Single-cell integration and multi-modal profiling reveals phenotypes and spatial organization of neutrophils in colorectal cancer

Valentin Marteau[1,20], Niloofar Nemati[1,20], Kristina Handler[2,20], Deeksha Raju[2], Erika Kvalem Soto[1], Georgios Fotakis[1], Sandro Carollo[1], Nina Boeck[1], Alessia Rossi[1], Alexander Kirchmair[1], Alexandra Scheiber[3], Arno Amann[3], Andreas Seeber[3], Elisabeth Gasser[4], Steffen Ormanns[5], Michael Günther[5], Agnieszka Martowicz[3,6], Zuzana Loncova[1,19], Giorgia Lamberti[1], Marie Theres Eling[14], Lena Horvath[3], Hassan Fazilaty[7], Tomas Valenta[7,10], Gregor Sturm[1,9], Dietmar Rieder[1], Sieghart Sopper[3], Andreas Pircher[3], Patrizia Stoitzner[11], Paul Ziegler[12], Markus Tschurtschenthaler[13], Florian Huemer[16], Daniel Neureiter[15], Richard Greil[16], Lukas Weiss[16,17], Marieke Ijsselsteijn[18], Noel F.C.C. de Miranda[18], Dominik Wolf [3], Isabelle C. Arnold[2,8,21*], Stefan Salcher[3,21*], Zlatko Trajanoski[1,21,22*]

[1]Biocenter, Institute of Bioinformatics, Medical University of Innsbruck, Austria
[2]Institute of Experimental Immunology, University of Zurich, Switzerland
[3]Department of Internal Medicine V, Haematology & Oncology, Comprehensive Cancer Center Innsbruck and Tyrolean Cancer Research Institute, Medical University of Innsbruck, Austria
[4]Department of Visceral, Transplant and Thoracic Surgery, Medical University Innsbruck, Austria
[5]Innpath, Tirol Kliniken, Medical University Innsbruck, Austria
[6]Tyrolpath Obrist Brunhuber GmbH, Zams, Austria
[7]Department of Molecular Life Sciences, University of Zürich, Switzerland
[8]Comprehensive Cancer Center Zürich, Zürich, Switzerland
[9]Boehringer Ingelheim International Pharma GmbH & Co KG, Biberach, Germany
[10]Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czech Republic
[11]Department of Dermatology, Venereology and Allergology, Medical University of Innsbruck, Austria
[12]Senckenberg Institute of Pathology, Goethe University Frankfurt, Germany
[13]Translational Cancer Research and Institute of Experimental Cancer Therapy, Klinikum rechts der Isar, School of Medicine & Health, Technical University of Munich, Germany
[14]Department of Therapeutic Radiology and Oncology, Medical University Innsbruck, Austria
[15]Institute of Pathology, Paracelsus Medical University, Salzburg, Austria
[16]Department of Internal Medicine III, Paracelsus Medical University, Salzburg, Austria
[17]Austrian Breast & Colorectal Cancer Study Group (ABCSG), Vienna, Austria
[18]Department of Pathology, Leiden University Medical Center, The Netherlands
[19]Present address: Institute of Pathology, Neuropathology and Molecular Pathology, Medical University of Innsbruck, Austria
[20]These authors contributed equally
[21]These authors contributed equally

[22]Lead contact

*Correspondence: stefan.salcher@i-med.ac.at, zlatko.trajanoski@i-med.ac.at

SUMMARY

Precision oncology including immunotherapy with checkpoint blockers for patients with colorectal cancer (CRC) continues to lag behind other cancer types. The immune composition of the tumor microenvironment (TME) has a major impact on the therapeutic response and clinical outcome. Here, we comprehensively characterize the TME at the single-cell level by first building a large-scale atlas that integrates 4.27 million single cells from 1,670 samples and then complementing the atlas by profiling single cells with low mRNA content. The analysis of the atlas allows refined tumor classification into four immune phenotypes: immune desert, B cell enriched, T cell enriched, and myeloid enriched subtypes. Within the myeloid compartment we uncover distinct subpopulations of neutrophils that acquire new functional properties in blood and in the TME, including antigen-presenting capabilities. Further, spatial multimodal single-cell profiling reveals that neutrophils are organized in clusters within distinct multi-cellular niches. Finally, using an orthotopic mouse model we show that cancer-derived systemic signals modify neutrophil production in the bone marrow, providing evidence for tumor-induced granulopoiesis. Our study provides a high-resolution resource for CRC TME and suggests novel therapeutic strategies targeting neutrophils.

KEYWORDS

Single-cell sequencing, spatial single-cell profiling, single-cell atlas, cellular niches

INTRODUCTION

Colorectal cancer (CRC) remains the most common gastrointestinal cancer and the second leading cause of cancer-related death according to the GLOBOCAN data[1]. The growing incidence in some countries reflects a modification in lifestyle and its consequences related with "Westernization" such as diet, obesity, physical inactivity, and alcohol consumption. Similarly, "Westernization" of diets and lifestyle have been attributed to the alarming growing incidence of early-onset CRCs (individuals <50 years of age)[2]. Despite the improvement of early detection screening policies and the development of many novel treatments, the median overall survival (OS) for patients with metastatic CRC (mCRC) is ~30 months[3]. Progress in targeted therapy for mCRC has been limited, relative to other solid cancers like lung cancer or melanoma[4]. The genetic heterogeneity[5] as well as paucity of drugs targeting common driver genes, pose considerable challenges for developing precision oncology approaches for patients with CRC. Moreover, CRC seems to be refractory to therapy with immune checkpoint inhibitors (ICI) with the notable exception of 4% of metastatic patients with tumors characterized by microsatellite instability (MSI) or *POLE* proofreading mutations[6]. This is somehow paradoxical since CRCs are known to be under immunological control, as we have shown in the past[7]. Thus, there is an urgent need to explore new immunomodulatory strategies for the vast majority of CRC patients with microsatellite stable (MSS) tumors. Strategies that convert the "immune cold" tumor microenvironment (TME) to "immune hot" could sensitize cancers to ICI therapy. A number of clinical trials with drug combinations have been carried out to overcome resistance to immunotherapy in MSS CRC[8,9]. The benefits of using a dual combination of ICIs (CTLA-4 and PD-1 blockade) or adding ICIs to chemotherapy appear to be very limited[8,10]. Combination strategies that exploit synergistic potential of ICI and targeted drugs showed increase of the anti-tumor activity, however, the objective response rates were disappointingly low[8]. Thus, to address this unmet clinical need a better understanding of the immunosuppressive TME and its heterogeneity is of utmost importance.

The intrinsic complexity of the interaction of the two interwoven systems, the tumor and the immune system, poses considerable challenges and requires comprehensive approaches to interrogate cancer immunity during tumor initiation and progression, and following therapeutic modulation thereof. A number of established and novel high-throughput technologies enable the generation of the necessary data and thereby provide the basis for mechanistic understanding. Specifically, next-generation sequencing (NGS) of cancer samples provide large data sets that can be mined for immunologically relevant parameters (see our reviews[11,12]). Using these tools we were able to chart the TME in CRC[13] and developed The Cancer Immunome Atlas[14]. While these attempts uncovered the enormous heterogeneity of MSS CRC and revealed unique cancer cell and TME subtypes with prognostic and therapeutic relevance, the advent of single-cell RNA sequencing (scRNA-seq) technologies opened fundamental new avenues to dissect cancer heterogeneity. We recently developed a novel computational approach to dissect the cellular diversity in the TME at high resolution by integrating expression profiles

of single cells from multiple studies using cutting-edge artificial intelligence (AI) methods[15]. Similar effort was recently carried out in the context of colorectal cancer[16], enabling stratification of patients. However, limited by the sampling bias of the commonly used 10x Genomics platform, this atlas is lacking neutrophils. Neutrophils are the most abundant cells circulating in the human body and are emerging as important regulators of cancer[17–20]. Neutrophils are highly abundant in intratumoral as well as stromal regions in CRC[18] and co-localize with $CD8^+$ T cells[19]. Neutrophils exhibit remarkable plasticity, show different pro-tumor and anti-tumor functions[20], and can acquire anti-tumor phenotype in response to immunotherapy[21]. The importance of neutrophil-T cell interactions is supported by the observation that $CD4^+$ PD-1^+ T cells in granulocyte-dominated cell neighborhoods are associated with favorable prognosis in CRC[22]. Hence, deep characterization of neutrophils in the TME, in normal tissue, and blood from CRC patients would provide novel insights into the myeloid compartment and the myeloid-lymphoid cell-cell interactions.

Here, we first integrated CRC single-cell RNA sequencing datasets spanning 4.27 million cells from 650 patients and 1,670 samples. We then carried out single-cell profiling of samples from CRC patients using optimized protocol and platform that efficiently recovers low mRNA-content immune cells. The analyses refined tumor classification into four immune phenotypes. Within the myeloid compartment we uncovered distinct subpopulations of neutrophils that acquire new functional properties in the TME including antigen-presenting capabilities (HLA-DR^+ neutrophils). Furthermore, spatial multimodal single-cell profiling revealed that neutrophils are organized in clusters within distinct multi-cellular niches. Finally, using an orthotopic mouse model we show that cancer-derived systemic signals modify neutrophil production in the bone marrow, providing evidence for tumor-induced granulopoiesis.

RESULTS

Generation of a large-scale CRC single-cell atlas

Based on our previous approach[15] we compiled CRC related scRNA-seq data from 48 studies and 76 datasets comprising 4.27 million cells from 650 patients and 1,670 samples (Figure 1A). This comprehensive CRC single-cell atlas integrates curated and quality-checked transcriptomic data from publicly available studies, complemented with extensive metadata curation, and encompasses the full spectrum of disease progression, from normal colon to polyps, primary tumors, and metastases, covering both early and advanced stages of CRC (see STAR Methods, Figures 1B-C and Supplementary Figure S1A-E). Study characteristics are summarized in Table S1. In total, the atlas includes transcriptomic data from 487 CRC patients (median: 58 years [22-91]), 99 patients with polyps (median: 60 years [40-75]), and 64 control individuals (median: 57 years [23-79]). Of the 487 CRC patients, 254 had colon cancer (COAD), 63 rectal cancer (READ), and 170 were not otherwise specified. Of these, 89 were classified as microsatellite instable (MSI), 268 as microsatellite stable (MSS), and 130 had an unknown microsatellite status. For disease stage 41 patients were classified as TNM stage I, 121 as stage II, 171 as stage III, and 110 as stage IV, while the stage was not reported for 44 patients. CRC samples include tissue of the primary tumor (n=674) or CRC liver metastasis (n=175) that were obtained by surgical resection. Among the control samples, 284 were derived from distant non-malignant tissue of colon tumor patients, of which 220 have a patient-matched tumor sample. 186 non-tumor control samples as well as 116 polyp and 90 matched normal samples were obtained from patients undergoing colonoscopy. Further, 103 blood samples of CRC patients as well as 24 samples derived from tumor affected lymph nodes, with another 18 non-tumor affected lymph node samples were included. Of the control patients none had a history of inflammatory bowel diseases.

Although neutrophils play a prominent role in cancer biology, they are underrepresented in most available scRNA-seq datasets generated with the widely used 10x Genomics technology[23] (Figure S1 F). To address this, we generated a scRNA-seq dataset from peripheral blood, adjacent normal tissue, and tumor tissue of 12 CRC patients using the BD Rhapsody platform and our optimized protocol[24] that allows capture of cells with very low mRNA content[23]. The data was then integrated into the final high-resolution single-cell atlas for CRC. Overall, the atlas integrates 4,264,946 single cells, which are annotated to 12 coarse cell type identities and 62 major cell subtypes or cell states (e.g. dividing cells) based on previously established canonical single-cell signatures including 1,146,196 epithelial cells, 1,552,108 T/NK cells, 804,042 B/Plasma cells, 385,589 myeloid cells, and 33,9767 stromal and endothelial cells (Figure 1D). Cancer cells (in total 600,740 cells, from primary tumor and metastatic tissue) in the atlas showed high heterogeneity in their transcriptomic profiles (Figure 1B). Due to the large patient cohort, we were able to apply high-resolution cancer cell classification based on their specific marker gene expression signatures (Figure S1D). We divided the following main clusters: 1) colonocyte-like (*FABP1, CEACAM7, ANPEP, SI*), 2) BEST4 (BEST4*, OTOP2, SPIB*), Goblet-like

153 (*MUC2, FCGBP, ZG16, ATOH1*), 3) Crypt-like (*LGR5, SMOC2*), and 4) TA-like (*TOP2A, UBE2C,*
154 *PCLAF, HELLS, TK1*) (Figure 1D, Figure S1C). There are highly mitotic/proliferative clusters (*LGR5,*
155 *TOP2A, MKI67*) of both COAD and READ that resemble highly aggressive and invasive cancer cells.
156 In addition, *EPCAM* expressing cells in the blood of CRC patients were annotated as circulating cancer
157 cells.
158
159 With a total of 2,741,739 immune cells, this single-cell atlas enables high-resolution classification of
160 immune cell types across both innate and adaptive subsets (Figure 1C, Figure S1A-C). These include T
161 cells - CD4 (*CD4*), CD8 (*CD8A*), and T regulatory cells (*FOXP3*) - as well as NK cells (*NKG7*,
162 *TYROBP*). B cells are represented as naïve (*IGHD*, *TCL1A*), memory (*CD27*,*CD44*), and activated
163 (*CD69*, *CD83*) subsets, while plasma cells (*JCHAIN*) express immunoglobulins such as IgA (*IGHA1*),
164 IgG (*IGHG1*), and IgM (*IGHM*). The myeloid lineage encompasses monocytes (*VCAN*, *CD14*),
165 macrophages (*C1QA*, *C1QB*), mast cells (*TPSB2*, *CPA3*), and dendritic cells, including classical
166 (*CLEC9A*, *CD1C*) and plasmacytoid subsets (*IL3RA*, *GZMB*), as well as eosinophils (*CLC*, *CCR3*) and
167 neutrophils (*CXCR2*, *FCGR3B*). Specialized populations such as gamma-delta T cells (*TRDV1*, *TRGC1*)
168 and innate lymphoid cells (ILCs; *IL4I1*, *KIT*) are also depicted, along with immune cell developmental
169 stages, such as progenitors (*CD24*, *AXL*) and stem-like cells (*LEF1*), and cycling states (*MKI67*, *CDK1*).
170
171 **Single-cell composition of the TME highlights distinct immune infiltration patterns in primary**
172 **CRC tumors.**
173 Utilizing the high-resolution atlas, we stratified 284 CRC untreated patients with primary tumor samples
174 based on the infiltration patterns of their tumor microenvironment (TME). The high number of patients
175 enabled us to perform unsupervised clustering of batch-corrected cell-type fractions, revealing four
176 major immune clusters with distinct subclusters (Figure 2A and Supplementary Figure 2A): 1) immune
177 deserted tumors (ID; i.e. no significant immune cell infiltration but high cancer cell fraction); 2) the
178 subtype of tumors with B cell dominance (B; B cell, plasma IgA and IgG cell); 3) the subtype of tumors
179 with myeloid dominance (M; macrophage/monocyte); and 4) subtype of tumors with T cell dominance
180 (T; $CD8^+$, $CD4^+$, T regulatory cells). Using logistic regression, we did not find any association of the
181 four different patient strata to tumor stages (early I/II *vs.* advanced III/IV), sex, or other features.
182 However, patients in the M (p=0.013) or B (p=0.089) subcluster were associated with a higher
183 likelihood of being in an early tumor stage (I/II). In addition, the CD8 (p=0.026) subcluster was
184 associated with a higher likelihood of the tumor being MSI compared to MSS.
185
186 In order to identify tumor-cell based TME imprinting characteristics, we next analyzed differentially
187 enriched pathways[25] in the cancer cells of each immune infiltration type. The immune B and M subtype
188 showed significant downregulation (FDR < 0.1) of the androgen pathway (Figure 2B).
189

190 **Integration of bulk RNA-seq data reveals genotype-immune phenotype associations**
191 The majority of the scRNA-seq studies lack both comprehensive cancer genotype information and
192 survival data. In contrast, the TCGA reference dataset includes this information, alongside bulk RNA-
193 sequencing data. Using the computational method SCISSOR[26], we integrated bulk RNA-seq data from
194 TCGA with single-cell transcriptomic signatures to assess the relationship between specific genotypes
195 and survival outcomes in CRC (Figure 3A). We evaluated the association of cell type transcriptomic
196 signatures derived from the atlas with genotype and survival data from the TCGA reference dataset,
197 which includes 629 patients (UICC I-IV)[5]. Building on our previous work that identified genomic
198 features influencing immune phenotypes in a pan-cancer study using bulk RNA-seq data[14], we leveraged
199 the high resolution of the single-cell CRC atlas for a more detailed analysis of these determinants. We
200 therefore analyzed the cell type transcriptomic signatures and their association with survival in 629
201 patients from the TCGA cohort. Overall, the presence of high levels of IgA^+ plasma cells, $CD4^+$ T cells
202 $CD8^+$ T cells and B cells showed a prominent association with improved survival in CRC patients.
203 Conversely, CRC patients with tumors enriched in Schwann cells were associated with worst survival
204 outcomes, as also reported previously[30]. Notably, high infiltration of various endothelial cells and
205 fibroblasts also emerged as strong negative predictors of survival (Figure 3A). Mutations in *BRAF* and
206 *KRAS* were associated with distinct immune infiltrate patterns (Figure 3B-C). For example, CRC
207 patients with *BRAF* mutations exhibited high infiltration of cDC1 and cDC2, as well as increased levels
208 of IgM^+ plasma cells. Similarly, IgM^+ plasma cell infiltration was also elevated in patients with *KRAS*
209 mutations. Conversely, *BRAF* mutations were linked to low levels of Plasma IgA infiltration, while
210 *KRAS* mutations were associated with high Plasma IgA infiltration. Additionally, MSI and MSS status
211 displayed opposing cell type infiltration patterns (Figure S3A and S3B).
212
213 Finally, we identified gene signatures based on the patient immune infiltration groups by selecting the
214 top 15 differentially expressed genes. These gene signatures were then used to conduct a survival
215 analysis utilizing data from the TCGA COAD-READ cohorts (n = 629). The Kaplan-Meier curves in
216 Figure 2D demonstrate that higher expression of the gene signatures of the B (p-value = 0.0015), M (p-
217 value = 0.0096), and T (p-value = 0.00078) subgroups is significantly correlated with improved overall
218 survival.
219
220
221
222 **Comprehensive characterization of neutrophils reveals highly diverse and plastic phenotypes**
223 We then carried out comprehensive characterization of the data we generated using the BD Rhapsody
224 platform. The obtained scRNA-seq dataset comprises 126,991 single cells categorized into 23 main cell
225 types based on specific marker genes, including 55,245 epithelial/tumor cells, 4,428 stromal/endothelial
226 cells, and 65,782 immune cells, including 30,924 neutrophils (Figure 4A). Compositional analysis using

227 scRNA-seq and flow cytometry showed high concordance when comparing blood, normal, and tumor
228 tissue samples (Figure 4B). As expected, neutrophils were the most abundant leukocyte subtype in blood
229 samples (60–70%), followed by T lymphocytes and monocytes (Figure 4C). Normal-adjacent tissue was
230 dominated by epithelial cells, while tumor tissue contained a markedly higher proportion of immune
231 cells. Notably, neutrophils comprised ~20% of all cells and ~40% of leukocytes in tumor tissue but were
232 present only in small numbers in normal-adjacent tissue (Figure 4C). Also within the granulocytes
233 lineage, neutrophils were the most abundant cells (26.2% of all cells) followed by eosinophils (1.2%)
234 and mast cells (0.4%).
235
236 We observed distinct transcriptomic differences between blood *versus* tumor-associated neutrophils
237 (TAN) (Figure 4D) with 2,214 genes differentially expressed (padj<0.05; Figure 4E). We evaluated
238 previously described marker genes distinguishing TAN from normal tissue neutrophils (NAN) across
239 cancers, including *OLR1*, *VEGFA*, *CXCR2*, and *PLAU*[27]. TAN markers (*OLR1*, *PLAU*, *VEGFA*) were
240 highly upregulated in TAN, whereas *CXCR2*, a marker of mature neutrophils, was highly expressed in
241 blood neutrophils but downregulated in TAN (Figure 4F). Spatial analysis of Lectin-type oxidized LDL
242 receptor 1 (LOX-1; encoded by *OLR1*) expression using multiplex immunofluorescence (IF) staining
243 demonstrated the presence of segregated LOX-1^{+} neutrophil cell clusters exclusively in CRC tumor
244 tissue, but not in normal-adjacent colonic tissue (Figure 4G; Supplementary Figure S4A) LOX-1^{+} cells
245 co-express the neutrophil marker CD15). Of note, LOX-1 has previously been identified as a marker for
246 TAN in lung cancer[15], indicating LOX-1 as a putative tumor-agnostic TAN marker.
247
248 Sub-clustering of neutrophils revealed three distinct blood neutrophil clusters (BN1-BN3) and four TAN
249 clusters (TAN1-TAN4) (Figure 4H). Trajectory analysis indicated a progressive transition from BN1 to
250 BN3 and subsequently to TAN1 (Figure 4I), demonstrating that neutrophil migration from the
251 bloodstream into tissue is accompanied by distinct transcriptomic changes. From TAN1, neutrophils
252 diverged into two trajectories: one leading to a putative pro-tumor TAN phenotype (TAN4)
253 characterized by high *IL1A* expression, a key driver of emergency myelopoiesis that fuels inflammation
254 and immunosuppression in the TME[28], and another toward a phenotype defined by high expression of
255 antigen-presentation-associated markers, such as *HLA-DRA* (TAN2/TAN3) (Figure 4J). Neutrophils
256 with antigen-presenting capabilities, referred to as hybrid neutrophils[27], have previously been identified
257 in early-stage lung cancer and are associated with an anti-tumor phenotype[15,29]. Interestingly, the
258 expression of *HLA-DRA* in blood or tissue-resident neutrophils was significantly higher in early-stage
259 CRC compared to advanced-stage CRC (Figure 4K), suggesting that these antigen-presenting
260 neutrophils may be diminished as the tumor progresses to more advanced stages, potentially contributing
261 to immune evasion. Finally, we validated HLA-DR protein expression using flow cytometry and
262 confirmed that the expression was markedly higher in TAN compared to both blood neutrophils and
263 NAN (Figure 4L).

264

265 **Spatial single-cell multi-modal profiling reveals neutrophil clusters within multicellular niches**

266 Increasing evidence indicates that immune cells in the TME occupy spatially distinct areas or patches
267 characterized by specific cellular compositions and discrete histological properties: so-called cellular
268 neighborhoods[22] or cell niches. Evidence for this type of segregated myeloid populations in CRC was
269 recently shown using a limited immunofluorescence (IF) panel[30] as well as in our data (Figure 4G). In
270 order to quantify spatial niches we analyzed the TME of MSS CRC patients using a spatial
271 transcriptomic technique with 380 genes and sub-cellular resolution (10x Genomics Xenium). The
272 spatial single-cell transcriptomic profiling revealed distinct spatial architecture and cell type
273 composition in both, center of the tumor and the invasive margin (Supplementary Figure 5A). Spatial
274 co-localization of T cell and B cell clusters pinpointed tertiary lymphoid structures in these samples
275 (Figure S5B-C). Neutrophils were the other immune cell type that was organized as clusters in the tumor
276 and the invasive margin (Figure 5B). This spatial organization of neutrophils was confirmed using
277 immunohistochemistry (IHC) stainings of the tissue from the same patients (Supplementary Figure 5D).

278 In order to quantify spatial multicellular interactions we used the graph deep-learning method
279 NicheCompass that utilizes existing knowledge of inter- and intracellular interaction pathways to
280 identify distinct cellular niches[31]. At the resolution level of two niches, we identified cancer and non-
281 cancer niches (Supplementary Figure 5E), whereas at the resolution of three niches we identified cancer,
282 boundary, and stroma niches (Supplementary Figure 5F). At the resolution level of ten we quantified
283 five cancer-cell dominated cell niches (CN1-CN5), two fibroblast-dominated niches (CN6-7) and a
284 plasma-cell-enriched niche (CN8) comprising the tumor-stroma boundary regions, a stromal niche
285 (CN9) and an immune-cell dominated niche that identified tertiary lymphoid structures (CN10) (Figure
286 5A). The cancer-cell dominated niche CN5 was enriched with macrophages and neutrophils (Figure
287 5B). Analysis of cancer cells within this niche revealed elevated expression of the chemokines *CXCL1*,
288 *CXCL3*, *CXCL10*, and *CXCL11*, along with the immunomodulatory enzyme *IDO1* (Figure 5C).
289 Neutrophils may amplify *IDO1* expression in cancer cells[32], contributing to an immunosuppressive
290 milieu. The identified chemokines, which recruit neutrophils through CXCR2 and CXCR3 signaling[33],
291 likely play a central role in driving cancer-cell mediated localized neutrophil aggregation and cluster
292 formation.

293 We then used an orthogonal method based on imaging mass-cytometry (IMC) and profiled FFPE
294 samples from patients with stage III colon cancer treated with adjuvant chemotherapy (CAPOX) and for
295 whom relapse and overall survival data were available (n=30). We previously developed and evaluated a
296 panel of antibodies for IMC on FFPE samples for a comprehensive overview of the TME and cancer-
297 immune cell interactions[34], including lineage and functional immune cell markers, surrogates of cancer
298 cell states (proliferation, apoptosis) and structural markers (epithelium, stroma, vessels). Following data-
299 driven identification of single-cell phenotypes segmentation, and image analysis, we quantified the

300 densities of major cell types (Figure 5D). The spatial protein profiling revealed similar cellular
301 composition of the TME as identified using single-cell mRNA profiling. Given the limited number of
302 features, quantification of cell niches based on protein expression was not possible. We therefore used
303 a statistical approach to define niches and applied cell neighborhood analysis by defining cell nuclei and
304 associating polygons (Voronoi diagram) to each nucleus[34] (see Methods), thereby allowing cells of
305 different sizes and distances to be assessed as neighbors. We used a permutation approach[34] to identify
306 pairwise interactions between cell phenotypes that occurred more or less frequently than expected by
307 chance. This spatial information revealed neutrophil clusters similar to the clusters quantified using
308 transcriptomic data (Figure 5E). It appears that the neutrophil clusters are more frequent in the tumor
309 compared to the invasive margin, albeit the differences in this cohort were not significant (Figure 5H).
310 Analyses of the survival data showed that the neutrophil clusters in both regions are associated with
311 worse prognosis (Figure 5I).

312 Overall, the multimodal profiling revealed that neutrophils are not randomly distributed but are localized
313 in spatial clusters. Moreover, neutrophil clusters are situated within distinct multicellular niches
314 composed of different fractions of cancer cells, endothelial cells, and other immune cells.

315

316 **Mouse neutrophils recapitulate human neutrophil phenotypes in various compartments**

317 To carry out functional studies and define underlying mechanisms, we deployed an orthotopic mouse
318 model based on tumor organoid transplantation (Figure 6A-B). AKPS organoids – harboring the four
319 most common CRC-driving mutations (*Apc*-LOF (loss of function), *p53*-LOF, *Kras*-GOF (gain of
320 function), *Smad4*-LOF) – were successfully engrafted following colonoscopy-guided injection into the
321 colonic mucosa of C57BL/6JRj mice[35] (Supplementary Figure 6A). Tumors developed within six to
322 eight weeks (Figure 6C) and progressively expanded through the colonic layers toward the peritoneal
323 cavity (Supplementary Figure 6B), mirroring some of the invasive characteristics of late-stage human
324 CRC. Notably, neutrophil infiltration was significantly higher in the tumor compared to the healthy
325 colon (Figure 6D-F). Interestingly, there was already an increase of neutrophils within the blood of
326 AKPS-tumor bearing compared to naive mice (Figure 6G) suggesting a systemic effect during tumor
327 development. Analysis of neutrophil activation markers like Siglec-F, Cd11b, CD80 and PD-L1 further
328 showed a progressive activation of neutrophils from the blood to the tumor (Figure 6G-J) while MHC-
329 II proteins were upregulated in tumor and colon tissue compared to blood neutrophils (Figure 6L).
330 Overall, the results confirm the validity of the ortothopic mouse model for human CRC.

331

332 **Cross-species analyses reveal tumor-induced granulopoiesis in CRC**

333 In order to carry out cross-species analyses of mouse and human CRC neutrophils, we performed
334 scRNA-seq on CD45-enriched leukocytes isolated from tumors and blood of mice six weeks post-
335 AKPS-injection. Colons and blood of naïve (control) mice were further included as controls.

336 Additionally CD45-enriched leukocytes were isolated from the bone-marrow of both tumor and non-
337 tumor bearing mice. Neutrophils from all tissues and phenotypes were extracted and integrated (Figures
338 7A andB). Clustering analysis identified 10 clusters with distinct transcriptional patterns (Figures 7C
339 and D) and a high concordance of BN and TAN marker genes derived from our human neutrophil dataset
340 (Supplementary Figure 7A), especially within the colon and tumor specific clusters. We then performed
341 pseudotime trajectory analysis, which showed a progressive transition of neutrophils from the bone-
342 marrow to the blood and colonic tissues (Figures 7E and F). Differential gene expression (DEG) analysis
343 between tumor and healthy bone-marrow and blood revealed that neutrophils already changed their
344 transcriptional phenotype prior to their recruitment into the tumor tissue (Figures 7G and H). Gene set
345 enrichment analysis (GSEA) of KEGG terms showed changes in the metabolic phenotypes
346 (Supplementary Figure 7C). Neutrophils from the tumor bone-marrow upregulated multiple genes
347 encoding for ribosomal proteins (*Rps2, Rpl32, Rpsa, Rps27l)* and proteins involved in cell proliferation[36]
348 (*Ccnb2, Mcm4, Pcna, Fen1*). Overall, the results from the single-cell profiling suggest tumor-induced
349 granulopoiesis which is likely driven by secreted factors.
350
351 To provide further evidence for tumor-induced granulopoiesis we carried out scRNA-seq profiling using
352 blood samples from healthy volunteers and compared the expression profiles of neutrophils with the
353 expression profiles of blood neutrophils from the CRC patients (Figure 7I). The results showed profound
354 changes in the expression with 2287 genes differentially expressed (padj<0.05; Figure 7J). GSEA on
355 KEGG terms indicated changes in the metabolic phenotypes. We then carried out functional experiments
356 with neutrophils isolated from blood of healthy donors and treated the neutrophils with supernatants
357 from a colorectal adenocarcinoma cell line (CaCO) and a non-small cell lung cancer cell line (SW900).
358 Seahorse analysis showed different O_2 consumption rates (Figure 7K), indicating context-dependent
359 metabolic rewiring of the neutrophils.
360
361 Together, these findings provide evidence of tumor-induced granulopoiesis and metabolic rewiring
362 which is likely driven by secreted factors from the TME.

363 **DISCUSSION**

364 We built a large-scale atlas of single-cell transcriptomes of CRC through integration of 4.27 million
365 cells from 650 patients and 1,670 samples representing 7 billion expression values. We then
366 complemented the atlas with single-cell data enriched with immune cells with low-mRNA content
367 including neutrophils, and carried out multi-modal spatail profiling at the single-cell level of the TME
368 in CRC. Our unique resource provides novel biological findings and has important translational
369 relevance.

370

371 First, we provide a high-resolution view of the TME in CRC with 61 major cell types/states and show
372 different cell type composition patterns including more precise functional transcriptomic classification
373 of malignant epithelial cells. The single-cell composition of the TME in CRC enabled refined tumor
374 classification and patient stratification into four immune phenotypes: immune deserted, myeloid, B cell,
375 and T cell subtypes. These findings may have important implications for enabling cancer
376 immunotherapy in CRC. For example, combination therapies that target both, myeloid cells and
377 lymphoid cells could represent an immunotherapeutic strategy to treat M subtypes as shown recently in
378 a melanoma mouse model[37]. Similarly, given the heterogeneity of the intratumoral B cells and the
379 importance of tertiary lymphoid structures[38], further analysis of the B cell subtypes might open
380 promising therapeutic avenues by additional refined B cell targeting.

381

382 Second, we characterized for the first time neutrophil subsets in various compartments in CRC patients
383 and show remarkable diversity and plasticity. Of particular interest for cancer immunotherapy is the
384 TAN phenotype with immunogenic antigen-presenting feature. This observation implies acquisition of
385 antigen presenting-like properties by neutrophils at the tumor site, as previously reported[39]. Such
386 conversion of neutrophils to antigen presenting cells may elicit anti-tumor immunity and has recently
387 been shown in a murine model[40].

388

389 Third, multi-modal spatial profiling revealed that neutrophils build clusters of cells within distinct
390 multicellular niches. Similar neutrophil clustering has been recently observed in the context of wound
391 healing using a zebrafish model[41]. The driving force for this type of myeloid structures remains unknown
392 and encourages further studies in the context of CRC and possibly other solid cancers.

393

394 Finally, using an orthotopic mouse model and cross-species analyses, we show that tumor-derived
395 signals modulate the neutrophils in the bone-marrow, implicating tumor-induced granulopoiesis.
396 Expression and functional data indicate metabolic rewiring in the bone marrow and imprinting the
397 neutrophil cell fate via progenitors.

398 Beyond these biological insights, the results from this study have also important translational
399 implications. As shown here, TANs can acquire antigen presenting properties and such conversion of

400 abundant neutrophils to antigen-presenting cells could overcome the limitations of the low abundance
401 of cross-presenting DCs. Additionally, our study highlights CRC as a systemic disease that requires
402 therapeutic strategies beyond the primary disease site. Such a strategy was recently shown for lung
403 cancer[42]. However, as we show here, the disease context has profound effects on the neutrophil
404 phenotypes and hence, requires further disease-specific investigations.
405
406 In conclusion, we provide a CRC atlas with single-cell resolution as well as a web portal that enables
407 interactive exploration of the dataset through cell-x-gene (https://crc.icbi.at) that allows visualization of
408 metadata and gene expression. The biological insights we present here and future discoveries arising
409 from the exploitation of the high-resolution CRC atlas could provide the basis for developing
410 combination therapies for CRC patients who are not sufficiently responding to immune checkpoint
411 blockers.

ACKNOWLEDGMENTS

This work was supported by the European Research Council (grant agreement No 786295 to ZT), by the Austrian Science Fund (FWF) (project I3978 to ZT). ZT is a member of the German Research Foundation (DFG) project TRR 241(INF). The results published here are in part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga. VM was funded by the Austrian Science Fund (FWF) 10.55776/DOC82. ICA was funded through an Eccellenza Professorial Fellowship from the Swiss National Science Foundation PCEFP3_187021 and through a Tandem Grant (2022 call) of the ISREC Foundation.

422 **FIGURE LEGENDS**

423

424 **Figure 1. Building a large-scale single-cell atlas for CRC.** (**A**) Summary of the data integration and
425 analysis workflow. (**B**) Number of samples for each type. (**C**) Overview of the cell fractions in different
426 tissues. (**D**) Overview of the CRC atlas and the epithelial, and stromal/endothelial components depicted
427 as UMAP plots. (**E**) Overview of the CRC atlas and the immune components depicted as UMAP plots.

428

429 **Figure 2. Tumor immune phenotypes in CRC.** (**A**) Patient characteristics and stratification of the
430 tumor immune phenotypes. (**B**) Differential activation of PROGENy cancer pathways across different
431 cell types in the TME. Heatmap colors indicate the deviation from the overall mean, independent of
432 tumor histology and stage. White dots indicate significant interactions at different false-discovery-rate
433 (FDR) thresholds. Only cytokine signatures with an FDR < 0.1 in at least one patient group are shown.
434 (**C**) Kaplan-Meier curves of the survival analysis according to the patient immune infiltration groups
435 (B, desert, M, T), performed using the molecular data of the TCGA COAD-READ cohorts (n = 461).

436

437 **Figure 3. Association of cellular composition and distinct genotypes and survival in the TCGA**
438 **data.** (**A-E**) *SCISSOR* analysis relating phenotypic information from bulk RNA-seq data from TCGA
439 with single cells. UMAP plots indicate the position of cells positively (blue) or negatively (red)
440 associated with mutation or better survival. A log2 ratio > 0 indicates a positive association with
441 mutation, or better survival, respectively. Shown are cell-types with a log2 ratio significantly different
442 from 0 at an FDR < 0.01 (paired Wilcoxon test). (**A**) Association of cellular composition with overall
443 survival. (**B**) Association of cellular composition with BRAF mutation in CRC patients (n = 136). (**C**)
444 Association of cellular composition with KRAS mutation in CRC patients (n = 489).

445

446 **Figure 4. Neutrophil characterization in CRC.** (**A**) UMAP of the generated CRC scRNA-seq dataset
447 (n = 12 patients; matched blood, tumor tissue, and normal-adjacent tissue) colored by cell type. (**B**) Cell-
448 type fractions obtained by scRNA-seq and by flow cytometry. (**C**) Neutrophil fractions (as percentage
449 of leukocytes) obtained by flow cytometry (n = 12 patients; matched blood, tumor tissue, and normal-
450 adjacent tissue). (**D**) UMAP of neutrophils from blood and tumor tissue analyzed by scRNA-seq. (**E**)
451 Volcano plot of the gene expression from blood and tumor-associated neutrophils (TANs). (**F**)
452 Expression of *OLR1*, *PLAU*, *CXCR2*, and *VEGFA* in neutrophils from blood and tumor tissue, analyzed
453 by scRNA-seq. (**G**) Selected multiplex immunofluorescence (mpIF) staining of LOX-1 (pink), pan-
454 cytokeratin (turquoise), and DAPI (blue) in tumor tissue and matched normal-adjacent colorectal cancer
455 (CRC) tissue. Scale bar: 100 μm. See also Figure S4. (**H**) Top 10 markers for BN1-BN3 and TAN1-
456 TAN4 sub-clusters. The marker gene quality is reflected by the area under the receiver operator
457 characteristics curve (AUROC; 1 = marker gene perfectly distinguishes the respective cluster from other
458 clusters in all patients; AUROC 0.5 = no better than random). (**I**) Partition-based graph abstraction

459 (PAGA) based on RNA velocities, projected on the UMAP plot of neutrophils from blood and tumor
460 tissue. Edges represent cell-type transitions called by PAGA. (**J**) Expression of *HLA-DRA* in neutrophils
461 from blood and tumor tissue, analyzed by scRNA-seq. **K**) Expression of *HLA-DRA* in neutrophils from
462 blood, tumor tissue, and normal tissue.. (**L**) Quantification of HLA-DR expression on neutrophils by
463 flow cytometry across blood, normal-adjacent tissue, and tumor tissue. Each dot represents the mean
464 value for an individual patient (n = 12).
465
466 **Figure 5. Spatial single-cell multi-modal profiling.** (**A**) Cell niches identified from imaging-based
467 spatial transcriptomics profiling of five CRC tumor sections (CN1-4: Intratumoral niches, CN5:
468 Intratumoral neutrophil-enriched niche, CN6-8: Tumor-stroma boundary, CN9: Stroma, CN10: TLS).
469 (**B**) Cell type composition of each niche. (**C**) Top 5 genes upregulated in cancer cells of CN5. (**D**)
470 Cellular neighborhoods of neutrophil clusters. (**E**) Fraction of all neutrophils and of neutrophils in
471 clusters within all cell types per niche.(**F**) UMAP of the imaging mass-cytometry data using 40 markers
472 and samples from 30 CRC patients.(**G**) Neutrophil niches identified using the imaging mass-cytometry
473 data at the tumor core and invasive margin. (**H**) Violin plot of the neutrophil niches by location. (**I**)
474 Kaplan-Meier curves for the neutrophil niches identified at the tumor core and invasive margin.
475
476 **Fig. 6. Neutrophils in a murine orthotopic organoid transplant model for CRC.** (**A-B**) Mouse AKPS
477 tumor organoid cells harbouring the four most common CRC driver mutations (APC^{KO}, $KRAS^{G12D}$,
478 $P53^{KO}$, $SMAD4^{KO}$) were injected intra-rectally into 8- to 10-week-old C57BL/6J males. At the study
479 endpoint (4-6-weeks post-injection), blood, tumor and adjacent colonic tissue were harvested for flow
480 cytometric analysis. (**C**) Survival curve 4 weeks post-injection. (**D**) Neutrophil numbers per milligram
481 tissue and frequency among live CD45+ leukocytes (right). (**E**) Representative plots of neutrophil
482 frequencies among CD45+ cells, as determined by flow cytometry. (**F**) Neutrophil frequencies among
483 CD45+ cells in the blood of naïve versus AKPS-injected mice. (**G**) Frequencies and (**H**) representative
484 flow-cytometry plots of Siglec-F-expressing neutrophils among total neutrophils in blood, colon and
485 tumour of AKPS-bearing mice. Mean Fluorescence Intensity (MFI) of cell-surface CD11b integrin (**I**)
486 or CD80 and PDL1 (**J**) expression in neutrophils. (**K**) MFI (left) and frequencies of MHC II expression
487 among neutrophils. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; as calculated by two-
488 tailed unpaired Student's t-test for comparison of two groups (**D, F**) or by one-way ANOVA for
489 statistical comparison of more than two groups (**G-K**). n=5, medians are displayed.
490
491 **Figure 7. Neutromap of mouse CRC.** Uniform manifold approximation and projection (UMAP) of
492 integrated neutrophils from tumor, colon, blood and bone marrow of AKPS tumor bearing and non-
493 tumor bearing mice. Cells are colored by (**A**) phenotype, (**B**) tissue type (colon is a combination of
494 colon, adjacent colon, tumor, and disseminated tumor) and (**C**) clusters generated via the Seurat
495 clustering algorithm. (**D**) Heatmap depicting the scaled average gene expression of the top 10

differentially expressed genes (DEGs) per cluster. Columns are clustered by tissue presence. (**E**) Heatmap showing the scaled average gene expression by cluster of the ortholog genes representing human neutrophil clusters. Columns are clustered by tissue presence. (**F**) UMAP representing RNA velocity with cells colored by tissue type. The arrows indicate the velocity flow. (A-F) n = pool of 6 mice each condition and tissue. (**G**) Volcano plot of gene expression data from bone-marrow neutrophils from tumor-bearing mice and control mice. (**H**) Volcano plot of gene expression data from blood neutrophils from tumor-bearing mice and control mice. (**I**) UMAP of blood neutrophils from CRC patients and controls. (**J**) Volcano plot of gene expression data from blood neutrophils from CRC patients and controls. (**K**) OCR (oxygen consumption rate) calculated from AUC of XF neutrophil activation assay kinetic traces of healthy human neutrophils activated with PMA, after 24h culture in SW900 supernatant, or CaCo2 supernatant or RPMI control medium. The data shown are the mean and SD, tested for differences using t test, n=3 biological replicates.

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Antibodies list (human samples)		See Table S2
Antibodies list (mouse samples)		See Table S2

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Resected colorectal tumors and adjacent healthy colorectal tissue (CRC) samples, naive/untreated	Tirol Kliniken Innsbruck	N/A
Whole blood samples from patients undergoing colorectal tumor resections	Tirol Kliniken Innsbruck	N/A
Human neutrophils from healthy donors	This study	N/A
Chemicals, peptides, and recombinant proteins		
StemPro hESC SFM	Gibco	Cat# A1000701
Penicillin-Streptomycin	Sigma-Aldrich	Cat# P4333
Primocin	InvivoGen	Cat# ant-pm-2
Fetal bovine serum (FBS), sterile-filtered	Sigma-Aldrich	Cat# F7524
Dulbecco's phosphate-buffered saline, without calcium and magnesium chloride, sterile-filtered, DPBS	Sigma-Aldrich	Cat# D8537
DMEM, high glucose, pyruvate, no glutamine	Gibco	Cat# 21969-035
L-glutamine, 200 mM, 100 mL	Sigma-Aldrich	Cat# G7513
Liberase DH, research grade	Roche	Cat# 5401054001
DNAse I, grade II, from bovine pancreas	Roche	Cat# 10104159001
OptiMEM	Gibco	Cat# 31985062
Matrigel	Corning	Cat# 354277
RPMI-1640 Medium	Gibco	Cat# 11875093
HEPES	Sigma-Aldrich	Cat# H0887
Collagenase type IV	Sigma-Aldrich	Cat# C5138
Collagenase type VIII	Sigma-Aldrich	Cat# C2139
Bovine Serum Albumin, BSA	Roche	Cat# 10738328103
Penicillin/streptomycin	Sigma-Aldrich	Cat# P0781
EDTA	SIgma-Aldrich	Cat# P0781
Percoll	Cytiva	Cat# 17089101
RPMI-1640 Medium	Sigma-Aldrich	Cat# R0883
Seahorse XF RPMI assay medium pack	Agilent Technologies	Cat# 103681
Seahorse XFp Cell Culture Miniplates	Agilent Technologies	Cat# 103022
Seahorse XFp FluxPak	Agilent Technologies	Cat# 103022
Poly-L-lysine	Sigma-Aldrich	Cat# P4832
Rotenone/Antimycin A	Agilent Technologies	Cat# 103010
Phorbol 12-myristate 13-acetate (PMA)	Fisher Scientific	Cat# J63916.MCR
Critical commercial assays		
Agencourt AMPure XP magnetic beads	Beckman Coulter	Cat# A63880
BD Human Single-Cell Multiplexing Kit	BD Biosciences	Cat# 633781
BD Rhapsody Cartridge Kit	BD Biosciences	Cat# 633733
BD Rhapsody cDNA Kit	BD Biosciences	Cat# 633773
BD Rhapsody Enhanced Cartridge Reagent Kit	BD Biosciences	Cat# 664887
BD Rhapsody WTA Amplification Kit	BD Biosciences	Cat# 633801
BD Stain Buffer (FBS)	BD Biosciences	Cat# 554656
High Sensitivity D1000 Reagents	Agilent	Cat# 5067-5585
High Sensitivity D1000 Screen Tape	Agilent	Cat# 5067-5584
Qubit dsDNA HS Assay Kit	Invitrogen	Cat# Q32854
BD Pharm Lyse Lysing Buffer	BD Biosciences	Cat# 555899
Calcein AM	Invitrogen	Cat# C1430
Draq7	BD Pharmingen	Cat# 564904
Anti-CD45 microbeads	Miltenyi Biotec	Cat# 130-052-301

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BD Mouse Single-Cell Multiplexing Kit	BD Biosciences	Cat# 633793
BD Rhapsody Whole Transcriptome Analysis Amplification Kit	BD Biosciences	Cat# 633801
Qubit dsDNA HS Kit	Thermo Fisher Scientific	Cat# Q32851
TapeStation 4200 HS D5000 tapes and reagents	Agilent Technologies	Cat# 5067-5592
NovaSeq 6000 SP Reagent Kit v1.5 (100 cycles)	Illumina	Cat# 20028401
Opal 6-plex Detection Kits	Akoya Biosciences	Cat# NEL821001KT
BOND Epitope Retrival 1	Leica Biosystems	Cat# AR9961
BOND Epitope Retrival 2	Leica Biosystems	Cat# AR9640
BOND Dewax Solution	Leica Biosystems	Cat# AR9222
BOND Wash Solution 10x	Leica Biosystems	Cat# AR9590
Spectral DAPI	Akoya Biosciences	Cat# FP1490
Prolong Diamond Antifade	Thermo Fisher	Cat# P36961
BOND Research Detection System	Leica Biosystems	Cat# DS9455
BOND Titration Kit	Leica Biosystems	Cat# OPT9049
Xenium Human Immuno-Oncology Profiling Panel	10x Genomics	Cat# 1000654
Xenium Slides & Sample Prep Reagents	10x Genomics	Cat# 1000460
Xenium Cell Segmentation Add-on Kit	10x Genomics	Cat# 1000662
Xenium Decoding Reagents	10x Genomics	Cat# 1000461
Xenium Cassettes	10x Genomics	Cat# 1000566
Xenium Decoding Consumables	10x Genomics	Cat# 1000487
MACSxpress Whole Blood Neutrophil Isolation Kit	Miltenyi Biotec	Cat# 130-104-434
Deposited data		
MUI_Innsbruck	This study	Zenodo link
Processed input data	This study	Zenodo link
Core atlas as cell-x-gene instance and h5ad/Seurat file formats with standardized metadata according to the cell-x-gene schema	This study	https://cellxgene.cziscience.com/collections/3a844375-60ea-474b-adf1-98e76928baee
Bian_2018_Science (fastq)	SRA	PRJNA382695
Borras_2023_Cell_Discov (processed)	synapse	syn41687327
Burclaff_2022_Cell_Mol_Gastroenterol_Hepatol (processed)	SRA	PRJNA768275
Che_2021_Cell_Discov (fastq)	SRA	PRJNA725335
Chen_2021_Cell (processed)	synapse	syn26044148
Chen_2024_Cancer_Cell (processed)	GEO	PRJNA991601
Conde_2022_Science (fastq)	ENA	PRJEB51634
Elmentaite_2021_Nature (fastq)	ENA	PRJEB40273
GarridoTrigo_2023_Nat_Commun (fastq)	SRA	PRJNA886695
Giguelay_2022_Theranostics (processed)	SRA	PRJNA666199
Guo_2022_JCI_Insight (fastq)	SRA	PRJNA779978
Harmon_2023_Nat_Cancer (processed)	SRA	PRJNA863562
He_2020_Genome_Biol (fastq)	SRA	PRJNA670909
HTAPP_HTAN (processed)	synapse	syn24181445
Huang_2024_Nat_Cancer (processed)	NGDC	PRJCA007578
James_2020_Nat_Immunol (fastq)	ENA	PRJEB32912
Ji_2024_Cancer_Lett (fastq)	SRA	PRJNA915211
Ji_2024_PLoS_Genet (processed)	GEO	GSE217774
Joanito_2022_Nat_Genet (fastq/processed)	EGA	EGAD00001008555

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Khaliq_2022_Genome_Biol (processed)	GEO	PRJNA827831
Kong_2023_Immunity (processed)	Broad Single Cell Portal	SCP1884
Lee_2020_Nat_Genet (fastq/processed)	GEO	PRJNA604751
Li_2017_Nat_Genet (processed)	GEO	PRJNA323703
Li_2023_Cancer_Cell (processed)	GEO	PRJNA846169
Liu_2024_Cancer_Res (processed)	GEO	PRJNA1028966
Mazzurana_2021_Cell_Res (processed)	GEO	PRJNA630996
Parikh_2019_Nature (processed)	SRA	PRJNA477814
Pelka_2021_Cell (processed)	GEO	PRJNA738517
Qi_2022_Nat_Commun (processed)	NGDC	HRA000979
Qian_2020_Cell_Res (processed)		http://blueprint.lambrechtslab.org/
Qin_2023_Cell_Rep_Med (processed)	CNGBdb	CNP0004138
Sathe_2023_Clin_Cancer_Res (processed)		https://dna-discovery.stanford.edu/research/datasets/
Scheid_2023_J_EXP_Med (processed)	Broad Single Cell Portal	SCP1690
Smillie_2019_Cell (processed)	Broad Single Cell Portal	SCP259
Terekhanova_2023_Nature (processed)	synapse	syn53214678
Thomas_2024_Nat_Med (processed)	GEO	PRJNA849977
Tian_2023_Nat_Med (processed)	Broad Single Cell Portal	SCP2079
Uhlitz_2021_EMBO_Mol_Med (processed)	GEO	PRJNA701264
deVries_2023_Nature (processed)	GEO	PRJNA894171
Wang_2020_J_Exp_Med (fastq)	SRA	PRJNA518129
Wang_2021_Adv_Sci (fastq)	SRA	PRJNA937723
Wang_2023_Sci_Adv (processed)	GEO	PRJNA937723
Wu_2022_Cancer_Discov (processed)		http://www.cancerdiversity.asia/scCRLM
Wu_2024_Cell (processed)		http://www.pancancer.cn/neu/
Yang_2023_Front_Oncol (fastq)	SRA	PRJNA972665
Zhang_2018_Nature (processed)	GEO	PRJNA429424
Zhang_2020_Cell (processed)	figshare	https://figshare.com/articles/dataset/CRC_CD45_rda/14820318
Zheng_2022_Signal_Transduct_Target_Ther (processed)	GEO	PRJNA676119
TCGA data RNA-seq and mutation data	GDC	https://portal.gdc.cancer.gov/
Cytosig signatures	Jiang *et al.*, 2021 (PMID:34594031)	https://github.com/data2intelligence/CytoSig/
Experimental models: Cell lines		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Vil-creER^{T2};APC^{KO};Trp53^{KO};Kras$^{G12D/WT}$Smad4^{KO} (AKPS) organoids	Dr. Tomas Valenta and Dr. Hassan Fazilaty (DMLS, University of Zürich, Basler Lab)[35]	N/A
SW900	ATCC	HTB-59
CaCo2	ATCC	HTB-37
Experimental models: Organisms/strains		
Mouse: C57BL/6J Strain #:000664	The Jackson Laboratory	RRID:IMSR_JAX:000664
Software and algorithms		
Seven Bridges - BD Rhapsody™ WTA Analysis Pipeline	BD Biosciences	v2.0
FlowJo™	BD Biosciences	v10.7
GraphPad Prism v9	GraphPad Software	www.graphpad.com
GraphPad Prism v10.4	GraphPad Software	www.graphpad .com
Cellranger	10x Genomics	v7.1.0
nf-core/scrnaseq pipeline	Philip et al., 2022	v2.4.1
nf-core/rnaseq pipeline	Philip et al., 2022	v3.12.0
Nextflow	Di Tommaso et al., 2017	v23.04.0
Singularity/Apptainer	Kurtzer et al., 2017	v1.3.1-1.el9
Nextflow workflow to reproduce this study	this study	https://github.com/icbi-lab/crc-atlas
Atlas web resource	this study	https://crc.icbi.at/
FlowJo Software	BD Life Sciences	v.10.8.0
Prism 10	GraphPad	v10
Bcl2fastq v2.20.0433	Illumina	
STAR v2.5.2b	(Dobin et al., 2023)[43]	
R v4.3	R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.	https://www.R-project.org/
Seurat v5.1.0		
decontX v1.0.0	(Yang et al., 2020)[44]	
SingleR v2.4.1	(Aran et al., 2019)[45]	
Celldex v1.12.0	(Heng et al., 2008)[46]	
Pheatmap v1.0.12		
Orthogene v1.8.0		
Slingshot v2.10.0	(Street et al., 2018)[47]	
Sceasy v0.0.7		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Fiji v1.54f	(Schindelin et al., 2012)[48]	https://imagej.net/software/fiji/downloads
QuPath v0.5.1	(Bankhead et al., 2017)[49]	https://qupath.github.io/
CellProfiler v4.2.8	(Stirling et al., 2021)[50]	https://cellprofiler.org/releases
R v4.4.2	R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.	https://www.R-project.org/
Seahorse XFe Wave Software	Agilent Technolgies	www.agilent-com
Compute node CPU	Intel	Intel(R) Xeon(R) Gold 6342 CPU @ 2.80GHz (2x)
GPU node GPU	Nvidia	NVIDIA A100 80GB PCIe
Other		
50 mL centrifuge tube	Sarstedt	Cat# 62.547.254
Stainless steel surgical blades, sterile, Nr. 22	VWR International GmbH	Cat# 233-54840
pluriStrainer, 100 mm strainer, sterile	pluriSelect	Cat# 43-50100-51
CytoOne 6 cm Petri dish, uncoated	Starlab	Cat# CC7672-3359
Mediware, disposable scalpels, stainless steel, sterile (for tissue dissection)	medizinbedarf.at	Cat# I1 09
Mediware, disposable forceps, sterile (for tissue dissection)	medizinbedarf.at	Cat# H7 301
Mediware pre-powdered gloves	medizinbedarf.at	Cat# 9321
Mediware sterile drapes	medizinbedarf.at	Cat# H2 50
gentleMACS Octo Dissociator with heaters	Miltenyi Biotec	Cat# 130-096-427
gentleMACS Octo Dissociator with heaters, C tubes	Miltenyi Biotec	Cat# 130-093-237
Filter tips, 10 mL, sterile	Starlab	Cat# S1121-3810
Filter tips, 20 mL, sterile	Starlab	Cat# S1120-1810
Filter tips, 100 mL, sterile	Starlab	Cat# S1123-1840
Filter tips, 200 mL, sterile	Starlab	Cat# S1120-8810
Filter tips, 1000 mL, sterile	Starlab	Cat# S1126-7810
Serological pipette, 5 mL	Sarstedt	Cat# 86.1253.001
Serological pipette, 10 mL	Sarstedt	Cat# 86.1254.001
Serological pipette, 25 mL	Sarstedt	Cat# 86.1685.001
Serological pipette, 50 mL	Sarstedt	Cat# 86.1256.001
Water bath	N/A	N/A
Vacuum pump, Vacusafe aspiration system with integrated pump	INTEGRA Biosciences	Cat# 158320
Qubit Fluorometer	Thermo Fisher Scientific	N/A
TapeStation 4200	Agilent Technologies	N/A
NovaSeq6000	Illumina	N/A
FACSymphony A5	BD Biosciences	N/A
LSRII Fortessa	BD Biosciences	N/A
Xenium Analyzer	10x Genomics	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Seahorse XF Analyzer	Agilent Technologies	N/A

528
529
530 **METHODS**
531
532 ***Patient recruitment***
533 Twelve naive tumor and matched adjacent normal tissues from CRC patients who underwent surgical
534 resections at the Department of Surgery, Medical University of Innsbruck, Austria were obtained.
535 Additionally, autologous whole blood samples were collected during surgery. Prior to study enrollment,
536 all patients signed the informed consent for participation in the study. The ethical committee of the
537 Medical University of Innsbruck approved the specimen collection and analysis of patients' data. Tumor
538 samples subjected to imaging mass cytometry (IMC) were obtained from patients in the Salzburg cohort.
539 These patients uniformly had R0-resected stage III colon cancer and received adjuvant CAPOX
540 treatment. This analysis was approved by the Ethics Committee of the Provincial Government of
541 Salzburg, Austria (415-E/2343/5-2018). Patients were otherwise unselected and treated at the
542 Department of Internal Medicine III, Paracelsus Medical University, Salzburg, Austria.
543
544 ***Dissociation of normal and tumor colorectal tissue***
545 Surgically resected human colorectal tumor and adjacent normal tissues were enzymatically and
546 mechanically dissociated into single cells as described in detail in the following publication[51]. Briefly,
547 after supervised sampling of tumor and adjacent normal CRC tissues from a pathologist, the specimens
548 were transported to the lab on ice, washed, weighed and minced into small pieces (<1 mm). The tissue
549 pieces were further enzymatically and mechanically digested by using DMEM medium (Gibco) with
550 Liberase DH (Roche) and DNAse I (Roche/Sigma-Aldrich) on a gentleMACS dissociator with heaters
551 (Miltenyi Biotec) at 37°C for ~45 min (selected program: 37C_h_TDK_1). The obtained single cell
552 solutions were sieved through a 100 µm cell strainer (pluriSelect) and kept on ice.
553
554 ***BD Rhapsody library preparation and sequencing***
555 Freshly isolated single-cells were immediately processed with the BD Rhapsody scRNA-seq platform
556 (BD Biosciences). Red blood cells were removed from dissociated tissues and from whole blood
557 samples using the BD Pharm Lyse lysing solution (BD Biosciences). Cell number and viability were
558 assessed with the BD Rhapsody scRNA-seq platform using Calcein-AM (Invitrogen) and Draq7 (BD
559 Biosciences). The BD Single-Cell Multiplexing Kit (BD Biosciences) was used to combine and load
560 multiple samples (human colorectal tumor tissue, adjacent normal tissue, blood) onto a single BD
561 Rhapsody™ cartridge (BD Biosciences). Sample-tag staining was performed according to the

562 manufacturer's protocol (sample-tag staining at room temperature for 20 min and washing by
563 centrifugation at 400 g for 5 min). Single-cell isolation in microwells (cell load: 20 min incubation at
564 room temperature, 20°C ± 2°C) with subsequent cell-lysis and capturing of poly-adenylated mRNA
565 molecules with barcoded, magnetic capture-beads was performed according to the manufacturer's
566 instructions. Beads were magnetically retrieved from the microwells, pooled into a single tube before
567 reverse transcription. Unique molecular identifiers (UMIs) were added to the cDNA molecules during
568 cDNA synthesis. Whole transcriptome amplification (WTA) and sample-tag sequencing libraries were
569 generated according to the BD Rhapsody single-cell whole-transcriptome amplification workflow. The
570 quantity and quality of the sequencing libraries was analyzed with the Qubit dsDNA HS (High
571 Sensitivity) assay kit (Invitrogen) and the 4200 TapeStation (Agilent) system. Libraries were sequenced
572 on the Novaseq 6000 system (Illumina) targeting a sequencing depth of 50.000 reads/cell.
573
574 ***Generation of the atlas***
575 The data for the core atlas was from previously published studies[52–98]. The selected studies were
576 published between May 2017 and July 2024 and the incorporated datasets were generated with six
577 different sequencing platforms, including the most commonly applied 10x Chromium (10x Genomics)
578 as well as Smart-seq2[99], SMARTer (C1), scTrio-seq2, GEXSCOPE (Singleron), TruDrop[100] and
579 DNBelab C4[101]. We further integrated our own data generated with the microwell-based BD Rhapsody
580 scRNA-seq platform (BD Biosciences). We specifically selected studies using comparable protocols for
581 sample processing and data generation, such as sequencing of whole cells. We did not exclude studies
582 that applied flow cytometry-based cell-sorting prior to sequencing, as these incorporate relevant
583 information on rare cell types[102,103,104]. From non-CRC studies we exclusively included those parts of
584 the published data that were relevant for our atlas: from the Garrido Trigo dataset[105] we only included
585 colon samples, from Burclaff[106], Kong[107], Mazzurana[108], Parikh[109], Smillie, Thomas, and Scheid[110]
586 datasets we only used the control samples.
587
588 ***Integration of scRNA-seq datasets***
589 Individual datasets were merged into a single AnnData object. Since genome annotations partly differed
590 between the datasets, we mapped gene symbols to the ensembl ids of the used annotation version and
591 updated the symbols to the latest version of HGNC. If gene ids were missing from a dataset, the values
592 were filled with zeros. Genes that were missing in more than 19 datasets (25%) were excluded
593 altogether.
594 We integrated the datasets using the scANVI algorithm[111,112], as it has been demonstrated to be one of
595 the top-performing methods for atlas-level integration and to scale to >1M cells[113]. Since scANVI
596 requires cell-type annotations for at least one of the input datasets, we manually annotated two "seed"
597 datasets based on unsupervised clustering as described below. Raw counts were used as input for
598 scANVI. Counts from plate-based assays such as Smart-seq2 were scaled by the gene length as

599 recommended on the scvi-tools website. The scANVI model was initialized with a pre-trained scVI
600 model[114], as recommended in the scvi-tools tutorial. The scVI model was trained on the 6000 most
601 highly variable genes as determined with scanpy's[115] *pp.highly_variable_genes* with parameters
602 *flavor="seurat_v3"* and *batch_key="dataset"*. Each sample was considered as an individual batch for
603 both scVI and scANVI. Other than that the algorithms were run with default parameters.
604
605 ***Unsupervised clustering and cell-type annotation***
606 We computed UMAP embeddings[116] and unsupervised Leiden-clustering[117] with scanpy[115], based on a
607 cell-cell neighborhood graph derived from scANVI latent space. Coarse, lineage-specific clusters were
608 iteratively sub-clustered to identify cell-types at a more fine-grained resolution. Cell type clusters were
609 annotated based on previously reported marker genes[105,118,119]
610
611 ***Comparing cell-type abundances***
612 Since comparing cell-type fractions between groups is challenging due to different characteristics of the
613 datasets and the inherent compositional nature of cell-type fractions, we applied the scCODA[120] model,
614 which addresses this issue. We were interested in the differences between conditions (COAD *vs.*
615 READ). To this end, we ran the scCODA model with the formula ~ *condition* + *tumor_stage* + *dataset*
616 with 500,000 iterations using "cancer cells" as the reference cell-type, where *tumor_stage* is a binary
617 vector classifying datasets into early (stages I-II) and advanced (stages III-IV), and dataset is a
618 categorical vector encoding the different datasets. The final result shows *credible effects* with a false-
619 discovery-rate (FDR) of 0.1.
620
621 ***Patient stratification***
622 We stratified patients into immune phenotypes based on immune cell-type fractions. We selected all
623 patients with primary tumor samples and excluded enriched datasets. Neutrophil fractions were
624 excluded, since they are not appropriately captured in the majority of datasets. Cell-type fractions of
625 primary tumor samples were loaded into a patient × cell-type AnnData container. Dataset-specific batch-
626 effects were removed using a linear model as implemented in *scanpy.pp.regress_out*. Patients were
627 clustered using graph-based Leiden clustering with the "correlation" distance metric for computing the
628 neighborhood graph. Patient clusters were labeled according to their predominant cell-types. In addition
629 to the histological subtypes based on the annotation of the original datasets, we annotated tumor types
630 based on the transcriptomics data according to the most abundant cancer cell cluster.
631
632 ***Differential gene expression testing***
633 We used DESeq2[121] on pseudo-bulk samples for differential expression testing which has been
634 demonstrated to perform well and properly correct for false discoveries[122]. For each cell-type and
635 patient, we summed up transcript counts for each gene. Pseudo-bulk samples consisting of fewer than

10 cells were discarded. We compared primary tumor samples from tumor *vs.* normal (*sample type*), primary tumor samples from the patient subtypes M/B/T/desert (*group*), NANs *vs.* TANs (*cell_type_tan_nan*), and Neutrophil clusters (BN1-3, TAN1-4), including the dataset as a covariate. For comparisons between multiple groups, we used contrasts with sum-to-zero coding. P-values were adjusted for multiple hypothesis testing with independent hypothesis weighting (IHW)[123].

Pathway, TF and cytokine signaling signatures

We performed pathway, transcription factor (TF), and cytokine signaling analysis on primary tumor samples with PROGENy[25,124], and CytoSig[125], respectively. Scores were computed using the *decoupler-py*[126] package.

Transcription factor activity was inferred from the Wald statistics of the differential expression analysis output, using a univariate linear model and the CollecTRI[127] database. Activities were defined as significantly activated or suppressed based on an FDR threshold of less than 0.1. The top 1,000 target genes of the progeny model were used, as recommended for single-cell data. The cytosig signature matrix was obtained from the *data2intelligence/CytoSig* GitHub repository and used with the scoring function implemented in the *progeny-py* package. The methods were run with the options *num_perm=0, center=True, norm=True scale=True,* and *min_size=5.* No permutations were used, as we perform statistics in a separate step at the level of biological replicates. Pathway-, transcription factor-, and cytosig scores were then compared between *condition* (tumor *vs*. normal) and patient *group* (T *vs*. B *vs.* M *vs.* deserted) using an ordinary least-squares (OLS) linear model, as implemented in the statsmodels package[115]. Sc[128]s were aggregated into pseudobulk samples by computing the mean of each variable for each patient and cell-type. Samples consisting of less than 10 cells were discarded. For each variable, we fitted a model with the formulas ~ *condition* + *dataset* + *tumor_stage* or ~ *group* + *dataset* + *condition* + *tumor_stage*, respectively. Coefficients were obtained from the linear model and p-values calculated with the f-test. P-values were adjusted for multiple testing with the Benjamini-Hochberg procedure.

SCISSOR analysis

We used SCISSOR[129] to associate phenotypic data from bulk RNA-seq experiments with our single-cell data. TCGA mutation and gene expression data was obtained from the GDC portal, survival data from SCISSOR[129] was run on primary tumor cells of each patient individually according to the SCISSOR tutorial using mutation data (logistic regression) and overall survival (cox-regression) as dependent variables. A grid search for the alpha-parameter was performed in $2^{-i/2}$ with $i \in [24, 23, \ldots, 2]$ and a cutoff parameter of 0.3. 62 of 674 samples with low overall cell count failed during SCISSOR's Seurat-preprocessing step and were excluded from the subsequent analysis. For each patient and cell-type, we computed the fraction of *scissor+ cells* (i.e. positively associated with a mutation or worse survival), *scissor- cells* (i.e. negatively associated), and *neutral cells* and added a pseudo-count of 0.01. A sample

673 was excluded from a cell-type if it contributed ≤10 cells. For each cell-type, we computed the log2-
674 ratio of scissor+ and scissor- cells as the mean fraction of scissor+ cells *vs.* the mean fraction of scissor-
675 cells. Significant differences were determined by comparing the fractions of scissor+ and scissor- cells
676 with a paired wilcoxon test with *zero_method="zsplit"* as implemented in the scipy package. P-values
677 were Benjamini-Hochberg-adjusted and considered significant at an FDR < 0.01.
678
679 ***Survival analysis***
680 Initially, gene lists were extracted from the patient groups according to their immune infiltration profiles
681 (B, desert, M and T). The gene lists consisted of differentially expressed genes, as returned from DESeq2
682 for the comparisons of each immune infiltration group with the rest, that were filtered using as thresholds
683 the adjusted p-values (< 0.05) and the logFold Change (greater than 0.5 or lower than -0.5).
684 Consequently, the gene lists were condensed into gene signatures expression values using the GSVA
685 (v1.50.5) R package. Once the gene signature expression values were calculated the TCGA COAD-
686 READ cohort (n = 629) was stratified into "high" and "low" expression, using as threshold the optimal
687 cutpoint as computed by the surv_cutpoint() function of the survminer (v0.4.9) R package. The survival
688 analysis was carried out using the survival (v3.7-0) R package and the Kaplan-Meier curves were plotted
689 using the ggsurvplot function of the survminer (v0.4.9) R package.
690
691 ***Quantification and statistical analysis***
692 Statistical analysis was performed using the statsmodels library in Python (scRNA-seq data) or
693 GraphPad Prism (flow cytometry and imaging data) using a linear model, t-test or wilcoxon test as
694 appropriate. Single cell-data were aggregated into pseudobulk samples by biological replicates.
695 Compositional analysis of cell-type fractions was performed using scCODA; survival analysis using the
696 Kaplan-Meier formula in R. P-values for untargeted analyses (DE genes, TFs, or pathways) were FDR-
697 adjusted. Significance levels and more details on the statistical tests are indicated in the figure captions.
698
699 ***Flow Cytometry staining of human samples***
700 Cells isolated from surgically resected CRC tumor and adjacent normal tissues as well as whole blood
701 obtained just prior surgery were stained with a cocktail of the following 20 antibodies (CD3, CD8, CD4,
702 CD14, CD16, CD19, CD15, CD28, CD34, CD31, CD38, CD45, CD56, CD90, CD123, CD161, CD193,
703 CD326, HLA-DR, TCRgd) to define all cell populations of interest. After washing and addition of 5 µL
704 7-AAD, cells were measured on a FACSymphony A5 flow cytometer (BD Biosciences). All source data
705 of the aforementioned antibodies is provided in the key resources table, applied fluorochromes are listed
706 in Table S2. Data were analyzed using FlowJo v10.7 software. For details of the gating strategy, see
707 Figure S8.
708

709 ***Multiplex immunofluorescence of human samples***

710 CRC tumor and tumor-adjacent tissue samples were fixed in 4% formalin and embedded in paraffin.
711 Three-micrometer sections were used for immunofluorescence staining. The immunofluorescence
712 staining on formalin-fixed paraffin-embedded (FFPE) tissue was conducted using the Opal 6-Plex
713 Detection Kit (Cat. No: NEL821001KT, Akoya Biosciences, Menlo Park, USA). A multiplex panel of
714 immune markers was established utilizing antibodies against CD15 (1:200, clone MMA, Abcam), LOX-
715 1/OLR1 (1:200, polyclonal, Sigma-Aldrich), and cytokeratin (1:1000, clone AE1/AE3, Dako; clone C-
716 11, Abcam). The staining procedure was performed using an automated staining system (BOND-RX,
717 Leica Biosystems). Markers were applied sequentially and paired with their respective Opal
718 fluorophores (CD15 – Opal650, LOX-1/OLR1 – Opal570, cytokeratin – Opal690). To visualize cell
719 nuclei, the tissue was stained with 4',6-diamidino-2-phenylindole (spectral DAPI, Akoya Biosciences).
720 The stained slides were scanned using the Mantra 2 Quantitative Pathology Workstation (Akoya
721 Biosciences), and representative images were acquired using the Mantra Snap software version 1.0.4.
722 Spectral unmixing, multispectral image analysis, and cell phenotyping were carried out using the inForm
723 Tissue Analysis Software version 2.4.10 (Akoya Biosciences). Pseudo-DAB staining projections were
724 achieved by converting fluorescent channel to a pseudo-brightfield images using InForm, where
725 fluorescent markers are displayed as pseudo-DAB staining and DAPI as hematoxylin.

726

727 ***Preparation of sections for Xenium spatial transcriptomics***

728 Sections were prepared according to manufacturer's instructions (Tissue Preparation Guide
729 Demonstrated Protocol CG000578) at the Department of Anatomy, Histology and Embryology, Medical
730 University of Innsbruck. 5 µm sections were taken from the FFPE tissue blocks with a microtome
731 (Epredia, HM355S), floated in a 45˚C water bath, and adhered to Xenium slides (10x Genomics).
732 Samples were baked at 42˚C for 3 hours and stored in a desiccator at room temperature, overnight,
733 followed by deparaffinization and decrosslinking, according to manufacturer's instructions
734 (Deparaffinization & Decrosslinking Demonstrated Protocol CG000580).

735

736 ***10x Xenium workflow and data acquisition***

737 Xenium samples were processed according to manufacturer's protocols (Xenium In Situ Gene
738 Expression with Morphology-based Cell Segmentation Staining User Guide CG0000749 and Decoding
739 & Imaging, User Guide CG000584). Samples were stained utilizing the predesigned Human Immuno-
740 Oncology panel from 10x genomics (PN 1000654), as they became available from the manufacturer.
741 Slides were stained with a Xenium Cell Segmentation kit according to manufacturer's instructions (User
742 Guide CG0000749). Briefly, probes of the immuno-oncology panel were hybridized overnight before
743 rolling circle amplification and native protein autofluorescence was reduced with a chemical
744 autofluorescence quencher. Slides were processed on a 10X Xenium Analyzer, with ROIs selected to
745 cover the selected regions. The Xenium Onboard Analysis pipeline v2.0.0 (10x Genomics) was run

746 directly on the instrument for imaging processing, cell segmentation, image registration, decoding,
747 deduplication, and secondary analysis, as previously described[130].
748
749 ***Post Xenium H&E staining***
750 After processing tissue sections through the Xenium workflow, the same sections on the Xenium slide
751 were used for H&E staining, according to the manufacturer's instructions (Post-Xenium Analyzer H&E
752 Staining CG000613) and imaged using a PANNORAMIC slide scanners (Institute for Pathology,
753 Neuropathology & Molecular Pathology, Medical University of Innsbruck).
754
755 ***Xenium data analysis***
756 Output files of the Xenium Onboard Analysis were further processed with scanpy[115], squidpy[131] and the
757 SpatialData framework[132] for region selection, quality control and calculation of spatial neighbor graphs.
758 We applied SpaNorm[133] to normalize transcript counts. Integration of samples was done with scvi on
759 the raw counts, followed by leiden clustering to define distinct cell types. We utilized tangram[134] in
760 cluster mode to annotate leiden clusters by mapping cell type annotations from a downsampled version
761 of the CRC atlas generated with scSampler[135]. NicheCompass[136] was run with OmniPath[137] and
762 NicheNet[138] as prior gene programs. Leiden clustering was then performed at different resolutions on
763 the niches latent space to identify distinct spatial niches. Neutrophil clusters were identified from the
764 spatial neighbor graph (neighbors=20, distance=25 µm) by first defining cluster seeds as cells of cell
765 type 'Neutrophil' with at least three homotypic neighbors. Cluster seeds were subsequently expanded
766 in five iterations to add all neutrophils adjacent to seed cells to the respective cluster. We further defined
767 cluster neighborhoods as all cells in the spatial neighbors graph (neighbors=20, distance=10 µm)
768 connected to a neutrophil cluster.
769
770 ***Imaging mass cytometry and data analysis***
771 Imaging mass cytometry was performed on tissue microarrays from microsatellite-stable CRC patients
772 as previously described[34], using the antibody panel provided in Table S2. Images were obtained from a
773 total of 37 patient samples, including intratumoral regions from 27 patients and tumor margin regions
774 from 29 patients. Among these, both intratumoral and tumor margin regions were imaged for 19 patients.
775 Per sample, one or two 1000x1000µm regions of interest were ablated, depending on tissue availability.
776 Images were visually inspected and exported to ome.tiff using MCD viewer (standard biotools). Image
777 normalization, cell segmentation and phenotype identification were performed as described previously[34]
778 and steps were validated throughout on the original images.
779
780 ***Spatial analysis of granulocyte niches***
781 To analyze the spatial organization in tumor and margin regions, we constructed Voronoi diagrams
782 based on the center coordinates of all detected nuclei. Each Voronoi region was then assigned the

783 corresponding cell type. The DBscan clustering algorithm was used to find niches consisting of at least
784 3 granulocytes in a maximum pairwise distance of 15 µm. For these niches all other cell types that share
785 a Voronoi ridge (edge) with a granulocyte were searched and added as "niche members". To determine
786 if the number of detected niches differed significantly from what would be expected by chance, a Monte
787 Carlo simulation with 1,000 iterations was conducted. In this simulation, the cell type labels for each
788 cell on the imaged slide was randomly assigned while maintaining the number of cells for each type
789 constant and preserving the overall spatial arrangement. A z-score and p-value were then calculated.
790
791 ***Tissue Microarrays and Image Acquisition***
792 Five tissue microarrays (TMAs) containing a total of 549 valid cores were processed and stained for
793 CD15 and HLA-DR using immunofluorescence (IF) and six TMAs containing in total 649 valid cores
794 were processed and stained with MPO and HLA-DR using IF. Images of the stained TMAs and whole
795 slides were acquired in .czi format.
796
797 ***Image Preprocessing and core extraction***
798 The .czi images were converted to a compatible format (.tif) for downstream analysis using Fiji (ImageJ,
799 v1.54f). Images were carefully inspected during this step to ensure no loss of resolution in any of the
800 RGB layers.
801 Individual TMA cores were extracted from the pre-processed images using QuPath (v0.5.1), an open-
802 source software platform for digital pathology. The cores were manually annotated to exclude damaged
803 or improperly stained regions, and only valid cores were included in subsequent analyses.
804
805 ***Cell Identification and Quantification***
806 A customized image analysis pipeline was developed in CellProfiler (v4.2.8) to identify and quantify
807 nuclei and marker-positive cells for CD15 and HLA-DR and in the second batch for MPO and HLA-
808 DR. The pipeline consisted of the following steps:
809 1. Nuclear Identification: Nuclei were segmented and identified based on fluorescence intensity.
810 2. Marker Co-Localization: Positive staining for CD15, MPO and HLA-DR was identified and
811 quantified, with nuclei serving as reference points for spatial localization and co-expression analysis.
812 Output: The pipeline produced quantitative data on nuclei counts, their spatial position and the number
813 of nuclei positive for each marker, based on staining intensity to distinguish specific marker expression
814 from background fluorescence.
815
816 ***Data Extraction and evaluation***
817 Quantitative data, including total nuclei count, all CD15-positive cells, all HLA-DR-positive cells, all
818 MPO-positive cells and co-expression counts, were exported from CellProfiler for statistical analysis.
819 Data evaluation and statistical analyses were conducted in R (v4.4.2).

820

821 ***Mice***

822 C57BL/6J (B6J, stock no. 000664) mice were obtained from The Jackson laboratory and used for
823 injections of AKPS organoids, or kept uninjected as non-tumour bearing control mice.

824 All mouse experiments were performed under the approval of Cantonal Veterinary Office of Zurich,
825 Switzerland and in accordance with Swiss guidelines. All mice were maintained on a 12h light and 12h
826 dark schedule, and water and chow were provided ad libitum.

827

828 ***Colonoscopy submucosal injection of AKPS organoids***

829 *Apc*-LOF, *p53*-LOF, *Kras*-GOF, *Smad4*-LOF (AKPS) tumor organoids were generated from AKP
830 organoids harbouring a RNP-mediated Smad4 mutation introduced by CRISPR-Cas9[35]. Organoids were
831 cultured in Matrigel and passaged every 2-3 days[139]. Colonoscopy-guided orthotopic cell transplantation
832 was done as described previously with modifications[140]. Briefly, following mechanical dissociation, the
833 equivalent of one Matrigel done was resuspended in 50 mL OptiMEM (Gibco) to equate to one injection
834 in one mouse. 8-to 10-week-old males were anesthetized with isoflurane and placed facing up on a 37°C
835 heating pad. Fecal content was flushed from the colon with warm DPBS and AKPS organoid solution
836 was injected carefully below the mucosa using a custom needle, syringe and colonoscope. Observation
837 of a visible injection bubble was noted as an indicator for successful injection in the correct location.
838 Mice were then regularly monitored according to official guidelines until experimental endpoint.

839

840 ***Tissue dissociation of mouse samples***

841 *AKPS tumors*

842 Tumors were collected at the study endpoint and cut into small pieces. Tissue was digested at 37 °C for
843 50 min in RPMI-1640 containing 15 mM HEPES (H0887 Sigma), 500 U/mL^{-1} type IV (C5138 Sigma)
844 and type VIII collagenase, (C2139 Sigma) and 0.05 mg/mL DNase I (10104159001 Roche). Non-
845 digested tissue was mechanically disrupted with a syringe and cell suspension was pushed through a 70
846 µm cell strainer. Flow-through was topped with DPBS buffer containing 2% BSA, 100 U mL/$^{-1}$
847 penicillin/streptomycin (P0781 Sigma) and 5 mM EDTA and cells were spun prior and resuspended
848 prior analysis.

849 *Blood*

850 Blood was collected from mice post-mortem by cardiac puncture in 2% BSA 5 mM EDTA, followed by
851 red blood cell lysis with distilled cold water for 30 sec.

852 *Colon from naive mice and adjacent colon to AKPS tumor*

853 Colonic tissue was collected, cleaned and opened longitudinally, prior to washing with PBS. Tissue was
854 cut into small pieces and washed twice while shaking for 25 min at 37 °C with HBSS 2% BSA, 100 U
855 mL/$^{-1}$ penicillin/streptomycin (P0781 Sigma) and 5 mM EDTA. Tissue was digested at 37 °C for 50 min
856 in RPMI-1640 containing 15 mM HEPES (H0887 Sigma), 500 U/mL^{-1} type IV (C5138 Sigma) and type
857 VIII collagenase, (C2139 Sigma) and 0.05 mg/mL DNase I (10104159001 Roche). Non-digested tissue
858 was mechanically disrupted with a syringe and cell suspension was pushed through a 70 µm cell strainer.
859 Cells were centrifuged for 8 min and layered onto a 40/80% Percoll (17089101 Cytiva) gradient (18
860 min, 2.100 *g*, 20 °C, no brake). The interphase was collected and washed in DPBS.

861 *Bone marrow*

862 Tibia and femur were collected and flushed with RPMI medium using a 23-gauge needle. The content
863 was filtered through a 40 µm strainer followed by red blood cell lysis via incubation in ice-cold distilled
864 water for 30 sec.

865

866 ***Flow cytometry staining of mouse samples***

867 *Staining*

868 Cells were surface stained in DPBS at 4 °C for 30 minutes with fixable viability dye eFluor 780, and
869 combinations of the following listed antibodies (1:200, all from BioLegend; unless otherwise stated):
870 CD11b BV510 (M1/70, 101263), CD45 BV650 (30-F11,103151), CD80 BV605 (1:100, 16-10A1,
871 104729), Ly6G Percp-Cy5.5 (1A8, 127616), MHCII AF700 (M5/114.15.2,107622), PD-L1 PE-Cy7
872 (1:100, 10F.9G2, 124314), Siglec-F BV421 (E50-2440, 552681 BD Biosciences). Samples were
873 acquired on a LSRII Fortessa (BD Biosciences).

874 *Data analysis*

875 Analysis of flow cytometry data was performed using FlowJo Software. Relative cell frequencies, cell
876 counts, or Mean Fluorescent Intensities (MFIs) were subsequently plotted in graphical formats using
877 GraphPad Prism.

878 *Statistical analysis*

879 Statistical analyses were performed using GraphPad Prism. For statistical comparisons of two groups,
880 two-tailed unpaired t-tests were used. For comparisons of more than two groups, a one-way analysis of
881 variance (ANOVA) was used. Differences were considered statistically significant when $P < 0.05$.

882

883 ***Single-cell RNA sequencing of mouse neutrophils - experimental***

884 *CD45+ cell enrichment using magnetic beads*

885 CD45 positive cells were enriched using anti-CD45 microbeads (130-052-301 Miltenyi Biotec)
886 according to the manufacturer's instructions.

887 *Cell hashing*

888 Samples were pooled for shared scRNAseq runs after labeling with sample tags (633793 BD Mouse
889 Single-Cell Multiplexing Kit). Thereby, 10^6 cells from each sample were incubated in a staining buffer
890 (DPBS, 1% BSA, 1% EDTA) and the respective sample tag for 20 min at room temperature (20°C ±
891 2°C). After incubation, cells were washed twice with staining buffer before counting. From each sample
892 10-20.000 cells were pooled to have a total of around 60-80.000 cells for single cell capture.

893 *Single cell capture, library preparation and sequencing*

894 Single cell capture was done using the BD Rhapsody Express Single-Cell Analysis System (BD
895 Biosciences) following the manufacturer's protocol. In short, cells were resuspended in 650 µL BD
896 Sample Buffer supplemented with Rnase inhibitors (1:1.000 SUPERase 20 U μL^{-1}, AM2694 Thermo
897 Fisher Scientific; NxGen Rnase Inhibitor 40 U μL^{-1}, 30281-2 Lucigen) and loaded on a cartridge. To
898 extract their RNA capture beads (v1 for tumor, disseminated and adjacent colon; enhanced beads for
899 colon, blood and bone marrow) were added, followed by cell lysis. After single cell capture, libraries
900 from the cDNA and the sample tags were prepared using the BD Rhapsody Whole Transcriptome
901 Analysis Amplification Kit (633801 BD Biosciences). Libraries were then quantified using a Qubit
902 Fluorometer (dsDNA HS kit: Q32851 Thermo Fisher Scientific) and a TapeStation 4200 (HS D500
903 types and reagents, 5067-5592 Agilent Technologies). Sequencing was performed on a NovaSeq6000
904 system (Illumina) with SP100 Reagents (100 cycles) and PhiX spike-in (20% for v1, 1% for enhanced
905 beads) with a depth of around 20.000 reads/cell for WTA and 500-1000 reads/cell for the sample tags.

906

907 ***Single-cell RNA sequencing of mouse neutrophils - data analysis***

908 *Pre-processing, normalization, clustering and data integration*

909 Demultiplexing was done using Bcl2fastq (v2.20.0433, Illumina). FASTQ files were then further
910 processed using the BD Rhapsody™ WTA analysis pipeline (v1.12.1) on the Seven Bridge Genomics
911 platform with default settings. Gene mapping was done with a STAR index generated using STAR[43]
912 (v2.5.2b) with the gene code GRCm38 v25. R (v4.3) and Python (v3) were used for the downstream
913 analysis. If not specified, the Seurat[141] (v5.1.0) R package was used. After the first quality assessment,
914 cells with fewer than 200 gene features, more than 5000 gene features and more than 25 % of
915 mitochondrial genes were removed from the dataset. Additionally, the R package decontX[44] (v1.0.0)
916 was applied to adjust the dataset for cell free RNA. This approach uses a Bayesian model to estimate
917 and remove cell free RNA contaminations. Data was log normalized and scaled with the Seurat functions

918 'NormalizeData' and 'ScaleData'. Thereby, the number of features, UMI counts and percentage of
919 mitochondrial genes were regressed out. The top 2.000 variable feature genes were used for principal
920 component analysis (PCA) followed by clustering using the Louvain algorithm with multi level
921 refinements. The colon and tumor samples were processed separate from the blood and bone marrow
922 samples. To account for batch effects in the colon and tumor samples the fast mutual nearest neighbors
923 (FastMNN) algorithm (Seurat function 'FastMNNIntegration') was applied. The blood and bone
924 marrow samples were merged without batch effect correction.

925

926 *Annotation, extraction and integration of neutrophils*

927 Neutrophil clusters were identified using on the one hand neutrophil specific marker genes (*S100a8,*
928 *S100a9*) and an automatic annotation approach called SingleR[45] (v2.4.1) with the murine ImmGen[46]
929 (Immunological Genome Project) as reference which was loaded with the celldex[45] (v1.12.0) R package.
930 Neutrophil clusters were then extracted and integrated from all tissues using the FastMNN approach
931 mentioned before.

932

933 ***DEG analysis***

934 To analyze DEGs between neutrophil clusters the Seurat function 'FindAllMarkers' with min.pct = 0.25
935 and logfc.threshold = 0.25 was used on log normalized data. This method applies a non-parametric
936 Wilcoxon rank sum test and corrects the p-values using the Bonferroni method. Genes of interest were
937 plotted in heatmap using pheatmap (v1.0.12).

938

939

940

941 ***Analysis of human neutrophil markers in mouse data***

942 To analyze the expression of marker genes identified in human neutrophil clusters, the genes were first
943 converted to mouse gene symbols using the orthogene Bioconductor package (v1.8.0) and then plotted
944 in a heatmap using pheatmap (v1.0.12).

945

946 ***Pseudotime trajectory analysis with Slingshot***

947 A pseudotime trajectory was inferred on the mnn.clusters of the combined mouse neutrophils dataset
948 (bone marrow, blood and colon) using the slingshot[47] (v2.10.0) R package. Therefore, the Seurat object
949 was first converted to a Single Cell Expression (SCE) object using sceasy R package (v0.0.7) , which
950 was then used for the slingshot analysis on the UMAP dimensional space. In brief, by minimum
951 spanning tree (MST) construction using the 'getLineages' function, a global lineage structure was

952 identified and pseudotime variables were inferred using the 'getCurves' function. To compare the
953 pseudotime between tissue, the cellular distribution was visualized in a density plot across pseudotime.
954
955 ***Neutrophil Activation Assay***
956 Human neutrophils were isolated from freshly collected peripheral blood from healthy donors.
957 Neutrophils were purified using the MACSxpress Whole Blood Neutrophil Isolation Kit (Miltenyi
958 Biotec) according to the manufacturer's instructions. Neutrophils were washed with DPBS (Sigma-
959 Aldrich) and centrifuged at 115 g for 10 min. Then, 2 million neutrophils were seeded in a 12-well plate
960 in 1 mL of either RPMI 1640 (Sigma-Aldrich) supplemented with 2 mM L-glutamine (Sigma-Aldrich),
961 10% FBS (Sigma-Aldrich) and 10 mL/L Penicillin-Streptomycin (Sigma-Aldrich), or with conditioned
962 medium obtained as supernatant of SW900 or CaCo2 cell lines after 48 hours of growth in RPMI
963 (supplemented with L-glutamine, FBS and Penicillin-Streptomycin) to approximately 70% confluency.
964 Neutrophils were cultured for 24 hours at 37°C and 5% CO_2. Neutrophils were collected, counted, and
965 centrifuged at 450 g for 10 min at room temperature (20°C ± 2°C). Cells were resuspended at a
966 concentration of $2.5x10^5$ cells/mL in Seahorse XF RPMI Medium pH 7.4 supplemented with 1 mM
967 pyruvate, 2 mM glutamine, 10 mM glucose (all from Agilent Technologies). Seahorse XFp Cell Culture
968 Miniplates (Agilent Technologies) were pre-coated with poly-L-lysine (Sigma-Aldrich) and $5x10^4$ cells
969 were plated into each well. Miniplates were centrifuged at 300 g for 1 min with low-brake deceleration.
970 Seahorse XP Neutrophil Activation assay was performed on an XFp instrument with sequential injection
971 of 0.5 µM Rotenone/Antimycin A (Agilent Technologies) and 100 ng/mL phorbol 12-myristate 13-
972 acetate (PMA) (Fisher Scientific) according to the manufacturer's protocol. Data was processed by
973 Seahorse XFe Wave software (Agilent Technologies) and by GraphPad Prism software (v 10.4;
974 GraphPad Software). O2 consumption was calculated from the area under the curve (AUC) of the kinetic
975 range between activation and return to basal rates.

976 **REFERENCES**

977 1. Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., and
978 Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and
979 mortality worldwide for 36 cancers in 185 countries. CA. Cancer J. Clin. *74*, 229–263.
980 https://doi.org/10.3322/caac.21834.
981 2. Akimoto, N., Ugai, T., Zhong, R., Hamada, T., Fujiyoshi, K., Giannakis, M., Wu, K.,
982 Cao, Y., Ng, K., and Ogino, S. (2021). Rising incidence of early-onset colorectal cancer
983 - a call to action. Nat. Rev. Clin. Oncol. *18*, 230–243. https://doi.org/10.1038/s41571-
984 020-00445-1.
985 3. Cremolini, C., Loupakis, F., Antoniotti, C., Lupi, C., Sensi, E., Lonardi, S., Mezi, S.,
986 Tomasello, G., Ronzoni, M., Zaniboni, A., et al. (2015). FOLFOXIRI plus bevacizumab
987 versus FOLFIRI plus bevacizumab as first-line treatment of patients with metastatic
988 colorectal cancer: updated overall survival and molecular subgroup analyses of the open-
989 label, phase 3 TRIBE study. Lancet Oncol. *16*, 1306–1315.
990 https://doi.org/10.1016/S1470-2045(15)00122-9.
991 4. Nicolantonio, F. (2021). Precision oncology in metastatic colorectal cancer - from
992 biology to medicine. Nat. Rev. Clin. Oncol. *18*, 506-525,.
993 https://doi.org/10.1038/s41571-021-00495-z.
994 5. Cancer Genome Atlas Network (2012). Comprehensive molecular characterization of
995 human colon and rectal cancer. Nature *487*, 330–337.
996 https://doi.org/10.1038/nature11252.
997 6. Le, D.T., Durham, J.N., Smith, K.N., Wang, H., Bartlett, B.R., Aulakh, L.K., Lu, S.,
998 Kemberling, H., Wilt, C., Luber, B.S., et al. (2017). Mismatch repair deficiency predicts
999 response of solid tumors to PD-1 blockade. Science *357*, 409–413.
1000 https://doi.org/10.1126/science.aan6733.
1001 7. Galon, J. (2006). Type, density, and location of immune cells within human colorectal
1002 tumors predict clinical outcome. Science *313*, 1960-1964,.
1003 https://doi.org/10.1126/science.1129139.
1004 8. Sahin, I.H., Ciombor, K.K., Diaz, L.A., Yu, J., and Kim, R. (2022). Immunotherapy for
1005 Microsatellite Stable Colorectal Cancers: Challenges and Novel Therapeutic Avenues.
1006 Am. Soc. Clin. Oncol. Educ. Book Am. Soc. Clin. Oncol. Annu. Meet. *42*, 1–12.
1007 https://doi.org/10.1200/EDBK_349811.
1008 9. Ros, J., Balconi, F., Baraibar, I., Saoudi Gonzalez, N., Salva, F., Tabernero, J., and Elez,
1009 E. (2023). Advances in immune checkpoint inhibitor combination strategies for
1010 microsatellite stable colorectal cancer. Front. Oncol. *13*, 1112276.
1011 https://doi.org/10.3389/fonc.2023.1112276.
1012 10. Antoniotti, C., Rossini, D., Pietrantonio, F., Catteau, A., Salvatore, L., Lonardi, S.,
1013 Boquet, I., Tamberi, S., Marmorino, F., Moretto, R., et al. (2022). Upfront FOLFOXIRI
1014 plus bevacizumab with or without atezolizumab in the treatment of patients with
1015 metastatic colorectal cancer (AtezoTRIBE): a multicentre, open-label, randomised,
1016 controlled, phase 2 trial. Lancet Oncol. *23*, 876–887. https://doi.org/10.1016/S1470-
1017 2045(22)00274-1.
1018 11. Hackl, H., Charoentong, P., Finotello, F., and Trajanoski, Z. (2016). Computational
1019 genomics tools for dissecting tumour-immune cell interactions. Nat. Rev. Genet. *17*,
1020 441–458. https://doi.org/10.1038/nrg.2016.67.
1021 12. Finotello, F., Rieder, D., Hackl, H., and Trajanoski, Z. (2019). Next-generation
1022 computational tools for interrogating cancer immunity. Nat. Rev. Genet. *20*, 724–746.
1023 https://doi.org/10.1038/s41576-019-0166-7.
1024 13. Angelova, M., Charoentong, P., Hackl, H., Fischer, M.L., Snajder, R., Krogsdam, A.M.,
1025 Waldner, M.J., Bindea, G., Mlecnik, B., Galon, J., et al. (2015). Characterization of the

1026 immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape
1027 mechanisms and novel targets for immunotherapy. Genome Biol. *16*, 64.
1028 https://doi.org/10.1186/s13059-015-0620-6.
1029 14. Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D.,
1030 Hackl, H., and Trajanoski, Z. (2017). Pan-cancer Immunogenomic Analyses Reveal
1031 Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint
1032 Blockade. Cell Rep. *18*, 248–262. https://doi.org/10.1016/j.celrep.2016.12.019.
1033 15. Salcher, S., Sturm, G., Horvath, L., Untergasser, G., Kuempers, C., Fotakis, G.,
1034 Panizzolo, E., Martowicz, A., Trebo, M., Pall, G., et al. (2022). High-resolution single-
1035 cell atlas reveals diversity and plasticity of tissue-resident neutrophils in non-small cell
1036 lung cancer. Cancer Cell *40*, 1503-1520.e8. https://doi.org/10.1016/j.ccell.2022.10.008.
1037 16. Chu, X., Li, X., Zhang, Y., Dang, G., Miao, Y., Xu, W., Wang, J., Zhang, Z., and Cheng,
1038 S. (2024). Integrative single-cell analysis of human colorectal cancer reveals patient
1039 stratification with distinct immune evasion mechanisms. Nat. Cancer *5*, 1409–1426.
1040 https://doi.org/10.1038/s43018-024-00807-z.
1041 17. Hirschhorn, D., Budhu, S., Kraehenbuehl, L., Gigoux, M., Schröder, D., Chow, A.,
1042 Ricca, J.M., Gasmi, B., De Henau, O., Mangarin, L.M.B., et al. (2023). T cell
1043 immunotherapies engage neutrophils to eliminate tumor antigen escape variants. Cell
1044 *186*, 1432-1447.e17. https://doi.org/10.1016/j.cell.2023.03.007.
1045 18. Väyrynen, J.P., Lau, M.C., Haruki, K., Väyrynen, S.A., Dias Costa, A., Borowsky, J.,
1046 Zhao, M., Fujiyoshi, K., Arima, K., Twombly, T.S., et al. (2020). Prognostic
1047 Significance of Immune Cell Populations Identified by Machine Learning in Colorectal
1048 Cancer Using Routine Hematoxylin and Eosin-Stained Sections. Clin. Cancer Res. Off.
1049 J. Am. Assoc. Cancer Res. *26*, 4326–4338. https://doi.org/10.1158/1078-0432.CCR-20-
1050 0071.
1051 19. Governa, V., Trella, E., Mele, V., Tornillo, L., Amicarella, F., Cremonesi, E., Muraro,
1052 M.G., Xu, H., Droeser, R., Däster, S.R., et al. (2017). The Interplay Between
1053 Neutrophils and CD8+ T Cells Improves Survival in Human Colorectal Cancer. Clin.
1054 Cancer Res. Off. J. Am. Assoc. Cancer Res. *23*, 3847–3858.
1055 https://doi.org/10.1158/1078-0432.CCR-16-2047.
1056 20. Hedrick, C.C., and Malanchi, I. (2022). Neutrophils in cancer: heterogeneous and
1057 multifaceted. Nat. Rev. Immunol. *22*, 173–187. https://doi.org/10.1038/s41577-021-
1058 00571-6.
1059 21. Gungabeesoon, J., Gort-Freitas, N.A., Kiss, M., Bolli, E., Messemaker, M., Siwicki, M.,
1060 Hicham, M., Bill, R., Koch, P., Cianciaruso, C., et al. (2023). A neutrophil response
1061 linked to tumor control in immunotherapy. Cell *186*, 1448-1464.e20.
1062 https://doi.org/10.1016/j.cell.2023.02.032.
1063 22. Schürch, C.M., Bhate, S.S., Barlow, G.L., Phillips, D.J., Noti, L., Zlobec, I., Chu, P.,
1064 Black, S., Demeter, J., McIlwain, D.R., et al. (2020). Coordinated Cellular
1065 Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive
1066 Front. Cell *183*, 838. https://doi.org/10.1016/j.cell.2020.10.021.
1067 23. Salcher, S., Heidegger, I., Untergasser, G., Fotakis, G., Scheiber, A., Martowicz, A.,
1068 Noureen, A., Krogsdam, A., Schatz, C., Schäfer, G., et al. (2024). Comparative analysis
1069 of 10X Chromium vs. BD Rhapsody whole transcriptome single-cell sequencing
1070 technologies in complex human tissues. Heliyon *10*, e28358.
1071 https://doi.org/10.1016/j.heliyon.2024.e28358.
1072 24. Scheiber, A., Trebo, M., Pittl, A., Heidegger, I., Hautz, T., Oberhuber, R., Trajanoski,
1073 Z., Augustin, F., Sopper, S., Wolf, D., et al. (2024). Profiling low-mRNA content cells
1074 in complex human tissues using BD Rhapsody single-cell analysis. STAR Protoc. *5*,
1075 103475. https://doi.org/10.1016/j.xpro.2024.103475.

25. Schubert, M. (2018). Perturbation-response genes reveal signaling footprints in cancer gene expression. Nat Commun *9, 20*. https://doi.org/10.1038/s41467-017-02391-6.
26. Choi, H.Y., Jo, H., Zhao, X., Hoadley, K.A., Newman, S., Holt, J., Hayward, M.C., Love, M.I., Marron, J.S., and Hayes, D.N. (2021). SCISSOR: a framework for identifying structural changes in RNA transcripts. Nat. Commun. *12*, 286. https://doi.org/10.1038/s41467-020-20593-3.
27. Horvath, L., Puschmann, C., Scheiber, A., Martowicz, A., Sturm, G., Trajanoski, Z., Wolf, D., Pircher, A., and Salcher, S. (2024). Beyond binary: bridging neutrophil diversity to new therapeutic approaches in NSCLC. Trends Cancer *10*, 457–474. https://doi.org/10.1016/j.trecan.2024.01.010.
28. Park, M.D., Le Berichel, J., Hamon, P., Wilk, C.M., Belabed, M., Yatim, N., Saffon, A., Boumelha, J., Falcomatà, C., Tepper, A., et al. (2024). Hematopoietic aging promotes cancer by fueling IL-1α-driven emergency myelopoiesis. Science *386*, eadn0327. https://doi.org/10.1126/science.adn0327.
29. Singhal, S., Bhojnagarwala, P.S., O'Brien, S., Moon, E.K., Garfall, A.L., Rao, A.S., Quatromoni, J.G., Stephen, T.L., Litzky, L., Deshpande, C., et al. (2016). Origin and Role of a Subset of Tumor-Associated Neutrophils with Antigen-Presenting Cell Features in Early-Stage Human Lung Cancer. Cancer Cell *30*, 120–135. https://doi.org/10.1016/j.ccell.2016.06.001.
30. Matusiak, M., Hickey, J.W., van IJzendoorn, D.G.P., Lu, G., Kidziński, L., Zhu, S., Colburg, D.R.C., Luca, B., Phillips, D.J., Brubaker, S.W., et al. (2024). Spatially Segregated Macrophage Populations Predict Distinct Outcomes in Colon Cancer. Cancer Discov. *14*, 1418–1439. https://doi.org/10.1158/2159-8290.CD-23-1300.
31. Birk, S., Bonafonte-Pardàs, I., Feriz, A.M., Boxall, A., Agirre, E., Memi, F., Maguza, A., Fan, R., Castelo-Branco, G., Bayraktar, O.A., et al. (2024). Large-scale characterization of cell niches in spatial atlases using bio-inspired graph learning. Preprint at bioRxiv, https://doi.org/10.1101/2024.02.21.581428 https://doi.org/10.1101/2024.02.21.581428.
32. Ouyang, Y., Zhong, W., Xu, P., Wang, B., Zhang, L., Yang, M., Chen, J., Li, H., Li, S., Chen, X., et al. (2024). Tumor-associated neutrophils suppress CD8+ T cell immunity in urothelial bladder carcinoma through the COX-2/PGE2/IDO1 Axis. Br. J. Cancer *130*, 880–891. https://doi.org/10.1038/s41416-023-02552-z.
33. Metzemaekers, M., Gouwy, M., and Proost, P. (2020). Neutrophil chemoattractant receptors in health and disease: double-edged swords. Cell. Mol. Immunol. *17*, 433–450. https://doi.org/10.1038/s41423-020-0412-0.
34. Plattner, C., Lamberti, G., Blattmann, P., Kirchmair, A., Rieder, D., Loncova, Z., Sturm, G., Scheidl, S., Ijsselsteijn, M., Fotakis, G., et al. (2023). Functional and spatial proteomics profiling reveals intra- and intercellular signaling crosstalk in colorectal cancer. iScience *26*, 108399. https://doi.org/10.1016/j.isci.2023.108399.
35. Dalessi, T., Valenta, T., Moro, G., Brügger, M.D., Meijs, A., Szczerba, B.M., Brunner, T., Lampart, S., Nater, M., Platt, R.J., et al. (2024). An oncoembryology approach uncovers SoxC-driven regulation of colon development and cancer. Preprint at bioRxiv, https://doi.org/10.1101/2024.10.08.617191 https://doi.org/10.1101/2024.10.08.617191.
36. Locard-Paulet, M., Palasca, O., and Jensen, L.J. (2022). Identifying the genes impacted by cell proliferation in proteomics and transcriptomics studies. PLoS Comput. Biol. *18*, e1010604. https://doi.org/10.1371/journal.pcbi.1010604.
37. Li, T., Li, X., Zamani, A., Wang, W., Lee, C.-N., Li, M., Luo, G., Eiler, E., Sun, H., Ghosh, S., et al. (2020). c-Rel Is a Myeloid Checkpoint for Cancer Immunotherapy. Nat. Cancer *1*, 507–517. https://doi.org/10.1038/s43018-020-0061-3.
38. Fridman, W.H., Meylan, M., Petitprez, F., Sun, C.-M., Italiano, A., and Sautès-Fridman,

C. (2022). B cells and tertiary lymphoid structures as determinants of tumour immune contexture and clinical outcome. Nat. Rev. Clin. Oncol. *19*, 441–457. https://doi.org/10.1038/s41571-022-00619-z.

39. Moffat, A., and Gwyer Findlay, E. (2024). Evidence for antigen presentation by human neutrophils. Blood *143*, 2455–2463. https://doi.org/10.1182/blood.2023023444.
40. Wu, Y., Ma, J., Yang, X., Nan, F., Zhang, T., Ji, S., Rao, D., Feng, H., Gao, K., Gu, X., et al. (2024). Neutrophil profiling illuminates anti-tumor antigen-presenting potency. Cell *187*, 1422-1439.e24. https://doi.org/10.1016/j.cell.2024.02.005.
41. Poplimont, H., Georgantzoglou, A., Boulch, M., Walker, H.A., Coombs, C., Papaleonidopoulou, F., and Sarris, M. (2020). Neutrophil Swarming in Damaged Tissue Is Orchestrated by Connexins and Cooperative Calcium Alarm Signals. Curr. Biol. CB *30*, 2761-2776.e7. https://doi.org/10.1016/j.cub.2020.05.030.
42. LaMarche, N.M., Hegde, S., Park, M.D., Maier, B.B., Troncoso, L., Le Berichel, J., Hamon, P., Belabed, M., Mattiuz, R., Hennequin, C., et al. (2024). An IL-4 signalling axis in bone marrow drives pro-tumorigenic myelopoiesis. Nature *625*, 166–174. https://doi.org/10.1038/s41586-023-06797-9.
43. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinforma. Oxf. Engl. *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
44. Yang, S., Corbett, S.E., Koga, Y., Wang, Z., Johnson, W.E., Yajima, M., and Campbell, J.D. (2020). Decontamination of ambient RNA in single-cell RNA-seq with DecontX. Genome Biol. *21*, 57. https://doi.org/10.1186/s13059-020-1950-6.
45. Aran, D., Looney, A.P., Liu, L., Wu, E., Fong, V., Hsu, A., Chak, S., Naikawadi, R.P., Wolters, P.J., Abate, A.R., et al. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat. Immunol. *20*, 163–172. https://doi.org/10.1038/s41590-018-0276-y.
46. Heng, T.S.P., Painter, M.W., and Immunological Genome Project Consortium (2008). The Immunological Genome Project: networks of gene expression in immune cells. Nat. Immunol. *9*, 1091–1094. https://doi.org/10.1038/ni1008-1091.
47. Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics *19*, 477. https://doi.org/10.1186/s12864-018-4772-0.
48. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.
49. Bankhead, P., Loughrey, M.B., Fernández, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., et al. (2017). QuPath: Open source software for digital pathology image analysis. Sci. Rep. *7*, 16878. https://doi.org/10.1038/s41598-017-17204-5.
50. Stirling, D.R., Swain-Bowden, M.J., Lucas, A.M., Carpenter, A.E., Cimini, B.A., and Goodman, A. (2021). CellProfiler 4: improvements in speed, utility and usability. BMC Bioinformatics *22*, 433. https://doi.org/10.1186/s12859-021-04344-9.
51. Nemati, N., Boeck, N., Lamberti, G., Lisandrelli, R., and Trajanoski, Z. (2024). Protocol for functional profiling of patient-derived organoids for precision oncology. STAR Protoc. *5*, 102887. https://doi.org/10.1016/j.xpro.2024.102887.
52. Wu, Y., Ma, J., Yang, X., Nan, F., Zhang, T., Ji, S., Rao, D., Feng, H., Gao, K., Gu, X., et al. (2024). Neutrophil profiling illuminates anti-tumor antigen-presenting potency. Cell *187*, 1422-1439.e24. https://doi.org/10.1016/j.cell.2024.02.005.
53. Bian, S., Hou, Y., Zhou, X., Li, X., Yong, J., Wang, Y., Wang, W., Yan, J., Hu, B., Guo,

H., et al. (2018). Single-cell multiomics sequencing and analyses of human colorectal cancer. Science *362*, 1060–1063. https://doi.org/10.1126/science.aao3791.

54. Borràs, D.M., Verbandt, S., Ausserhofer, M., Sturm, G., Lim, J., Verge, G.A., Vanmeerbeek, I., Laureano, R.S., Govaerts, J., Sprooten, J., et al. (2023). Single cell dynamics of tumor specificity vs bystander activity in CD8+ T cells define the diverse immune landscapes in colorectal cancer. Cell Discov. *9*, 114. https://doi.org/10.1038/s41421-023-00605-4.
55. Burclaff, J., Bliton, R.J., Breau, K.A., Ok, M.T., Gomez-Martinez, I., Ranek, J.S., Bhatt, A.P., Purvis, J.E., Woosley, J.T., and Magness, S.T. (2022). A Proximal-to-Distal Survey of Healthy Adult Human Small Intestine and Colon Epithelium by Single-Cell Transcriptomics. Cell. Mol. Gastroenterol. Hepatol. *13*, 1554–1589. https://doi.org/10.1016/j.jcmgh.2022.02.007.
56. Che, L.-H., Liu, J.-W., Huo, J.-P., Luo, R., Xu, R.-M., He, C., Li, Y.-Q., Zhou, A.-J., Huang, P., Chen, Y.-Y., et al. (2021). A single-cell atlas of liver metastases of colorectal cancer reveals reprogramming of the tumor microenvironment in response to preoperative chemotherapy. Cell Discov. *7*, 80. https://doi.org/10.1038/s41421-021-00312-y.
57. Chen, Y., Wang, D., Li, Y., Qi, L., Si, W., Bo, Y., Chen, X., Ye, Z., Fan, H., Liu, B., et al. (2024). Spatiotemporal single-cell analysis decodes cellular dynamics underlying different responses to immunotherapy in colorectal cancer. Cancer Cell *42*, 1268-1285.e7. https://doi.org/10.1016/j.ccell.2024.06.009.
58. Domínguez Conde, C., Xu, C., Jarvis, L.B., Rainbow, D.B., Wells, S.B., Gomes, T., Howlett, S.K., Suchanek, O., Polanski, K., King, H.W., et al. (2022). Cross-tissue immune cell analysis reveals tissue-specific features in humans. Science *376*, eabl5197. https://doi.org/10.1126/science.abl5197.
59. Elmentaite, R., Kumasaka, N., Roberts, K., Fleming, A., Dann, E., King, H.W., Kleshchevnikov, V., Dabrowska, M., Pritchard, S., Bolt, L., et al. (2021). Cells of the human intestinal tract mapped across space and time. Nature *597*, 250–255. https://doi.org/10.1038/s41586-021-03852-1.
60. Garrido-Trigo, A., Corraliza, A.M., Veny, M., Dotti, I., Melón-Ardanaz, E., Rill, A., Crowell, H.L., Corbí, Á., Gudiño, V., Esteller, M., et al. (2023). Macrophage and neutrophil heterogeneity at single-cell spatial resolution in human inflammatory bowel disease. Nat. Commun. *14*, 4506. https://doi.org/10.1038/s41467-023-40156-6.
61. Giguelay, A., Turtoi, E., Khelaf, L., Tosato, G., Dadi, I., Chastel, T., Poul, M.-A., Pratlong, M., Nicolescu, S., Severac, D., et al. (2022). The landscape of cancer-associated fibroblasts in colorectal cancer liver metastases. Theranostics *12*, 7624–7639. https://doi.org/10.7150/thno.72853.
62. Guo, W., Zhang, C., Wang, X., Dou, D., Chen, D., and Li, J. (2022). Resolving the difference between left-sided and right-sided colorectal cancer by single-cell sequencing. JCI Insight *7*, e152616. https://doi.org/10.1172/jci.insight.152616.
63. Harmon, C., Zaborowski, A., Moore, H., St Louis, P., Slattery, K., Duquette, D., Scanlan, J., Kane, H., Kunkemoeller, B., McIntyre, C.L., et al. (2023). γδ T cell dichotomy with opposing cytotoxic and wound healing functions in human solid tumors. Nat. Cancer *4*, 1122–1137. https://doi.org/10.1038/s43018-023-00589-w.
64. He, S., Wang, L.-H., Liu, Y., Li, Y.-Q., Chen, H.-T., Xu, J.-H., Peng, W., Lin, G.-W., Wei, P.-P., Li, B., et al. (2020). Single-cell transcriptome profiling of an adult human cell atlas of 15 major organs. Genome Biol. *21*, 294. https://doi.org/10.1186/s13059-020-02210-0.
65. Huang, C., Wang, X., Wang, Y., Feng, Y., Wang, X., Chen, S., Yan, P., Liao, J., Zhang, Q., Mao, C., et al. (2024). Sirpα on tumor-associated myeloid cells restrains antitumor

1226 immunity in colorectal cancer independent of its interaction with CD47. Nat. Cancer *5*,
1227 500–516. https://doi.org/10.1038/s43018-023-00691-z.
1228 66. James, K.R., Gomes, T., Elmentaite, R., Kumar, N., Gulliver, E.L., King, H.W., Stares,
1229 M.D., Bareham, B.R., Ferdinand, J.R., Petrova, V.N., et al. (2020). Distinct microbial
1230 and immune niches of the human colon. Nat. Immunol. *21*, 343–353.
1231 https://doi.org/10.1038/s41590-020-0602-z.
1232 67. Ji, T., Fu, H., Wang, L., Chen, J., Tian, S., Wang, G., Wang, L., and Wang, Z. (2024).
1233 Single-cell RNA profiling reveals classification and characteristics of mononuclear
1234 phagocytes in colorectal cancer. PLoS Genet. *20*, e1011176.
1235 https://doi.org/10.1371/journal.pgen.1011176.
1236 68. Ji, L., Fu, G., Huang, M., Kao, X., Zhu, J., Dai, Z., Chen, Y., Li, H., Zhou, J., Chu, X., et
1237 al. (2024). scRNA-seq of colorectal cancer shows regional immune atlas with the
1238 function of CD20+ B cells. Cancer Lett. *584*, 216664.
1239 https://doi.org/10.1016/j.canlet.2024.216664.
1240 69. Joanito, I., Wirapati, P., Zhao, N., Nawaz, Z., Yeo, G., Lee, F., Eng, C.L.P., Macalinao,
1241 D.C., Kahraman, M., Srinivasan, H., et al. (2022). Single-cell and bulk transcriptome
1242 sequencing identifies two epithelial tumor cell states and refines the consensus
1243 molecular classification of colorectal cancer. Nat. Genet. *54*, 963–975.
1244 https://doi.org/10.1038/s41588-022-01100-4.
1245 70. Khaliq, A.M., Erdogan, C., Kurt, Z., Turgut, S.S., Grunvald, M.W., Rand, T., Khare, S.,
1246 Borgia, J.A., Hayden, D.M., Pappas, S.G., et al. (2022). Refining colorectal cancer
1247 classification and clinical stratification through a single-cell atlas. Genome Biol. *23*, 113.
1248 https://doi.org/10.1186/s13059-022-02677-z.
1249 71. Kong, L., Pokatayev, V., Lefkovith, A., Carter, G.T., Creasey, E.A., Krishna, C.,
1250 Subramanian, S., Kochar, B., Ashenberg, O., Lau, H., et al. (2023). The landscape of
1251 immune dysregulation in Crohn's disease revealed through single-cell transcriptomic
1252 profiling in the ileum and colon. Immunity *56*, 444-458.e5.
1253 https://doi.org/10.1016/j.immuni.2023.01.002.
1254 72. Lee, H.-O., Hong, Y., Etlioglu, H.E., Cho, Y.B., Pomella, V., Van den Bosch, B.,
1255 Vanhecke, J., Verbandt, S., Hong, H., Min, J.-W., et al. (2020). Lineage-dependent gene
1256 expression programs influence the immune landscape of colorectal cancer. Nat. Genet.
1257 *52*, 594–603. https://doi.org/10.1038/s41588-020-0636-z.
1258 73. Li, H., Courtois, E.T., Sengupta, D., Tan, Y., Chen, K.H., Goh, J.J.L., Kong, S.L., Chua,
1259 C., Hon, L.K., Tan, W.S., et al. (2017). Reference component analysis of single-cell
1260 transcriptomes elucidates cellular heterogeneity in human colorectal tumors. Nat. Genet.
1261 *49*, 708–718. https://doi.org/10.1038/ng.3818.
1262 74. Li, J., Wu, C., Hu, H., Qin, G., Wu, X., Bai, F., Zhang, J., Cai, Y., Huang, Y., Wang, C.,
1263 et al. (2023). Remodeling of the immune and stromal cell compartment by PD-1
1264 blockade in mismatch repair-deficient colorectal cancer. Cancer Cell *41*, 1152-1169.e7.
1265 https://doi.org/10.1016/j.ccell.2023.04.011.
1266 75. Liu, X., Wang, X., Yang, Q., Luo, L., Liu, Z., Ren, X., Lei, K., Li, S., Xie, Z., Zheng,
1267 G., et al. (2024). Th17 Cells Secrete TWEAK to Trigger Epithelial-Mesenchymal
1268 Transition and Promote Colorectal Cancer Liver Metastasis. Cancer Res. *84*, 1352–1371.
1269 https://doi.org/10.1158/0008-5472.CAN-23-2123.
1270 76. Mazzurana, L., Czarnewski, P., Jonsson, V., Wigge, L., Ringnér, M., Williams, T.C.,
1271 Ravindran, A., Björklund, Å.K., Säfholm, J., Nilsson, G., et al. (2021). Tissue-specific
1272 transcriptional imprinting and heterogeneity in human innate lymphoid cells revealed by
1273 full-length single-cell RNA-sequencing. Cell Res. *31*, 554–568.
1274 https://doi.org/10.1038/s41422-020-00445-x.
1275 77. Parikh, K., Antanaviciute, A., Fawkner-Corbett, D., Jagielowicz, M., Aulicino, A.,

Lagerholm, C., Davis, S., Kinchen, J., Chen, H.H., Alham, N.K., et al. (2019). Colonic epithelial cell diversity in health and inflammatory bowel disease. Nature *567*, 49–55. https://doi.org/10.1038/s41586-019-0992-y.

78. Pelka, K., Hofree, M., Chen, J.H., Sarkizova, S., Pirl, J.D., Jorgji, V., Bejnood, A., Dionne, D., Ge, W.H., Xu, K.H., et al. (2021). Spatially organized multicellular immune hubs in human colorectal cancer. Cell *184*, 4734-4752.e20. https://doi.org/10.1016/j.cell.2021.08.003.
79. Qi, J., Sun, H., Zhang, Y., Wang, Z., Xun, Z., Li, Z., Ding, X., Bao, R., Hong, L., Jia, W., et al. (2022). Single-cell and spatial analysis reveal interaction of FAP+ fibroblasts and SPP1+ macrophages in colorectal cancer. Nat. Commun. *13*, 1742. https://doi.org/10.1038/s41467-022-29366-6.
80. Qian, J., Olbrecht, S., Boeckx, B., Vos, H., Laoui, D., Etlioglu, E., Wauters, E., Pomella, V., Verbandt, S., Busschaert, P., et al. (2020). A pan-cancer blueprint of the heterogeneous tumor microenvironment revealed by single-cell profiling. Cell Res. *30*, 745–762. https://doi.org/10.1038/s41422-020-0355-0.
81. Qin, P., Chen, H., Wang, Y., Huang, L., Huang, K., Xiao, G., Han, C., Hu, J., Lin, D., Wan, X., et al. (2023). Cancer-associated fibroblasts undergoing neoadjuvant chemotherapy suppress rectal cancer revealed by single-cell and spatial transcriptomics. Cell Rep. Med. *4*, 101231. https://doi.org/10.1016/j.xcrm.2023.101231.
82. Sathe, A., Mason, K., Grimes, S.M., Zhou, Z., Lau, B.T., Bai, X., Su, A., Tan, X., Lee, H., Suarez, C.J., et al. (2023). Colorectal Cancer Metastases in the Liver Establish Immunosuppressive Spatial Networking between Tumor-Associated SPP1+ Macrophages and Fibroblasts. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. *29*, 244–260. https://doi.org/10.1158/1078-0432.CCR-22-2041.
83. Scheid, J.F., Eraslan, B., Hudak, A., Brown, E.M., Sergio, D., Delorey, T.M., Phillips, D., Lefkovith, A., Jess, A.T., Duck, L.W., et al. (2023). Remodeling of colon plasma cell repertoire within ulcerative colitis patients. J. Exp. Med. *220*, e20220538. https://doi.org/10.1084/jem.20220538.
84. Smillie, C.S., Biton, M., Ordovas-Montanes, J., Sullivan, K.M., Burgin, G., Graham, D.B., Herbst, R.H., Rogel, N., Slyper, M., Waldman, J., et al. (2019). Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. Cell *178*, 714-730.e22. https://doi.org/10.1016/j.cell.2019.06.029.
85. Terekhanova, N.V., Karpova, A., Liang, W.-W., Strzalkowski, A., Chen, S., Li, Y., Southard-Smith, A.N., Iglesia, M.D., Wendl, M.C., Jayasinghe, R.G., et al. (2023). Epigenetic regulation during cancer transitions across 11 tumour types. Nature *623*, 432–441. https://doi.org/10.1038/s41586-023-06682-5.
86. Thomas, M.F., Slowikowski, K., Manakongtreecheep, K., Sen, P., Samanta, N., Tantivit, J., Nasrallah, M., Zubiri, L., Smith, N.P., Tirard, A., et al. (2024). Single-cell transcriptomic analyses reveal distinct immune cell contributions to epithelial barrier dysfunction in checkpoint inhibitor colitis. Nat. Med. *30*, 1349–1362. https://doi.org/10.1038/s41591-024-02895-x.
87. Tian, J., Chen, J.H., Chao, S.X., Pelka, K., Giannakis, M., Hess, J., Burke, K., Jorgji, V., Sindurakar, P., Braverman, J., et al. (2023). Combined PD-1, BRAF and MEK inhibition in BRAFV600E colorectal cancer: a phase 2 trial. Nat. Med. *29*, 458–466. https://doi.org/10.1038/s41591-022-02181-8.
88. Uhlitz, F. (2021). Mitogen-activated protein kinase activity drives cell trajectories in colorectal cancer. EMBO Mol Med *13, e14123*. https://doi.org/10.15252/emmm.202114123.
89. Chen, B., Scurrah, C.R., McKinley, E.T., Simmons, A.J., Ramirez-Solano, M.A., Zhu, X., Markham, N.O., Heiser, C.N., Vega, P.N., Rolong, A., et al. (2021). Differential pre-

1326 malignant programs and microenvironment chart distinct paths to malignancy in human
1327 colorectal polyps. Cell *184*, 6262-6280.e26. https://doi.org/10.1016/j.cell.2021.11.031.
1328 90. Wang, Y., Song, W., Wang, J., Wang, T., Xiong, X., Qi, Z., Fu, W., Yang, X., and Chen,
1329 Y.-G. (2020). Single-cell transcriptome analysis reveals differential nutrient absorption
1330 functions in human intestine. J. Exp. Med. *217*, e20191130.
1331 https://doi.org/10.1084/jem.20191130.
1332 91. Wang, H., Gong, P., Chen, T., Gao, S., Wu, Z., Wang, X., Li, J., Marjani, S.L., Costa, J.,
1333 Weissman, S.M., et al. (2021). Colorectal Cancer Stem Cell States Uncovered by
1334 Simultaneous Single-Cell Analysis of Transcriptome and Telomeres. Adv. Sci. Weinh.
1335 Baden-Wurtt. Ger. *8*, 2004320. https://doi.org/10.1002/advs.202004320.
1336 92. Wang, F., Long, J., Li, L., Wu, Z.-X., Da, T.-T., Wang, X.-Q., Huang, C., Jiang, Y.-H.,
1337 Yao, X.-Q., Ma, H.-Q., et al. (2023). Single-cell and spatial transcriptome analysis
1338 reveals the cellular heterogeneity of liver metastatic colorectal cancer. Sci. Adv. *9*,
1339 eadf5464. https://doi.org/10.1126/sciadv.adf5464.
1340 93. Wu, Y., Yang, S., Ma, J., Chen, Z., Song, G., Rao, D., Cheng, Y., Huang, S., Liu, Y.,
1341 Jiang, S., et al. (2022). Spatiotemporal Immune Landscape of Colorectal Cancer Liver
1342 Metastasis at Single-Cell Level. Cancer Discov. *12*, 134–153.
1343 https://doi.org/10.1158/2159-8290.CD-21-0316.
1344 94. Yang, M., Yang, C., Ma, D., Li, Z., Zhao, W., and Yang, D. (2023). Single-cell analysis
1345 reveals cellular reprogramming in advanced colon cancer following FOLFOX-
1346 bevacizumab treatment. Front. Oncol. *13*, 1219642.
1347 https://doi.org/10.3389/fonc.2023.1219642.
1348 95. Zhang, L., Yu, X., Zheng, L., Zhang, Y., Li, Y., Fang, Q., Gao, R., Kang, B., Zhang, Q.,
1349 Huang, J.Y., et al. (2018). Lineage tracking reveals dynamic relationships of T cells in
1350 colorectal cancer. Nature *564*, 268–272. https://doi.org/10.1038/s41586-018-0694-x.
1351 96. Zhang, L., Li, Z., Skrzypczynska, K.M., Fang, Q., Zhang, W., O'Brien, S.A., He, Y.,
1352 Wang, L., Zhang, Q., Kim, A., et al. (2020). Single-Cell Analyses Inform Mechanisms
1353 of Myeloid-Targeted Therapies in Colon Cancer. Cell *181*, 442-459.e29.
1354 https://doi.org/10.1016/j.cell.2020.03.048.
1355 97. Zheng, X., Song, J., Yu, C., Zhou, Z., Liu, X., Yu, J., Xu, G., Yang, J., He, X., Bai, X.,
1356 et al. (2022). Single-cell transcriptomic profiling unravels the adenoma-initiation role of
1357 protein tyrosine kinases during colorectal tumorigenesis. Signal Transduct. Target. Ther.
1358 *7*, 60. https://doi.org/10.1038/s41392-022-00881-8.
1359 98. de Vries, N.L., van de Haar, J., Veninga, V., Chalabi, M., Ijsselsteijn, M.E., van der
1360 Ploeg, M., van den Bulk, J., Ruano, D., van den Berg, J.G., Haanen, J.B., et al. (2023).
1361 γδ T cells are effectors of immunotherapy in cancers with HLA class I defects. Nature
1362 *613*, 743–750. https://doi.org/10.1038/s41586-022-05593-1.
1363 99. Picelli, S., Björklund, Å.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R.
1364 (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat.
1365 Methods *10*, 1096–1098. https://doi.org/10.1038/nmeth.2639.
1366 100. Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L.,
1367 Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell
1368 transcriptomics applied to embryonic stem cells. Cell *161*, 1187–1201.
1369 https://doi.org/10.1016/j.cell.2015.04.044.
1370 101. Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I.,
1371 Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-
1372 wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–
1373 1214. https://doi.org/10.1016/j.cell.2015.05.002.
1374 102. Goveia, J., Rohlenova, K., Taverna, F., Treps, L., Conradi, L.-C., Pircher, A., Geldhof,
1375 V., de Rooij, L.P.M.H., Kalucka, J., Sokol, L., et al. (2020). An Integrated Gene

Expression Landscape Profiling Approach to Identify Lung Tumor Endothelial Cell Heterogeneity and Angiogenic Candidates. Cancer Cell *37*, 21-36.e13. https://doi.org/10.1016/j.ccell.2019.12.001.

103. Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., Kang, B., Liu, Z., Jin, L., Xing, R., et al. (2018). Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. Nat. Med. *24*, 978–985. https://doi.org/10.1038/s41591-018-0045-3.
104. Maier, B., Leader, A.M., Chen, S.T., Tung, N., Chang, C., LeBerichel, J., Chudnovskiy, A., Maskey, S., Walker, L., Finnigan, J.P., et al. (2020). A conserved dendritic-cell regulatory program limits antitumour immunity. Nature *580*, 257–262. https://doi.org/10.1038/s41586-020-2134-y.
105. Madissoon, E., Wilbrey-Clark, A., Miragaia, R.J., Saeb-Parsy, K., Mahbubani, K.T., Georgakopoulos, N., Harding, P., Polanski, K., Huang, N., Nowicki-Osuch, K., et al. (2019). scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. Genome Biol. *21*, 1. https://doi.org/10.1186/s13059-019-1906-x.
106. Adams, T.S., Schupp, J.C., Poli, S., Ayaub, E.A., Neumark, N., Ahangari, F., Chu, S.G., Raby, B.A., DeIuliis, G., Januszyk, M., et al. (2020). Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. Sci. Adv. *6*, eaba1983. https://doi.org/10.1126/sciadv.aba1983.
107. Reyfman, P.A., Walter, J.M., Joshi, N., Anekalla, K.R., McQuattie-Pimentel, A.C., Chiu, S., Fernandez, R., Akbarpour, M., Chen, C.-I., Ren, Z., et al. (2019). Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. Am. J. Respir. Crit. Care Med. *199*, 1517–1536. https://doi.org/10.1164/rccm.201712-2410OC.
108. Habermann, A.C., Gutierrez, A.J., Bui, L.T., Yahn, S.L., Winters, N.I., Calvi, C.L., Peter, L., Chung, M.-I., Taylor, C.J., Jetter, C., et al. (2020). Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. Sci. Adv. *6*, eaba1972. https://doi.org/10.1126/sciadv.aba1972.
109. Vieira Braga, F.A., Kar, G., Berg, M., Carpaij, O.A., Polanski, K., Simon, L.M., Brouwer, S., Gomes, T., Hesse, L., Jiang, J., et al. (2019). A cellular census of human lungs identifies novel cell states in health and in asthma. Nat. Med. *25*, 1153–1163. https://doi.org/10.1038/s41591-019-0468-5.
110. Mayr, C.H., Simon, L.M., Leuschner, G., Ansari, M., Schniering, J., Geyer, P.E., Angelidis, I., Strunz, M., Singh, P., Kneidinger, N., et al. (2021). Integrative analysis of cell state changes in lung fibrosis with peripheral protein biomarkers. EMBO Mol. Med. *13*, e12871. https://doi.org/10.15252/emmm.202012871.
111. Xu, C., Lopez, R., Mehlman, E., Regier, J., Jordan, M.I., and Yosef, N. (2021). Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. Mol. Syst. Biol. *17*, e9620. https://doi.org/10.15252/msb.20209620.
112. Gayoso, A., Lopez, R., Xing, G., Boyeau, P., Valiollah Pour Amiri, V., Hong, J., Wu, K., Jayasuriya, M., Mehlman, E., Langevin, M., et al. (2022). A Python library for probabilistic analysis of single-cell omics data. Nat. Biotechnol. *40*, 163–166. https://doi.org/10.1038/s41587-021-01206-w.
113. Luecken, M.D., Büttner, M., Chaichoompu, K., Danese, A., Interlandi, M., Mueller, M.F., Strobl, D.C., Zappia, L., Dugas, M., Colomé-Tatché, M., et al. (2022). Benchmarking atlas-level data integration in single-cell genomics. Nat. Methods *19*, 41–50. https://doi.org/10.1038/s41592-021-01336-8.
114. Lopez, R., Regier, J., Cole, M.B., Jordan, M.I., and Yosef, N. (2018). Deep generative modeling for single-cell transcriptomics. Nat. Methods *15*, 1053–1058. https://doi.org/10.1038/s41592-018-0229-2.

115. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15. https://doi.org/10.1186/s13059-017-1382-0.
116. Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. https://doi.org/10.1038/nbt.4314.
117. Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. *9*, 5233. https://doi.org/10.1038/s41598-019-41695-z.
118. Schupp, J.C., Adams, T.S., Cosme, C., Raredon, M.S.B., Yuan, Y., Omote, N., Poli, S., Chioccioli, M., Rose, K.-A., Manning, E.P., et al. (2021). Integrated Single-Cell Atlas of Endothelial Cells of the Human Lung. Circulation *144*, 286–302. https://doi.org/10.1161/CIRCULATIONAHA.120.052318.
119. Sikkema, L., Strobl, D., Zappia, L., Madissoon, E., Markov, N.S., Zaragosi, L., Ansari, M., Arguel, M., Apperloo, L., Bécavin, C., et al. (2022). An integrated cell atlas of the human lung in health and disease. Preprint at bioRxiv, https://doi.org/10.1101/2022.03.10.483747 https://doi.org/10.1101/2022.03.10.483747.
120. Büttner, M., Ostner, J., Müller, C.L., Theis, F.J., and Schubert, B. (2021). scCODA is a Bayesian model for compositional single-cell data analysis. Nat. Commun. *12*, 6876. https://doi.org/10.1038/s41467-021-27150-6.
121. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.
122. Squair, J.W., Gautier, M., Kathe, C., Anderson, M.A., James, N.D., Hutson, T.H., Hudelle, R., Qaiser, T., Matson, K.J.E., Barraud, Q., et al. (2021). Confronting false discoveries in single-cell differential expression. Nat. Commun. *12*, 5692. https://doi.org/10.1038/s41467-021-25960-2.
123. Ignatiadis, N., Klaus, B., Zaugg, J.B., and Huber, W. (2016). Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. Nat. Methods *13*, 577–580. https://doi.org/10.1038/nmeth.3885.
124. Holland, C.H., Tanevski, J., Perales-Patón, J., Gleixner, J., Kumar, M.P., Mereu, E., Joughin, B.A., Stegle, O., Lauffenburger, D.A., Heyn, H., et al. (2020). Robustness and applicability of transcription factor and pathway analysis tools on single-cell RNA-seq data. Genome Biol. *21*, 36. https://doi.org/10.1186/s13059-020-1949-z.
125. Jiang, P., Zhang, Y., Ru, B., Yang, Y., Vu, T., Paul, R., Mirza, A., Altan-Bonnet, G., Liu, L., Ruppin, E., et al. (2021). Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. Nat. Methods *18*, 1181–1191. https://doi.org/10.1038/s41592-021-01274-5.
126. Badia-i-Mompel, P., Vélez Santiago, J., Braunger, J., Geiss, C., Dimitrov, D., Müller-Dott, S., Taus, P., Dugourd, A., Holland, C.H., Ramirez Flores, R.O., et al. (2022). decoupleR: ensemble of computational methods to infer biological activities from omics data. Bioinforma. Adv. *2*, vbac016. https://doi.org/10.1093/bioadv/vbac016.
127. Müller-Dott, S., Tsirvouli, E., Vazquez, M., Ramirez Flores, R.O., Badia-I-Mompel, P., Fallegger, R., Türei, D., Lægreid, A., and Saez-Rodriguez, J. (2023). Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities. Nucleic Acids Res. *51*, 10934–10949. https://doi.org/10.1093/nar/gkad841.
128. Seabold, S., and Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. scipy. https://doi.org/10.25080/Majora-92bf1922-011.
129. Sun, D., Guan, X., Moran, A.E., Wu, L.-Y., Qian, D.Z., Schedin, P., Dai, M.-S.,

Danilov, A.V., Alumkal, J.J., Adey, A.C., et al. (2022). Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. Nat. Biotechnol. *40*, 527–538. https://doi.org/10.1038/s41587-021-01091-3.

130. Oliveira, M.F., Romero, J.P., Chung, M., Williams, S., Gottscho, A.D., Gupta, A., Pilipauskas, S.E., Mohabbat, S., Raman, N., Sukovich, D., et al. (2024). Characterization of immune cell populations in the tumor microenvironment of colorectal cancer using high definition spatial profiling. Preprint at bioRxiv, https://doi.org/10.1101/2024.06.04.597233 https://doi.org/10.1101/2024.06.04.597233.
131. Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A.C., Kuemmerle, L.B., Rybakov, S., Ibarra, I.L., Holmberg, O., Virshup, I., et al. (2022). Squidpy: a scalable framework for spatial omics analysis. Nat. Methods *19*, 171–178. https://doi.org/10.1038/s41592-021-01358-2.
132. Marconato, L., Palla, G., Yamauchi, K.A., Virshup, I., Heidari, E., Treis, T., Vierdag, W.-M., Toth, M., Stockhaus, S., Shrestha, R.B., et al. (2024). SpatialData: an open and universal data framework for spatial omics. Nat. Methods. https://doi.org/10.1038/s41592-024-02212-x.
133. Salim, A., Bhuva, D.D., Chen, C., Yang, P., Davis, M.J., and Yang, J.Y. (2024). SpaNorm: spatially-aware normalisation for spatial transcriptomics data. Preprint at bioRxiv, https://doi.org/10.1101/2024.05.31.596908 https://doi.org/10.1101/2024.05.31.596908.
134. Biancalani, T., Scalia, G., Buffoni, L., Avasthi, R., Lu, Z., Sanger, A., Tokcan, N., Vanderburg, C.R., Segerstolpe, Å., Zhang, M., et al. (2021). Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. Nat. Methods *18*, 1352–1362. https://doi.org/10.1038/s41592-021-01264-7.
135. Song, D., Xi, N.M., Li, J.J., and Wang, L. (2022). scSampler: fast diversity-preserving subsampling of large-scale single-cell transcriptomic data. Bioinforma. Oxf. Engl. *38*, 3126–3127. https://doi.org/10.1093/bioinformatics/btac271.
136. Birk, S., Bonafonte-Pardàs, I., Feriz, A.M., Boxall, A., Agirre, E., Memi, F., Maguza, A., Fan, R., Castelo-Branco, G., Theis, F.J., et al. (2024). Quantitative characterization of cell niches in spatial atlases. Preprint at bioRxiv, https://doi.org/10.1101/2024.02.21.581428 https://doi.org/10.1101/2024.02.21.581428.
137. Türei, D., Valdeolivas, A., Gul, L., Palacio-Escat, N., Klein, M., Ivanova, O., Ölbei, M., Gábor, A., Theis, F., Módos, D., et al. (2021). Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. Mol. Syst. Biol. *17*, e9923. https://doi.org/10.15252/msb.20209923.
138. Browaeys, R., Saelens, W., and Saeys, Y. (2020). NicheNet: modeling intercellular communication by linking ligands to target genes. Nat. Methods *17*, 159–162. https://doi.org/10.1038/s41592-019-0667-5.
139. Sato, T., and Clevers, H. (2013). Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. Science *340*, 1190–1194. https://doi.org/10.1126/science.1234852.
140. Fumagalli, A., Suijkerbuijk, S.J.E., Begthel, H., Beerling, E., Oost, K.C., Snippert, H.J., van Rheenen, J., and Drost, J. (2018). A surgical orthotopic organoid transplantation approach in mice to visualize and study colorectal cancer progression. Nat. Protoc. *13*, 235–247. https://doi.org/10.1038/nprot.2017.137.
141. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., et al. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat. Biotechnol. *42*, 293–304. https://doi.org/10.1038/s41587-023-01767-y.

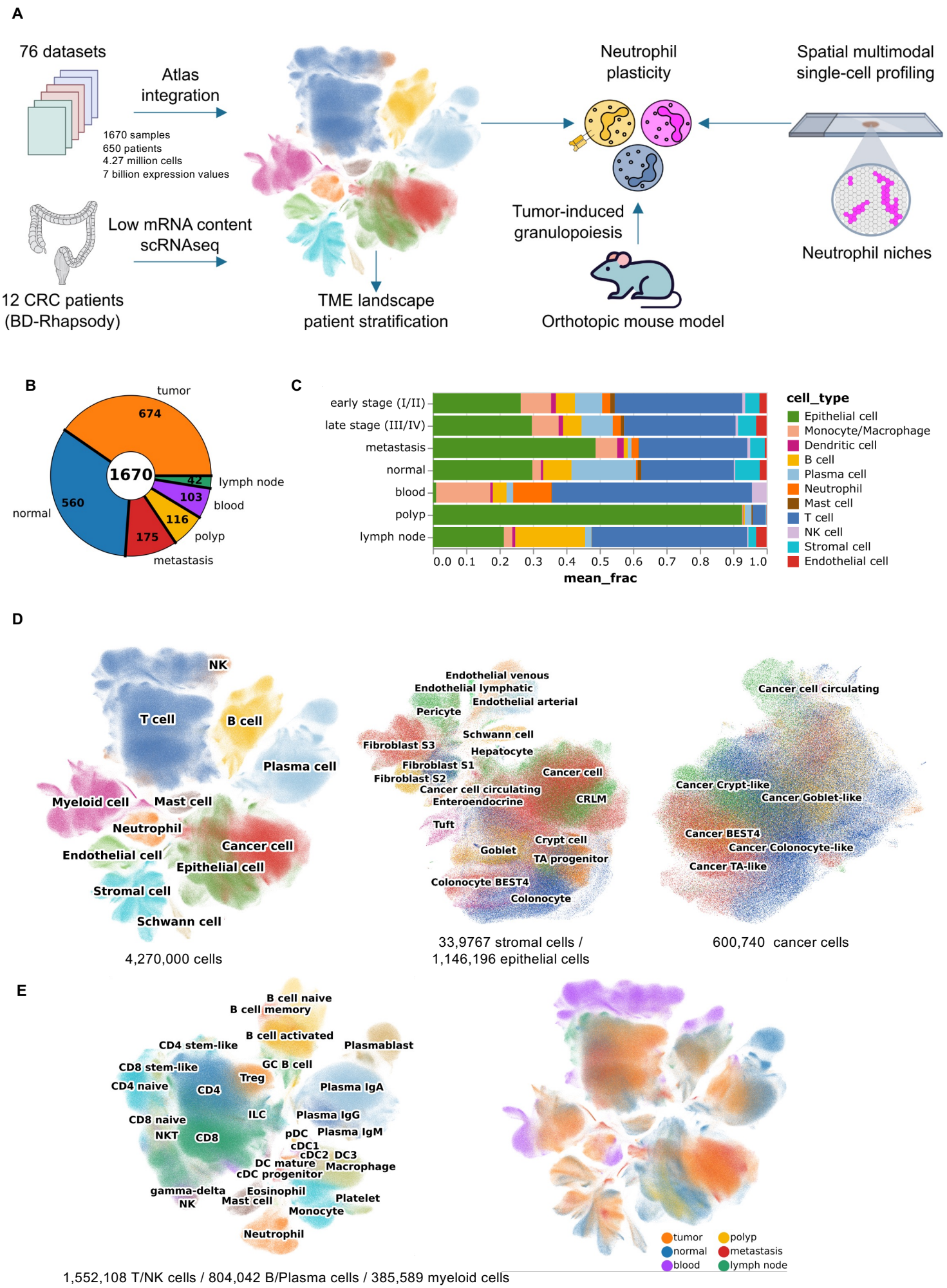

A
76 datasets
Atlas integration
1670 samples
650 patients
4.27 million cells
7 billion expression values
Low mRNA content scRNAseq
12 CRC patients (BD-Rhapsody)
TME landscape patient stratification
Neutrophil plasticity
Spatial multimodal single-cell profiling
Tumor-induced granulopoiesis
Neutrophil niches
Orthotopic mouse model
B
tumor
674
1670
42
lymph node
103
blood
560
normal
116
polyp
175
metastasis
C
early stage (I/II)
late stage (III/IV)
metastasis
normal
blood
polyp
lymph node
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
mean_frac
cell_type
Epithelial cell
Monocyte/Macrophage
Dendritic cell
B cell
Plasma cell
Neutrophil
Mast cell
T cell
NK cell
Stromal cell
Endothelial cell
D
NK
T cell
B cell
Plasma cell
Myeloid cell
Mast cell
Neutrophil
Cancer cell
Endothelial cell
Epithelial cell
Stromal cell
Schwann cell
4,270,000 cells
Endothelial venous
Endothelial lymphatic
Endothelial arterial
Pericyte
Schwann cell
Fibroblast S3
Hepatocyte
Fibroblast S1
Cancer cell
Fibroblast S2
Cancer cell circulating
Enteroendocrine
CRLM
Tuft
Crypt cell
Goblet
TA progenitor
Colonocyte BEST4
Colonocyte
33,9767 stromal cells /
1,146,196 epithelial cells
Cancer cell circulating
Cancer Crypt-like
Cancer Goblet-like
Cancer BEST4
Cancer Colonocyte-like
Cancer TA-like
600,740 cancer cells
E
B cell naive
B cell memory
B cell activated
CD4 stem-like
Plasmablast
CD8 stem-like
GC B cell
Treg
CD4 naive
CD4
Plasma IgA
ILC
Plasma IgG
CD8 naive
NKT
CD8
pDC
Plasma IgM
cDC1
cDC2
DC3
DC mature
Macrophage
cDC progenitor
gamma-delta
Eosinophil
NK
Mast cell
Platelet
Monocyte
Neutrophil
tumor
normal
blood
polyp
metastasis
lymph node
1,552,108 T/NK cells / 804,042 B/Plasma cells / 385,589 myeloid cells

Figure 1

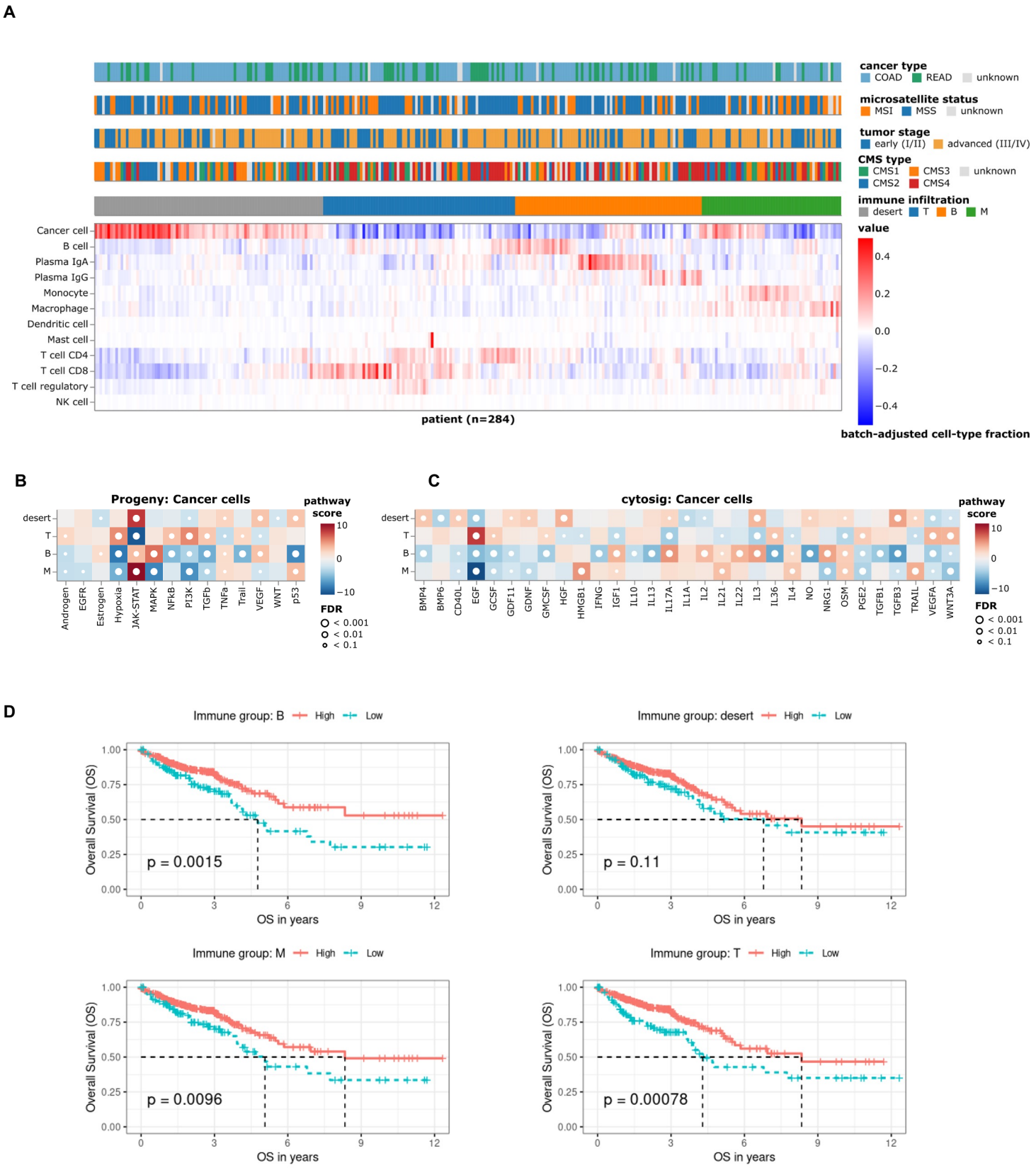
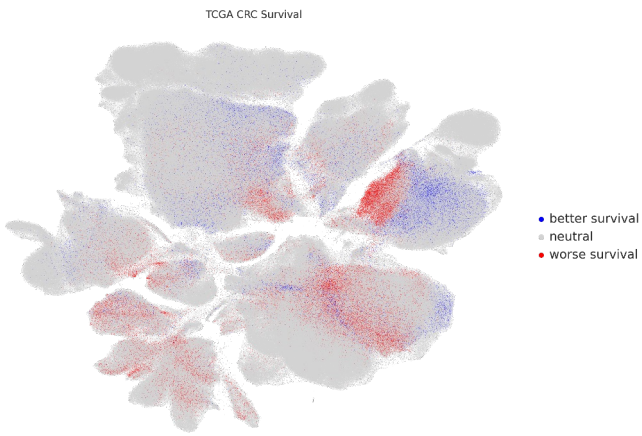
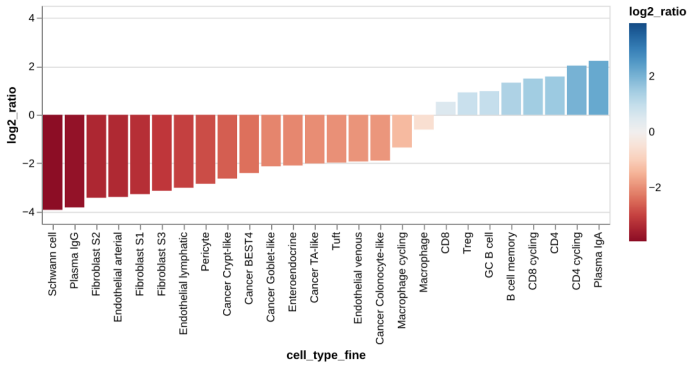

A
cancer type
COAD READ unknown
microsatellite status
MSI MSS unknown
tumor stage
early (I/II) advanced (III/IV)
CMS type
CMS1 CMS3 unknown
CMS2 CMS4
immune infiltration
desert T B M
value
0.4
0.2
0.0
−0.2
−0.4
batch-adjusted cell-type fraction
Cancer cell
B cell
Plasma IgA
Plasma IgG
Monocyte
Macrophage
Dendritic cell
Mast cell
T cell CD4
T cell CD8
T cell regulatory
NK cell
patient (n=284)
B
Progeny: Cancer cells
pathway score
10
0
−10
desert
T
B
M
Androgen
EGFR
Estrogen
Hypoxia
JAK-STAT
MAPK
NFkB
PI3K
TGFb
TNFa
Trail
VEGF
WNT
p53
FDR
< 0.001
< 0.01
< 0.1
C
cytosig: Cancer cells
pathway score
10
0
−10
desert
T
B
M
BMP4
BMP6
CD40L
EGF
GCSF
GDF11
GDNF
GMCSF
HGF
HMGB1
IFNG
IGF1
IL10
IL13
IL17A
IL1A
IL2
IL21
IL22
IL3
IL36
IL4
NO
NRG1
OSM
PGE2
TGFB1
TGFB3
TRAIL
VEGFA
WNT3A
FDR
< 0.001
< 0.01
< 0.1
D
Immune group: B High Low
Overall Survival (OS)
p = 0.0015
OS in years
Immune group: desert High Low
Overall Survival (OS)
p = 0.11
OS in years
Immune group: M High Low
Overall Survival (OS)
p = 0.0096
OS in years
Immune group: T High Low
Overall Survival (OS)
p = 0.00078
OS in years

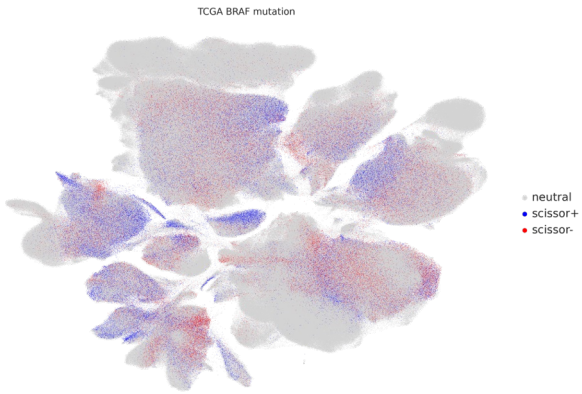
Figure 2

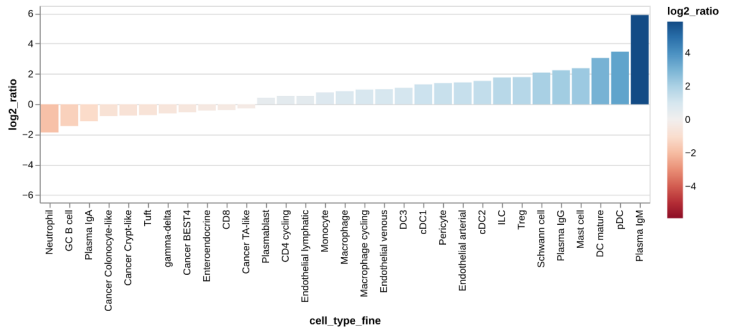
A


TCGA CRC Survival
better survival
neutral
worse survival

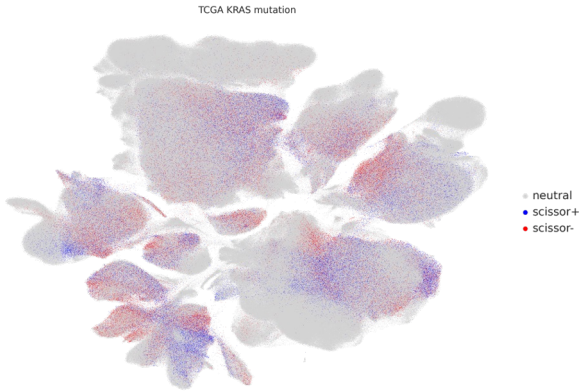

log2_ratio
cell_type_fine
Schwann cell
Plasma IgG
Fibroblast S2
Endothelial arterial
Fibroblast S1
Fibroblast S3
Endothelial lymphatic
Pericyte
Cancer Crypt-like
Cancer BEST4
Cancer Goblet-like
Enteroendocrine
Cancer TA-like
Tuft
Endothelial venous
Cancer Colonocyte-like
Macrophage cycling
Macrophage
CD8
Treg
GC B cell
B cell memory
CD8 cycling
CD4
CD4 cycling
Plasma IgA

B


TCGA BRAF mutation
neutral
scissor+
scissor-


log2_ratio
cell_type_fine
Neutrophil
GC B cell
Plasma IgA
Cancer Colonocyte-like
Cancer Crypt-like
Tuft
gamma-delta
Cancer BEST4
Enteroendocrine
CD8
Cancer TA-like
Plasmablast
CD4 cycling
Endothelial lymphatic
Monocyte
Macrophage
Macrophage cycling
Endothelial venous
DC3
cDC1
Pericyte
Endothelial arterial
cDC2
ILC
Treg
Schwann cell
Plasma IgG
Mast cell
DC mature
pDC
Plasma IgM

C


TCGA KRAS mutation
neutral
scissor+
scissor-

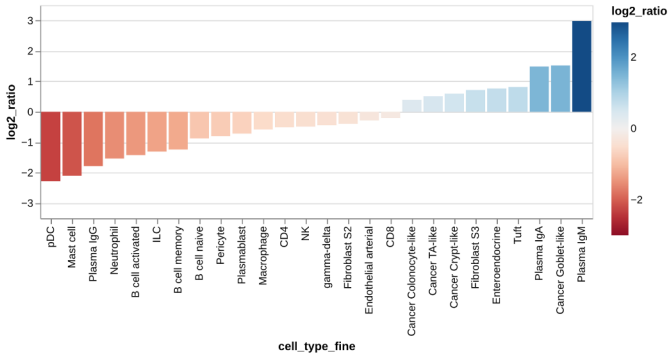

log2_ratio
cell_type_fine
pDC
Mast cell
Plasma IgG
Neutrophil
B cell activated
ILC
B cell memory
B cell naive
Pericyte
Plasmablast
Macrophage
CD4
NK
gamma-delta
Fibroblast S2
Endothelial arterial
CD8
Cancer Colonocyte-like
Cancer TA-like
Cancer Crypt-like
Fibroblast S3
Enteroendocrine
Tuft
Plasma IgA
Cancer Goblet-like
Plasma IgM

Figure 3

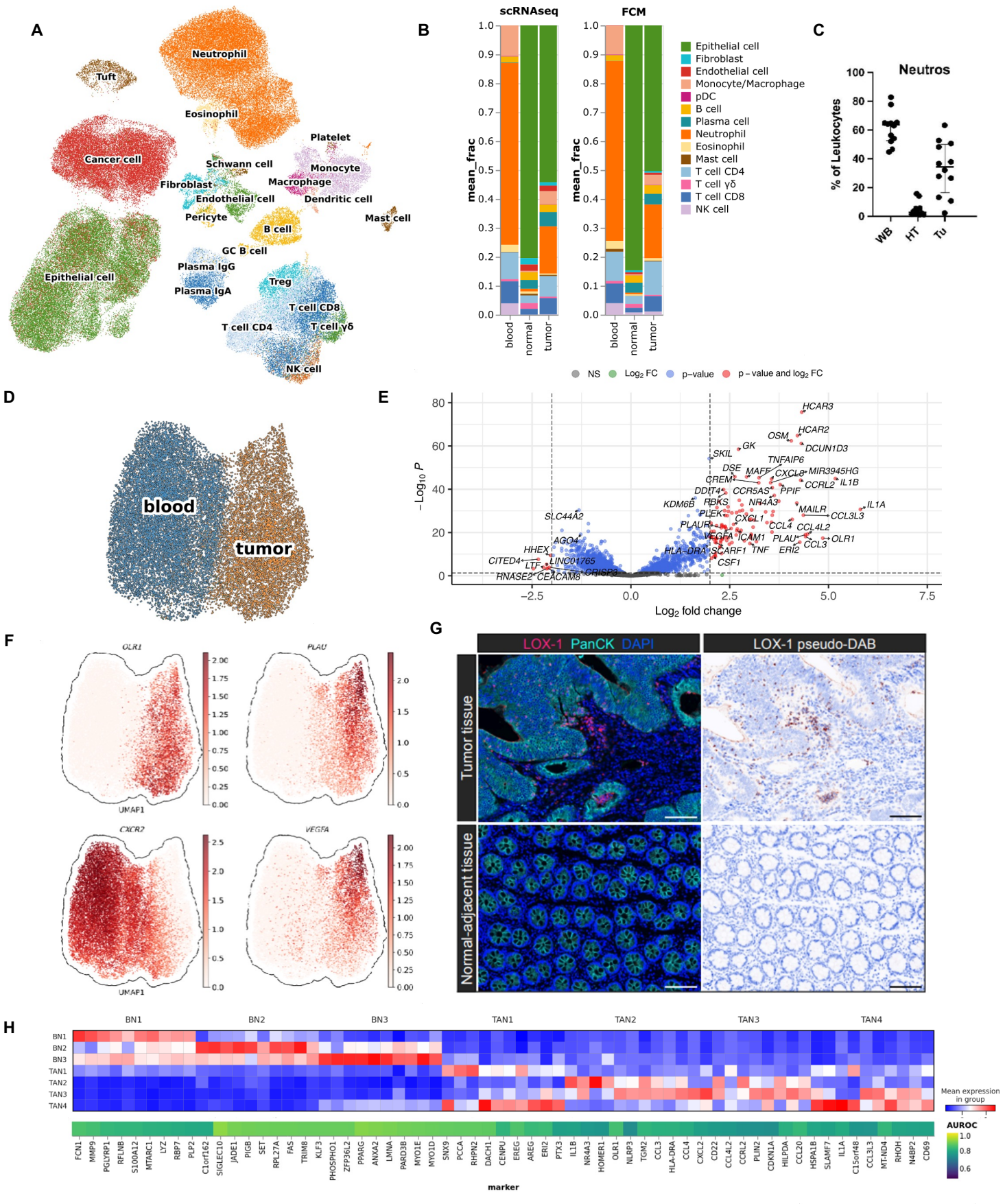

A
Neutrophil
Tuft
Eosinophil
Platelet
Cancer cell
Schwann cell
Monocyte
Fibroblast
Macrophage
Endothelial cell
Dendritic cell
Pericyte
Mast cell
B cell
GC B cell
Plasma IgG
Epithelial cell
Treg
Plasma IgA
T cell CD8
T cell CD4
T cell γδ
NK cell
B
scRNAseq
FCM
mean_frac
blood
normal
tumor
Epithelial cell
Fibroblast
Endothelial cell
Monocyte/Macrophage
pDC
B cell
Plasma cell
Neutrophil
Eosinophil
Mast cell
T cell CD4
T cell γδ
T cell CD8
NK cell
NS
Log2 FC
p-value
p-value and log2 FC
C
Neutros
% of Leukocytes
WB
HT
Tu
D
blood
tumor
E
−Log10 P
Log2 fold change
HCAR3
HCAR2
OSM
DCUN1D3
GK
SKIL
TNFAIP6
DSE
MAFF
CXCL8
MIR3945HG
CREM
CCRL2
IL1B
DDIT4
CCR5AS
PPIF
KDM6B
RBKS
NR4A3
IL1A
MAILR
SLC44A2
PLEK
CXCL1
CCL3L3
PLAUR
CCL4
CCL4L2
VEGFA
ICAM1
PLAU
OLR1
AGO4
HHEX
HLA-DRA
SCARF1
TNF
ERI2
CCL3
CITED4
LINC01765
CSF1
LTF
CRISP3
RNASE2
CEACAM8
F
OLR1
PLAU
CXCR2
VEGFA
UMAP1
G
LOX-1 PanCK DAPI
LOX-1 pseudo-DAB
Tumor tissue
Normal-adjacent tissue
H
BN1
BN2
BN3
TAN1
TAN2
TAN3
TAN4
Mean expression in group
AUROC
FCN1
MMP9
PGLYRP1
RFLNB
S100A12
MTARC1
LYZ
RBP7
PLP2
C1orf162
SIGLEC10
JADE1
PIGB
SET
RPL27A
FAS
TRIM8
KLF3
PHOSPHO1
ZFP36L2
PPARG
ANXA2
LMNA
PARD3B
MYO1E
MYO1D
SNX9
PCCA
RHPN2
DACH1
CENPU
EREG
AREG
ERI2
PTX3
IL1B
NR4A3
HOMER1
OLR1
NLRP3
TGM2
CCL3
HLA-DRA
CCL4
CXCL2
CD22
CCL4L2
CCRL2
PLIN2
CDKN1A
HILPDA
CCL20
HSPA1B
SLAMF7
IL1A
C15orf48
CCL3L3
MT-ND4
RHOH
N4BP2
CD69
marker
I
BN1
BN2
BN3
TAN1
TAN2
TAN3
TAN4
J
HLA-DRA
K
blood
normal
tumor
I
IIA
IIIB
IVA
HLA-DRA
Fraction of cells in group (%)
5
10
Mean expression in group
0.1
0.15
L
HLA-DR MFI
HLA-DR MFI on Neutrophils
Blood
HT
Tumor

Figure 4

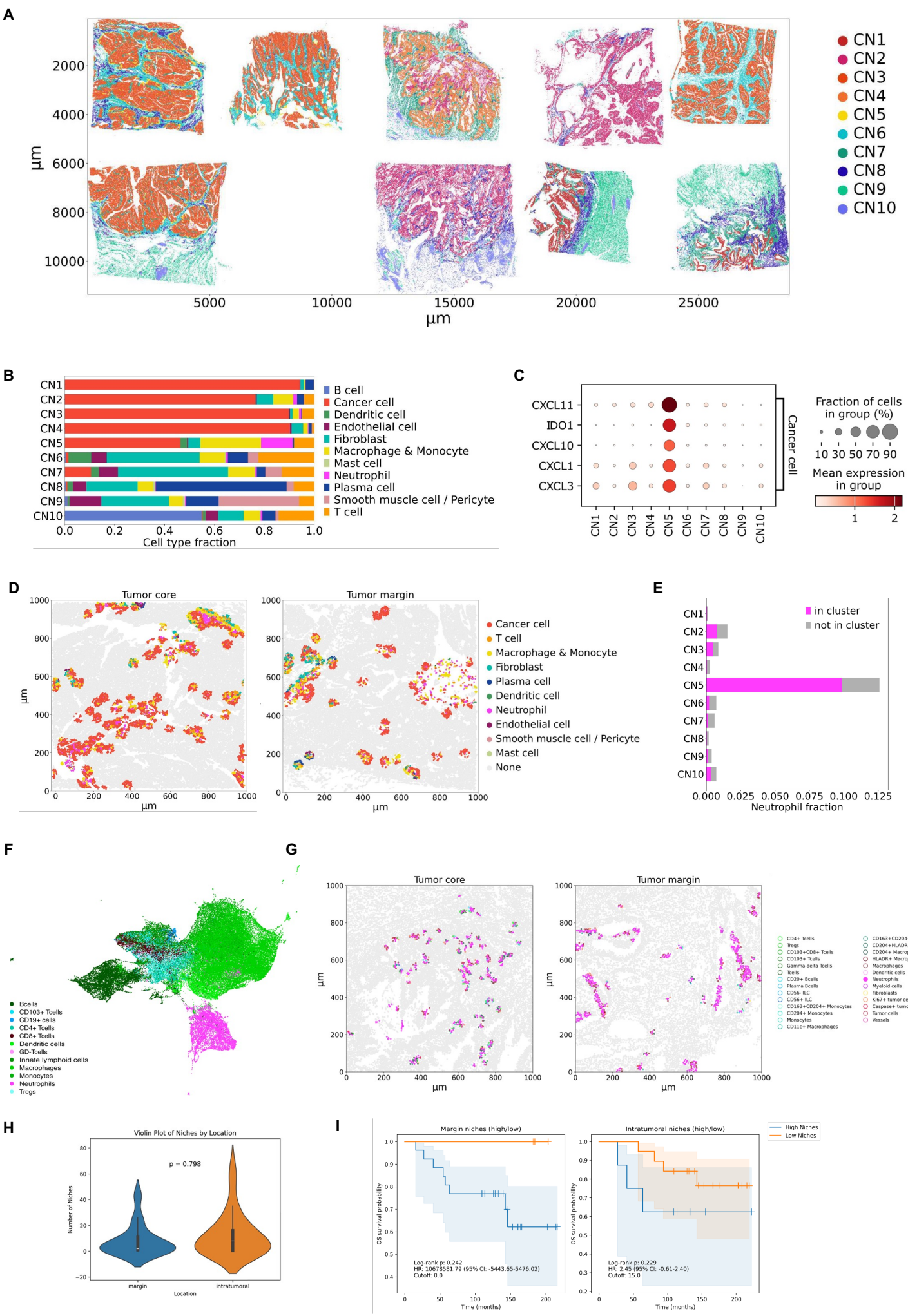

A
CN1
CN2
CN3
CN4
CN5
CN6
CN7
CN8
CN9
CN10
µm
B
B cell
Cancer cell
Dendritic cell
Endothelial cell
Fibroblast
Macrophage & Monocyte
Mast cell
Neutrophil
Plasma cell
Smooth muscle cell / Pericyte
T cell
Cell type fraction
C
CXCL11
IDO1
CXCL10
CXCL1
CXCL3
Cancer cell
Fraction of cells in group (%)
10 30 50 70 90
Mean expression in group
D
Tumor core
Tumor margin
Cancer cell
T cell
Macrophage & Monocyte
Fibroblast
Plasma cell
Dendritic cell
Neutrophil
Endothelial cell
Smooth muscle cell / Pericyte
Mast cell
None
E
in cluster
not in cluster
Neutrophil fraction
F
Bcells
CD103+ Tcells
CD19+ cells
CD4+ Tcells
CD8+ Tcells
Dendritic cells
GD-Tcells
Innate lymphoid cells
Macrophages
Monocytes
Neutrophils
Tregs
G
Tumor core
Tumor margin
H
Violin Plot of Niches by Location
p = 0.798
Number of Niches
margin
intratumoral
Location
I
Margin niches (high/low)
Intratumoral niches (high/low)
High Niches
Low Niches
OS survival probability
Log-rank p: 0.242
HR: 106785881.79 (95% CI: -5443.65-5476.02)
Cutoff: 0.0
Log-rank p: 0.229
HR: 2.45 (95% CI: -0.61-2.40)
Cutoff: 15.0
Time (months)

Figure 5

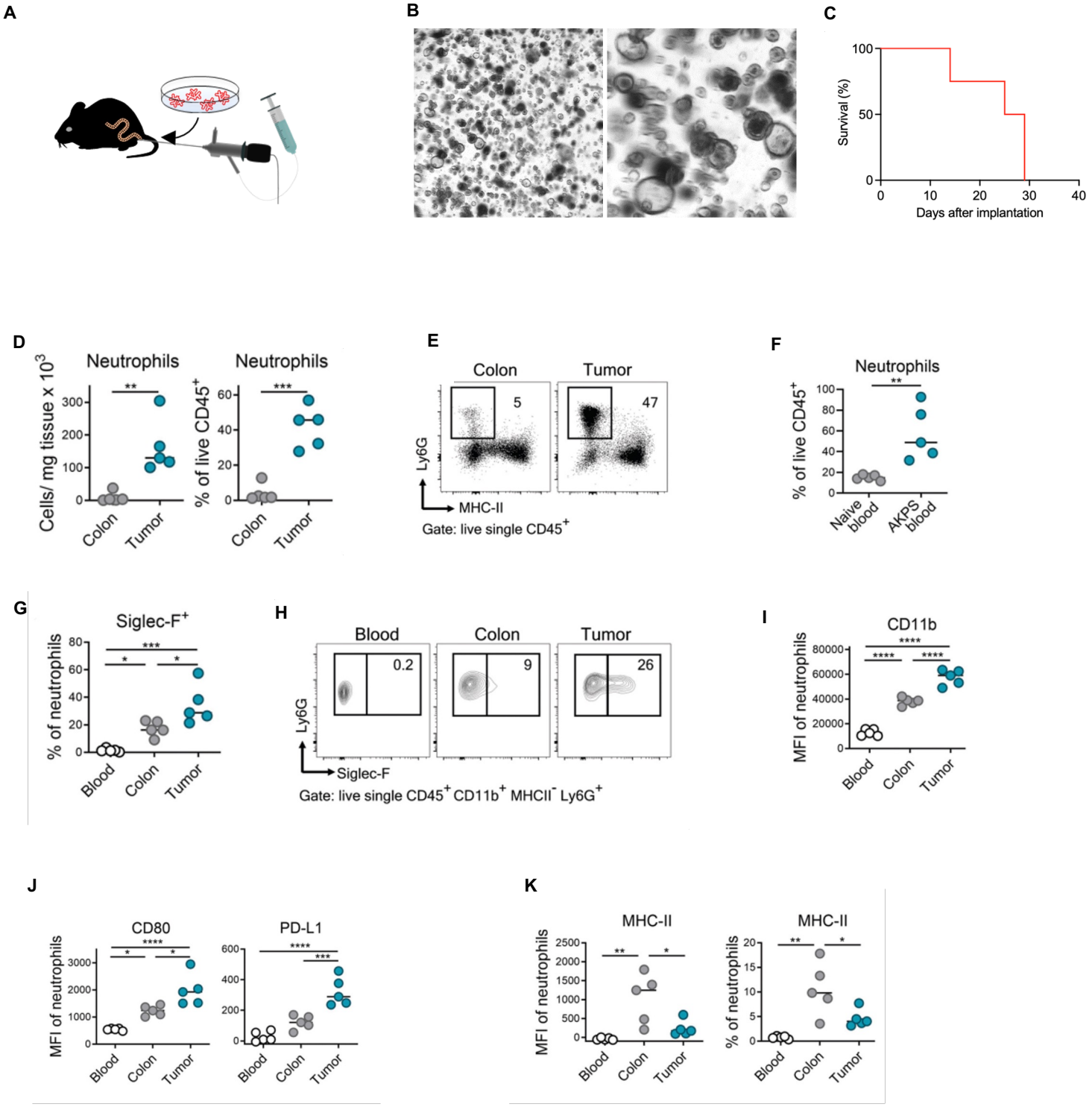

A
B
C
Survival (%)
Days after implantation
D
Neutrophils
Cells/ mg tissue x 10^3
% of live CD45+
Colon
Tumor
E
Colon
Tumor
Ly6G
MHC-II
Gate: live single CD45+
F
Neutrophils
% of live CD45+
Naive blood
AKPS blood
G
Siglec-F+
% of neutrophils
Blood
Colon
Tumor
H
Blood
Colon
Tumor
Ly6G
Siglec-F
Gate: live single CD45+ CD11b+ MHCII- Ly6G+
I
CD11b
MFI of neutrophils
J
CD80
PD-L1
MFI of neutrophils
K
MHC-II
MFI of neutrophils
% of neutrophils

Figure 6

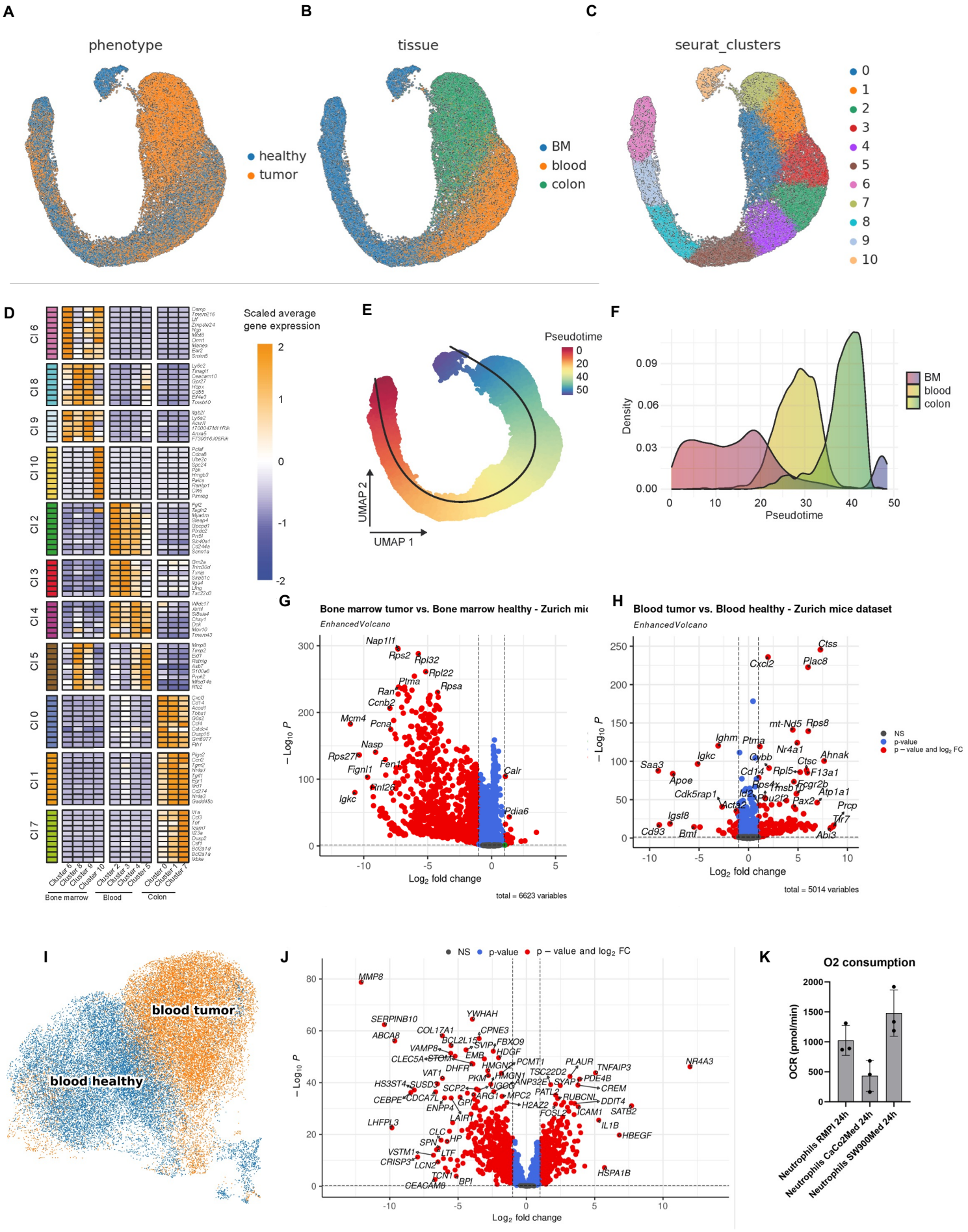

A
phenotype
healthy
tumor
B
tissue
BM
blood
colon
C
seurat_clusters
0
1
2
3
4
5
6
7
8
9
10
D
Scaled average gene expression
Bone marrow
Blood
Colon
E
Pseudotime
UMAP 2
UMAP 1
F
Density
Pseudotime
BM
blood
colon
G
Bone marrow tumor vs. Bone marrow healthy - Zurich mi
EnhancedVolcano
Log2 fold change
total = 6623 variables
H
Blood tumor vs. Blood healthy - Zurich mice dataset
EnhancedVolcano
Log2 fold change
total = 5014 variables
NS
p-value
p – value and log2 FC
I
blood tumor
blood healthy
J
NS
p-value
p – value and log2 FC
Log2 fold change
K
O2 consumption
OCR (pmol/min)
Neutrophils RMPI 24h
Neutrophils CaCo2Med 24h
Neutrophils SW900Med 24h

Figure 7

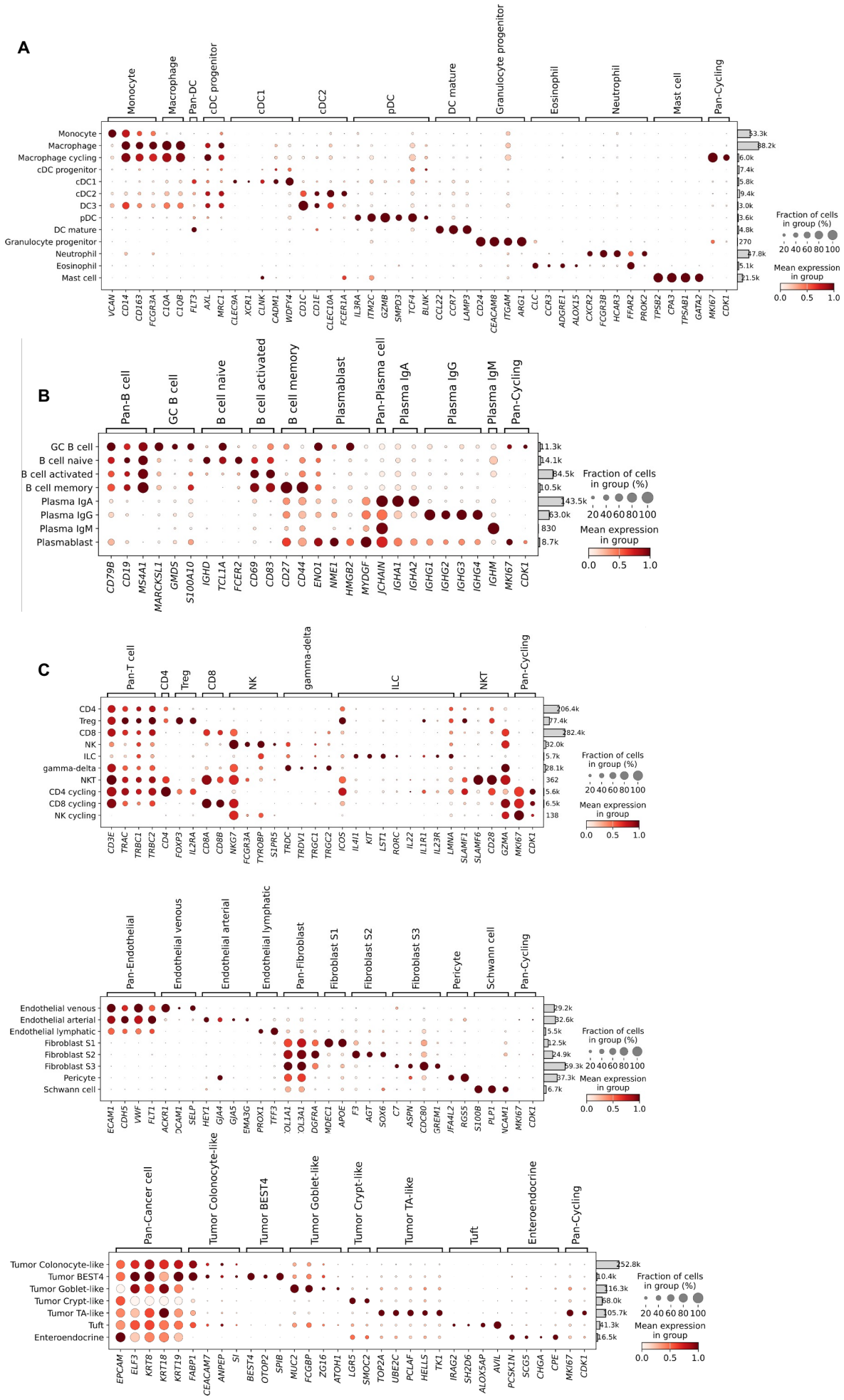


Figure S1: related to Figure 1. Composition of CRC atlas

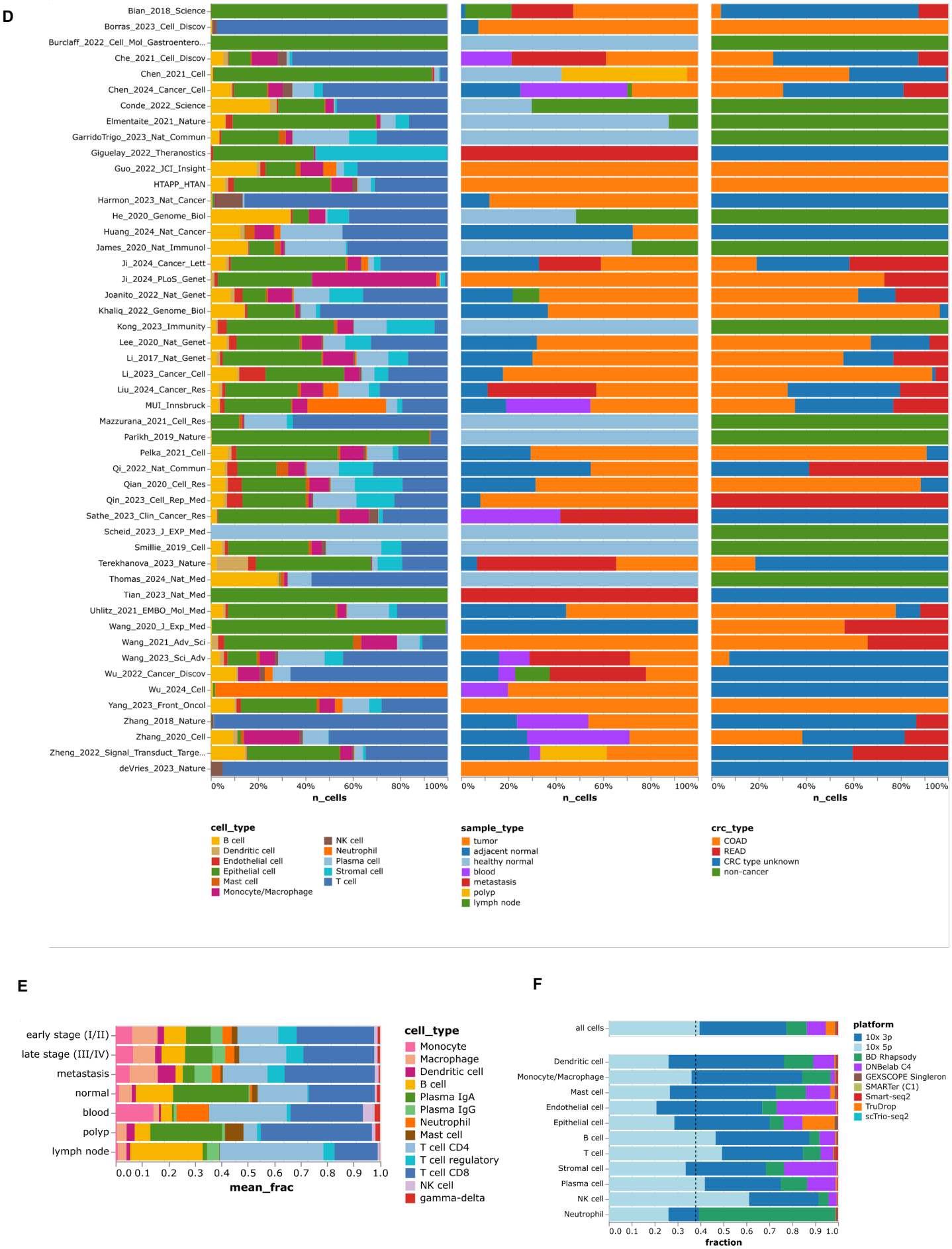


Figure S1: related to Figure 1. Composition of CRC atlas

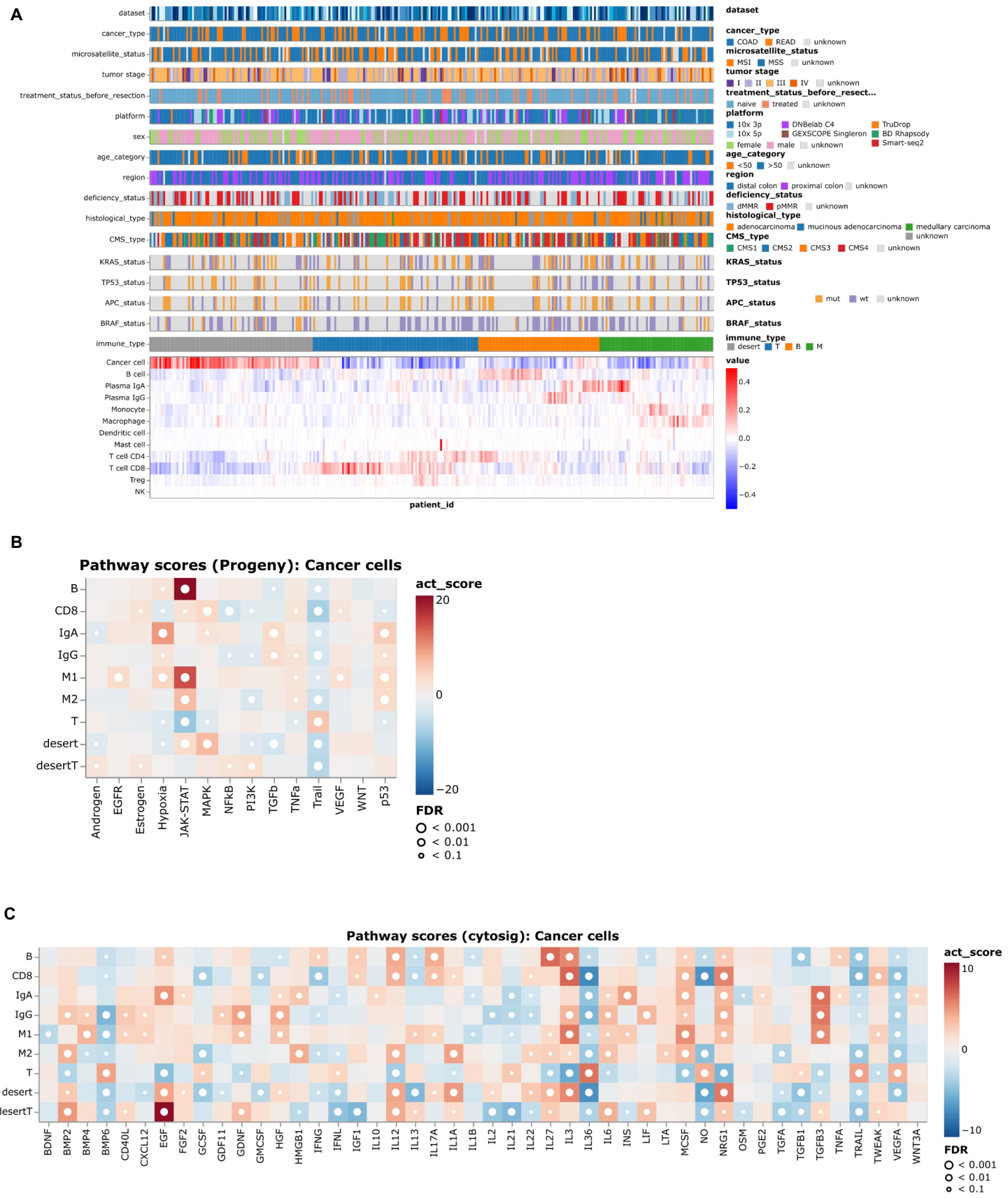
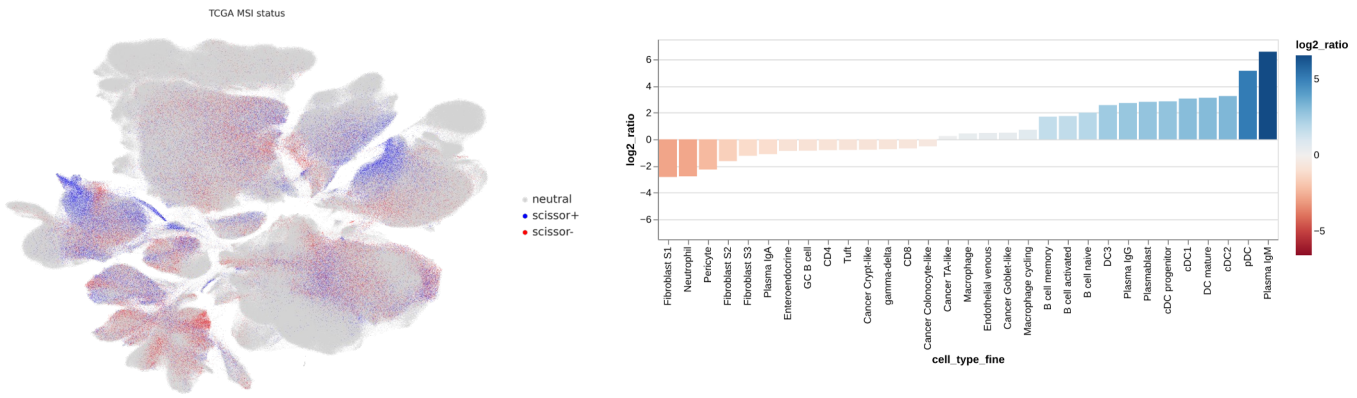


Figure S2: related to Figure 2. Tumor immune phenotypes in CRC. (**A**) Immune phenotypes and clinical, genetic features. (**B**) Results of the Progeny analysis. (**C**) Results of the cytosig analysis.

A



B

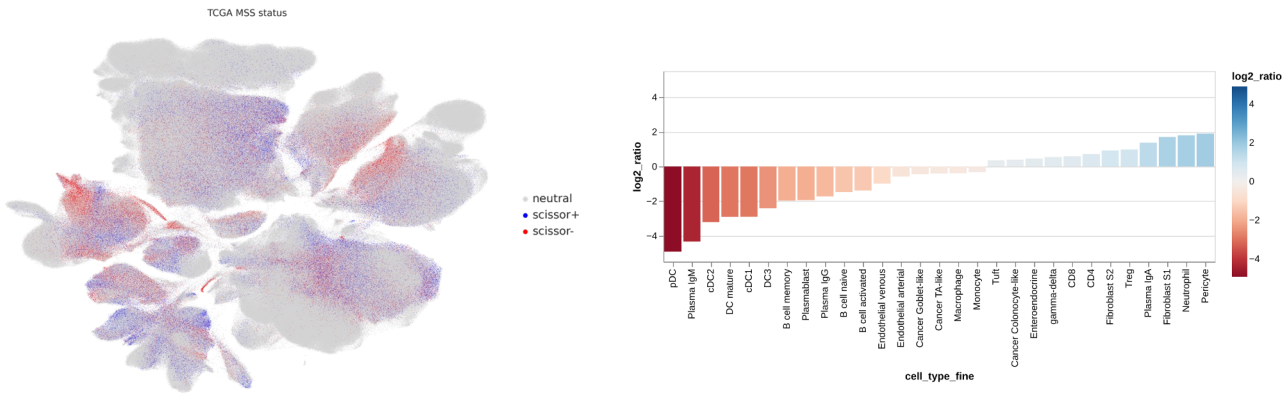


Figure S3: related to Figure 3. Association of cellular composition and distinct genotypes and survival in the TCGA data. (**A**) Association of cellular composition with MSI status in CRC patients (n = 177). (**B**) Association of cellular composition with MSS status in CRC patients (n = 412).

A

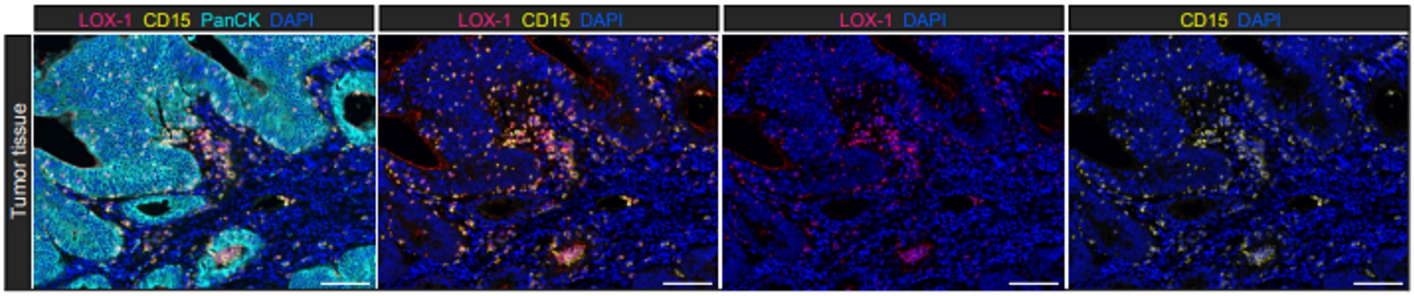


Figure S4: related to Figure 4. Characterization of neutrophils in CRC. (**A**) Selected multiplex immunofluorescence (mpIF) staining of LOX-1 (pink), CD15 (yellow), pan-cytokeratin (turquoise), and DAPI (blue) in colorectal cancer (CRC) tumor tissue. Scale bar: 100 µm.

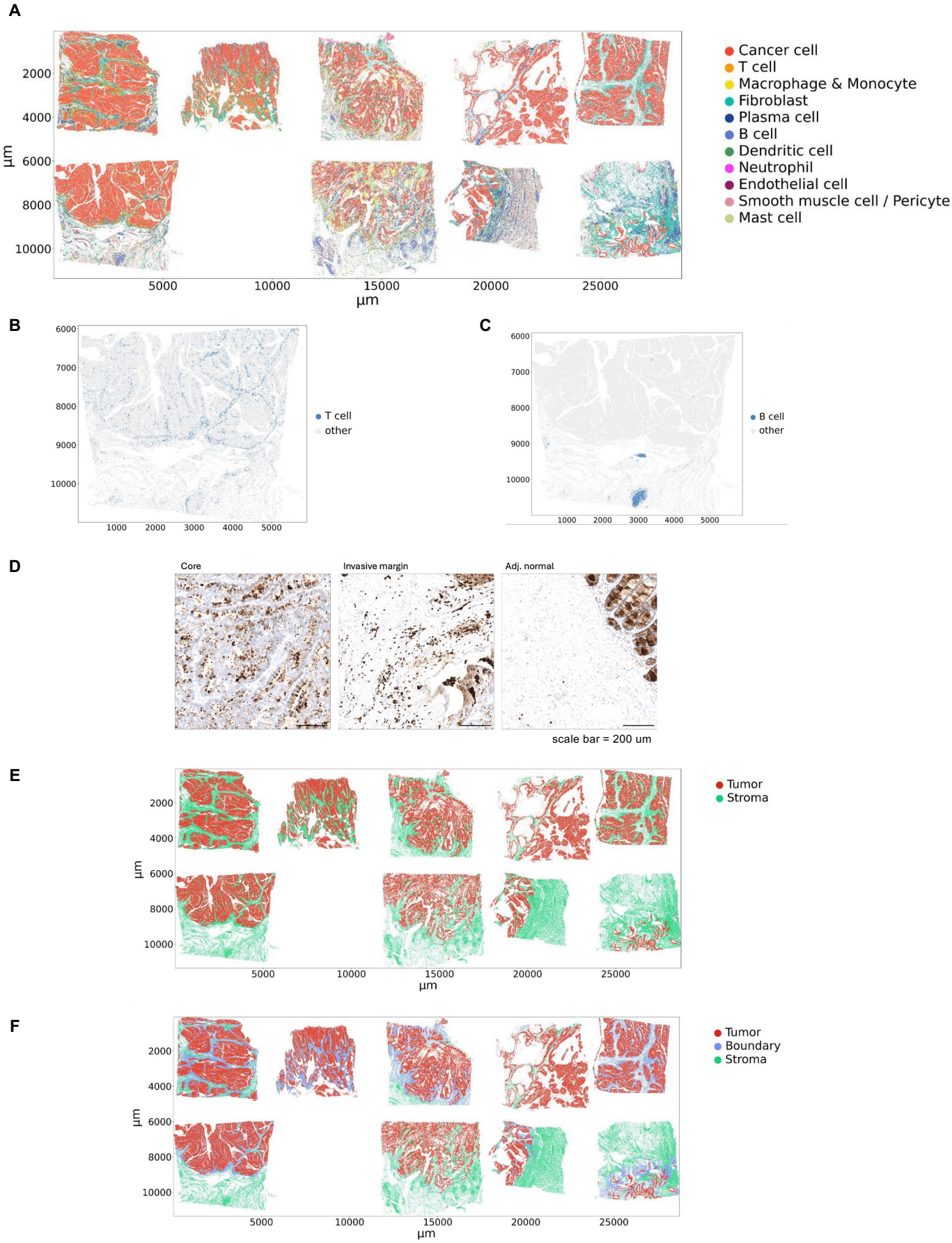
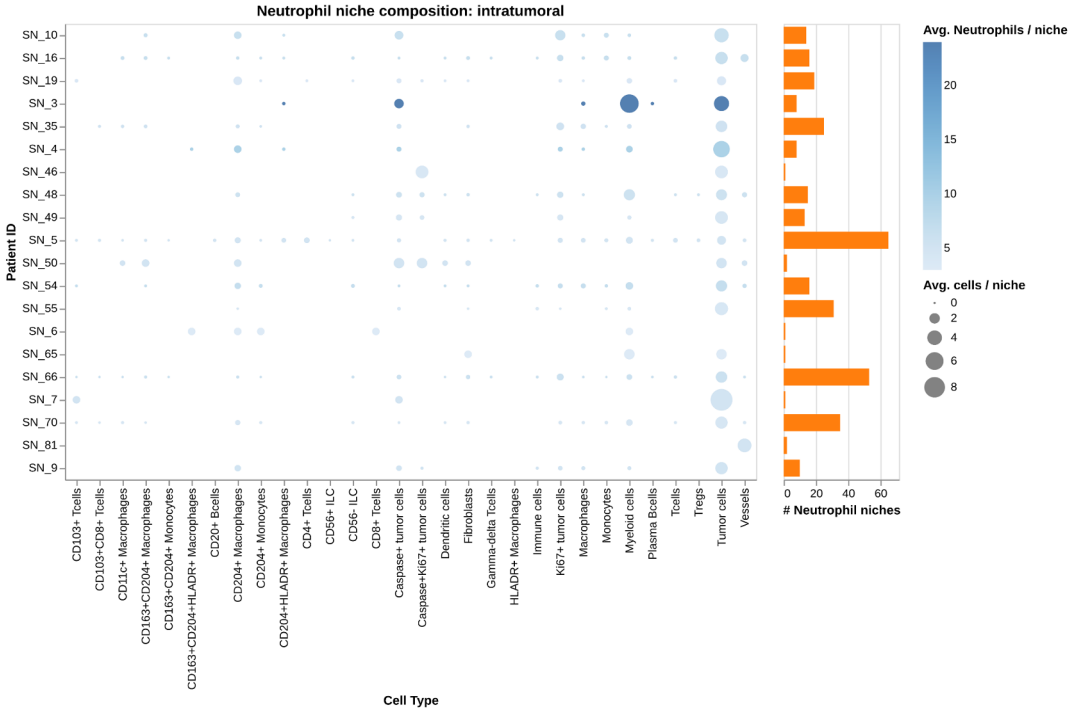


Figure S5: related to Figure 5. Spatial multi-modal profiling of neutrophils. (**A**) Cell type annotation of the spatial transcriptomics data. (**B-C**) Localization of the T cells and B cells (representatively shown for one patient). (**D**) IHC of neutrophils (**F-E**) Cell niches at a resolution of two and 3, respectively.

G



H

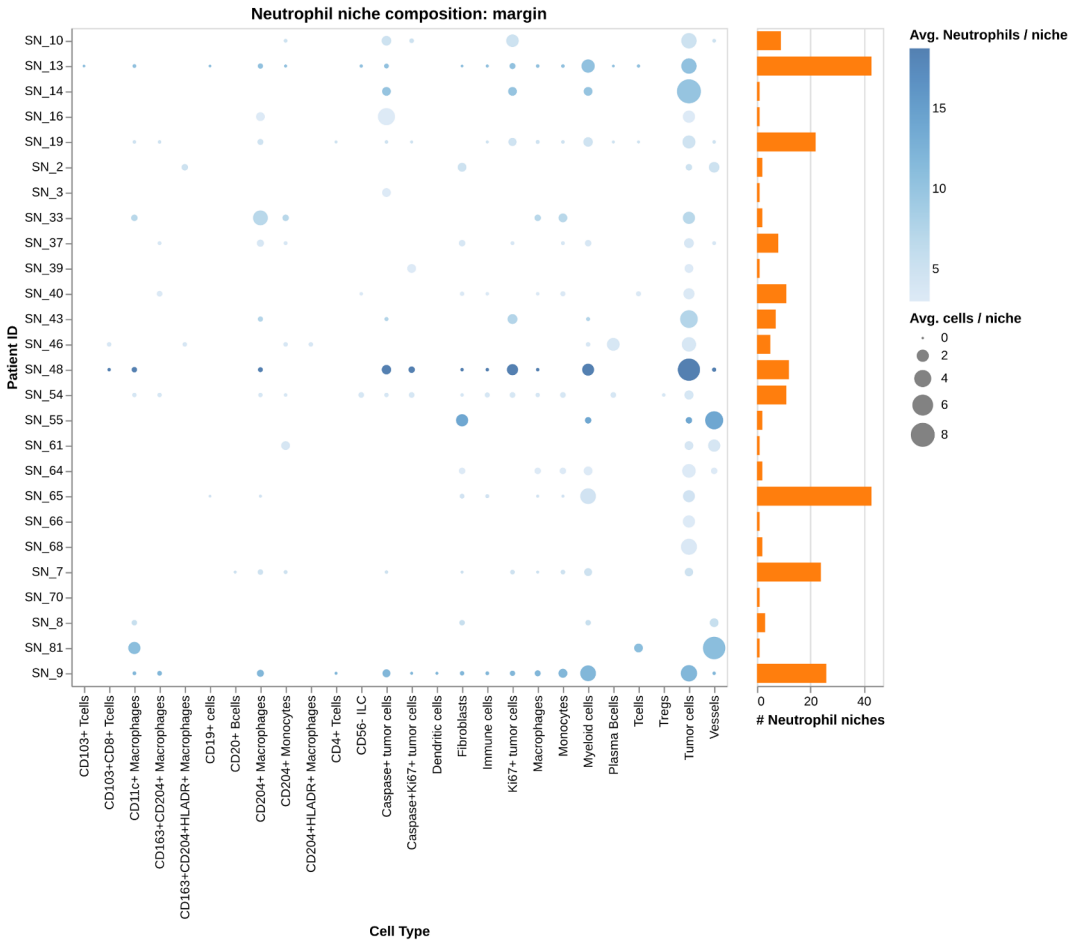
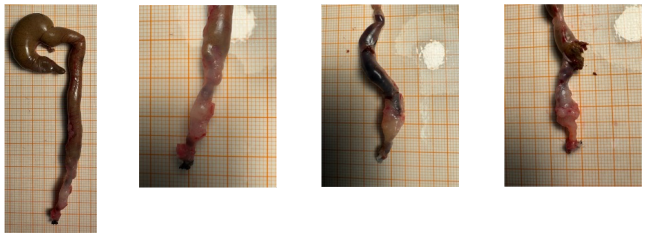
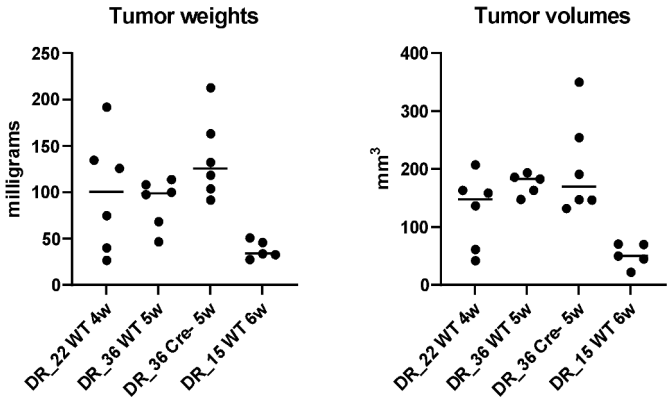


Figure S5: related to Figure 5. Spatial multi-modal profiling of neutrophils. (**G**) Neutrophil niche composition in the tumor core. (**H**) Neutrophil niche composition at the invasive margin.

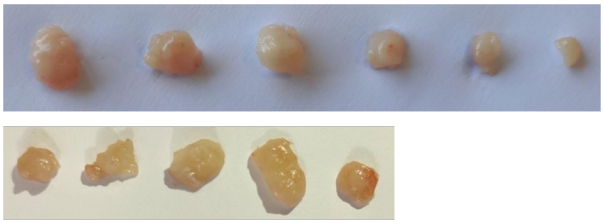
A



B



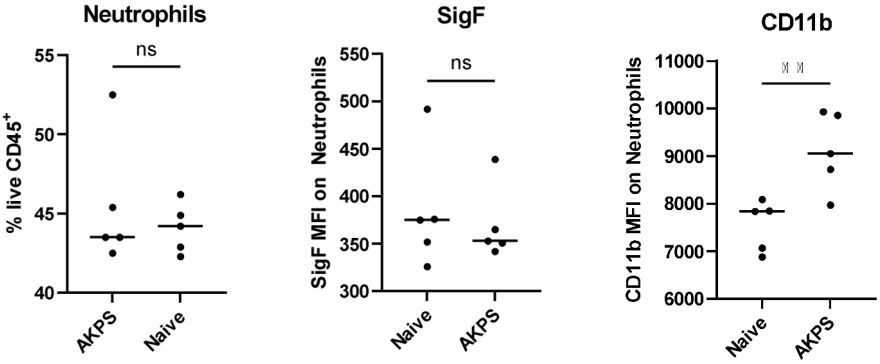
C



4 week AKPS

6 week AKPS

D



E

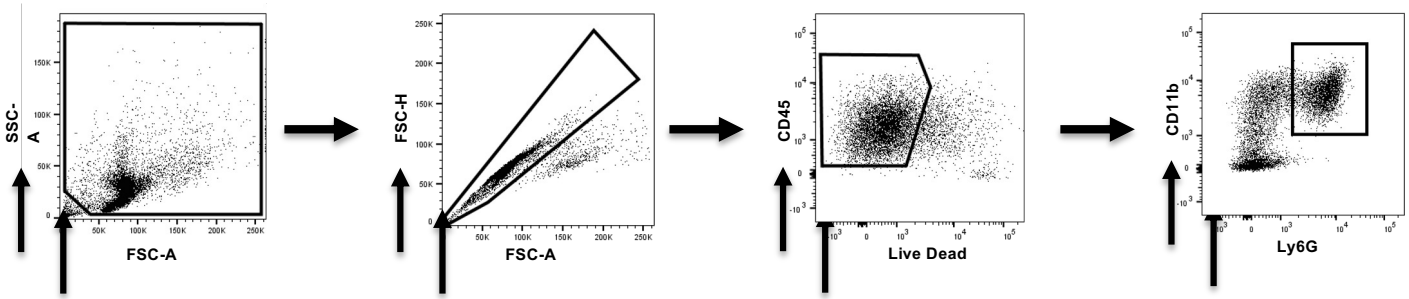


Figure S6: related to Figure 6. Orthotopic mouse model for CRC. (**A**) Orthotopic transplantation of AKPS organoids. (**B**) Tumor weights and tumor volumes of the mice. (**C**) Picture of the respective tumors. (**D**) Neutrophils in the bone marrow of the mice after 6 weeks. (**E**) Representative plots of the neutrophils in the bone marrow

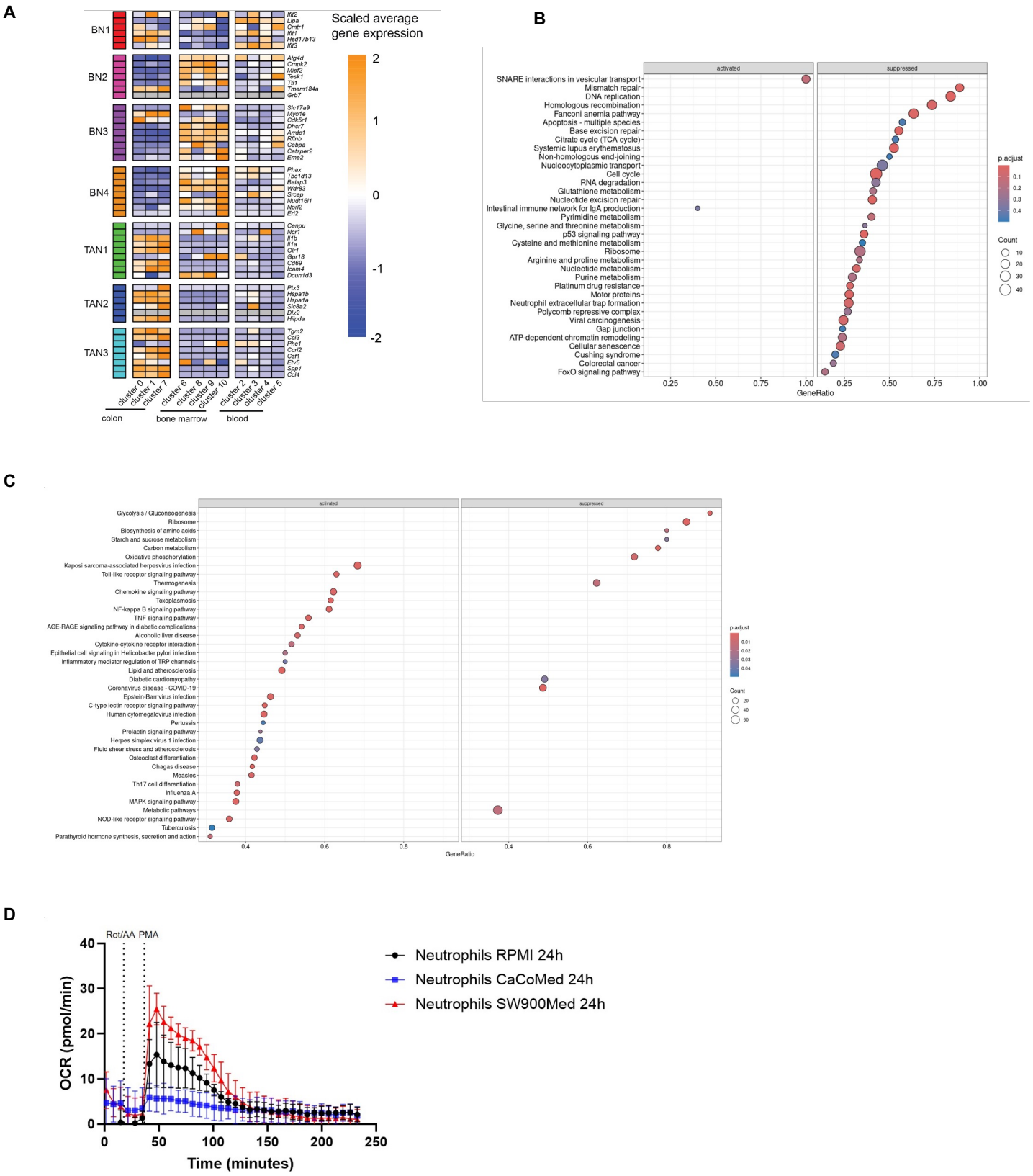


Figure S7: related to Figure 7. Tumors are reprogramming neutrophil metabolism. (**A**) Human-mouse comparisons of the expression profiles. Gene Set Enrichment Analysis (GSEA) on KEGG terms for (**B**) the expression profiles of bone marrow neutrophils from AKPC mice and controls, and (**C**) the expression profiles of blood neutrophils from healthy volunteers and patients with CRC. (**D**) XF Neutrophil Activation Assay kinetic traces of oxygen consumption rate (OCR) of healthy human neutrophils after 24h culture in SW900 supernatant, or CaCo2 supernatant or RPMI control medium. Mitochondrial function is blocked by Rot/AA injection, oxidative burst is activated by PMA injection. The data shown are the mean and SD, n=3 biological replicates.

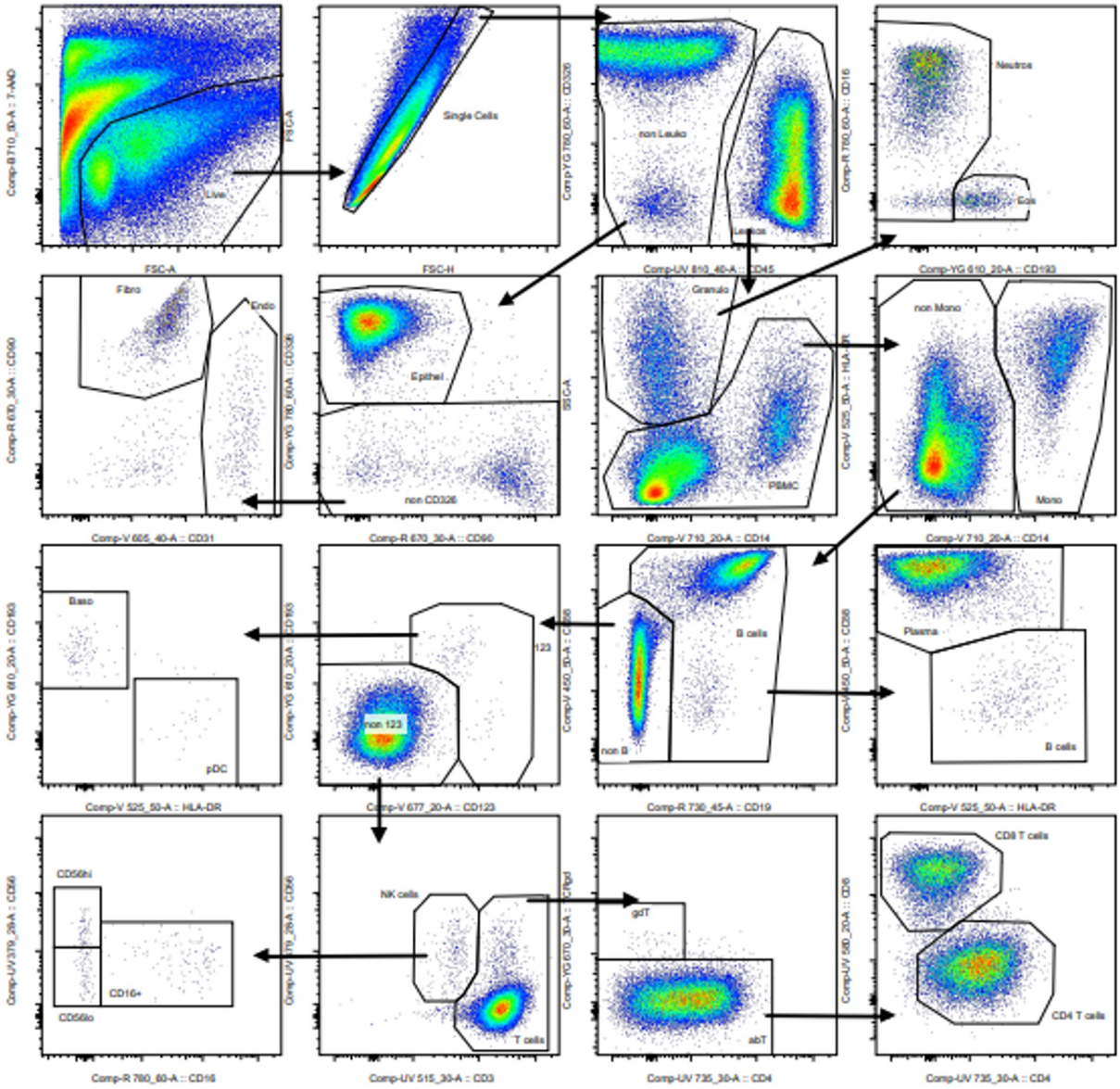


Figure S8: related to STAR Methods. Flow cytometry gating strategy to define cell populations from CRC tumor tissue, normal adjacent tissue and whole blood samples. In initial cleaning steps dead cells, debris and doublets were removed using 7-AAD staining and scatter characteristics. Leukocytes were defined by CD45 staining and sequentially gated into subtypes including neutrophils, monocytes, T cells and B cells. Non-CD45^{+} cells were gated into epithelial cells, endothelial cells and fibroblasts.