



Universiteit
Leiden
The Netherlands

Towards implementation of the tumour-stroma ratio in colorectal cancer

Polack, M.

Citation

Polack, M. (2025, December 10). *Towards implementation of the tumour-stroma ratio in colorectal cancer*. Retrieved from <https://hdl.handle.net/1887/4285284>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4285284>

Note: To cite this publication please use the final published version (if applicable).



Chapter 6

Self-Supervised Learning Reveals Clinically Relevant Histomorphological Patterns for Therapeutic Strategies in Colon Cancer

Meaghan Polack+, Bojing Liu+, Nicolas Coudray#, Adalberto
Claudio Quiros#, Theodore Sakellaropoulos, Hortense Le, Afreen
Karimkhan, Augustinus S.L.P. Crobach, J. Han J.M. van Krieken,
Ke Yuan, Rob A.E.M. Tollenaar, Wilma E. Mesker, and Aristotelis
Tsirigos

+, # author with equal contribution

•

Abstract

Self-supervised learning (SSL) automates the extraction and interpretation of histopathology features on unannotated hematoxylin-eosin-stained whole slide images (WSIs). We train an SSL Barlow Twins encoder on 435 colon adenocarcinoma WSIs from The Cancer Genome Atlas to extract features from small image patches (tiles). Leiden community detection groups tiles into histomorphological phenotype clusters (HPCs). HPC reproducibility and predictive ability for overall survival is confirmed in an independent clinical trial (N=1213 WSIs). This unbiased atlas results in 47 HPCs displaying unique and shared clinically significant histomorphological traits, highlighting tissue type, quantity, and architecture, especially in the context of tumor stroma. Through in-depth analysis of these HPCs, including immune landscape and gene set enrichment analysis, and associations to clinical outcomes, we shine light on the factors influencing survival and responses to treatments of standard adjuvant chemotherapy and experimental therapies. Further exploration of HPCs may unveil additional insights and aid decision-making and personalized treatments for colon cancer patients.

Introduction

Traditionally, the diagnosis of colon cancer is confirmed by microscopic assessment of resection specimens on hematoxylin and eosin-stained (H&E) slides by pathologists. For each patient, a personalized treatment strategy is tailored through a multidisciplinary meeting, following guidelines consisting of risk assessments based on clinicopathological characteristics, including the tumor-node-metastasis (TNM) classification and additional biomarkers [1–4]. However, due to an aging population and increasing amount of biomarker research, diagnosing and predicting the prognosis of colon cancer patients can be time consuming, or complicated and resource-demanding, especially when incorporating screening for mutational variants [2,3,5,6].

In modern digital pathology, scanning H&E slides into high-resolution whole slide images (WSIs) has enabled the applications of deep learning (DL) [7]. Deep convolutional neural networks, in particular, have benefited the diagnostic process, initially by minimizing inter-rater disagreement and workload [7–9]. In colorectal cancer, supervised DL models also showed the ability to predict molecular pathways (i.e. mutation density, microsatellite instability [MSI], chromosomal instability) and key mutations like BRAF and KRAS [10,11]. DL even has intriguing potential in predicting complicated prognostic outcomes such as patient survival [12–14]. Moreover, integrating multi-omic data with the associated H&E slides, i.e. multimodal data integration, led to improvements in prognostic prediction on overall 49 survival (OS) for most cancer types [15,16].

Previous DL studies primarily focused on training models to extract features from WSIs under supervision of potentially extensive and time-consuming human-derived annotations on slide or pixel level, i.e. supervised learning [10,12,15,17]. Self-supervised learning (SSL) on the other hand, has gained significantly increasing attention for its capacity to automatically capture image features from unlabeled data [11]. Applications of SSL models, including general-purpose foundation models [18,19], have demonstrated superior performance in various downstream cancer classification, survival, and molecular phenotype prediction tasks compared to traditional supervised learning models [20–24]. Barlow Twins, an SSL model designed to learn non-redundant image features, has several advantages over other SSL learning models (e.g. contrastive learning models), including not requiring extensive batch sizes nor asymmetry between the network "twins" [25].

Despite the efficacy in decision-making, DL models are often labeled as "black boxes", posing significant challenges in terms of interpretability. Supervised attention-based multiple instance learning

is a common interpretive method, enabling DL models to concentrate on informative segments within WSI according to predefined training labels [26]. Another approach involves employing unsupervised clustering algorithms to organize extracted features into clinically relevant and interpretable clusters which can be subsequently linked to diverse patient-related outcomes [23,27,28]. These clustering methods offer significant advantages, including the prediction of various clinical outcomes, intuitive visualization for pathologists, and interpretation and correlation with a range of molecular data.

In this work, our objective is twofold: first, to automatically and reliably extract clinically relevant histologic patterns from WSIs, which can be interpreted by pathologists, and second, to investigate the connections between these patterns and patient outcomes as well as molecular phenotypes across different treatment groups within a large clinical trial for colon cancer. To achieve this, we apply an SSL pipeline [23] involving the Barlow Twins encoder for feature extraction, followed by a community detection algorithm to construct an unbiased atlas of histomorphologic phenotype clusters (HPCs). This algorithm is exclusively trained on public data from the colon adenocarcinoma cohort within The Cancer Genome Atlas (TCGA) multi-institutional database [29]. Remarkably, the identified HPCs generalizes well in unseen WSIs obtained from the clinical Bevacizumab-Avastin® adjuVANT (AVANT) trial [30]. Subsequently, HPCs are linked to patient OS. An HPC-based classifier trained using TCGA data for OS, demonstrates prognostic significance in the external validation of the AVANT study, even when considering key clinical and demographic factors typically employed in clinical settings. Notably, by conducting comprehensive analyses of the distinct histomorphologic features of each HPC and their associations with immune and genetic profiles, we provide insight into morphological and molecular determinants of patient survival upon different treatments (e.g. standard-of-care adjuvant chemotherapy and experimental targeted therapies). Exploring these features further could yield additional insights into other histopathology diagnostics, supporting shared decision-making and advancing personalized treatment options for colon cancer patients in the future.

Results

Self-supervised learning of WSI features using the multi-institutional TCGA dataset

We trained the self-supervised algorithm using data exclusively from the TCGA colon adenocarcinoma (TCGA-COAD) set, eliminating the need for annotations by pathologists (Figure 1). A total of 435 WSIs (428 patients) obtained from the TCGA-COAD dataset (see Methods: Study population for details) was first divided into smaller image patches (224-by-224 pixels), also known as image tiles, at a magnification level of 10x (Figure 1a) (see Methods: Data pre-processing for details). To identify features on these patches, we trained an SSL Barlow Twins feature extractor using a random subset of tiles (Figure 1a) (see Methods: Extracting image features using Barlow Twins for details). The Barlow Twins was trained with the objective function to evaluate the cross-correlation matrix between the embeddings (feature vector z) of two identical backbone networks, which were fed distorted variants of a batch of image tiles. The objective function was optimized by minimizing the deviation of the crosscorrelation matrix from the identity matrix. This led to increased similarity among the embedding of z vectors of the distorted sample versions, while reducing redundancy among the individual components of these vectors. As a result, each tile was described as a vector of 128 extracted features that can subsequently be used to group tiles into clusters by similarity.

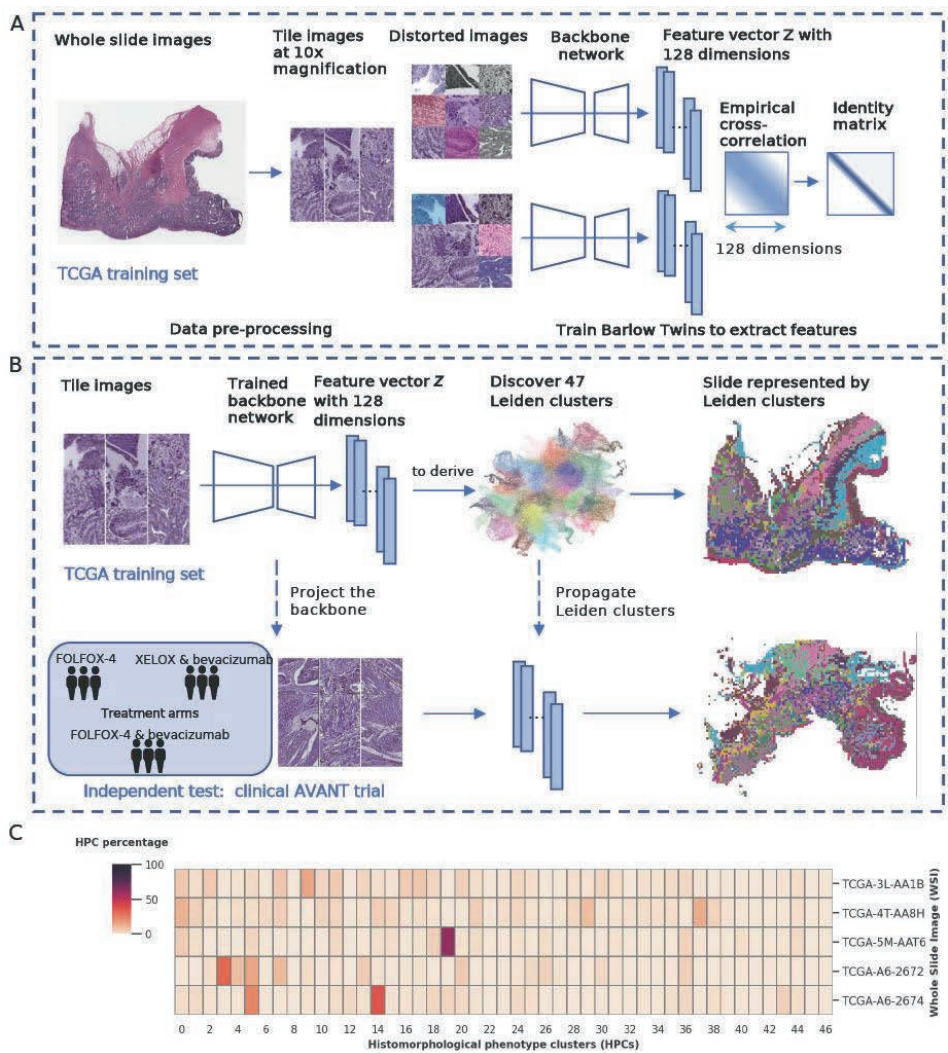


Figure 1. Overview of the Model Architecture: Training Barlow Twins and deriving Histomorphological Phenotype Clusters. (A) Training Barlow Twins with TCGA. WSIs from TCGA were processed to extract image tiles and normalize stain colours. The Barlow Twins network was employed to learn 128-dimensional z vectors from these image tiles. (B) Deriving HPCs. The tiles from TCGA were projected into z vector representations obtained from the trained Barlow Twins network. HPCs were defined by applying Leiden community detection to the nearest neighbour graph of z tile vector representations. Each WSI was represented by a compositional vector of the derived HPCs, indicating the percentage of each HPC with respect to the total tissue area. The Barlow Twins model and HPCs were then projected and integrated into the external AVANT trial. (C) Whole Slide Image Representation. The compositional HPC data represented the WSIs in the study. AVANT, Bevacizumab-Avastin® adjuvant trial. HPC, histomorphological phenotype cluster. TCGA, The Cancer Genome Atlas. WSI, whole slide image.

Construction of an unbiased atlas of histologic patterns through community detection

We applied the Leiden community detection algorithm to derive HPCs, i.e. the clusters with similar histologic patterns (Figure 1b) (see Methods: Identification of HPCs for details). The process began by first projecting the trained Barlow Twins onto the entire TCGA-COAD dataset, extracting 128 dimensional feature representations for each image tile. Subsequently, we utilized Leiden community detection on a nearest neighbor graph constructed from these tile vector representations (Figure 1b). Tiles with similar vector representations were clustered into a group and assigned a specific HPC ID number. The optimization of the Leiden configuration was achieved through an unsupervised process (see Methods: Identification of HPCs for details, Supplementary Figure 1a), resulting in the identification of a total of 47 HPCs, visually represented in a dimensionality reduction plot (Uniform Manifold Approximation and Projection; UMAP plot) (Figure 2a).

As external dataset, we analyzed a total of 1213 colon cancer patients with diagnostic pathology H&E WSIs (one WSI per patient), a subset of the clinical AVANT trial [30,31] (see Methods: Study population for details). We harnessed the optimized SSL Barlow Twins model to generate embeddings of the unseen AVANT WSI tiles. The assignment of identified HPCs to the unseen AVANT was achieved using the K-nearest neighbors approach. The HPC label of each tile in the AVANT data was determined based on majority votes from its K-nearest neighbors in the TCGA training set (Figure 1b). As a result, we obtained comprehensive visual representations of the WSIs where the WSI tiles are colored by their corresponding HPC (Figure 1b). Additionally, we were able to capture the characteristics and heterogeneity of WSIs using the compositional data derived from the HPCs, i.e. the percentage of the area on a WSI covered by each HPC, thus facilitating downstream analyses and modeling (Figure 1c).

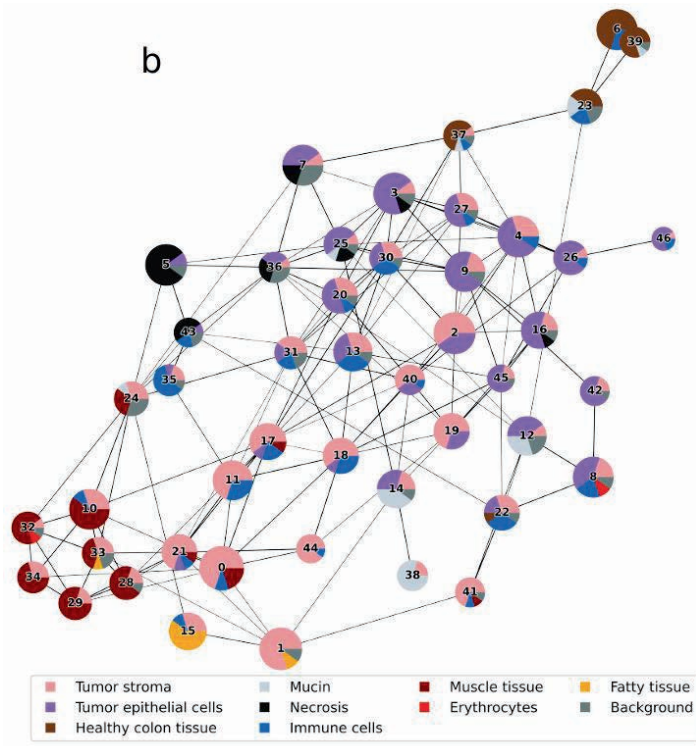
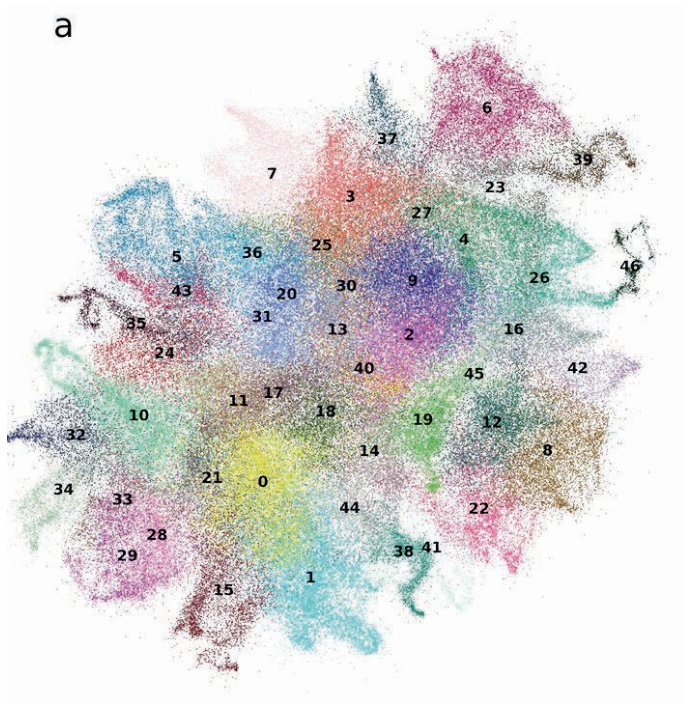
Histopathological assessment and characterization of HPCs

Each HPC underwent histopathological analysis on a randomly selected set of 32 tiles per cluster within TCGA, independently evaluated by two pathologists (ASLPC and JHJMvK) and a researcher (MP)(see Methods: Interpretation of HPCs for details). Tissue types, as observed on the tiles, were described with specific attention to tumor epithelium, tumor stroma and immune cells. Other unique histopathological features or patterns, such as tumor differentiation grade and stromal organization, were noted as well. All present tissue types were scored in percentages (Supplementary Table 1) and depicted using pie charts (Figure 2b). We plotted the interconnections of 47 HPCs using partition-based graph abstraction (PAGA) [32], with the pie charts reflecting their tissue compositions (Figure 2b). Interestingly, distinct larger groups of clusters, or "super-clusters", could be observed based on the similarity of tissue composition, interconnectedness, and topology of HPCs in the PAGA plot (Figure 2b).

In total, we identified eight super-clusters: (1) healthy and dysplastic colon tissue, (2) necrosis, (3) mucinous areas, (4) immune cells, (5) muscle tissue (longitudinal/axial), (6) fatty tissue, (7) tumor stroma, and, (8) tumor epithelium, in no particular order, formed by groups of HPCs shown in Figure 2c. Common histopathological characteristics were noted among HPCs within each designated super-cluster, while HPCs encompassing varied tissue types across multiple super-clusters were often situated at their intersections. For instance, HPC 12, containing not only tumor epithelium but also mucinous tumor, was located between HPCs belonging to the two super-clusters of mucinous tumor and tumor epithelium in the PAGA plot. Moreover, HPC 23, marked by dysplastic colon tissue, formed a bridge between the healthy colon tissue HPC 39 and the tumor epithelium-containing super-cluster HPCs (e.g. HPCs 4, 26, 46), suggesting a potential chronological pathogenesis. In summary, the derived HPCs displayed distinctive histopathologic characteristics. Moreover, HPCs located in close proximity on the UMAP and PAGA plot, based on the extracted features, demonstrated common traits, hinting at potential pattern relationships, mixed phenotypes, or pathogenic trajectories.

Assessment of HPC consistency within and across TCGA and AVANT cohorts

Although SSL methods have been applied recently in histopathology, there is usually no systematic analysis of consistency of the histologic patterns discovered by these methods within and across datasets [21,27]. Here, we address this potential pitfall by incorporating several qualitative and quantitative assessments.



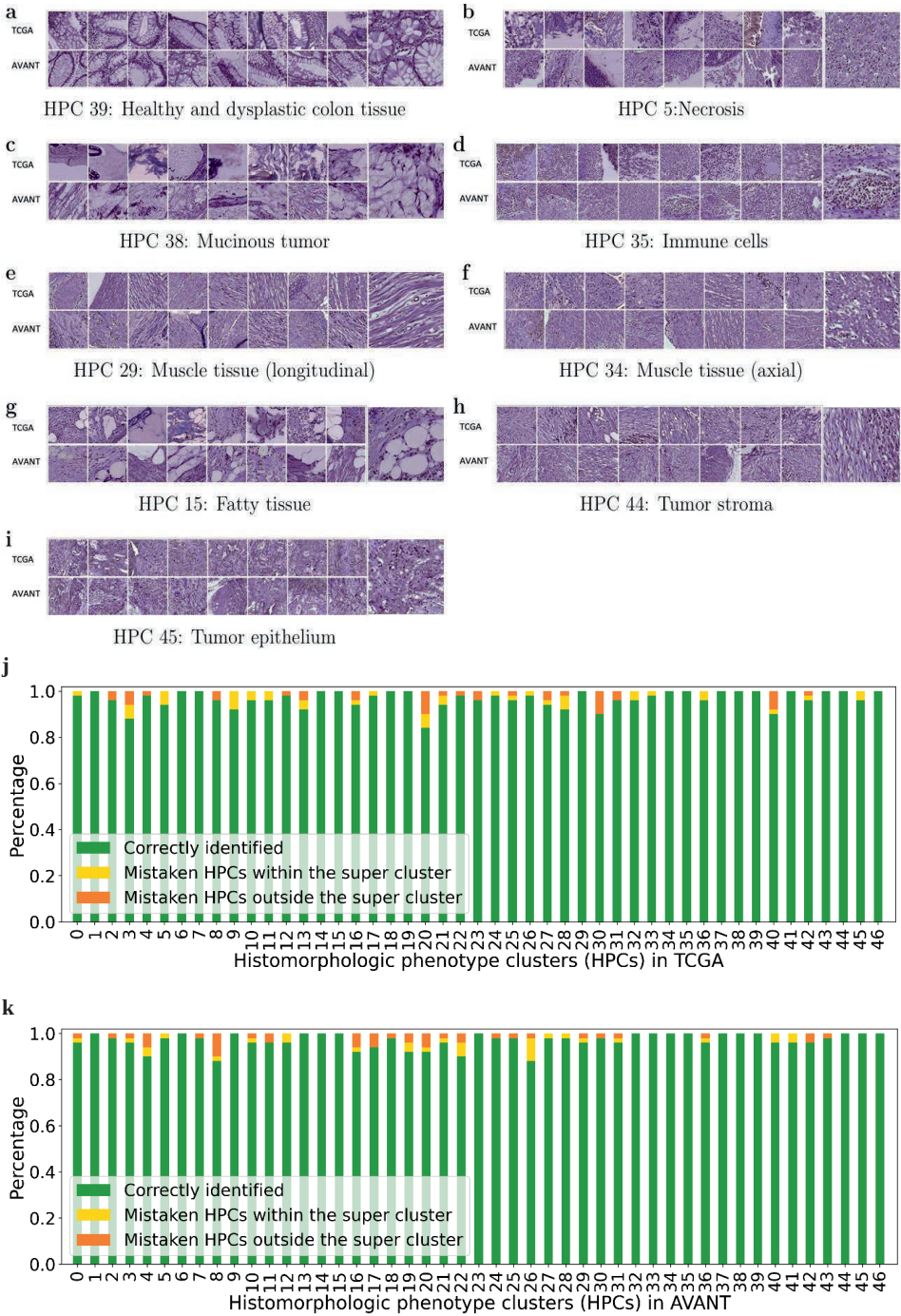
C

Healthy and dysplastic colon tissue	39 Healthy colon tissue	6 Inflamed healthy colon tissue	23 Dysplastic colon tissue	37 High grade dysplasia-adenoma	Necrosis	5 Necrotic	43 Necrotic	7 Avital tumor with necrosis	36 Avital tumor with necrosis	25 Moderately differentiated tumor necrotic	
											
Mucinous	14 Mucinous tumor	38 Mucinous tumor stroma	12 Mucinous tumor		Immune cells	13 Infiltrated stroma	18 Infiltrated stroma	35 Immune cells	30 Mixed, immune-high	22 Infiltrated stroma	31 Mixed, infiltrated stroma
											
Muscle tissue (longitudinal/axial fibres)	29 Muscle (longitudinal)	28 Muscle (longitudinal)	33 Muscle (longitudinal)	10. Muscle (longitudinal)	Fatty tissue	15 Fatty tissue					
											
Tumor stroma	0 Disorganized stroma	1 Edematous stroma	41 Disorganized and loose stroma	19 Edematous disorganized stroma	Tumor epithelium	4 Moderately differentiated tumor - aligned stroma	27 Moderately differentiated tumor - aligned stroma	3 Moderately differentiated tumor	9 Mixed tumor epithelial - infiltrated stroma	8. Mixed tumor epithelial - infiltrated stroma	
											
	11 Vascularized stroma-immune cells	21 Vascularized stroma	17 Infiltrated aligned stroma-high	44 Disorganized, immune-low stroma (desert-type)		16 Well-differentiated tumor	45 Poor-differentiated tumor	42 Moderately differentiated tumor	46 Infiltrated undifferentiated tumor	26 Infiltrated poor-undifferentiated	
											
	2 Aligned stroma - well differentiated tumor	40 Disorganized stroma, growing infiltratively				20 Disorganized and vascularized stroma-low					
											

Figure 2. Identification of HPCs in TCGA and subsequent classification into super-clusters. (A) UMAP showing 47 HPCs identified from the TCGA dataset, each scatter representing an image tile. (B) PAGA plot of HPCs. Each node represented an HPC with edges representing connections between HPCs based on their vector representation similarity. The pie chart of each node represented the tissue composition for each HPC. (C) Grouping of HPCs into super-clusters according to histopathology tissue similarities. Representative tiles for each HPC were labelled with ID and a brief description. HPC, histomorphological phenotype cluster. PAGA, partition-based abstraction graph. TCGA, The Cancer Genome Atlas. UMAP, uniform manifold approximation and projection plot.

First, qualitative assessment was conducted to evaluate the within-cluster and between-cluster heterogeneity of 47 HPCs derived in TCGA-COAD. Based on 32 randomly selected tiles from each HPC in the TCGA-COAD, three experts (ASLPC, JHJMvK, and MP) independently assessed each HPC by comparing tissue type quantity and architectures (histopathological assessment procedure stated above). Overall, all raters reached the general consensus that there was a noteworthy level of within-cluster morphological similarity and a significant diversity among the 47 HPCs (Figure 3 [a-i]), although phenotypic similarities varied across HPCs, implying that some HPCs may appear more similar than others. To delve deeper into the within-cluster and between-cluster heterogeneity, we carried out quantitative objective blinded tests within TCGA and AVANT tiles separately. This was to ascertain whether the morphological patterns identified by each HPC could also be recognized by human experts.

In this test, the assessor (MP) was shown three groups of image tiles, each containing five tiles. Two groups were from the same HPC, and the third was from a randomly selected other HPC, also called the "odd HPC". The assessor was required to identify the "odd HPC" (Supplementary Figure 2a, see Methods: Pathologist assessment of HPCs for details). Each of the 47 HPCs underwent 50 tests to determine the success rate. Within TCGA, we found that 17 out of the 47 HPCs achieved 100% identification rate, while the remaining 30 HPCs had a correct identification percentage ranging from 84% to 98%. Similarly, within in AVANT, 17 out of 47 HPCs achieved a perfect accuracy, while the rest had an accuracy ranging from 88% to 98% (Figure 3j). In general, HPCs in close proximity to each other in the PAGA plot or belonging to the same super-cluster were more prone to erroneous assignment.



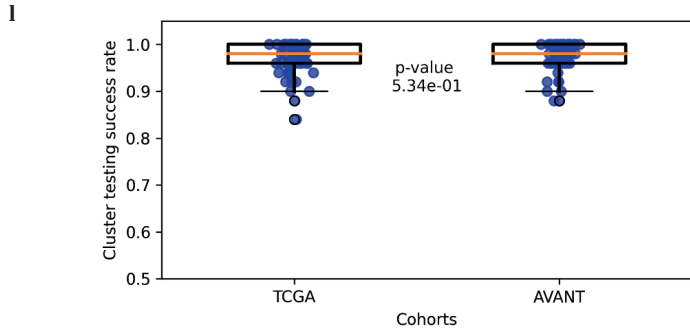


Figure 3. Verification of HPCs in the TCGA training set and the external clinical AVANT trial (a-i) Example tiles from TCGA (upper row) and AVANT (lower row) showcase the eight super-clusters with a zoomed-in representative tile. The muscle tissue super-cluster is further divided into longitudinal and axial subgroups. (j-k) Stacked bar plots illustrate instances of misclassification for each HPC in TCGA training set and AVANT external test set. Green bars represent the percentage of correctly identified odd clusters, yellow bars indicate misclassifications within the tested HPC’s super-cluster, orange bars show misclassifications outside the super-cluster. (l) Box plots display similar distributions of success test rates (corresponding to the green bars in panels j and k) for HPCs in TCGA and AVANT cohorts (two-sided Wilcoxon signed-rank test, $p = 0.534$). Each blue point within each box plot represents the success test rate for a single HPC, calculated based on 50 tests per HPC. The central orange line within each box represents the median, while the bounds of the box indicate the 25th and 75th percentiles (interquartile range). Whiskers extend to the minima and maxima within 1.5 times the interquartile range. Points outside the whiskers represent outliers. HPC, histomorphological phenotype cluster. TCGA, The Cancer Genome Atlas. AVANT, Bevacizumab-Avastin® adjuvant trial.

To evaluate the effectiveness of transferring morphological patterns from the TCGA to the external AVANT test set, three experts (ASLPC, JHJMvK, and MP) independently reviewed a randomly chosen set of 32 tiles from the TCGA-COAD subset and another 32 tiles randomly selected from the AVANT trial. This qualitative comparison concluded a remarkable resemblance between the TCGA and AVANT tiles within their respective HPCs (Figure 3 [a-i]). In comparing the objective test results from TCGA and AVANT, we found an 80% overlap in the misclassified HPCs between the two datasets and 65% overlap in correctly classified HPCs (Figure 3j, 3k). These results indicate that the robust morphological features extracted from the training set can be effectively transferred to an independent unseen test set.

HPC-based classifier was associated to OS in patients treated with standard-of-care and AVANT-experimental treatment

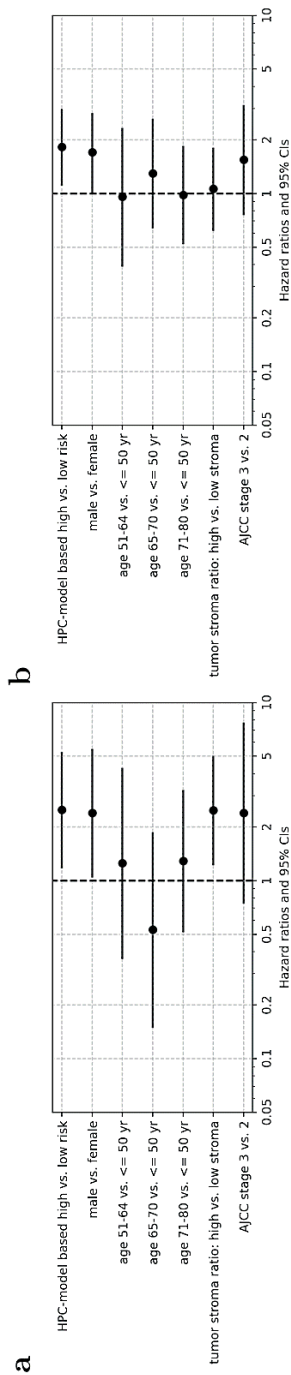
We explored the prognostic significance of HPCs on OS. The OS prediction model was developed within TCGA-COAD (see Methods: Optimization of HPCs and the prediction of OS using Cox regressions with L2 regularization for details). For external validation, we utilized the control group from the AVANT trial who had only received standard adjuvant chemotherapy (i.e. FOLFOX-4). The AVANT trial aimed to investigate whether combining bevacizumab, a humanized anti-vascular endothelial growth factor (VEGF) monoclonal antibody, with standard chemotherapy would improve survival among colon cancer patients [30,33]. The trial had three treatment arms: FOLFOX-4, bevacizumab+FOLFOX-4, bevacizumab+XELOX. The study was prematurely terminated due to the adverse effect in patient survival associated with bevacizumab [30]. Given the unique bevacizumab experimental treatment and its adverse effects, we hypothesized that the survival model trained on TCGA-COAD patients may not generalize well in the AVANT bevacizumab-treated group. We referred to our source population, represented by the TCGA-COAD, as the "standard-of-care group", in contrast to the unique bevacizumab treatment in the AVANT trial. We then opted to validate the OS prediction model, trained on TCGA-COAD data, primarily on AVANT patients who exclusively received standard FOLFOX-4 chemotherapy. This subset is henceforth referred to as the "AVANT control group", serving as an independent test set for validation.

We modeled HPCs on OS using Cox regression with L2 regularization trained within the TCGA-COAD incorporated all 47 HPCs as predictors. The model was optimized through five-fold cross validation (CV) on the TCGA training set (Supplementary Figure 3a). The optimized regularized Cox model was then tested in the independent test set of the AVANT control group. We observed a test set c-index of 0.65 (bootstrap 95% confidence interval [CI]=0.55-0.74) (Supplementary Figure 3b). The HPC-based classifier (i.e. high risk versus low risk) was determined by the median predicted hazard obtained in the TCGA-COAD. The HPC-based model also outperformed a clinical baseline model trained on age, sex, tumor-stroma ratio (TSR), and AJCC TNM stage (Supplementary Figure 3c). To investigate whether our HPC-based classifier provides additional prognostic value to the existing important clinical predictors, a regular multivariable Cox regression was fitted within the external AVANT control test set. The model included HPC-based classifier as well as important clinical and demographic variables (Figure 4a). Notably, the HPC-based risk classifier demonstrated significance as an independent prognostic factor (hazard ratio [HR] = 2.50, 95% CI=1.18 – 5.31), along with male sex 213 (HR = 2.42, 95% CI = 1.07 – 5.47) and the TSR (HR = 2.49, 95% CI = 1.23 – 5.04). The 20 most important HPCs associated to OS were summarized using the interpretable SHapley Additive exPlanations (SHAP) (Figure 4c).

Harnessing the well-defined experimental protocols in AVANT, we were granted the unique opportunity to examine the influence of HPCs on colon cancer OS in the bevacizumab treatment groups (i.e. bevacizumab+FOLFOX-4 or bevacizumab+XELOX). Given the proven comparable therapeutical efficacy of FOLFOX-4 and XELOX [33–35], we consolidated the bevacizumab+FOLFOX-4 and bevacizumab+XELOX cohorts into a unified "AVANT-experimental group". Similar to the analysis stated above, we trained Cox regressions with L2 regularization encompassing all 47 HPCs within the AVANT-experimental patients using 5-fold CV (Supplementary Figure 3d). The HPC-based classifier remained of independent prognostic value (HR = 1.82, 95% CI = 1.11 – 2.99) after adjusting for age, sex, TNM tumor staging, and TSR (Figure 4b). The top 20 most influential HPCs on OS prediction were shown in the SHAP summary plot (Figure 4d).

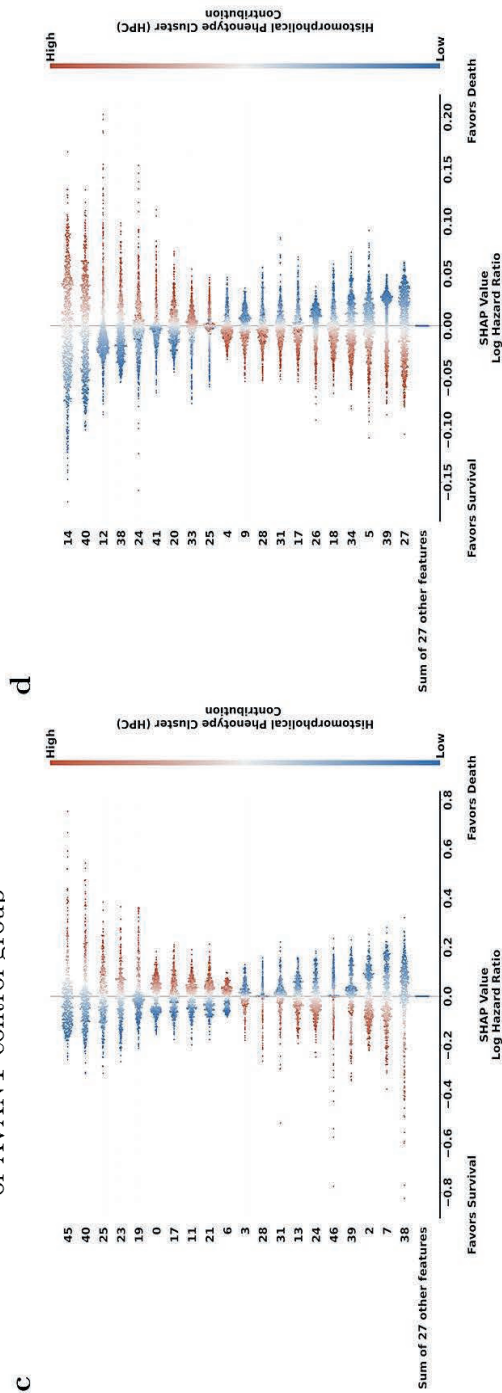
Pathological assessment of OS-associated HPCs in the standard-of-care and AVANT-experimental groups

We highlighted the top 20 most influential HPCs on OS prediction reflecting the SHAP summary plot (Figure 4c, 4d) in the PAGA plots for the standard-of-care and AVANT-experimental groups (Figure 5). The survival-favorable HPCs were highlighted in a shade of red and the survival-unfavorable HPCs were highlighted in blue. Preliminary findings showed that HPCs containing proportionally more healthy colon tissue or immune cells appeared associated with an improved OS, and HPCs comprising mucinous tumor, tumor stroma, and poor-to-undifferentiated (high-grade) tumor epithelium were associated with worse OS in both groups. An overview of the HPCs and their associations to OS, including pathological explanations and references, were given in Supplementary Table 2. We will highlight some OS-associated HPCs in depth here.



Standard-of-care group: external test set
of AVANT control group

AVANT-experimental group



Top 20 OS-associated HPCs
trained in standard-of-care TCGA

Top 20 OS-associated HPCs
trained in AVANT-experimental groups

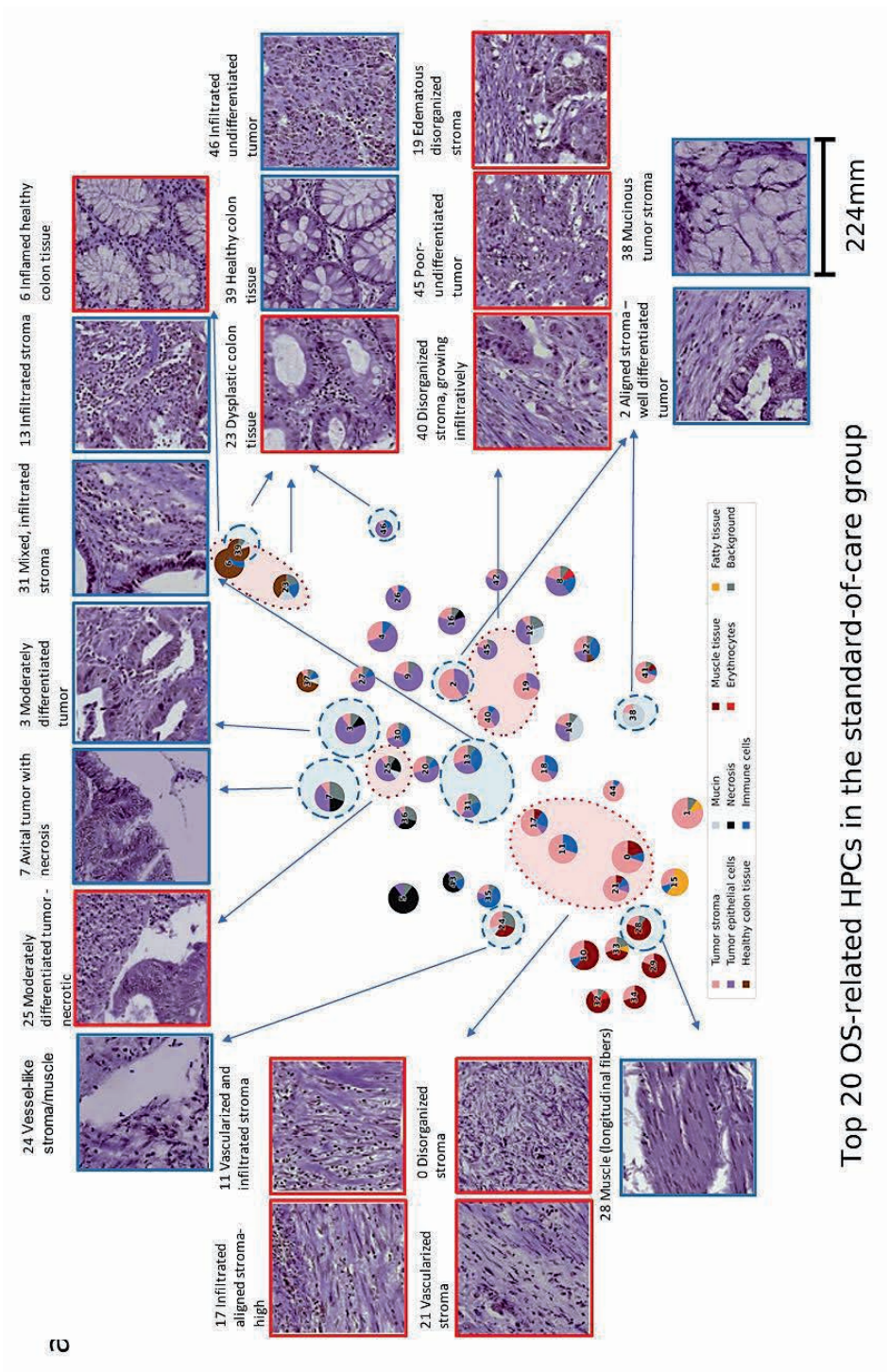
(previous page) **Figure 4.** HPC-based classifier was associated with OS in patients treated with standard-of-care and AVANT-experimental treatment. (a) Ordinary Cox regression for OS, incorporating the HPC-based risk classifier, along with sex, age categories, tumor-stroma ratio, and AJCC TNM staging, was conducted within the external test set of the AVANT control group (N=379 patients after excluding those with missing clinical information). Each point represents the point estimate of HR, and the horizontal whiskers depict the 95% CI. The HPC model-based classifier stands as an independent prognostic factor (HR = 2.50, 95% CI = 1.18 – 5.31) for OS. (b) Ordinary Cox regression for OS, incorporating the HPC-based risk classifier, along with sex, age categories, tumor-stroma ratio, and AJCC TNM staging, was conducted within the AVANT experimental group (N=751 patients after excluding those with missing clinical information). Each point represents the point estimate of HR, and the horizontal whiskers depict the 95% CI. The HPC model-based classifier stands as an independent prognostic factor (HR = 1.82, 95% CI = 1.11 – 2.99) for OS. (c and d) The SHAP summary plots depict the relationship between the center-log-transformed compositional value of an HPC and its impact on death hazard prediction. Statistics estimated from AVANT control (N=405 after excluding those with missing survival data) (c) and experimental groups (N=780 patients after excluding those with missing survival data) (d). The color bar indicates the relative compositional value of an HPC, with red indicating higher and blue indicating lower composition. Higher compositions of the top 10 HPCs were associated with worse OS, while higher compositions of the bottom 10 HPCs were linked to improved OS. AJCC TNM, American Joint Committee on Cancer tumor-node-metastasis classification. AVANT, Bevacizumab-Avastin® adjuvant trial. HPC, histomorphological phenotype cluster. OS, overall survival. SHAP, SHapley Additive exPlanations. TCGA, The Cancer Genome Atlas.

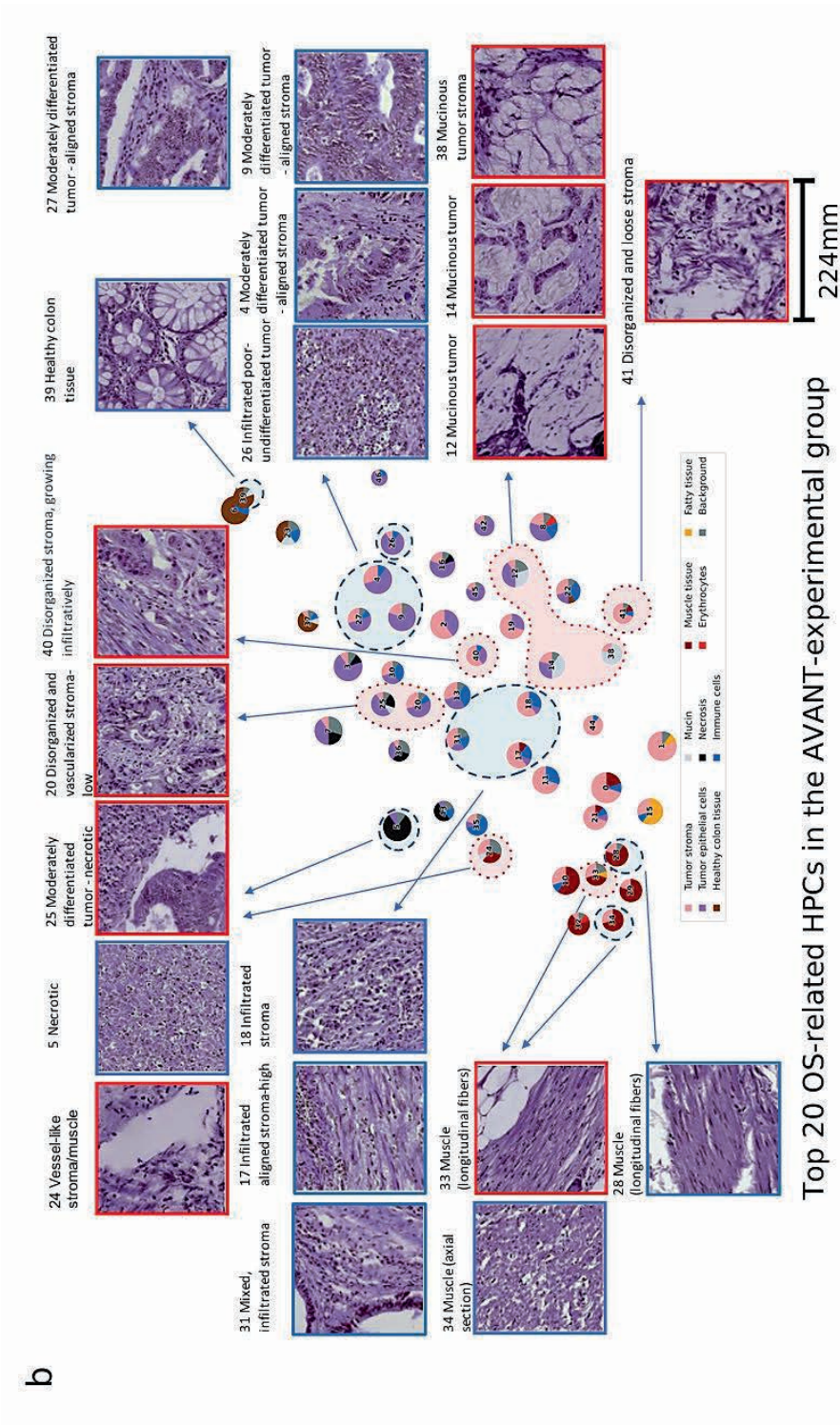
The healthy colon tissue-containing HPC 39 was among the top survival-favorable HPCs as indicated by larger SHAP values in both groups (Figure 4c, 4d). In the standard-of-care group, HPC 23 characterized by dysplastic and low-grade tumor epithelium and HPC 6 marked by stroma-infiltrated healthy colon tissue indicating inflammation, were associated with a worse OS. HPC 31, predominately composed of immune cells, was associated with an improved OS in both treatment groups. A similar association was also noted for the infiltrated-stroma HPC 13, which is part of the immune super-cluster, in the standard-of-care group (Figure 5a) as well as HPC 18 (infiltrated stroma, containing aligned tumor stroma and immune cells) and HPC 17 (infiltrated aligned stroma-high; organized tumor stroma with an immune component) in the AVANT-experimental group (Figure 5b). In the AVANT-experimental group, the mucinous super-cluster (HPCs 12, 14 and 38) was associated with a poor survival (Figure 5b), while conversely, HPC 38 (mucinous tumor stroma) led to a better survival in the standard-of-care group (Figure 5a).

Upon closer examination, we noted that the algorithm captured a distinction in stroma organization within the stromal tissue presented in the tumor stroma and tumor epithelium super-clusters. HPCs in these two super-clusters all contained some component of stromal tissue. However, HPCs with disorganized or heterogeneous tumor stroma with neovascularization (HPC 40 in both groups; 0, 11 and 21 in standard-of-care, and 41 in the AVANT-experimental group) were associated with a poor survival, whereas aligned and organized tumor stromal "strands" were often observed among the top survival-favorable HPCs (HPC 2 in the standard-of-care and HPC 27 in the AVANT-experimental group).

Moreover, the tumor differentiation grade was another parameter correlating with HPCs and their impact on OS. Poor-to-undifferentiated (high-grade) tumor epithelium, e.g. HPC 45, in the standard-of-care group, was linked to a worse survival, while well-to-moderately differentiated (low-grade) tumor led to an improved survival (e.g. HPCs 3 in the standard-of-care group and HPC 4 in the AVANT-experimental group). A unique pattern was furthermore observed in survival-favorable HPCs 46 (standard-of-care group) and 26 (AVANT-experimental group), which was characterized by a predominance of poor-to-undifferentiated tumor epithelium but accompanied by a notable influx of immune cells.

Additional findings were related to the survival-favorable HPC 7 (avital tumor with necrosis) in the standard-of-care group (Figure 5a) and the positive association between necrosis-dominated HPC 5 and better survival in the AVANT-experimental group (Figure 5b). The associations between HPCs containing muscle tissue (e.g. HPC 24, 33 and 34) and survival were generally inconsistent, possibly due to the similarity of muscle fibers and tumor stroma, caused for instance by the organization of muscle fibers and vascularization.





(previous pages) **Figure 5.** PAGA plots highlighted with important HPCs related to OS in the standard-of-care and experimental treated group (a) Standard treated group: HPCs colored in the red are linked to worse survival and HPCs colored in blue are linked to better survival. (b) AVANT-experimental treated group: HPCs colored in the red are linked to worse survival and HPCs colored in blue are linked to better survival. AVANT, Bevacizumab-Avastin® adjuVANT trial. HPC, histomorphological phenotype cluster. PAGA, partition-based graph abstraction.

Additional findings were related to the survival-favorable HPC 7 (avital tumor with necrosis) in the standard-of-care group (Figure 5a) and the positive association between necrosis-dominated HPC 5 and better survival in the AVANT-experimental group (Figure 5b). The associations between HPCs containing muscle tissue (e.g. HPC 24, 33 and 34) and survival were generally inconsistent, possibly due to the similarity of muscle fibers and tumor stroma, caused for instance by the organization of muscle fibers and vascularization.

Outcome-associated HPCs are linked to diverse immune features in the tumor microenvironment

Spearman correlation coefficients were calculated between the top 20 survival-related HPCs on one hand, and the TCGA immune landscape on the other hand (see Methods: Linkage between survival-related HPCs, immune landscape, and gene expression data for details). The correlation heatmap was plotted with bi-directional hierarchical clustering (Figure 6a). Interestingly, in the standard-of-care group (Figure 6a), HPCs 0, 11, 17 and 21, all part of the phenotypic tumor stroma super-cluster, were identified by the genotypic immune analysis as having a positive correlation with the stroma-high category. Moreover, their SHAP values indicated an positive association with worse survival. Conversely, HPCs 13 and 31 in the immune cell super-cluster were correlated with higher leukocyte fraction. Survival prediction model suggested the immune cell super-cluster was in general associated with a better survival. Further anticipated findings included the validation of HPCs 2 and 24, marked by stroma-high tiles, through immune feature correlation within the stroma-high category.

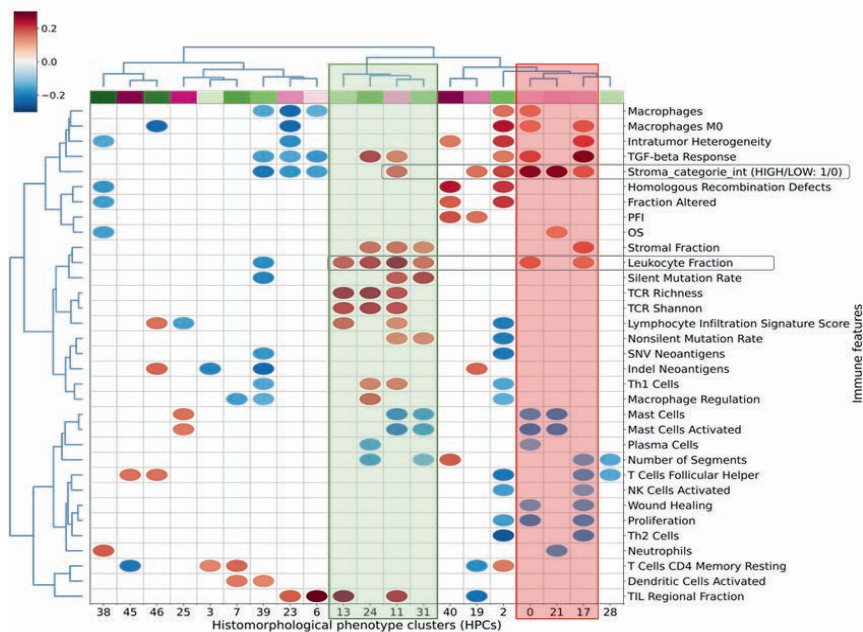
In the AVANT-experimental group (Figure 6b), survival-favorable HPCs 5, 17, 31, 18 correlated with an increased leukocyte fraction. Notably, HPCs 17 and 31 aligned with an elevated immune cell composition (Figure 5b). Additionally, HPCs 17 and 18 were also associated with a higher expression of stroma, which was consistent with the observed stroma-high morphology (Figure 5b). On the contrary, survival-unfavorable HPCs 14, 40 and 20 exhibited higher genomic instability. For instance,

HPCs 40 and 20 were positively correlated with homologous recombination defects, intratumor heterogeneity, and HPC 14 was positively linked to nonsilent mutation rate, single nucleotide variants, and indel neoantigens. Although HPC 14 also showed a positive correlation with the leukocyte fraction, a more pronounced association with genomic instability through multiple pathways seemed to play a greater role in its negative impact on OS.

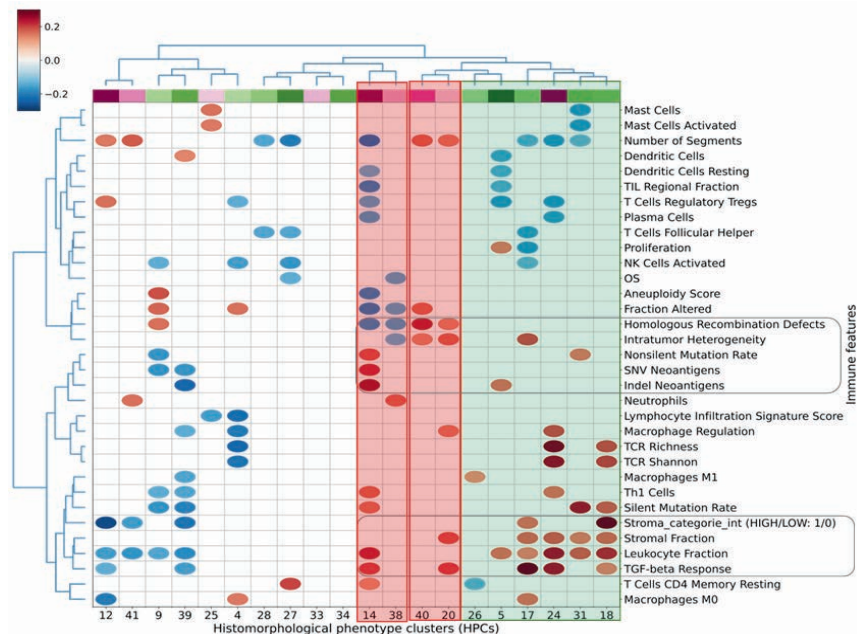
The resulting correlations with immune landscape data aligned with the observed morphologies of HPCs, particularly in stromal and immune features. Furthermore, the data suggest a potential role in genomic instabilities within the AVANT-experimental group. Taken together, these results demonstrate that HPCs can capture the remarkable heterogeneity of the tumor microenvironment.

Outcome-associated HPCs are linked to oncogenic pathways and bevacizumab's mechanism of action

Next, we performed gene set enrichment analysis (GSEA) to discern associations between top OS-related HPCs from both the standard-of-care and AVANT-experimental groups and key cancer hallmark pathways (Figure 6c and d). In the AVANT-experimental cohort, survival-related HPCs showed striking alignment with the enrichment observed in oncogenic hallmark pathways (Figure 6d), which such alignment was overall much less pronounced in standard-of-care group (Figure 6c). Still, in the standard-of-care group, pathways encoding epithelial-to-mesenchymal transition, leading to an increased tumor-stromal percentage, were enriched in survival-unfavorable HPCs 11, 17, 40 and 19. HPCs related to inflammatory response pathways showed primarily positive associations with OS, e.g. with HPCs 13 and 31 related to better survival.



Standard-of-care group: immune landscape



AVANT-experimental group: immune landscape

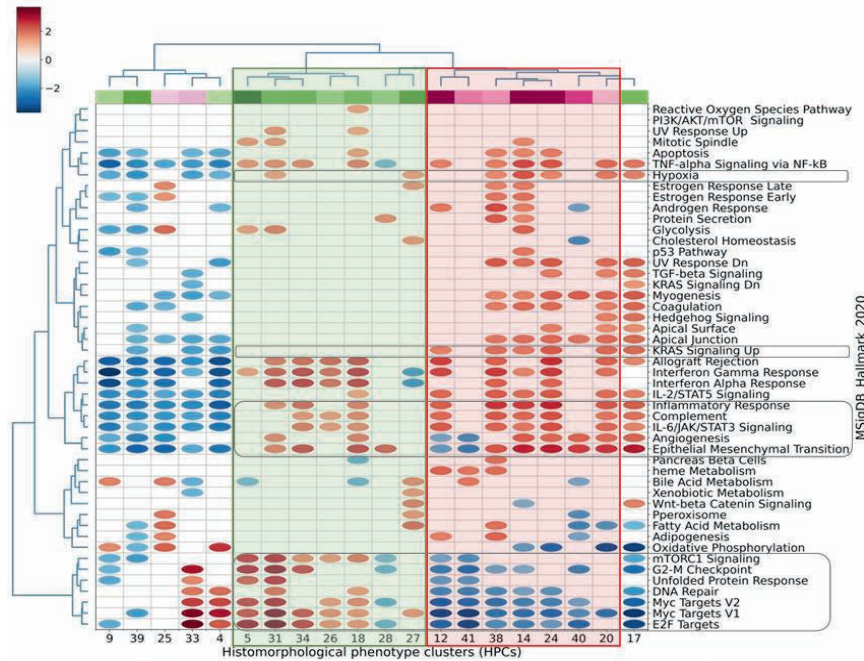
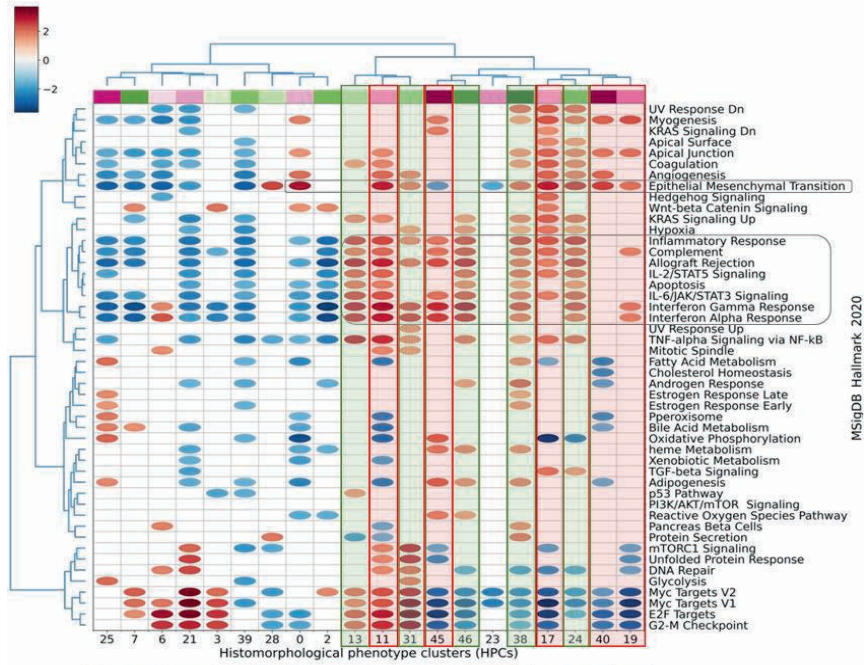


Figure 6. Survival-associated HPCs in relation to immune and genetic profile. (a) Standard-of-care group: Spearman's correlations between top 20 OS-related HPCs and immune landscape features. HPCs (columns of the matrix) were colored according to the beta-coefficients estimated from the optimized regularised Cox regression, with red indicating HPCs related to worse survival and green indicating HPCs related to better survival. The color bar at the upper left corner indicates the value of correlation coefficients with red denoting positive and blue denoting negative correlations. (b) AVANT-experimental treated group: Spearman's correlations between top 20 OS-related HPCs and immune landscape features. (c) Standard-of-care group GSEA between the top OS-related HPCs and major cancer hallmark pathways. HPCs (columns of the matrix) were colored according to the beta984 coefficients estimated from the optimized regularized Cox regression, with red indicating HPCs related to worse survival and green indicating HPCs related to better survival. The color bar at the upper left corner indicates the value of the correlation coefficients with red denoting enrichment and blue denoting underrepresentation in a gene pathway. (d) AVANT-experimental treated group GSEA for the top 20 OS-related HPCs. The immune landscape analysis (N=355 patients) and GSEA analysis (N=265 patients) were performed using data available from TCGA. AVANT, Bevacizumab-Avastin® adjuvant trial. GSEA, gene set enrichment analysis. HPC, histomorphological phenotype cluster. OS, overall survival.

In the AVANT-experimental group (Figure 6d), we observed a strong correlation between hierarchical clustering based on oncogenic enrichment scores and survival-related HPCs. Several HPCs exhibited enrichment in pathways that potentially be specific to bevacizumab through its target of *VEGFa* expression and *KRAS* signaling-up pathway. HPC 5, characterized by necrosis, was linked to elevated *VEGFa* expression ($\rho=0.163$, two-sided $p=0.005$), indicating that patients with higher pre-treatment necrosis levels may have benefited from the bevacizumab treatment. HPC 27, 18, and 31 were associated with enriched hypoxia (HPCs 31, 27) and angiogenesis (HPCs 18, 31), and unfolded protein (HPCs 31, 1, 5) pathways which involve the *VEGFa* gene. In addition, other survival-favorable HPCs may not be specific to bevacizumab but related to the standard cytotoxic chemotherapy of XELOX and FOLFOX-4. Survival-unfavorable HPCs (12, 38, 14, 24, 20) were linked to the *KRAS* signaling-up pathway, which may have a negative impact on the prognosis of patients treated with FOLFOX. Certain survival-favorable HPCs exhibited enrichment in pathways related to cell cycle regulation, signaling, DNA repair, and growth, including G2/M checkpoint (34, 31, 5), E2F targets (5, 31, 34, 26, 18), Myc targets (5, 31, 34, 26, 18), and mTORC1 signaling (5, 31, 34, 26, 18), DNA repair (18, 31). In contrast, survival-unfavorable HPCs demonstrated depletion in these pathways.

Discussion

In this study, we derived and independently validated a total of 47 distinct HPCs that were extracted from colon cancer H&E WSIs using a self-supervised algorithm. These HPCs possess distinctive histomorphologic features carefully identified and assessed by the pathologists and were also linked several immune features and oncogenic pathways. The HPCs showed state-of-the-art performance on OS prediction. Furthermore, the HPC-based risk classifier was an independent prognostic factor after adjusting for important clinical and demographic variables, suggesting additional insight beyond the current clinical prognostication. The unique AVANT trial enabled us to endeavor in identifying possible mechanisms of response to bevacizumab and standard chemotherapy using the HPCs.

Moreover, we hereby emphasize the importance of the tumor microenvironment, or tumor stroma, and its effect on survival. Tumor stroma is composed of extracellular matrix, vasculature, immune cells and cancer-associated fibroblasts, forming a complex and close interaction with tumor epithelial cells [36, 37]. Subject to increasing research the past decades, this dynamic entity has been found to modulate tumor behavior through its cross-talk, and ultimately influence patient-related outcomes. Specifically, regarding the amount of stroma and stromal architecture, i.e. alignment or categorization of the desmoplastic reaction, and immune infiltrate in the stroma, our results corroborate previous literature [31,38–47]. However, such assessments have not been implemented in standard pathology diagnostics.

We identified common histopathological patterns associated with survival as observed in both the standard and bevacizumab-treated group. In line with our results, immune cell infiltration (e.g. HPCs 13, 31) [40,46] and aligned, stroma-low tumors (e.g. HPCs 2, 27) [43,47] were associated with better survival, while poor-to-undifferentiated (high-grade) tumor epithelium (e.g. HPC 45) or mucinous tumor epithelium (e.g. HPCs 12, 14) [1,3,48,49], and disorganized stroma-high tumors (e.g. HPCs 40, 0) [43,47] were linked to worse survival. This pathological phenotype correlated with the corresponding genetic immune profile (e.g. increased leukocyte fractions correlated with HPCs 13 and 31; and a stroma-high category was seen in HPCs 0, 21, 17, 11) and enrichment in oncogenic pathways (e.g. epithelial-to-mesenchymal transition pathway, contributing to tumor stroma amounts [36], correlated with HPCs 17 and 11). We also observed an association, though imperfect, between mucinous tumor, poor-to-undifferentiated tumor epithelium, and survival. One explanation may arise from the absence of contextual information in small images. The differentiation between well-differentiated and undifferentiated, or adenocarcinoma and more mucinous tumor types, is established based on whether each tissue type constitutes more than 50% of the total tumor, underscoring the importance of considering the overall context [1,3,48,49].

Another interesting discovery emerges from the survival-favoring HPCs 26 (AVANT-experimental group) and 46 (standard-of-care group). These HPCs contain primarily poor-to-undifferentiated tumor epithelium but with a high influx of immune cells. Such a histopathological pattern is frequently observed in MSI tumors [50]. MSI-high tumors have been linked to a favorable prognosis [2,51], characterized by lower differentiation grade, increased T-cell infiltration, and reduced susceptibility to invasiveness and *KRAS* mutation [51], however, are commonly identified through separate MSI analysis and/or additional immunohistochemical staining for mismatch repair enzymes in pathology diagnostics [2,3,51].

Furthermore, we also noted HPC 39 containing predominately healthy colon tissue associated with better survival. Interpretation hereof lies in the nature of the multivariable analysis, where the 47 HPCs were modeled simultaneously. One can interpret this result as, while holding the other 46 HPCs constant, patients with relatively more healthy colon tissue, showed an improved survival. The higher proportion of healthy colon tissue may thus indicate relatively smaller or less aggressive tumors. Indeed, within routine TNM assessments, lower pathological T-stage is known to lead to an improved survival [1,3]. Additionally, the favorable OS could potentially also be linked to an absence of the cancer field effect, leaving this healthy colon tissue unaffected from the cancerous lesion [52].

In the GSEA analysis, we noted a remarkable concurrence in AVANT-experimental group between clustering based on cancer hallmark pathways and outcome-related HPCs, while such alignment was much less pronounced in the standard-of-care group. One possible explanation is the heterogeneity of patients in the TCGA-COAD dataset. In contrast to the well-defined treatment protocol in AVANT, TCGA-COAD patients encompass diverse disease and demographic profiles. Consequently, this diversity led to a wide spectrum of treatments, including surgical, neoadjuvant, and adjuvant therapies. The survival-related HPCs discovered in the TCGA-COAD could therefore be, if at all, related to multiple distinctive biological pathways. A general alignment observed in the AVANT-experimental group was hence not anticipated in the standard-of-care group.

In the AVANT-experimental group however, several survival-favoring HPCs either directly correlated with *VEGFA* expression or were associated to enrichment in oncogenic pathways involving *VEGFA* gene, indicating a favorable responses to the contentious bevacizumab treatment. In particular, HPC 5, primarily characterized by necrosis, emerged as a significant contributor to enhanced survival, displaying a positive correlation with *VEGFA* expression, which is the target of bevacizumab. Necrosis

promotes the expression of *VEGF α* , as dying tumor cells release signals that stimulate the growth of new blood vessels [33,53]. In addition, survival-favoring HPCs 18, 27 and 31 were associated with enriched hypoxia and angiogenesis pathways, which also involve the *VEGF α* gene. We hypothesize that patients exhibiting a higher abundance of HPC 5, 18, 27, and 31 may correspondingly express elevated levels of bevacizumab's target *VEGF α* , which in turn results in a more favorable response to this treatment.

Oxaliplatin plus 5-fluorouracil-based regimens of XELOX and FOLFOX-4 are standard chemotherapy for colorectal cancer [2]. Nonetheless, the treatment response is still modest with an estimated rate of approximately 50% [1,54], and the prediction of which patients will respond to this adjuvant chemotherapy remains challenging. We observed enrichment in oncogenic pathways that may be within the context of XELOX and FOLFOX-4 treatments in the AVANT trial. Survival-unfavorable HPCs (12, 38, 14, 24, 20) were linked to the *KRAS* signaling-up pathway. In line with previous literature, several *KRAS* mutations activate downstream signaling pathways and the *KRAS G12D* mutation was predictive of an inferior response to FOLFOX [55]. HPCs 12, 14, and 38 contained mucinous tumors tissue, which has also been linked to *KRAS* mutational burden [50,56]. Survival-favoring HPCs (5, 31, 34, 26, 18, 27) were associated with enrichment in pathways of cell cycle, signaling, DNA repair, and growth (i.e. G2/M checkpoint, E2F targets, Myc targets, and mTORC1 signaling), while survival-unfavorable HPCs (12, 41, 38, 14, 24, 40, 20) linked to a depletion of those pathways. Interestingly, patients with altered DNA repair capacity showed greater benefits from treatment with oxaliplatin [54,57]. Moreover, carriers of *MNAT1* gene, which is one of the leading genes in the G2/M checkpoint pathway [58], were linked to better treatment outcome of FOLFOX [54]. A plausible mechanism is that patients exhibiting activated oncogenic activities within these pathways might harbor a greater abundance of targets suitable for oxaliplatin-based cytotoxic chemotherapy.

Another finding was our HPC-based prediction on OS outperforming the clinical baseline model. This HPC-based risk classifier remained an independent prognostic factor (HR=2.50, 95% CI=1.18-5.31) after adjusting for crucial clinical and demographic variables including TSR and tumor stage. This finding aligns with recent findings reported by Jiang et al. [13]. Of interest, although both HPCs and the TSR scores were derived from H&E slides, they appeared to encapsulate distinct non-overlapping information. A potential explanation could be that the HPC-based classifier captures intricate details, while TSR assessment requires a broader contextual understanding, not fully attainable with small image patches [37,38,59]. Comparing our OS prediction directly with previous studies poses further challenges due to differences in cancer types (e.g., colorectal [12,13,60] instead of colon cancer only), varied outcomes (e.g., 5-year disease-free survival [DFS] [12]), diverse statistical measures (e.g., hazard

ratios [60]), and absence of independent test sets [15] in prior studies. Nevertheless, a recently study reported a test set c-index of 0.65 for OS prediction in colorectal cancer [13] aligning with our reported c-index in the context of colon cancer.

Building upon the aforementioned findings, this study showcases the prospective clinical utility of AI-generated HPCs (Figure 7). Cancer WSIs were preprocessed into image patches and subsequently used to train SSL encoders and to form HPCs. These HPCs serve as condensed representations of the original WSIs, ready to be inspected by pathologists and enabling flexible linkage to multimodal omics data. These HPCs hold promise in classifying various tumor characteristics, potentially predicting patient prognosis and discerning distinct sensitivity groups to various therapies. Although this study was already trained on multicenter TCGA data and validated in an external, clinical multicenter cohort, we plan to conduct additional validations in other population-based external cohorts to strengthen its clinical applicability. In a pathology lab, this could be implemented as followed: Alongside the routine pathology report, an AI-generated report would provide personalized prognostic risk quantification (e.g., based on the patient's HPC composition and SHAP values). The report would also include tissue composition descriptions for each HPC, granting pathologists a complete overview per patient. Subsequently, with these reports used in the multidisciplinary team meetings, a colon cancer patient can be granted an optimal personalized treatment strategy.

Given the significant potential implications of these HPCs on OS and therapy response, implementation of this AI-based analysis may be advocated to international guideline organizations, such as the TNM evaluation committee of the Union for International Cancer Control (UICC). Our HPC-based analysis not only summarizes the TNM classification, but also correlates to parameters assessed in standard pathology diagnostics, such as the International Collaboration on Cancer Reporting (ICCR) [61]. Of note, ICCR parameters pertain to a global assessment on whole slide or patient level, while HPCs can provide a local assessment of histopathological characteristics on the tile level. However, these tile-level histopathological characteristics can still be linked to ICCR parameters. For example, a core element of the ICCR, 'Histological tumor grade', can be correlated to HPCs containing tumor epithelium with different differentiation statuses. HPCs 2 and 3 are for instance characterized by well-moderately differentiated (or 'low-grade') tumor epithelium, while HPC 45 is formed by poorly-to-undifferentiated (or 'high-grade') tumor epithelium-containing tiles. Moreover, 'Histological tumor type' in the ICCR can be linked to HPC 12, which contains predominantly mucinous tumors. Lastly, the TNM classification, and particularly tumor size, in the ICCR reporting guidelines may be linked to HPCs containing healthy colon tissue (e.g. HPC 39), since WSIs with more of the healthy HPC 39 may correspond to a fractionally smaller tumor size (i.e. lower T-stage).

Despite the interesting findings, the study also has several limitations. The identification of HPCs was based on small image tiles as is imperial to model training, while information regarding the larger context is likely lost, as also stated above. For example, it is often challenging to distinguish aligned and organized tumor stromal strands from muscle tissue through traditional microscopic assessment [59]. Therefore, pathologists typically make this distinction based on contextual cues, color variances, or use additional immunohistochemical stainings, all of which is not available within the small image patches. Another limitation pertains to using TCGA as the training set. Although the TCGA is a large open-sourced database, it depends on the availability of registered clinicopathological data. This introduces potential bias and variability in data quality across participating institutions. Moreover, due to data availability, we were only able to focus on predicting OS rather than DFS which may better reflect tumor behavior and biology [29,62]. Lastly, due to the AVANT treatment regimen design, separate analysis regarding bevacizumab or oxaliplatin-based chemotherapy was not attainable. Nonetheless, we made efforts to differentiate between their distinct mechanisms by conducting histopathological inspections and correlating the findings with immune landscape and oncogenic pathways.

In conclusion, our study employed a self-supervised approach to identify and validate histopathological features in colon cancer that are recognizable by human eyes and relevant to prognosis. These features were interpreted through a pathology-focused perspective. Our results highlighted the clinical significance of tumor tissue type, stromal amount and architecture, and the involvement of immune cells. Integration of histopathological features with genetic and gene expression data unveiled potential insights into oncogenic pathways and their relation to patient survival. Utilizing data from the clinical AVANT trial, we proposed mechanisms influencing patient sensitivity to diverse treatments. Future research should focus on refining prediction accuracy and validating the proposed mechanisms regarding the therapeutic strategies in colon cancer.

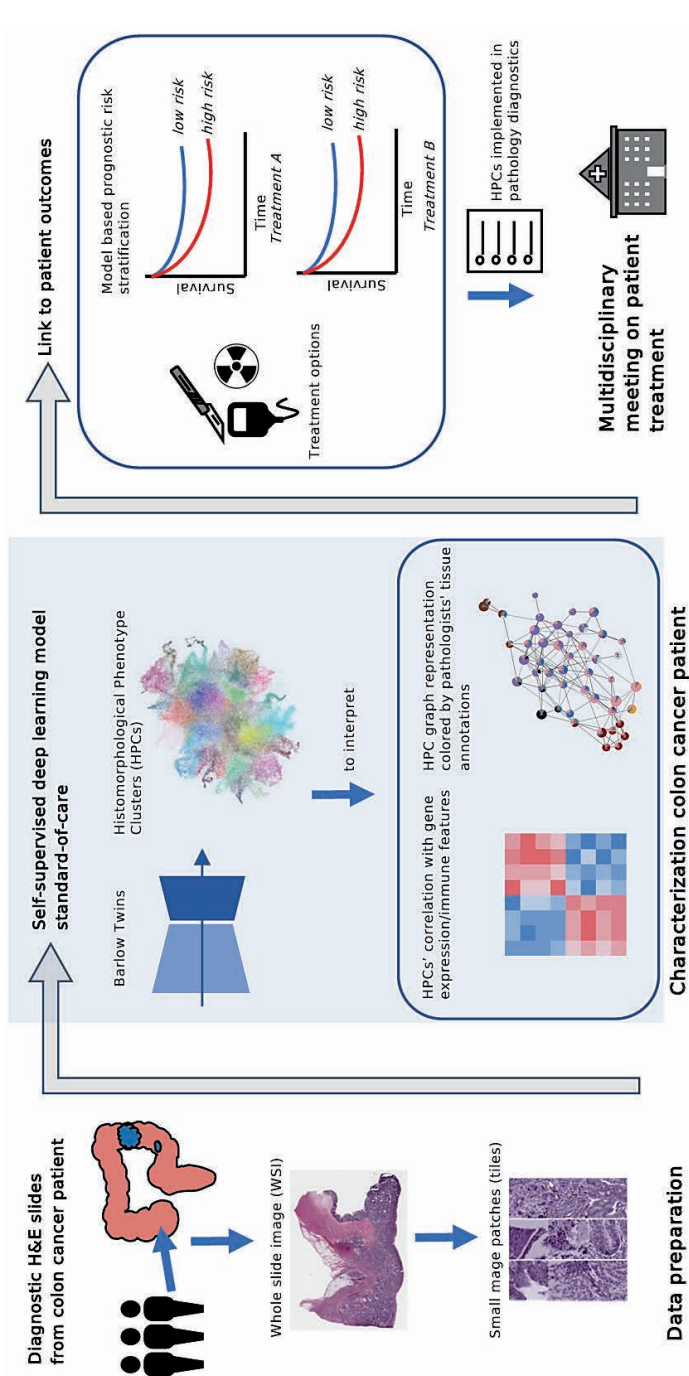


Figure 7. Clinical application of AI-derived HPCs in prediction of patient outcomes. The clinical algorithm consists of three key stages: data preparation, cancer patient characterization, and AI-supported multidisciplinary treatment meetings. Data preparation involves collecting histopathology WSIs, segmenting them into small image tiles. Patient characterization encompasses SSL model training, yielding HPCs via clustering. HPCs are easily interpretable by pathologists, linkable to omic data. Most importantly, HPCs are valuable for predicting diagnosis, patient outcomes, and treatment responses. In treatment-related outcomes, AI-predicted high/low risk groups aid multidisciplinary meetings, enabling personalized treatment plans by oncologists, pathologists, and other physicians. AI, artificial intelligence. HPC, histomorphological phenotype cluster. SSL, self-supervised learning, WSI, whole slide image.

Acknowledgements

The work is supported by the Swedish Research Council (BL, 2019-06360), NCI/NIH Cancer Center Support Grant (AT, P30CA016087), the Stichting Fonds Oncologie Holland and the Bollenstreekfonds, Hillegom, Netherlands (WEM, no grant numbers). We thank the team of NYU Langone High Performance Computing (HPC) Core's resources supporting us to perform the analysis. We would like to thank the NYU Applied Bioinformatics Laboratories (ABL) for providing bioinformatics support and helping with analysis of the data. ABL is a shared resource partially supported by the Cancer Center Support Grant P30CA016087 at the Laura and Isaac Perlmutter Cancer Center. We would also like to thank Leslie Solorzano for her input in data preprocessing. The AVANT-trial (BO17920) was originally financed by Genentech Inc., Roche, Switzerland. The authors thank Nikolas Jan Rakebrandt as liaison from Roche.

References

- James D Brierley, Mary K Gospodarowicz, and Christian Wittekind. *TNM classification of malignant tumours*. John Wiley & Sons, 2017.
- G. Argilés, J. Tabernero, R. Labianca, D. Hochhauser, R. Salazar, T. Iveson, P. Laurent-Puig, P. Quirke, T. Yoshino, J. Taieb, E. Martinelli, and D. Arnold. Localised colon cancer: Esmo clinical practice guidelines for diagnosis, treatment and follow-up. *Annals of Oncology*, 31(10):1291–1305, 2020.
- Martin R Weiser. *Ajcc 8th edition: colorectal cancer*. *Annals of surgical oncology*, 25:1454–1455, 2018.
- A Cervantes, R Adam, S Roselló, D Arnold, N Normanno, J Taieb, J Seligmann, T De Baere, P Osterlund, T Yoshino, et al. Metastatic colorectal cancer: Esmo clinical practice guideline for diagnosis, treatment and follow-up. *Annals of Oncology*, 34(1):10–32, 2023.
- Eileen Morgan, Melina Arnold, A Gini, V Lorenzoni, CJ Cabasag, Mathieu Laversanne, Jerome Vignat, Jacques Ferlay, Neil Murphy, and Freddie Bray. Global burden of colorectal cancer in 2020 and 2040: Incidence and mortality estimates from globocan. *Gut*, 72(2):338–344, 2023.
- Aoife Maguire and Kieran Sheahan. Controversies in the pathological assessment of colorectal cancer. *World journal of gastroenterology: WJG*, 20(29):9850, 2014.
- Kaustav Bera, Kurt A Schalper, David L Rimm, Vamsidhar Velcheti, and Anant Madabhushi. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nature reviews Clinical oncology*, 16(11):703–715, 2019.
- Hamid R Tizhoosh, Phedias Diamandis, Clinton JV Campbell, Amir Safarpour, Shivam Kalra, Danial Maleki, Abtin Riasatian, and Morteza Babaie. Searching images for consensus: can ai remove observer variability in pathology? *The American journal of pathology*, 191(10):1702–1708, 2021.
- Wouter Bulten, Hans Pinckaers, Hester van Boven, Robert Vink, Thomas de Bel, Bram van Ginneken, Jeroen van der Laak, Christina Hulsbergen-van de Kaa, and Geert Litjens. Automated deep-learning system for gleason grading of prostate cancer using biopsies: a diagnostic study. *The Lancet Oncology*, 21(2):233–241, 2020.
- Mohsin Bilal, Shan E Ahmed Raza, Ayesha Azam, Simon Graham, Mohammad Ilyas, Ian A Cree, David Snead, Fayyaz Minhas, and Nasir M Rajpoot. Development and validation of a weakly supervised deep learning framework to predict the status of molecular pathways and key mutations in colorectal cancer from routine histology images: a retrospective study. *The Lancet Digital Health*, 3(12):e763–e772, 2021.
- Veenu Rani, Syed Tufael Nabi, Munish Kumar, Ajay Mittal, and Krishan Kumar. Self-supervised learning: A succinct review. *Archives of Computational Methods in Engineering*, pages 1–15, 2023.
- Ellery Wulczyn, David F Steiner, Melissa Moran, Markus Plass, Robert Reihs, Fraser Tan, Isabelle Flament-Auvigne, Trissia Brown, Peter Regitnig, Po-Hsuan Cameron Chen, et al. Interpretable survival prediction for colorectal cancer using deep learning. *NPJ digital medicine*, 4(1):71, 2021.
- Xiaofeng Jiang, Michael Hoffmeister, Hermann Brenner, Hannah Sophie Muti, Tanwei Yuan, Sebastian Foersch, Nicholas P West, Alexander Brobeil, Jitendra Jonnagaddala, Nicholas Hawkins, et al. End-to-end prognostication in colorectal cancer by deep learning: a retrospective, multicentre study. *The Lancet Digital Health*, 6(1):e33–e43, 2024.
- Han Xiao, Zongpeng Weng, Kaiyu Sun, Jingxian Shen, Jie Lin, Shuling Chen, Bin Li, Yiyu Shi Ming Kuang, Xinming Song, et al. Predicting 5-year recurrence risk in colorectal cancer: development and validation of a histology-based deep learning approach. *British Journal of Cancer*, pages 1–10, 2024.
- Richard J Chen, Ming Y Lu, Drew FK Williamson, Tiffany Y Chen, Jana Lipkova, Zahra Noor, Muhammad Shaban, Maha Shady, Mane Williams, Bumjin Joo, et al. Pan-cancer integrative histology-genomic analysis via multimodal deep learning. *Cancer Cell*, 40(8):865–878, 2022.
- Omar SM El Nahhas, Chiara ML Loeffler, Zunamys I Carrero, Marko van Treeck, Fiona R Kolbinger, Katherine J Hewitt, Hannah S Muti, Mara Graziani, Qinghe Zeng, Julien Calderaro, et al. Regression-based deep-learning predicts molecular biomarkers from pathology slides. *Nature communications*, 15(1):1253, 2024.
- Nicolas Coudray and Aristotelis Tsirigos. Deep learning links histology, molecular signatures and prognosis in cancer. *Nature Cancer*, 1(8):755–757, 2020.
- Richard J Chen, Tong Ding, Ming Y Lu, Drew FK Williamson, Guillaume Jaume, Andrew H Song, Bowen Chen, Andrew Zhang, Daniel Shao, Muhammad Shaban, et al. Towards a general-purpose foundation model for computational pathology. *Nature Medicine*, 30(3):850–862, 2024.
- Eugene Vorontsov, Alican Bozkurt, Adam Casson, George Shaikovski, Michal Zelechowski, Kristen Severson, Eric Zimmermann, James Hall, Neil Tenenholtz, Nicolo Fusi, et al. A foundation model for clinical-grade computational pathology and rare cancers detection. *Nature Medicine*, pages 1–12, 2024.
- Mingu Kang, Heon Song, Seonwook Park, Donggeun Yoo, and Sérgio Pereira. Benchmarking self-supervised learning on diverse pathology datasets. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3344–3354, 2023.
- Richard J Chen and Rahul G Krishnan. Self-supervised vision transformers learn visual concepts in histopathology. *arXiv preprint arXiv:2203.00585*, 2022.
- Sophia J Wagner, Daniel Reisenbüchler, Nicholas P West, Jan Moritz Niehues, Jiefu Zhu, Sebastian Foersch, Gregory Patrick Veldhuizen, Philip Quirke, Heike I Grabsch, Piet A van den Brandt, et al. Transformer-based biomarker prediction from colorectal cancer histology: A large-scale multicentric study. *Cancer Cell*, 41(9):1650–1661, 2023.

23. Adalberto Claudio Quiros, Nicolas Coudray, Anna Yeaton, Xinyu Yang, Bojing Liu, Hortense Le, Luis Chiriboga, Afreen Karimkhan, Navneet Narula, David A Moore, et al. Mapping the landscape of histomorphological cancer phenotypes using self-supervised learning on unannotated pathology slides. *Nature Communications*, 15(1):4596, 2024
24. Francesco Cisternino, Sara Ometto, Soumick Chatterjee, Edoardo Giacomuzzi, Adam P Levine, and Craig A Glastonbury. Self-supervised learning for characterising histomorphological diversity and spatial mra expression prediction across 23 human tissue types. *Nature Communications*, 15(1):5906, 2024.
25. Jure Zbontar, Li Jing, Ishan Misra, Yann LeCun, and Stéphane Deny. Barlow twins: Self-supervised809 learning via redundancy reduction. In *International Conference on Machine Learning*, pages 12310–810 12320. PMLR, 2021
26. Maximilian Ilse, Jakub Tomczak, and Max Welling. Attention-based deep multiple instance learning. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 2127–2136. PMLR, 10–15 Jul 2018
27. Chengkuan Chen, Ming Y Lu, Drew FK Williamson, Tiffany Y Chen, Andrew J Schaumberg, and Faisal Mahmood. Fast and scalable search of whole-slide images via self-supervised deep learning. *Nature Biomedical Engineering*, 6(12):1420–1434, 2022.
28. Corentin Gueréndel, Phil Arnold, and Ben Torben-Nielsen. Creating small but meaningful representations of digital pathology images. In Manfredo Atzori, Nikolay Burlutskiy, Francesco Ciompi, Zhang Li, Fayyaz Minhas, Henning Müller, Tingying Peng, Nasir Rajpoot, Ben Torben-Nielsen, Jeroen van der Laak, Mitko Veta, Yinyin Yuan, and Inti Zlobec, editors, *Proceedings of the MICCAI Workshop on Computational Pathology*, volume 156 of *Proceedings of Machine Learning Research*, pages 206–215. PMLR, 27 Sep 2021
29. The cancer genome atlas program (tcga). <https://www.cancer.gov/ccg/research/genome-sequencing/tcga>. Accessed: 2023-05-09.
30. Aimery de Gramont, Eric Van Cutsem, Hans-Joachim Schmoll, Josep Tabernero, Stephen Clarke, Malcolm J Moore, David Cunningham, Thomas H Cartwright, J Randolph Hecht, Fernando Rivera, et al. Bevacizumab plus oxaliplatin-based chemotherapy as adjuvant treatment for colon cancer (avant): a phase 3 randomised controlled trial. *The lancet oncology*, 13(12):1225–1233, 2012
31. Stéphanie M Zunder, Gabi W van Pelt, Hans J Gelderblom, Christoph Mancao, Hein Putter, Rob A Tollenaar, and Wilma E Mesker. Predictive potential of tumour-stroma ratio on benefit from adjuvant bevacizumab in high-risk stage ii and stage iii colon cancer. *British Journal of Cancer*, 119(2):164–169, 2018.
32. F Alexander Wolf, Fiona K Hamey, Mireya Plass, Jordi Solana, Joakim S Dahlin, Berthold Göttgens, Nikolaus Rajewsky, Lukas Simon, and Fabian J Theis. Paga: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome biology*, 20:1–9, 2019.
33. Herbert Hurwitz, Louis Fehrenbacher, William Novotny, Thomas Cartwright, John Hainsworth, William Heim, Jordan Berlin, Ari Baron, Susan Griffing, Eric Holmgren, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *New England journal of medicine*, 350(23):2335–2342, 2004.
34. Leonard B Saltz, Stephen Clarke, Eduardo Díaz-Rubio, Werner Scheithauer, Arie Figer, Ralph Wong, Sheryl Koski, Mikhail Lichinitser, Tsai-Shen Yang, Fernando Rivera, et al. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase iii study. *Journal of clinical oncology*, 26(12):2013–2019, 2008.
35. Fairouz F Kabbinavar, Julie Hambleton, Robert D Mass, Herbert I Hurwitz, Emily Bergsland, and Somnath Sarkar. Combined analysis of efficacy: the addition of bevacizumab to fluorouracil/leucovorin improves survival for patients with metastatic colorectal cancer. *Journal of clinical oncology*, 23(16):3706–3712, 2005
36. Tessa P Sandberg, Maaiké PME Stuart, Jan Oosting, Rob AEM Tollenaar, Cornelis FM Sier, and Wilma E Mesker. Increased expression of cancer-associated fibroblast markers at the invasive front and its association with tumor-stroma ratio in colorectal cancer. *BMC cancer*, 19(1):1–9, 2019
37. Gabi W van Pelt, Tessa P Sandberg, Hans Morreau, Hans Gelderblom, J Han JM van Krieken, Rob AEM Tollenaar, and Wilma E Mesker. The tumour–stroma ratio in colon cancer: the biological role and its prognostic impact. *Histopathology*, 73(2):197–206, 2018.
38. Wilma E Mesker, Jan Junggebur, Karoly Szuhai, Pieter de Heer, Hans Morreau, Hans J Tanke, and Rob AEM Tollenaar. The carcinoma–stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. *Analytical Cellular Pathology*, 29(5):387–398, 2007.
39. Peter Friedl and Stephanie Alexander. Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell*, 147(5):992–1009, 2011.
40. Alexander Bagaev, Nikita Kotlov, Krystle Nomie, Viktor Svekolkina, Azamat Gafurov, Olga Isaeva, Nikita Osokin, Ivan Kozlov, Felix Frenkel, Olga Gancharova, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer cell*, 39(6):845–865, 2021.
41. TJA Dekker, A Charehbili, VTHBM Smit, Peter ten Dijke, E Meershoek-Klein Kranenbarg, CJH van de Velde, JWR Nortier, RAEM Tollenaar, WE Mesker, and JR Kroep. Disorganised stroma determined on pre-treatment breast cancer biopsies is associated with poor response to neoadjuvant chemotherapy: Results from the neozotac trial. *Molecular oncology*, 9(6):1120–1128, 2015

42. MTA Strous, TKE Faes, ALHM Gubbels, RLA van der Linden, WE Mesker, K Bosscha, CM Bronkhorst, MLG Janssen-Heijnen, FJ Vogelaar, and AP de Bruïne. A high tumour-stroma ratio (tsr) in colon tumours and its metastatic lymph nodes predicts poor cancer-free survival and chemo resistance. *Clinical and Translational Oncology*, 24(6):1047–1058, 2022.
43. Stéphanie Zunder, Priscilla Van der Wilk, Hans Gelderblom, Tim Dekker, Christoph Mancao, Anna Kiialainen, Hein Putter, Rob Tollenaar, and Wilma Mesker. Stromal organization as predictive biomarker for the treatment of colon cancer with adjuvant bevacizumab; a post-hoc analysis of the avant trial. *Cellular Oncology*, 42:717–725, 2019
44. Hideki Ueno, Yukihide Kanemitsu, Shigeki Sekine, Megumi Ishiguro, Eisaku Ito, Yojiro Hashiguchi, Fukuo Kondo, Hideyuki Shimazaki, Yoshiki Kajiwara, Koichi Okamoto, et al. A multicenter study of the prognostic value of desmoplastic reaction categorization in stage ii colorectal cancer. *The American journal of surgical pathology*, 43(8):1015–1022, 2019
45. Hideki Ueno, Yoshiki Kajiwara, Yoichi Ajioka, Tamotsu Sugai, Shigeki Sekine, Megumi Ishiguro, Atsuo Takashima, and Yukihide Kanemitsu. Histopathological atlas of desmoplastic reaction characterization in colorectal cancer. *Japanese Journal of Clinical Oncology*, 51(6):1004–1012, 2021
46. Franck Pagès, Bernhard Mlecnik, Florence Marliot, Gabriela Bindea, Fang-Shu Ou, Carlo Bifulco, Alessandro Lugli, Inti Zlobec, Tilman T Rau, Martin D Berger, et al. International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study. *The Lancet*, 391(10135):2128–2139, 2018.
47. M Polack, MA Smit, GW van Pelt, AGH Roodvoets, E Meershoek-Klein Kranenbarg, H Putter, H Gelderblom, ASLP Crobach, V Terpstra, G Petrushevska, et al. Results from the united study: a multicenter study validating the prognostic effect of the tumor–stroma ratio in colon cancer. *ESMO open*, 9(4):102988, 2024.
48. Hideki Ueno, Yoshiki Kajiwara, Hideyuki Shimazaki, Eiji Shinto, Yojiro Hashiguchi, Kuniaki Nakanishi, Kazunari Maekawa, Yuka Katsurada, Takahiro Nakamura, Hidetaka Mochizuki, et al. New criteria for histologic grading of colorectal cancer. *The American journal of surgical pathology*, 36(2):193–201, 2012.
49. ID Nagtegaal, Mark Arends, and Robert Odze. Tumours of the colon and rectum: Who classification of tumours of the colon and rectum, tnm staging of carcinomas of the colon and rectum and the introduction. In *World Health Organization Classification of Tumours of the Digestive System*, pages 157–162. IARC press, 2019.
50. Jinru Shia, Nikolaus Schultz, Deborah Kuk, Efsevia Vakiani, Sumit Middha, Neil H Segal, Jaclyn F Hechtman, Michael F Berger, Zsofia K Stadler, Martin R Weiser, et al. Morphological characterization of colorectal cancers in the cancer genome atlas reveals distinct morphology–molecular associations: clinical and biological implications. *Modern pathology*, 30(4):599–609, 2017.
51. C Richard Boland and Ajay Goel. Microsatellite instability in colorectal cancer. *Gastroenterology*, 138(6):2073–2087, 2010
52. Paul Lochhead, Andrew T Chan, Reiko Nishihara, Charles S Fuchs, Andrew H Beck, Edward Giovannucci, and Shuji Ogino. Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. *Modern Pathology*, 28(1):14–29, 2015.
53. Napoleone Ferrara. Vegf and the quest for tumour angiogenesis factors. *Nature Reviews Cancer*, 2(10):795–803, 2002
54. Elisabeth J Kap, Petra Seibold, Swantje Richter, Dominique Scherer, Nina Habermann, Yesilda Balavarca, Lina Jansen, Natalia Becker, Katrin Pfütz, O Popanda, et al. Genetic variants in dna repair genes as potential predictive markers for oxaliplatin chemotherapy in colorectal cancer. *The pharmacogenomics journal*, 15(6):505–512, 2015
55. Gongmin Zhu, Lijiao Pei, Hongwei Xia, Qiulin Tang, and Feng Bi. Role of oncogenic kras in the prognosis, diagnosis and treatment of colorectal cancer. *Molecular cancer*, 20(1):1–17, 2021.
56. Shih-Ching Chang, Pei-Ching Lin, Jen-Kou Lin, Chien-Hsing Lin, Shung-Haur Yang, Wen-Yi Liang, Wei-Shone Chen, and Jeng-Kai Jiang. Mutation spectra of common cancer-associated genes in different phenotypes of colorectal carcinoma without distant metastasis. *Annals of surgical oncology*, 23:849–855, 2016.
57. Pavel Vodicka, Sona Vodenkova, Tomas Buchler, and Ludmila Vodickova. Dna repair capacity and response to treatment of colon cancer. *Pharmacogenomics*, 20(17):1225–1233, 2019
58. Arthur Liberzon, Chet Birger, Helga Thorvaldsdóttir, Mahmoud Ghandi, Jill P Mesirov, and Pablo Tamayo. The molecular signatures database hallmark gene set collection. *Cell systems*, 1(6):417–425, 2015.
59. Gabi W van Pelt, Sanne Kjær-Frifeldt, J Han JM van Krieken, Raed Al Dieri, Hans Morreau, Rob AEM Tollenaar, Flemming B Sørensen, and Wilma E Mesker. Scoring the tumor-stroma ratio in colon cancer: procedure and recommendations. *Virchows Archiv*, 473:405–412, 2018
60. Jakob Nikolas Kather, Johannes Krisam, Pornpimol Charoentong, Tom Luedde, Esther Herpel, Cleo-Aron Weis, Timo Gaiser, Alexander Marx, Nektarios A Valous, Dyke Ferber, et al. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS medicine*, 16(1):e1002730, 2019.
61. MB Loughrey, M Arends, I Brown, LJ Burgart, C Cunningham, JF Flejou, S Kakar, R Kirsch, M Kojima, A Lugli, et al. Colorectal cancer histopathology reporting guide. International Collaboration on Cancer Reporting. Sydney, Australia, 2020
62. Jianfang Liu, Tara Lichtenberg, Katherine A Hoadley, Laila M Poisson, Alexander J Lazar, Andrew D Cherniack, Albert J Kovatich, Christopher C Benz, Douglas A Levine, Adrian V Lee, et al. An integrated tcga pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell*, 173(2):400–416, 2018
63. Carmen J Allegra, Greg Yothers, Michael J O'Connell, Saima Sharif, Nicholas J Petrelli, Linda H Colangelo, James N Atkins, Thomas E Seay, Louis Fehrenbacher, Richard M Goldberg, et al. Phase iii trial assessing bevacizumab in stages ii and iii carcinoma of the colon: results of nsabp protocol c-08. *Journal of Clinical Oncology*, 29(1):11, 2011
64. Nicolas Coudray, Paolo Santiago Ocampo, Theodore Sakellaropoulos, Navneet Narula, Matija Snuderl, David Fenyo, Andre L. Moreira, Narges Razavian, and Aristotelis Tsirigos. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature Medicine*, 24, 2018.

65. Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley. Color transfer between images. *IEEE Computer Graphics and Applications*, 21, 2001
66. Vincent A Traag, Ludo Waltman, and Nees Jan Van Eck. From louvain to leiden: guaranteeing well-connected communities. *Scientific reports*, 9(1):5233, 2019
67. Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017
68. Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, 3(29), 2018.
69. Aravind Subramanian, Pablo Tamayo, Vamsi K Mootha, Sayan Mukherjee, Benjamin L Ebert, Michael A Gillette, Amanda Paulovich, Scott L Pomeroy, Todd R Golub, Eric S Lander, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550, 2005
70. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E. The immune landscape of cancer. *Immunity*. Apr 17;48(4):812-30. 2018

Methods

Ethics statement

Our research complies with all relevant ethical regulations. The study protocol was approved by Applied Bioinformatics Laboratories of New York University Grossman School of Medicine and the Department of Surgery of Leiden University Medical Center. The analyses were performed using anonymized archival material, not necessitating additional informed consents. Data from TCGA-COAD was open-accessed, ensuring patient anonymity without risk of patient identification. All institutions contributing annotated biospecimens provided documentation to the TCGA, and have obtained ethical approvals to use the sample and data according to the human subjects protection and data access policies in TCGA program. Archival material derived from the AVANT-trial (BO17920) was performed in accordance with the declaration of Helsinki [30,31]. Protocol approval was obtained from the local medical ethics review committees or institutional review boards at participating sites.

Study population

The TCGA-COAD dataset was used for training and extracting features and histologic patterns using SSL. This dataset consisted of 451 WSIs from 444 unique patients [29] with matched genetic and transcriptomic information. We excluded duplications and WSIs with erroneous resolution that were not suitable for the analyses (i.e. only several kilobytes in size). The final TCGA training set included 435 WSIs from 428 patients with a diagnosed pathological TNM-stage I-IV colon carcinoma (333 alive, 94 dead, and 1 missing vital status). We referred to the source population which TCGA-COAD representing as the "standard-of-care" group, contrasting it with the clinical trial external test data described below.

As external dataset we leveraged a study comprising 1213 colon cancer patients with available diagnostic H&E WSIs (one WSI per patient) as part of the clinical Bevacizumab-Avastin® adjuvant (AVANT) trial [30,31]. Bevacizumab, a VEGF monoclonal antibody, had initially been shown to improve the OS in patients with metastatic colon cancer when jointly used with the standard chemotherapy [33]. In the phase III AVANT trial, with an intent-to-treat population of 3451 patients, an open-label design was used [30]. Patients were randomly assigned in a 1:1:1 ratio to three different treatment regimens: FOLFOX-4 (intravenous 5-fluorouracil/folinic acid plus oxaliplatin), bevacizumab-FOLFOX-4, and bevacizumab-

XELOX (oral capecitabine plus intravenous oxaliplatin) [30]. The study aimed to investigate whether adding bevacizumab to the standard oxaliplatin-based adjuvant chemotherapy could improve DFS among patients with stage II-high risk and III colon cancer [30]. However, the trial was prematurely terminated due to the serious adverse effect on the patient's OS in the bevacizumab-treated group [30]. The AVANT trial was chosen also due to the previously studied potential correlation of VEGF, the stromal compartment (e.g. TSR), and patient prognosis [31,38,43]. For a detailed overview of the trial and patient characteristics, see Zunder et al [31].

Given the unique treatment regimen of bevacizumab and its adverse effect in non-metastatic colon cancer, as also proven by a predecesing clinical trial, the NSABP protocol C-08 trial [63], we decided to primarily validate OS prediction, trained in the TCGA-COAD, within the control group who only received FOLFOX-4 (without bevacizumab) and refer to it as the "AVANT control group". Subsequently, as several phase III trials had demonstrated that FOLFOX and XELOX are comparable in the context of metastatic colorectal cancer [33–35], we thus combined the bevacizumab+FOLFOX-4 and bevacizumab+XELOX groups into a unified "AVANT-experimental group" and conducted a separate analysis to predict OS within the bevacizumab-treated patients.

Data pre-processing

Tissue segmentation and image tiling

We used the preprocessing methods described in our previous study [64]. Tissue areas in WSIs were segmented against background at 10x magnification level (pixel size approximate 1.0 μ m). WSIs were divided into non-overlapping image tiles of size 224x224 pixels. The selection of a 10x magnification was based on two key considerations. Firstly, it aligns with the standard magnification utilized by pathologists during microscopic assessments in clinical practice. Secondly, we conducted visual inspections of tiles at 20x, 10x, and 5x magnifications, and found that tiles at 10x magnification provided an optimal balance of capturing sufficient detail while also offering a reasonably sized overview of the morphological structure. In addition, to overcome the variability of color stains from different scan facilities in the TCGA and AVANT cohorts, we further applied the color normalization [65]. In total, we obtained 1,117,796 tiles in the TCGA training set (i.e. TCGA-COAD), and 4,827,055 tiles in the AVANT external test set (AVANT-COAD, consisting of the standard and experimental treatment groups).

Extracting image features using Barlow Twins

We trained the SSL Barlow Twins feature extractor based on 250,000 image tiles randomly selected from the TCGA-COAD dataset. The Barlow Twins extracted unique latent vectors (128 dimension) from the preprocessed image tiles in TCGA. The model is based on ResNet-like architecture consisted of several ResNet layers and one self-attention layer [23]. In essence, the Barlow Twins calculates the cross-correlation matrix between the embedding outputs of two identical twin networks, both fed with distorted versions of the same image tile [25]. It is optimized to make the correlation matrix close to the identity matrix [25]. We used the batch size of 64 trained on a single NVIDIA® Tesla V100 GPU for 60 epochs. The Barlow Twins feature extractor was frozen after the training and used to project image tiles into latent representation in the entire TCGA training set. To facilitate the downstream analysis, we also applied five-fold CV partition in the TCGA-COAD training set on patient level balanced the on AJCC TNM stage, OS outcomes (i.e. death or censor), and binned survival time categories.

Leiden community detection algorithm

The Leiden clustering algorithm is a graph-based clustering algorithm that aims to identify distinct communities or clusters within graph data [66]. In brief, it optimizes the modularity function (Equation 1), in such way to maximize the difference between the actual number of edges and the expected number of edges in a community [66]. This modularity function also includes a resolution parameter γ , with higher values leading to more clusters and lower values leading to fewer clusters. Leiden clustering is initiated by assigning each node in the graph to its own individual cluster, treating them as separate communities, then iteratively optimizes the modularity function by moving a node from its current cluster to a neighboring cluster or by merging clusters until the algorithm converged and revealing distinct communities or clusters. We employed the Leiden clustering in a particular workflow. We began by constructing a neighborhood graph using the $K=250$ nearest neighbors from a pool of 200,000 randomly selected latent image vectors from a training set. Subsequently, we applied the Leiden algorithm to identify clusters within this neighborhood graph. These cluster labels were then propagated to each individual image tile across the entire dataset, once more utilizing the K-nearest neighbors approach.

$$\mathcal{H} = \frac{1}{2m} \sum_c \left(e_c - \gamma \frac{K_c^2}{2m} \right)$$

Equation 1. Modularity measures the difference between the actual number of edges in a community and the expected number of such edges. e_c denotes the actual number of edges in community c and the expected number of edges is expressed as $\frac{K_c^2}{2m}$, where K_c is the sum of the edges of the nodes in community c and m is the total number of edges in the network. γ is the resolution parameter, with higher values leading to more clusters and lower values leads to fewer clusters.

Identification of HPCs

We identified HPCs as clusters obtained from the Leiden community detection algorithm operated on 128-dimensional image features extracted through the Barlow Twins encoder. The Leiden method was also used for quality control to eliminate artifacts and underfocused image tiles. Quality control was carried out after training the Barlow Twins and extracting image features. In the overall process, we initially generated a substantial number of Leiden clusters. Next, we visually examined sample image tiles from each cluster and removed clusters exhibiting artifacts such as air bubbles, foreign objects, etc., or those containing under-focused images. In particular, within the training set of a randomly selected CV fold (fold 0), we randomly selected 200,000 latent image vectors to generate Leiden clusters. We obtained 125 clusters at a high resolution of $\gamma=6$. The Leiden labels were propagated into the entire TCGA set using again the KNN methods. Next, we inspected randomly selected sample tiles (N=32) from each cluster and identified 12 clusters containing predominately artifacts or underfocused images. Image tiles labeled by these 12 clusters were subsequently removed from the further analyses. The HPCs were newly derived in this cleaned dataset by re-running the Leiden clustering. Optimization of the HPCs was conducted using primarily unsupervised methods and secondarily confirmed using supervised methods. Importantly, both the two approaches converged on the same optimal Leiden resolution, as elaborated below.

Optimization of HPCs using unsupervised methods

Leiden resolutions (i.e. $\gamma=0.4, 0.7, 1.0, 1.5, 2.5$, and 3.0) were optimized using three unsupervised statistical tests: the Disruption score, Silhouette score, and Daves-Boundin index. Due to the potentially high variance in data from the different institutions in the TCGA, all three scores were weighted by the mean percentage of the institution presence in each cluster. We consistently identified the optimal Leiden resolution as 1.5 through the three aforementioned statistical tests (Supplementary Figure 1a).

Optimization of HPCs and the prediction of OS using Cox regressions with L2 regularization

To identify optimal HPC configurations and their associations with patient OS, we trained L2 regularized Cox regressions for OS prediction using 5-fold CV. The Cox regressions were trained separately among standard-of-care colon cancer patients (i.e. TCGA-COAD) and among the AVANT-experimental group.

OS prediction in the standard-of-care group

Prediction of OS from HPCs among the standard-of-care group was trained within TCGA-COAD using 5-fold CV and tested in the independent AVANT control group. Specifically, for each CV fold within TCGA-COAD, we began by generating a range 587 of Leiden clustering configurations at various resolutions, including gamma values of $0.4, 0.7, 1.0, 1.5, 2.5$, and 3.0 , using the method described earlier. Next, at each Leiden resolution, we calculated the compositional representation of HPCs for each WSI (see Main Figure 1c), followed by a center-log-ratio transformation (Equation 2). This transformation was designed to mitigate inter-dependencies among HPCs, ensuring that the independence assumptions required for subsequent Cox regression analysis were met. The L2 regularized Cox regressions were then trained at the patient level, with one WSI per patient considered. At each Leiden resolution, we performed a multivariable L2 regularized Cox regression, incorporating all center-log-ratio-transformed HPCs specific to that resolution. We finetuned L2 regularizer (alpha) through an iterative process involving 50 steps, spanning the alpha range from 10^{-4} to 10^4 . This sequence of steps was repeated across all five cross-validation folds. The optimal Leiden resolution and L2 regularizer was selected based the CV C-index.

Through this optimization process, the optimal Leiden resolution was determined to be 1.5 , and the L2 regularizer alpha was fixed at 0.1842 . Of the note, this optimal Leiden resolution of 1.5 concurred with the

result from the unsupervised approaches. An HPC-based classifier was determined by the median predicted hazard obtained in the TCGA-COAD. Once the Cox model is optimized, we evaluated the final model performance in the external test set consisting of the AVANT-standard care group. First, the 47 HPCs were integrated into the AVANT-standard care group by employing the K-nearest neighbors method (K=250), where each AVANT tile's HPC label was determined based on the majority votes from its nearest neighbors in the TCGA training set. Next, the trained Cox model, with optimized regularization and parameter estimates for the HPCs, was then applied to the AVANT-standard care group to test the prediction of OS (Supplementary Figure 3b).

Furthermore, employing the same CV method, we trained a clinical baseline model on OS in TCGA-COAD using L2 regularized Cox regression incorporating age, sex, TNM staging, and TSR as predictors. We observed a c-index of 0.58 (bootstrap 95% CI=0.49-0.67) in the independent AVANT-standard care group and the model-based risk classifier did not reach the statistical significance level (Supplementary Figure 3c). This baseline model illustrates a simulation of decision-making in clinical practice as control, using the most readily available and relevant clinical and demographic variables. Our HPC-based model outperformed this clinical baseline model. In addition, we explored whether the HPC-based classifier add additional prognostic value to the existing important clinical predictors. We fitted an ordinary multivariable Cox regression within the external AVANT-standard care group, including HPC-based classifier, sex, age, TSR, and AJCC TNM stage (Main Figure 4a).

$$\text{clr}(x_i) = \ln \left(\frac{x_i}{g} \right), x_i = \frac{\text{counting of tiles in } C_i}{\text{total number of tiles in each WSI}}, \quad g = \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}}$$

Equation 2. The center-log-ratio transformation (clr) calculates the natural logarithm of the ratio (x_i) of compositional data for a specific cluster (C_i) to the geometric mean (g) of the compositional data across all clusters. The compositional data (x_i) is obtained by dividing the counting of tiles in cluster C_i by the total number of tiles in each WSI. The geometric mean (g) is computed by multiplying the compositional data values (x_i) for each cluster (C_i), ranging from $i = 1$ to n (where n is the total number of clusters), and raising the resulting product to the power of $\frac{1}{n}$.

OS prediction in the AVANT-experimental group

Bevacizumab, a unique intervention investigated in the AVANT trial, was unlikely accessed by patients from the TCGA-COAD cohort. Considering the significant poor prognostic outcome associated with bevacizumab in non-metastatic colorectal cancer, we postulated that the relationship between HPCs on OS might be influenced by this intervention. Furthermore, we hypothesized that estimates of HPCs on OS trained from the general COAD population, such as TCGA, might not be applicable to bevacizumab-treated patients.

To address these hypotheses, we conducted a separate 5-fold CV estimating the relationship between HPCs and OS within the bevacizumab-treated patients (AVANT-experimental group). Similarly, we generated 5-fold train-validation split stratified by TNM stage and survival time. We used the same sets of HPCs obtained at the previously optimized Leiden 1.5 resolution. Similarly, we modelled the center-log-transformed compositional data of HPCs on OS using Cox regressions with L2 regularization. We fine-tuned the L2 regularizer specifically for AVANT-experimental group. The model performance of HPCs on OS prediction was evaluated using c-index in the 5-fold CV validation sets (Supplementary Figure 3d).

To understand the importance of each HPC on OS, we employed SHAP values [67]. SHAP values measure the marginal contribution of a HPC towards the predicted OS, considering all possible combinations of features. We highlighted top 20 important HPCs favoring and hindering survival for both standard-of-care and AVANT-experimental treated COAD patients.

Interpretation of HPCs*Plotting HPCs using UMAP and PAGA plot*

We applied UMAP dimensionality reduction [68] to TCGA-COAD tile vector representations (128 dimensional vectors), color-coded by 47 HPC IDs using optimized Leiden clustering configuration from the prior step. Next, we generated a PAGA plot where each HPC is a node connected by lines based on vector similarity. Pie charts within each HPC node showed annotated tissue type percentages annotated independently by three experts (ASLPC, JHJMvK, and MP) (see below). We analyzed this plot to identify and describe significant interconnected clusters (Main Figure 2b). Based on the PAGA plot, we also defined

the "super-clusters" (Main Figure 2c) according to the interconnections among HPCs 646 and tissue composition.

Pathologist assessment of HPCs

Histopathological assessment and characterization of HPCs

The histomorphological features of 47 HPCs derived in TCGA training set were comprehensively and independently analyzed by two expert pathologists (ASLPC and JHJMvK) and one medical researcher (MP). The analysis was exclusively conducted using image tiles and the assessors were kept blinded to results from other analyses. Specifically, we randomly selected a set of 32 tiles from each HPC within TCGA and each individual tile was examined with specific focus on tumor epithelium, tumor stroma, and immune cells. Attention was also paid to general tumor differentiation grade, tumor stromal amount, stromal classification, i.e. aligned or disorganized, stromal neovascularization or other notable patterns (e.g. dysplastic tissue, fatty tissue, muscle tissue fibers, blood vessels, erythrocytes, etc.) [3,43,59]. Each assessor evaluated the tissue composition based on 32 randomly selected tiles per HPC, providing an average tissue composition for each HPC. In cases of conflicting tissue annotations, the assessors carried out discussions to achieve consensus. In addition, a short label was given to each HPC based on tissue annotation either

according to the first and second predominant tissue types and patterns (e.g. "aligned stroma – well differentiated tumor epithelium"), or as the single most dominant tissue type covering $\geq 70\%$ area of an average tile (e.g. "necrotic"). The tissue description, composition, and short labels of all 47 HPCs were displayed in Supplementary Table 1.

Assessment of HPC consistency within and across TCGA and AVANT cohorts

We performed various qualitative and quantitative analyses to evaluate the within-cluster and between-cluster heterogeneity of HPCs, as well as their transferability from TCGA to AVANT cohorts. Initially, qualitative visual assessments were independently conducted by three experts (ASLPC, JHJMvK, and MP) to evaluate the concordance within and discordance between HPCs. This evaluation followed the established protocol for tissue type analysis, utilizing the previously randomly selected 32 tiles per HPC. The assessors noted consistent histomorphological patterns within HPCs and diverse patterns across various HPCs.

In addition, we conducted quantitative objective tests within TCGA and AVANT tiles separately to determine if the learned morphological features of the HPCs could be replicated by human eyes. We displayed three rows, each containing five tiles, which is referred to as a "question". Among these rows, two belonged to the same HPC, while the third row, also referred to as the "odd HPC", were tiles randomly selected from a different HPC. All example tiles were randomly selected within each HPC. The researcher (MP) was required to identify the "odd HPC", and upon doing so, the next question was presented (example shown in Supplementary Figure 2). In order to calculate the success rate, we ran 50 questions per HPC and we repeated the experiment for all 47 HPCs. The test was conducted on a closed website accessible through login and was conducted separately for tiles randomly selected from TCGA and AVANT. The conducted experiment exhibited an average completion time of under 10 seconds per question. The success rate was determined for each HPC by dividing the number of incorrect answers by the total number of questions (50 questions per HPC). We hypothesized that most HPCs would be distinguishable. However, some HPCs may be challenging to differentiate due to tissue similarities, either within the same super-cluster or through similar tissue morphology (e.g., muscle tissue and tumor stroma). The results showed that HPCs in close proximity to each other in the PAGA or belonging to the same super-cluster were indeed more prone to being mistaken.

To validate whether HPCs derived from the TCGA training set generalized well in the external AVANT test set, three assessors independently carried out visual examination of tissue characteristics in 32 randomly selected tiles from each HPC in both the AVANT trial and the TCGA set, comparing their histomorphological features regarding tissue types and composition established in the previous step. Assessors confirmed that the consistent and meaningful patterns, as learned by the Barlow Twins, could be replicated across other clinical cohorts like the AVANT trial. To validate our observation, we compared the objective test results between TCGA and AVANT, focusing on the overlap in both misclassified and correctly classified HPCs across the datasets.

Linkage between survival-related HPCs, immune landscape, and gene expression data

We calculated Spearman correlations between HPCs and RNASeq-derived immune features from TCGA-COAD data, correcting for false discovery rate (0.05) using Benjamini-Hochberg correction. We created

bi-hierarchical clusters for immune features and top 20 significant HPCs linked to OS in both standard-of-care and AVANT-experimental groups. Clustering was based on correlation coefficients using Euclidean distance and Ward's aggregation.

In addition, we performed GSEA to explore the potential associations between HPCs and the major cancer hallmark pathways, as previously described [69]. The analysis was conducted within 282 TCGA-COAD patients with available gene expression data for 20,530 genes. For each HPC, we calculated the Spearman correlation coefficients with the expression data of these 20,530 genes and sorted the resulting correlation coefficients to rank the genes in a descending order. Subsequently, enrichment scores for 50 major cancer hallmark pathways, as defined by the "MSigDB_Hallmark_2020" gene set [58], were computed for each HPC based on the generated ranked gene list. A positive enrichment score indicated higher composition of a HPC was associated with gene enrichment in a pathway and a negative value

suggested higher composition of a HPC was associated with underrepresentation of a pathway. We plotted the cancer hallmark enrichment scores for the top 20 important HPCs related to OS, for both the general and experimental treatment groups. We set a significance level of 0.01 of the false discovery q value. The analysis was conducted using the gseapy (1.0.4) Python package. Two-way hierarchical clustering was based on the Euclidean distance of the enrichment scores and Ward's aggregation method.

Lastly, we created a Supplementary Table 2 to grant a detailed overview of the significant HPCs associated with OS, including the direction of the association (worse or better OS) for both treatment groups (TCGA and AVANT standard, or AVANT-experimental) and in-depth histopathological analyses potentially explaining these associations with appropriate references.

Data and Code Availability Statement

The publicly available TCGA-COAD can be accessed at Genomic Data Commons portal (<https://gdc.cancer.gov/>). The immune landscape signatures for TCGA-COAD is available in Thorsson et al. [70]. The AVANT data that support the findings of this study are available from Genentech Inc., Roche. However, access to these data is restricted as they were used under license for the current study and are not for commercial use to protect patient privacy. Data may however be available from the authors upon request

for research purposes with permission from Genentech Inc., Roche. Data analyses are based upon publicly available Python software packages and codes are available from our previous publication [23] (<https://github.com/AdalbertoCq/Histomorphological-Phenotype-Learning>). The remaining data are available within the Article, Supplementary Information or Source Data file. Source code and Python software packages are available from our previous publication [23] (<https://github.com/AdalbertoCq/Histomorphological-Phenotype-Learning>).

- Chapter 6

Supplementary Material

Supplementary Table 1. Tissue composition for all 47 HPCs

HPC Super-cluster	Label	General description	Tumor epithelium	Tumor stroma	Immune cells	Other	HPC tissue type composition
0	Tumor stroma	Mostly disorganized stroma, sometimes some aligned strands and sometimes some muscle tissue, immune cells in between, with a few isolated tumor cells	A couple isolated tumor cells	Primarily tumor stroma on tiles. Stroma-high tiles, more disorganized than aligned stroma	Some immune cells in hotspots, immune-high overall	Some muscle tissue	Stroma 70%, Muscle tissue 20%, Immune cells 10%
1	Tumor stroma	Loose, edematous and disorganized stroma with background, sometimes vessels and an artefact. Most likely not tumor-induced stroma but healthy serosa as there is also some fatty tissue	A couple isolated tumor cells	Primarily loose, edematous and disorganized stroma-high tiles with background, sometimes vessels with some erythrocytes. Most likely not tumor-induced stroma but healthy serosa looking at pattern, and there is also some fatty tissue	A couple immune cells, immune-low	Some fatty tissue	Stroma 80%, Fatty tissue 10%, Background 10%
2	Tumor stroma	Aligned stroma Mostly up until 50% stroma in aligned, straight strands. Well differentiated tumor in between, with a little dirty necrosis	Mostly well differentiated tumor epithelial cells	Mostly aligned, straight strands of tumor stroma on the stroma-low tiles	A couple immune cells, and some (dirty) necrosis, immune-low	N/A	Stroma 40%, Tumor epithelial cells 60%
3	Tumor epithelium	Moderately differentiated tumor - well differentiated tumor	Much well-moderately differentiated tumor epithelium	Little tumor stroma, stroma-low tiles	A couple immune cells, and some (dirty) necrosis, immune-low	Some high-grade dysplasia, some background. A couple vessels with some erythrocytes	Tumor epithelial cells 70%, Stroma 10%, Necrosis 10%, Background 10%
4	Tumor epithelium	Moderately differentiated tumor - aligned in between strands of stroma, some immune cells and a few spots of dirty necrosis	Well-moderately differentiated tumor	Some strands of aligned tumor stroma, stroma-low tiles	Some immune cells, and a few spots of dirty necrosis, immune-high overall	N/A	Tumor epithelial cells 60%, Stroma 30%, Immune cells 10%
5	Necrosis	Necrotic Mostly (dirty) necrosis, some tumor epithelial cells and some vessels, with background	A couple isolated tumor cells	Some aligned tumor stroma strands, stroma-containing low tiles	Primarily tiles stroma-containing (dirty) necrosis	Some vessels with dirty necrosis and some background or	Necrosis 80%, Tumor epithelial cells 10%, Background 10%

			similarly fatty tissue or mucin	
6	Healthy and inflamed dysplastic colon tissue	Healthy and some dysplastic colon tissue, with many immune cells in proliferative and often vascularized stroma, aspect of inflammation	No tumor epithelium No tumor stroma, only inflamed and vascularized 'healthy' stroma	Many immune cells Healthy or dysplastic colon tissue 70%, Immune cells 30%
7	Necrosis	Avital tumor with necrosis	Mostly moderately differentiated tumor epithelium	Some immune cells Much background Tumor epithelial cells 40%, Background 30%, Stroma 10%, Necrosis 20%
8	Tumor epithelium	Mixed tumor epithelial - infiltrated stroma	Mostly moderately differentiated tumor epithelium	Part immune cells, immune-high Some background and spots of erythrocytes Tumor epithelial cells 40%, Stroma 20%, Immune cells 20%, Background 10%, Erythrocytes 10%
9	Tumor epithelium	Moderately differentiated tumor - aligned stroma	Mostly well-moderately differentiated tumor epithelium	A few immune cells, immune-low Some background, some mucin Tumor epithelial cells 70%, Stroma 20%, Background 10%
10	Muscle tissue (longitudinal fibers)	Muscle tissue fiber strands as well as similar looking aligned strands of stroma. Some vessels with erythrocytes, and immune cells in between	No tumor epithelium	Some tiles contain strands of aligned stroma, then the tile is stroma-high Mostly muscle tissue, stroma longitudinal fibers and some vessels Muscle 60%, Stroma 30%, Immune cells 10%
11	Tumor stroma	Vascularized and infiltrated stroma	Few isolated tumor epithelial cells	High amount of immune cells, potentially part of lymphoid tissue. Immune-high A couple tiles with a bit background or fatty tissue Stroma 70%, Immune cells 30%

12	Mucinous	Mucinous tumor	Some background in otherwise mucinous tumor containing tiles. Notably high contrast in colors, little to no stroma	Predominantly mucinous tumor epithelium	Some stroma, disorganized spots. Stroma-low	Some immune cells, immune-low	Mostly mucus, some background and one tile with fatty tissue	Tumor epithelial cells 40%, Mucin 30%, Background 20%, Stroma 10%
13	Immune cells	Infiltrated stroma	The mostly aligned tumor stroma is highly infiltrated with immune cells, with some well-moderately differentiated tumor epithelial tissue and some background	Well differentiated tumor epithelium, high-grade dysplasia	Mostly aligned, straight strands of tumor stroma on the stroma-low tiles	High amount of immune cells in stroma, potentially part of lymphoid tissue. Also some dirty necrosis. Immune-high	N/A	Immune cells 30%, Stroma 30%, Tumor epithelial cells 30%, Necrosis 10%
14	Mucinous	Mucinous tumor	Tumor epithelial cells with some stroma, on tiles consisting approx. half of mucin, some background	Predominantly mucinous tumor epithelium	A part disorganized tumor stroma, some aligned strands. Officially stroma-high tiles, excluding the mucus	A few immune cells, immune-low	Mostly mucus, some background	Mucin 40%, Tumor epithelial cells 30%, Stroma 20%, Background 10%
15	Fatty tissue	Fatty tissue	Much fatty tissue, some vessels (also round shape), strands of mostly aligned stroma, some immune cells	No tumor epithelium	Strands of aligned stroma, potentially not only tumor-induced but some healthy as well	A part immune cells, some lymphoid tissue, in stroma then immune-high overall	Predominantly fatty tissue on the tiles, and some other round shapes like vessels, some with erythrocytes, and a spot of mucin	Fatty tissue 60%, Stroma 30%, Immune cells 10%
16	Tumor epithelium	Well-differentiated tumor	Mostly well-moderately differentiated tumor, some stroma in between, and dirty necrosis	Well differentiated tumor epithelium, high-grade dysplasia	Some aligned strands of stroma, tiles are stroma-low	Some immune cells, some dirty necrosis, immune-low	Some background	Tumor epithelial cells 60%, Stroma 20%, Necrosis 10%, Background 10%
17	Tumor stroma	Infiltrated aligned stroma-high	Some immune cells and tumor epithelial cells in the much most prevalent aligned strands of stroma (stroma-high). Sometimes a neural bundle and a few spots more similar to muscle fibers	Some parts tumor epithelium, well-differentiated	Predominantly tumor stroma, mostly aligned strands on stroma-high tiles	Part immune cells background, a couple tiles with some muscle tissue. A sporadic neural bundle	Some spots	Stroma 60%, Immune cells 20%, Tumor epithelial cells 10%, Muscle tissue 10%
18	Immune cells	Infiltrated stroma	Cluster abundant in cells. Much stroma, mostly disorganized, with many immune cells and some isolated tumor cells	Part isolated tumor epithelial cells	Predominantly tumor stroma, mostly disorganized stroma-high tiles	Many immune cells infiltrated in tumor stroma, immune-high	N/A	Stroma 60%, Immune cells 30%, Tumor epithelial cells 10%

19	Tumor stroma	Edematous disorganized stroma, almost no immune cells but with tumor epithelium, some necrosis	Much edematous tumor stroma, mostly disorganized stroma-high tiles	Some parts tumor epithelium, well-differentiated	Few immune cells, some (dirty) necrosis, immune-low	A few spots mucin	Stroma 70%, Tumor epithelial cells 30%
20	Tumor epithelium	Disorganized and vascularized stroma-low	Mostly disorganized stroma, stroma-low tiles, some vascularization	Well-moderately differentiated tumor	A part immune cells, some (dirty) necrosis, immune-low mostly	Some background mucin, a slight signet cell component	Tumor epithelial cells 50%, Stroma 30%, Immune cells 10%, Background 10%
21	Tumor stroma	Vascularized stroma	Aligned strands of much loose stroma with neovascularization (or small white shapes similar to vessels), sometimes some muscle fibers, with immune cells and isolated tumor cells in between	A couple isolated tumor cells	Predominantly aligned tumor stroma strands, infiltrated in tumor stroma, some parts vascularized (or other round shapes)	Some background, a few spots of fatty tissue. A couple tiles muscle tissue in stead of stroma	Stroma 70%, Immune cells 10%, Tumor epithelial cells 10%, Muscle tissue 10%
22	Immune cells	Infiltrated stroma	Many immune cells in tumor-induced and healthy stroma, with some tumor or healthy colon tissue and some background	Some well differentiated tumor epithelium, high-grade dysplasia	Healthy and tumor stroma, which is stroma-high	High number of immune cells, some necrosis, immune-high	Immune cells 30%, Stroma 30%, Tumor epithelial cells 20%, Background 10%, Healthy and dysplastic colon tissue 10%
23	Healthy and dysplastic colon tissue	Dysplastic colon tissue	Mostly dysplastic colon tissue (adenoma), with many immune cells in between and much background and mucin	A couple isolated tumor cells	Dysplastic stroma	Many immune cells in dysplastic stroma	Healthy and dysplastic colon tissue 40%, Immune cells 20%, Mucin 20%, Background 20%
24	Muscle tissue	Vessel-like stroma/muscle	Approx. half of the tiles contain background, some mucin, the other half is stroma and/or muscle tissue of vessels; large, white round shapes. There is some dirty necrosis	A couple isolated tumor cells	Stroma in the shape of vessels, these are mostly aligned tumor stroma and stroma-high	Muscle tissue, mostly the walls from vessels and background, a spot of erythrocytes within these vessels, some mucin	Muscle tissue 30%, Stroma 30%, Background 30%, Mucin 10%

25	Necrosis	Moderately differentiated tumor - necrotic	Moderately differentiated tumor epithelial with (dirty) necrosis, and some stroma in between. Some background and mucin as well	Mostly well differentiated tumor epithelial cells	Some strands of aligned tumor stroma, stroma-low tiles	Part immune cells in (dirty) necrosis, otherwise immune-low	In general busy and disorganized tiles, round shapes, some background, some mucin	Tumor epithelial cells 50%, Necrosis 20%, Stroma 10%, Mucin 10%, Background 10%
26	Tumor epithelium	Infiltrated poor-differentiated tumor	Cluster rich in cells. Tiles undifferentiated with solid growing poor-undifferentiated tumor, others immune cells and some stroma in between. Potentially microsatellite instability	Tiles with solid growing poor-undifferentiated epithelium	A couple strands of aligned tumor stroma-low tiles	Many immune cells, some tiles only containing immune cells, immune-high	Cluster rich in cells, one tile a slight signet cell component, a spot background	Tumor epithelial cells 80%, Immune cells 10%, Stroma 10%
27	Tumor epithelium	Moderately differentiated tumor - aligned stroma	Some background or mucin. Well-moderately differentiated tumor - aligned epithelial cells in some aligned strands of stroma and some immune cells	Well differentiated tumor epithelium, high-grade dysplasia	Some strands of aligned tumor stroma, stroma-low tiles	Some immune cells, also in (dirty) necrosis, overall immune-high	Some background	Tumor epithelial cells 50%, Stroma 30%, Immune cells 10%, Background 10%
28	Muscle tissue	Muscle (longitudinal fibers)	Mostly muscle tissue, longitudinal sections, sometimes aligned stroma strands and some background. A few tumor epithelial cells and neural bundles	A couple isolated tumor cells	A couple tiles contain strands of aligned stroma, these are stroma-high	A few immune cells, immune-low in stroma tiles	Predominantly longitudinal fibers, some background or a spot of fatty tissue. A sporadic neural bundle	Muscle tissue 70%, Stroma 20%, Background 10%
29	Muscle tissue	Muscle (longitudinal fibers)	Mostly longitudinal sectioned muscle tissue and/or stroma strands, a bit edematous, some vessels, a neural bundle, not many immune cells	A couple isolated tumor cells	A couple tiles contain strands of aligned stroma, then the tile is stroma-high	A few immune cells, immune-low in stroma tiles	Predominantly longitudinal fibers, some background or a spot of fatty tissue. A sporadic neural bundle. Small vessels often in between	Muscle tissue 80%, Stroma 20%
30	Immune cells	Mixed, immune-high	Aligned stroma strands and well-moderately differentiated tumor epithelial cells, immune cells in between, some dirty necrosis	Well differentiated tumor epithelium, high-grade dysplasia	Aligned stroma strands on mostly stroma-low tiles	Many immune cells infiltrated in tumor stroma, some spots of (dirty) necrosis	N/A	Tumor epithelial cells 30%, Stroma 30%, Immune cells 30%, Necrosis 10%

31	Immune cells Mixed, infiltrated stroma	Often aligned stroma with many immune cells and moderately differentiated tumor epithelium. Often some background, some necrosis	Well differentiated tumor epithelium, high-grade dysplasia	Aligned, loose stroma strands on mostly stroma-low tiles	Part immune cells often infiltrated in tumor stroma, immune-high	Some background, a couple spots of mucin or fatty tissue	Stroma 40%, Immune cells 20%, Tumor epithelial cells 20%, Background 20%
32	Muscle tissueMuscle (axial section)	Axial sections of muscle tissue, some disorganized strands of stroma and some background. Also, some tiles that contain erythrocytes, similar looking to the muscle fiber sections	No tumor epithelium	No tumor stroma	A few immune cells	Tiles contain mostly muscle tissue, with the fibers in axial sections. Some are part of a vessel wall, with sometimes also erythrocytes in vessels. Some background	Muscle tissue 70%, Erythrocytes 10%, Stroma 10%, Background 10%
33	Muscle tissueInvasive tumor buds in muscle, fatty tissue	Muscle fibers in buds in muscle, longitudinal sections with fatty tissue, some vessels and only few immune cells. Some tumor budding present	Some tumor buds	A couple tiles contain strands of aligned stroma, similar to muscle tissue fibers, then the tile is stroma-high	A couple immune cells, immune-low in stroma tiles	Fatty tissue, in between loose bundles of longitudinal muscle fibers, muscle tissue as part of vessels	Muscle tissue 50%, Stroma 20%, Fatty tissue 20%, Background 10%
34	Muscle tissueMuscle (axial section)	Mostly muscle tissue in axial plane, some tiles disorganized tumor stroma, with some vessels; overall few nuclei or immune cells	A couple isolated tumor cells	A couple tiles contain strands of disorganized tumor stroma, these are tiles stroma-high	A few immune cells, immune-low in stroma tiles	Predominantly axial sections of muscle tissue, a few spots of background or mucus	Muscle tissue 70%, Stroma 30%
35	Immune cells Immune cells	Mostly immune cells in stroma or lymphoid tissue, with some tumor epithelial cells or necrosis	A couple isolated tumor cells	Tumor- and healthy stroma, stroma-high tiles	Highly infiltrated (tumor) stroma or necrosis, sometimes part of lymphoid tissue, immune-high	Some background, a few spots of mucin or fatty tissue, a few spots of dysplastic colon tissue	Immune cells 60%, Stroma 20%, Tumor epithelial cells 10%, Background 10%
36	Necrosis Avital tumor with necrosis	Loose and pieces of tissue, mostly avital tumor epithelium or necrosis, in tiles with much background, some strands of stroma	Loose and pieces of tumor epithelium, well-moderately differentiated	Some spots of stroma, mostly disorganized, stroma-low tiles	Immune cells in much present necrosis	Much background, a few spots of mucin	Necrosis 30%, Tumor epithelial cells 30%, Background 30%, Stroma 10%

37	Healthy and dysplastic colon tissue	High-grade dysplasia - adenoma	Adenoma or well differentiated tumor, with much background, some mucin, and few spots of dirty necrosis	Adenoma, high-grade dysplasia or well differentiated tumor epithelium	A few strands of aligned (tumor) stroma, stroma-low	Some immune cells in dysplastic and tumor stroma, a spot of dirty necrosis	Some background and/or mucin	Healthy or dysplastic colon tissue 60%, Stroma 10%, Immune cells 10%, Background 10%, Mucin 10%
38	Mucinous	Mucinous tumor stroma	Mucin, mostly from tumors with disorganized tumor stroma, some background	A couple isolated tumor cells	Some disorganized strands of tumor stroma in between the mucin, tiles are stroma-high	A few immune cells, immune-low	Predominantly mucin	Mucin 80%, Stroma 20%
39	Healthy and dysplastic colon tissue	Healthy colon tissue	Mostly clear, healthy colon tissue with mucin in goblet cells and the villi, a small component dysplasia, some background and immune cells in between	No tumor epithelium	No tumor stroma	Some immune cells in healthy stroma, normal healthy colon tissue, some background and/or mucin	Predominantly some healthy colon tissue, some background and/or mucin	Healthy or dysplastic colon tissue 80%, Background 10%, Mucin 10%
40	Tumor stroma	Disorganized stroma, growing infiltratively	Polymorphic and poorly differentiated tumors in much prevalent edematous and disorganized stroma, growing infiltratively, a few immune cells	Part polymorphic and poorly differentiated tumor epithelium, growing infiltratively	Much prevalent edematous and disorganized tumor stroma, often vascularized, stroma-high	Part immune cells, immune-low	N/A	Stroma 60%, Tumor epithelial cells 30%, Immune cells 10%
41	Tumor stroma	Disorganized and loose stroma	Heterogeneous cluster: Some background and/or (loose) stromal tissue, not all tumor-associated, with some vessels and immune cells, a few spots of mucin	A couple isolated tumor cells	Predominantly loose aligned stroma fibers, some tumor-induced, then stroma-high. Vascularized	Part immune cells, immune-low	Some tiles contain muscle tissue from vessels, some background, a few spots of mucin	Stroma 70%, Muscle tissue 10%, Background 10%, Immune cells 10%
42	Tumor epithelium	Moderately differentiated tumor	Mostly moderately differentiated tumor epithelial with a few aligned strands of stroma in between. Some background and a few artefacts	Mostly well-moderately differentiated tumor epithelium	Some aligned strands of stroma, tiles are stroma-low	Few immune cells, some spots (dirty) necrosis, immune-low	Some background	Tumor epithelial cells 70%, Stroma 20%, Background 10%
43	Necrosis	Necrotic	Necrosis with anucleic cells, or fields with mostly immune cells. Much background and some tumor epithelial cells	Some moderately tumor epithelium, some isolated tumor epithelial cells	A few strands of aligned (tumor) stroma	Immune cells in much stroma present necrosis	Some background	Necrosis 50%, Immune cells 20%, Background 20%, Tumor epithelial cells 10%

44	Tumor stroma	Disorganized, immune-low stroma (desert-type)	Mostly aligned stroma, desert type with few nuclei and immune cells, potentially some not tumor-induced with a few isolated tumor epithelial cells	A couple isolated tumor cells	Much prevalent edematous and disorganized tumor stroma, often vascularized, stroma-high	A few immune cells, immune-low	A couple tiles contain longitudinal muscle fibers	Stroma 90%, Immune cells 10%
45	Tumor epithelium	Poor-undifferentiated tumor	Poor-undifferentiated tumor epithelium, with some disorganized stroma, some background	Much poor-undifferentiated tumor epithelium	Some disorganized tumor stroma, stroma-low	A few immune cells, immune-low	Some background	Tumor epithelial cells 70%, Stroma 20%, Background 10%
46	Tumor epithelium	Infiltrated undifferentiated tumor	Cluster rich in cells. Solid growing, undifferentiated tumor, with some immune cells. Potentially micro-satellite instability. Component signet cell carcinoma and squamous cells	Much poor-undifferentiated tumor epithelium	A few strands of aligned (tumor) stroma, stroma-low	Many immune cells, also a couple tiles with only immune cells, immune-high	A tile with slight component signet cell, one with necrosis	Tumor epithelial cells 70%, Immune cells 20%, Stroma 10%

The table shows all HPCs with tissue compositions, including labels, super-clusters, and detailed description of aspect and percentage of tissue types. HPC, histomorphological phenotype cluster. N/A, not applicable.

Supplementary Table 2. Summary of the associations between HPCs and overall survival

HPC	Association	Treatment group	Potential associations with OS	References
0	Worse survival	TCGA and AVANT-standard	Disorganized stroma indicates active remodeling and an intense, yet failing, host response to aggressive invasive tumor cells. The tiles are also stroma-high, consistent with results from the immune landscape analysis, also leading to a worse survival.	Mesker et al (2007), Friedl et al (2011), Zunder et al (2018), Zunder et al (2020), Ueno et al (2021), Strous et al (2022), Polack et al (2024)
2	Better survival	TCGA and AVANT-standard	Aligned stroma indicates a successful defense of the host against less invasive tumor cells. Moreover, the tiles are mostly stroma-low with well-differentiated tumor epithelium leading to a better survival.	Mesker et al (2007), Friedl et al (2011), Weiser et al (2011), Ueno et al (2012), Zunder et al (2018), Zunder et al (2020), Ueno et al (2021), Strous et al (2022), Polack et al (2024)
3	Better survival	TCGA and AVANT-standard	Well-to-moderately differentiated (low-grade) tumor epithelium, along with some high-grade dysplasia and low levels of stromal contents, are well-established indicators of a better survival.	Mesker et al (2007), Weiser et al (2011), Ueno et al (2012), Brierley et al (2016), Weiser et al (2018), Zunder et al (2018), Strous et al (2022), Polack et al (2024)
4	Better survival	AVANT-experimental	This cluster consists mostly of well-to-moderately differentiated (low-grade) tumor epithelium, with stroma-low tiles and an influx of immune cells, which is associated with a favorable survival.	Mesker et al (2007), Weiser et al (2011), Ueno et al (2012), Brierley et al (2016), Pages et al (2018), Weiser et al (2018), Zunder et al (2018), Bagaev et al (2021), Strous et al (2022), Polack et al (2024)
5	Better survival	AVANT-experimental	Much necrosis, promoting the expression of VEGFa as target of bevacizumab, is present in this cluster, leading to a better response to bevacizumab. Enrichment in pathways of cell cycle and DNA repair may also indicate more targets to the cytotoxic chemotherapy and promote survival.	Ferrara et al (2002), Hurwitz et al (2004), Kap et al (2015), Roos et al (2016), Vodicka et al (2019)
6	Worse survival	TCGA and AVANT-standard	This cluster contains a significant amount of inflamed and partially dysplastic colon tissue, indicating a large amount of remodeling. The presence of potential chronic inflammation predisposes the development of cancer as well as gives rise to a possible cancer field effect, hence leading to a worse survival.	Ullman et al (2011), Lockhead et al (2015), Brierley et al (2016), Weiser et al (2018)
7	Better survival	TCGA and AVANT-standard	Avital tumor and necrosis in this cluster are signs of poor neovascularization and thus potentially indicating an less aggressive tumor. Additionally, downregulation in epithelial-to-mesenchymal transition pathways could lead to favorable outcomes.	Ferrara et al (2002), Brierley et al (2016), Weiser et al (2018), Sandberg et al (2019)
9	Better survival	AVANT-experimental	Besides relatively favorable histopathology aspects similar to some previous HPCs (e.g. HPC 4, low-grade tumor epithelium), downregulation of pathways related to epithelial-to-mesenchymal transition and angiogenesis, as well as lower tumor necrosis factor-alpha could lead to an improved survival.	Ferrara et al (2002), Zins et al (2007), Weiser et al (2011), Ueno et al (2012), Weiser et al (2018), Sandberg et al (2019)

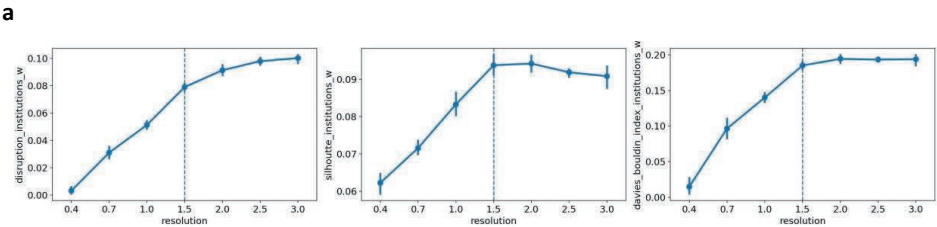
11	Worse survival	TCGA and AVANT-standard	Neovascularization in stroma indicates active host response with wound healing features, known to have tumor promoting properties. Tiles are stroma-high and linked to an epithelial-to-mesenchymal transition in the gene set enrichment analysis, associated to worse outcomes.	Ferrara et al (2002), Mesker et al (2007), Zunder et al (2018), Sandberg et al (2019), Strouss et al (2022), Polack et al (2024)
12	Worse survival	AVANT-experimental	Mucinous tumors are well-known to lead to worse prognosis and can be linked to a higher mutational burden. <i>KRAS</i> signaling-up pathways and depletion in cell cycle regulatory pathways such as DNA repair, overall lead to worse survival.	Kap et al (2015), Brierley et al (2016), Chang et al (2016), Roos et al (2016), Shia et al (2017), Weiser et al (2018), Nagtegaal et al (2019), Vodicka et al (2019), Zhu et al (2021)
13	Better survival	TCGA and AVANT-standard	The immune cell influx has a positive correlation with survival, indicating a good host response; an increased leukocyte fraction was seen in the immune landscape analysis as well.	Pages et al (2018), Bagaev et al (2021)
14	Worse survival	AVANT-experimental	Similar to HPC 12, mucinous tumors are well-known to lead to a worse prognosis and can be linked to a higher mutational burden. <i>KRAS</i> signaling-up pathway and depletion in cell cycle regulatory pathways such as DNA repair, overall lead to worse survival.	Kap et al (2015), Brierley et al (2016), Chang et al (2016), Roos et al (2016), Shia et al (2017), Weiser et al (2018), Nagtegaal et al (2019), Vodicka et al (2019), Zhu et al (2021)
17	Worse survival (TCGA and AVANT-standard) AND Better survival (AVANT-experimental)	Both treatment groups	Mixed associations to survival are seen here. This could be due to high amount of stroma and enrichment in epithelial-to-mesenchymal transition gene pathways, which are poor prognostic factors and lead to a worse survival among the standard-treated patients. However, both the aligned stroma with immune infiltration and the enrichment in the hypoxia pathway (involving the expression of VEGFa gene as the target of bevacizumab), lead to a better survival in experimentally-treated patients.	Ferrara et al (2002), Hurwitz et al (2004), Mesker et al (2007), Friedl et al (2011), Pages et al (2018), Zunder et al (2018), Sandberg et al (2019), Zunder et al (2020), Bagaev et al (2021), Ueno et al (2021), Strouss et al (2022), Polack et al (2024)
18	Better survival	AVANT-experimental	The immune cell component may indicate an effective response from the host and thus lead to a favorable survival. Enrichment in the angiogenesis pathway (involving VEGFa as the target of bevacizumab) may indicate an improved survival in patients receiving bevacizumab.	Hurwitz et al (2004), Kap et al (2015), Pages et al (2018), Bagaev et al (2021)
19	Worse survival	TCGA and AVANT-standard	Stroma-high tiles categorized by disorganized stroma, including an absence of immune cells, indicate an active remodeling with a failing host response, which collectively may lead to a poor survival outcome.	Friedl et al (2011), Kap et al (2015), Pages et al (2018), Zunder et al (2018), Zunder et al (2020), Bagaev et al (2021), Ueno et al (2021), Strouss et al (2022), Polack et al (2024)
20	Worse survival	AVANT-experimental	Disorganized stroma can be a sign of insufficient host response and matrix remodeling. Enrichment in pathways of <i>KRAS</i> signaling-up and hypoxia, along with depletion in cell cycle regulation pathways, may contribute to an unfavorable survival through several plausible mechanisms.	Friedl et al (2011), Kap et al (2015), Chang et al (2016), Roos et al (2016), Vodicka et al (2019), Zunder et al (2020), Ueno et al (2021), Zhu et al (2021)

21	Worse survival	TCGA and AVANT-standard	Neovascularization in stroma-high tiles, as indicator of the failure of host response, and enrichment in oncogenic pathways may cause an unfavorable survival in the standard-treated patients.	Ferrera et al (2002), Kap et al (2015), Pages et al (2018), Bagaev et al (2021)
23	Worse survival	TCGA and AVANT-standard	Tiles are characterized by precancerous dysplastic colon tissue and chronic inflammation, both of which predispose to cancer development and suggest a cancer field effect, potentially leading a worse survival.	Ullman et al (2011), Lockhead et al (2015), Brierley et al (2016), Weiser et al (2018)
24	Better survival (AVANT-standard) AND Worse survival (AVANT-experimental)	Both treatment groups	The underlying mechanisms of the mixed results between the standard and experimental treatment groups remains unclear. However, we hypothesize that this may be due to the complex interactions between mixed histopathological patterns and gene expression patterns. On one hand, the worse survival may be attributed to the stroma-high histopathological features of stromal strands in the muscle fibers. This is further confirmed by the upregulation of epithelial-to-mesenchymal transition pathway that has been linked to a poor survival. Enrichment in KRAS signaling-up and depletion cell cycle regulatory pathways may be associated with a worse response to chemotherapy. On the other hand, enrichment in interferon-gamma response and inflammatory response pathways may indicate a strong host immune response which may lead to a better survival.	Ferrera et al (2002), Hurwitz et al (2004), Zins et al (2007), Friedl et al (2011), Kap et al (2015), Chang et al (2016), Roos et al (2016), Pages et al (2018), Sandberg et al (2019), Vodicka et al (2019), Jorgovanovic et al (2020), Zunder et al (2020), Bagaev et al (2021), Zhu et al (2021), Polack et al (2024)
25	Worse survival	Both treatment groups	Necrotic and mucinous components may be associated with a worse survival in general. We also observed depletion in expression of genes in several immune related pathways (e.g. interferon-gamma response). In addition, no enrichment was observed in the pathways involving VEGFa, potentially indicating a low number of targets for bevacizumab.	Ferrara et al (2002), Brierley et al (2016), Pages et al (2018), Weiser et al (2018), Nagtegaal et al (2019), Jorgovanovic et al (2020), Bagaev et al (2021)
26	Better survival	AVANT-experimental	Tiles are high in immune cells despite the presence of poorly-to-undifferentiated (high-grade) tumor epithelium, which has been linked to tumors with microsatellite instability. This is known to lead to a better survival. Enrichment in cell cycle regulatory pathways may indicate more targets for cytotoxic chemotherapy, potentially enhancing the survival.	Boland et al (2010), Brierley et al (2016), Shia et al (2017), Pages et al (2018), Weiser et al (2018), Nagtegaal et al (2019), Bagaev et al (2021)
27	Better survival	AVANT-experimental	Aligned stroma and influx of immune cells are both predictors for a good response to cytotoxic chemotherapy. Enrichment in the hypoxia pathway, which involves the VEGFa gene, potentially indicate more targets and a better response to bevacizumab. Together, these factors may result in an improved outcome in the experimentally treated patients.	Hurwitz et al (2004), Pages et al (2018), Kap et al (2015), Bagaev et al (2021), Ueno et al (2021)
28	Better survival	Both treatment groups	Higher fractions of this muscle tissue-abundance cluster may indicate more healthy tissue, which may lead a better prognosis. No tumor buds was observed (in contrast to HPC 33), which may indicate a less aggressive tumor.	Brierley et al (2016), Lugli et al (2017), Weiser et al (2018), Nagtegaal et al (2019)

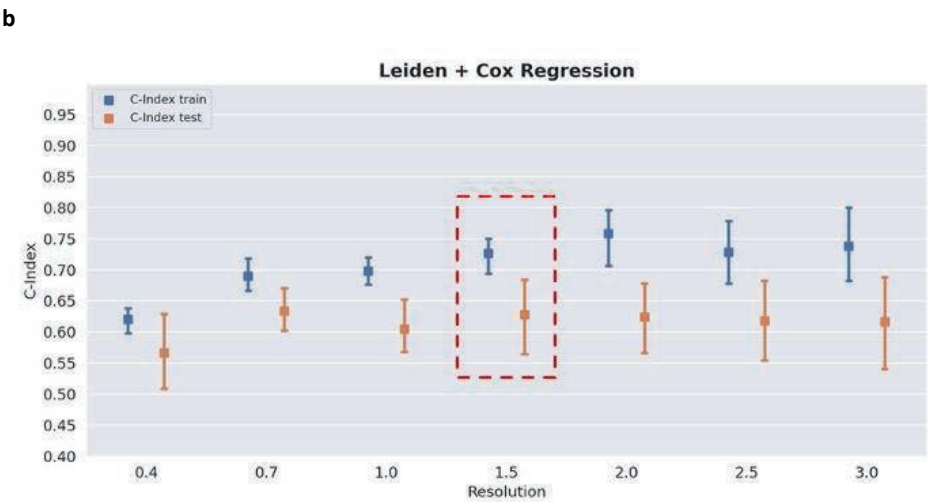
31	Better survival	Both treatment groups	Aligned stroma and an immune component are both favorable factors regarding response to cytotoxic therapy. This is consistent with the high leukocyte fraction observed in the immune landscape analysis. Moreover, enrichment in pathways of hypoxia and angiogenesis, which involves VEGFa, may indicate more targets for bevacizumab. Enrichments in cell cycle regulation (e.g. G3/M checkpoint, mTORC1) may suggest more targets for the chemotherapy.	Ferrara et al (2002), Hurwitz et al (2004), Roos et al (2016), Pages et al (2018), Vodicka et al (2019), Zunder et al (2020), Bagaev et al (2021), Ueno et al (2021)
33	Worse survival	AVANT-experimental	The worse survival may be related to tumor buds invading the serosa and fatty tissue, a known indicator for aggressive tumors. Fatty tissue is also known as a cancer inducer, providing energy for cancer cells. The relative absence of immune cells observed in the tiles suggests a weak host response, which is further reflected in the depletion in inflammatory and angiogenesis pathways. Together, these factors may contribute to poor survivals.	Ferrara et al (2015), Lugli et al (2017), Pages et al (2018), Bagaev et al (2021), Brown et al (2023)
34	Better survival	AVANT-experimental	Muscle tissue abundance may suggest fractionally more healthy tissue, which in turn suggests a smaller tumor size and thus a better prognosis. Moreover, we observed neovascularization in the tiles, which in the experimentally-treated group could enhance survival. Enrichment in pathways related to cell cycle regulations may indicate more targets to the cytotoxic chemotherapy.	Ferrera et al (2002), Kap et al (2015), Brierley et al (2016), Roos et al (2016), Weiser et al (2018), Nagtegaal et al (2019), Vodicka et al (2019)
38	Better survival (AVANT-standard) AND Worse survival (AVANT-experimental)	Both treatment groups	Mechanisms underlying these different associations remain unclear. On one hand, mucinous tumors are known to lead to a worse prognosis. KRAS signaling-up pathways and depletion in cell cycle regulatory pathways may suggest a poor response to chemotherapy. On the other hand, enrichment in several inflammatory response pathways may suggest a strong host immune response that results in a better survival.	Kap et al (2015), Brierley et al (2016), Chang et al (2016), Roos et al (2016), Pages et al (2018), Weiser et al (2018), Nagtegaal et al (2019), Vodicka et al (2019), Jorgovanovic et al (2020), Bagaev et al (2021), Zhu et al (2021)
39	Better survival	Both treatment groups	Tiles contain healthy colon tissue. Higher fraction of healthy colon tissue tiles may indicate relatively smaller tumor size and potentially less cancer field effect in a larger tissue area (i.e. little disturbances in the healthy tissue) and leads to a better survival.	Lockhead et al (2015), Brierley et al (2016), Weiser et al (2018), Nagtegaal et al (2019)
40	Worse survival	Both treatment groups	Disorganized stroma, indicating an insufficient host response, and poor-to-undifferentiated (high-grade) tumor epithelium could both lead to an unfavorable survival. A cancer field effect negatively influences surrounding tissue, adding to the poor association. Depletion of cell cycle pathways may suggest few targets for cytotoxic chemotherapy and thus worse outcomes.	Friedl et al (2011), Weiser et al (2011), Ueno et al (2012), Kap et al (2015), Lockhead et al (2015), Roos et al (2016), Nagtegaal et al (2019), Vodicka et al (2019), Zunder et al (2020), Ueno et al (2021)
41	Worse survival	AVANT-experimental	The combination of disorganized and loose stroma as well as absence of immune components indicates active remodeling and failure of the immune response. Depletion in cell cycle pathways may suggest fewer targets for cytotoxic chemotherapy and result in a worse survival.	Friedl et al (2011), Kap et al (2015), Roos et al (2016), Vodicka et al (2019), Zunder et al (2020), Bagaev et al (2021), Ueno et al (2021)

45	Worse survival	TCGA and AVANT-standard	Poor-to-undifferentiated (high-grade) tumor epithelium is a well-known indicator of poor survival.	Weiser et al (2011), Ueno et al (2012), Brierley et al (2016), Weiser et al (2018), Nagtegaal et al (2019)
46	Better survival	TCGA and AVANT-standard	Abundance in immune cell components is positively correlated with a better survival. Poorly-to-undifferentiated (high-grade) tumor epithelium together with enrichment in immune-related gene pathways may potentially indicate microsatellite instability which is known to indicate a better survival.	Boland et al (2010), Brierley et al (2016), Shia et al (2017), Pages et al (2018), Weiser et al (2018), Nagtegaal et al (2019), Bagaev et al (2021)

The table summarizes HPCs that were significantly associated with OS per treatment group, including references. Color legend: Orange, HPCs associated with worse survival in one group. Red, HPCs associated with worse survival in both groups. Blue, HPCs associated with improved survival in one group. Green, HPCs associated with improved survival in both groups. Grey, HPCs with mixed associations with OS in the two treatment groups. AVANT, Bevacizumab-Avastin® adjuVANT trial. HPC, histomorphological phenotype cluster. OS, overall survival. TCGA, The Cancer Genome Atlas. VEGF, vascular endothelial growth factor.



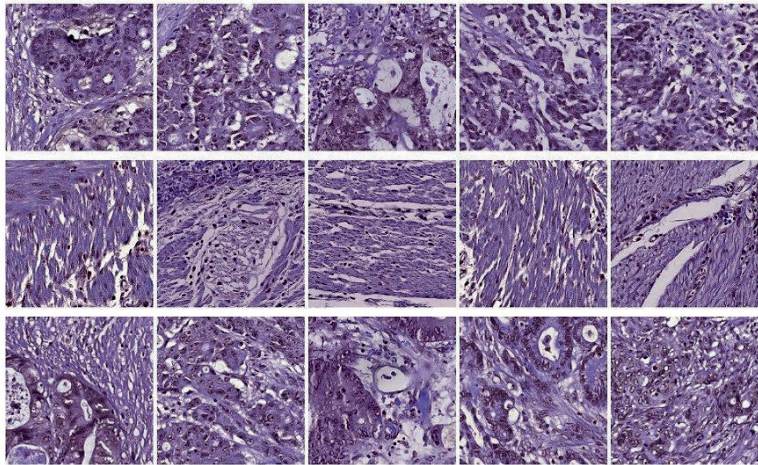
Leiden resolution optimization using unsupervised methods in TCGA



Leiden resolution optimization on OS prediction from regularized Cox regression in TCGA

Supplementary Figure 1: Optimization of Leiden resolution (a) Disruption score, Silhouette score, and Daves-Boundin index weighted by mean percentage of the institution presence in each HPC reached a consensus on optimal Leiden resolution at 1.5 in TCGA (n=405 patients). (b) We trained regularized Cox regressions including all 47 HPCs for each Leiden resolution using 5-fold CV in TCGA (n=405 patients). The Leiden resolution 0.7 and 1.5 showed the highest validation c-index. We chose Leiden 1.5 as the optimal resolution because it led to more explorable HPCs. CV, cross-validation. HPC, histomorphological phenotype cluster. OS, overall survival. TCGA, The Cancer Genome Atlas. n, number of patients included in the analyses after those with missing clinical data. Source data are provided as a Source Data file.

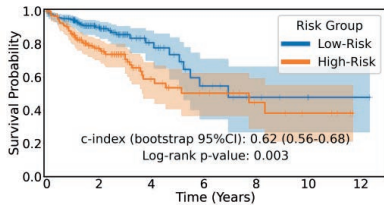
Cluster: 19 / 47, Test: 42 / 50



Supplementary Figure 2. Objective test for all HPCs. Screenshot of the online objective test. Three rows of tiles were presented at each test, with 2 rows were from the same HPC and the other row from a randomly selected different HPC.

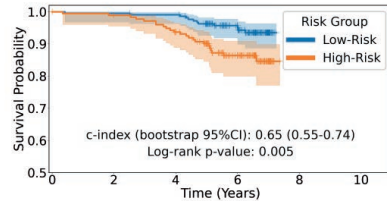
a HPC-based classifier in five-fold CV

in TCGA



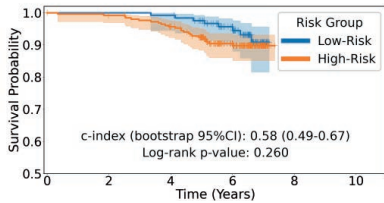
b HPC-based classifier in the external

AVANT control group



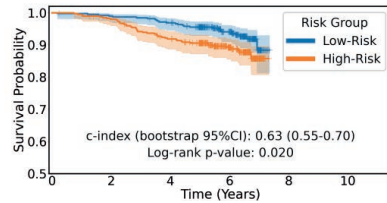
c Clinical baseline classifier in the ex-

ternal AVANT control group



d HPC-based classifier in five-fold CV in

AVANT-experimental group



Supplementary Figure 3. Regularized Cox regression for OS prediction. (a) Kaplan-Meier plot showing the HPC-based classifier stratifying OS risk in the five-fold CV in TCGA (two-sided log-rank $p=0.003$, $n=405$ patients). (b) The HPC-based classifier stratifying OS risk in the external AVANT control test set (two-sided log-rank $p=0.005$, $n=391$ patients). (c) The clinical baseline model (predictors including age, sex, tumor-stroma ratio, AJCC TNM staging) stratifying OS risk in the independent AVANT control test set (two-sided log-rank $p=0.260$, $n=378$ patients). (d) The HPC-based classifier stratified OS risk AVANT-experimental group (two-sided log-rank $p=0.020$, $n=780$ patients). AVANT, Bevacizumab-Avastin® adjuVANT trial. CV, cross-validation. HPC, histomorphological phenotype cluster. OS, overall survival. TCGA, The Cancer Genome Atlas. n , number of patients included in the analyses after excluding those with missing clinical data. Source data are provided as a Source Data file.

References

- Alexander Bagaev, Nikita Kotlov, Krystle Nomie, Viktor Svelkolkin, Azamat Gafurov, Olga Isaeva, Nikita Osokin, Ivan Kozlov, Felix Frenkel, Olga Gancharova, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer cell*, 39(6):845–865, 2021.
- C Richard Bolland and Ajay Goel. Microsatellite instability in colorectal cancer. *Gastroenterology*, 138(6):2073–2087, 2010.
- James D Brierley, Mary K Gospodarowicz, and Christian Wittekind. *TNM classification of malignant tumours*. John Wiley & Sons, 2017.
- Kristy A Brown and Philipp E Scherer. Update on adipose tissue and cancer. *Endocrine reviews*, 44(6):961–974, 2023.
- Shih-Ching Chang, Pei-Ching Lin, Jen-Kou Lin, Chien-Hsing Lin, Shung-Haur Yang, Wen-Yi Liang, Wei-Shone Chen, and Jeng-Kai Jiang. Mutation spectra of common cancer-associated genes in different phenotypes of colorectal carcinoma without distant metastasis. *Annals of surgical oncology*, 23:849–855, 2016.
- Napoleone Ferrara. Vegf and the quest for tumour angiogenesis factors. *Nature Reviews Cancer*, 2(10):795–803, 2002.
- Peter Friedl and Stephanie Alexander. Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell*, 147(5):992–1009, 2011.
- Herbert Hurwitz, Louis Fehrenbacher, William Novotny, Thomas Cartwright, John Hainsworth, William Heim, Jordan Berlin, Ari Baron, Susan Griffing, Eric Holmgren, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *New England journal of medicine*, 350(23):2335–2342, 2004.
- Dragica Jorgovanovic, Mengjia Song, Liping Wang, and Yi Zhang. Roles of ifn- γ in tumor progression and regression: a review. *Biomarker research*, 8:1–16, 2020.
- Elisabeth J Kap, Petra Seibold, Swantje Richter, Dominique Scherer, Nina Habermann, Yesilda Balavarcia, Lina Jansen, Natalia Becker, Katrin Pfütze, O Popanda, et al. Genetic variants in dna repair genes as potential predictive markers for oxaliplatin chemotherapy in colorectal cancer. *The pharmacogenomics journal*, 15(6):505–512, 2015.
- Paul Lochhead, Andrew T Chan, Reiko Nishihara, Charles S Fuchs, Andrew H Beck, Edward Giovannucci, and Shuji Ogino. Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. *Modern Pathology*, 28(1):14–29, 2015.
- Alessandro Lugli, Richard Kirsch, Yoichi Ajioka, Fred Bosman, Gieri Cathomas, Heather Dawson, Hala El Zimaity, Jean-François Fléjou, Tine Plato Hansen, Arndt Hartmann, et al. Recommendations for reporting tumor budding in colorectal cancer based on the international tumor budding consensus conference (itbcc) 2016. *Modern pathology*, 30(9):1299–1311, 2017.
- Wilma E Mesker, Jan Junggeburst, Karoly Szuhai, Pieter de Heer, Hans Morreau, Hans J Tanke, and Rob AEM Tollenaar. The carcinoma–stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. *Analytical Cellular Pathology*, 29(5):387–398, 2007.
- ID Nagtegaal, Mark Arends, and Robert Odze. Tumours of the colon and rectum: Who classification of tumours of the colon and rectum, tm staging of carcinomas of the colon and rectum and the introduction. In *World Health Organization Classification of Tumours of the Digestive System*, pages 157–162. IARC press, 2019.
- Franck Pagès, Bernhard Mlecnik, Florence Marliot, Gabriela Bindea, Fang-Shu Ou, Carlo Bifulco, Alessandro Lugli, Inti Zlobec, Tilman T Rau, Martin D Berger, et al. International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study. *The Lancet*, 391(10135):2128–2139, 2018.
- M Polack, MA Smit, GW van Pelt, AGH Roodvoets, E Meershoek-Klein Kranenbarg, H Putter, H Gelderblom, ASLP Crobach, V Terpsstra, G Petrushevska, et al. Results from the united study: a multicenter study validating the prognostic effect of the tumor-stroma ratio in colon cancer. *ESMO open*, 9(4):102988, 2024.
- Wynand P Roos, Adam D Thomas, and Bernd Kaina. Dna damage and the balance between survival and death in cancer biology. *Nature Reviews Cancer*, 16(1):20–33, 2016.
- Tessa P Sandberg, Maaïke PME Stuart, Jan Oosting, Rob AEM Tollenaar, Cornelis FM Sier, and Wilma E Mesker. Increased expression of cancer-associated fibroblast markers at the invasive front and its association with tumor-stroma ratio in colorectal cancer. *BMC cancer*, 19(1):1–9, 2019.
- Jinru Shia, Nikolaus Schultz, Deborah Kuk, Efsevia Vakiani, Sumit Middha, Neil H Segal, Jaclyn F Hechtman, Michael F Berger, Zsofia K Stadler, Martin R Weiser, et al. Morphological characterization of colorectal cancers in the cancer genome atlas reveals distinct morphology–molecular associations: clinical and biological implications. *Modern pathology*, 30(4):599–609, 2017.
- MTA Strous, TKE Faes, ALHM Gubbels, RLA van der Linden, WE Mesker, K Bosscha, CM Bronkhorst, MLG Janssen-Heijnen, FJ Vogelaar, and AP de Bruïne. A high tumour-stroma ratio (tsr) in colon tumours and its metastatic lymph nodes predicts poor cancer-free survival and chemo resistance. *Clinical and Translational Oncology*, 24(6):1047–1058, 2022.
- Hideki Ueno, Yoshiki Kajiwara, Yoichi Ajioka, Tamotsu Sugai, Shigeki Sekine, Megumi Ishiguro, Atsuo Takashima, and Yukihide Kanemitsu. Histopathological atlas of desmoplastic reaction characterization in colorectal cancer. *Japanese Journal of Clinical Oncology*, 51(6):1004–1012, 2021.
- Hideki Ueno, Yoshiki Kajiwara, Hideyuki Shimazaki, Eiji Shinto, Yojiro Hashiguchi, Kuniaki Nakanishi, Kazunari Maekawa, Yuka Katsurada, Takahiro Nakamura, Hidetaka Mochizuki, et al. New criteria for histologic grading of colorectal cancer. *The American journal of surgical pathology*, 36(2):193–201, 2012.
- Thomas A Ullman and Steven H Itzkowitz. Intestinal inflammation and cancer. *Gastroenterology*, 140(6):1807–1816, 2011.
- Pavel Vodicka, Sona Vodenkova, Tomas Buchler, and Ludmila Vodickova. Dna repair capacity and response to treatment of colon cancer. *Pharmacogenomics*, 20(17):1225–1233, 2019.

25. Martin R Weiser. Ajcc 8th edition: colorectal cancer. *Annals of surgical oncology*, 25:1454–1455, 2018.
26. Martin R Weiser, Mithat Gönen, Joanne F Chou, Michael W Kattan, and Deborah Schrag. Predicting survival after curative colectomy for cancer: individualizing colon cancer staging. *Journal of Clinical Oncology*, 29(36):4796, 2011.
27. Gongmin Zhu, Lijiao Pei, Hongwei Xia, Qiulin Tang, and Feng Bi. Role of oncogenic kras in the prognosis, diagnosis and treatment of colorectal cancer. *Molecular cancer*, 20(1):1–17, 2021.
28. Karin Zins, Dietmar Abraham, Mouldy Sioud, and Seyedhossein Aharinejad. Colon cancer cell–derived tumor necrosis factor- α mediates the tumor growth–promoting response in macrophages by up-regulating the colony-stimulating factor-1 pathway. *Cancer research*, 67(3):1038–1045, 2007.
29. Stéphanie M Zunder, Hans Gelderblom, Rob A Tollenaar, and Wilma E Mesker. The significance of stromal collagen organization in cancer tissue: An in-depth discussion of literature. *Critical Reviews in Oncology/Hematology*, 151:102907, 2020.
30. Stéphanie M Zunder, Gabi W van Pelt, Hans J Gelderblom, Christoph Mancao, Hein Putter, Rob A Tollenaar, and Wilma E Mesker. Predictive potential of tumour-stroma ratio on benefit from adjuvant bevacizumab in high-risk stage ii and stage iii colon cancer. *British Journal of Cancer*, 119(2):164–169, 2018.

