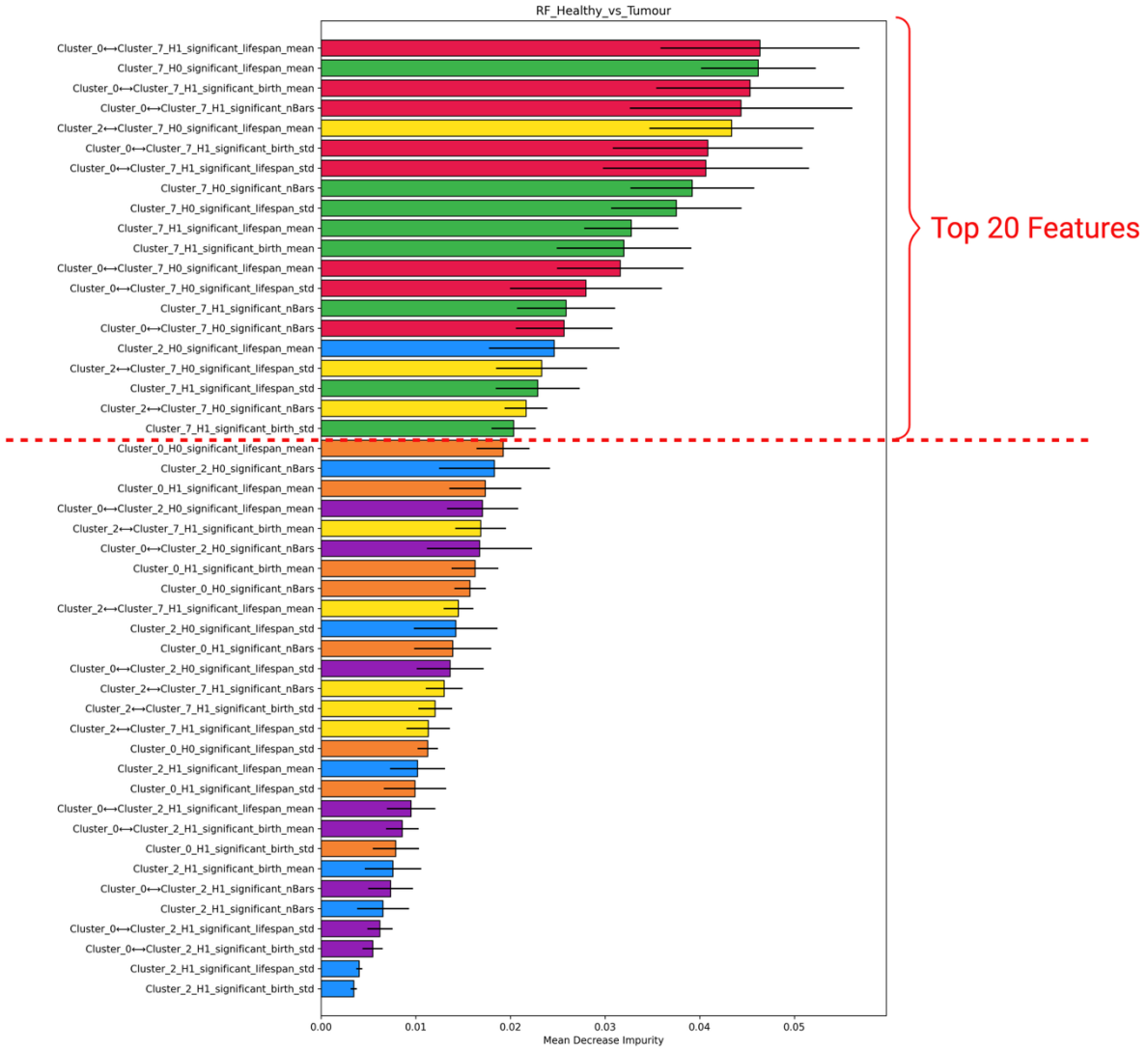
Figure 38. Topological features distinguishing peritumour and healthy regions in CRLM: feature importance and statistical significance (CosMx). A. A ranked bar plot of the TDA features used to discriminate between peritumour and healthy regions, as identified by random forest classifier feature importance (mean decrease in impurity, MDI). The top 20 features, as ranked by average MDI, are indicated above a dashed red line. B. A volcano plot comparing the distribution of the same TDA features between regions. Each point represents one TDA feature, with the top 20 features highlighted and labelled, and those not passing FDR significance are noted as non-significant in the legend.

Healthy versus Tumour

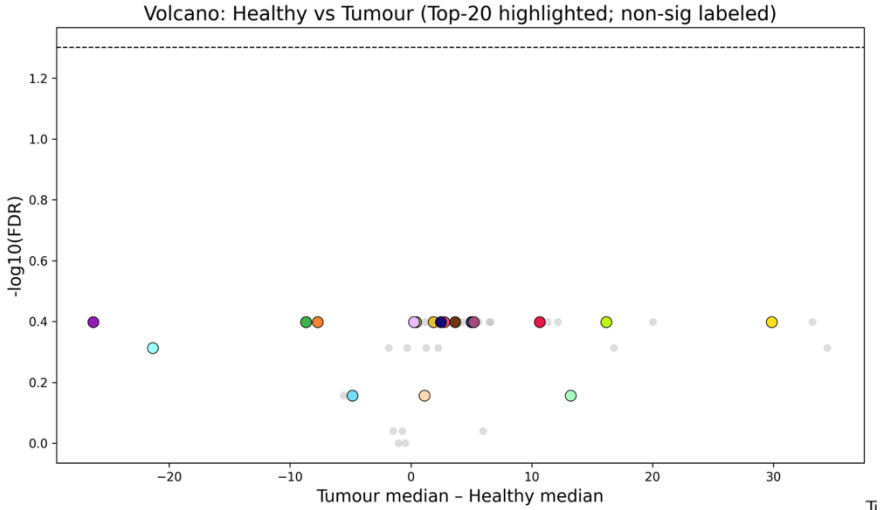
The variables that best separated tumour from healthy tissue again centred on the fibroblastic-immune interface (NC7) but now included its crosstalk with the myeloid/T-reg niche (NC0) (Fig.39, A). RF ranked NC0↔NC7 H_1 lifespan mean, NC0↔NC7 H_1 birth mean and s.d., and NC0↔NC7 H_1 lifespan s.d. alongside NC7 H_0 counts and lifespans. When comparing tumour to healthy tissue in the volcano plot, two main patterns became apparent (Fig.39, B). Features reflecting the abundance of NC7 structures or the enrichment of topological loops tended to be elevated in tumour regions. These included higher numbers of NC7 H_0 bars, increased NC2–NC7 H_0 counts, and several NC7 H_1 loop metrics, all pointing to more numerous and longer-lived fibroblastic–immune interface micro-regions occurring within the tumour core. In contrast, measures that captured the spatial separation of these structures, such as the mean and s.d. of NC7 H_0 lifespans and similar metrics for NC2–NC7, were greater in healthy liver, indicating that the fewer fibroblastic–immune interface elements presented, and they were more widely dispersed. The higher statistics of NC0–NC7 H_1 features in the tumour region further suggested that fibroblastic–immune interfaces were more frequently interleaved with myeloid and T-reg neighbourhoods, consistent with a more immunosuppressive microenvironment. Altogether, these findings portrayed the tumour core as having a more complex and fragmented fibroblastic–immune interface, with many long-lived NC7 islands and visible admixture of NC0 elements. In contrast, the healthy liver showed a simpler topological landscape with fewer but more widely separated NC7 structures.

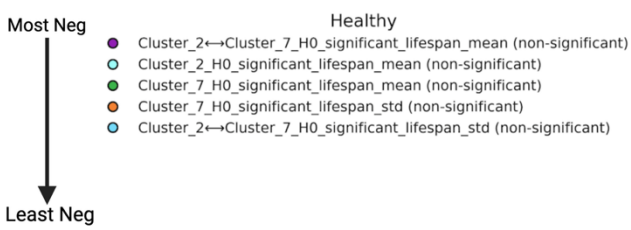
CosMx thus recapitulated the spatial logic seen with Cell DIVE while shifting emphasis toward intratumoural immune suppression. At the peritumour region, adaptive/APC neighbourhoods (NC2) exhibit more numerous and compact components, which are sometimes braided with NC7, whereas the tumour interior was dominated by interface-centred topology enriched for NC7 and its pairing with the myeloid/T-reg niche NC0. The healthy liver exhibited the fewest interface components, which were widely separated, and showed little evidence of either NC2 activity or NC0↔NC7 crosstalk, resulting in a comparatively simple topological landscape. Although none of these slide-level differences reached statistical significance in this small cohort, the qualitative agreement between RF feature rankings, UMAP structure, and volcano plot directions supported the findings of earlier sections, which highlighted a margin-centred concentration of adaptive immunity and an intratumoural admixture of fibroblastic and immunosuppressive programmes.

A.


RF_Healthy_vs_Tumour
Cluster_0⟷Cluster_7_H1_significant_lifespan_mean
Cluster_7_H0_significant_lifespan_mean
Cluster_0⟷Cluster_7_H1_significant_birth_mean
Cluster_0⟷Cluster_7_H1_significant_nBars
Cluster_2⟷Cluster_7_H0_significant_lifespan_mean
Cluster_0⟷Cluster_7_H1_significant_birth_std
Cluster_0⟷Cluster_7_H1_significant_lifespan_std
Cluster_7_H0_significant_nBars
Cluster_7_H0_significant_lifespan_std
Cluster_7_H1_significant_lifespan_mean
Cluster_7_H1_significant_birth_mean
Cluster_0⟷Cluster_7_H0_significant_lifespan_mean
Cluster_0⟷Cluster_7_H0_significant_lifespan_std
Cluster_7_H1_significant_nBars
Cluster_0⟷Cluster_7_H0_significant_nBars
Cluster_2_H0_significant_lifespan_mean
Cluster_2⟷Cluster_7_H0_significant_lifespan_std
Cluster_7_H1_significant_lifespan_std
Cluster_2⟷Cluster_7_H0_significant_nBars
Cluster_7_H1_significant_birth_std
Cluster_0_H0_significant_lifespan_mean
Cluster_2_H0_significant_nBars
Cluster_0_H1_significant_lifespan_mean
Cluster_0⟷Cluster_2_H0_significant_lifespan_mean
Cluster_2⟷Cluster_7_H1_significant_birth_mean
Cluster_0⟷Cluster_2_H0_significant_nBars
Cluster_0_H1_significant_birth_mean
Cluster_0_H0_significant_nBars
Cluster_2⟷Cluster_7_H1_significant_lifespan_mean
Cluster_2_H0_significant_lifespan_std
Cluster_0_H1_significant_nBars
Cluster_0⟷Cluster_2_H0_significant_lifespan_std
Cluster_2⟷Cluster_7_H1_significant_nBars
Cluster_2⟷Cluster_7_H1_significant_birth_std
Cluster_2⟷Cluster_7_H1_significant_lifespan_std
Cluster_0_H0_significant_lifespan_std
Cluster_2_H1_significant_lifespan_mean
Cluster_0_H1_significant_lifespan_std
Cluster_0⟷Cluster_2_H1_significant_lifespan_mean
Cluster_0⟷Cluster_2_H1_significant_birth_mean
Cluster_0_H1_significant_birth_std
Cluster_2_H1_significant_birth_mean
Cluster_0⟷Cluster_2_H1_significant_nBars
Cluster_2_H1_significant_nBars
Cluster_0⟷Cluster_2_H1_significant_lifespan_std
Cluster_0⟷Cluster_2_H1_significant_birth_std
Cluster_2_H1_significant_lifespan_std
Cluster_2_H1_significant_birth_std
0.00
0.01
0.02
0.03
0.04
0.05
Mean Decrease Impurity
Top 20 Features

B.


Volcano: Healthy vs Tumour (Top-20 highlighted; non-sig labeled)
-log10(FDR)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−20
−10
0
10
20
30
Tumour median − Healthy median


Most Neg
Least Neg
Healthy
Cluster_2⟷Cluster_7_H0_significant_lifespan_mean (non-significant)
Cluster_2_H0_significant_lifespan_mean (non-significant)
Cluster_7_H0_significant_lifespan_mean (non-significant)
Cluster_7_H0_significant_lifespan_std (non-significant)
Cluster_2⟷Cluster_7_H0_significant_lifespan_std (non-significant)

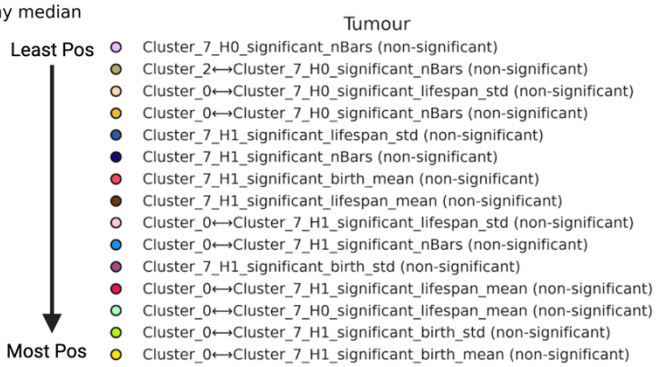

Tumour
Least Pos
Most Pos
Cluster_7_H0_significant_nBars (non-significant)
Cluster_2⟷Cluster_7_H0_significant_nBars (non-significant)
Cluster_0⟷Cluster_7_H0_significant_lifespan_std (non-significant)
Cluster_0⟷Cluster_7_H0_significant_nBars (non-significant)
Cluster_7_H1_significant_lifespan_std (non-significant)
Cluster_7_H1_significant_nBars (non-significant)
Cluster_7_H1_significant_birth_mean (non-significant)
Cluster_7_H1_significant_lifespan_mean (non-significant)
Cluster_0⟷Cluster_7_H1_significant_lifespan_std (non-significant)
Cluster_0⟷Cluster_7_H1_significant_nBars (non-significant)
Cluster_7_H1_significant_birth_std (non-significant)
Cluster_0⟷Cluster_7_H1_significant_lifespan_mean (non-significant)
Cluster_0⟷Cluster_7_H0_significant_lifespan_mean (non-significant)
Cluster_0⟷Cluster_7_H1_significant_birth_std (non-significant)
Cluster_0⟷Cluster_7_H1_significant_birth_mean (non-significant)

Figure 39. Topological features distinguishing healthy and tumour regions in CRLM: feature importance and statistical significance (CosMx). A. A ranked bar plot of the TDA features used to discriminate between healthy and tumour regions, as identified by random forest classifier feature importance (mean decrease in impurity, MDI). The top 20 features, as ranked by average MDI, are indicated above a dashed red line. B. A volcano plot comparing the distribution of the same TDA features between regions. Each point represents one TDA feature, with the top 20 features highlighted and labelled, and those not passing FDR significance are noted as non-significant in the legend.

Discussion

Progress in spatial biology techniques, including multiplex tissue imaging and spatial transcriptomics platforms, has enhanced our analysis of the TME, allowing researchers to comprehensively map cellular positions inside a tissue and apply sophisticated bioinformatics methods to examine spatial interactions among relevant cells (Lewis *et al.,* 2021). Platforms such as MIBI (Angelo *et al.,* 2014), CODEX (Goltsev *et al.,* 2018), and Cell DIVE (Gerdes *et al.,* 2013), alongside next-generation spatial transcriptomics (Ståhl *et al.,* 2016), permit simultaneous readouts of dozens of protein markers or thousands of genes *in situ*, retaining the coordinates necessary to interrogate proximity, neighbourhood structure, and higher-order geometry. In the field of oncology, these technologies have established that immune control and evasion are fundamentally encoded in tissue architecture. Previous studies have shown how recurrent cellular neighbourhoods, gradients of immune cells at tumour borders, and distance-dependent cell–cell interactions serve as spatial signatures that predict patient outcomes and therapy response in diverse malignancies such as breast cancer (Keren *et al.,* 2018), primary CRC (Schürch *et al.,* 2020; Pelka *et al.,* 2021), melanoma (Cabrita *et al.,* 2020), non-small cell lung cancer (Zhang *et al.,* 2022), and pancreatic ductal adenocarcinoma (Moncada *et al.,* 2020), highlighting the broad clinical relevance of spatial tissue profiling across cancer types. More recently, spatial-biology analytic pipelines have been applied directly to CRLM. Comparative profiling revealed that liver metastases are immunologically distinct from their primary tumours, showing an enrichment of immunosuppressive cell populations. Most notably, the unique spatial arrangement of $CD68^{-}CD163^{+}$ macrophages within these metastases is associated with a poor response to immunotherapy (He *et al.,* 2022b). Spatiotemporal single-cell and spatial mapping traced a shift during metastatic progression from inflamed lesions toward myeloid-dominated niches, where highly metabolically active M2-like macrophages emerged as a key suppressive feature within metastatic sites (Wu *et al.,* 2022a). Preoperative chemotherapy was linked to broad reprogramming of the metastatic TME, including activation of B-cell compartments, reduced diversity and maturation of tumour-associated macrophages, lower frequencies of dysfunctional T-cells, decreased ECM-remodelling CAFs, and accumulation of myofibroblasts (Che *et al.,* 2021). Wood et al. (2023) further deconvoluted histopathological growth patterns into molecular–immune subgroups, showing that desmoplastic rims carry adaptive immune signatures and better outcomes compared with replacement-type lesions. While these studies provide critical insights, a more general, multi-scale deconstruction of the architectural rules governing immune resistance in CRLM is still needed. In this study, we built on that paradigm for CRLM, leveraging two complementary spatial modalities, delicate tissue region definition/separation strategy, and a multiscale analytic pipeline that can capture cellular organisation beyond abundance to extract interpretable rules that connect liver-specific ecology, spatial compartmentalisation, and actionable immunobiology.

Following advice from the MuSpAn developers (Bull *et al.,* 2024; Bull *et al.,* 2025), our analyses were designed around small datasets of ROIs rather than using each whole-slide dataset directly. This choice had two practical benefits. First, cross-PCF and other second-order spatial statistics assumed the analysis window had roughly homogeneous cell density. However, whole-slide spatial data could be highly heterogeneous, as cell density can vary dramatically between the tumour core, peritumour, and healthy liver regions. If we compute a single global estimate across the entire slide, the normalisation is based on one overall density, which averages across these differences and can mask or distort true local clustering. By restricting the analysis to ROIs of the same size, we ensured that the density used for normalisation in each calculation was more consistent. This approach made the spatial statistics more stable and the results easier to interpret. For TDA, using the ROI-level dataset also ensured that we did not create artificial void-like structures when the selected ROIs were randomly positioned. Second, distance-based measures (e.g., cross-PCFs) and TDA scale quickly with the number of cells. Partitioning slides into ROIs greatly reduces runtime and memory demands while preserving the local patterns we aim to study in datasets with millions of cells.

Careful statistical designs were applied for reliable inference in this study, where multiple subregions are sampled from each patient. Statistically, we explicitly avoided pseudo-replication by treating each slide (patient) as the unit of inference. Since each patient (independent whole-slide sample) contributed multiple ROIs per region, treating ROIs as independent datapoints would artificially inflate the sample size and underestimate variability due to intra-slide correlations. To enable robust statistical comparisons, for every spatial feature within a region, we averaged ROI-level values within each slide and then compared regions using paired Wilcoxon signed-rank tests across slides (N=13 Cell DIVE; N=4 CosMx), controlling multiplicity as described elsewhere. This provides a conservative, reproducible framework for spatial-omics studies with many sub-samples per individual. For TDA, when evaluating discriminative performance, we trained classifiers using ROI-level feature matrices and validated their performance with grouped cross-validation, where all ROIs from a single patient remained together in the same fold. This approach respects the nested nature of the data, as ROIs from the same slide share staining conditions, imaging artefacts, tissue processing, and patient-specific biology. Without this grouping, ordinary random cross-validation could place ROIs from the same patient into both the training and test sets. This would allow the model to learn slide-specific signals, leading to an overestimate of accuracy because it measures performance on familiar patient data rather than new individuals. By grouping ROIs by slide, we avoid this information leakage and ensure that the reported accuracy reflects true out-of-patient generalisation rather than within-patient redundancy (i.e., how well the model generalises to entirely new patients, rather than just to other regions within the same sample) (Pedregosa *et al.,* 2011; Varoquaux *et al.,* 2017).

Within that context, the present work added a multi-scale map of how immune failure was organised in space in CRLM, and our results predominantly demonstrated TIME compartmentalisation, where regional shifts could be observed when spatial pattern metrics were used as regional features. Key biological mechanisms with therapeutic potential could be derived from each tissue region (Fig.38).

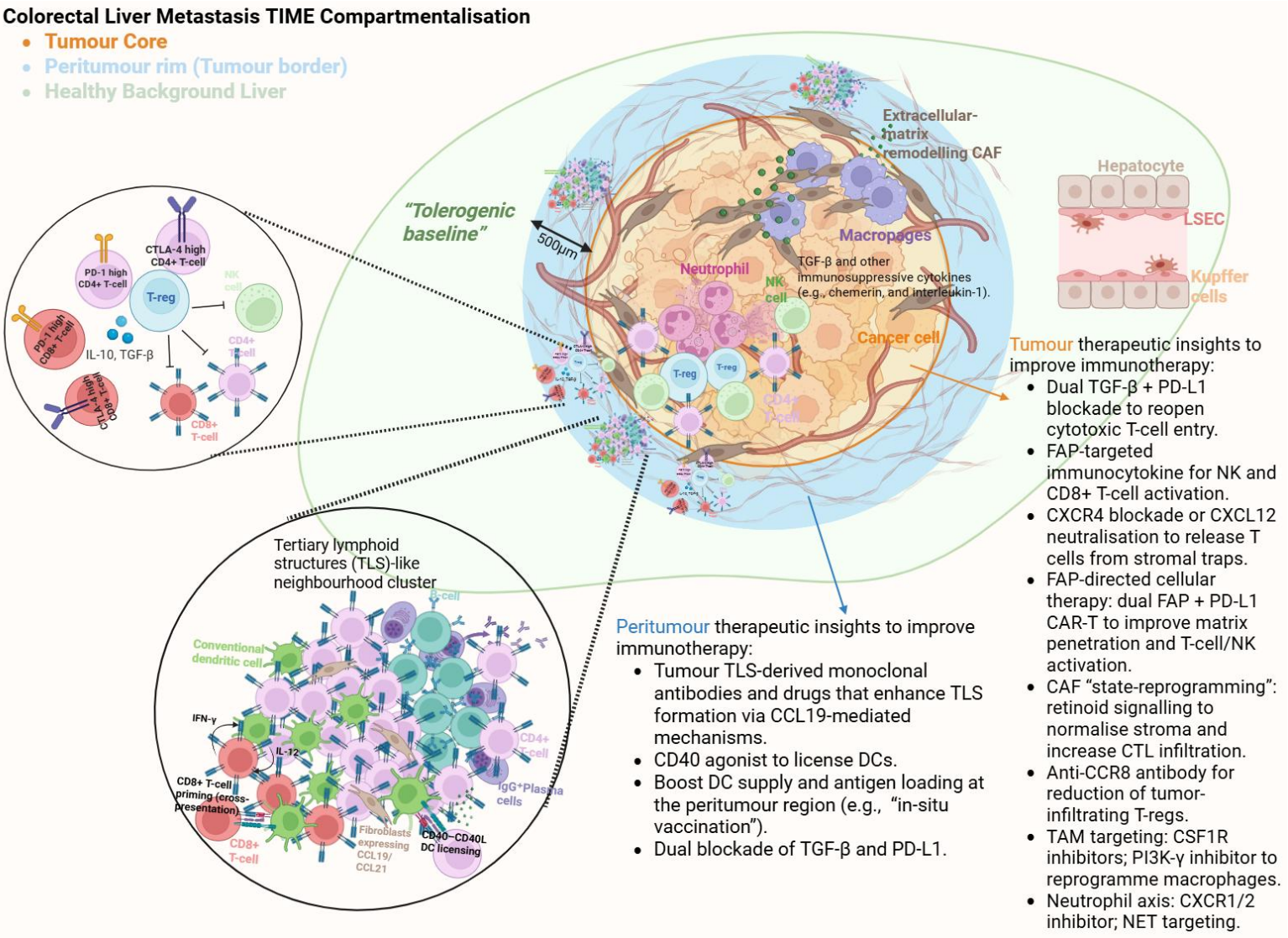


Figure 38. Spatial profiling techniques revealed TIME compartmentalisation in CRLM, providing potential mechanisms underly immunotherapy resistance and therapeutic opportunities.

Inside the peritumour rim, an outstanding organisational unit appeared to be NC2, a CD4-centred module that couples T helper cells to cDCs and B-lineage cells, sometimes including specialised fibroblasts expressing chemokines such as CCL19 and CCL21, which are well-established attractants for lymphoid cells and DCs, fostering the assembly of TLS at sites of chronic inflammation and within tumours (Zhang *et al.,* 2024; Pimenta and Barnes, 2014). Notably, a recent multi-omics study in CRLM showed that both peri- and intra-tumour TLSs correlated with improved overall and relapse-free survival after liver resection, and most TLSs localised within around 2 mm of the tumour margin, where they form the "front line" of anti-tumour immunity, supporting that grading TLSs as a single entity across peri- and intra-tumour compartments for prognostic use in CRLM (Zhang *et al.,* 2024). Biologically, the good prognosis may be due to rim-TLSs in CRLM containing fibroblastic reticular cell-like CAFs that produce CCL19, recruiting CCR7+ T-cells, B-cells, and cDCs into organised niches. And exogenous CCL19 was found to promote TLS neogenesis and tumour control in vivo (Zhang *et al.,* 2024; Filderman *et al.,* 2021; Chang *et al.,* 2019). Moreover, the TLS-positive lesions were found to be enriched for IgG+ plasma cells, which coated tumour cells, associated with epithelial apoptosis signatures, and inhibited human CRC growth in mice. Whereas TLS-negative cases preferentially

harboured IgA$^+$ plasma cells, which would induce activations of myeloid-derived suppressor cells in tumours (Zhang *et al.,* 2024; Chen *et al.,* 2023c). Further investigations into tumour TLS-derived monoclonal antibodies and drugs that enhance TLS formation via CCL19-mediated mechanisms may potentially revolutionise the immunotherapeutic approaches for CRLM, particularly in desmoplastic lesions (Zhang *et al.,* 2024). In fact, TLS was already found to correlate with improved immunotherapy response and survival in other solid tumours, such as in melanoma (anti-CTLA4 and anti-PD-1 antibodies), triple-negative breast cancer (anti-PD-L1 antibodies), and soft-tissue sarcoma (anti-PD-1 antibodies) (Cabrita *et al.,* 2020; Zhang *et al.,* 2021; Italiano *et al.,* 2022). Although the prognostic value of TLS in CRLM still needs to be further confirmed, these findings, together, argue that peritumoural CRLM is partially immune-active, and the TLS-like structures within it may provide both B-cell–mediated effector functions and a scaffold for local T-cell priming.

A central axis linking these TLS structures to effective anti-tumour cytotoxic control is the CD4→DC→CD8 axis. Foundational work established that CD4$^+$ T-cells "license" DCs, especially the cDC1 subtype, through the CD40-CD40L interaction, so that these activated DCs can efficiently prime, sustain, and guide CD8$^+$ T-cell responses (Schoenberger *et al.,* 1998; Wu *et al.,* 2022b; Bennett *et al.,* 1998). Later studies confirmed that the Batf3-dependent cDC1 lineage was indispensable for cross-presentation, T-cell recruitment, and tumour rejection (Hildner *et al.,* 2008; Böttcher *et al.,* 2018b). More recently, in vivo evidence showed that early priming of CD4$^+$ T-cells against tumour-derived antigens also depended on cDC1, as selective deletion of MHC-II in cDC1 alone was sufficient to abrogate early CD4$^+$ T-cell priming (Ferris *et al.,* 2020). Consistent with these mechanisms, effective PD-1 blockade depends on intratumoural crosstalk between DCs and T cells, mediated through IL-12 and IFN-γ signalling, indicating that DC licensing and antigen cross-presentation constitute key bottlenecks for successful checkpoint therapy (Garris *et al.* 2018). In our present dataset, short-range enrichment of CD8–DC, CD4–DC, and CD4–CD8 interactions at peritumour was captured by ROI-based cross-PCF metrics, NC2 TDA metrics, and whole-slide permutation-based KNN analyses (albeit with some dilution), consistent with a model in which CD4$^+$ T-cell-organised neighbourhoods, particularly those resembling TLS, create favourable conditions for CD8$^+$ T-cell infiltration and activation. These insights point to peritumour-rim-targeted therapeutic strategies that expand and activate cross-priming DCs while reinforcing CD4–DC licensing. CD40 agonist antibodies mimic CD4$^+$ T-cell help through CD40–CD40L, converting cDCs into potent initiators of CTL responses. Early clinical studies combining CD40 agonists (e.g., Sotigalimab) with anti-PD-1 antibodies (e.g., Nivolumab) in advanced melanoma and pancreatic ductal adenocarcinoma showed reduced immunosuppressive features within patients, improved tumour sensitivity to anti-PD-1 immunotherapy, and an emerging safety–efficacy profile (Weiss *et al.,* 2024; Padrón *et al.,* 2022; Bullock, 2022). Next-generation bi-specific CD40/CD47 antibodies are being engineered to enhance tumour-local activity while limiting systemic toxicity (Andersson *et al.,* 2024). A complementary approach is to boost DC supply and antigen loading directly at the peritumour region. Recombinant human FLT3 ligand (FLT3L) can increase circulating and tissue DC subsets in humans (Anandasabapathy *et al.,* 2015). "In-situ vaccination" regimens based on this concept applied intratumoural FLT3L to recruit/expand DCs, pattern-recognition agonists (e.g., poly-ICLC) to provide maturation signals, and low-dose radiotherapy to release tumour antigens (Marron *et al.,* 2016; Svensson-Arvelund *et al.,* 2022). Phase

I/II trials pairing this cross-priming-boosting vaccine with Pembrolizumab are currently under active evaluation in lymphoma, metastatic breast cancer, and head and neck squamous cell carcinoma (ClinicalTrials.gov (NCT03789097), 2019).

Despite the concentration of TLS-like CD4–cDC–B modules at the peritumour rim, this border region is not uniformly permissive. Immune activation coexists with potent checkpoints that limit cytotoxic responses. Cross-PCFs in our data showed short-range adjacency between T-regs and both $CD4^+$ and $CD8^+$ subsets at the margin, consistent with a locally immunosuppressive microenvironment where T-regs constrain helper and effector T-cell activities through direct cell–cell interactions and/or the production of suppressive cytokines such as IL-10 and TGF-β (Taylor *et al.,* 2006; Wang *et al.,* 2023b). The higher proportions of exhausted T-cells (both PD-1-high and CTLA-4-high) at the peritumour region and short-range proximities between T-reg and (activated and exhausted) T-cells detected by the Cell DIVE dataset supported this immunosuppressive axis further. In parallel, CAFs form a second regulatory layer. Previous in vivo studies using metastatic CRC models showed that T-reg-mediated immunosuppression at invasive fronts was reinforced by TGF-β signalling in fibroblasts, a feature associated with poor prognosis in CRC (Tauriello *et al.,* 2018; Calon *et al.,* 2015; Calon *et al.,* 2012). Consistent cross-tumour evidence further showed that dual blockade of TGF-β and PD-L1 reduced stromal TGF-β activity, permitted $CD8^+$ T-cell penetration into previously excluded tumour cores, and elicited durable antitumour responses (Tauriello *et al.,* 2018; Mariathasan *et al.,* 2018). Extending this concept, a recent in vivo work demonstrated that combined PD-L1/TGF-β inhibition not only improved infiltration but also expanded and differentiated stem-like ($TCF1^+$) $CD8^+$ T-cells capable of replenishing exhausted pools (Castiglioni *et al.,* 2023). Furthermore, in the Cell DIVE data, we observed both $aSMA^+$ and FAP^+ CAFs presented at the peritumour region. Functionally, FAP^+ CAFs can secrete CXCL12 to trap $CXCR4^+$ T-cells within stromal scaffolds and away from tumour nests, disrupting this axis with CXCR4 blockade or CXCL12 neutralisation could release T-cells into tumour cores and enhance anti-PD-L1 efficacy in vivo (Feig *et al.,* 2013; Fearon *et al.,* 2021; Zboralski *et al.,* 2017).

Across the two platforms, the tumour core displayed an immunosuppressive environment characterised by enrichment of extracellular-matrix remodelling CAFs, T-reg enriched neighbourhoods, and myeloid-CAF interaction programmes, which maintain T-cell exclusion and dysfunction. And topologically, these immunosuppressive features seemed to be spatially clustered within the tumour core. Like we mentioned above, when mitigating CAFs that deposit and remodel collagen-rich matrices, amplify solid stress, and limit lymphocyte engagement with malignant nests (Sahai *et al.,* 2020), dual inhibition of TGF-β and PD-L1 could be a therapeutic way to reopen access routes for CTL and NK cells, and convert previously immune-suppressive lesions into inflamed tissues (Castiglioni *et al.,* 2023; Mariathasan *et al.,* 2018; Viel *et al.,* 2016). Or to focus on CXCL12 production by CAFs, which keeps $CXCR4^+$ T-cells away from tumour islets. Neutralising CXCL12 with an L-RNA aptamer called NOX-A12 increased intratumoural T-cell accumulation and improved the activity of ICIs (e.g., Pembrolizumab) in an early phase clinical trial for microsatellite stable metastatic CRC and pancreatic cancer (Suarez-Carmona *et al.,* 2021; Zboralski *et al.,* 2017). Similarly, combining CXCR4 antagonist with anti-PD-1 antibodies in a trial for advanced pancreatic cancer also demonstrated increased overall survival and CTL tumour infiltration, together with reduced circulating T-regs and myeloid-derived suppressor cells (Bockorny *et al.,* 2020). Beyond targeting cytokine, multiple FAP-targeted modalities have shown promise. One emerging way to turn the stroma barrier into a local

immune amplifier is FAP-IL2v, which couples an anti-FAP antibody to an IL-2 variant engineered to abolish IL-2Rα binding while retaining IL-2Rβγ signalling. This directs IL-2 activity to FAP^+ stroma and preferentially expands $CD8^+$ T-cells and NK cells over T-regs. In preclinical models, it strengthened the activity of ICIs and CD40 agonist (Waldhauer *et al.,* 2021). Clinically, FAP-IL2v plus Atezolizumab showed manageable safety with antitumour activity in early-phase trials for solid tumours (Prenen *et al.,* 2024; Steeghs *et al.,* 2024). FAP-directed cellular therapy is another promising route. In CRC models, dual chimeric antigen receptor (CAR) T-cells co-targeting FAP and PD-L1 showed increased activation, boosted IFN-γ/TNF-α secretion with reduced IL-10, improved matrix penetration, and minimal toxicity to normal peripheral blood mononuclear cells in vitro (Gao *et al.,* 2025). However, FAP is also expressed by bone-marrow-derived mesenchymal stem cells and skeletal muscles, and early reports of cachexia with broad FAP targeting emphasised the need for safety-gated designs that confine activity to the tumour compartment (Liu *et al.,* 2019). In parallel, CAF "state-reprogramming" aims to normalise rather than ablate stroma. Retinoid signalling can push CAFs toward quiescent states and increase CTL infiltration in pancreatic cancer models. Clinically, all-trans-retinoic acid plus chemotherapy showed a safe profile in a phase I trial in Pancreatic cancer (Ene-Obong *et al.,* 2013; Kocher *et al.,* 2020). Complementing these stromal-focused approaches, selectively depleting intratumoural T-regs via CCR8 may act as another potent future treatment for CRLM, since CCR8 was preferentially expressed on tumour-infiltrating T-regs and higher CCR8 levels correlated with poorer survival. In CRC ex vivo models, adding anti-CCR8 therapy reduced the frequency and suppressive phenotype of $CCR8^+$ T-regs, reversed exhaustion markers on $CD4^+$ and $CD8^+$ T-cells, and restored antitumour activity, while in vivo models further revealed slowed tumour growth and complete tumour regressions, supporting combination with PD-1-based regimens in T-reg-rich tumours (Chen *et al.,* 2024).

The myeloid-cell accumulations in the tumour core may further amplify immune suppression. TAM are abundant and phenotypically skewed toward the metabolically active $MRC1^+$ $CCL18^+$ M2-like subtype that suppresses T-cell function and correlates with immune dysfunction during metastatic progression of CRLM (Wu *et al.,* 2022a). Another transcriptionally distinct macrophage state marked by SPP1, and these macrophages accumulate in hypoxic, necrotic tumour regions, promoting matrix remodelling and T-cell dysfunction through osteopontin–CD44 signalling, and consistently correlates with poor outcomes in CRC cohorts (Li *et al.,* 2023; Matusiak *et al.,* 2024). Moreover, our permutation and KNN-based short-distance neighbourhood analyses revealed clustering between FAP^+/$aSMA^+$ fibroblasts and macrophages within the tumour core, echoing previous findings that TAMs accumulate near CAF-rich regions and that $SPP1^+$ macrophages frequently co-localise with FAP^+ fibroblasts at intratumoural stromal–immune interfaces enriched for TGF-β and other immunosuppressive cytokines (Liu *et al.,* 2019; Qi *et al.,* 2020). High expression of $SPP1^+$ and FAP^+ in CRLM patients has also been linked to poorer responses to anti-PD-L1 therapy (Qi *et al.,* 2020). Meanwhile, αSMA^+ CAFs were found to release TGF-β, M-CSF, IL-6, IL-8, IL-10 and CCL2, recruiting monocytes and skewing macrophages toward an M2 phenotype. In turn, TAMs further activate CAFs, reinforcing a feed-forward circuit of immunosuppression and matrix remodelling (Liu *et al.,* 2019). Within the neutrophil axis, accumulation of tumour-associated neutrophils (TAN) in the tumour core may reinforce T-cell dysfunction. In colorectal and other cancers, neutrophils polarise along an N1 versus N2 spectrum with TGF-β driving pro-tumour phenotypes, and high systemic or intratumoural neutrophil signals often portend adverse outcomes, although context-specific exceptions exist. (Yan *et al.,* 2022; Rao *et al.,* 2012). Neutrophils also generate extracellular traps in CRLM that physically cover tumour cells, impair contacts with CTLs, awaken dormant hepatic micrometastases through

PAD4-dependent citrullination of matrix, and promote mesenchymal-to-epithelial transition (Yuzhalin *et al.,* 2018; Li *et al.,* 2022a). Although challenges exist, this myeloid axis might still be therapeutically actionable in CRLM. Targeting CSF1R in preclinical models reprogrammed or depleted M2-like suppressive macrophages, increased antigen presentation, and restored sensitivity to T-cell checkpoint therapy (Cannarile *et al.,* 2017; Wen *et al.,* 2022). Notably, CSF1R inhibitor (Pexidartinib) has already been approved by the FDA for systemic treatment of tenosynovial giant cell tumour (Monestime *et al.,* 2020). Inhibition of myeloid PI3K-γ with Eganelisib reconditioned TAMs toward an inflammatory state, improved responses to PD-1 blockade in multiple models and showed immune-remodelling with encouraging safety in early clinical studies for solid tumours (O'Connell *et al.,* 2024; Hong *et al.,* 2023). For neutrophil and polymorphonuclear myeloid-derived suppressor cells trafficking, inhibiting CXCR1 and CXCR2 on them with SX-682 reduced tumour recruitment of suppressive granulocytes, and enhanced the efficacy of T cell-based immunotherapy, such as adoptive T-cell/NK cell transfer and ICIs (Greene *et al.,* 2020; Sun *et al.,* 2019). This CXCR1/2 antagonist is also being tested with Nivolumab in microsatellite-stable metastatic CRC in early-phase trials (Johnson *et al.,* 2022). Lastly, strategies that dismantle neutrophil extracellular traps, including DNase-based degradation and PAD4 inhibition, reduced the formation of neutrophil extracellular traps and the growth of CRLM in vivo (Yuzhalin *et al.,* 2018; Rayes *et al.,* 2019).

Previous studies have identified an additional layer of prognostically relevant intratumoural interaction involving NK cells, which can attract cDC1s into tumours through the secretion of CCL5 and XCL1. The close spatial association of NK cells and cDC1s within tumours has been linked to improved patient survival (Böttcher *et al.,* 2018a; Böttcher *et al.,* 2018b). However, our CosMx data revealed considerable inter-patient variability in the short-range clustering of NK cells and cDCs within the CRLM tumour core, suggesting that this potentially favourable prognostic marker is not consistently observed in CRLM, hence may be unstable in predicting patients' prognoses in CRLM.

Healthy background liver ROIs displayed the expected parenchymal–vascular segregation with modest immune–stromal interface complexity, a pattern that reflects the liver's baseline tolerogenic program. In homeostasis, antigens in the liver are taken up within the slow-flowing sinusoidal network, where local hepatic, myeloid, and stromal cells influence lymphocytes to remain restrained instead of triggering inflammation (Gottwick *et al.,* 2022). This may help explain the comparatively simple topology and muted immune–CAF interleaving observed here. Recent studies emphasised the liver's role as an immune "hub" that filters blood-borne antigens but dampens effector priming, producing systemic tolerance to self and gut-derived ligands (Ronca *et al.,* 2025; Tiegs and Lohse, 2010). A central element of this tolerance is the sinusoidal endothelium. LSECs are highly endocytic cells that cross-present antigens on MHC-I and MHC-II, which preferentially drive naïve $CD8^+$ T-cell anergy rather than cytotoxic differentiation (DeLeve *et al.,* 2018; Limmer *et al.,* 2000). This depends in part on the upregulation of PD-L1 (B7-H1) expression on LSECs, while co-stimulatory molecules (CD80/CD86) for T-cell activation remain low (Diehl *et al.,* 2008). An in vivo study conducted by Höchst et al. (2012) further confirmed that circulating tumour antigens, such as carcinoembryonic antigen (CEA) from CRC cells, were preferentially scavenged and cross-presented by LSECs, rendering antigen-specific $CD8^+$ T-cells non-responsive via PD-L1 and thereby preventing the control of CEA^+

tumour growth. Moreover, subsequent studies showed that LSECs could also impair DC-mediated $CD8^+$ T-cell activation and cytokine secretion (e.g., IL-2 and TNF) via potential antigen-presentation competition and could drive CD8 T-cell exhaustion via chronic stimulation (Kremer *et al.* 2024; Zheng and Tian, 2019). Liver-resident macrophages (Kupffer cells) reinforce the inherent immunosuppressive environment. Unlike conventional APCs, Kupffer cells express lower levels of T-cell co-stimulatory molecules and MHC-II, while producing high levels of IL-10, TGF-β and prostaglandins that suppress pro-inflammatory $CD4^+$ T-cell activation, skew effector differentiation, and favour T-reg expansion (Gottwick *et al.,* 2022; Zheng *et al.,* 2023; You *et al.,* 2008). However, recent evidence showed that certain subsets of Kupffer cells can reverse $CD8^+$ T-cell tolerance in the liver when exposed to IL-2, suggesting that although Kupffer cells are powerful inducers of immune tolerance under homeostatic conditions, their functional plasticity in different environments makes them less suitable as mediators for antigen-specific immunotherapy (De Simone *et al.,* 2021; Gottwick *et al.,* 2022; Li *et al.,* 2022b).

Overall, the tolerogenic state of the liver explains why immune activation in CRLM is primarily localised at the peritumour region. Overcoming the default tolerance of the liver requires robust pro-inflammatory responses and structured antigen presentation, such as those provided by TLS-like niches. In contrast, the tumour core, layered upon the baseline liver tolerance, often demands not just an increased supply of effector cells but the active dismantling of stromal and myeloid barriers to enable cytotoxic cell entry. Therefore, therapeutic strategies that enhance DC licensing alongside alleviating the suppressive effects of CAFs, T-regs, and myeloid cells are more likely to promote effective antitumour immunity within these CRLM lesions (Zheng and Tian, 2019).

Limitations and Future Perspectives

While our dual-cohort, multi-scale analysis provides a robust spatial atlas of the CRLM TME, several limitations should be acknowledged. The chief limitation is the small cohort size for the spatial transcriptomics CosMx dataset (N=4). This reduces the statistical power for slide-level comparisons, resulting in some CosMx-specific inferences being qualitative and primarily hypothesis-generating. Therefore, findings derived solely from the CosMx data should be considered hypothesis-generating. Nonetheless, the strong concordance of key architectural patterns, such as immune exclusion, peritumour T-cell enrichment, and a myeloid-skewed tumour core, between the protein-level (Cell DIVE, N=13) and transcript-level (CosMx) readouts provides crucial cross-modality validation, boosting confidence that these are robust biological signals rather than platform-specific artefacts. Second, while our ROI-centric strategy is a principled and necessary choice for robust local spatial statistics, it inherently trades a fraction of the whole-organ context for local interpretability. This approach samples the tissue rather than analysing it whole, which could miss larger, slide-scale architectural gradients. The manual annotation and selection of ROIs from the Cell DIVE dataset, though guided by biological intent, could also introduce selection bias, potentially over- or under-representing specific microenvironments. To counteract these potential biases, our statistical framework was anchored at the slide level. For broad characterisations, such as quantifying overall cell type and neighbourhood proportions, we averaged metrics across all available ROIs within a given

region for each patient. This approach provides a comprehensive regional summary. In contrast, for more targeted biological questions, such as analysing the composition of immune-rich niches or the topology of specific neighbourhood clusters, we first applied a principled filtering step. Only ROIs meeting predefined criteria, such as sufficient immune cell infiltration or the presence of a target neighbourhood, were included in the subsequent slide-level averaging. This ensures that our analysis of specialised microenvironments is not diluted by irrelevant or sparsely populated regions. In both scenarios, by ensuring the patient remains as the unit of biological replication, we mitigated some intra-patient correlations, and the results reflected inter-patient variability more. Third, spatial metrics, especially statistics like cross-PCFs, are sensitive to technical nuances such as boundary handling and intensity heterogeneity. Our use of alpha-shapes for edge correction provides a more accurate boundary definition than a simple bounding box, but it remains an approximation that can influence distance calculations near the ROI periphery. Similarly, per-ROI computation helps control for large-scale density variations between regions, but micro-heterogeneity within an ROI can still affect local normalisation. While our approach represents a robust standard practice to minimise these concerns, they are inherent challenges in the field.

Fourth, although we used cross-platform datasets for results validation, the protein markers used for Cell DIVE and the transcriptomic markers used for CosMx did not fully overlap. This partial concordance prevents a comprehensive one-to-one cross-validation of all identified cell types. For example, our Cell DIVE panel did not include protein markers for DCs, NK cells, and B cells. This is a significant omission, as these cells are critical players in orchestrating anti-tumour immunity. The absence of these markers in our proteomics panel means that although we could identify TLS-like structures in the CosMx data, we could not fully validate their protein-level composition or functional state across the larger cohort. Fifth, our analysis was constrained by the granularity of the marker panels. For Cell DIVE, we lacked markers that would allow the resolution of functionally distinct myeloid states, such as $SPP1^+$ TAMs, *MRC1*$^+$ *CCL18*$^+$ M2-like macrophages, $CD163^+$ M2 macrophages, and $C1QC^+$ resident tissue macrophages, each of which has been implicated in immunosuppression, stromal crosstalk, or immunotherapy prognosis in metastatic CRC (Sathe *et al.,* 2023; Zhang *et al.,* 2025; Wu *et al.,* 2022a). We also could not delineate neutrophil subtypes such as $CXCR2^+$ TANs or populations prone to form neutrophil extracellular traps (Zhao *et al.,* 2017). Similarly, the Cell DIVE panel also lacked key functional markers for CTL activation, such as Granzyme B, which is a canonical mediator of killing (Hay *et al.,* 2022). For the CosMx dataset, although the transcriptomic depth was greater, our clustering did not resolve these specific myeloid subtypes, representing an area for further in-depth analysis.

Sixth, our study did not stratify patients by the histopathological growth pattern of their metastases. Liver metastases from CRC can be classified at the tumour–liver interface into distinct growth patterns, most commonly the desmoplastic pattern and the replacement pattern, which carry different prognostic values and reflect different biology (van Dam *et al.,* 2017; Höppener *et al.,* 2024). The desmoplastic pattern is angiogenic and forms a fibrotic rim that often contains higher densities of tumour-infiltrating lymphocytes, hence this pattern associates with improved survival after resection (Galjart *et al.,* 2019). In CRLM exhibiting the replacement growth pattern, cancer cells co-

opt pre-existing blood vessels from the surrounding liver tissue rather than inducing sprouting angiogenesis. This pattern confers resistance to anti-angiogenic therapies such as Bevacizumab and is associated with poorer clinical outcomes (Frentzas *et al.,* 2016). By not correlating our spatial findings with the growth pattern of each lesion, we potentially introduced selection biases when describing the general microenvironment of each region. Also, we missed an opportunity to test whether specific neighbourhood signatures concentrate within desmoplastic rims versus replacement fronts. Seventh, to convey a general view of the CRLM microenvironment, our patient metadata showed high heterogeneity in many recorded features (Table.1), such as prior treatments received before liver resection, primary CRC site, dissemination site, etc. Hence, we could not capture the more precise microenvironment that potentially would only appear in a few special cases. Eighth, and most critically, spatial proximity is not equivalent to functional interaction. Our findings are descriptive and establish strong correlations that now require functional validation to prove causality. Cellular communication occurs through both direct cell-cell contact (juxtacrine signalling) and the diffusion of secreted factors over short distances (paracrine signalling). While our proximity analyses are well-suited to generating hypotheses about juxtacrine interactions, they cannot fully capture the impact of paracrine signalling, which can influence cells that are not immediate neighbours. Observing that a T-reg is the nearest neighbour to a CD8+ T cell strongly suggests a contact-dependent suppressive interaction, but does not confirm it, nor does it preclude the influence of suppressive cytokines like IL-10 or TGF-β acting over a wider area. Computational ligand-receptor inference from spatial transcriptomics data can add a layer of mechanistic plausibility by identifying potential signalling axes (Wang *et al.,* 2022), but the next essential step is to move from geometry to causal mechanisms through wet-lab experiments. Finally, our two-dimensional analysis provides only a cross-section of a complex three-dimensional reality, potentially obscuring the true topology of infiltration paths, the full volume of a TLS, or the contiguous nature of the fibrotic barrier.

These limitations illuminate an exciting path forward. A recommended next step is to transition from spatial correlation to functional causality using advanced *ex vivo* models. Patient-derived organoid co-culture systems, which integrate tumour organoids with patient-matched immune and stromal cells, provide a powerful platform to mechanistically dissect the interactions we have mapped (Thorel *et al.,* 2024; Neal *et al.,* 2018). In such "TME-in-a-dish" models, researchers can directly test whether CAFs from an immune-suppressive tumour are sufficient to block T-cell infiltration and killing of organoids, and whether this effect can be reversed by targeting specific pathways like TGF-β. This approach enables robust, patient-specific validation and preclinical screening of barrier-disrupting therapies. To overcome the constraints of two-dimensional analysis, the field is rapidly advancing toward volumetric atlases of the TME. One strategy reconstructs serial sections into cohesive three-dimensional models using tools such as STitch3D and Spateo, which align multi-slice spatial transcriptomics and generate 3D molecular holograms of tissues (Shu *et al.,* 2024; Qiu *et al.,* 2024).
The complementary strategy performs direct high-resolution imaging of intact, centimetre-scale tissue blocks using light-sheet fluorescence microscopy on cleared specimens (Schueth *et al.,* 2023).
These three-dimensional approaches will reveal the true, continuous topology of the fibro-myeloid barrier, map the tortuous paths of T-cell migration, and define the complete volume of functional immune hubs like TLS. Ultimately, the goal is to integrate these advanced imaging modalities with

other data layers, including spatial genomics, epigenomics, and metabolomics, to build comprehensive, multi-omics spatial atlases of CRLM. Large-scale collaborative initiatives like the MOSAIC project are already underway to generate such big pan-cancer datasets for thousands of patients, which provide substrates for AI-driven discovery of novel biological mechanisms and therapeutic targets (Lehar *et al.,* 2023). This deeper, multi-modal view will catalyse the emergence of next-generation spatial biomarkers, such as quantifying tissue morphological features or cell–cell proximity signatures, which have already demonstrated decent predictive power for immunotherapy response compared to conventional markers like PD-L1 expression level and tumour mutation burden (Williams *et al.,* 2024; Shamai *et al.,* 2022; Park *et al.,* 2022; Hu *et al.,* 2021). Finally, the prospective validation of these spatial signatures in well-designed clinical trials will be essential to translate these architectural insights into clinically actionable tools that can guide the rational design of TME-remodelling therapies and truly personalise the future of cancer treatment.

Acknowledgement

The author would like to express sincere gratitude to the following people for their invaluable support and contributions:

Dr. Alex Gordon-Weeks, for his dedicated guidance and insightful feedback throughout the year, as well as his essential support at every stage of this work;
Dr. Areeb Mian, for generously sharing access to the Cell DIVE dataset;
Dr. Peter Wan, for the generous provision of the CosMx dataset;
Jue Ban, for being a supportive friend and lab mate;
Joanne Russell, for her unfailing assistance with all detailed programme-related questions;
The entire education team of the Department of Oncology, for organising events and overseeing progress throughout the year;
and all the family members and dear friends, for their unwavering emotional support.

References

Aitchison, J. (1982) 'The statistical analysis of compositional data', *Journal of the Royal Statistical Society: Series B (Methodological),* 44(2), pp. 139–160.

Ali, D., Asaad, A., Jimenez, M.-J., Nanda, V., Paluzo-Hidalgo, E. and Soriano-Trigueros, M. (2023) 'A survey of vectorization methods in topological data analysis', *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 45(12), pp. 14069–14080.

Anandasabapathy, N., Breton, G., Hurley, A., Caskey, M., Trumpfheller, C., Sarma, P., Pring, J., Pack, M., Buckley, N., Matei, I., Lyden, D., Green, J., Hawthorne, T., Marsh, H. C., Yellin, M., Davis, T., Keler, T. and Schlesinger, S. J. (2015) 'Efficacy and safety of CDX-301, recombinant human Flt3L, at expanding dendritic cells and hematopoietic stem cells in healthy human volunteers', *Bone Marrow Transplant,* 50(7), pp. 924-30.

Andersson, H., Nyesiga, B., Hermodsson, T., Enell Smith, K., Hägerbrand, K., Lindstedt, M. and Ellmark, P. (2024) 'Next-generation CD40 agonists for cancer immunotherapy', *Expert Opinion on Biological Therapy,* 24(5), pp. 351-363.

Angelo, M., Bendall, S. C., Finck, R., Hale, M. B., Hitzman, C., Borowsky, A. D., Levenson, R. M., Lowe, J. B., Liu, S. D., Zhao, S., Natkunam, Y. and Nolan, G. P. (2014) 'Multiplexed ion beam imaging of human breast tumors', *Nat Med,* 20(4), pp. 436-42.

Bagaev, A., Kotlov, N., Nomie, K., Svekolkin, V., Gafurov, A., Isaeva, O., Osokin, N., Kozlov, I., Frenkel, F., Gancharova, O., Almog, N., Tsiper, M., Ataullakhanov, R. and Fowler, N. (2021) 'Conserved pan-cancer microenvironment subtypes predict response to immunotherapy', *Cancer Cell,* 39(6), pp. 845-865.e7.

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I. W. H., Ng, L. G., Ginhoux, F. and Newell, E. W. (2019) 'Dimensionality reduction for visualizing single-cell data using UMAP', *Nature Biotechnology,* 37(1), pp. 38–44.

Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *Journal of the Royal Statistical Society. Series B (Methodological),* 57(1), pp. 289–300.

Bennett, S. R., Carbone, F. R., Karamalis, F., Flavell, R. A., Miller, J. F. and Heath, W. R. (1998) 'Help for cytotoxic-T-cell responses is mediated by CD40 signalling', *Nature,* 393(6684), pp. 478-80.

Binnewies, M., Roberts, E. W., Kersten, K., Chan, V., Fearon, D. F., Merad, M., Coussens, L. M., Gabrilovich, D. I., Ostrand-Rosenberg, S., Hedrick, C. C., Vonderheide, R. H., Pittet, M. J., Jain, R. K., Zou, W., Howcroft, T. K., Woodhouse, E. C., Weinberg, R. A. and Krummel, M. F. (2018)

'Understanding the tumor immune microenvironment (TIME) for effective therapy', *Nature Medicine,* 24(5), pp. 541-550.

Bockorny, B., Semenisty, V., Macarulla, T., Borazanci, E., Wolpin, B. M., Stemmer, S. M., Golan, T., Geva, R., Borad, M. J., Pedersen, K. S., Park, J. O., Ramirez, R. A., Abad, D. G., Feliu, J., Muñoz, A., Ponz-Sarvise, M., Peled, A., Lustig, T. M., Bohana-Kashtan, O., Shaw, S. M., Sorani, E., Chaney, M., Kadosh, S., Vainstein Haras, A., Von Hoff, D. D. and Hidalgo, M. (2020) 'BL-8040, a CXCR4 antagonist, in combination with pembrolizumab and chemotherapy for pancreatic cancer: the COMBAT trial', *Nat Med,* 26(6), pp. 878-885.

Bokemeyer, C., Bondarenko, I., Makhson, A., Hartmann, J. T., Aparicio, J., de Braud, F., Donea, S., Ludwig, H., Schuch, G., Stroh, C., Loos, A. H., Zubel, A. and Koralewski, P. (2009) 'Fluorouracil, leucovorin, and oxaliplatin with and without cetuximab in the first-line treatment of metastatic colorectal cancer', *J Clin Oncol,* 27(5), pp. 663-71.

Bolhuis, K., Kos, M., van Oijen, M. G. H., Swijnenburg, R. J. and Punt, C. J. A. (2020) 'Conversion strategies with chemotherapy plus targeted agents for colorectal cancer liver-only metastases: A systematic review', *Eur J Cancer,* 141, pp. 225-238.

Böttcher, J. P. and Reis e Sousa, C. (2018b) 'The Role of Type 1 Conventional Dendritic Cells in Cancer Immunity', *Trends Cancer,* 4(11), pp. 784-792.

Böttcher, J. P., Bonavita, E., Chakravarty, P., Blees, H., Cabeza-Cabrerizo, M., Sammicheli, S., Rogers, N. C., Sahai, E., Zelenay, S. and Reis e Sousa, C. (2018a) 'NK Cells Stimulate Recruitment of cDC1 into the Tumor Microenvironment Promoting Cancer Immune Control', *Cell,* 172(5), pp. 1022-1037.e14.

Breiman, L. (2001) 'Random Forests', *Machine Learning,* 45(1), pp. 5–32.

Buisman, F. E., Galjart, B., van der Stok, E. P., Balachandran, V. P., Boerner, T., Drebin, J. A., Grünhagen, D. J., Jarnagin, W. R., Kingham, T. P., Verhoef, C., D'Angelica, M. I. and Groot Koerkamp, B. (2020) 'Recurrence Patterns After Resection of Colorectal Liver Metastasis are Modified by Perioperative Systemic Chemotherapy', *World J Surg,* 44(3), pp. 876-886.

Bull, J. A., Moore, J. W., Mulholland, E. J., Leedham, S. J. and Byrne, H. M. (2025) 'MuSpAn: A Toolbox for Multiscale Spatial Analysis', *bioRxiv*, pp. 2024.12.06.627195.

Bull, J. A., Mulholland, E. J., Moore, J. W., Bosque, J. J., Stolz, B. J., Boen, J., Eggington, H. R., Belnoue-Davis, H. L., Jones, H., Gatenbee, C. D., Anderson, A. R. A., Easton, A., Todd, P., Cunningham, C., Taylor, S., Byrne, H. M. and Leedham, S. (2024) 'Integrating diverse statistical methods to analyse stage-discriminatory cell interactions in colorectal neoplasia', *bioRxiv*, pp. 2024.06.02.597010.

Bullock, T. N. J. (2022) 'CD40 stimulation as a molecular adjuvant for cancer vaccines and other immunotherapies', *Cellular & Molecular Immunology,* 19(1), pp. 14-22.

Cabrita, R., Lauss, M., Sanna, A., Donia, M., Skaarup Larsen, M., Mitra, S., Johansson, I., Phung, B., Harbst, K., Vallon-Christersson, J., van Schoiack, A., Lövgren, K., Warren, S., Jirström, K., Olsson, H., Pietras, K., Ingvar, C., Isaksson, K., Schadendorf, D., Schmidt, H., Bastholt, L., Carneiro, A., Wargo, J. A., Svane, I. M. and Jönsson, G. (2020) 'Tertiary lymphoid structures improve immunotherapy and survival in melanoma', *Nature,* 577(7791), pp. 561-565.

Calon, A., Espinet, E., Palomo-Ponce, S., Tauriello, D. V., Iglesias, M., Céspedes, M. V., Sevillano, M., Nadal, C., Jung, P., Zhang, X. H., Byrom, D., Riera, A., Rossell, D., Mangues, R., Massagué, J., Sancho, E. and Batlle, E. (2012) 'Dependency of colorectal cancer on a TGF-β-driven program in stromal cells for metastasis initiation', *Cancer Cell,* 22(5), pp. 571-84.

Calon, A., Lonardo, E., Berenguer-Llergo, A., Espinet, E., Hernando-Momblona, X., Iglesias, M., Sevillano, M., Palomo-Ponce, S., Tauriello, D. V., Byrom, D., Cortina, C., Morral, C., Barceló, C., Tosi, S., Riera, A., Attolini, C. S., Rossell, D., Sancho, E. and Batlle, E. (2015) 'Stromal gene expression defines poor-prognosis subtypes in colorectal cancer', *Nat Genet,* 47(4), pp. 320-9.

Cannarile, M. A., Weisser, M., Jacob, W., Jegg, A.-M., Ries, C. H. and Rüttinger, D. (2017) 'Colony-stimulating factor 1 receptor (CSF1R) inhibitors in cancer therapy', *Journal for ImmunoTherapy of Cancer,* 5(1), pp. 53.

Castiglioni, A., Yang, Y., Williams, K., Gogineni, A., Lane, R. S., Wang, A. W., Shyer, J. A., Zhang, Z., Mittman, S., Gutierrez, A., Astarita, J. L., Thai, M., Hung, J., Yang, Y. A., Pourmohamad, T., Himmels, P., De Simone, M., Elstrott, J., Capietto, A.-H., Cubas, R., Modrusan, Z., Sandoval, W., Ziai, J., Gould, S. E., Fu, W., Wang, Y., Koerber, J. T., Sanjabi, S., Mellman, I., Turley, S. J. and Müller, S. (2023) 'Combined PD-L1/TGFβ blockade allows expansion and differentiation of stem cell-like CD8 T cells in immune excluded tumors', *Nature Communications,* 14(1), pp. 4703.

Chang, J. E., Buechler, M. B., Gressier, E., Turley, S. J. and Carroll, M. C. (2019) 'Mechanosensing by Peyer's patch stroma regulates lymphocyte migration and mucosal antibody responses', *Nature Immunology,* 20(11), pp. 1506-1516.

Che, L.-H., Liu, J.-W., Huo, J.-P., Luo, R., Xu, R.-M., He, C., Li, Y.-Q., Zhou, A.-J., Huang, P., Chen, Y.-Y., Ni, W., Zhou, Y.-X., Liu, Y.-Y., Li, H.-Y., Zhou, R., Mo, H. and Li, J.-M. (2021) 'A single-cell atlas of liver metastases of colorectal cancer reveals reprogramming of the tumor microenvironment in response to preoperative chemotherapy', *Cell Discovery,* 7(1), pp. 80.

Chen, E. X., Jonker, D. J., Loree, J. M., Kennecke, H. F., Berry, S. R., Couture, F., Ahmad, C. E., Goffin, J. R., Kavan, P., Harb, M., Colwell, B., Samimi, S., Samson, B., Abbas, T., Aucoin, N., Aubin,

F., Koski, S. L., Wei, A. C., Magoski, N. M., Tu, D. and O'Callaghan, C. J. (2020) 'Effect of Combined Immune Checkpoint Inhibition vs Best Supportive Care Alone in Patients With Advanced Colorectal Cancer: The Canadian Cancer Trials Group CO.26 Study', *JAMA Oncol,* 6(6), pp. 831-838.

Chen, E. X., Loree, J. M., Titmuss, E., Jonker, D. J., Kennecke, H. F., Berry, S., Couture, F., Ahmad, C. E., Goffin, J. R., Kavan, P., Harb, M., Colwell, B., Samimi, S., Samson, B., Abbas, T., Aucoin, N., Aubin, F., Koski, S., Wei, A. C., Tu, D. and O'Callaghan, C. J. (2023a) 'Liver Metastases and Immune Checkpoint Inhibitor Efficacy in Patients With Refractory Metastatic Colorectal Cancer: A Secondary Analysis of a Randomized Clinical Trial', *JAMA Network Open,* 6(12), pp. e2346094-e2346094.

Chen, Q., Shen, M., Yan, M., Han, X., Mu, S., Li, Y., Li, L., Wang, Y., Li, S., Li, T., Wang, Y., Wang, W., Wei, Z., Hu, C. and Jin, A. (2024) 'Targeting tumor-infiltrating CCR8(+) regulatory T cells induces antitumor immunity through functional restoration of CD4(+) T(convs) and CD8(+) T cells in colorectal cancer', *J Transl Med,* 22(1), pp. 709.

Chen, Z., Han, F., Du, Y., Shi, H. and Zhou, W. (2023b) 'Hypoxic microenvironment in cancer: molecular mechanisms and therapeutic interventions', *Signal Transduction and Targeted Therapy,* 8(1), pp. 70.

Chen, Z., Zhang, G., Ren, X., Yao, Z., Zhou, Q., Ren, X., Chen, S., Xu, L., Sun, K., Zeng, Q., Kuang, M., Kuang, D.-M. and Peng, S. (2023c) 'Cross-talk between Myeloid and B Cells Shapes the Distinct Microenvironments of Primary and Secondary Liver Cancer', *Cancer Research,* 83(21), pp. 3544-3561.

ClinicalTrials.gov**.** (2019). *In Situ Vaccination With Flt3L, Radiation, and Poly-ICLC Combined With Pembrolizumab in Patients With Non-Hodgkin's Lymphoma, Metastatic Breast Cancer, and Head and Neck Squamous Cell Carcinoma.* ClinicalTrials.gov (Identifier: NCT03789097). Available at: https://clinicaltrials.gov/study/NCT03789097 (Accessed 24 August 2025).

De Greef, K., Rolfo, C., Russo, A., Chapelle, T., Bronte, G., Passiglia, F., Coelho, A., Papadimitriou, K. and Peeters, M. (2016) 'Multisciplinary management of patients with liver metastasis from colorectal cancer', *World J Gastroenterol,* 22(32), pp. 7215-25.

De Simone, G., Andreata, F., Bleriot, C., Fumagalli, V., Laura, C., Garcia-Manteiga, J. M., Di Lucia, P., Gilotto, S., Ficht, X., De Ponti, F. F., Bono, E. B., Giustini, L., Ambrosi, G., Mainetti, M., Zordan, P., Bénéchet, A. P., Ravà, M., Chakarov, S., Moalli, F., Bajenoff, M., Guidotti, L. G., Ginhoux, F. and Iannacone, M. (2021) 'Identification of a Kupffer cell subset capable of reverting the T cell dysfunction induced by hepatocellular priming', *Immunity,* 54(9), pp. 2089-2100.e8.

DeLeve, L. D. and Maretti-Mira, A. C. (2017) 'Liver Sinusoidal Endothelial Cell: An Update', *Semin Liver Dis,* 37(4), pp. 377-387.

Diehl, L., Schurich, A., Grochtmann, R., Hegenbarth, S., Chen, L. and Knolle, P. A. (2008) 'Tolerogenic maturation of liver sinusoidal endothelial cells promotes B7-homolog 1-dependent CD8+ T cell tolerance', *Hepatology,* 47(1), pp. 296-305.

Dou, L., Ono, Y., Chen, Y. F., Thomson, A. W. and Chen, X. P. (2018) 'Hepatic Dendritic Cells, the Tolerogenic Liver Environment, and Liver Disease', *Semin Liver Dis,* 38(2), pp. 170-180.

Ene-Obong, A., Clear, A. J., Watt, J., Wang, J., Fatah, R., Riches, J. C., Marshall, J. F., Chin-Aleong, J., Chelala, C., Gribben, J. G., Ramsay, A. G. and Kocher, H. M. (2013) 'Activated pancreatic stellate cells sequester CD8+ T cells to reduce their infiltration of the juxtatumoral compartment of pancreatic ductal adenocarcinoma', *Gastroenterology,* 145(5), pp. 1121–32.

Eng, C., Kim, T. W., Bendell, J., Argilés, G., Tebbutt, N. C., Di Bartolomeo, M., Falcone, A., Fakih, M., Kozloff, M., Segal, N. H., Sobrero, A., Yan, Y., Chang, I., Uyei, A., Roberts, L. and Ciardiello, F. (2019) 'Atezolizumab with or without cobimetinib versus regorafenib in previously treated metastatic colorectal cancer (IMblaze370): a multicentre, open-label, phase 3, randomised, controlled trial', *Lancet Oncol,* 20(6), pp. 849-861.

Esfahani, K., Roudaia, L., Buhlaiga, N., Del Rincon, S. V., Papneja, N. and Miller, W. H., Jr. (2020) 'A review of cancer immunotherapy: from the past, to the present, to the future', *Curr Oncol,* 27(Suppl 2), pp. S87-s97.

Fakih, M., Raghav, K. P. S., Chang, D. Z., Bendell, J. C., Larson, T., Cohn, A. L., Huyck, T. K., Cosgrove, D., Fiorillo, J. A., Garbo, L. E., Ravimohan, S., Potter, V., D'Adamo, D., Sharma, N., Wang, Y. A., Coppieters, S., Herpers, M., Soares Viana de Oliveira, C. and Paulson, A. S. 'Single-arm, phase 2 study of regorafenib plus nivolumab in patients with mismatch repair-proficient (pMMR)/microsatellite stable (MSS) colorectal cancer (CRC)', *Journal of Clinical Oncology,* 39(15_suppl), pp. 3560-3560.

Fares, J., Fares, M. Y., Khachfe, H. H., Salhab, H. A. and Fares, Y. (2020) 'Molecular principles of metastasis: a hallmark of cancer revisited', *Signal Transduction and Targeted Therapy,* 5(1), pp. 28.

Fearon, D. T. and Janowitz, T. (2021) 'AMD3100/Plerixafor overcomes immune inhibition by the CXCL12-KRT19 coating on pancreatic and colorectal cancer cells', *Br J Cancer,* 125(2), pp. 149-151.
Feig, C., Jones, J. O., Kraman, M., Wells, R. J. B., Deonarine, A., Chan, D. S., Connell, C. M., Roberts, E. W., Zhao, Q., Caballero, O. L., Teichmann, S. A., Janowitz, T., Jodrell, D. I., Tuveson, D. A. and Fearon, D. T. (2013) 'Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts

synergizes with anti–PD-L1 immunotherapy in pancreatic cancer', *Proceedings of the National Academy of Sciences,* 110(50), pp. 20212-20217.

Ferris, S. T., Durai, V., Wu, R., Theisen, D. J., Ward, J. P., Bern, M. D., Davidson, J. T. t., Bagadia, P., Liu, T., Briseño, C. G., Li, L., Gillanders, W. E., Wu, G. F., Yokoyama, W. M., Murphy, T. L., Schreiber, R. D. and Murphy, K. M. (2020) 'cDC1 prime and are licensed by CD4(+) T cells to induce anti-tumour immunity', *Nature,* 584(7822), pp. 624-629.

Filderman, J. N., Appleman, M., Chelvanambi, M., Taylor, J. L. and Storkus, W. J. (2021) 'STINGing the Tumor Microenvironment to Promote Therapeutic Tertiary Lymphoid Structure Development', *Front Immunol,* 12, pp. 690105.

Folprecht, G., Gruenberger, T., Bechstein, W. O., Raab, H. R., Lordick, F., Hartmann, J. T., Lang, H., Frilling, A., Stoehlmacher, J., Weitz, J., Konopke, R., Stroszczynski, C., Liersch, T., Ockert, D., Herrmann, T., Goekkurt, E., Parisi, F. and Köhne, C. H. (2010) 'Tumour response and secondary resectability of colorectal liver metastases following neoadjuvant chemotherapy with cetuximab: the CELIM randomised phase 2 trial', *Lancet Oncol,* 11(1), pp. 38-47.

Frentzas, S., Simoneau, E., Bridgeman, V. L., Vermeulen, P. B., Foo, S., Kostaras, E., Nathan, M. R., Wotherspoon, A., Gao, Z.-h., Shi, Y., Van den Eynden, G., Daley, F., Peckitt, C., Tan, X., Salman, A., Lazaris, A., Gazinska, P., Berg, T. J., Eltahir, Z., Ritsma, L., van Rheenen, J., Khashper, A., Brown, G., Nyström, H., Sund, M., Van Laere, S., Loyer, E., Dirix, L., Cunningham, D., Metrakos, P. and Reynolds, A. R. (2016) 'Vessel co-option mediates resistance to anti-angiogenic therapy in liver metastases', *Nature Medicine,* 22(11), pp. 1294-1302.

Galjart, B., Nierop, P. M. H., van der Stok, E. P., van den Braak, R. R. J. C., Höppener, D. J., Daelemans, S., Dirix, L. Y., Verhoef, C., Vermeulen, P. B. and Grünhagen, D. J. (2019) 'Angiogenic desmoplastic histopathological growth pattern as a prognostic marker of good outcome in patients with colorectal liver metastases', *Angiogenesis,* 22(2), pp. 355-368.

Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pagès, C., Tosolini, M., Camus, M., Berger, A., Wind, P., Zinzindohoué, F., Bruneval, P., Cugnenc, P. H., Trajanoski, Z., Fridman, W. H. and Pagès, F. (2006) 'Type, density, and location of immune cells within human colorectal tumors predict clinical outcome', *Science,* 313(5795), pp. 1960-4.

Gao, Y., Luo, C., Yang, H., Xie, Q., He, H., Li, J. and Miao, J. (2025) 'Enhanced efficacy of dual chimeric antigen receptor-T cells targeting programmed death-ligand 1 and cancer-associated fibroblasts in colorectal cancer in vitro', *Cytojournal,* 22, pp. 29.

Garris, C. S., Arlauckas, S. P., Kohler, R. H., Trefny, M. P., Garren, S., Piot, C., Engblom, C., Pfirschke, C., Siwicki, M., Gungabeesoon, J., Freeman, G. J., Warren, S. E., Ong, S., Browning, E., Twitty, C. G., Pierce, R. H., Le, M. H., Algazi, A. P., Daud, A. I., Pai, S. I., Zippelius, A., Weissleder, R. and Pittet,

M. J. (2018) 'Successful Anti-PD-1 Cancer Immunotherapy Requires T Cell-Dendritic Cell Crosstalk Involving the Cytokines IFN-γ and IL-12', *Immunity,* 49(6), pp. 1148-1161.e7.

Geng, Y., Feng, J., Huang, H., Wang, Y., Yi, X., Wei, S., Zhang, M., Li, Z., Wang, W. and Hu, W. (2022) 'Single-cell transcriptome analysis of tumor immune microenvironment characteristics in colorectal cancer liver metastasis', *Ann Transl Med,* 10(21), pp. 1170.

Gerdes, M. J., Sevinsky, C. J., Sood, A., Adak, S., Bello, M. O., Bordwell, A., Can, A., Corwin, A., Dinn, S., Filkins, R. J., Hollman, D., Kamath, V., Kaanumalle, S., Kenny, K., Larsen, M., Lazare, M., Li, Q., Lowes, C., McCulloch, C. C., McDonough, E., Montalto, M. C., Pang, Z., Rittscher, J., Santamaria-Pang, A., Sarachan, B. D., Seel, M. L., Seppo, A., Shaikh, K., Sui, Y., Zhang, J. and Ginty, F. (2013) 'Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue', *Proceedings of the National Academy of Sciences,* 110(29), pp. 11982-11987.

Goltsev, Y., Samusik, N., Kennedy-Darling, J., Bhate, S., Hale, M., Vazquez, G., Black, S. and Nolan, G. P. (2018) 'Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging', *Cell,* 174(4), pp. 968-981.e15.

Gordon, C. R., Rojavin, Y., Patel, M., Zins, J. E., Grana, G., Kann, B., Simons, R. and Atabek, U. (2009) 'A review on bevacizumab and surgical wound healing: an important warning to all surgeons', *Ann Plast Surg,* 62(6), pp. 707-9.
Gottwick, C., Carambia, A. and Herkel, J. (2022) 'Harnessing the liver to induce antigen-specific immune tolerance', *Semin Immunopathol,* 44(4), pp. 475-484.

Greene, S., Robbins, Y., Mydlarz, W. K., Huynh, A. P., Schmitt, N. C., Friedman, J., Horn, L. A., Palena, C., Schlom, J., Maeda, D. Y., Zebala, J. A., Clavijo, P. E. and Allen, C. (2020) 'Inhibition of MDSC Trafficking with SX-682, a CXCR1/2 Inhibitor, Enhances NK-Cell Immunotherapy in Head and Neck Cancer Models', *Clin Cancer Res,* 26(6), pp. 1420-1431.

Gregorutti, B., Michel, B. and Saint-Pierre, P. (2017) 'Correlation and variable importance in random forests', *Statistics and Computing,* 27(3), pp. 659–678.

Hay, Z. L. Z. and Slansky, J. E. (2022) 'Granzymes: The Molecular Executors of Immune-Mediated Cytotoxicity', *Int J Mol Sci,* 23(3).

He, S., Bhatt, R., Brown, C., Brown, E. A., Buhr, D. L., Chantranuvatana, K., Danaher, P., Dunaway, D., Garrison, R. G., Geiss, G., Gregory, M. T., Hoang, M. L., Khafizov, R., Killingbeck, E. E., Kim, D., Kim, T. K., Kim, Y., Klock, A., Korukonda, M., Kutchma, A., Lewis, Z. R., Liang, Y., Nelson, J. S., Ong, G. T., Perillo, E. P., Phan, J. C., Phan-Everson, T., Piazza, E., Rane, T., Reitz, Z., Rhodes, M., Rosenbloom, A., Ross, D., Sato, H., Wardhani, A. W., Williams-Wietzikoski, C. A., Wu, L. and Beechem, J. M. (2022a) 'High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging', *Nat Biotechnol,* 40(12), pp. 1794-1806.

He, Y., Han, Y., Fan, A. h., Li, D., Wang, B., Ji, K., Wang, X., Zhao, X. and Lu, Y. (2022b) 'Multi-perspective comparison of the immune microenvironment of primary colorectal cancer and liver metastases', *Journal of Translational Medicine,* 20(1), pp. 454.

Helmink, B. A., Reddy, S. M., Gao, J., Zhang, S., Basar, R., Thakur, R., Yizhak, K., Sade-Feldman, M., Blando, J., Han, G., Gopalakrishnan, V., Xi, Y., Zhao, H., Amaria, R. N., Tawbi, H. A., Cogdill, A. P., Liu, W., LeBleu, V. S., Kugeratski, F. G., Patel, S., Davies, M. A., Hwu, P., Lee, J. E., Gershenwald, J. E., Lucci, A., Arora, R., Woodman, S., Keung, E. Z., Gaudreau, P. O., Reuben, A., Spencer, C. N., Burton, E. M., Haydu, L. E., Lazar, A. J., Zapassodi, R., Hudgens, C. W., Ledesma, D. A., Ong, S., Bailey, M., Warren, S., Rao, D., Krijgsman, O., Rozeman, E. A., Peeper, D., Blank, C. U., Schumacher, T. N., Butterfield, L. H., Zelazowska, M. A., McBride, K. M., Kalluri, R., Allison, J., Petitprez, F., Fridman, W. H., Sautès-Fridman, C., Hacohen, N., Rezvani, K., Sharma, P., Tetzlaff, M. T., Wang, L. and Wargo, J. A. (2020) 'B cells and tertiary lymphoid structures promote immunotherapy response', *Nature,* 577(7791), pp. 549-555.

Heymann, F., Peusquens, J., Ludwig-Portugall, I., Kohlhepp, M., Ergen, C., Niemietz, P., Martin, C., van Rooijen, N., Ochando, J. C., Randolph, G. J., Luedde, T., Ginhoux, F., Kurts, C., Trautwein, C. and Tacke, F. (2015) 'Liver inflammation abrogates immunological tolerance induced by Kupffer cells', *Hepatology,* 62(1), pp. 279-91.

Hildner, K., Edelson, B. T., Purtha, W. E., Diamond, M., Matsushita, H., Kohyama, M., Calderon, B., Schraml, B. U., Unanue, E. R., Diamond, M. S., Schreiber, R. D., Murphy, T. L. and Murphy, K. M. (2008) 'Batf3 deficiency reveals a critical role for CD8alpha+ dendritic cells in cytotoxic T cell immunity', *Science,* 322(5904), pp. 1097-100.

Hinshaw, D. C. and Shevde, L. A. (2019) 'The Tumor Microenvironment Innately Modulates Cancer Progression', *Cancer Research,* 79(18), pp. 4557-4566.

Höchst, B., Schildberg, F. A., Böttcher, J., Metzger, C., Huss, S., Türler, A., Overhaus, M., Knoblich, A., Schneider, B., Pantelis, D., Kurts, C., Kalff, J. C., Knolle, P. and Diehl, L. (2012) 'Liver sinusoidal endothelial cells contribute to CD8 T cell tolerance toward circulating carcinoembryonic antigen in mice', *Hepatology,* 56(5), pp. 1924-33.

Hollander, M., Wolfe, D. A. and Chicken, E. (2013) *Nonparametric statistical methods.* John Wiley & Sons.

Hong, D. S., Postow, M., Chmielowski, B., Sullivan, R., Patnaik, A., Cohen, E. E. W., Shapiro, G., Steuer, C., Gutierrez, M., Yeckes-Rodin, H., Ilaria, R., O'Connell, B., Peng, J., Peng, G., Zizlsperger, N., Tolcher, A. and Wolchok, J. D. (2023) 'Eganelisib, a First-in-Class PI3Kγ Inhibitor, in Patients with Advanced Solid Tumors: Results of the Phase 1/1b MARIO-1 Trial', *Clin Cancer Res,* 29(12), pp. 2210-2219.

Höppener, D. J., Aswolinskiy, W., Qian, Z., Tellez, D., Nierop, P. M. H., Starmans, M., Nagtegaal, I. D., Doukas, M., de Wilt, J. H. W., Grünhagen, D. J., van der Laak, J. A. W. M., Vermeulen, P., Ciompi, F. and Verhoef, C. (2024) 'Classifying histopathological growth patterns for resected colorectal liver metastasis with a deep learning analysis', *BJS Open,* 8(6).

Hu, J., Cui, C., Yang, W., Huang, L., Yu, R., Liu, S. and Kong, Y. (2021) 'Using deep learning to predict anti-PD-1 response in melanoma and lung cancer patients from histopathology images', *Transl Oncol,* 14(1), pp. 100921.

Italiano, A., Bessede, A., Pulido, M., Bompas, E., Piperno-Neumann, S., Chevreau, C., Penel, N., Bertucci, F., Toulmonde, M., Bellera, C., Guegan, J. P., Rey, C., Sautès-Fridman, C., Bougoüin, A., Cantarel, C., Kind, M., Spalato, M., Dadone-Montaudie, B., Le Loarer, F., Blay, J. Y. and Fridman, W. H. (2022) 'Pembrolizumab in soft-tissue sarcomas with tertiary lymphoid structures: a phase 2 PEMBROSARC trial cohort', *Nature Medicine,* 28(6), pp. 1199-1206.

Johnson, B., Kopetz, S., Hwang, H., Yuan, Y., DePinho, R. A., Zebala, J. and Overman, M. J. (2022) 'STOPTRAFFIC-1: A phase I/II trial of SX-682 in combination with nivolumab for refractory RAS-mutated microsatellite stable (MSS) metastatic colorectal cancer (mCRC)', *Journal of Clinical Oncology,* 40(16_suppl), pp. TPS3638-TPS3638.

Johnson, W. E., Li, C. and Rabinovic, A. (2006) 'Adjusting batch effects in microarray expression data using empirical Bayes methods', *Biostatistics,* 8(1), pp. 118–127.

Kalluri, R. (2016) 'The biology and function of fibroblasts in cancer', *Nature Reviews Cancer,* 16(9), pp. 582-598.

Keren, L., Bosse, M., Marquez, D., Angoshtari, R., Jain, S., Varma, S., Yang, S.-R., Kurian, A., Van Valen, D., West, R., Bendall, S. C. and Angelo, M. (2018) 'A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging', *Cell,* 174(6), pp. 1373-1387.e19.

Kremer, K. N., Khammash, H. A., Miranda, A. M., Rutt, L. N., Twardy, S. M., Anton, P. E., Campbell, M. L., Garza-Ortiz, C., Orlicky, D. J., Pelanda, R., McCullough, R. L. and Torres, R. M. (2024) 'Liver sinusoidal endothelial cells regulate the balance between hepatic immunosuppression and immunosurveillance', *Front Immunol,* 15, pp. 1497788.

Kocher, H. M., Basu, B., Froeling, F. E. M., Sarker, D., Slater, S., Carlin, D., deSouza, N. M., De Paepe, K. N., Goulart, M. R., Hughes, C., Imrali, A., Roberts, R., Pawula, M., Houghton, R., Lawrence, C., Yogeswaran, Y., Mousa, K., Coetzee, C., Sasieni, P., Prendergast, A. and Propper, D. J. (2020) 'Phase I clinical trial repurposing all-trans retinoic acid as a stromal targeting agent for pancreatic cancer', *Nature Communications,* 11(1), pp. 4841.

Lehar, J., Madissoon, E., Chevallier, J., Schiratti, J. B., Kamburov, A., Barnes, R., Haignere, C., Joy, A., Dodacki, A., Hoffmann, C., Lastra, S. L., Perez, A. E., Bayard, Q., Loga, K. v., Chaperon, H., Vert, J.-P., Durand, E., Soumelis, V., Wainrib, G. and Clozel, T. (2023) 'MOSAIC: Multi-Omic Spatial Atlas in Cancer, effect on precision oncology', *Journal of Clinical Oncology,* 41(16_suppl), pp. e15076-e15076.

Lewis, S. M., Asselin-Labat, M.-L., Nguyen, Q., Berthelet, J., Tan, X., Wimmer, V. C., Merino, D., Rogers, K. L. and Naik, S. H. (2021) 'Spatial omics and multiplexed imaging to explore cancer biology', *Nature Methods,* 18(9), pp. 997-1012.

Li, D., Shao, J., Cao, B., Zhao, R., Li, H., Gao, W., Chen, P., Jin, L., Cao, L., Ji, S. and Dong, G. (2022a) 'The Significance of Neutrophil Extracellular Traps in Colorectal Cancer and Beyond: From Bench to Bedside', *Front Oncol,* 12, pp. 848594.

Li, N., Zhu, Q., Tian, Y., Ahn, K. J., Wang, X., Cramer, Z., Jou, J., Folkert, I. W., Yu, P., Adams-Tzivelekidis, S., Sehgal, P., Mahmoud, N. N., Aarons, C. B., Roses, R. E., Thomas-Tikhonenko, A., Furth, E. E., Stanger, B. Z., Rustgi, A., Haldar, M., Katona, B. W., Tan, K. and Lengner, C. J. (2023) 'Mapping and modeling human colorectal carcinoma interactions with the tumor microenvironment', *Nature Communications,* 14(1), pp. 7915.

Li, W., Chang, N. and Li, L. (2022b) 'Heterogeneity and Function of Kupffer Cells in Liver Injury', *Front Immunol,* 13, pp. 940867.

Limmer, A., Ohl, J., Kurts, C., Ljunggren, H. G., Reiss, Y., Groettrup, M., Momburg, F., Arnold, B. and Knolle, P. A. (2000) 'Efficient presentation of exogenous antigen by liver endothelial cells to CD8+ T cells results in antigen-specific T-cell tolerance', *Nat Med,* 6(12), pp. 1348-54.

Lin, N., Li, J., Yao, X., Zhang, X., Liu, G., Zhang, Z. and Weng, S. (2022) 'Prognostic value of neutrophil-to-lymphocyte ratio in colorectal cancer liver metastasis: A meta-analysis of results from multivariate analysis', *International Journal of Surgery,* 107, pp. 106959.

Liu, Q. L., Zhou, H., Zhou, Z. G. and Chen, H. N. (2023) 'Colorectal cancer liver metastasis: genomic evolution and crosstalk with the liver microenvironment', *Cancer Metastasis Rev,* 42(2), pp. 575-587.

Liu, T., Han, C., Wang, S., Fang, P., Ma, Z., Xu, L. and Yin, R. (2019) 'Cancer-associated fibroblasts: an emerging target of anti-cancer immunotherapy', *J Hematol Oncol,* 12(1), pp. 86.

Mantovani, A., Marchesi, F., Malesci, A., Laghi, L. and Allavena, P. (2017) 'Tumour-associated macrophages as treatment targets in oncology', *Nat Rev Clin Oncol,* 14(7), pp. 399-416.

Mariathasan, S., Turley, S. J., Nickles, D., Castiglioni, A., Yuen, K., Wang, Y., Kadel Iii, E. E., Koeppen, H., Astarita, J. L., Cubas, R., Jhunjhunwala, S., Banchereau, R., Yang, Y., Guan, Y., Chalouni, C., Ziai, J., Şenbabaoğlu, Y., Santoro, S., Sheinson, D., Hung, J., Giltnane, J. M., Pierce, A. A., Mesh, K., Lianoglou, S., Riegler, J., Carano, R. A. D., Eriksson, P., Höglund, M., Somarriba, L., Halligan, D. L., van der Heijden, M. S., Loriot, Y., Rosenberg, J. E., Fong, L., Mellman, I., Chen, D. S., Green, M., Derleth, C., Fine, G. D., Hegde, P. S., Bourgon, R. and Powles, T. (2018) 'TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells', *Nature,* 554(7693), pp. 544-548.

Marron, T., Bhardwaj, N., Hammerich, L., George, F., Kim-Schulze, S., Keler, T., Davis, T., Crowley, E., Salazar, A. and Brody, J. (2016) 'A Novel In Situ Vaccine: Intratumoral Flt3L and Poly-ICLC with Low-Dose Radiotherapy', *Clinical Lymphoma, Myeloma and Leukemia,* 16, pp. S97-S98.

Matusiak, M., Hickey, J. W., van, I. D. G. P., Lu, G., Kidziński, L., Zhu, S., Colburg, D. R. C., Luca, B., Phillips, D. J., Brubaker, S. W., Charville, G. W., Shen, J., Loh, K. M., Okwan-Duodu, D. K., Nolan, G. P., Newman, A. M., West, R. B. and van de Rijn, M. (2024) 'Spatially Segregated Macrophage Populations Predict Distinct Outcomes in Colon Cancer', *Cancer Discov,* 14(8), pp. 1418-1439.

Maughan, T. S., Adams, R. A., Smith, C. G., Meade, A. M., Seymour, M. T., Wilson, R. H., Idziaszczyk, S., Harris, R., Fisher, D., Kenny, S. L., Kay, E., Mitchell, J. K., Madi, A., Jasani, B., James, M. D., Bridgewater, J., Kennedy, M. J., Claes, B., Lambrechts, D., Kaplan, R. and Cheadle, J. P. (2011) 'Addition of cetuximab to oxaliplatin-based first-line combination chemotherapy for treatment of advanced colorectal cancer: results of the randomised phase 3 MRC COIN trial', *Lancet,* 377(9783), pp. 2103-14.

McInnes, L., Healy, J. and Melville, J. (2018) 'Umap: Uniform manifold approximation and projection for dimension reduction', *arXiv preprint arXiv:1802.03426*.

Mielgo, A. and Schmid, M. C. (2020) 'Liver Tropism in Cancer: The Hepatic Metastatic Niche', *Cold Spring Harb Perspect Med,* 10(3).

Moncada, R., Barkley, D., Wagner, F., Chiodin, M., Devlin, J. C., Baron, M., Hajdu, C. H., Simeone, D. M. and Yanai, I. (2020) 'Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas', *Nature Biotechnology,* 38(3), pp. 333-342.

Monestime, S. and Lazaridis, D. (2020) 'Pexidartinib (TURALIO™): The First FDA-Indicated Systemic Treatment for Tenosynovial Giant Cell Tumor', *Drugs R D,* 20(3), pp. 189-195.

Moretto, R., Germani, M. M., Borelli, B., Conca, V., Rossini, D., Boraschi, P., Donati, F., Urbani, L., Lonardi, S., Bergamo, F., Cerma, K., Ramondo, G., D'Amico, F. E., Salvatore, L., Valente, G.,

Barbaro, B., Giuliante, F., Di Maio, M., Masi, G. and Cremolini, C. (2024) 'Predicting early recurrence after resection of initially unresectable colorectal liver metastases: the role of baseline and pre-surgery clinical, radiological and molecular factors in a real-life multicentre experience', *ESMO Open,* 9(4).

Muralidharan, V., Nguyen, L., Banting, J. and Christophi, C. (2014) 'The Prognostic Significance of Lymphatics in Colorectal Liver Metastases', *HPB Surgery,* 2014(1), pp. 954604.

Neal, J. T., Li, X., Zhu, J., Giangarra, V., Grzeskowiak, C. L., Ju, J., Liu, I. H., Chiou, S.-H., Salahudeen, A. A., Smith, A. R., Deutsch, B. C., Liao, L., Zemek, A. J., Zhao, F., Karlsson, K., Schultz, L. M., Metzner, T. J., Nadauld, L. D., Tseng, Y.-Y., Alkhairy, S., Oh, C., Keskula, P., Mendoza-Villanueva, D., De La Vega, F. M., Kunz, P. L., Liao, J. C., Leppert, J. T., Sunwoo, J. B., Sabatti, C., Boehm, J. S., Hahn, W. C., Zheng, G. X. Y., Davis, M. M. and Kuo, C. J. (2018) 'Organoid Modeling of the Tumor Immune Microenvironment', *Cell,* 175(7), pp. 1972-1988.e16.

Nirmal, A. J. and Sorger, P. K. (2024) 'SCIMAP: A Python Toolkit for Integrated Spatial Analysis of Multiplexed Imaging Data', *J Open Source Softw,* 9(97).

Nirmal, A. J., Maliga, Z., Vallius, T., Quattrochi, B., Chen, A. A., Jacobson, C. A., Pelletier, R. J., Yapp, C., Arias-Camison, R., Chen, Y. A., Lian, C. G., Murphy, G. F., Santagata, S. and Sorger, P. K. (2022) 'The Spatial Landscape of Progression and Immunoediting in Primary Melanoma at Single-Cell Resolution', *Cancer Discov,* 12(6), pp. 1518-1541.

Nordlinger, B., Sorbye, H., Glimelius, B., Poston, G. J., Schlag, P. M., Rougier, P., Bechstein, W. O., Primrose, J. N., Walpole, E. T., Finch-Jones, M., Jaeck, D., Mirza, D., Parks, R. W., Mauer, M., Tanis, E., Van Cutsem, E., Scheithauer, W. and Gruenberger, T. (2013) 'Perioperative FOLFOX4 chemotherapy and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC 40983): long-term results of a randomised, controlled, phase 3 trial', *Lancet Oncol,* 14(12), pp. 1208-15.

O'Connell, B. C., Hubbard, C., Zizlsperger, N., Fitzgerald, D., Kutok, J. L., Varner, J., Ilaria, R., Cobleigh, M. A., Juric, D., Tkaczuk, K. H. R., Elias, A., Lee, A., Dakhil, S., Hamilton, E., Soliman, H. and Peluso, S. (2024) 'Eganelisib combined with immune checkpoint inhibitor therapy and chemotherapy in frontline metastatic triple-negative breast cancer triggers macrophage reprogramming, immune activation and extracellular matrix reorganization in the tumor microenvironment', *Journal for ImmunoTherapy of Cancer,* 12(8), pp. e009160.

Okines, A., Puerto, O. D., Cunningham, D., Chau, I., Van Cutsem, E., Saltz, L. and Cassidy, J. (2009) 'Surgery with curative-intent in patients treated with first-line chemotherapy plus bevacizumab for metastatic colorectal cancer First BEAT and the randomised phase-III NO16966 trial', *Br J Cancer,* 101(7), pp. 1033-8.

Otter, N., Porter, M. A., Tillmann, U., Grindrod, P. and Harrington, H. A. (2017) 'A roadmap for the computation of persistent homology', *EPJ Data Science,* 6(1), pp. 17.

Padrón, L. J., Maurer, D. M., O'Hara, M. H., O'Reilly, E. M., Wolff, R. A., Wainberg, Z. A., Ko, A. H., Fisher, G., Rahma, O., Lyman, J. P., Cabanski, C. R., Yu, J. X., Pfeiffer, S. M., Spasic, M., Xu, J., Gherardini, P. F., Karakunnel, J., Mick, R., Alanio, C., Byrne, K. T., Hollmann, T. J., Moore, J. S., Jones, D. D., Tognetti, M., Chen, R. O., Yang, X., Salvador, L., Wherry, E. J., Dugan, U., O'Donnell-Tormey, J., Butterfield, L. H., Hubbard-Lucey, V. M., Ibrahim, R., Fairchild, J., Bucktrout, S., LaVallee, T. M. and Vonderheide, R. H. (2022) 'Sotigalimab and/or nivolumab with chemotherapy in first-line metastatic pancreatic cancer: clinical and immunologic analyses from the randomized phase 2 PRINCE trial', *Nature Medicine,* 28(6), pp. 1167-1177.

Paget, S. (1989) 'The distribution of secondary growths in cancer of the breast. 1889', *Cancer Metastasis Rev,* 8(2), pp. 98-101.

Pan, Y., Yu, Y., Wang, X. and Zhang, T. (2020) 'Tumor-Associated Macrophages in Tumor Immunity', *Front Immunol,* 11, pp. 583084.

Park, S., Ock, C. Y., Kim, H., Pereira, S., Park, S., Ma, M., Choi, S., Kim, S., Shin, S., Aum, B. J., Paeng, K., Yoo, D., Cha, H., Park, S., Suh, K. J., Jung, H. A., Kim, S. H., Kim, Y. J., Sun, J. M., Chung, J. H., Ahn, J. S., Ahn, M. J., Lee, J. S., Park, K., Song, S. Y., Bang, Y. J., Choi, Y. L., Mok, T. S. and Lee, S. H. (2022) 'Artificial Intelligence-Powered Spatial Analysis of Tumor-Infiltrating Lymphocytes as Complementary Biomarker for Immune Checkpoint Inhibition in Non-Small-Cell Lung Cancer', *J Clin Oncol,* 40(17), pp. 1916-1928.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R. and Dubourg, V. (2011) 'Scikit-learn: Machine learning in Python', *the Journal of machine Learning research,* 12, pp. 2825–2830.

Pelka, K., Hofree, M., Chen, J. H., Sarkizova, S., Pirl, J. D., Jorgji, V., Bejnood, A., Dionne, D., Ge, W. H., Xu, K. H., Chao, S. X., Zollinger, D. R., Lieb, D. J., Reeves, J. W., Fuhrman, C. A., Hoang, M. L., Delorey, T., Nguyen, L. T., Waldman, J., Klapholz, M., Wakiro, I., Cohen, O., Albers, J., Smillie, C. S., Cuoco, M. S., Wu, J., Su, M. J., Yeung, J., Vijaykumar, B., Magnuson, A. M., Asinovski, N., Moll, T., Goder-Reiser, M. N., Applebaum, A. S., Brais, L. K., DelloStritto, L. K., Denning, S. L., Phillips, S. T., Hill, E. K., Meehan, J. K., Frederick, D. T., Sharova, T., Kanodia, A., Todres, E. Z., Jané-Valbuena, J., Biton, M., Izar, B., Lambden, C. D., Clancy, T. E., Bleday, R., Melnitchouk, N., Irani, J., Kunitake, H., Berger, D. L., Srivastava, A., Hornick, J. L., Ogino, S., Rotem, A., Vigneau, S., Johnson, B. E., Corcoran, R. B., Sharpe, A. H., Kuchroo, V. K., Ng, K., Giannakis, M., Nieman, L. T., Boland, G. M., Aguirre, A. J., Anderson, A. C., Rozenblatt-Rosen, O., Regev, A. and Hacohen, N. (2021) 'Spatially organized multicellular immune hubs in human colorectal cancer', *Cell,* 184(18), pp. 4734–4752.e20.

Pimenta, E. M. and Barnes, B. J. (2014) 'Role of Tertiary Lymphoid Structures (TLS) in Anti-Tumor Immunity: Potential Tumor-Induced Cytokines/Chemokines that Regulate TLS Formation in Epithelial-Derived Cancers', *Cancers (Basel),* 6(2), pp. 969-97.

Prenen, H., Deva, S., Keam, B., Lindsay, C. R., Lugowska, I., Yang, J. C., Longo, F., de Miguel, M., Ponz-Sarvise, M., Ahn, M. J., Gumus, M., Champiat, S., Italiano, A., Salas, S., Perets, R., Arslan, C., Cho, B. C., Evers, S., Boetsch, C., Marbach, D., Dejardin, D., Sleiman, N., Ardeshir, C., Richard, M., Charo, J., Kraxner, A., Keshelava, N., Teichgräber, V. and Moreno, V. (2024) 'Phase II Study to Determine the Antitumor Activity and Safety of Simlukafusp Alfa (FAP-IL2v) Combined with Atezolizumab in Esophageal Cancer', *Clin Cancer Res,* 30(14), pp. 2945-2953.

Qi, J., Sun, H., Zhang, Y., Wang, Z., Xun, Z., Li, Z., Ding, X., Bao, R., Hong, L., Jia, W., Fang, F., Liu, H., Chen, L., Zhong, J., Zou, D., Liu, L., Han, L., Ginhoux, F., Liu, Y., Ye, Y. and Su, B. (2022) 'Single-cell and spatial analysis reveal interaction of FAP+ fibroblasts and SPP1+ macrophages in colorectal cancer', *Nature Communications,* 13(1), pp. 1742.

Qiao, T., Yang, W., He, X., Song, P., Chen, X., Liu, R., Xiao, J., Yang, X., Li, M., Gao, Y., Chen, G., Lu, Y., Zhang, J., Leng, J. and Ren, H. (2023) 'Dynamic differentiation of F4/80+ tumor-associated macrophage and its role in tumor vascularization in a syngeneic mouse model of colorectal liver metastasis', *Cell Death Dis,* 14(2), pp. 117.

Qiu, X., Zhu, D. Y., Lu, Y., Yao, J., Jing, Z., Min, K. H., Cheng, M., Pan, H., Zuo, L., King, S., Fang, Q., Zheng, H., Wang, M., Wang, S., Zhang, Q., Yu, S., Liao, S., Liu, C., Wu, X., Lai, Y., Hao, S., Zhang, Z., Wu, L., Zhang, Y., Li, M., Tu, Z., Lin, J., Yang, Z., Li, Y., Gu, Y., Ellison, D., Ryan Lu, Y., Hu, Q., Hu, Y., Chen, A., Liu, L., Weissman, J. S., Ma, J., Xu, X., Liu, S. and Bai, Y. (2024) 'Spatiotemporal modeling of molecular holograms', *Cell,* 187(26), pp. 7351-7373.e61.

Rao, H. L., Chen, J. W., Li, M., Xiao, Y. B., Fu, J., Zeng, Y. X., Cai, M. Y. and Xie, D. (2012) 'Increased intratumoral neutrophil in colorectal carcinomas correlates closely with malignant phenotype and predicts patients' adverse prognosis', *PLoS One,* 7(1), pp. e30806.

Rayes, R. F., Mouhanna, J. G., Nicolau, I., Bourdeau, F., Giannias, B., Rousseau, S., Quail, D., Walsh, L., Sangwan, V., Bertos, N., Cools-Lartigue, J., Ferri, L. E. and Spicer, J. D. (2019) 'Primary tumors induce neutrophil extracellular traps with targetable metastasis promoting effects', *JCI Insight,* 5(16).

Rodriguez, A. B., Peske, J. D., Woods, A. N., Leick, K. M., Mauldin, I. S., Meneveau, M. O., Young, S. J., Lindsay, R. S., Melssen, M. M., Cyranowski, S., Parriott, G., Conaway, M. R., Fu, Y. X., Slingluff, C. L., Jr. and Engelhard, V. H. (2021) 'Immune mechanisms orchestrate tertiary lymphoid structures in tumors via cancer-associated fibroblasts', *Cell Rep,* 36(3), pp. 109422.

Ronca, V., Gerussi, A., Collins, P., Parente, A., Oo, Y. H. and Invernizzi, P. (2025) 'The liver as a central “hub” of the immune system: pathophysiological implications', *Physiological Reviews,* 105(2), pp. 493-539.

Sahai, E., Astsaturov, I., Cukierman, E., DeNardo, D. G., Egeblad, M., Evans, R. M., Fearon, D., Greten, F. R., Hingorani, S. R., Hunter, T., Hynes, R. O., Jain, R. K., Janowitz, T., Jorgensen, C., Kimmelman, A. C., Kolonin, M. G., Maki, R. G., Powers, R. S., Puré, E., Ramirez, D. C., Scherz-Shouval, R., Sherman, M. H., Stewart, S., Tlsty, T. D., Tuveson, D. A., Watt, F. M., Weaver, V., Weeraratna, A. T. and Werb, Z. (2020) 'A framework for advancing our understanding of cancer-associated fibroblasts', *Nature Reviews Cancer,* 20(3), pp. 174-186.

Salachan, P. V., Rasmussen, M., Ulhøi, B. P., Jensen, J. B., Borre, M. and Sørensen, K. D. (2023) 'Spatial whole transcriptome profiling of primary tumor from patients with metastatic prostate cancer', *Int J Cancer,* 153(12), pp. 2055-2067.

Sathe, A., Mason, K., Grimes, S. M., Zhou, Z., Lau, B. T., Bai, X., Su, A., Tan, X., Lee, H., Suarez, C. J., Nguyen, Q., Poultsides, G., Zhang, N. R. and Ji, H. P. (2023) 'Colorectal Cancer Metastases in the Liver Establish Immunosuppressive Spatial Networking between Tumor-Associated SPP1+ Macrophages and Fibroblasts', *Clin Cancer Res,* 29(1), pp. 244-260.

Sautès-Fridman, C., Petitprez, F., Calderaro, J. and Fridman, W. H. (2019) 'Tertiary lymphoid structures in the era of cancer immunotherapy', *Nat Rev Cancer,* 19(6), pp. 307-325.

Schoenberger, S. P., Toes, R. E., van der Voort, E. I., Offringa, R. and Melief, C. J. (1998) 'T-cell help for cytotoxic T lymphocytes is mediated by CD40-CD40L interactions', *Nature,* 393(6684), pp. 480-3.

Schueth, A., Hildebrand, S., Samarska, I., Sengupta, S., Kiessling, A., Herrler, A., zur Hausen, A., Capalbo, M. and Roebroeck, A. (2023) 'Efficient 3D light-sheet imaging of very large-scale optically cleared human brain and prostate tissue samples', *Communications Biology,* 6(1), pp. 170.

Schürch, C. M., Bhate, S. S., Barlow, G. L., Phillips, D. J., Noti, L., Zlobec, I., Chu, P., Black, S., Demeter, J., McIlwain, D. R., Kinoshita, S., Samusik, N., Goltsev, Y. and Nolan, G. P. (2020) 'Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front', *Cell,* 182(5), pp. 1341-1359.e19.

Shamai, G., Livne, A., Polónia, A., Sabo, E., Cretu, A., Bar-Sela, G. and Kimmel, R. (2022) 'Deep learning-based image analysis predicts PD-L1 status from H&E-stained histopathology images in breast cancer', *Nat Commun,* 13(1), pp. 6753.

Shasha, T., Gruijs, M. and van Egmond, M. (2022) 'Mechanisms of colorectal liver metastasis development', *Cellular and Molecular Life Sciences,* 79(12), pp. 607.

Shu, H., Chen, J., Hu, J., Wang, Y., Peng, J., Shang, X. and Wang, T. (2024) 'Efficient integration of multiple spatial transcriptomics data for 3D domain detection, matching, and alignment with stMSA', *bioRxiv*, pp. 2024.07.29.605604.

Sojka, D. K., Huang, Y. H. and Fowell, D. J. (2008) 'Mechanisms of regulatory T-cell suppression - a diverse arsenal for a moving target', *Immunology,* 124(1), pp. 13-22.

Stachtea, X., Loughrey, M. B., Salvucci, M., Lindner, A. U., Cho, S., McDonough, E., Sood, A., Graf, J., Santamaria-Pang, A., Corwin, A., Laurent-Puig, P., Dasgupta, S., Shia, J., Owens, J. R., Abate, S., Van Schaeybroeck, S., Lawler, M., Prehn, J. H. M., Ginty, F. and Longley, D. B. (2022) 'Stratification of chemotherapy-treated stage III colorectal cancer patients using multiplexed imaging and single-cell analysis of T-cell populations', *Modern Pathology,* 35(4), pp. 564-576.

Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J. O., Huss, M., Mollbrink, A., Linnarsson, S., Codeluppi, S., Borg, Å., Pontén, F., Costea, P. I., Sahlén, P., Mulder, J., Bergmann, O., Lundeberg, J. and Frisén, J. (2016) 'Visualization and analysis of gene expression in tissue sections by spatial transcriptomics', *Science,* 353(6294), pp. 78-82.

Steeghs, N., Gomez-Roca, C., Rohrberg, K. S., Mau-Sørensen, M., Robbrecht, D., Tabernero, J., Ahmed, S., Rodríguez-Ruiz, M. E., Ardeshir, C., Schmid, D., Sleiman, N., Watson, C., Piper-Lepoutre, H., Dejardin, D., Evers, S., Boetsch, C., Charo, J., Teichgräber, V. and Melero, I. (2024) 'Safety, Pharmacokinetics, Pharmacodynamics, and Antitumor Activity from a Phase I Study of Simlukafusp Alfa (FAP-IL2v) in Advanced/Metastatic Solid Tumors', *Clinical Cancer Research,* 30(13), pp. 2693-2701.

Su, T., Yang, Y., Lai, S., Jeong, J., Jung, Y., McConnell, M., Utsumi, T. and Iwakiri, Y. (2021) 'Single-Cell Transcriptomics Reveals Zone-Specific Alterations of Liver Sinusoidal Endothelial Cells in Cirrhosis', *Cell Mol Gastroenterol Hepatol,* 11(4), pp. 1139-1161.

Suarez-Carmona, M., Williams, A., Schreiber, J., Hohmann, N., Pruefer, U., Krauss, J., Jäger, D., Frömming, A., Beyer, D., Eulberg, D., Jungelius, J. U., Baumann, M., Mangasarian, A. and Halama, N. (2021) 'Combined inhibition of CXCL12 and PD-1 in MSS colorectal and pancreatic cancer: modulation of the microenvironment and clinical effects', *J Immunother Cancer,* 9(10).

Sun, L., Clavijo, P. E., Robbins, Y., Patel, P., Friedman, J., Greene, S., Das, R., Silvin, C., Van Waes, C., Horn, L. A., Schlom, J., Palena, C., Maeda, D., Zebala, J. and Allen, C. T. (2019) 'Inhibiting myeloid-derived suppressor cell trafficking enhances T cell immunotherapy', *JCI Insight,* 4(7).

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A. and Bray, F. (2021) 'Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries', *CA Cancer J Clin,* 71(3), pp. 209-249.

Svensson-Arvelund, J., Cuadrado-Castano, S., Pantsulaia, G., Kim, K., Aleynick, M., Hammerich, L., Upadhyay, R., Yellin, M., Marsh, H., Oreper, D., Jhunjhunwala, S., Moussion, C., Merad, M., Brown, B. D., García-Sastre, A. and Brody, J. D. (2022) 'Expanding cross-presenting dendritic cells enhances oncolytic virotherapy and is critical for long-term anti-tumor immunity', *Nature Communications,* 13(1), pp. 7149.

Tamiya, A., Yamazaki, K., Boku, N., Machida, N., Kojima, T., Taku, K., Yasui, H., Fukutomi, A., Hironaka, S. and Onozawa, Y. (2009) 'Safety of bevacizumab treatment in combination with standard chemotherapy for metastatic colorectal cancer: a retrospective review of 65 Japanese patients', *Int J Clin Oncol,* 14(6), pp. 513-7.

Tauriello, D. V. F., Palomo-Ponce, S., Stork, D., Berenguer-Llergo, A., Badia-Ramentol, J., Iglesias, M., Sevillano, M., Ibiza, S., Cañellas, A., Hernando-Momblona, X., Byrom, D., Matarin, J. A., Calon, A., Rivas, E. I., Nebreda, A. R., Riera, A., Attolini, C. S.-O. and Batlle, E. (2018) 'TGFβ drives immune evasion in genetically reconstituted colon cancer metastasis', *Nature,* 554(7693), pp. 538-543.

Taylor, A., Verhagen, J., Blaser, K., Akdis, M. and Akdis, C. A. (2006) 'Mechanisms of immune suppression by interleukin-10 and transforming growth factor-beta: the role of T regulatory cells', *Immunology,* 117(4), pp. 433-42.
Thomson, A. W. and Knolle, P. A. (2010) 'Antigen-presenting cell function in the tolerogenic liver environment', *Nature Reviews Immunology,* 10(11), pp. 753-766.

Tiegs, G. and Lohse, A. W. (2010) 'Immune tolerance: What is unique about the liver', *Journal of Autoimmunity,* 34(1), pp. 1-6.

Thorel, L., Perréard, M., Florent, R., Divoux, J., Coffy, S., Vincent, A., Gaggioli, C., Guasch, G., Gidrol, X., Weiswald, L.-B. and Poulain, L. (2024) 'Patient-derived tumor organoids: a new avenue for preclinical research and precision medicine in oncology', *Experimental & Molecular Medicine,* 56(7), pp. 1531-1551.

Tsilimigras, D. I., Brodt, P., Clavien, P. A., Muschel, R. J., D'Angelica, M. I., Endo, I., Parks, R. W., Doyle, M., de Santibañes, E. and Pawlik, T. M. (2021) 'Liver metastases', *Nat Rev Dis Primers,* 7(1), pp. 27.

Tveit, K. M., Guren, T., Glimelius, B., Pfeiffer, P., Sorbye, H., Pyrhonen, S., Sigurdsson, F., Kure, E., Ikdahl, T., Skovlund, E., Fokstuen, T., Hansen, F., Hofsli, E., Birkemeyer, E., Johnsson, A., Starkhammar, H., Yilmaz, M. K., Keldsen, N., Erdal, A. B., Dajani, O., Dahl, O. and Christoffersen,

T. (2012) 'Phase III trial of cetuximab with continuous or intermittent fluorouracil, leucovorin, and oxaliplatin (Nordic FLOX) versus FLOX alone in first-line treatment of metastatic colorectal cancer: the NORDIC-VII study', *J Clin Oncol,* 30(15), pp. 1755-62.

Ülger, B. V., Hatipoğlu, E. S., Ertuğrul, Ö., Tuncer, M. C., Özmen, C. A. and Gül, M. (2018) 'Variations in the vascular and biliary structures of the liver: a comprehensive anatomical study', *Acta Chirurgica Belgica,* 118(6), pp. 354-371.

Vadisetti, S. N., Kazi, M., Patkar, S., Mundhada, R., Desouza, A., Saklani, A. and Goel, M. (2024) 'Patterns and Predictors of Recurrence After Curative Resection of Colorectal Liver Metastasis (CRLM)', *J Gastrointest Cancer,* 55(4), pp. 1559-1568.

Van Cutsem, E., Köhne, C. H., Láng, I., Folprecht, G., Nowacki, M. P., Cascinu, S., Shchepotin, I., Maurel, J., Cunningham, D., Tejpar, S., Schlichting, M., Zubel, A., Celik, I., Rougier, P. and Ciardiello, F. (2011) 'Cetuximab plus irinotecan, fluorouracil, and leucovorin as first-line treatment for metastatic colorectal cancer: updated analysis of overall survival according to tumor KRAS and BRAF mutation status', *J Clin Oncol,* 29(15), pp. 2011-9.

Van Cutsem, E., Nordlinger, B. and Cervantes, A. (2010) 'Advanced colorectal cancer: ESMO Clinical Practice Guidelines for treatment', *Ann Oncol,* 21 Suppl 5, pp. v93-7.

van Dam, P.-J., van der Stok, E. P., Teuwen, L.-A., Van den Eynden, G. G., Illemann, M., Frentzas, S., Majeed, A. W., Eefsen, R. L., Coebergh van den Braak, R. R. J., Lazaris, A., Fernandez, M. C., Galjart, B., Laerum, O. D., Rayes, R., Grünhagen, D. J., Van de paer, M., Sucaet, Y., Mudhar, H. S., Schvimer, M., Nyström, H., Kockx, M., Bird, N. C., Vidal-Vanaclocha, F., Metrakos, P., Simoneau, E., Verhoef, C., Dirix, L. Y., Van Laere, S., Gao, Z.-h., Brodt, P., Reynolds, A. R. and Vermeulen, P. B. (2017) 'International consensus guidelines for scoring the histopathological growth patterns of liver metastasis', *British Journal of Cancer,* 117(10), pp. 1427-1441.

van der Geest, L. G., Lam-Boer, J., Koopman, M., Verhoef, C., Elferink, M. A. and de Wilt, J. H. (2015) 'Nationwide trends in incidence, treatment and survival of colorectal cancer patients with synchronous metastases', *Clin Exp Metastasis,* 32(5), pp. 457-65.

Varoquaux, G., Raamana, P. R., Engemann, D. A., Hoyos-Idrobo, A., Schwartz, Y. and Thirion, B. (2017) 'Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines', *NeuroImage,* 145, pp. 166-179.

Vidal-Vanaclocha, F. (2008) 'The prometastatic microenvironment of the liver', *Cancer Microenviron,* 1(1), pp. 113-29.

Viel, S., Marçais, A., Guimaraes, F. S., Loftus, R., Rabilloud, J., Grau, M., Degouve, S., Djebali, S., Sanlaville, A., Charrier, E., Bienvenu, J., Marie, J. C., Caux, C., Marvel, J., Town, L., Huntington, N.

D., Bartholin, L., Finlay, D., Smyth, M. J. and Walzer, T. (2016) 'TGF-β inhibits the activation and functions of NK cells by repressing the mTOR pathway', *Sci Signal,* 9(415), pp. ra19.

Virtanen, P. and Gommers, R. and Oliphant, T. E. and Haberland, M. and Reddy, T. and Cournapeau, D. and Burovski, E. and Peterson, P. and Weckesser, W. and Bright, J. and van der Walt, S. J. and Brett, M. and Wilson, J. and Millman, K. J. and Mayorov, N. and Nelson, A. R. J. and Jones, E. and Kern, R. and Larson, E. and Carey, C. J. and Polat, İ. and Feng, Y. and Moore, E. W. and VanderPlas, J. and Laxalde, D. and Perktold, J. and Cimrman, R. and Henriksen, I. and Quintero, E. A. and Harris, C. R. and Archibald, A. M. and Ribeiro, A. H. and Pedregosa, F. and van Mulbregt, P. and Vijaykumar, A. and Bardelli, A. P. and Rothberg, A. and Hilboll, A. and Kloeckner, A. and Scopatz, A. and Lee, A. and Rokem, A. and Woods, C. N. and Fulton, C. and Masson, C. and Häggström, C. and Fitzgerald, C. and Nicholson, D. A. and Hagen, D. R. and Pasechnik, D. V. and Olivetti, E. and Martin, E. and Wieser, E. and Silva, F. and Lenders, F. and Wilhelm, F. and Young, G. and Price, G. A. and Ingold, G.-L. and Allen, G. E. and Lee, G. R. and Audren, H. and Probst, I. and Dietrich, J. P. and Silterra, J. and Webber, J. T. and Slavič, J. and Nothman, J. and Buchner, J. and Kulick, J. and Schönberger, J. L. and de Miranda Cardoso, J. V. and Reimer, J. and Harrington, J. and Rodríguez, J. L. C. and Nunez-Iglesias, J. and Kuczynski, J. and Tritz, K. and Thoma, M. and Newville, M. and Kümmerer, M. and Bolingbroke, M. and Tartre, M. and Pak, M. and Smith, N. J. and Nowaczyk, N. and Shebanov, N. and Pavlyk, O. and Brodtkorb, P. A. and Lee, P. and McGibbon, R. T. and Feldbauer, R. and Lewis, S. and Tygier, S. and Sievert, S. and Vigna, S. and Peterson, S. and More, S. and Pudlik, T. and Oshima, T. and Pingel, T. J. and Robitaille, T. P. and Spura, T. and Jones, T. R. and Cera, T. and Leslie, T. and Zito, T. and Krauss, T. and Upadhyay, U. and Halchenko, Y. O. and Vázquez-Baeza, Y. and SciPy, C. (2020) 'SciPy 1.0: fundamental algorithms for scientific computing in Python', *Nature Methods,* 17(3), pp. 261–272.

Waldhauer, I., Gonzalez-Nicolini, V., Freimoser-Grundschober, A., Nayak, T. K., Fahrni, L., Hosse, R. J., Gerrits, D., Geven, E. J. W., Sam, J., Lang, S., Bommer, E., Steinhart, V., Husar, E., Colombetti, S., Van Puijenbroek, E., Neubauer, M., Cline, J. M., Garg, P. K., Dugan, G., Cavallo, F., Acuna, G., Charo, J., Teichgräber, V., Evers, S., Boerman, O. C., Bacac, M., Moessner, E., Umaña, P. and Klein, C. (2021) 'Simlukafusp alfa (FAP-IL2v) immunocytokine is a versatile combination partner for cancer immunotherapy', *MAbs,* 13(1), pp. 1913791.

Wang, S., Zheng, H., Choi, J. S., Lee, J. K., Li, X. and Hu, H. (2022) 'A systematic evaluation of the computational tools for ligand-receptor-based cell–cell interaction inference', *Briefings in Functional Genomics,* 21(5), pp. 339-356.

Wang, J., Zhao, X. and Wan, Y. Y. (2023b) 'Intricacies of TGF-β signaling in Treg and Th17 cell biology', *Cellular & Molecular Immunology,* 20(9), pp. 1002-1022.

Wang, Y., Zhong, X., He, X., Hu, Z., Huang, H., Chen, J., Chen, K., Zhao, S., Wei, P. and Li, D. (2023a) 'Liver metastasis from colorectal cancer: pathogenetic development, immune landscape

of the tumour microenvironment and therapeutic approaches', *Journal of Experimental & Clinical Cancer Research,* 42(1), pp. 177.

Weiss, S. A., Sznol, M., Shaheen, M., Berciano-Guerrero, M., Couselo, E. M., Rodríguez-Abreu, D., Boni, V., Schuchter, L. M., Gonzalez-Cao, M., Arance, A., Wei, W., Ganti, A. K., Hauke, R. J., Berrocal, A., Iannotti, N. O., Hsu, F. J. and Kluger, H. M. (2024) 'A Phase II Trial of the CD40 Agonistic Antibody Sotigalimab (APX005M) in Combination with Nivolumab in Subjects with Metastatic Melanoma with Confirmed Disease Progression on Anti-PD-1 Therapy', *Clin Cancer Res,* 30(1), pp. 74-81.

Wen, J., Wang, S., Guo, R. and Liu, D. (2023) 'CSF1R inhibitors are emerging immunotherapeutic drugs for cancer treatment', *Eur J Med Chem,* 245(Pt 1), pp. 114884.

Williams, H. L., Frei, A. L., Koessler, T., Berger, M. D., Dawson, H., Michielin, O. and Zlobec, I. (2024) 'The current landscape of spatial biomarkers for prediction of response to immune checkpoint inhibition', *NPJ Precis Oncol,* 8(1), pp. 178.

Wolf, F. A., Angerer, P. and Theis, F. J. (2018) 'SCANPY: large-scale single-cell gene expression data analysis', *Genome Biol,* 19(1), pp. 15.

Wood, C. S., Pennel, K. A. F., Leslie, H., Legrini, A., Cameron, A. J., Melissourgou-Syka, L., Quinn, J. A., van Wyk, H. C., Hay, J., Roseweir, A. K., Nixon, C., Roxburgh, C. S. D., McMillan, D. C., Biankin, A. V., Sansom, O. J., Horgan, P. G., Edwards, J., Steele, C. W. and Jamieson, N. B. (2023) 'Spatially Resolved Transcriptomics Deconvolutes Prognostic Histological Subgroups in Patients with Colorectal Cancer and Synchronous Liver Metastases', *Cancer Research,* 83(8), pp. 1329-1344.

Wu, R., Ohara, R. A., Jo, S., Liu, T. T., Ferris, S. T., Ou, F., Kim, S., Theisen, D. J., Anderson, D. A., 3rd, Wong, B. W., Gershon, T., Schreiber, R. D., Murphy, T. L. and Murphy, K. M. (2022b) 'Mechanisms of CD40-dependent cDC1 licensing beyond costimulation', *Nat Immunol,* 23(11), pp. 1536-1550.

Wu, Y., Yang, S., Ma, J., Chen, Z., Song, G., Rao, D., Cheng, Y., Huang, S., Liu, Y., Jiang, S., Liu, J., Huang, X., Wang, X., Qiu, S., Xu, J., Xi, R., Bai, F., Zhou, J., Fan, J., Zhang, X. and Gao, Q. (2022a) 'Spatiotemporal Immune Landscape of Colorectal Cancer Liver Metastasis at Single-Cell Level', *Cancer Discovery,* 12(1), pp. 134-153.

Xiao, J., Yu, X., Meng, F., Zhang, Y., Zhou, W., Ren, Y., Li, J., Sun, Y., Sun, H., Chen, G., He, K. and Lu, L. (2024) 'Integrating spatial and single-cell transcriptomics reveals tumor heterogeneity and intercellular networks in colorectal cancer', *Cell Death & Disease,* 15(5), pp. 326.
Xiao, Y. and Yu, D. (2021) 'Tumor microenvironment as a therapeutic target in cancer', *Pharmacology & Therapeutics,* 221, pp. 107753.

Xiao, Z., Todd, L., Huang, L., Noguera-Ortega, E., Lu, Z., Huang, L., Kopp, M., Li, Y., Pattada, N., Zhong, W., Guo, W., Scholler, J., Liousia, M., Assenmacher, C.-A., June, C. H., Albelda, S. M. and Puré, E. (2023) 'Desmoplastic stroma restricts T cell extravasation and mediates immune exclusion and immunosuppression in solid tumors', *Nature Communications,* 14(1), pp. 5110.

Yan, M., Zheng, M., Niu, R., Yang, X., Tian, S., Fan, L., Li, Y. and Zhang, S. (2022) 'Roles of tumor-associated neutrophils in tumor metastasis and its clinical applications', *Front Cell Dev Biol,* 10, pp. 938289.

Yerke, A., Fry Brumit, D. and Fodor, A. A. (2024) 'Proportion-based normalizations outperform compositional data transformations in machine learning applications', *Microbiome,* 12(1), pp. 45.
You, Q., Cheng, L., Kedl, R. M. and Ju, C. (2008) 'Mechanism of T cell tolerance induction by murine hepatic Kupffer cells', *Hepatology,* 48(3), pp. 978-90.

Yuan, Z., Li, Y., Zhang, S., wwu, X., Dou, H., Yu, X., Zhang, Z., Yang, S. and Xiao, M. (2023) 'Extracellular matrix remodeling in tumor progression and immune escape: from mechanisms to treatments', *Mol Cancer,* 22(1), pp. 48.

Yuzhalin, A. E., Gordon-Weeks, A. N., Tognoli, M. L., Jones, K., Markelc, B., Konietzny, R., Fischer, R., Muth, A., O'Neill, E., Thompson, P. R., Venables, P. J., Kessler, B. M., Lim, S. Y. and Muschel, R. J. (2018) 'Colorectal cancer liver metastatic growth depends on PAD4-driven citrullination of the extracellular matrix', *Nature Communications,* 9(1), pp. 4783.

Zboralski, D., Hoehlig, K., Eulberg, D., Frömming, A. and Vater, A. (2017) 'Increasing Tumor-Infiltrating T Cells through Inhibition of CXCL12 with NOX-A12 Synergizes with PD-1 Blockade', *Cancer Immunology Research,* 5(11), pp. 950-956.

Zhang, H., Hong, L., Dong, Z., Xin, S., Lin, B., Cheng, J., Tian, W., Li, B., Wang, J., Liu, X., Liu, C., Jin, Y., Feng, Y., Su, G., Sun, X., Liu, Q., Dai, X., Gao, Y., Tong, Z., Liu, L., Zhu, X., Zheng, Y., Zhao, P., Guo, T., Fang, W. and Bao, X. (2025) 'Spatially resolved C1QC+ macrophage-CD4+ T cell niche in colorectal cancer microenvironment: implications for immunotherapy response', *Cell Discovery,* 11(1), pp. 60.

Zhang, Q., Abdo, R., Iosef, C., Kaneko, T., Cecchini, M., Han, V. K. and Li, S. S.-C. (2022) 'The spatial transcriptomic landscape of non-small cell lung cancer brain metastasis', *Nature Communications,* 13(1), pp. 5983.

Zhang, Y., Chen, H., Mo, H., Hu, X., Gao, R., Zhao, Y., Liu, B., Niu, L., Sun, X., Yu, X., Wang, Y., Chang, Q., Gong, T., Guan, X., Hu, T., Qian, T., Xu, B., Ma, F., Zhang, Z. and Liu, Z. (2021) 'Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer', *Cancer Cell,* 39(12), pp. 1578-1593.e8.

Zhang, Y., Liu, G., Zeng, Q., Wu, W., Lei, K., Zhang, C., Tang, M., Zhang, Y., Xiang, X., Tan, L., Cui, R., Qin, S., Song, X., Yin, C., Chen, Z. and Kuang, M. (2024) 'CCL19-producing fibroblasts promote tertiary lymphoid structure formation enhancing anti-tumor IgG response in colorectal cancer liver metastasis', *Cancer Cell,* 42(8), pp. 1370-1385.e9.

Zhao, J., Ou, B., Feng, H., Wang, P., Yin, S., Zhu, C., Wang, S., Chen, C., Zheng, M., Zong, Y., Sun, J. and Lu, A. (2017) 'Overexpression of CXCR2 predicts poor prognosis in patients with colorectal cancer', *Oncotarget,* 8(17), pp. 28442-28454.

Zheng, M. and Tian, Z. (2019) 'Liver-Mediated Adaptive Immune Tolerance', *Front Immunol,* 10, pp. 2525.

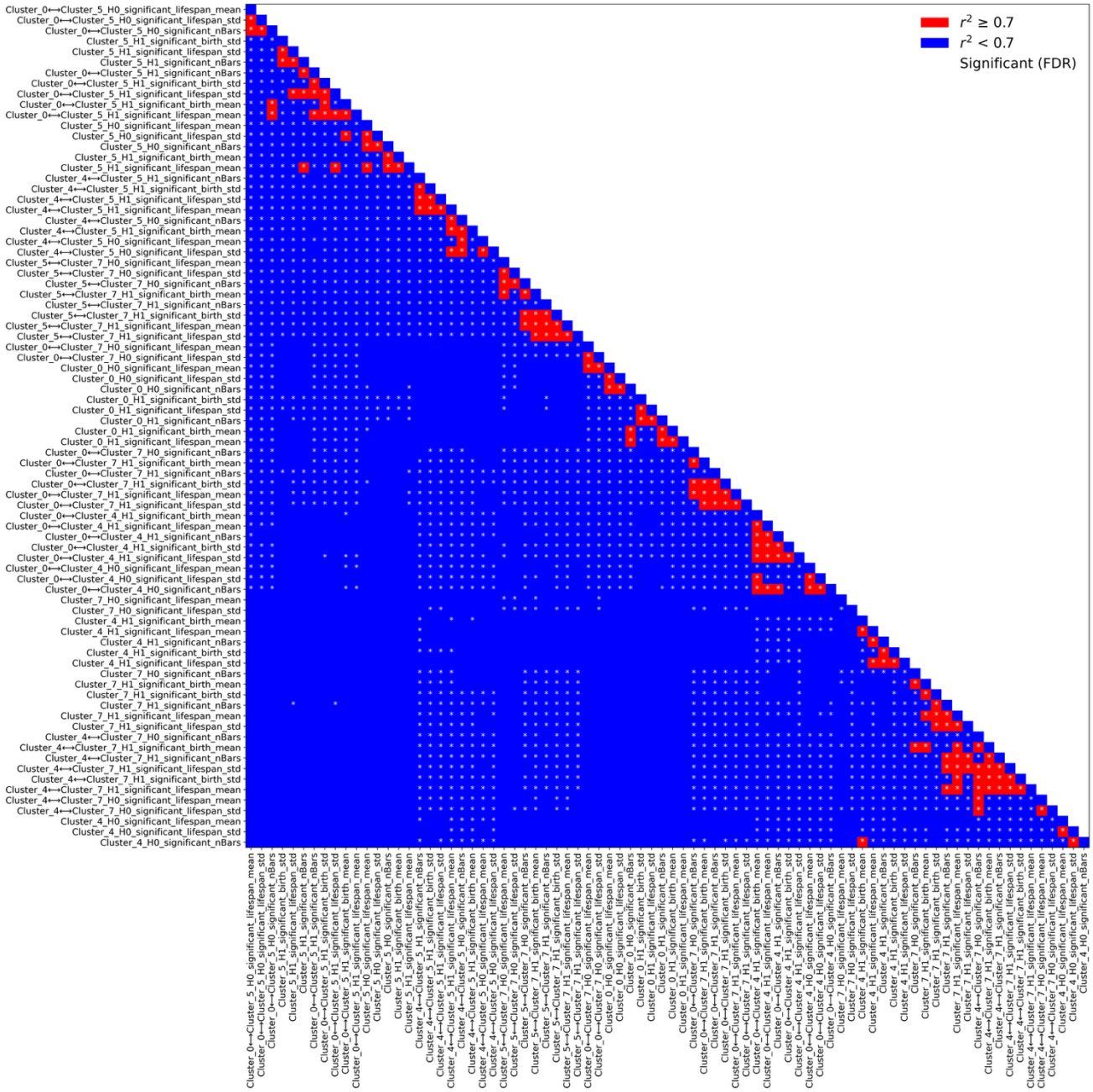
Zheng, W., Yang, L., Jiang, S., Chen, M., Li, J., Liu, Z., Wu, Z., Gong, J. and Chen, Y. (2023) 'Role of Kupffer cells in tolerance induction after liver transplantation', *Front Cell Dev Biol,* 11, pp. 1179077.

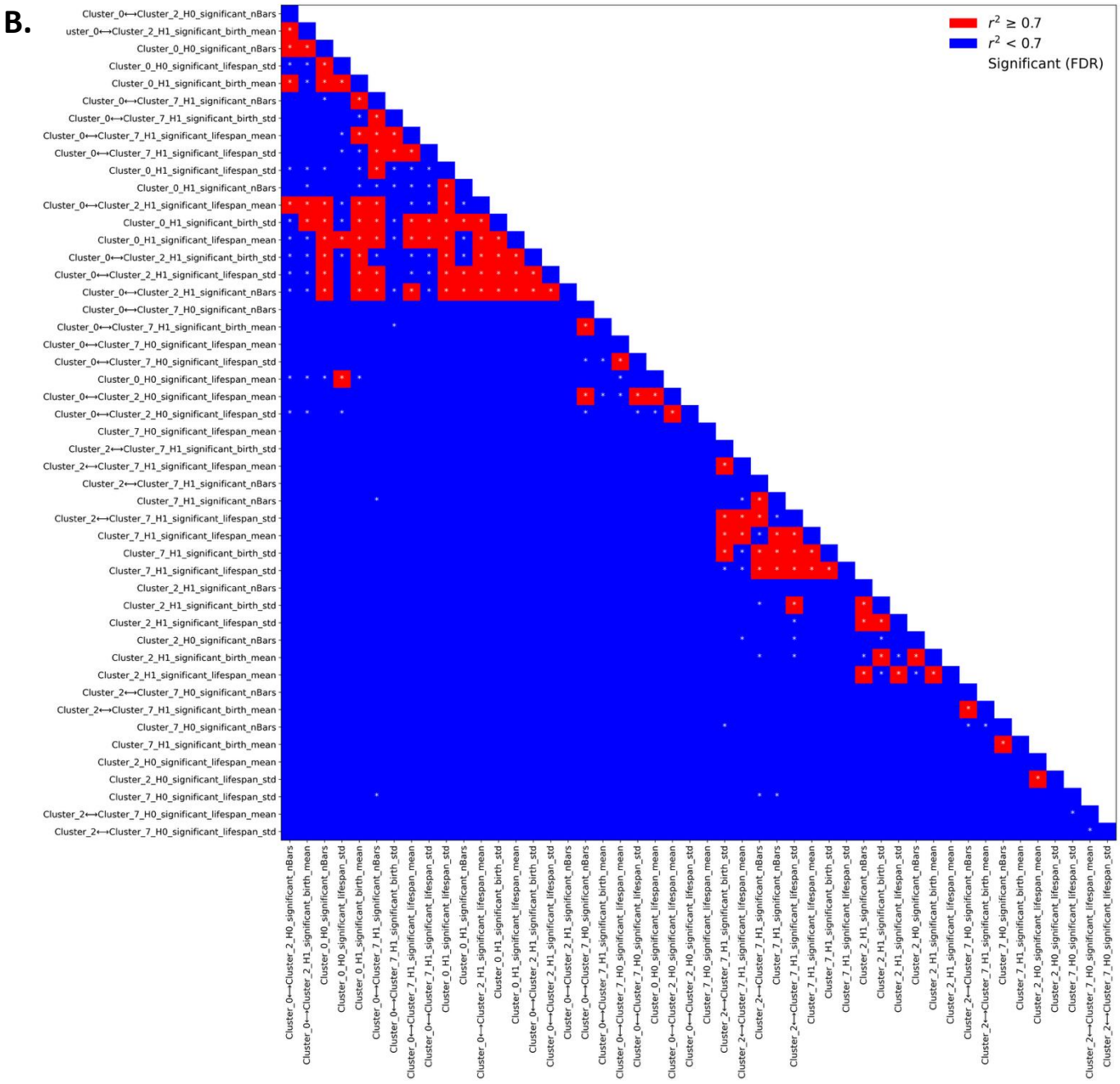
Zhou, Y., Shen, G., Zhou, X. and Li, J. (2025) 'Therapeutic potential of tumor-associated neutrophils: dual role and phenotypic plasticity', *Signal Transduction and Targeted Therapy,* 10(1), pp. 178.

Supplementary Figures

Supplementary Figure 1.

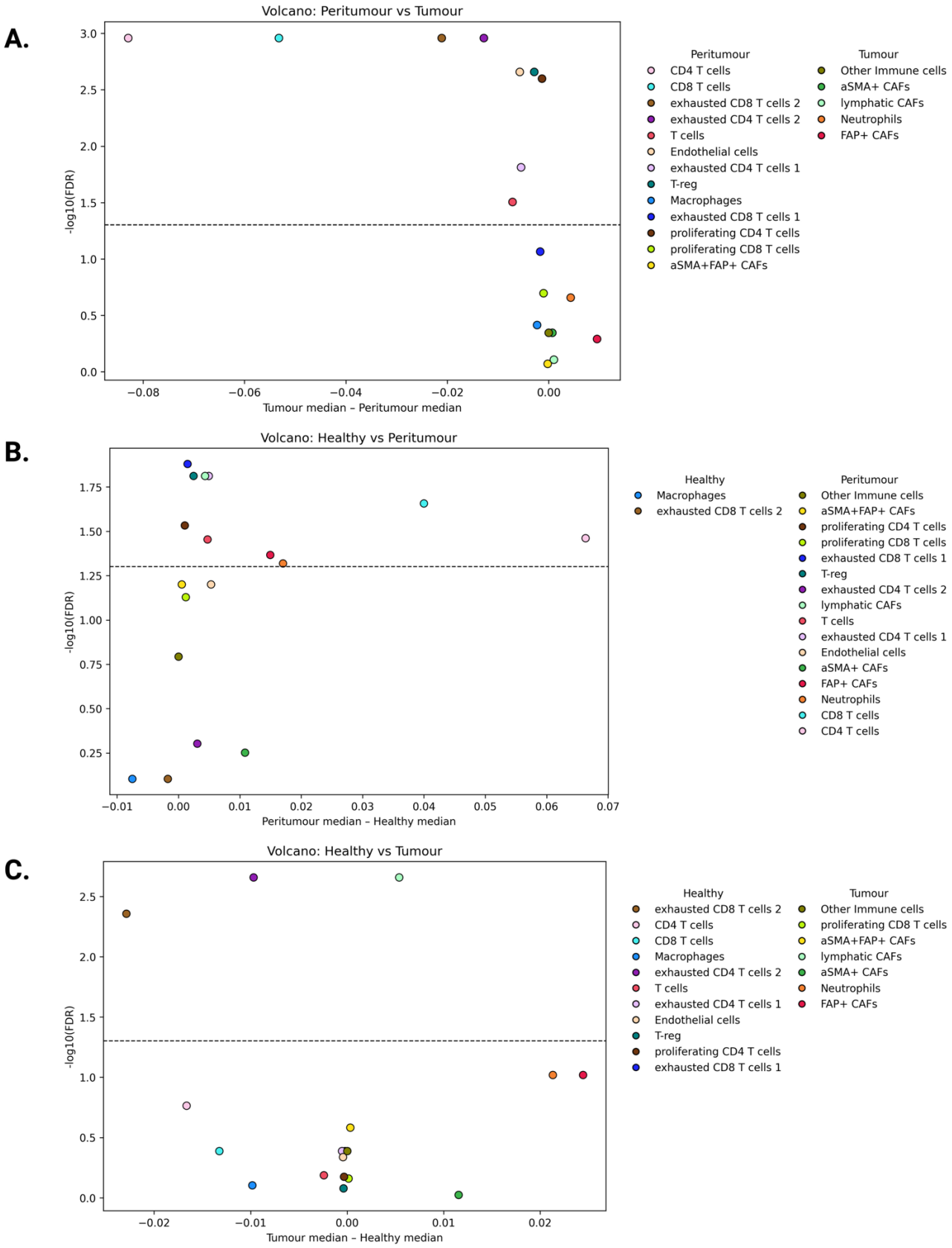
A.





Feature–feature correlation matrices for selected TDA summary features in Cell DIVE (A) and CosMx (B) datasets. Lower-triangular heatmaps displaying pairwise correlations between all slide-level, region-aggregated TDA summary features used for classification in the Cell DIVE (A) and CosMx (B) datasets, respectively. Each axis lists the selected set of significant-persistence TDA features, calculated as described in Section 2.5. Each square indicates the squared Spearman rank correlation (R^2) between a given feature pair, calculated across all slide-region combinations. Red squares: Feature pairs with strong correlation ($R^2 \geq 0.7$), considered redundant. Blue squares: Feature pairs with weaker correlation ($R^2 < 0.7$), indicating feature independence. White *, correlations with FDR-corrected p-value < 0.05.

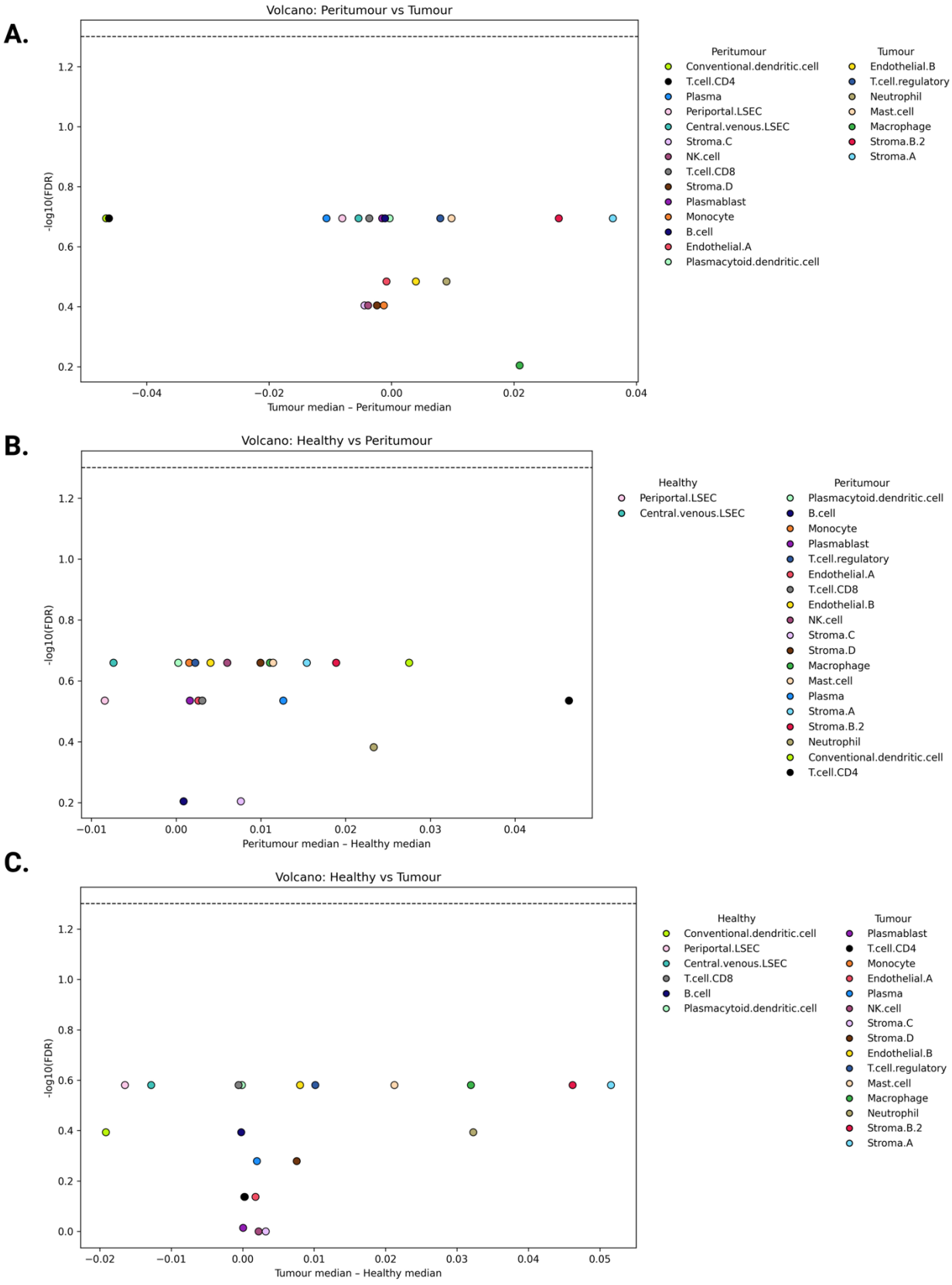
Supplementary Figure 2.



Volcano plots for region-wise differences in normalised cell-type count in the Cell DIVE cohort. A. Peritumour vs Tumour: Each point represents a targeted cell type, showing the effect size (difference in slide-level region medians, tumour minus peritumour) on the x-axis and the negative log10 of the FDR-adjusted p-value (paired Wilcoxon signed-rank test) on the y-axis. Significant cell types (above the FDR threshold, dashed line) are labelled and colour-coded by their region of

enrichment, as defined in the legend. B. Healthy vs Peritumour: Effect size is defined as peritumour median minus healthy median. C. Healthy vs Tumour: Effect size is tumour median minus healthy median.

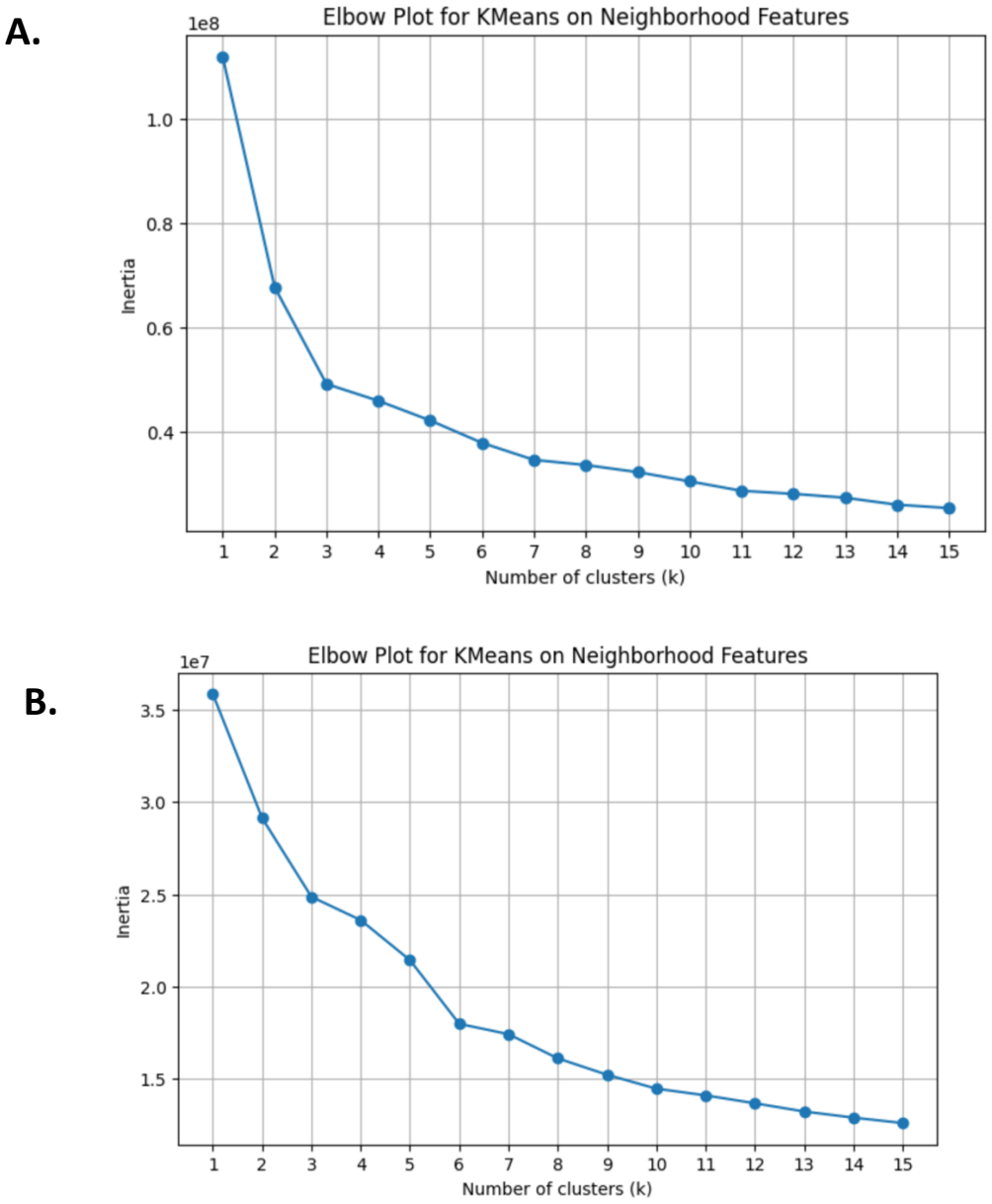
Supplementary Figure 3.



Volcano plots for region-wise differences in normalised cell-type count in the CosMx cohort. A. Peritumour vs Tumour: Each point represents a targeted cell type, showing the effect size (difference in slide-level region medians, tumour minus peritumour) on the x-axis and the negative log10 of the FDR-adjusted p-value (paired Wilcoxon signed-rank test) on the

y-axis. Significant cell types (above the FDR threshold, dashed line) are labelled and colour-coded by their region of enrichment, as defined in the legend. B. Healthy vs Peritumour: Effect size is defined as peritumour median minus healthy median. C. Healthy vs Tumour: Effect size is tumour median minus healthy median.

Supplementary Figure 4.



Elbow plot for selection of K-means cluster number (k) for cellular neighbourhood composition in Cell DIVE cohort (A) and CosMx cohort (B). The line plot shows the within-cluster sum of squared distances (inertia) as a function of the number of clusters (k), based on batch-corrected 10-nearest neighbour (10-NN) neighbourhood composition vectors for all cells. A. The inflection point at k=9 (elbow) was selected as the optimal cluster number for downstream identification of cellular neighbourhood clusters (NCs) in the Cell DIVE dataset. B. The inflection point at k=10 (elbow) was selected as the optimal cluster number for downstream identification of cellular NCs in the CosMx dataset.