GZMK^{high} CD8^{+} T effector memory cells are associated with CD15^{high} neutrophil abundance in early-stage colorectal tumors and predict poor clinical outcome.

Silvia Tiberti[1], Carlotta Catozzi[1], Caterina Scirgolea[2], Ottavio Croci[3], Mattia Ballerini[1], Danilo Cagnina[1], Chiara Soriani[1], Carina B. Nava Lauson[1], Angeli D. Macandog[1], Giovanni Bertalot[4], Wanda L. Petz[5], Simona P. Ravenda[6], Valerio Licursi[7], Paola Paci[7], Marco Rasponi[8], Nicola Fazio[6], Guangwen Ren[9], Uberto Fumagalli-Romario[5], Martin H. Shaefer[1], Stefano Campaner[3], Enrico Lugli[2,10], Luigi Nezi[1*] and Teresa Manzo[1*#].

[1] Department of Experimental Oncology, European Institute of Oncology- IRCCS, Milano, Italy

[2] Laboratory of Translational Immunology, Humanitas Clinical and Research Center-IRCCS, Milan, Italy

[3] Center for Genomic Science of IIT, CGS@SEMM (Istituto Italiano di Tecnologia at European School of Molecular Medicine), Fondazione Istituto Italiano di Tecnologia (IIT), Milan, Italy.

[4] APSS Trento Hospital & Health Care, Trento, Italy

[5] Gastrointestinal Surgery Division, European Institute of Oncology-IRCCS

[6] Division of Gastrointestinal Medical Oncology and Neuroendocrine Tumors, European Institute of Oncology-IRCCS

[7] Department of Computer, Control and Management Engineering, Sapienza - University of Rome

[8] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy

[9] The Jackson Laboratory Cancer Center, Bar Harbor, ME, USA.

[10] Humanitas Flow Cytometry Core, Humanitas Clinical and Research Center – IRCCS, Milan, Italy.

[*] LN and TM contributed equally.

[#] Leading contact: teresa.manzo@ieo.it

ABSTRACT

Tumor contexture has emerged as a major prognostic determinant and tumor infiltrating $CD8^+$ T cells have been associated with a better prognosis in several solid tumors, including early-stage colorectal cancer (CRC). However, the tumor immune infiltrate is highly heterogeneous and understanding how the interplay between different immune cell compartments impacts on the clinical outcome is still in its infancy.

Here, we describe in a prospective cohort a novel $CD8^+$ T effector memory population, which is characterized by high levels of Granzyme K ($GZMK^{high}$ $CD8^+$ T_{EM}) and is correlated with $CD15^{high}$ tumor infiltrating neutrophils. We provide both *in vitro* and *in vivo* evidence of the role of stromal cell-derived factor 1 (CXCL12/SDF-1) in driving functional changes on neutrophils at the tumor site, promoting their retention and increasing the crosstalk with $CD8^+$ T cells. Mechanistically, as a consequence of the interaction with neutrophils, $CD8^+$ T cells are skewed towards a $CD8^+$ T_{EM} phenotype and produce high levels of GZMK, which in turn decreases E-cadherin pathway. The correlations of $GZMK^{high}$ $CD8^+$ T_{EM} and neutrophils with both tumor progression in mice and early relapse in CRC patients demonstrate the role of $GZMK^{high}$ $CD8^+$ T_{EM} in promoting malignancy. Indeed, a gene signature defining $GZMK^{high}$ $CD8^+$ T_{EM} was associated with worse prognosis on a larger independent cohort of CRC patients and a similar analysis was extended to lung cancer (TCGA).

Overall, our results highlight the emergence of $GZMK^{high}$ $CD8^+$ T_{EM} in early-stage CRC tumors as a hallmark driven by the interaction with neutrophils, which could implement current patient stratification and be targeted by novel therapeutics.

KEYWORDS

$CD8^+$ T effector memory cells

Granzyme K

Neutrophils

Colorectal cancer

Tumor microenvironment

BACKGROUND

Colorectal cancer (CRC) is the third most commonly diagnosed tumor worldwide and a leading cause of death for cancer, with the efficacy of treatment and subsequent survival largely dictated by the tumor-node-metastasis (TNM) stage at diagnosis. Tumor contexture has emerged as a major determinant to establish prognosis and guide novel therapies (Galon et al., 2014; Pagès et al., 2009a). However, despite favorable TNM staging and significant lymphocyte infiltration (Ogino et al., 2009), 30-40% of patients treated for early-stage CRC relapse and ultimately die from metastatic disease (American Cancer Society, 2018), demonstrating the inadequacy of current stratification.

In this regard, tumor infiltration by cytotoxic $CD8^+$ T cells has been associated with a better prognosis in several solid tumors, including CRC (Fountzilas et al., 2018; Koch et al., 2006) and high levels of memory $CD8^+$ T cells prevent early metastatic invasion and are associated with better survival (Pagès et al., 2009b), supporting the idea that the intratumor immune landscape contributes to tumor control and prevention of recurrence. However, the tumor immune infiltrate is highly heterogeneous and, although the influence of various myeloid cell populations on T cell activity has been reported (Binnewies et al., 2018; Lavin et al., 2017; Salmon et al., 2016), our understanding of the interplay between different cell compartments at the tumor site and their impact on the clinical outcome is still in its infancy. In particular, tumor-associated neutrophils represent a significant fraction of the inflammatory cells in the tumor microenvironment (TME) of many types of cancers, including CRC (Berry et al., 2017; Carus et al., 2013; Ilie et al., 2014; Li et al., 2018; Rao et al., 2012; Silvestre-Roig et al., 2019). The concept of neutrophil phenotypic and functional diversity has emerged from murine tumor models (Ballesteros et al., 2020), but the existence of subsets of infiltrating neutrophils with different functions in humans is still debated. As a result, the contribution of these cells to inhibiting or promoting tumor progression by reshaping anti-tumor immune

response remains largely unexplored and often contradictory in humans (Gentles et al., 2015; Singhal et al., 2016).

Here, we sought to gain insight into the crosstalk between the TME and immune cells to elucidate their contribution to relapse and implement patient's stratification for subsequent immune-based therapeutics. We have estimated heterogeneity of immune cells within and across CRC patients by employing multiparametric flow cytometry and single-cell transcriptomics, obtaining functional information. This combined approach has revealed an association of neutrophils infiltrating the tumor with a unique immune signature of tumor infiltrating lymphocytes (TILs), which is characterized by high levels of Granzyme K (GZMK) expression and leads to worse prognosis.

METHODS

Study approval.

Tumors, adjacent normal tissue and peripheral blood samples from patients diagnosed with CRC at European Institute of Oncology (IEO), were collected with written informed consent as approved by the Internal Research Center Review Board. All patients were treatment-naïve at the time of sample collection (Table 1) and fresh tissue was collected the day of the surgery. All donors provided their consent in accordance with the Declaration of Helsinki.

Mice were obtained from Charles River, housed and bred in a specific pathogen free animal facility. The experiments described in this manuscript were performed in accordance with the European Union Guideline on Animal Experiments. Mouse protocols were approved by the Italian Ministry of Health, the IEO Committee and the Institutional Animal Care and Use Committee of The Jackson Laboratory. Mice were matched for age and sex in each experiment.

Tissue dissociation and cells.

Primary tumors (T) and Normal Adjacent Tissue (NAT, collected at 10cm from the tumor site) were dissociated into single-cell suspension by mechanical and enzymatic disaggregation. The samples were then filtered through 70-µm cell strainers and washed in PBS without Ca^{2+} and Mg^{2+}. Interstitial fluid was prepared from T and NAT. Briefly, tissues were minced in cold PBS and incubated on ice until cellular debris was precipitated. The supernatant was collected and spinned at 1200g at 4ºC and then used fresh or stored at -80ºC.

Peripheral blood mononuclear cells (PBMCs) were isolated via density-gradient separation with Ficoll. Total $CD8^+$ T cells were enriched by negative magnetic separation using the Miltenyi $CD8^+$ T Cell Isolation Kit Human. Human polymorphonuclear leukocytes (PMNs) enriched pellets were diluted with PBS and 3% dextran solution in HBSS was added in a 1:1 ratio. After centrifugation, the supernatant was processed for red blood cell lysis according to manufacturer's instructions (Red Blood Cell Lysis Solution, Miltenyi). Cells were either used fresh or cryopreserved according to standard procedures. Mouse neutrophils and $CD8^+$ T cells were purified according to manufacturer's instructions using anti-Ly6G and $CD8^+$ T Cell Isolation Kit Mouse, respectively (Miltenyi Biotech). Purity of isolated human and mouse immune populations were checked by FACS (>95%).

***In vitro* $CD8^+$ T Cells and Neutrophils co-culture.**

$CD8^+$ T cells activated with plate-bound anti-human CD3 (2mg/mL) and anti-human CD28 (2mg/mL) and freshly isolated neutrophils were co-cultured at different ratios in complete RPMI-1640 media.

***In vitro* GZMK and SDF1 treatments.**

HT-29 and CACO2 cells were cultured in transwells in a ratio 1:9 (CACO2:HT-29) for 16 days when trans-epithelial electrical resistance (TEER) was measured to assess the epithelial formation. Cells were treated with rHu-GZMK 10uM for 24h. For treatment of neutrophils, freshly isolated cells were treated with rHu-GZMK 10uM or SDF1 100ng/ml and analyzed at different time points as indicated in the relative figure legend.

Neutrophils migration assay.

Neutrophil migration assays were performed using a custom-made microfluidic device in PDMS composed by three parallel channels (height 200µm, length 1.1cm, width central channel 1300µm, lateral channels 700µm) separated by two lines of trapezoidal pillars. The central channel was used to confine a collagen matrix gel, the right channel for cell injection and the left channel for treatments. PBS 10X, HEPES 80mM, NaOH 1N, 8.04mg/ml rat tail Collagen type I (Corning) and RPMI culture medium (without phenol red) were mixed on ice to obtain a collagen matrix solution. The collagen solution was injected in the central channel of the microfluidic chip and incubated at 37°C to allow the complete reticulation of the collagen matrix. The collagen matrix was hydrated by injecting culture medium into the lateral channels of the device. Culture medium was added to the reservoirs to avoid evaporation and chips were incubated at 37°C. Purified human neutrophils ($3x10^6$/ml) from healthy donor were labeled with calcein-AM (green), Hoechst 33342 (blue), and DRAQ7 (red) and injected in the right channel of the microfluidic devices. The microfluidic chips were incubated at 37°C for 20 minutes on an angled surface (30 degrees) to allow neutrophils to lie at the interface between the lateral channel and the collagen gel. Chemoattractant solutions or culture medium were added to the left channel and time lapse experiments were performed using Nikon Eclipse Ti (3 fields for chip, 1 image every 2.30min for 4h at 10X magnification) to evaluate invasion of neutrophils in the collagen gel. Interstitial fluid from T and NAT of LN and HN patients, purified CXCL6 (400ng/ml), SDF-1 (100ng/ml) and IL-8

(100ng/ml) were tested. Migrating cells were counted and analyzed using ImageJ version 2.0.

Neutrophils Gelatinase assay.

The gelatinase activity was measured from control or SDF1 (100 ng/mL) treated neutrophils in triplicate, using DQ gelatin (molecular Probes, 100 µg/mL final concentration) following manufacturer's recommendations. The fluorescent output was measured every 15 min at room temperature on a GloMax Discover (Promega Instrument) microplate reader (Ex= ~495 nm, Em= ~515 nm) and corrected for background. *Clostridium histolyticum* collagenase (0.15 U/mL) was used to generate an internal calibration curve.

Neutrophils Adhesion assay.

Purified human neutrophils were labeled with calcein-AM (green), Hoechst 33342 (blue) for 20 minutes at 37°C, as previously described (Htwe et al., 2018). After washing, neutrophils were incubated in RPMI media with or without SDF1 (100 ng/mL) for 1 hour at 37°C in a 96-well on a monolayer of human micro-endothelial cells (HMEC1). After two gentle washes with PBS to remove non-adherent cells, images of three fields were recorded for each well using an EVOS microscope (EVOS FL Cell Imaging System) at 4x magnification. Adherent neutrophils were measured by counting calcein positive neutrophils using ImageJ version 2.0.0.

Multiparametric flow cytometry and sorting.

High-dimensional flow cytometry was performed on T, NAT and PB. Conjugated antibodies used for flow cytometry are shown in Table 2. Briefly, single-cell suspensions were stained fresh or after thawing in pre-warmed RPMI with 10% FBS. Cells were washed in staining

buffer (PBS with 2%FBS and 2mM EDTA) and incubated for 30min at 4°C with surface antibody cocktail in staining buffer, washed twice in staining buffer, fixed and permeabilized using Fixation/Permeabilization Solution Kit (BD, Cat # 554714), and stained intracellularly for 1h at 4°C with an antibody cocktail in the permeabilization solution. Samples were acquired using a FACSymphony A5 or a FACSCelesta BVYG equipped with FACSDiva software version 8.0.1 (all from BD Biosciences). Compensation Beads (ThermoFisher, Cat# 01-2222-41) were used to prepare single-stained controls for electronic compensation. Dead cells were excluded using Fixable Viability Stain (BD). To evaluate ex vivo cytokines production, CD8+ T cells were stimulated 3 h with PMA (20ng/mL), ionomycin (1ug/mL) and GolgiPlug protein-transport inhibitor (brefeldin A, 1:1000). CD8 T cell subsets were sorted to high purity using FACSAria III (BD Biosciences).

All the FCS files generated were analyzed and visualized with software FlowJo v10 (TreeStar). RPhenograph (Levine et al., 2015) package was used to analyze the multiparametric data from CRC patients derived T, NAT and PB. CD8 analysis was performed on 3000 events per sample after manual gating isolation of singlet, LD negative, CD45+CD3+CD8+ T cells. All samples were concatenated by the “cytof_exprsMerge function”. The number of nearest neighbors identified in the first iteration of the algorithm, K value, was set to 100. UMAP and tSNE representation were generated and visualized using FlowJo version 10 (FlowJo). Under-represented clusters (<0.5%) were discarded in subsequent analysis. The balloon plots, performed using the ggplots2 R package, generat a further metaclustering classification, interpolating MFI and frequency of each marker per cluster. MFI were multiplied to derive the integrated MFI (iMFI, rescaled to values from 0 to 100). Euclidean distance and Ward-linkage methods via cytofkit package were applied to generate hierarchical metaclustering of Phenograph generated clusters.

Bead-based multiplexed ELISA.

Multiplexed ELISA on interstitial fluid from matched T and NAT fresh tissues were performed on a Luminex 200 platform (Luminex Inc.,) using custom kits of pre-mixed antibody-coated beads (R&D System Inc., MN), which included the following analytes: CCL11_eotaxin, CCL13_MCP4, CCL17_TARC, CCL2_MCP1, CCL22_MDC, CCL26_EOTAXIN3,, CCL5_RANTES, CCL8_MCP2, CD25_IL2Ra, CX3CL1_FRACTALKINE, CXCL1_GROa, CXCL10_IP10, CXCL11_ITAC1, CXCL13_BLC_BCA1, CXCL4_PF4, CXCL5_ENA78, CXCL6_GCP2, EGF, IFNg, GMCSF, HGF, IL10, IL1ra_IL1F3, IL7, IL8_CXCL8, TRAIL, VEGFA, IL1b_IL1F2, IL5, IL6, IL-17F, IL22, IL23, TNFa, LXSAHM-01, CCL3_MIP1a, CCL4_MIP1b, FGFbasic_FGF2, GCSF, IL1a_IL1F1, IL2, IL4, IL-12 p70, IL13, IL15, IL-17/IL17A, CXCL12_SDF1 , TGFB1, TGFB3. The assay was runned based on manufacturer reccomendations. Briefly, 50ul of samples together with kit standards were added to each well in duplicate and incubated with the diluted Microparticle Cocktail at 4C, ON, on a shaker at 850rpm. Unbound soluble molecules were removed by washing the plate. The Biotin-Antibody Cocktail specific to the analytes of interest was added to each well for 1h at RT. After washing again, the Streptavidin-Phycoeriythrin conjugated was added for 30 minutes at RT. After the final washing steps, the microparticles are resuspended in kit buffer and read on a Luminex 200 platform. The outputs (pg/mL) were visualized and statistically analyzed in R upon centering and scaling using the scale function in R (SD from mean pg/mL). Data were visualized together with sample annotations using the ComplexHeatmap package. Group comparisons were visualized as boxplots using the ggpubr package and statistically analyzed by applying the non-parametric Wilcoxon signed-rank test on the values, where significance is set at p-value < 0.05. Soluble molecule data were associated with other experimental data by Spearman correlation and visualized using the corrplot package.

Immunohistochemistry (IHC) and Immunofluorescence (IF).

Slides (5um) from formalin-fixed paraffin embedded (FFPE) samples were processed for deparaffinization and rehydration. Heat-induced antigen retrieval (citrate buffer pH 6 or 9 – Thermofisher) was performed using the microwave. 3% H_2O_2 incubation was used for IHC preparation. Slides were incubated ON at 4°C with mouse anti-human CD66b antibody (BioLegend, 305102, 1:100 for both IHC and IF), anti-human CD8 (Invitrogen, 53-0008-82, 1:100), rabbit anti-human GRZMK (Invitrogen, LS-C119554-50, 1:100), rabbit anti-human E-cadherin (Abcam, Ab1416, 1:100), rat anti-human Ki67 (Ebioscience, SolA15, 1:200), rabbit anti-human EPCAM (Abcam, Ab32394, 1:400), mouse anti-human SDF1 (R&D, MAB350, 1:100), anti-human αSMA (Abcam, Ab8211, 1:200) in a blocking solution composed of 3% BSA, 5% goat serum (Invitrogen, 10000C) and 0,1% Triton in PBS. After incubation with HRP-secondary antibody (30min) and Aminoethyl Carbazole (AEC) + High Sensitivity Substrate Chromogen (Dako) (for IHC) or fluorophore conjugated antibodies (Invitrogen, A32723, goat anti-mouse 488; Invitrogen, A21247, goat anti-rabbit 647; Invitrogen, A21206, goat anti-rabbit 488; Invitrogen, A21424, goat anti-mouse 555; 1:500; Invitrogen, A21434, goat anti-rat 555, all 1:500 for 1 hour at RT), slides were mounted and acquired by using Aperio or SP8 confocal microscope (Leica), respectively.

Imaging analysis.

For quantification of GZMK signal on human colorectal cancers, sections were labeled with DAPI, anti-GZMK, anti-CD66b and Alexa Fluor 488-conjugated anti-CD8 primary antibodies and images captured with a Nikon CSU-W1 spinning disk (Nikon Europe BV) using a 40X/1.15 NA water immersion objective lens (pixel size 0.1625 x 0.1625 um^2). For the tumor area identification, previews of tissue sections on the DAPI channel were acquired with a 10X/0.3 NA dry objective lens. Inside the tumor areas, 3 to 4 regions were randomly chosen for the acquisition with a 40X/1.15 NA water immersion objective lens. Each region was made of 12 adjacent tiles covering a total area > 4 mm^2 per tissue section.

The quantification of the GZMK signal in $CD8^+$ cells was done with a custom-made Fiji macro (Schindelin et al., 2012). Briefly, images were pre-processed with denoising and background subtraction, and nuclei were segmented on the DAPI channel with the plugin StarDist (Schmidt et al., 2018), using the built-in model Versatile (fluorescent nuclei) and the intensity parameters, including “mean intensity” and “raw integrated density”, of CD8, GZMK and CD66b fluorescence were quantified in a band of thickness 1 um around each nucleus.

Statistical analysis was performed in R. $CD8^+$ cells were identified from the database by setting criteria on the intensity parameters of the specific fluorescence. Particularly, a cell was considered as positive for CD8 fluorescence channel if its “raw integrated density” was above a threshold value, set by the comparison in images of the “raw integrated density” values of cells expressing CD8 fluorescence and the “raw integrated density” of cells expressing only autofluorescence signal. Moreover, a further correction regarding the cells that had the band around the nucleus overlapping with red blood cells was applied: the ratio of the “raw integrated density” in the green channel versus the red channel was computed and only cells with a ratio green/red > 4 were considered as positive for CD8. The same procedure described above was used for the detection of CD66b positive cells (for red blood cells correction a ratio red/green > 2 was set to consider cells as CD66b positive).

Differences between two groups were calculated by unpaired, two-tailed Student's t test.

The analysis of GZMK fluorescence in CD8+ cells lying in the proximity of a neutrophil ($CD66b^+$) was performed as follows: the distance between a $CD8^+$ cell and its nearest neutrophil was computed per each field of view. Only distances below 10 um were taken into account to select pairs of CD8+ cell-neutrophil that were close to each other. The mean intensity of GZMK fluorescence in CD8+ cells close to neutrophils was compared to the GZMK mean intensity of cells close to neutrophils that were not positive for CD8.

For the quantification of SDF-1 in cells positive for αSMA, human colorectal cancers sections labeled with DAPI, anti-SDF-1, anti-αSMA and anti-EPCAM primary antibodies, and

captured with a Nikon Eclipse Ti widefield microscope (Nikon Europe BV) using a 20X/0.75 NA dry objective lens (pixel size 0.32 x 0.32 um^2). For the tumor area's identification, previews of tissue sections were acquired with a 4X/0.2 NA dry objective lens. Inside the tumor areas, regions were randomly chosen for the acquisition with 20X/0.75 NA dry objective lens for a total acquired area > 35 mm^2 per tissue section.

The quantification of SDF-1 signal in αSMA^+ cells was done with Qupath v. 0.2.3 (Bankhead et al., 2017) through a custom-made script. Briefly, cell nuclei were detected on DAPI channel with the StarDist extension and a cell expansion of 1 um was applied in order to measure fluorescence in the cytoplasmic/membrane cell's area. For the nuclei segmentation with StarDist, a custom model was trained and tested using the StarDist (2D) network in ZeroCostDL4Mic (von Chamier et al., 2021): a training dataset was generated by the manual annotation of nuclei in QuPath and labeled images were exported for the deep-learning training; the model's training and quality control assessment were performed with the StarDist (2D) notebook. For the detection of pixels positive for αSMA staining, a pixel classifier was created with the QuPath's function Create thresholder and the quantification of the positive area inside each segmented cell was added as a measurement parameter in the detection's results table. Other two-pixel classifiers in the red and far-red channels were similarly created to be used in the statistical analysis step. The expression of SDF1 was quantified as the mean intensity of the corresponding fluorescence signal in the band around the nucleus.

Statistical analysis was performed in R. Cells with an area > 3 um2 (estimated as the 10% of the area of a band created around a nucleus with area size 46 um2) of pixels classified as positive for αSMA were considered as αSMA+. Cells overlapping with red blood cells were discarded from the analysis when presenting positive area > 0 um2 for the pixel classifiers in green, red and far-red channels. Differences between two groups were calculated by unpaired, two-tailed Student's t test.

scRNA-seq.

T and NAT single cell suspension after counting was resuspended in PBS without Ca2+ and Mg2+ with 0.04% BSA. Approximately 2,000 cells/ul from each sample were used for the analysis. Briefly, every sample was loaded into one channel of Single Cell Chip A using a Chromium Single Cell 3′ v2 Reagent Kit (10x Genomics). After capture and lysis, complementary DNA was synthesized and amplified over 14 cycles according to the manufacturer's protocol (10x Genomics). Libraries were prepared from 50 ng amplified cDNA. Sequencing was performed using a NovaSeq 6000 System (Illumina). An average sequencing depth of at least 50,000 reads per cell was obtained for each sample.

scRNAseq analysis.

Single cells of samples from 11 patients (each one sequenced in the far tumor (D) and tumor site (T)) were sequenced with the 3' v3 chemistry and aligned using Cellranger count v. 3.1 on human genome hg38. Only droplets with a minimum number of unique molecular identifiers (UMI) were considered as "cells"; this was done by default by the Cellranger program.

We carried out a preliminary analysis using data of all the samples collected from the 11 patients (D and T, total of 89715 cells, data not shown). Based on total cells in each patient and the distribution of each patient's cells within groups identified by UMAP, we excluded patients 3, 8 and 6 from further analyses. Thus, 52316 cells were processed.

We used the Seurat v. 3 R package to analyse single cell data (Stuart et al., 2019). To identify cell populations or sub-populations, uniform manifold approximation and projection (UMAP) was carried out from the first principal components of the PCA based on the expression of the most variable genes, normalized using SCT transformation as suggested in Seurat documentation ("SCTransform" function, (Hafemeister and Satija, 2019). The

different clusters in UMAPs were found using nearest neighbor clustering (SNN) with the function “FindClusters” of the Seurat package. In particular:

a) For global cell population (52316 cells, fig. Supp. 6A): for UMAP, the first 30 principal components were used, and n.neighbors=60, min.dist=0.005 were set. For SNN clustering, top 20 principal components were used, with other parameters k.param=10, prune.SNN=1/10. The resolution used in the “FindClusters” function was set to 0.11.

b) For T/NK cells (8716 cells, fig. Supp. 6E): cells belonging to the “T/NK" cluster of the previous step (a) were used. The UMAP was built using the first 30 principal components and n.neighbors=40, min.dist=0.002 were set as parameters. SNN clustering was carried out as above (a) but with a resolution of 0.09 in the “FindClusters" function. Finally, cells with the first UMAP dimension (UMAP_1) > 8 were considered as outliers and removed.

c) To further characterize T CD8+ cell subtypes (fig. 6A), cells belonging to the clusters 3 and 4 were extracted from the previous step (b) and the UMAP was built using the first 15 principal components from the expression of the following cell-specific markers that presumably would identify the expected subtypes:

TRDC, PRF1, CD3D, CD3E, TRGC1, TRGC2, LEF1, IL6R, NOSIP, SELL, CCR7, CD4, BATF, TIGIT, FOXP3, TNFRSF4, CD8A, CD8B, GZMH, FGFBP2, GZMB, GNLY, ZEB2, HLA-DRB5, CX3CR1, GZMA, IFNG, CCL4, SBF2, KLRB1, LTB, IL7R, FOS, CD40LG, DPP4, IRF1, GZMM, GZMK, HLA-DRB1, HLA-DRA, CCL5, CTLA4, PDCD1, HAVCR2, LAG3, CD38, ENTPD1, LAYN, CXCL13, TOX. Other parameters for the UMAP were n.neighbors=20, min.dist=0.01. For SNN clustering, the top 15 principal components were used, with other parameters being k.param=10, prune.SNN=1/10. The resolution parameter in the “FindClusters” function was set to 0.5.

Differential expression of genes between different cell subgroups was carried out with the “FindMarkers” function in the Seurat package, using the Wilcoxon test. Thresholds were set to log2FC=0.25 and adjusted p value=0.05. For Figures 6C, 6F and Supp. 7D and E the

size of the points corresponds to the score of the specific signature; this score is defined, for each cell, as the sum of the counts of the genes belonging to a specific signature / the total sum of the counts of that cell. The points with the biggest size correspond to a score higher than 80% of the distribution of the scores of all the cells.

Cell types in the entire population (Fig. Supp 6A) were identified with the help of scMCA (single cell mouse cell atlas) package (Sun et al., 2019) and then evaluating the expression of the cell type specific markers. Pseudotime analysis (fig. 6E) was done using Monocle v. 3.

***In vivo* animal experiments.**

C57BL/6J were subcutaneously injected with 1×10^6 MC38 cells/mouse. Animals were euthanized when the tumor reached the volume of $<1cm^3$ and $>1cm^3$ to evaluate $CD8^+$ T cell and neutrophil infiltration during progression. Blood was collected and spleen and tumors harvested and dissociated to a single cell suspension. Tumors were digested (1,5 mg/mL collagenase, 0,75 mg/mL hyaluronidase and 0,1 mg/mL DNaseI) in RPMI 10% FBS at 37°C. Cell suspension was filtered through 70-µm cell strainers, washed, counted and stained for multiparametric flow cytometry.

Survival analysis methods.

We investigated the interplay between immune cytolytic activity and neutrophils infiltration on overall (OS) and disease-free survival (DFS) in the colon adenocarcinoma cohort (COAD) of The Cancer Genome Atlas (TCGA) (Weinstein et al., 2013) by performing a Kaplan-Meier analysis. The cytolytic activity index was quantified by computing the geometric mean of granzyme A and perforin expression (Rooney et al., 2015) and samples were split according to the Maximally Selected Ranks Statistics cutoff (“maxstat”, version 0.7-25). Neutrophils relative abundance was obtained by running CIBERSORTx (Rooney et

al., 2015) with a validated leukocyte gene signature matrix (LM22) on RNAseq data of COAD samples. Samples were divided into two groups according to the presence or absence of Neutrophils. Similarly, CIBERSORTx-derived gene expression signature was employed to infer the GZMK^{high} cell subtype' abundance and analyze its association with survival. First, the single-cell raw counts matrix of CD8^{+}Tem population was split into two subgroups expressing "high" or "low" values of GZMK and used it to feed CIBERSORTx, building a gene signature able to discriminate the two cell sub-populations. Secondly, we collected survival tables of TCGA-COAD and TCGA-LUAD cohorts from cBioPortal (Cerami et al., 2012) and ran Kaplan-Meier analysis stratifying patients with "high" and "low" infiltration of CD8^{+} Tem cells expressing high levels of GZMK. Survival analyses were carried out on R software version 4.0.5), by using "survival" (version 3-2-11) and "survminer" (version 0.4.9) packages.

Statistics and Data visualization.

Results were analyzed and visualized on Prism version 9.2.0 (GraphPad). All statistical analyses were conducted using GraphPad Prism version 9.2.0 or R software version 4.0.2. Differences between two groups were calculated by two-tailed Student's t test. Multiple comparisons were performed by one-way ANOVA followed by Bonferroni's multiple comparison tests. Significance was set at P values ≤ 0.05. For all figures: *, P ≤ 0.05–0.01; **, P ≤ 0.01–0.001; ***, P ≤ 0.001–0.0001; ****, P ≤ 0.0001. Unless noted in the figure legend, all data are shown as mean ± SEM. The n numbers for each experiment and the numbers of experiments are noted in the figure legends.

RESULTS

Early-stage CRC patients can be stratified based on the abundance of CD15^{high} neutrophils in the tumor microenvironment.

Despite consistent T cell infiltration, over a third of early-stage CRC patient relapse, suggesting that the anti-tumor immune response is somehow “leaky”. In this regard, characterizing the heterogeneity of immune cells within and across CRC patients would be essential to fully understand the mechanisms behind immune evasion. We profiled peripheral blood (PB), tumor (T) and normal adjacent tissue (NAT >10 cm from tumor tissue) from treatment-naïve patients (n= 42) with early-stage CRC (8th ed. AJCC stages I, II and III) undergoing surgical resection (Figure 1A). Details of all patients involved in the study are summarized in Table 1. Among the infiltrating innate immune cell populations, we observed no significant difference on dendritic cells (DCs), lower proportion of NK cells in T compared to NAT and enrichment in macrophages in T, albeit they were present at low frequencies (Figure Supp. 1A). On the other hand, neutrophils were higher in T compared to NAT both in number and frequency, as assessed by IHC (CD66b^{+}, Figure Supp. 1B) and FACS (Figures 1B) (Bronte et al., 2016; Gustafson et al., 2015), respectively. Neutrophils were identified as CD45^{+}CD11b^{+}CD15^{+} cells along with intermediate/low expression of CD33 and negative staining for HLA-DR (Figure Supp. 1C) (Gustafson et al., 2015). Analysis of CD15 expression on the surface of intratumor neutrophils (Pillay et al., 2013) revealed two distinct neutrophil populations identified by high and low levels of CD15 (referred to as CD15^{high} and CD15^{low}, respectively, Figure Supp. 1D). Despite being overall similar in NAT, CD15^{high} neutrophils were enriched in tumors in a fraction of patients (Figure Supp. 1E). Thus, based on the CD15^{high} and CD15^{low} neutrophil populations, we divided our cohort in 2 groups, namely low (LN) and high CD15^{high} (HN) neutrophils (Figures 1C and D).

In line with previous studies, our data suggested that CD15^{high} neutrophils, which predominantly infiltrate HN tumors, were in a N2-like tumor-promoting functional state (Fridlender et al., 2009). Indeed, we reported that CXCR2^{+}CXCR4^{+} double positive cells

were significantly higher in the CD15^{high} compared to the CD15^{low} neutrophil compartment (Figure 1E). Accordingly, HN tumors showed increased frequencies of N2-like tumor-promoting CXCR2^{+}CXCR4^{+} neutrophils (Figure 1F). In line with a tumor-promoting phenotype, CD15^{high} neutrophils expressed significantly higher levels of CD10, a cell surface marker that was proposed to distinguish T-cell suppressive from T-cell stimulatory neutrophils (Figure Supp. 1F) (Fridlender et al., 2009; Marini et al., 2017). Moreover, exposure of neutrophils to interstitial fluid from HN tumors induced higher production of total reactive oxygen (tROS) species (Figure Supp. 1I). Of note, the frequencies of CXCR4^{+}CXCR2^{-} neutrophils, previously defined as aged (Jaillon et al., 2020a), were comparable between CD15^{high} and CD15^{low} neutrophils as well as in HN and LN tumors (Figures Supp. 1G-H). Overall, these data revealed that CRC patients can be stratified based on the abundance of a CD15^{high} neutrophil population, which might have tumor promoting functions.

Neutrophil homeostasis in peripheral tissues results from the balance between clearance by macrophages and recirculation regulated by soluble factors (Cox et al., 1995; Grigg et al., 1991) (Buckley et al., 2006; Robertson et al., 2014; Wang et al., 2017; Woodfin et al., 2011). Indeed, frequency of macrophages anticorrelated with CD15^{high} neutrophils (Figure Supp. 1J) and was significantly lower in HN compared to LN (Figure Supp. 1K), suggesting that lower clearance in HN tumors might contribute to higher abundance of neutrophils, as previously reported (Peng et al., 2021a).

Next, we asked which was the contribution of soluble chemoattractants in determining the CD15^{high} neutrophil difference between HN and LN. This was tested on a microfluidic device that enabled measurement of the ability of freshly isolated neutrophils to migrate through a 3D collagen matrix. Surprisingly, migration was higher when neutrophils were exposed to LN compared to HN tumor interstitial fluids (Figures 1G), suggesting that the higher CD15^{high} neutrophils observed in HN compared to LN did not reflect an increase in their recruitment

into the TME. On the contrary, exposure of neutrophils to interstitial fluid from HN tumors increased expression of CD15 when compared to LN (Figures Supp. 1L), suggesting an active role of the TME in driving neutrophil function. Therefore, we profiled 49 inflammatory soluble factors in the interstitial fluid of T and NAT from CRC patients, of which 33 gave consistently measurable results (Figure 1H). While not significantly associated neither with demographics nor tumor grade, we found out that a subset of these factors was directly correlated with the frequency of $CD15^{high}$ neutrophils (Figure 1I). In particular, $CD15^{high}$ neutrophils were positively correlated with IL-17E/IL-25, which plays important roles in inflammation occurring in the gut (Caruso et al., 2009; Owyang et al., 2006; Wang et al., 2014), it has been shown to upregulate the expression of CXCL1 and CXCL10 (Senra et al., 2019) and promotes a Th2-like responses by inducing IL-4, IL-5 and IL-13 expression (Fort et al., 2001). In agreement, we found that CXCL1, CXCL10, IL-4 and IL-13 were positively correlated with $CD15^{high}$ neutrophils. Interestingly, IL-4/IL-13–stimulated neutrophils have been reported to have diminished ability to form NETs and migrate toward CXCL8 (Impellizzieri et al., 2019), suggesting that neutrophils from HN and LN might be functionally different. Enrichment of G-CSF and IL-10 in HN further supported the role of the TME in driving neutrophil polarization toward N2-like functional state (Ferrante; Fridlender et al., 2009; Gerlini et al., 2004; Gholamin et al., 2009; Joshita et al., 2009; Kyo et al., 2000; Moore et al., 2001). Among the detected factors, CXCL12/SDF-1 (from now on referred to as SDF-1) was of particular interest because of its role in neutrophil trafficking and retention at inflammatory site through its binding to CXCR4 (C et al., 2003; Filippo and Rankin, 2018), which we found to be highly expressed on $CD15^{high}$ neutrophils together with CXCR2 (Figures 1E-F). Imaging analysis confirmed that overall levels of SDF-1 were higher in HN compared to LN tumors and this was due to the combination of a larger number of αSMA^{+} cancer associated fibroblasts in HN and their higher expression of SDF-1 (Figure Supp.1M-N).

Despite surface expression of CXCR4, SDF-1 treatment did not foster neutrophil's migration *in vitro.* On the contrary, IL-8 and CXCL6, well-known chemo-attractants found to negatively correlate with CD15^{high} TANs *in vivo*, showed a pro-migratory effect (Figures Supp. 1O-P). On the other hand, neutrophils treated *in vitro* with SDF-1 showed higher frequency CXCR2^{+}CXCR4^{+} double positive cells (Figures 1J and Supp. 1Q-R), an elevated level of total ROS (Figure 1K), produced more gelatinase (Figure 1L) and adhered more to micro-endothelial cells compared to untreated controls (Figure 1M), which recapitulates, at least in part, *in vivo* observations. The slight increase in CD62L upon SDF-1 treatment supported the idea that these neutrophils were not aged (Figure Supp.1S) (Peng et al., 2021b). In summary, these results showed that tumor infiltration of CD15^{high} neutrophils distinguished between LN and HN patients, and suggested that elevated concentrations of SDF-1 changed neutrophil functional state rather than their recruitment.

The TME of CRC is infiltrated by two main CD8^{+} T cell populations identified by differential expression of CD39 and GZMK.

We next hypothesized that the different abundance of CD15^{high} neutrophils might alter the immune landscape in CRC TME and, in particular, CD8^{+} T cells. CD8^{+} T cell infiltration is unanimously accepted as a favorable prognostic marker in CRC (Chiba et al., 2004; Galon et al., 2014; Naito et al., 1998; Pagès et al., 2009b; Prall et al., 2004; Zlobec et al., 2010). However, querying the collection of non-metastatic CRC patients in the TCGA database with a cytolytic CD8^{+} T cell transcriptional signature did not result in an effective stratification (DFS, p=0.12 Figure Supp. 2A), suggesting that other factors might influence CD8^{+} T cell-mediated response and contribute to determining the clinical outcome. Therefore, we employed high dimensional flow cytometry to carry out a detailed assessment of the diversity in the CD8^{+} T cell compartment, including markers of memory and effector differentiation (CD45RA, CCR7, CD27, CD28, CD127, CX3CR1, CD161), activation (OX40, CD25, HLA-

DR), inhibitory receptors (PD1 and TIGIT), tissue residency and tumor reactivity (CD69, CD103 and CD39) and effector molecules (GZMB and GZMK). The whole list is reported in Table 2. Within the TME, $CD3^+$ T cells represented over 60% of the immune infiltrate (Figure Supp 2B) and were enriched in T compared to NAT across patients (Figure 2A). In the T, the composition of the $CD3^+$ T cells included $CD4^+$ Th (T_{Th}) cells, $CD8^+$ T cells, T regulatory (T_{Reg}) cells and T gamma-delta (Tγδ) cells, with $CD4^+$ Th and $CD8^+$ T cells being predominant (Figure 2B).

Focusing on the $CD8^+$ T cell compartment, naive (T_N identified by $CD45RA^+CCR7^+$) and central memory (T_{CM} identified by $CD45RA^-CCR7^+$) $CD8^+$ T cells were rare both in the T and in NAT compared to PB (Figure Supp. 2C). In line with previous reports (Pagès et al., 2009b), the majority (about 80%) of the infiltrating $CD8^+$ T cells were effector memory cells (T_{EM}), characterized by $CCR7^-CD45RA^-$ expression (Figures 2C and Supp. 2C). Therefore, based on differential CD27 and CD28 expression (Figure Supp. 2D), we further divided the T_{EM} into 4 phenotypically and functionally distinct subpopulations called EM_1 ($CD27^+CD28^+$), EM_2 ($CD27^+CD28^-$), EM_3 ($CD27^-CD28^-$) and EM_4 ($CD27^-CD28^+$) (Romero et al., 2007). The T_{EM3} and T_{EM4} T cell compartment, which represented effector-like and memory-like phenotypes respectively, were the most represented in T (Figures 2D). We detected low frequencies of terminally differentiated effector memory cells re-expressing CD45RA (T_{EMRA}) (Figures 2C and Supp. 2C), a phenotype associated with terminal T cell differentiation (Thome and Farber, 2015). In accordance, only a minority of $CD8^+$ T cells exhibited traits that have been previously associated with “T cell exhaustion” (McLane et al., 2019), such as co-expression of PD-1, LAG3 and TIM3 (Figure Supp. 2E).

Analysis with t-distributed stochastic neighbor embedding (tSNE) confirmed that $CD8^+$ T cell profiles were substantially different in the PB, NAT and T compartments also at single-cell level (Figures 2E-F), independently from other factors like patient's age, gender or tumor stage (Figure 2G). Based on the calculated mean fluorescence intensity (MFI) and

frequency of different markers (Figures Supp. 2F-G), we identified 20 different clusters (CL) in the $CD8^+$ T cell compartment (Figures Supp. 2G-H, CL 9 and 16 were excluded from further analyses because their frequencies were below 0.5%). A number of clusters were enriched in PB compared to NAT and T and were characterized by lack of expression of residency and activation markers (CD69, CD103 and CD39). These included: CL3 - annotated as naïve (T_N) and characterized by high levels of CD45RA, CCR7 and CD127 expression and lack of activation markers; CL1, 5 and 18 were annotated as T_{EMRA} being characterized by low CCR7 level, general lack of expression of memory markers along with high levels of the cytolytic molecule GRZB; CL6 - annotated as central memory (T_{CM}) because of the expression of several memory-associated markers (CCR7, CD127, CD27, CD28) but absence of activation and cytotoxic markers (Figures 2H and Supp. 2G). Four other clusters, instead, were enriched in T and NAT compared to PB, indicating that they were tissue-infiltrating $CD8^+$ T cell subsets. These were: CL14 and 20 mainly composed by $CD45RA^{neg}CCR7^{neg}CD28^{high}CD27^{low}CX3CR1^{pos}$ T_{EM} cells, and CL8 and 11 annotated as tissue-resident memory T cells (T_{RM}) since they displayed high expression of markers of memory (CD127 and CD28) and tissue residency (CD69 and CD103). These two tissue-infiltrating T_{RM} subsets expressed intermediate levels of CD161 and were mostly present in NAT (Figures 2F and Supp. 2G). In line with previous reports (Fergusson et al., 2015), this population appeared to be different from the $CD161^{high}$ mucosal associated invariant T cells (MAIT) in CL19, which appeared to be equally represented in all three different compartments in early CRC patients.

We next focused on CL12 and CL4, which were specifically enriched in T compared to NAT and PB, and found that they were mainly composed by a T cell population reminiscent of $CD8^+$ T_{EM} cells ($CD45RA^{low}CCR7^{low}CD28^{int}CD27^{low}CD8^+$). They showed low or intermediate levels of CD127 and CX3CR1, and differential expression of activation, exhaustion and cytotoxicity markers (Figures 2H and Supp. 2G). In particular, CL12 identified $CD8^+$ T cells

with negligible expression of GZMK and high levels of CD39, a marker that has been recently associated with tumor antigen encounters (Duhen et al., 2018; Simoni et al., 2018). On the contrary, CL4 was primarily composed by $CD8^+$ T cells characterized by lack of CD39 and elevated levels of GZMK (Figure 2I). Collectively, these data showed that CRC TME is mainly infiltrated by two distinct $CD8^+$ T_{EM} cells populations that could be distinguished based on GZMK and CD39 expression ($GrzK^{high}CD39^{neg}$ and $GrzK^{low}CD39^{pos}$).

Using a manual gating strategy, we observed that $CD8^+$ T_{EM} compartment was predominantly occupied by $CD39^{neg}$ T cells (Figure 3A). However, the expression of CD69 and CD103 in the $CD39^{neg}$ $CD8^+$ T_{EM} cells indicated that they were not derived from blood contamination (Figures Supp. 3A-B). Thus, we decided to focus our attention on the characterization of this newly emerged population of $CD8^+$ T_{EM}, annotated as CL4 and identified by the unique expression of CD39 and GZMK. Manual gating analysis showed that $GZMK^{high}$ $CD8^+$ T_{EM} cells were almost entirely $CD39^{neg}$ (Figure 3B) and were characterized by significantly lower expression of PD1 and higher GZMB when compared to their $GZMK^{low}$ counterpart (Figure 3C). When FACS-sorted based on CD39 expression and assessed for their ability to produce effector molecules, $CD39^{neg}$ $CD8^+$ T cells produced significantly higher levels of GZMK and retained the ability to produce IFNγ, TNFα and GZMβ compared to their $CD39^{high}$ counterparts (Figure Supp. 3D). In conclusion, we reported that the TME in early-CRC patients was highly infiltrated by a novel $GZMK^{high}$ $CD8^+$ T_{EM} population, which lacked CD39 expression and was distinct from the "exhausted-like" $GZMK^+$ $CD8^+$ T cell subpopulations recently reported by others (Hornburg et al., 2021; Mogilenko et al., 2021; Zhang et al., 2018).

Neutrophils modulate CRC infiltrating $CD8^+$ T cells increasing their GZMK expression.

Neutrophils influence $CD8^+$ T cell responses both in animal models and humans. However, evidences are still sparse and often contrasting (Eruslanov et al., 2014; Jaillon et al., 2020a;

Li et al., 2020; Sangaletti et al., 2014a), and even the stratification of non-metastatic CRC patients in the TCGA database was not effective when combining neutrophils to the cytolytic $CD8^+$ T cell transcriptional signature (Figure Supp. 4A, p=0.44). Thus, since CRC patients can be distinguished in two distinct subgroups based on the abundance of $CD15^{high}$ neutrophils, we sought to better understand their link with $CD8^+$ T cells. Strikingly, Spearman correlation analysis revealed that frequencies of intra-tumor $CD15^{high}$ neutrophils were positively correlated with frequencies of $GZMK^{high}$ $CD8^+$ T_{EM} cells (Figure 4A). Of note, no significant correlation was found between $CD15^{high}$ neutrophils and other $CD8^+$ T cell differentiation states or activation markers (Figures Supp. 4B and C). Since $CD4^+$ T cells were the predominant population in the $CD3^+$ T cell compartment, we sought to address their potential contribution to GZMK production. Our data showed that $CD4^+$ T cells produced low amounts of GZMK and they were neither correlated with $CD15^{high}$ neutrophils abundance nor differentially present in the TME of HN and LN (Figures Supp. 4D-F). Together, these data prompted us to investigate the specific cross-talk between $CD15^{high}$ neutrophils and $GZMK^{high}$ $CD8^+$ T_{EM} cells.

Despite comparable frequencies of $GZMK^{high}$ $CD8^+$ T_{EM} cells in NAT and PB of LN and HN patients (Figure Supp. 4G), intratumor $CD8^+$ T_{EM} cells expressed higher levels of GZMK in HN compared to LN patients, as assessed both by FACS (Figure 4B-C) and quantitative confocal microscopy (Figure 4D), confirming that the higher proportion of $GZMK^{high}$ $CD8^+$ T_{EM} cells was a tumor-specific hallmark in HN CRC patients.

Next, we investigated if neutrophils were directly involved in inducing a $GZMK^{high}$ phenotype. In CRC patients, $CD8^+$ T cells in proximity to neutrophils (distance r ≤10um) expressed higher GZMK levels compared to other cell types (Figure 4E), suggesting that the interaction might be contact-mediated and the effect direct. Indeed, co-culturing of optimally activated $CD8^+$ T cells and neutrophils derived from healthy donors demonstrated that the frequency of $GZMK^{high}$ $CD8^+$ T cells increased in presence of neutrophils and this was dependent from

both the ratio of the two cell types and their physical contact (Figure 4F). These data revealed that skewing of $CD8^+$ T cells toward a $GZMK^{high}$ phenotype resulted from direct cross-talk between neutrophils and $CD8^+$ T cells at the tumor site.

$GZMK^{high}$ $CD8^+$ T_{EM} cells promote relapse in early CRC.

To gain insight into the contribution of $GZMK^{high}$ $CD8^+$ T_{EM} and their interaction with neutrophils to the anti-tumor response in CRC, we analyzed the immune infiltrate during tumor progression in the syngeneic MC38 mouse model of colon cancer (Corbett et al., 1975). Similarly to what we observed on CRC patients, intratumor $GZMK^{high}$ $CD8^+$ T cells were positively correlated with the percentage of intratumor neutrophils (defined as $CD11b^+Ly6g^+CXCR2^+$) in the MC38 model (Figure 4G). In addition, murine $CD8^+$ T cells increased GZMK expression upon co-culturing with neutrophils from bone marrow and, to a larger extent, with tumor-infiltrating neutrophils (Figures 4H). Finally, the frequencies of intratumor $GZMK^{high}$ $CD8^+$ T cells and tumor-infiltrating neutrophils were positively correlated with the size of tumors (Figures 4I). Overall, these results confirmed the existence of a cross-talk between neutrophils and $CD8^+$ T cells resulting in an increased GZMK production, and supported a model where $GZMK^{high}$ $CD8^+$ T cells might foster tumor progression.

Emerging evidence suggests that GZMK might promote inflammation and untargeted tissue damage (Wensink et al., 2015). This has been related to the activity of GZMK as an extracellular protease and its ability to promote epithelial-to-mesenchymal transition (EMT) by remodeling components of the extracellular matrix (Bouwman et al., 2021). We tested this hypothesis *in vitro* by exposing a functional human intestinal epithelial model to the neutrophils-$CD8^+$ T cells co-culture, which we showed to induce GZMK production by $CD8^+$ T cells. Compared to each of the two cell types alone, the co-culture induced a significant decrease in E-Cadherin expression (Figure 4J), a feature associated to EMT and tumor

malignancy (MA et al., 2005; Zeisberg and Neilson, 2009). Similar results were obtained when the intestinal epithelial model was treated with purified GZMK (Figure 4K), providing evidence that release of GZMK would be sufficient to induce tissue remodeling. In line with this, HN CRC tumors showed reduced expression of E-Cadherin compared to LN (Figure 4L), suggesting that infiltration of GZMK^{high} CD8^{+} T cells could be linked to tumor progression and, possibly, contribute to relapse.

To test this link, we profiled patients in our cohort that encountered early relapse (Rel, n=5) and we observed that frequencies of CD15^{high} neutrophils and GZMK^{high} CD8^{+} T_{EM} cells in their T were comparable to HN and significantly higher than LN (Figure 5A-B). In agreement, T from relapsed patients were enriched in CL4 (mainly composed of CD8^{+} T_{EM} GZMK^{high}CD39^{neg}), while CL12 (mostly CD8^{+} T_{EFF} GZMK^{low} CD39^{pos}) was poorly represented (Figures 1H-I and 5C-E). All together, these results support the existence of a CD8-neutrophil crosstalk at the tumor site, which significantly contributes to the clinical outcome (Pagès et al., 2009b, 2009c). Thus, evaluation of the different subtypes of cytotoxic CD8^{+} T cells could provide a mean to predict tumor recurrence and patient survival.

GZMK expression correlates with a distinct transcriptional program in CD8^{+} T_{EM}.

Aiming to build a signature that could address the prognostic potential of infiltrating GZMK^{high}CD8^{+} T_{EM} cells, paired T and NAT tissues from 8 CRC (stage I-III) HN patients were profiled by single cell RNA sequencing (scRNAseq). We collected data from a total of 52316 cells. Mean reads per cell were 94183 (min:37500; max:235694) with a median number of genes detected per cell of 3779 (min:1433; max:5392). Sequencing saturation was >49% for all samples, indicating comprehensive sampling of the available transcripts. After processing and normalization, dimensionality reduction was performed with uniform manifold approximation and projection (UMAP). Shared nearest neighbor (SNN) clustering distinguished eleven different clusters that were annotated to nine broad cell lineages,

including epithelial (*EPCAM*$^+$), endothelial (*CD34*$^+$), mesenchymal (*MMP2*$^+$), B cells (*MZB1*$^+$), and T/NK (*CD3D*$^+$) (Figures Supp. 5A-C). Notwithstanding the expected variability among patients, all the annotated clusters were represented across the patients (Figure Supp. 5D). Expression analysis across all the annotated cell subtypes highlighted the T/NK cluster as the main source of *GZMK* expression in the TME of HN (Figure Supp. 5E).

Zooming in on the 8523 cells composing the T/NK cluster, we generated a new UMAP that further distinguished seven T cell subtypes (Figure Supp. 5F), among which cluster 1 was enriched in CD4$^+$ while clusters 3 and 4 were enriched in CD8$^+$ T cells (Figures Supp. 5G-H). Interestingly, *GZMK* expressing cells were mostly restricted to the CD8$^+$ clusters 3 and 4, confirming results from our phenotypic analysis. Within the CD8$^+$-enriched T cells clusters 3 and 4, we could identify eleven sub-clusters (Figures 6A). These clusters were annotated with manually curated gene-sets (Table Supp. 1) based on recent literature (Li et al., 2019). While confirming the presence of previously described putative T cell sub-populations, this analysis highlighted a remarkable transcriptional heterogeneity. The eleven clusters included minor fractions of CD4$^+$ T cells falling in the memory (CD4_T_{EM}), the *FOXP3* expressing regulatory (Treg), *CD40LG* positive mucosal associated invariant T (MAIT) and exhausted (Tex) T cell clusters. Co-expression of *TRDC* and *TRGC1* enabled discrimination of gamma delta T cells (Tγδ), while higher expression of *AKR1C3*, *CXC3R1*, *FCER1G*, *NKG7* and *PRF1* were distinctive transcriptional traits of the natural killer (NK) cluster. Naive-like T cells were assigned based on expression of *IL7R*, *CCR7* and *SELL* signature genes (Figures Supp. 6A and C). Most of the retrieved T cells fell within the CD8$^+$ T effector memory subgroup (CD8_T_{EM}, Figure Supp. 6B) characterized by low expression of *GZMB* and high expression of *GZMK* (Figures 6B), in agreement with the GZMK^{high} T cells identified by FACS analysis. On the other hand, marginal expression of genes associated to exhaustion, such as *CTLA-4*, *PDCD-1*, *CXCL13*, *HAVCR2*, *ENTPD1* and *LAYN*, and differential expression analysis revealed that CD8_T_{EM} were considerably distinct from T_{EX} cells

(Figures 6C-D). In addition, pseudotime analysis showed that CD8_T_{EM} were closer to Naive-like T cells than to terminally differentiated effectors (T_{TE}, Figures 6E and Supp. 6D). The borders between CD8_T_{EM} and the clusters labeled as CD8_T_{EM}-PD1 and HLADR_CD8 were diffuse, instead, even though hierarchical clustering of differentially expressed genes confirmed the robustness of the model. While suggesting that transcriptional gradients may contribute to T cell heterogeneity, this observation also marks intermediate states on the transition of CD8_T_{EM} toward exhaustion, for example, through the expression of genes like *PDCD1* and *CXCL13*. Functionally, pathway enrichment analysis indicated that, despite having lower cytolytic potential (padj= 0.005) compared to T_{TE} and NK cells, both CD8_T_{EM} and HLADR_CD8 (both expressing high levels of *GZMK* levels) were not impaired in proliferation (Figure Supp. 6E). In particular, CD8_T_{EM} had lower expression of lymphocyte activation and migration and higher cell-cell adhesion signatures (Figure Supp. 6E). Noteworthy, similar expression levels of effector molecule *TNFα* characterized the two subsets of CD8_T_{EM} grouped based on *GZMK* expression (Figures Supp. 6F-G), further supporting that *GZMK* expression is not associated with a dysfunctional state in CD8_T_{EM} (see also Figure Supp. 3D). Finally, querying of publicly available CRC datasets on The Cancer Genome Atlas (TCGA) using a transcriptional signature that distinguished CD8_T_{EM} highly expressing *GZMK* among those defined by scRNAseq (Figure Supp. 6I, see Methods for details), showed better prognosis in patients with lower infiltration of CD8 T_{EM} expressing high *GZMK* (Figure 6G). This signature provided a better prognostic value than the functional signature related to the cytolytic activity of $CD8^+$ T cells and it was effective in stratifying patients with other carcinoma types, like the TCGA lung adenocarcinoma (Figure Supp. 6J).

In conclusion, our functional and transcriptional characterization of the tumor immune infiltrates resulted in an effective stratification of patients and predicted a different risk of

relapse in early-stage CRC patients. Our findings have the potential to be extended to other cancer types, including lung carcinoma.

DISCUSSION

We conducted unbiased analysis of the immune cells within and across CRC patients by combining multiparametric flow cytometry with functional assays and single cell transcriptomics. Our data reveal that CRC patients can be stratified based on high (HN) or low (LN) $CD15^{high}$ neutrophil infiltration and highlight the emergence of a novel $GZMK^{high}$ $CD8^{+}$ T_{EM} cell population as a hallmark of HN tumors (Figure 7). The predominance of the $GZMK^{high}$ $CD8^{+}$ T_{EM} cells in CRC patients that encountered an early relapse and the correlation with poor clinical outcome on an independent cohort underline the clinical relevance of the cell population and its therapeutic potential.

Despite several attempts to profile the immune cell content of CRC (Pagès et al., 2009b; Zhang et al., 2018), the prognostic value of immune cell infiltration is not conclusive, especially regarding neutrophils (Shaul and Fridlender, 2019). We provide evidence that neutrophils are the most abundant innate cell population infiltrating the tumor bed in early-stage CRC. Pro-tumorigenic roles have been ascribed to neutrophils across different tumor types (Galdiero et al., 2018; Mantovani et al., 2011), including promotion of angiogenesis and immunosuppression (Bronte et al., 2016; Coffelt et al., 2016), but it is not clear whether distinct neutrophil populations gather the tumor or if phenotypic changes occur inside the TME. The presented analysis of the canonical neutrophil marker CD15 (Pillay et al., 2013) reveals phenotypic changes specifically in the population associated with tumors. Intratumor neutrophils from LN patients have significantly lower expression of CD15 compared to healthy- and patient-derived peripheral neutrophils. Conversely, neutrophils infiltrating the TME of HN patients maintain high levels of CD15 expression, are enriched in

$CXCR2^+CXCR4^+$ cells and tend to produce more ROS, in line with a so called N2-like tumor-promoting functional state (Fridlender et al., 2009). Interestingly, treatment of freshly isolated neutrophils with interstitial fluid from HN tumors led to higher CD15 expression but, differently from interstitial fluid from LN, did not foster migration, suggesting that specific soluble factors in the TME might polarize neutrophils toward a $CD15^{high}$ state rather than increasing their recruitment. Neutrophils are recruited early to the site of inflammation, where their not-specific antimicrobial functions can endanger the integrity of the host tissue in case of inappropriate retention (Nathan and Ding, 2010; Rock et al., 2010). To avoid it, in addition to clearance by macrophages-directed efferocytosis (Cox et al., 1995; Grigg et al., 1991), neutrophils redistribute into surrounding tissues or vasculature through reverse migration (Buckley et al., 2006; Robertson et al., 2014; Wang et al., 2017; Woodfin et al., 2011), which has been suggested to follow desensitization to local chemotactic gradients (Holmes et al., 2012; Powell et al., 2017; Robertson et al., 2014). In this regard, SDF-1/CXCR4 signaling plays an important role in the retention of neutrophils at inflammatory sites (Isles et al., 2019). Since clearance by resident tissue macrophages has been associated with reduction of CD15 surface expression on neutrophils (Shelley et al., 2018), we can speculate that the enrichment of $CD15^{high}$ neutrophils in HN tumors might be due, at least in part, to inefficient clearance, as previously demonstrated in other tumor contexts (Peng et al., 2021a). On the other hand, our analysis of differentially abundant soluble factors pointed at SDF-1, which correlates *in vivo* with frequency of $CD15^{high}$ neutrophils and, *in vitro,* increased adhesion and release of TME's remodeling factors rather than driving chemotaxis. Overall, these data suggest that elevated concentrations of SDF-1 may promote the reprogramming of neutrophils infiltrating HN tumors. Future studies would need to address if this may lead to an increased retention of neutrophils in the tumor, altering their homeostatic turnover.

Neutrophils influence $CD8^+$ T cell-mediated responses both in animal models and humans. However, results are still sparse and often conflicting (Eruslanov et al., 2014; Governa et

al., 2017; Jaillon et al., 2020b; Li et al., 2020; Sangaletti et al., 2014b; Singhal et al., 2016; Wu et al., 2014). We provide evidence that direct interaction between neutrophils and $CD8^+$ T cells leads to increased GZMK expression, both *in vivo* and *in vitro*. When compared to their $GZMK^{neg}$ counterpart, $GZMK^{high}$ $CD8^+$ T_{EM} cells harbor significantly lower PD-1 expression and higher levels of Gzmβ, posing these cells as distinct compared to the $GZMK^+$ $CD8^+$ T cell subpopulations previously identified and described as pre-dysfunctional (Guo et al., 2018; Hornburg et al., 2021; Mogilenko et al., 2021; Zhang et al., 2018). Moreover, their abundance in the TME is independent from tumor stage and aging ($P > 0.05$, ANOVA) (Mogilenko et al., 2021). Of particular relevance, $GZMK^{high}$ $CD8^+$ T_{EM} cells lack CD39 expression, which might suggest that they are contributing to a general inflammatory state (Chiou et al., 2021; Corridoni et al., 2020) rather than targeting specifically tumor associated antigens (Scheper et al., 2019; Simoni et al., 2018). Thus, the high proportion of $GZMK^{high}$ $CD39^{neg}$ $CD8^+$ T_{EM} cells found in HN CRC patients raises the possibility that they could be involved in a cross-reactive response, favoring cancer progression. The alloreactive-like transcriptional signature associated with $GZMK^{high}$ $CD8^+$ T_{EM} favors this model, with upregulation of CRTAM and CCL5/RANTES promoting retention of activated $GZMK^{high}$ $CD8^+$ T_{EM} cells (Arase et al., 2005; Cortez et al., 2014; Galibert et al., 2005) and recruitment and immunosuppressive polarization of myeloid cells (Aldinucci et al., 2020; Schall et al., 1990; Seo et al., 2020; Swanson et al., 2002; Walzer et al., 2003), respectively. However, only future analysis of the TCR clonality of $GZMK^{high}$ $CD39^{neg}$ $CD8^+$ T_{EM} will allow testing this hypothesis unequivocally.

We show that, at least in HN CRC patients, $CD8^+$ T_{EM} are the main source of GZMK. Further studies are needed to determine whether $CD8^+$ T cells recruited to the tumor are skewed by neutrophils toward a $GZMK^{high}$ phenotype or, conversely, the presence of $GZMK^{high}$ $CD8^+$ T cells impacts neutrophil homeostasis and function, as already proved for other immune cell populations (Joeckel et al., 2011). Regardless of the order of events, GZMK might play a

pathogenic role by contributing to inflammation and tissue damage. This relates to its non-cytotoxic activity as extracellular protease and its ability to promote cellular migration and epithelial-to-mesenchymal transition (EMT) by remodeling components of the extracellular matrix (Joeckel and Bird, 2014; Wensink et al., 2015), facilitating cytokine responses (M et al., 2016; Wensink et al., 2014) In line with this, HN tumors showed reduced expression of E-Cadherin compared to LN, which is a feature of EMT and tumor malignancy. Along with tangible therapeutic implications for future interventions, high expression of GZMK by infiltrating $CD8^+$ T_{EM} in response to $CD15^{high}$ neutrophils links them to tumor progression rather than anti-tumor function, suggesting that integrating functional information resulting from the crosstalk of $CD8^+$ T cells with different components of the immune contexture and the stroma might implement the prognostic value of current biomarkers, which are mostly based on limited phenotypic or transcriptional $CD8^+$ T cell characteristics.

LIST OF ABBREVIATIONS

Colorectal Cancer (CRC)

Granzyme K (GZMK)

Stromal Cell-Derived Factor 1 (CXCL12/SDF-1)

Tumor-Node-Metastasis (TNM)

Tumor Microenvironment (TME)

Tumor Infiltrating Lymphocytes (Tils)

Peripheral Blood (PB)

Tumor (T)

Normal Adjacent Tissue (NAT)

Dendritic Cells (DCs)

High $CD15^{high}$ Neutrophils (HN)

Low $CD15^{high}$ Neutrophils (LN)

Total Reactive Oxygen Species (TRos)

$CD4^{+}$ Th Cells (T_{th})

T Regulatory Cells (T_{reg})

T Gamma-Delta Cells (Tγδ)

Naive T Cells (T_N)

Central Memory T Cells (T_{CM})

Effector Memory T Cells (T_{EM})

Terminally Differentiated Effector Memory T Cells Re-Expressing CD45RA (T_{EMRA})

T-Distributed Stochastic Neighbor Embedding (Tsne)

Mean Fluorescence Intensity (MFI)

Clusters (CL)

Tissue-Resident Memory T Cells (T_{RM})

Mucosal Associated Invariant T Cells (MAIT)

Epithelial-To-Mesenchymal Transition (EMT)

Early Relapse (Rel)

Single Cell RNA Sequencing (scRNAseq)

Uniform Manifold Approximation And Projection (UMAP)

Shared Nearest Neighbor (SNN)

Exhausted T Cells (T_{ex})

Natural Killer Cells (NK)

Terminally Differentiated Effectors Cells (T_{TE})

The Cancer Genome Atlas (TCGA)

Peripheral Blood Mononuclear Cells (PBMCs)

Polymorphonuclear Leukocytes (PMNs)

Trans-Epithelial Electrical Resistance (TEER)

CACO2

Human Micro-Endothelial Cells (HMEC1)

Integrated MFI (Imfi)

Immunohistochemistry (IHC)

Formalin-Fixed Paraffin Embedded (FFPE)

Unique Molecular Identifiers (UMI)

Single Cell Mouse Cell Atlas (SCMCA)

Overall Survival (OS)

Disease-Free Survival (DFS)

Colon Adenocarcinoma Cohort (COAD)

FIGURE LEGENDS

Figure 1. Early-stage colorectal cancer (CRC) patients can be stratified based on the abundance of CD15^{high} neutrophils in the tumor microenvironment (TME). A. Study design. Immune analysis of Tumor tissues (T), Normal Adjacent Tissues (NAT) (10 cm from T) and Peripheral Blood (PB) of early-stage CRC patients (n=42). **B.** Frequency of Tumor Associated Neutrophils (TAN) defined as CD45^{+} CD11b^{+} HLADR^{-} CD56^{-} CD66b^{+} within CD45^{+} immune compartment in T compared with NAT samples in CRC patients (total n=22). Bars indicate 90-10 percentile. **C.** Density plot of CD15 expression within TAN defining a bimodal distribution in the CRC cohort of patients. Patients were clustered as low (LN) or high (HN) based on CD15 expression. **D.** Frequency of CD15^{high} neutrophils in LN and HN patients. Bars indicate 90-10 percentile. **E.** Bar plot of CXCR4^{+} CXCR2^{+} cell frequency within CD15^{low} and CD15^{high} neutrophils. **F.** Bar plot of CXCR4^{+} CXCR2^{+} cell frequency within LN and HN neutrophils. **G.** Representative brightfield images of neutrophils migration assay on

the microfluidic device. Interstitial Fluid of T and NAT from HN and LN patients were tested. Bar plot of number of migrating neutrophils on chip upon stimulation with IF from HN (N=6) and LN (N=5) patients. Scale bar 100mm. **H.** Heatmap showing the abundance of soluble molecules in from IF of T (tumor, light blue) and NAT (far-tumor, pink) tissues. For all the samples, tissue type, tumor grade, sex, age and Microsatellite Instability (MSI) status are indicated. **I.** Spearman correlation between soluble molecules in T IF and $CD15^{high}$ neutrophils. **J.** Frequency of $CXCR4^{+}$ $CXCR2^{+}$ in Healthy Donors (HD) fresh-isolated neutrophils treated with SDF-1. **K.** Bar plot representation of total ROS (tRos) MFI within SDF-1 treated-neutrophils from healthy donors (HD). **L.** Gelatinase activity assay on neutrophils freshly-isolated from HD and treated with SDF-1 for the indicated time. Bar plot representation of fold change over untreated control. **M.** Quantification of adherent neutrophils on endothelial cells after 2 hours treatment with SDF-1.

2. High-dimensional single cell analysis of $CD8^{+}$ T cells identifies a tumor-specific $CD39^{-}$ $GzmK^{+}$ population enriched in early-stage CRC. A. CD3 Lymphocytes frequency within $CD45^{+}$ in T compared with NAT. Bars indicate 90-10 percentile. **B-D.** UMAP Representation of concatenated $CD3^{+}$ T cells from T sample with bar plot quantification. **B.** Cytotoxic $CD8^{+}$ T cells (blue), Th $CD4^{+}$ (purple), Treg $CD4^{+}$, $CD25^{+}$, $CD127^{-}$ (pink) and γδT cells (orange) distribution. **C.** Naive T cells (light green), central memory T cells (dark green), effector memory T cells (dark blue), effector memory cells re-expressing CD45RA (light blue) distribution. **D.** effector memory T cells subtypes (EM_{1-4}), see main text for description. **E.** tSNE representation of Phenograph algoritm identified 18 clusters of concatenated CD8+ T cells (3,000 cells/sample) from T (n=34), NAT (n=17) and PB (n = 22) samples from CRC patients. **F.** Cluster representation in different tissues: PB, NAT, T. **G.** Hierarchical metaclustering of CRC patients derived samples based on Phenograph identified clusters

using Ward's minimum variance method. For all the samples, sample type, grade, sex, age are indicated. **H.** Balloon plot of the average expression levels and expression frequencies of marker of differentiation, residency, memory, cytotoxicity, tumor reactivity, activation and exhaustion in clusters identified in (E-F) and enriched in T. **I.** Bar plot of $CD39^+$ $GzmK^-$ and $CD39^-$ $GzmK^+$ frequency within Cluster 4 and Cluster 12, respectively. *, $P < 0.05$**, $P < 0.01$ ***, $P < 0.001$.

Figure 3. Phenotypic characterization of the $CD39^-$ $GZMK^+$ $CD8^+$ T cell population infiltrating early stage CRC patients.

A. Representative contour plot and quantification of $CD39^-$ and $CD39^+$ frequency within $CD8^+$ T_{EM}. **B.** Representative contour plot and quantification of $CD39^-$ and $CD39^+$ frequency within $GZMK^+$ $CD8^+$ T_{EM}. **C.** Bar plot showing frequencies of the indicated markers within $GZMK^+$ (dark blue) and $GZMK^-$ (light gray) $CD8^+$ T_{EM}. ***; $P < 0.001$.

Figure 4. $CD15^{high}$ neutrophil accumulation induces GZMK expression in $CD8^+$ T cells, leading to tissue remodeling. A. Spearman correlation between $CD15^{high}$ neutrophils and expression of GZMK within $CD8^+$ T_{EM} cells on each individual CRC patient (black dots). Shadowed area (dark blue) includes High Neutrophils (HN) patients. **B.** Representative contour plots of $CD8^+$ T_{EM} GZMK expression in Low Neutrophils (LN, turquoise) and High Neutrophils (HN, purple) patients. Percentage of $CD15^{high}$ neutrophils are indicated. **C.** Box plot of $GZMK^+$ $CD8^+$ T_{EM} cells in HN (n=10) and LN (n=6) patients. **D-E**. Representative confocal images and quantification of GZMK Mean Fluorescence Intensity (MFI) in (**D**) HN (n=9) and LN (n=6) tumors. Scale bar 10mm. **(E)** $CD8^+$ and $CD8^-$ cells within a 10 μm

distance from neutrophils ($CD66b^+$). Scale bar 5 mm. (See Methods for details). **F.** Co-culture of CD8 and neutrophils isolated from peripheral blood of healthy donors. Bar plot of frequency of $GZMK^+$ within $CD8^+$ T cells after co-culture, in plate or in Transwell (TRANS) with neutrophils derived from healthy donors (HD). Neutrophils-to-$CD8^+$ T cells ratio in the co-cultures are indicated. **G-I.** MC38 tumor bearing mice were euthanized and neutrophils content and GZMK expression in tumors evaluated by flow cytometry. **(G)** Spearman correlation between frequency of $GZMK^+$ within $CD8^+$ T cells and frequency of $CXCR2^+$ cells within $Ly6G^+$ $CD11b^+$ neutrophils in MC38 tumors. **(H)** Bar plot of $GZMK^+$ MFI within activated $CD8^+$ T cells co-cultured in a 1:1 ratio with neutrophils isolated from Bone Marrow (BM) or tumors from MC38 tumor bearing mice. **(I)** Spearman correlation between frequency of $GZMK^+$ within $CD8^+$ T cells in MC38 tumors and tumor size (mm^3). **J-K.** Representative confocal images and quantification of CACO-HT29 cells co-cultured on transwells and then (**J**) exposed for 24h to neutrophils (NEU), $CD8^+$ T cells (CD8) or neutrophils/$CD8^+$ T cells in 1:1 ratio (1:1), or (**K**) treated for 24h with control media (CTR) or active recombinant human GzmK (GzmK). Staining for DAPI in blue and E-Cadherin in green. Scale bar 100µm. **L.** Representatives confocal images and quantification of FFPE sections of LN and HN tumors from CRC patients. Staining for DAPI in blue and E-Cadherin in green. Scale bar 50µm.

Figure 5. CRC patients encountering early relapse have high infiltration of $GZMK^+$ $CD8^+$ T effector memory (T_{EM}) cells. A. Representative histogram and quantification of GZMK frequency within T_{EM} $CD8^+$ Tumor Infiltrating Lymphocytes (TILs) in High Neutrophils (HN), Low Neutrophils (LN) and Early-Relapsed (Rel) patients. **B.** Representative histogram and quantification of $CD15^{high}$ neutrophils in HN, LN, Rel tumors. **C.** Bar plot of Cluster 12 (CL12, green) and Cluster 4 (CL4, purple) in HN, LN and Rel tumors. **D.** Representative

contour plot of GZMK and CD39 expression within $CD8^+$ T_{EM} in in HN, LN and Rel tumors. **E.** Box plot of $CD39^-$ $GZMK^+$ within $CD8^+$ T_{EM} cells in HN, LN and Rel tumors.

Figure 6. Single-cell RNAseq analysis on CRC patients defines an effective prognostic signature based on $CD8^+$ $GZMK^+$ T_{EM} abundance. UMAP projection of $CD8^+$ T cells. Colors represent different T cell subtypes. Key genes used for manual annotation are indicated. **B.** Violin plots showing the normalized expression of GZMB and GZMK in the different T $CD8^+$ cell subtypes. **C.** UMAP showing the expression of CD8_Tem and CD8_Tex signatures for each cell. Cells are color coded based on SNN-clustering, dot size reflects the expression of the signature for each cell. **D.** Violin plots showing the expression of genes associated with T-cell exhaustion within the indicated cell subtype. **E.** Monocle pseudotime analysis of the $CD8^+$ T cell subtypes. The pseudotime line connects the indicated T cell subtypes identified in (A). **F.** UMAP as in (C) showing the expression of Cytolysis and Allograft rejection signatures. **G.** Kaplan-Meyer analysis of the association of $GZMK^{high}$ cell subtype' abundance with disease free survival (DFS) on the TCGA-COAD cohort (see Methods and Table Suppl. 1 for details). High signature in red, low in turquoise.

Figure 7. Simplified graphical summary. Abundance of $CD15^{high}$ neutrophils in the tumor identifies High (HN) and Low (LN) neutrophil subgroups of early-stage CRC patients (**1**). Elevated SDF-1 levels in HN tumors reshape the functional state of infiltrating neutrophils (**2**), promoting their retention (**3**) and activation (**4**) at the tumor site. The interaction with neutrophils (**5**) pushes effector memory $CD8^+$ T (T_{EM}) to produce high levels of Granzyme K (GZMK) (**6**), which in turn remodels the tumor microenvironment (TME) fostering EMT (**7**). A $GZMK^{high}$ T_{EM} transcriptional signature effectively stratify non-metastatic early CRC patients and predict poor clinical outcome.

SUPPLEMENTARY INFORMATIONS

Availability of supporting data

All data supporting the findings of this study are available from the corresponding authors upon reasonable request.

Authors' contributions

LN and TM contribute equally. Correspondence to Teresa Manzo (teresa.manzo@ieo.it). ST, TM and LN designed. ST, CC, CS, MB, CBNL, GR performed experiments; ST, CC, CS, OC, MB, DC, CS, CBNL, ADM, GB, VL, PP, GR, MS, SC, LN and TM analyzed data; MR, EL provided critical expertise and resources; WP, SPR, NF and UFR obtained tissues from patients; LN and TM coordinated the whole study; ST, CC, LN and TM wrote and edited the manuscript. All the authors discussed and read the manuscript.

Acknowledgements

LN and TM dedicate this work to Leonardo and Lara Nezi. We thank Phil Greenberg for thoughtful discussion. The authors thanks IEO Biobank B4Me, all the nurses, the Molecular Pathology, Imaging and Genomic Unit at the Department of Experimental Oncology at IEO, patients abd their families involved in the study.

BIBLIOGRAPHY

Aldinucci, D., Borghese, C., and Casagrande, N. (2020). The CCL5/CCR5 Axis in Cancer Progression. Cancers 2020, Vol. 12, Page 1765 *12*, 1765.

American Cancer Society (2018). Colorectal Cancer Screening Test Use* (%), Adults 50 Years and Older by State, 2018.

Arase, N., Takeuchi, A., Unno, M., Hirano, S., Yokosuka, T., Arase, H., and Saito, T. (2005).

Heterotypic interaction of CRTAM with Necl2 induces cell adhesion on activated NK cells and CD8+ T cells. Int. Immunol. *17*, 1227–1237.

Ballesteros, I., Rubio-Ponce, A., Genua, M., Lusito, E., Kwok, I., Fernández-Calvo, G., Khoyratty, T.E., van Grinsven, E., González-Hernández, S., Nicolás-Ávila, J.Á., et al. (2020). Co-option of Neutrophil Fates by Tissue Environments. Cell *183*, 1282-1297.e18.

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. Reports 2017 71 *7*, 1–7.

Berry, R.S., Xiong, M.J., Greenbaum, A., Mortaji, P., Nofchissey, R.A., Schultz, F., Martinez, C., Luo, L., Morris, K.T., and Hanson, J.A. (2017). High levels of tumor-associated neutrophils are associated with improved overall survival in patients with stage II colorectal cancer. PLoS One *12*.

Binnewies, M., Roberts, E.W., Kersten, K., Chan, V., Fearon, D.F., Merad, M., Coussens, L.M., Gabrilovich, D.I., Ostrand-Rosenberg, S., Hedrick, C.C., et al. (2018). Understanding the tumor immune microenvironment (TIME) for effective therapy. Nat. Med. *24*, 541–550.

Bouwman, A.C., van Daalen, K.R., Crnko, S., ten Broeke, T., and Bovenschen, N. (2021). Intracellular and Extracellular Roles of Granzyme K. Front. Immunol. *12*.

Bronte, V., Brandau, S., Chen, S.-H., Colombo, M.P., Frey, A.B., Greten, T.F., Mandruzzato, S., Murray, P.J., Ochoa, A., Ostrand-Rosenberg, S., et al. (2016). Recommendations for myeloid-derived suppressor cell nomenclature and characterization standards. Nat. Commun. 2016 71 *7*, 1–10.

Buckley, C.D., Ross, E.A., McGettrick, H.M., Osborne, C.E., Haworth, O., Schmutz, C., Stone, P.C.W., Salmon, M., Matharu, N.M., Vohra, R.K., et al. (2006). Identification of a phenotypically and functionally distinct population of long-lived neutrophils in a model of reverse endothelial migration. J. Leukoc. Biol. *79*, 303–311.

C, M., PC, B., G, B., JC, G.-R., TJ, W., and SM, R. (2003). Chemokines acting via CXCR2

and CXCR4 control the release of neutrophils from the bone marrow and their return following senescence. Immunity *19*, 583–593.

Carus, A., Ladekarl, M., Hager, H., Pilegaard, H., Nielsen, P.S., and Donskov, F. (2013). Tumor-associated neutrophils and macrophages in non-small cell lung cancer: No immediate impact on patient outcome. Lung Cancer *81*, 130–137.

Caruso, R., Sarra, M., Stolfi, C., Rizzo, A., Fina, D., Fantini, M.C., Pallone, F., MacDonald, T.T., and Monteleone, G. (2009). Interleukin-25 Inhibits Interleukin-12 Production and Th1 Cell-Driven Inflammation in the Gut. Gastroenterology *136*, 2270–2279.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discov. *2*, 401–404.

von Chamier, L., Laine, R.F., Jukkala, J., Spahn, C., Krentzel, D., Nehme, E., Lerche, M., Hernández-Pérez, S., Mattila, P.K., Karinou, E., et al. (2021). Democratising deep learning for microscopy with ZeroCostDL4Mic. Nat. Commun. 2021 121 *12*, 1–18.

Chiba, T., Ohtani, H., Mizoi, T., Naito, Y., Sato, E., Nagura, H., Ohuchi, A., Ohuchi, K., Shiiba, K., Kurokawa, Y., et al. (2004). Intraepithelial CD8+ T-cell-count becomes a prognostic factor after a longer follow-up period in human colorectal carcinoma: Possible association with suppression of micrometastasis. Br. J. Cancer *91*, 1711–1717.

Chiou, S.H., Tseng, D., Reuben, A., Mallajosyula, V., Molina, I.S., Conley, S., Wilhelmy, J., McSween, A.M., Yang, X., Nishimiya, D., et al. (2021). Global analysis of shared T cell specificities in human non-small cell lung cancer enables HLA inference and antigen discovery. Immunity *54*, 586-602.e8.

Coffelt, S.B., Wellenstein, M.D., and De Visser, K.E. (2016). Neutrophils in cancer: Neutral no more. Nat. Rev. Cancer *16*, 431–446.

Corbett, T.H., Griswold Jr., D.P., Roberts, B.J., Peckham, J.C., and Schabel Jr., F.M.

(1975). Tumor induction relationships in development of transplantable cancers of the colon in mice for chemotherapy assays, with a note on carcinogen structure. Cancer Res *35*, 2434–2439.

Corridoni, D., Antanaviciute, A., Gupta, T., Fawkner-Corbett, D., Aulicino, A., Jagielowicz, M., Parikh, K., Repapi, E., Taylor, S., Ishikawa, D., et al. (2020). Single-cell atlas of colonic CD8+ T cells in ulcerative colitis. Nat. Med. 2020 269 *26*, 1480–1490.

Cortez, V.S., Cervantes-Barragan, L., Song, C., Gilfillan, S., McDonald, K.G., Tussiwand, R., Edelson, B.T., Murakami, Y., Murphy, K.M., Newberry, R.D., et al. (2014). CRTAM controls residency of gut CD4+CD8+ T cells in the steady state and maintenance of gut CD4+ Th17 during parasitic infection. J. Exp. Med. *211*, 623–633.

Cox, G., Crossley, J., and Xing, Z. (1995). Macrophage engulfment of apoptotic neutrophils contributes to the resolution of acute pulmonary inflammation in vivo. Am. J. Respir. Cell Mol. Biol. *12*, 232–237.

Duhen, T., Duhen, R., Montler, R., Moses, J., Moudgil, T., De Miranda, N.F., Goodall, C.P., Blair, T.C., Fox, B.A., McDermott, J.E., et al. (2018). Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. Nat. Commun. *9*.

Eruslanov, E.B., Bhojnagarwala, P.S., Quatromoni, J.G., Stephen, T.L., Ranganathan, A., Deshpande, C., Akimova, T., Vachani, A., Litzky, L., Hancock, W.W., et al. (2014). Tumor-associated neutrophils stimulate T cell responses in early-stage human lung cancer. J. Clin. Invest. *124*, 5466–5480.

Fergusson, J.R., Hühn, M.H., Swadling, L., Walker, L.J., Kurioka, A., Llibre, A., Bertoletti, A., Holländer, G., Newell, E.W., Davis, M.M., et al. (2015). CD161intCD8+ T cells: a novel population of highly functional, memory CD8+ T cells enriched within the gut. Mucosal Immunol. 2016 92 *9*, 401–413.

Ferrante, A. CD8+ T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer.

Filippo, K. De, and Rankin, S.M. (2018). CXCR4, the master regulator of neutrophil trafficking in homeostasis and disease. Eur. J. Clin. Invest. *48*.

Fort, M.M., Cheung, J., Yen, D., Li, J., Zurawski, S.M., Lo, S., Menon, S., Clifford, T., Hunte, B., Lesley, R., et al. (2001). IL-25 Induces IL-4, IL-5, and IL-13 and Th2-Associated Pathologies In Vivo. Immunity *15*, 985–995.

Fountzilas, E., Kotoula, V., Tikas, I., Manousou, K., Papadopoulou, K., Poulios, C., Karavasilis, V., Efstratiou, I., Pectasides, D., Papaparaskeva, K., et al. (2018). Prognostic significance of tumor genotypes and CD8+ infiltrates in stage I-III colorectal cancer. Oncotarget *9*, 35623–35638.

Fridlender, Z.G., Sun, J., Kim, S., Kapoor, V., Cheng, G., Ling, L., Worthen, G.S., and Albelda, S.M. (2009). Polarization of Tumor-Associated Neutrophil Phenotype by TGF-β: “N1” versus “N2” TAN. Cancer Cell *16*, 183–194.

Galdiero, M.R., Varricchi, G., Loffredo, S., Mantovani, A., and Marone, G. (2018). Roles of neutrophils in cancer growth and progression. J. Leukoc. Biol. *103*, 457–464.

Galibert, L., Diemer, G.S., Liu, Z., Johnson, R.S., Smith, J.L., Walzer, T., Comeau, M.R., Rauch, C.T., Wolfson, M.F., Sorensen, R.A., et al. (2005). Nectin-like Protein 2 Defines a Subset of T-cell Zone Dendritic Cells and Is a Ligand for Class-I-restricted T-cell-associated Molecule * ♦. J. Biol. Chem. *280*, 21955–21964.

Galon, J., Mlecnik, B., Bindea, G., Angeli, H., Lagorce, C., Marincola, F., Fox, B., and Pagès, F. (2014). Towards the introduction of the “Immunoscore” in the classification of malignant tumours. J. Pathol. *232*, 199–209.

Gentles, A.J., Newman, A.M., Liu, C.L., Bratman, S. V., Feng, W., Kim, D., Nair, V.S., Xu, Y., Khuong, A., Hoang, C.D., et al. (2015). The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat. Med. *21*, 938–945.

Gerlini, G., Tun-Kyi, A., Dudli, C., Burg, G., Pimpinelli, N., and Nestle, F.O. (2004). Metastatic Melanoma Secreted IL-10 Down-Regulates CD1 Molecules on Dendritic Cells in

Metastatic Tumor Lesions. Am. J. Pathol. *165*, 1853.

Gholamin, M., Moaven, O., Memar, B., Farshchian, M., Naseh, H., Malekzadeh, R., Sotoudeh, M., Rajabi-Mashhadi, M.T., Forghani, M.N., Farrokhi, F., et al. (2009). Overexpression and interactions of interleukin-10, transforming growth factor β, and vascular endothelial growth factor in esophageal squamous cell carcinoma. World J. Surg. *33*, 1439–1445.

Governa, V., Trella, E., Mele, V., Tornillo, L., Amicarella, F., Cremonesi, E., Muraro, M.G., Xu, H., Droeser, R., Däster, S.R., et al. (2017). The Interplay Between Neutrophils and CD8(+) T Cells Improves Survival in Human Colorectal Cancer. Clin Cancer Res *23*, 3847–3858.

Grigg, J.M., Silverman, M., Savill, J.S., Sarraf, C., and Haslett, C. (1991). Neutrophil apoptosis and clearance from neonatal lungs. Lancet *338*, 720–722.

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.

Gustafson, M.P., Lin, Y., Maas, M.L., Keulen, V.P. Van, Johnston, P.B., Peikert, T., Gastineau, D.A., and Dietz, A.B. (2015). A Method for Identification and Analysis of Non-Overlapping Myeloid Immunophenotypes in Humans. PLoS One *10*, e0121546.

Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019 201 *20*, 1–15.

Holmes, G.R., Anderson, S.R., Dixon, G., Robertson, A.L., Reyes-Aldasoro, C.C., Billings, S.A., Renshaw, S.A., and Kadirkamanathan, V. (2012). Repelled from the wound, or randomly dispersed? Reverse migration behaviour of neutrophils characterized by dynamic modelling. J. R. Soc. Interface *9*, 3229–3239.

Hornburg, M., Desbois, M., Lu, S., Guan, Y., Lo, A.A., Kaufman, S., Elrod, A., Lotstein, A.,

DesRochers, T.M., Munoz-Rodriguez, J.L., et al. (2021). Single-cell dissection of cellular components and interactions shaping the tumor immune phenotypes in ovarian cancer. Cancer Cell *39*, 928-944.e6.

Htwe, S.S., Wake, H., Liu, K., Teshigawara, K., Stonestreet, B.S., Lim, Y.P., and Nishibori, M. (2018). Inter-a inhibitor proteins maintain neutrophils in a resting state by regulating shape and reducing ROS production. Blood Adv. *2*, 1923–1934.

Ilie, M., Hofman, V., Long-Mira, E., Selva, E., Vignaud, J.-M., Padovani, B., Mouroux, J., Marquette, C.-H., and Hofman, P. (2014). “Sentinel” Circulating Tumor Cells Allow Early Diagnosis of Lung Cancer in Patients with Chronic Obstructive Pulmonary Disease. PLoS One *9*, 111597.

Impellizzieri, D., Ridder, F., Raeber, M.E., Egholm, C., Woytschak, J., Kolios, A.G.A., Legler, D.F., and Boyman, O. (2019). IL-4 receptor engagement in human neutrophils impairs their migration and extracellular trap formation. J. Allergy Clin. Immunol. *144*, 267-279.e4.

Isles, H.M., Herman, K.D., Robertson, A.L., Loynes, C.A., Prince, L.R., Elks, P.M., and Renshaw, S.A. (2019). The CXCL12/CXCR4 Signaling Axis Retains Neutrophils at Inflammatory Sites in Zebrafish. Front. Immunol. *0*, 1784.

Jaillon, S., Ponzetta, A., Di Mitri, D., Santoni, A., Bonecchi, R., and Mantovani, A. (2020a). Neutrophil diversity and plasticity in tumour progression and therapy. Nat. Rev. Cancer *20*, 485–503.

Jaillon, S., Ponzetta, A., Di Mitri, D., Santoni, A., Bonecchi, R., and Mantovani, A. (2020b). Neutrophil diversity and plasticity in tumour progression and therapy. Nat Rev Cancer *20*, 485–503.

Joeckel, L.T., and Bird, P.I. (2014). Are all granzymes cytotoxic in vivo? Biol. Chem. *395*, 181–202.

Joeckel, L.T., Wallich, R., Martin, P., Sanchez-Martinez, D., Weber, F.C., Martin, S.F., Borner, C., Pardo, J., Froelich, C., and Simon, M.M. (2011). Mouse granzyme K has pro-

inflammatory potential. Cell Death Differ. *18*, 1112–1119.

Joshita, S., Nakazawa, K., Sugiyama, Y., Kamijo, A., Matsubayashi, K., Miyabayashi, H., Furuta, K., Kitano, K., and Kawa, S. (2009). Granulocyte-colony stimulating factor-producing pancreatic adenosquamous carcinoma showing aggressive clinical course. Intern. Med. *48*, 687–691.

Koch, M., Beckhove, P., Op Den Winkel, J., Autenrieth, D., Wagner, P., Nummer, D., Specht, S., Antolovic, D., Galindo, L., Schmitz-Winnenthal, F.H., et al. (2006). Tumor infiltrating T lymphocytes in colorectal cancer: Tumor-selective activation and cytotoxic activity in situ. Ann. Surg. *244*, 986–992.

Kyo, S., Kanaya, T., Takakura, M., and Inoue, M. (2000). A case of cervical cancer with aggressive tumor growth: Possible autocrine growth stimulation by G-CSF and Il-6. Gynecol. Oncol. *78*, 383–387.

Lavin, Y., Kobayashi, S., Leader, A., Amir, E. ad D., Elefant, N., Bigenwald, C., Remark, R., Sweeney, R., Becker, C.D., Levine, J.H., et al. (2017). Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. Cell *169*, 750-765.e17.

Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir, E.A.D., Tadmor, M.D., Litvin, O., Fienberg, H.G., Jager, A., Zunder, E.R., et al. (2015). Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. Cell *162*, 184–197.

Li, H., van der Leun, A.M., Yofe, I., Lubling, Y., Gelbard-Solodkin, D., van Akkooi, A.C.J., van den Braber, M., Rozeman, E.A., Haanen, J.B.A.G., Blank, C.U., et al. (2019). Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. Cell *176*, 775-789.e18.

Li, P., Lu, M., Shi, J., Gong, Z., Hua, L., Li, Q., Lim, B., Zhang, X.H.F., Chen, X., Li, S., et al. (2020). Lung mesenchymal cells elicit lipid storage in neutrophils that fuel breast cancer lung metastasis. Nat. Immunol. *21*, 1444–1455.

Li, Z., Zhao, R., Cui, Y., Zhou, Y., and Wu, X. (2018). The dynamic change of neutrophil to

lymphocyte ratio can predict clinical outcome in stage I-III colon cancer. Sci. Reports 2018 81 *8*, 1–8.

M, S., Y, M., S, R., LG, P., Y, S., D, C., and DJ, G. (2016). Extracellular granzyme K mediates endothelial activation through the cleavage of protease-activated receptor-1. FEBS J. *283*, 1734–1747.

MA, H., N, K., and H, B. (2005). Molecular requirements for epithelial-mesenchymal transition during tumor progression. Curr. Opin. Cell Biol. *17*, 548–558.

Mantovani, A., Cassatella, M.A., Costantini, C., and Jaillon, S. (2011). Neutrophils in the activation and regulation of innate and adaptive immunity. Nat. Rev. Immunol. *11*, 519–531.

Marini, O., Costa, S., Bevilacqua, D., Calzetti, F., Tamassia, N., Spina, C., De Sabata, D., Tinazzi, E., Lunardi, C., Scupoli, M.T., et al. (2017). Mature CD10+ and immature CD10- neutrophils present in G-CSF-treated donors display opposite effects on T cells. Blood *129*, 1343–1356.

McLane, L.M., Abdel-Hakeem, M.S., and Wherry, E.J. (2019). CD8 T Cell Exhaustion During Chronic Viral Infection and Cancer. Annu. Rev. Immunol. *37*, 457–495.

Mogilenko, D.A., Shpynov, O., Andhey, P.S., Arthur, L., Swain, A., Esaulova, E., Brioschi, S., Shchukina, I., Kerndl, M., Bambouskova, M., et al. (2021). Comprehensive Profiling of an Aging Immune System Reveals Clonal GZMK+ CD8+ T Cells as Conserved Hallmark of Inflammaging. Immunity *54*, 99-115.e12.

Moore, K.W., De Waal Malefyt, R., Coffman, R.L., and O'Garra, A. (2001). Interleukin-10 and the interleukin-10 receptor. Annu. Rev. Immunol. *19*, 683–765.

Naito, Y., Saito, K., Shiiba, K., Ohuchi, A., Saigenji, K., Nagura, H., and Ohtani, H. (1998). CD8+ T Cells Infiltrated within Cancer Cell Nests as a Prognostic Factor in Human Colorectal Cancer. Cancer Res. *58*.

Nathan, C., and Ding, A. (2010). Nonresolving Inflammation. Cell *140*, 871–882.

Ogino, S., Nosho, K., Irahara, N., Meyerhardt, J.A., Baba, Y., Shima, K., Glickman, J.N.,

Ferrone, C.R., Mino-Kenudson, M., Tanaka, N., et al. (2009). Lymphocytic reaction to colorectal cancer is associated with longer survival, independent of lymph node count, microsatellite instability, and CpG island methylator phenotype. Clin. Cancer Res. *15*, 6412–6420.

Owyang, A.M., Zaph, C., Wilson, E.H., Guild, K.J., McClanahan, T., Miller, H.R.P., Cua, D.J., Goldschmidt, M., Hunter, C.A., Kastelein, R.A., et al. (2006). Interleukin 25 regulates type 2 cytokine-dependent immunity and limits chronic inflammation in the gastrointestinal tract. J. Exp. Med. *203*, 843–849.

Pagès, F., Kirilovsky, A., and ...Galon, J. (2009a). In situ cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. J. Clin. Oncol. *27*, 5944–5951.

Pagès, F., Berger, A., Camus, M., Sanchez-Cabo, F., Costes, A., Molidor, R., Mlecnik, B., Kirilovsky, A., Nilsson, M., Damotte, D., et al. (2009b). Effector Memory T Cells, Early Metastasis, and Survival in Colorectal Cancer. Http://Dx.Doi.Org/10.1056/NEJMoa051424 *353*, 2654–2666.

Pagès, F., Kirilovsky, A., Mlecnik, B., Asslaber, M., Berger, A., Fridman, W., and Galon, J. (2009c). In situ cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. J. Clin. Oncol. *27*, 5944–5951.

Peng, Z., Liu, C., Victor, A.R., Cao, D.Y., Veiras, L.C., Bernstein, E.A., Khan, Z., Giani, J.F., Cui, X., Bernstein, K.E., et al. (2021a). Tumors exploit CXCR4hiCD62Llo aged neutrophils to facilitate metastatic spread. Oncoimmunology *10*.

Peng, Z., Liu, C., Victor, A.R., Cao, D.-Y., Veiras, L.C., Bernstein, E.A., Khan, Z., Giani, J.F., Cui, X., Bernstein, K.E., et al. (2021b). Tumors exploit CXCR4hiCD62Llo aged neutrophils to facilitate metastatic spread. Https://Doi.Org/10.1080/2162402X.2020.1870811 *10*.

Pillay, J., Tak, T., Kamp, V.M., and Koenderman, L. (2013). Immune suppression by neutrophils and granulocytic myeloid-derived suppressor cells: Similarities and differences.

Cell. Mol. Life Sci. *70*, 3813–3827.

Powell, D., Tauzin, S., Hind, L.E., Deng, Q., Beebe, D.J., and Huttenlocher, A. (2017). Chemokine Signaling and the Regulation of Bidirectional Leukocyte Migration in Interstitial Tissues. Cell Rep. *19*, 1572–1585.

Prall, F., Dührkop, T., Weirich, V., Ostwald, C., Lenz, P., Nizze, H., and Barten, M. (2004). Prognostic role of CD8+ tumor-infiltrating lymphocytes in stage III colorectal cancer with and without microsatellite instability. Hum. Pathol. *35*, 808–816.

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

Robertson, A.L., Holmes, G.R., Bojarczuk, A.N., Burgon, J., Loynes, C.A., Chimen, M., Sawtell, A.K., Hamza, B., Willson, J., Walmsley, S.R., et al. (2014). A zebrafish compound screen reveals modulation of neutrophil reverse migration as an anti-inflammatory mechanism. Sci. Transl. Med. *6*.

Rock, K.L., Latz, E., Ontiveros, F., and Kono, H. (2010). The sterile inflammatory response. Annu. Rev. Immunol. *28*, 321–342.

Romero, P., Zippelius, A., Kurth, I., Pittet, M.J., Touvrey, C., Iancu, E.M., Corthesy, P., Devevre, E., Speiser, D.E., and Rufer, N. (2007). Four Functionally Distinct Populations of Human Effector-Memory CD8 + T Lymphocytes. J. Immunol. *178*, 4112–4119.

Rooney, M.S., Shukla, S.A., Wu, C.J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell *160*, 48–61.

Salmon, H., Idoyaga, J., Rahman, A., Leboeuf, M., Remark, R., Jordan, S., Casanova-Acebes, M., Khudoynazarova, M., Agudo, J., Tung, N., et al. (2016). Expansion and Activation of CD103+ Dendritic Cell Progenitors at the Tumor Site Enhances Tumor Responses to Therapeutic PD-L1 and BRAF Inhibition. Immunity *44*, 924–938.

Sangaletti, S., Tripodo, C., Vitali, C., Portararo, P., Guarnotta, C., Casalini, P., Cappetti, B., Miotti, S., Pinciroli, P., Fuligni, F., et al. (2014a). Defective Stromal Remodeling and Neutrophil Extracellular Traps in Lymphoid Tissues Favor the Transition from Autoimmunity to Lymphoma. Cancer Discov. *4*, 110–129.

Sangaletti, S., Tripodo, C., Vitali, C., Portararo, P., Guarnotta, C., Casalini, P., Cappetti, B., Miotti, S., Pinciroli, P., Fuligni, F., et al. (2014b). Defective stromal remodeling and neutrophil extracellular traps in lymphoid tissues favor the transition from autoimmunity to lymphoma. Cancer Discov *4*, 110–129.

Schall, T.J., Bacon, K., Toy, K.J., and Goeddel, D. V. (1990). Selective attraction of monocytes and T lymphocytes of the memory phenotype by cytokine RANTES. Nature *347*, 669–671.

Scheper, W., Kelderman, S., Fanchi, L.F., Linnemann, C., Bendle, G., de Rooij, M.A.J., Hirt, C., Mezzadra, R., Slagter, M., Dijkstra, K., et al. (2019). Low and variable tumor reactivity of the intratumoral TCR repertoire in human cancers. Nat. Med. *25*, 89–94.

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 2012 97 *9*, 676–682.

Schmidt, U., Weigert, M., Broaddus, C., and Myers, G. (2018). Cell detection with star-convex polygons. In Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), (Springer, Cham), pp. 265–273.

Senra, L., Mylonas, A., Kavanagh, R.D., Fallon, P.G., Conrad, C., Borowczyk-Michalowska, J., Wrobel, L.J., Kaya, G., Yawalkar, N., Boehncke, W.H., et al. (2019). IL-17E (IL-25) Enhances Innate Immune Responses during Skin Inflammation. J. Invest. Dermatol. *139*, 1732-1742.e17.

Seo, W., Shimizu, K., Kojo, S., Okeke, A., Kohwi-Shigematsu, T., Fujii, S., and Taniuchi, I.

(2020). Runx-mediated regulation of CCL5 via antagonizing two enhancers influences immune cell function and anti-tumor immunity. Nat. Commun. 2020 111 *11*, 1–16.

Shaul, M.E., and Fridlender, Z.G. (2019). Tumour-associated neutrophils in patients with cancer. Nat. Rev. Clin. Oncol. 2019 1610 *16*, 601–620.

Shelley, M., Ross, C., and Nizet, V. (2018). The Ontogeny of a Neutrophil: Mechanisms of Granulopoiesis and Homeostasis. Microbiol. Mol. Biol. Rev. *82*.

Silvestre-Roig, C., Fridlender, Z.G., Glogauer, M., and Scapini, P. (2019). Neutrophil Diversity in Health and Disease. Trends Immunol. *40*, 565–583.

Simoni, Y., Becht, E., Fehlings, M., Loh, C.Y., Koo, S.L., Teng, K.W.W., Yeong, J.P.S., Nahar, R., Zhang, T., Kared, H., et al. (2018). Bystander CD8+ T cells are abundant and phenotypically distinct in human tumour infiltrates. Nature *557*, 575–579.

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.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e21.

Sun, H., Zhou, Y., Fei, L., Chen, H., and Guo, G. (2019). scMCA: A tool to define mouse cell types based on single-cell digital expression. In Methods in Molecular Biology, (Methods Mol Biol), pp. 91–96.

Swanson, B.J., Murakami, M., Mitchell, T.C., Kappler, J., and Marrack, P. (2002). RANTES production by memory phenotype T cells is controlled by a posttranscriptional, TCR-dependent process. Immunity *17*, 605–615.

Thome, J.J.C., and Farber, D.L. (2015). Emerging concepts in tissue-resident T cells: Lessons from humans. Trends Immunol. *36*, 428–435.

Walzer, T., Marçais, A., Saltel, F., Bella, C., Jurdic, P., and Marvel, J. (2003). Cutting Edge: Immediate RANTES Secretion by Resting Memory CD8 T Cells Following Antigenic Stimulation. J. Immunol. *170*, 1615–1619.

Wang, A.J., Smith, A., Li, Y., Urban, J.F., Ramalingam, T.R., Wynn, T.A., Lu, N., Shea-Donohue, T., Yang, Z., and Zhao, A. (2014). Genetic deletion of IL-25 (IL-17E) confers resistance to dextran sulfate sodium-induced colitis in mice. Cell Biosci. *4*.

Wang, J., Hossain, M., Thanabalasuriar, A., Gunzer, M., Meininger, C., and Kubes, P. (2017). Visualizing the function and fate of neutrophils in sterile injury and repair. Science (80-.). *358*, 111–116.

Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R.M., Ozenberger, B.A., Ellrott, K., Sander, C., Stuart, J.M., Chang, K., Creighton, C.J., et al. (2013). The cancer genome atlas pan-cancer analysis project. Nat. Genet. *45*, 1113–1120.

Wensink, A.C., Kemp, V., Fermie, J., Laorden, M.I.G., Van Der Poll, T., Hack, C.E., and Bovenschen, N. (2014). Granzyme K synergistically potentiates LPS-induced cytokine responses in human monocytes. Proc. Natl. Acad. Sci. U. S. A. *111*, 5974–5979.

Wensink, A.C., Hack, C.E., and Bovenschen, N. (2015). Granzymes Regulate Proinflammatory Cytokine Responses. J. Immunol. *194*, 491–497.

Woodfin, A., Voisin, M.B., Beyrau, M., Colom, B., Caille, D., Diapouli, F.M., Nash, G.B., Chavakis, T., Albelda, S.M., Rainger, G.E., et al. (2011). The junctional adhesion molecule JAM-C regulates polarized transendothelial migration of neutrophils in vivo. Nat. Immunol. *12*, 761–769.

Wu, P., Wu, D., Ni, C., Ye, J., Chen, W., Hu, G., Wang, Z., Wang, C., Zhang, Z., Xia, W., et al. (2014). γδT17 cells promote the accumulation and expansion of myeloid-derived suppressor cells in human colorectal cancer. Immunity *40*, 785–800.

Zeisberg, M., and Neilson, E.G. (2009). Biomarkers for epithelial-mesenchymal transitions. J. Clin. Invest. *119*, 1429–1437.

Zhang, L., Yu, X., Zheng, L., Zhang, Y., Li, Y., Fang, Q., Gao, R., Kang, B., Zhang, Q., Huang, J.Y., et al. (2018). Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature *564*, 268–272.

Zlobec, I., Karamitopoulou, E., Terracciano, L., Piscuoglio, S., Iezzi, G., Muraro, M.G., Spagnoli, G., Baker, K., Tzankov, A., and Lugli, A. (2010). TIA-1 Cytotoxic Granule-Associated RNA Binding Protein Improves the Prognostic Performance of CD8 in Mismatch Repair-Proficient Colorectal Cancer. PLoS One *5*, e14282.

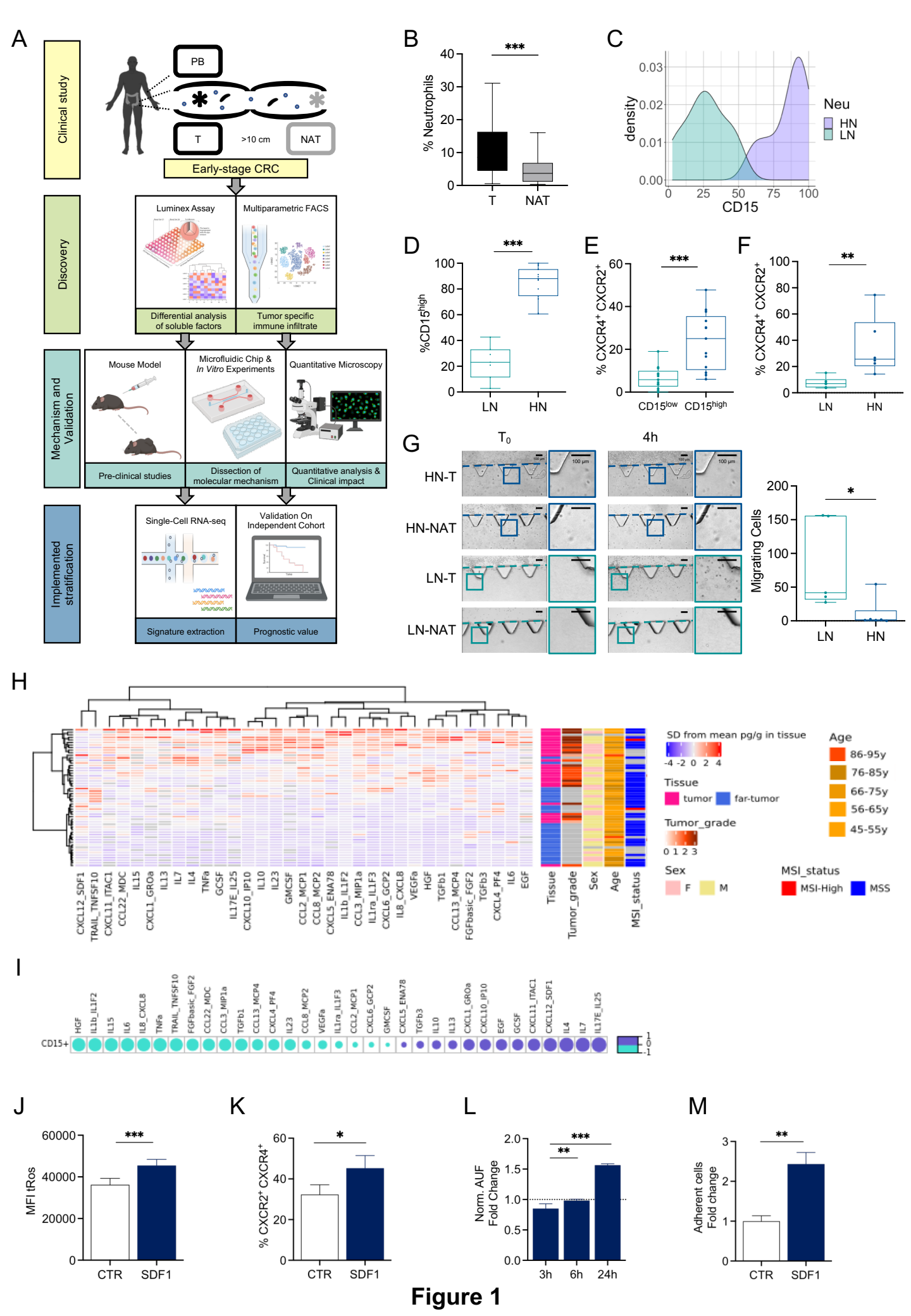
A
Clinical study
PB
T
>10 cm
NAT
Early-stage CRC
Discovery
Luminex Assay
Multiparametric FACS
Differential analysis of soluble factors
Tumor specific immune infiltrate
Mechanism and Validation
Mouse Model
Microfluidic Chip & In Vitro Experiments
Quantitative Microscopy
Pre-clinical studies
Dissection of molecular mechanism
Quantitative analysis & Clinical impact
Implemented stratification
Single-Cell RNA-seq
Validation On Independent Cohort
Signature extraction
Prognostic value
B
% Neutrophils
T
NAT
C
density
CD15
Neu
HN
LN
D
%CD15high
LN
HN
E
% CXCR4+ CXCR2+
CD15low
CD15high
F
% CXCR4+ CXCR2+
LN
HN
G
T0
4h
HN-T
HN-NAT
LN-T
LN-NAT
Migrating Cells
LN
HN
H
SD from mean pg/g in tissue
Tissue
tumor
far-tumor
Tumor_grade
Age
86-95y
76-85y
66-75y
56-65y
45-55y
Sex
F
M
MSI_status
MSI-High
MSS
CXCL12_SDF1
TRAIL_TNFSF10
CXCL11_ITAC1
CCL22_MDC
IL15
CXCL1_GROa
IL13
IL7
IL4
TNFa
GCSF
IL17E_IL25
CXCL10_IP10
IL10
IL23
GMCSF
CCL2_MCP1
CCL8_MCP2
CXCL5_ENA78
IL1b_IL1F2
CCL3_MIP1a
IL1ra_IL1F3
CXCL6_GCP2
IL8_CXCL8
VEGFa
HGF
TGFb1
CCL13_MCP4
FGFbasic_FGF2
TGFb3
CXCL4_PF4
IL6
EGF
Tissue
Tumor_grade
Sex
Age
MSI_status
I
CD15+
HGF
IL1b_IL1F2
IL15
IL6
IL8_CXCL8
TNFa
TRAIL_TNFSF10
FGFbasic_FGF2
CCL22_MDC
CCL3_MIP1a
TGFb1
CCL13_MCP4
CXCL4_PF4
IL23
CCL8_MCP2
VEGFa
IL1ra_IL1F3
CCL2_MCP1
CXCL6_GCP2
GMCSF
CXCL5_ENA78
TGFb3
IL10
IL13
CXCL1_GROa
CXCL10_IP10
EGF
GCSF
CXCL11_ITAC1
CXCL12_SDF1
IL4
IL7
IL17E_IL25
J
MFI tRos
CTR
SDF1
K
% CXCR2+ CXCR4+
CTR
SDF1
L
Norm. AUF Fold Change
3h
6h
24h
M
Adherent cells Fold change
CTR
SDF1

Figure 1

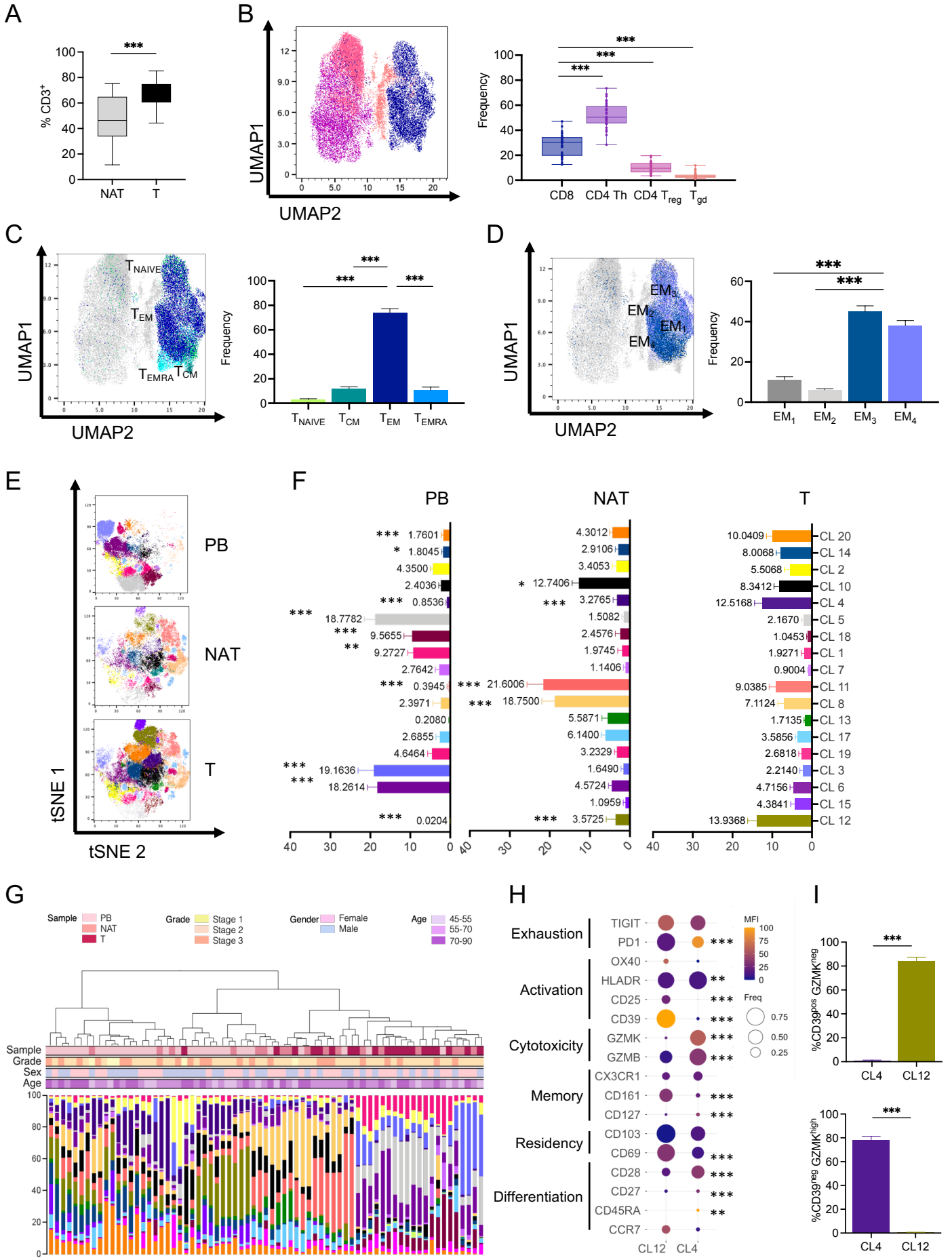
A
% CD3+
NAT
T
B
UMAP1
UMAP2
Frequency
CD8
CD4 Th
CD4 Treg
Tgd
C
TNAIVE
TEM
TEMRA
TCM
D
EM1
EM2
EM3
EM4
E
PB
NAT
T
tSNE 1
tSNE 2
F
PB
NAT
T
CL 20
CL 14
CL 2
CL 10
CL 4
CL 5
CL 18
CL 1
CL 7
CL 11
CL 8
CL 13
CL 17
CL 19
CL 3
CL 6
CL 15
CL 12
G
Sample
PB
NAT
T
Grade
Stage 1
Stage 2
Stage 3
Gender
Female
Male
Age
45-55
55-70
70-90
Sex
H
Exhaustion
TIGIT
PD1
Activation
OX40
HLADR
CD25
CD39
Cytotoxicity
GZMK
GZMB
Memory
CX3CR1
CD161
CD127
Residency
CD103
CD69
Differentiation
CD28
CD27
CD45RA
CCR7
CL12
CL4
MFI
Freq
I
%CD39pos GZMKneg
%CD39neg GZMKhigh
CL4
CL12

Figure 2

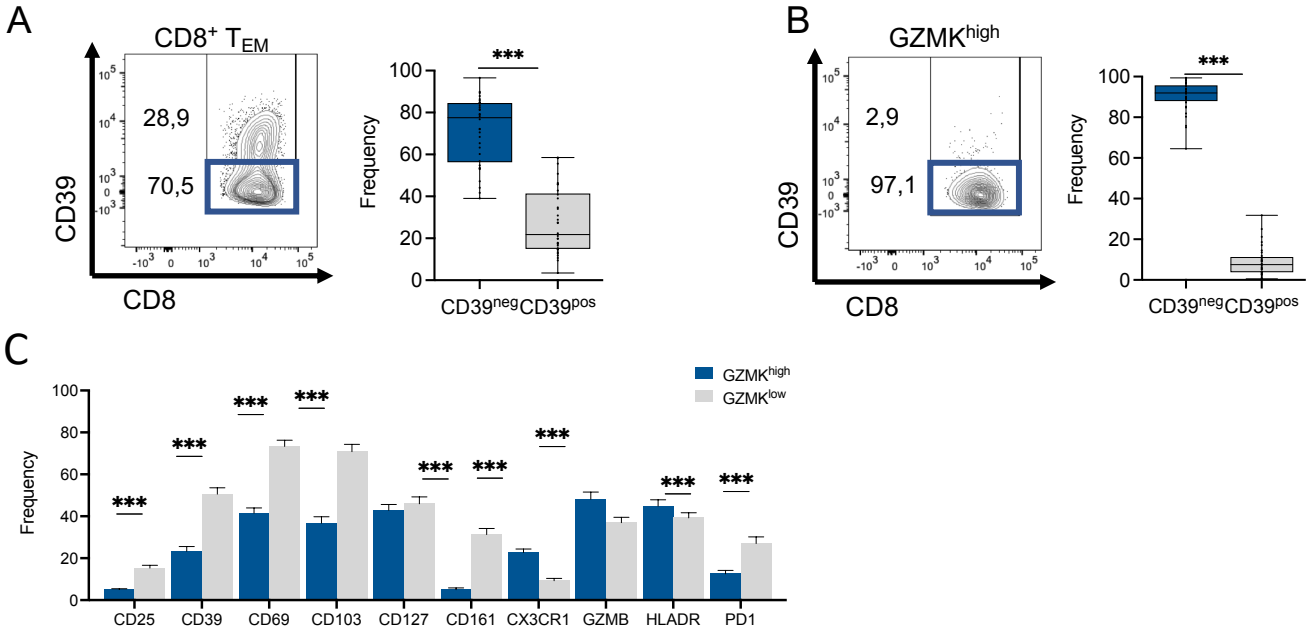
A
CD8+ TEM
28,9
70,5
CD39
CD8
Frequency
***
CD39negCD39pos
B
GZMKhigh
2,9
97,1
CD39
CD8
Frequency
***
CD39negCD39pos
C
GZMKhigh
GZMKlow
Frequency
CD25
CD39
CD69
CD103
CD127
CD161
CX3CR1
GZMB
HLADR
PD1

Figure 3

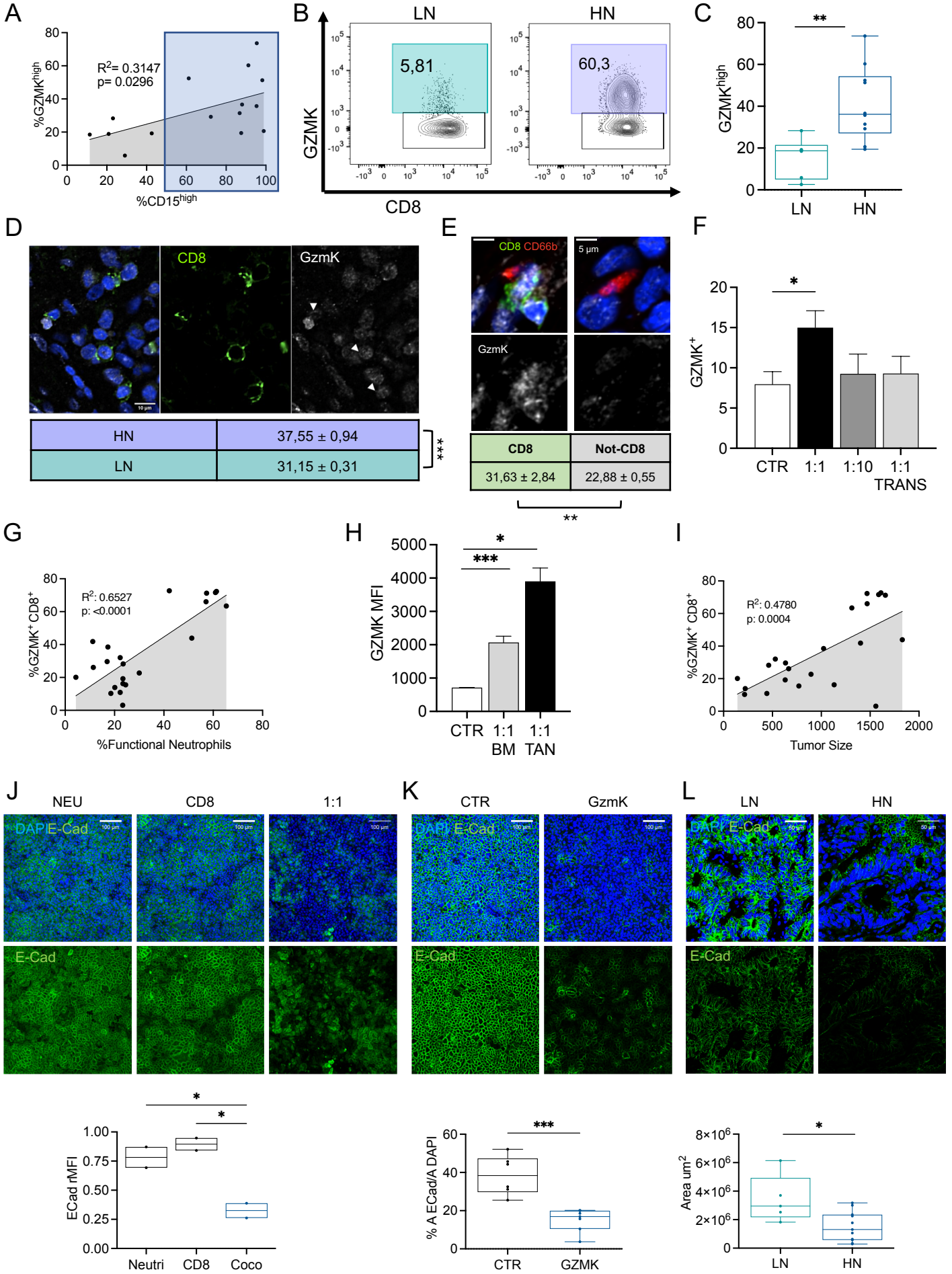
A
R2= 0.3147
p= 0.0296
%GZMKhigh
%CD15high
B
LN
HN
5,81
60,3
GZMK
CD8
C
GZMKhigh
LN
HN
D
CD8
GzmK
HN 37,55 ± 0,94
LN 31,15 ± 0,31
***
E
CD8 CD66b
5 µm
GzmK
CD8 Not-CD8
31,63 ± 2,84 22,88 ± 0,55
**
F
GZMK+
CTR 1:1 1:10 1:1 TRANS
G
R2: 0.6527
p: <0.0001
%GZMK+ CD8+
%Functional Neutrophils
H
GZMK MFI
CTR 1:1 BM 1:1 TAN
I
R2: 0.4780
p: 0.0004
%GZMK+ CD8+
Tumor Size
J
NEU
CD8
1:1
DAPIE-Cad
E-Cad
ECad rMFI
Neutri CD8 Coco
K
CTR
GzmK
DAPI E-Cad
E-Cad
% A ECad/A DAPI
CTR GZMK
L
LN
HN
DAPI E-Cad
E-Cad
Area um2
LN HN

Figure 4

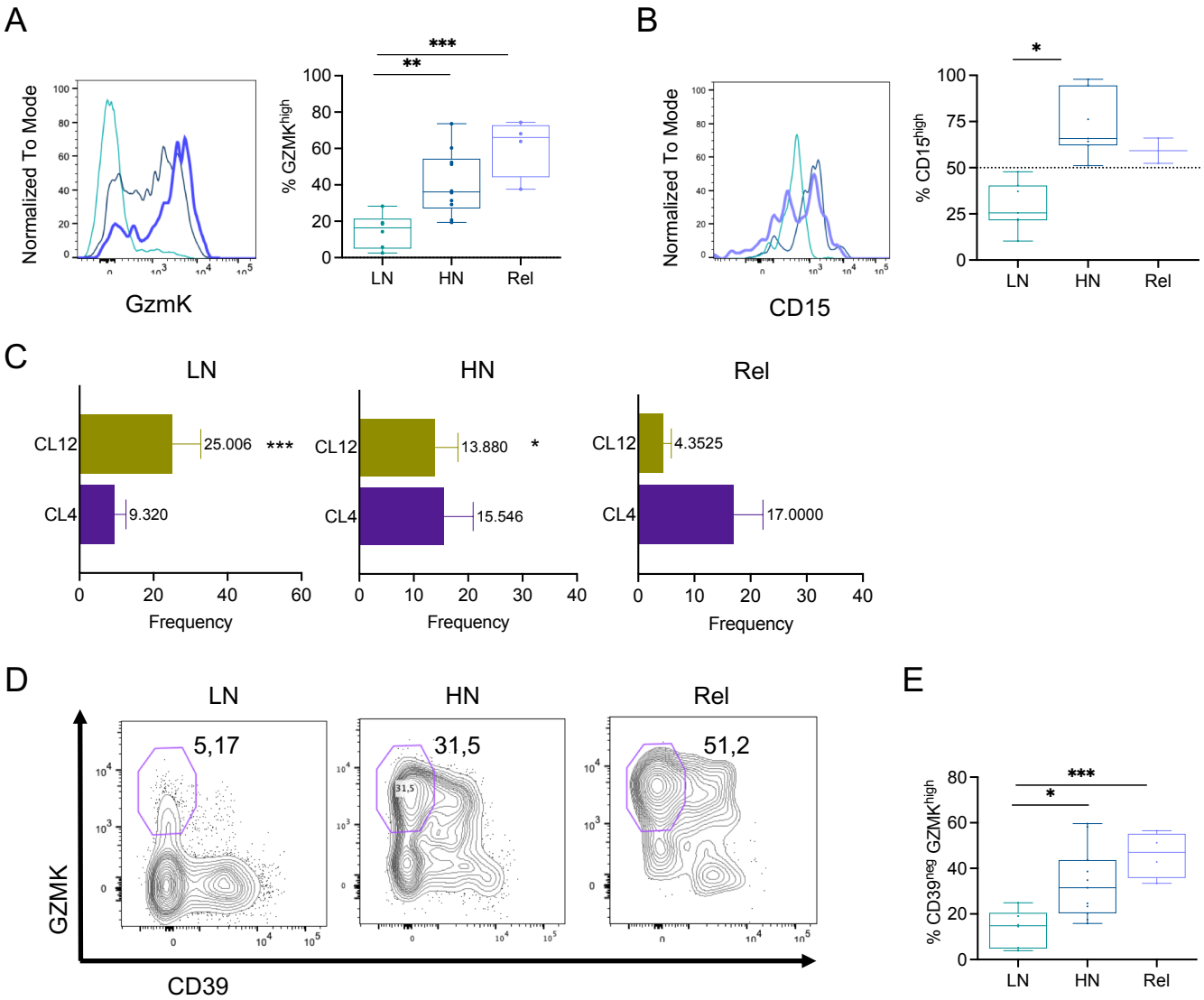
A
Normalized To Mode
GzmK
% GZMK^high
***
**
LN
HN
Rel
B
Normalized To Mode
CD15
% CD15^high
*
LN
HN
Rel
C
LN
HN
Rel
CL12
CL4
25.006
***
9.320
13.880
*
15.546
4.3525
17.0000
Frequency
D
LN
HN
Rel
5,17
31,5
51,2
GZMK
CD39
E
% CD39^neg GZMK^high
***
*
LN
HN
Rel

Figure 5

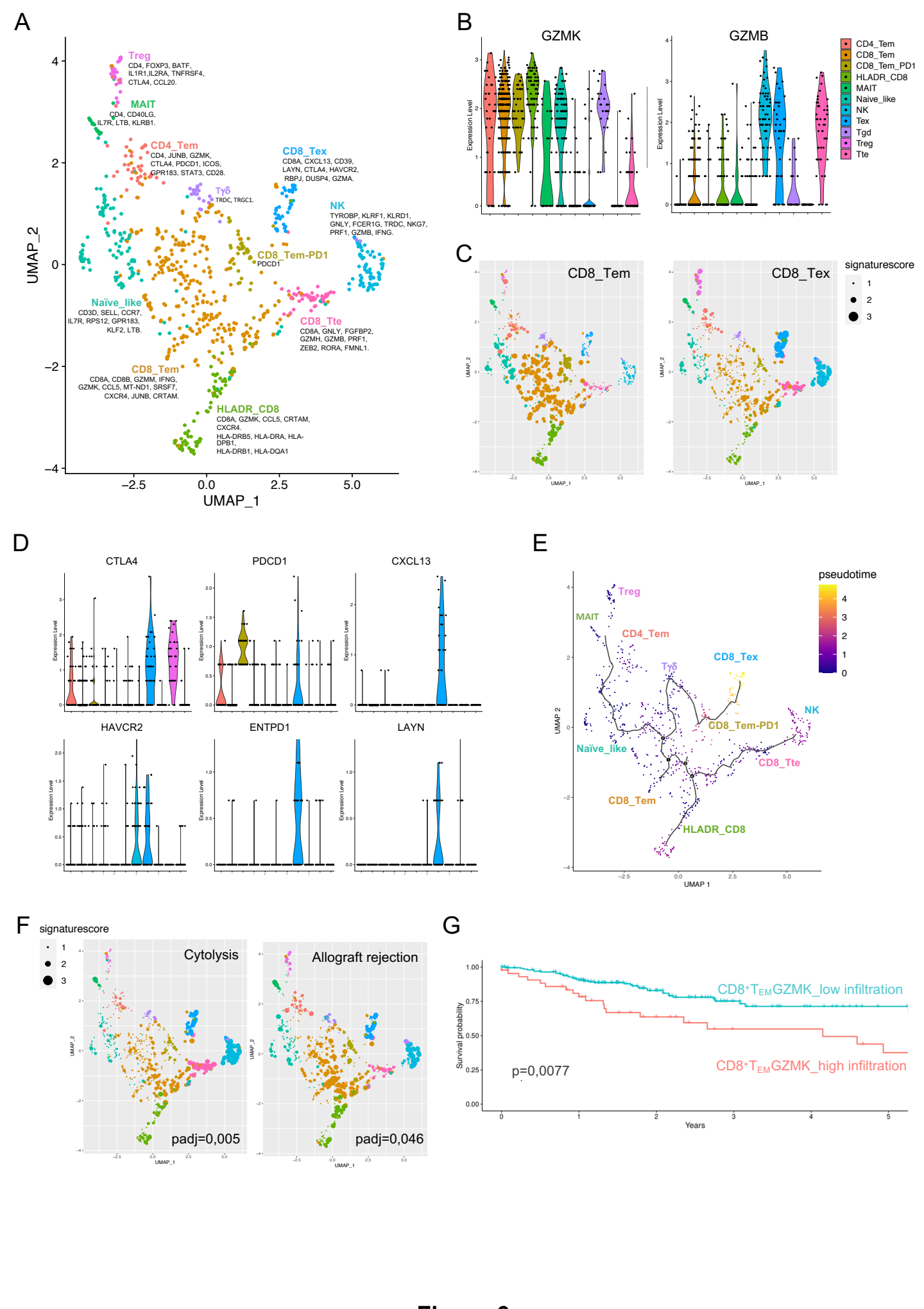
A
Treg
CD4, FOXP3, BATF, IL1R1,IL2RA, TNFRSF4, CTLA4, CCL20.
MAIT
CD4, CD40LG, IL7R, LTB, KLRB1.
CD4_Tem
CD4, JUNB, GZMK, CTLA4, PDCD1, ICOS, GPR183, STAT3, CD28.
Tγδ
TRDC, TRGC1.
CD8_Tex
CD8A, CXCL13, CD39, LAYN, CTLA4, HAVCR2, RBPJ, DUSP4, GZMA.
NK
TYROBP, KLRF1, KLRD1, GNLY, FCER1G, TRDC, NKG7, PRF1, GZMB, IFNG.
CD8_Tem-PD1
PDCD1
Naïve_like
CD3D, SELL, CCR7, IL7R, RPS12, GPR183, KLF2, LTB.
CD8_Tte
CD8A, GNLY, FGFBP2, GZMH, GZMB, PRF1, ZEB2, RORA, FMNL1.
CD8_Tem
CD8A, CD8B, GZMM, IFNG, GZMK, CCL5, MT-ND1, SRSF7, CXCR4, JUNB, CRTAM.
HLADR_CD8
CD8A, GZMK, CCL5, CRTAM, CXCR4.
HLA-DRB5, HLA-DRA, HLA-DPB1,
HLA-DRB1, HLA-DQA1
UMAP_2
UMAP_1
B
GZMK
GZMB
Expression Level
CD4_Tem
CD8_Tem
CD8_Tem_PD1
HLADR_CD8
MAIT
Naive_like
NK
Tex
Tgd
Treg
Tte
C
CD8_Tem
CD8_Tex
signaturescore
1
2
3
UMAP_2
UMAP_1
D
CTLA4
PDCD1
CXCL13
HAVCR2
ENTPD1
LAYN
Expression Level
E
pseudotime
Treg
MAIT
CD4_Tem
Tγδ
CD8_Tex
NK
CD8_Tem-PD1
Naïve_like
CD8_Tte
CD8_Tem
HLADR_CD8
UMAP 2
UMAP 1
F
signaturescore
1
2
3
Cytolysis
padj=0,005
Allograft rejection
padj=0,046
UMAP_2
UMAP_1
G
Survival probability
CD8⁺T_EM GZMK_low infiltration
CD8⁺T_EM GZMK_high infiltration
p=0,0077
Years

Figure 6

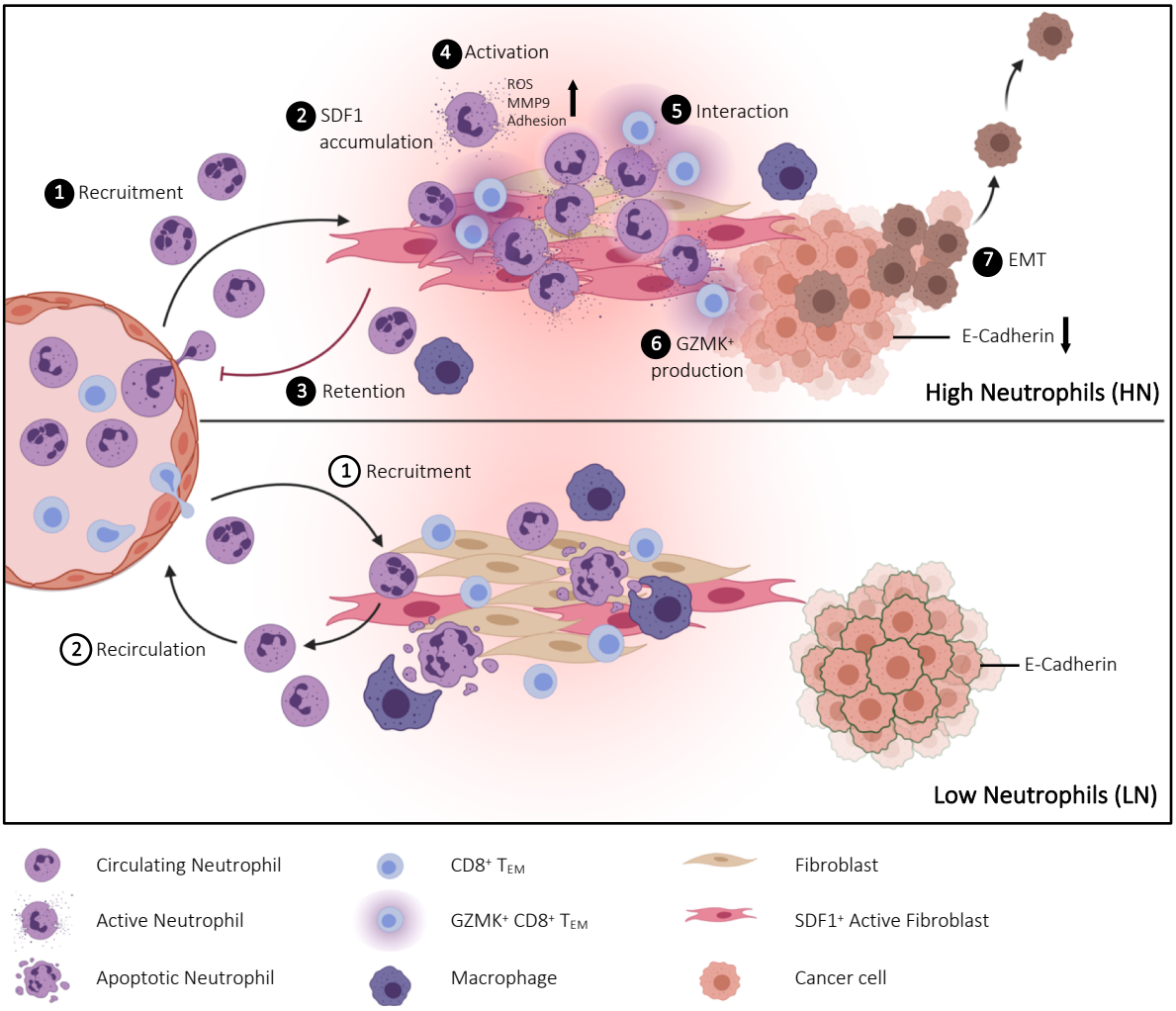
4 Activation
ROS
MMP9
Adhesion
2 SDF1
accumulation
5 Interaction
1 Recruitment
7 EMT
6 GZMK+
production
E-Cadherin
3 Retention
High Neutrophils (HN)
1 Recruitment
2 Recirculation
E-Cadherin
Low Neutrophils (LN)
Circulating Neutrophil
Active Neutrophil
Apoptotic Neutrophil
CD8+ TEM
GZMK+ CD8+ TEM
Macrophage
Fibroblast
SDF1+ Active Fibroblast
Cancer cell

Figure 7

Study Cohort		
	Count	Frequency (%)
Total of patients	42	100
Relapse (2 years)	4	9,5
MSI	6	14,3
Age		
<50	2	4,8
50-70	22	52,4
>70	18	42,9
Gender		
F	20	47,6
M	22	52,4
Stage first diagnosis		
I	9	21,4
II	15	35,7
III	18	42,9
BMI		
Normal	19	45,2
Overweight	15	35,7
Obese	7	16,7
NA	1 (NT14)	2,4
Site of primary tumor		
Left	18	42,8
Right	23	54,8
Medium Transverse Colon	1	2,4
Tumor Type		
Adenocarcinoma	39	92,9
Mucinous adenocarcinoma	3	7,1

Table 1. CRC patients' characteristics. Age, Gender, Stage at first diagnosis, Body Mass Index (BMI), Site of primary tumor and Tumor type are shown for the study cohort.

Marker	Fluorophore	Manufacturer	Cat. #
Zombie Aqua		BioLegend	423101
TCRgd	PerCP-Cy5.5	BioLegend	331224
NKG2A	FITC	Miltenyi	130-113-568
CD39	APC-H7	BioLegend	328226
TIGIT	APC	BioLegend	372706
CD25	BV786	BD	741035
CCR7	BV711	BD	566602
OX40	BV650	BD	563658
CD161	BV605	Biolegend	339916
CD27	BV570	BioLegend	302825
CD11b	BV510	Biolegend	301334
PD1	BV480	BD	566112
CD103	BV421	BioLegend	350214
CD8	BUV805	BD	564912
CD28	BUV737	BD	564438
HLADR	BUV661	BD	565073
CD4	BUV615	BD	624297
CD45RA	BUV563	BD	565702
CD3	BUV496	BD	564809
CD69	BUV395	BD	564364
CD45	PE-Cy7	Biolegend	304016304016
CD56	PE-CY5.5	eBioscience	35-0567-42
CD127	PE-CY5	eBioscience	15-1278-42
CX3CR1	PECF594	Biolegend	341624
GZMB	APC-R700	BD	560213
GZMK	PE	Santa Cruz	sc-56125

Table 2. Markers Multiparametric Flow Cytometry panel. Markers with the corresponding Fluorophore, Manufacturer and Catalog ID (Cat. #) are shown.