Novel mediators of anti-tumor immunity: dissecting intratumoral immune responses at the single-cell level

Vries, N.L. de

Novel mediators of anti-tumor immunity:

dissecting intratumoral immune responses at the single-cell level

Natasja L. de Vries

Novel mediators of anti-tumor immunity:

dissecting intratumoral immune responses at the single-cell level

Natasja L. de Vries

Novel mediators of anti-tumor immunity:
dissecting intratumoral immune responses at the single-cell level

ISBN:	9789464195460
Cover design by:	Erik Heeres
Layout by:	Ilse Modder I www.ilsemodder.nl
Printing by:	Gildeprint Enschede I www.gildeprint.nl

The research presented in this thesis was performed at the Department of Immunology and the Department of Pathology of Leiden University Medical Center in the Netherlands.

Novel mediators of anti-tumor immunity:
dissecting intratumoral immune responses at the single-cell level

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 6 oktober 2022
klokke 11.15 uur

door

Natasja Lieuwke de Vries
geboren te Leiden
in 1992

Promotor:	Prof.dr. F. Koning
Co-promotor:	Dr. N.F.C.C. de Miranda
Leden promotiecommissie:	Prof.dr. F.J.T. Staal Prof.dr. S.H. van der Burg Prof.dr. I.J.M. de Vries (Radboud UMC) Prof.dr. Z. Trajanoski (Medical University of Innsbruck)

"De adelaar zweeft graag hoog boven de wereld,
niet om op de mensen neer te zien,
maar om hen aan te sporen omhoog te kijken."

Uit: De cirkel van het leven, Kübler-Ross, E.

Ter nagedachtenis aan mijn moeder

Voor mijn ouders

TABLE OF CONTENTS

Chapter 1	General introduction and outline of this thesis	13
Chapter 2	Unraveling the complexity of the cancer microenvironment with multidimensional genomic and cytometric technologies	27
Chapter 3	High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of anti-tumor immunity	55
Chapter 4	γδ T cells are effectors of immune checkpoint blockade in mismatch repair-deficient colon cancers with antigen presentation defects	91
Chapter 5	The ILC compartment in mismatch repair-deficient colorectal cancers is dominated by CD127-negative ILC1-like cells	135
Chapter 6	T cells with tumor-specific phenotypes largely originate from pericolic lymph nodes in colorectal cancer	167
Chapter 7	Local and systemic immune profiles of human pancreatic ductal adenocarcinoma revealed by single-cell mass cytometry	191
Chapter 8	General discussion and future perspectives	227
Appendices	Summary in English	238
	Nederlandse samenvatting	242
	Curriculum vitae and List of publications	250
	Acknowledgements / Dankwoord	254

ABBREVIATIONS

B2M/β2m	β2-microglobulin
$B2M^{KO}$	β2-microglobulin-knockout
$B2M^{MUT}$	β2-microglobulin-mutant
$B2M^{WT}$	β2-microglobulin-wildtype
BTN	Butyrophilin
CAR	Chimeric antigen receptor
CDR3	Third complementarity-determining region
CIN	Chromosomal instability
CM	Central memory
CML	Chronic myelogenous leukemia
COAD	Colon adenocarcinoma
CRC	Colorectal cancer
CyTOF	Cytometry by time-of-flight
DC	Dendritic cell
DRUP	Drug rediscovery protocol
EM	Effector memory
GMS	Gaussian mean shift
HLA	Human leukocyte antigen
HSNE	Hierarchical stochastic neighbor embedding
ICB	Immune checkpoint blockade
ILC	Innate lymphoid cell
IMC	Imaging mass cytometry
KIR	Killer-cell immunoglobulin-like receptor
LN	Lymph node
MHC	Major histocompatibility complex
MIBI-TOF	Multiplexed ion beam imaging by time-of-flight
MMR	Mismatch repair
MMR-d	Mismatch repair-deficient
MMR-p	Mismatch repair-proficient
MSI-H	Microsatellite instability-high
MSS	Microsatellite stable
NCR	Natural cytotoxicity receptor
PBMC	Peripheral blood mononuclear cell
PDAC	Pancreatic ductal adenocarcinoma
PDTO	Patient-derived tumor organoid
NK	Natural killer
scRNA-seq	single-cell RNA-sequencing
STAD	Stomach adenocarcinoma
TCR	T cell receptor
TCRβ	T cell receptor β-chain
TGF-β	Transforming growth factor-β

TIL	Tumor-infiltrating lymphocyte
TLS	Tertiary lymphoid structure
Treg	Regulatory T cell
t-SNE	t-distributed stochastic neighbor embedding
UCEC	Uterus corpus endometrium carcinoma
UMAP	Uniform manifold approximation and projection

1 General introduction

GENERAL INTRODUCTION

The complexities of cancer pathogenesis

In 1914, the German zoologist Theodor Boveri was the first to propose a genetic basis for cancer development and speculated that cancer is the result of chromosome alterations in cells.[1] Ever since, a cancer cell-centric vision has long dominated the field of cancer biology. Cancer was thought to form exclusively by multistep alterations in the genome of cells, leading to the progressive transformation of normal cells into cancer cells along with the acquisition of malignant features such as self-sufficiency in growth signals, evasion from apoptosis, and limitless replicative potential, all constituting important "hallmarks of cancer".[2] However, during the last decades it has become increasingly recognized that the study of cancer must also include other components of the cancer microenvironment, such as stromal cells, fibroblasts, and immune cells, to fully capture cancer biology.[3,4] Hence, for malignant transformation cells not only have to undergo alterations that result in their proliferation and migration, but also have to evade from control mechanisms, including those provided by cells surrounding the cancer cells such as immune cells.

The role of the immune system in cancer

Malignant transformation is accompanied by genetic alterations that eventually translate to modified proteins. The latter can be recognized during the process of immune surveillance[5], in which the immune system constantly surveys the body for transformed cells and eradicates such cells. The importance of the immune system in recognizing and eliminating cancer cells gained recognition by observations of an increased incidence of cancer among transplant patients using immunosuppressive drugs,[6] patients with immunodeficiencies,[7] and mutant mouse models lacking key components of the immune system[8-10]. Cancer immune surveillance incorporates contributions of both innate and adaptive immunity, which are each equipped with mechanisms to recognize and eliminate cancer cells (**Figure 1**). Innate lymphoid cells (ILCs), including natural killer (NK) cells, can be activated in response to cellular stress or DNA damage occurring in cancer cells, leading to increased expression of stress-induced self ligands on the surface of cancer cells that can be recognized by innate immune receptors.[11] In addition, ILCs and NK cells can be activated in response to the absence of self human leukocyte antigen (HLA) class I molecules,[12] the human counterpart of the major histocompatibility complex (MHC), on the surface of cancer cells. Myeloid cells, such as dendritic cells, macrophages, monocytes and granulocytes, also belong to the innate immune system and play important roles in the phagocytosis of cancer cells and the secretion of inflammatory cytokines.[13] By a highly orchestrated process dictated by multiple cues (e.g. inflammatory signals) that are provided by the innate immune system, cells of the adaptive immune system can become activated. The adaptive immune compartment comprises B and T lymphocytes. The latter can be separated into $CD8^+$ (cytotoxic) or $CD4^+$ (helper) T cells. In contrast to innate immune cells, B and T cells express diverse rearranged receptors which are specific for an antigen, and generate long-term immunological memory.

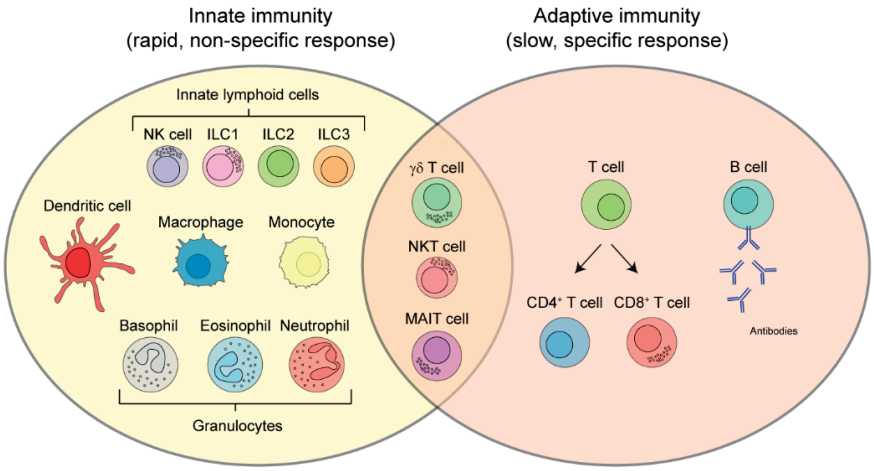


Figure 1. Cellular players of innate and adaptive immunity.
A simplified overview of the main cellular players of innate and adaptive immune responses. The innate immune system mounts a rapid and non-specific response, and comprises innate lymphoid cells, dendritic cells, macrophages, monocytes, and granulocytes. The adaptive immune response is slower to develop, but shows antigenic specificity and long-term immunological memory. It consists of T cells ($CD4^+$ and $CD8^+$ T cells) and B cells. γδ T cells, NKT cells, and MAIT cells are cytotoxic lymphocytes that span the interface of innate and adaptive immunity. Adapted from Dranoff *et al.* (2004)[14].

The genetic and epigenetic alterations that define cancer allow the immune system to generate anti-tumor T cell responses. Cancer cells express a range of aberrantly expressed proteins, including proteins originating from somatic mutations in cancer genomes (neoantigens[15]) and proteins expressed at elevated levels in tumor cells that can also be present on non-malignant cells (tumor-associated antigens, differentiation antigens[16]). During carcinogenesis, dying cancer cells release antigens that can initiate a stepwise process referred to as the cancer-immunity cycle[17] (**Figure 2**). For a successful anti-tumor T cell response, such cancer antigens need to be recognized and taken up during the process of immune surveillance via interactions with antigen-presenting cells such as dendritic cells accompanied by pro-inflammatory signals. Upon uptake of cancer antigens, dendritic cells migrate toward regional lymph nodes where they present captured antigens on HLA class I and II molecules to naïve $CD8^+$ and $CD4^+$ T cells, respectively. The process of uptake, processing, and presentation of antigens by dendritic cells to $CD8^+$ T cells is called cross-presentation. As a result, T cells against the cancer antigens are primed, activated, and start to clonally expand. Co-stimulatory signals provided by the interaction between co-stimulatory molecules expressed on dendritic cells (e.g. B7.1 and B7.2) and co-stimulatory receptors on T cells (e.g. CD28) further fuel anti-tumor T cell responses. Activated $CD8^+$ T cells migrate to and infiltrate the tumor, where they can recognize cancer cells via the interaction between their T cell receptor (TCR) and the cognate antigen bound to HLA class I molecules on the surface of cancer cells. $CD8^+$ T cells can eliminate cancer cells by two main pathways: i) the release of cytolytic granules such as perforin and granzymes, and ii) cell

death pathway-mediated apoptosis by the Fas-Fas ligand or TRAIL-TRAIL-R1/2 pathways.[18] In addition, production of the pro-inflammatory cytokine IFN-γ, by CD8+ T cells among others, can increase the expression of HLA class I antigens by cancer cells, thereby enhancing their targeting and killing. Activated CD4+ T cells provide help to promote CD8+ T cell responses via the secretion of effector cytokines (e.g. IL-2) and can, under specific circumstances, exert cytotoxicity against cancer cells (e.g. via IFN-γ, TNF-α). To oppose continued amplification of T cell responses, immune checkpoints and inhibitors are in place that negatively regulate T cells.[17] Last, in addition to the innate and adaptive immune arms, a variety of immune cell populations have been discovered that are in between innate and adaptive immunity, also called unconventional T cells. Unconventional T cells include γδ T cells, NKT cells, and mucosal-associated invariant T (MAIT) cells, and often reside at mucosal tissues where they can rapidly respond in an innate-like fashion. Unlike conventional CD8+ and CD4+ T cells, most unconventional T cells do not recognize peptide antigens associated with HLA molecules, but a broad spectrum of non-polymorphic ligands ranging from metabolite antigens such as phosphoantigens that can be expressed by cancer cells, tumor-derived lipid antigens, to stress-induced molecules on the surface of cancer cells, among others.[19]

The immune system and cancer immune evasion

The immune system plays a crucial role in the protection against cancer development and progression. However, the generation of diversity in tumors may enable the emergence of cancer cell clones that can escape the recognition and elimination by the immune system.[20,21] Cancer cells can escape immune surveillance through different mechanisms. One of the most studied cancer immune evasion mechanisms involves the loss of cancer antigen expression through downregulation or loss of HLA class I expression.[22] This would preclude HLA class I-mediated antigen presentation, thus rendering these tumors insensitive to CD8+ T cell-mediated immunity. Furthermore, cancer cells can escape immune cell killing by enhanced resistance to the cytotoxic effects of anti-tumor immunity via the induction of anti-apoptotic mechanisms or the loss of pro-apoptotic factors.[23] Third, cancer immune evasion can occur through the development of an immunosuppressive cancer microenvironment via the secretion of immunosuppressive cytokines (e.g. TGF-β, IL-10, VEGF) by cancer cells and/or the recruitment of immune cell types with immunosuppressive functions such as regulatory T cells and subsets of myeloid cells.[24] Hence, the immune system plays a dichotomous role in the development and progression of cancer, where different cells can antagonize or promote carcinogenesis. The type, density, and spatial localization of tumor-infiltrating lymphocytes (TILs) are well-known determinants for the prognosis of cancer patients.[25,26] High densities of CD8+ T cells within the center of different types of tumors correlated with improved patient prognosis, whereas high densities of CD8+ T cells at the tumor margin had no effect on survival.[26-28] Such intratumoral immune responses were found to have enhanced prognostic value for patients with colorectal cancer (CRC) than previous pathological criteria for tumor staging.[29] In addition to CD8+ T cells, high levels of intratumoral γδ T cells and subsets of CD4+ T cells represented favorable prognostic signatures in a collection of different cancer types, while high levels of intratumoral myeloid cell populations (polymorphonuclear cells, eosinophils, macrophages) primarily correlated with poor survival.[30]

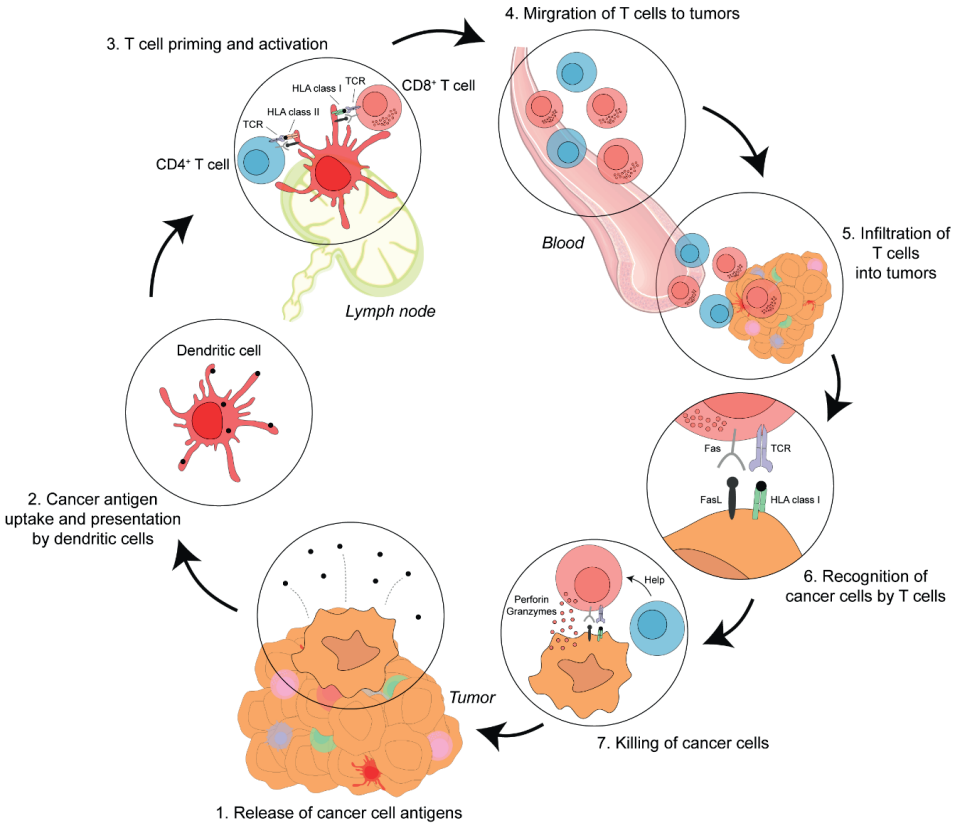


Figure 2. The generation of anti-tumor T cell responses.
A simplified overview of the main cellular players in the generation of anti-tumor T cell responses. Antigens released from dying cancer cells are taken up by antigen-presenting cells such as dendritic cells. Upon migration toward regional lymph nodes, dendritic cells activate naïve CD4⁺ (helper) and CD8⁺ (cytotoxic) T cells through interactions with antigen-HLA complexes on the surface of dendritic cells with the TCR on T cells (HLA class I-mediated antigen cross-presentation for CD8⁺ T cells and HLA class II-mediated antigen presentation for CD4⁺ T cells). The interaction of co-stimulatory molecules on the cell surface of dendritic cells with co-stimulatory receptors on T cells provides full activation. Upon priming and activation, CD8⁺ and CD4⁺ T cells migrate to and infiltrate the tumor. Here, CD8⁺ T cells can recognize cancer cells expressing the cognate antigen bound to HLA class I molecules on their surface, leading to the elimination of cancer cells via the release of perforin and granzymes or via the cell death pathway (e.g. Fas-FasL). CD4⁺ T cells provide help to promote CD8⁺ T cell responses via the secretion of effector cytokines (e.g. IL-2) and can, under specific circumstances, exert cytotoxicity against cancer cells (e.g. via IFN-γ, TNF-α). Adapted from Chen *et al.* (2013)[17].

The revolution of cancer immunotherapy

Remarkable progress in understanding cancer immunity has led to novel approaches for immune-based therapies. In contrast to chemotherapy, radiation, and targeted therapy that focus on eliminating the cancer cells themselves, cancer immunotherapy is aimed at stimulating the body's own immune system to recognize and eliminate the cancer cells or by engineering features of the immune system into therapeutic products. Cytotoxic T cells are viewed as central players in anti-tumor immunity, and T cell-mediated immunotherapies have emerged

as a successful treatment against cancer. The different steps involved in the generation of anti-tumor T cell responses provide a wide range of potential immunotherapeutic targets. Here we will discuss three main categories of cancer immunotherapeutic strategies, ranging from counteracting T cell inhibitory mechanisms to stimulating T cell effector mechanisms. First, cancer immunotherapies have been developed that intervene in the control of T cell activation by targeting negative regulators on T cells. T cells can express inhibitory immune checkpoints such as PD-1 and CTLA-4 that can bind to their ligands PD-L1/PD-L2 and B7.1/B7.2, respectively, on the surface of cancer cells (but also on other immune cell subsets).[31,32] Therapeutic blockade of these interactions with antibodies promotes the activation and expansion of T cells (**Figure 3**). This form of cancer immunotherapy is referred to as immune checkpoint blockade, and earned James P. Allison and Tasuku Honjo the Nobel Prize in Physiology or Medicine 2018. Immune checkpoint blockade therapies targeting the PD-1/PD-L1 and CTLA-4 axis are particularly effective in melanoma, non-small cell lung cancer, and DNA mismatch repair-deficient cancers.[33-37] These cancers are characterized by a high mutation burden, which increases the probability that neoantigens are presented in complex with HLA class I at the surface of cancer cells, allowing for enhanced immunogenicity.[38] As a consequence of their immunogenic character, these cancers generally show a dense infiltration by cytotoxic T cells, currently viewed as the main effector cell of immune checkpoint blockade-induced anti-tumor immunity.[39-41] Another immunotherapeutic strategy involves the modification of T cells to boost the recognition of cancer cells by adoptive cellular therapy with TILs, or genetically modified T cells expressing recombinant TCRs or chimeric antigen receptors (CARs). The most well-developed of such strategies is the use of CAR T cells. CAR T cell therapy involves the isolation of T cells from patients, which are genetically altered by transduction with a construct engineered to express CARs. CARs are composed of an extracellular antibody-binding domain, which can recognize cancer-specific antigens, fused to an intercellular TCR-signaling domain for T cell activation.[42] Further, CARs include co-stimulatory receptors to provide T cell co-stimulatory signals.[43] The modified T cells are subsequently infused back into the patient. CAR T cell therapy can yield substantial clinical benefit, mainly in B cell malignancies.[44] The last category, therapeutic cancer vaccines, primary promotes cancer antigen presentation. It is still a largely experimental therapy. Therapeutic cancer vaccines generally involve exogenous administration of selected cancer antigens, determined by sequencing of the expressed cancer genome, combined with adjuvants that activate dendritic cells.[45] The aim of therapeutic cancer vaccines is to activate or introduce cancer antigen-specific T cell responses.[46-48] Although successful responses can be achieved, therapeutic cancer vaccines generally fail to generate durable T cell responses.[45]

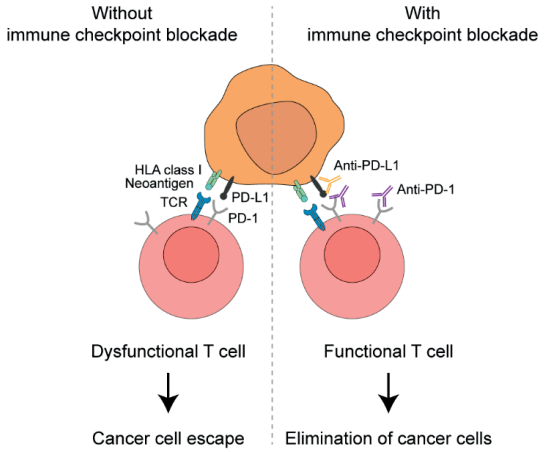


Figure 3. The effect of immune checkpoint blockade on anti-tumor T cell responses.
Inhibitory immune checkpoint protein PD-1 suppresses T cell response and function by engaging with its ligands (PD-L1, PD-L2) that can be expressed by cancer cells but also other immune cell subsets. Therapeutic antibodies directed against immune checkpoint molecule PD-1, or its ligands, block the interaction of PD-1 with PD-L1/PD-L2. As a result, the inhibitory signal is blocked and functional T cells can eliminate cancer cells upon the recognition of cancer antigens presented by HLA class I molecules on the surface of cancer cells with their TCR, accompanied by appropriate costimulatory signals.

The studies described in this thesis

This thesis has studied the tumor microenvironment of CRC and pancreatic ductal adenocarcinoma (PDAC). CRC is the third most frequently diagnosed cancer type worldwide.[49] Currently, more than 1.9 million new cases of CRC per year arise throughout the world, with an estimated 935.000 deaths annually.[49] The five-year survival rate is approximately 65%, but the prognosis of patients largely depends on the tumor stage at the time of diagnosis.[50] CRC is a heterogeneous disease with different pathways of carcinogenesis. One of the pathways of genomic instability involved in the pathogenesis of CRC is microsatellite instability-high (MSI-H). This phenotype is caused by a deficient DNA mismatch repair (MMR) system, involving MMR proteins MLH1, MSH2, MLH6, and PMS2.[51,52] MMR-deficiency occurs in approximately 15-20% of CRCs. As a consequence of the deficient MMR system in these tumors, somatic mutations (insertions, deletions) often accumulate at DNA microsatellite sequences. MMR-deficient cancers are characterized by the presence of numerous TILs.[53,54] The majority of CRCs are MMR-proficient (or microsatellite stable (MSS)), with chromosomal instability (CIN) as the most common form of genetic instability. The CIN pathway is characterized by structural chromosomal aberrations.[55] As compared to MMR-deficient tumors, MMR-proficient tumors generally show a lower mutation burden and less dense infiltration by immune cells.[56,57] The type of genomic instability observed in CRC is known to correlate with the clinical prognosis of patients, where MMR-deficient cancers show a reduced risk of recurrence and improved disease-free survival as compared to MMR-proficient cancers.[56] In addition, the type of genomic instability in CRC is a determining factor for the response to immunotherapy. Immune checkpoint blockade therapy, such as PD-1

blockade, has proven to be effective at reinvigorating T cell-mediated immune responses in patients with MMR-deficient CRC, while MMR-proficient CRCs are largely unresponsive to this therapy.[37,58-60] However, only a minority of MMR-deficient CRCs respond to immune checkpoint blockade. Immune evasion mechanisms counteracting T cell-mediated immune responses are common in MMR-deficient CRCs. For instance, loss of HLA class I expression is described to occur in the majority (78%) of MMR-deficient tumors.[61] In theory, HLA class I-mediated antigen presentation would be an essential component for the clinical activity of $CD8^+$ T cell-based immunotherapies. Surprisingly, the majority of HLA class I-negative MMR-deficient CRCs show durable clinical responses to PD-1 blockade,[62] suggesting that immune cell subsets other than $CD8^+$ T cells may contribute to these responses. Here, different types of innate and unconventional immune cells may come into play as they are capable of HLA class I-unrestricted tumor killing.

PDAC is one of the most lethal cancer types in the industrialized world, with a five-year survival rate of 10%.[63] The majority of PDAC patients present with locally advanced or metastatic disease, with only around 20% of patients eligible for surgical resection of their tumor.[64] Because of its poor prognosis, pancreatic cancer accounts for nearly as many deaths (466.000) as cases (496.000) per year worldwide.[49] The accumulation of genetic alterations in oncogenes and tumor suppressor genes drives the stepwise progression through non-invasive precursor lesions, including pancreatic intraepithelial neoplasia (i.e. PanIN), intraductal papillary mucinous neoplasia (i.e. IPMN), and mucinous cystic neoplasms (i.e. MCN), to invasive cancer.[64] PDAC is non-immunogenic and characterized by a lack of naturally occurring immune responses due to a generally low mutation burden as well as a large stromal compartment consisting of few and mostly immunosuppressive immune cells.[65] Defects in DNA repair systems such as MMR-deficiency are infrequent in PDAC, occurring in approximately 1% of cases.[66] Although patients with higher levels of infiltration by $CD8^+$ and/or $CD4^+$ T cells have significantly improved prognosis,[67] the numbers of tumor-infiltrating T cells are generally low and the cells mainly locate in the stromal compartment of the PDAC microenvironment.[68] Studies have reported minimal clinical benefit in response to immune checkpoint blockade in PDAC patients, with the exception of MMR-deficient PDAC tumors.[69,70] The low mutation burden, low numbers of cytotoxic T cells, large stromal compartment, and immunosuppressive immune cells may render pancreatic cancer insensitive to immunotherapy. In contrast to CRC, mutations in HLA, β2m, or other components of the antigen presentation machinery are not commonly found in PDAC.[71] Currently, combination therapies with immune checkpoint blockade are being examined that seek to enhance anti-tumor immunity.[72]

The need for high-dimensional analyses of the cancer microenvironment

The cancer immune microenvironment plays a critical role in the course of natural disease progression and the clinical prognosis of patients, but can also determine response to cancer immunotherapy. With the emergence of cancer immunotherapies, significant advances have been made that changed the clinical management of cancer treatment for a significant number of patients. However, durable clinical responses to cancer immunotherapy

treatment are observed in only a minority of patients, and the development of resistance to therapy remains an important complication for advanced cancer patients. To understand the mechanisms that are at play in the cancer immune microenvironment affecting responses to current immunotherapies, and for the development of novel, alternative immunotherapeutic strategies, it is crucial to characterize the cancer microenvironment with high-dimensional approaches (**Figure 4**). By simultaneously dissecting immune cell populations across multiple lineages, high-dimensional approaches allow the identification and characterization of intratumoral innate and adaptive immune landscapes and may discriminate immune cell subsets that can be exploited in an immunotherapeutic setting.

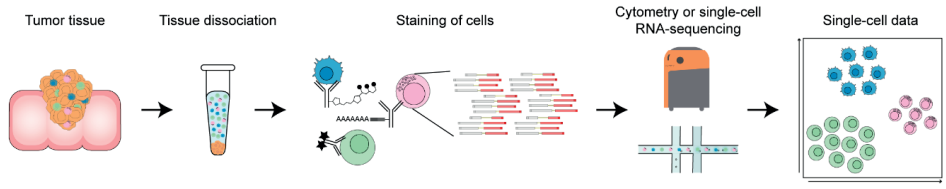


Figure 4. Immunophenotyping of tumor samples with high-dimensional approaches.
Tumor tissue samples are first dissociated into single-cell suspensions, and enriched for immune cells by density-gradient centrifugation. Single-cell suspensions are subsequently stained with a cocktail of antibodies, against different immune cell markers, that can be coupled to fluorochromes, heavy metal isotopes, or oligonucleotides. Stained cells are measured by flow cytometry, mass cytometry, or single-cell RNA-sequencing. Clustering of the obtained single-cell data based on their complete immune cell profile reveals the variety of immune cell types present in the cancer microenvironment.

OUTLINE OF THIS THESIS

In **Chapter 2**, we discuss the heterogeneity of cancer and the need for multidimensional analyses of the cancer microenvironment. We describe cutting-edge multidimensional single-cell technologies that have been paramount in studying the cancer microenvironment, and their strengths and weaknesses. We deliberate on potential avenues for the integration of multi-omics data, and how such integrated data will help to understand the complex role of the microenvironment in cancer.

In **Chapter 3**, we provide a blueprint of innate and adaptive immune cell populations in tumor and non-malignant tissues of patients with CRC analyzed by single-cell mass cytometry. We discover novel mediators of anti-tumor immunity in colorectal tumors, including PD-1$^+$ γδ T cells and a previously unappreciated ILC1-like population. We show that both immune subsets displayed cytotoxic activity and were particularly frequent in MMR-deficient CRCs. In addition, we find tumor-resident CD8$^+$ and CD4$^+$ T cell populations with highly similar activated (CD38$^+$PD-1$^+$) tissue-resident (CD103$^+$) phenotypes that are infrequent in non-malignant tissues. These observations formed the basis for new research questions that are studied in **Chapter 4**, **Chapter 5**, and **Chapter 6**.

In **Chapter 4**, we investigate the involvement of PD-1$^+$ γδ T cells in anti-tumor immune responses and as effectors of immune checkpoint blockade therapy in MMR-deficient colon cancers with loss of HLA class I-mediated antigen presentation. We apply a combination of transcriptomic and imaging approaches for an in-depth analysis of MMR-deficient tumor samples before and after immune checkpoint blockade, accompanied by *in vitro* functional studies, to provide evidence indicating that γδ T cells mediate responses to HLA class I-negative, MMR-deficient tumors during immune checkpoint blockade therapy.

In **Chapter 5**, we explore the full spectrum of ILC subsets and their functional differences in MMR-deficient CRCs by single-cell RNA-sequencing with mass- and flow cytometric characterization of immunophenotypic markers and transcription factors. We demonstrate that the majority of ILCs in MMR-deficient CRCs consist of CD127$^-$CD103$^+$ ILC1-like cells, and provide evidence for an active role for these cells in the recognition and cytotoxicity against MMR-deficient cancers.

In **Chapter 6**, we study TCR repertoires of distinct CD8$^+$ and CD4$^+$ T cell populations in colorectal tumors, defined by the expression of markers of activation (CD38, PD-1) and tissue-residency (CD103) along with their non-activated and non-tissue-resident counterparts. By comparing those TCR repertoires to T cells from adjacent healthy tissues, pericolic lymph nodes, and peripheral blood, we provide insights into the origin of tissue-resident activated (CD103$^+$CD38$^+$PD-1$^+$) CD8$^+$ and CD4$^+$ T cells and their clonal enrichment, a characteristic of antigenic responses.
In **Chapter 7**, we investigate the immune composition in tumors and non-malignant tissues of patients with PDAC by single-cell mass cytometry. We describe a suppressive immune landscape of PDAC, deprived of tissue-resident memory CD8$^+$ T cells with cytotoxic potential, and remarkable increased frequencies of B cells and regulatory T cells as compared to non-malignant pancreatic tissue. In addition, we find a tumor tissue-specific ILC1-like population resembling the one described in CRC (**Chapter 3** and **Chapter 5**). To some extent, the immune cell composition of PDAC could be reflected in portal vein blood, suggesting a regional enrichment of immune cells involved in the anti-tumor immune response.

In **Chapter 8**, we discuss the main findings of this thesis in a broader perspective and provide implications for future research and clinical practice.

REFERENCES

1 Boveri, T. *Zur frage der entstehung maligner tumoren*. (Gustav Fischer, 1914).

2 Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* 100, 57-70, doi:10.1016/s0092-8674(00)81683-9 (2000).

3 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* 144, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).

4 Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov* 12, 31-46, doi:10.1158/2159-8290.Cd-21-1059 (2022).

5 Burnet, F. M. The concept of immunological surveillance. *Prog Exp Tumor Res* 13, 1-27, doi:10.1159/000386035 (1970).

6 Penn, I. Malignant Tumors in Organ Transplant Recipients. *Springer-Verlag, New York* (1970).

7 Gatti, R. A. & Good, R. A. Occurrence of malignancy in immunodeficiency diseases. A literature review. *Cancer* 28, 89-98, doi:10.1002/1097-0142(197107)28:1<89::aid-cncr2820280117>3.0.co;2-q (1971).

8 Dighe, A. S., Richards, E., Old, L. J. & Schreiber, R. D. Enhanced in vivo growth and resistance to rejection of tumor cells expressing dominant negative IFN gamma receptors. *Immunity* 1, 447-456, doi:10.1016/1074-7613(94)90087-6 (1994).

9 Kaplan, D. H. *et al.* Demonstration of an interferon gamma-dependent tumor surveillance system in immunocompetent mice. *Proc Natl Acad Sci U S A* 95, 7556-7561, doi:10.1073/pnas.95.13.7556 (1998).

10 Shankaran, V. *et al.* IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature* 410, 1107-1111, doi:10.1038/35074122 (2001).

11 Lanier, L. L. NK cell recognition. *Annu Rev Immunol* 23, 225-274, doi:10.1146/annurev.immunol.23.021704.115526 (2005).

12 Kärre, K., Ljunggren, H. G., Piontek, G. & Kiessling, R. Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. *Nature* 319, 675-678, doi:10.1038/319675a0 (1986).

13 Engblom, C., Pfirschke, C. & Pittet, M. J. The role of myeloid cells in cancer therapies. *Nat Rev Cancer* 16, 447-462, doi:10.1038/nrc.2016.54 (2016).

14 Dranoff, G. Cytokines in cancer pathogenesis and cancer therapy. *Nat Rev Cancer* 4, 11-22, doi:10.1038/nrc1252 (2004).

15 Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* 348, 69-74, doi:10.1126/science.aaa4971 (2015).

16 Ilyas, S. & Yang, J. C. Landscape of Tumor Antigens in T Cell Immunotherapy. *J Immunol* 195, 5117-5122, doi:10.4049/jimmunol.1501657 (2015).

17 Chen, D. S. & Mellman, I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* 39, 1-10, doi:10.1016/j.immuni.2013.07.012 (2013).

18 Martínez-Lostao, L., Anel, A. & Pardo, J. How Do Cytotoxic Lymphocytes Kill Cancer Cells? *Clin Cancer Res* 21, 5047-5056, doi:10.1158/1078-0432.Ccr-15-0685 (2015).

19 Mayassi, T., Barreiro, L. B., Rossjohn, J. & Jabri, B. A multilayered immune system through the lens of unconventional T cells. *Nature* 595, 501-510, doi:10.1038/s41586-021-03578-0 (2021).

20 Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol* 3, 991-998, doi:10.1038/ni1102-991 (2002).

21 Schreiber, R. D., Old, L. J. & Smyth, M. J. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science (New York, N.Y.)* 331, 1565-1570, doi:10.1126/science.1203486 (2011).

22 Campoli, M., Chang, C. C. & Ferrone, S. HLA class I antigen loss, tumor immune escape and immune selection. *Vaccine* 20 Suppl 4, A40-45, doi:10.1016/s0264-410x(02)00386-9 (2002).

23 Andersen, M. H., Reker, S., Kvistborg, P., Becker, J. C. & thor Straten, P. Spontaneous immunity against Bcl-xL in cancer patients. *J Immunol* 175, 2709-2714, doi:10.4049/jimmunol.175.4.2709 (2005).

24 Vesely, M. D., Kershaw, M. H., Schreiber, R. D. & Smyth, M. J. Natural innate and adaptive immunity to cancer. *Annu Rev Immunol* 29, 235-271, doi:10.1146/annurev-immunol-031210-101324 (2011).

25 Binnewies, M. *et al.* Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med* 24, 541-550, doi:10.1038/s41591-018-0014-x (2018).

26 Fridman, W. H., Zitvogel, L., Sautes-Fridman, C. & Kroemer, G. The immune contexture in cancer prognosis and treatment. *Nature reviews. Clinical oncology* 14, 717-734, doi:10.1038/nrclinonc.2017.101 (2017).

27 Naito, Y. *et al.* CD8+ T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer. *Cancer Res* 58, 3491-3494 (1998).

28 Fridman, W. H., Pages, F., Sautes-Fridman, C. & Galon, J. The immune contexture in human

tumours: impact on clinical outcome. *Nat Rev Cancer* 12, 298-306, doi:10.1038/nrc3245 (2012).

29 Galon, J. *et al.* Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 313, 1960-1964, doi:10.1126/science.1129139 (2006).

30 Gentles, A. J. *et al.* The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med* 21, 938-945, doi:10.1038/nm.3909 (2015).

31 Leach, D. R., Krummel, M. F. & Allison, J. P. Enhancement of antitumor immunity by CTLA-4 blockade. *Science* 271, 1734-1736, doi:10.1126/science.271.5256.1734 (1996).

32 Okazaki, T. & Honjo, T. PD-1 and PD-1 ligands: from discovery to clinical application. *Int Immunol* 19, 813-824, doi:10.1093/intimm/dxm057 (2007).

33 Hodi, F. S. *et al.* Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 363, 711-723, doi:10.1056/NEJMoa1003466 (2010).

34 Topalian, S. L. *et al.* Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med* 366, 2443-2454, doi:10.1056/NEJMoa1200690 (2012).

35 Rizvi, N. A. *et al.* Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124-128, doi:10.1126/science.aaa1348 (2015).

36 Kelderman, S., Schumacher, T. N. & Kvistborg, P. Mismatch Repair-Deficient Cancers Are Targets for Anti-PD-1 Therapy. *Cancer Cell* 28, 11-13, doi:10.1016/j.ccell.2015.06.012 (2015).

37 Le, D. T. *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409-413, doi:10.1126/science.aan6733 (2017).

38 Germano, G. *et al.* Inactivation of DNA repair triggers neoantigen generation and impairs tumour growth. *Nature* 552, 116-120, doi:10.1038/nature24673 (2017).

39 Dolcetti, R. *et al.* High prevalence of activated intraepithelial cytotoxic T lymphocytes and increased neoplastic cell apoptosis in colorectal carcinomas with microsatellite instability. *Am J Pathol* 154, 1805-1813, doi:10.1016/s0002-9440(10)65436-3 (1999).

40 Tumeh, P. C. *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 515, 568-571, doi:10.1038/nature13954 (2014).

41 Taube, J. M. *et al.* Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res* 20, 5064-5074, doi:10.1158/1078-0432.Ccr-13-3271 (2014).

42 Eshhar, Z., Waks, T., Gross, G. & Schindler, D. G. Specific activation and targeting of cytotoxic lymphocytes through chimeric single chains consisting of antibody-binding domains and the gamma or zeta subunits of the immunoglobulin and T-cell receptors. *Proc Natl Acad Sci U S A* 90, 720-724, doi:10.1073/pnas.90.2.720 (1993).

43 Sadelain, M., Rivière, I. & Riddell, S. Therapeutic T cell engineering. *Nature* 545, 423-431, doi:10.1038/nature22395 (2017).

44 Park, J. H. *et al.* Long-Term Follow-up of CD19 CAR Therapy in Acute Lymphoblastic Leukemia. *N Engl J Med* 378, 449-459, doi:10.1056/NEJMoa1709919 (2018).

45 Saxena, M., van der Burg, S. H., Melief, C. J. M. & Bhardwaj, N. Therapeutic cancer vaccines. *Nat Rev Cancer* 21, 360-378, doi:10.1038/s41568-021-00346-0 (2021).

46 Gubin, M. M. *et al.* Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* 515, 577-581, doi:10.1038/nature13988 (2014).

47 Carreno, B. M. *et al.* Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. *Science* 348, 803-808, doi:10.1126/science.aaa3828 (2015).

48 Ott, P. A. *et al.* An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* 547, 217-221, doi:10.1038/nature22991 (2017).

49 Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 71, 209-249, doi:10.3322/caac.21660 (2021).

50 Rawla, P., Sunkara, T. & Barsouk, A. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. *Prz Gastroenterol* 14, 89-103, doi:10.5114/pg.2018.81072 (2019).

51 Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D. & Perucho, M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363, 558-561, doi:10.1038/363558a0 (1993).

52 Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487, 330-337, doi:10.1038/nature11252 (2012).

53 Smyrk, T. C., Watson, P., Kaul, K. & Lynch, H. T. Tumor-infiltrating lymphocytes are a marker for microsatellite instability in colorectal carcinoma. *Cancer* 91, 2417-2422 (2001).

54 de Miranda, N. F. *et al.* Infiltration of Lynch colorectal cancers by activated immune cells associates with early staging of the primary tumor and absence of lymph node metastases. *Clin Cancer Res* 18, 1237-1245, doi:10.1158/1078-0432.CCR-11-1997 (2012).

55 Pino, M. S. & Chung, D. C. The chromosomal instability pathway in colon cancer. *Gastroenterology* 138, 2059-2072, doi:10.1053/j.gastro.2009.12.065 (2010).

56 Domingo, E. *et al.* Somatic POLE proofreading domain mutation, immune response, and prognosis in colorectal cancer: a retrospective, pooled biomarker study. *The lancet. Gastroenterology & hepatology* 1, 207-216, doi:10.1016/s2468-1253(16)30014-0 (2016).

57 Giannakis, M. *et al.* Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell Rep* 15, 857-865, doi:10.1016/j.celrep.2016.03.075 (2016).

58 Overman, M. J. *et al.* Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *The Lancet. Oncology* 18, 1182-1191, doi:10.1016/s1470-2045(17)30422-9 (2017).

59 Overman, M. J. *et al.* Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 36, 773-779, doi:10.1200/jco.2017.76.9901 (2018).

60 Chalabi, M. *et al.* Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat Med* 26, 566-576, doi:10.1038/s41591-020-0805-8 (2020).

61 Ijsselsteijn, M. E. *et al.* Revisiting immune escape in colorectal cancer in the era of immunotherapy. *Br J Cancer* 120, 815-818, doi:10.1038/s41416-019-0421-x (2019).

62 Middha, S. *et al.* Majority of B2M-Mutant and -Deficient Colorectal Carcinomas Achieve Clinical Benefit From Immune Checkpoint Inhibitor Therapy and Are Microsatellite Instability-High. *JCO precision oncology* 3, doi:10.1200/po.18.00321 (2019).

63 Rawla, P., Sunkara, T. & Gaduputi, V. Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. *World J Oncol* 10, 10-27, doi:10.14740/wjon1166 (2019).

64 Kamisawa, T., Wood, L. D., Itoi, T. & Takaori, K. Pancreatic cancer. *Lancet* 388, 73-85, doi:10.1016/s0140-6736(16)00141-0 (2016).

65 Feig, C. *et al.* The pancreas cancer microenvironment. *Clin Cancer Res* 18, 4266-4276, doi:10.1158/1078-0432.Ccr-11-3114 (2012).

66 Hu, Z. I. *et al.* Evaluating Mismatch Repair Deficiency in Pancreatic Adenocarcinoma: Challenges and Recommendations. *Clin Cancer Res* 24, 1326-1336, doi:10.1158/1078-0432.Ccr-17-3099 (2018).

67 Ino, Y. *et al.* Immune cell infiltration as an indicator of the immune microenvironment of pancreatic cancer. *Br J Cancer* 108, 914-923, doi:10.1038/bjc.2013.32 (2013).

68 Blando, J. *et al.* Comparison of immune infiltrates in melanoma and pancreatic cancer highlights VISTA as a potential target in pancreatic cancer. *Proc Natl Acad Sci U S A* 116, 1692-1697, doi:10.1073/pnas.1811067116 (2019).

69 Royal, R. E. *et al.* Phase 2 trial of single agent Ipilimumab (anti-CTLA-4) for locally advanced or metastatic pancreatic adenocarcinoma. *J Immunother* 33, 828-833, doi:10.1097/CJI.0b013e3181eec14c (2010).

70 Brahmer, J. R. *et al.* Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med* 366, 2455-2465, doi:10.1056/NEJMoa1200694 (2012).

71 Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* 32, 185-203.e113, doi:10.1016/j.ccell.2017.07.007 (2017).

72 Brouwer, T. P., Vahrmeijer, A. L. & de Miranda, N. Immunotherapy for pancreatic cancer: chasing the light at the end of the tunnel. *Cell Oncol (Dordr)* 44, 261-278, doi:10.1007/s13402-021-00587-z (2021).

Unraveling the complexity of the cancer microenvironment with multidimensional genomic and cytometric technologies

2

Natasja L. de Vries[1,2], Ahmed Mahfouz[3,4,5], Frits Koning[2], Noel F.C.C. de Miranda[1]

[1]Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands. [2]Department of Immunology, Leiden University Medical Center, Leiden, the Netherlands. [3]Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands. [4]Delft Bioinformatics Lab, Delft University of Technology, Delft, the Netherlands. [5]Leiden Computational Biology Center, Leiden University Medical Center, Leiden, the Netherlands.

Front Oncol. 2020; 10: 1254.
doi: 10.3389/fonc.2020.01254

ABSTRACT

Cancers are characterized by extensive heterogeneity that occurs intratumorally, between lesions, and across patients. To study cancer as a complex biological system, multidimensional analyses of the tumor microenvironment are paramount. Single-cell technologies such as flow cytometry, mass cytometry, or single-cell RNA-sequencing have revolutionized our ability to characterize individual cells in great detail and, with that, shed light on the complexity of cancer microenvironments. However, a key limitation of these single-cell technologies is the lack of information on spatial context and multicellular interactions. Investigating spatial contexts of cells requires the incorporation of tissue-based techniques such as multiparameter immunofluorescence, imaging mass cytometry, or *in situ* detection of transcripts. In this Review, we describe the rise of multidimensional single-cell technologies and provide an overview of their strengths and weaknesses. In addition, we discuss the integration of transcriptomic, genomic, epigenomic, proteomic, and spatially-resolved data in the context of human cancers. Lastly, we will deliberate on how the integration of multi-omics data will help to shed light on the complex role of cell types present within the human tumor microenvironment, and how such system-wide approaches may pave the way toward more effective therapies for the treatment of cancer.

INTRODUCTION

Heterogeneity of cancer and need for multidimensional approaches

A genetic basis for cancer development was first proposed by the German zoologist Theodor Boveri who speculated that malignant tumors might be the result of abnormal chromosome alterations in cells.[1] By then, a cancer cell-centric vision dominated, where tumorigenesis was thought to be exclusively driven by multistep alterations in cellular genomes. During the last decades, however, it has become increasingly apparent that the study of cancers must also encompass other constituents of the cancer microenvironment including immune cells, fibroblasts, and other stromal components, to capture all aspects of cancer biology.[2] The immune system, for example, plays a dichotomous role in cancer development and progression, as different cells can antagonize or promote tumorigenesis.[3] The mapping and understanding of the interplay between cancer cells and other constituents of the cancer microenvironment is thus fundamental for the clinical management of this disease.

The study of cancers as complex systems is further complicated by cancer heterogeneity that can occur at different levels; intratumorally, between lesions, and across patients. Intratumoral heterogeneity involves the near-stochastic generation of both genetic (e.g. mutations, chromosomal aberrations) and epigenetic (e.g. DNA methylation, chromatin remodeling, post-translational modification of histones) modifications. Within tumors, distinct niches can favor the outgrowth of different cancer cell clones that acquired characteristics compatible with regional microenvironments (e.g. nutrient and oxygen availability, exposure to immune cells). Other intrinsic sources of heterogeneity such as self-renewal of cancer cells and cell differentiation processes contribute further to intratumoral heterogeneity.[4,5] In addition, the immune system is a major part of the tumor microenvironment and contains many different types of adaptive (e.g. $CD4^+$ and $CD8^+$ T lymphocytes) and innate (e.g. macrophages and innate lymphoid cells) immune cells that also contribute to cancer heterogeneity.[6] Their location within a tumor has been shown to significantly impact their anti- or pro-tumorigenic effects.[7] In addition, the density of immune cell infiltration in tumors is a well-known determinant for the prognosis of cancer patients.[8] Inter-lesional heterogeneity can be observed between multiple primary tumors at time of diagnosis, between a primary tumor and metastasis, and between different metastatic niches in individual patients. They can be a result of the outgrowth of subclones that can be (epi)genetically distinguished by mutations or structural variations.[9] Moreover, the structure and composition of the cancer microenvironment can vary between the primary tumor and metastases. Upon extravasation, cancer cells from primary tumors are exposed to different types of immune cells, stromal cells, platelets, and metabolic stress, and have to adapt to the new tissue microenvironment. As such, the metastatic tissue ("soil") plays a critical role in regulating the growth of metastases ("seed").[10] Finally, interpatient heterogeneity is, on top of the aforementioned variables, also fueled by distinct germline genetic backgrounds and environmental and stochastic factors that can affect cancer progression but also immunity. Major challenges in the field of cancer research are the identification of predictive biomarkers to select patients that are likely to respond to specific treatments, detection of mechanisms of resistance to therapy, and the development of novel treatments to improve cancer survival.

Here, we review the rise of cutting-edge multidimensional technologies such as spectral flow cytometry, multiparameter immunofluorescence, (imaging) mass cytometry, single-cell RNA-sequencing (scRNA-seq), and RNA spatial profiling that may play a crucial role to address the former problems. We will discuss how multi-omics data of dissociated cells as well as of spatial data can be obtained (**Figure 1A**) and the importance of integrating them to reveal the full cellular landscape of the cancer microenvironment (**Figure 1B,C**). For example, single-cell data of dissociated cells can be used as guide for cell type identification in spatial data[11] and, vice versa, spatial data can be used to predict the location of dissociated cells based on the similarity of their expression profiles to spatially-mapped data[12-14] (**Figure 1B**). In addition, mapping can be used to predict the spatial profile of genes or proteins which have not been experimentally measured to expand the coverage of spatial data[15-17] (**Figure 1C**).

Multidimensional single-cell technologies and their strengths and weaknesses

Single-cell DNA- and RNA-sequencing

Next-generation sequencing (NGS) approaches have revolutionized our ability to generate high-throughput genomic data where individual RNA and DNA molecules are represented by sequencing reads thereby retaining information on genotypes, phenotypes, cellular states, and sub-clonal alterations. Traditional molecular profiling has, until recently, largely relied on the analysis of bulk cell populations. Deep sequencing of DNA and RNA from tissues enabled reconstruction of "average" genomes and "average" transcriptomes that could then be deconstructed by employing bioinformatic algorithms to perform clonal evolution analysis or determine the composition of cancer microenvironments.[18-21] For an unbiased and systematic characterization of cells, high-throughput single-cell DNA- and RNA-sequencing have emerged as powerful tools. With single-cell DNA-sequencing, the genomic heterogeneity of tissues can be explored in detail. It can be used to detect nucleotide variations and chromosomal copy number alterations as well as more complex genomic rearrangements and cellular fractions. Single-cell genome sequencing involves whole-genome amplification of single cells, of which the three main methods are MDA[22], MALBAC[23], and DOP-PCR[24]. In 2011, the first study of DNA-sequencing of human breast cancer single cells was published[25], which was followed by many single-cell studies charting genetic heterogeneity within individual tumors as well as between primary tumors and their metastases, thereby allowing for a detailed understanding of the evolution processes occurring in a tumor. Single-cell DNA-sequencing has myriad applications in cancer research including examining intratumoral heterogeneity[26-28], investigating clonal evolution during tumorigenic processes[25,29-32], tracing metastatic dissemination[33], genomic profiling of circulating tumor cells[34-36], measuring mutation rates[37], and gain insight into resistance to therapy[38]. By defining, in detail, the genetic composition of tumors, the rationalization of targeted cancer therapies is made possible. However, drawbacks of single-cell DNA-sequencing methods are non-uniform coverage and allelic dropout events as well as artifacts introduced during genomic amplification, all of which contribute to a high rate of false negative and false positive findings.[39]

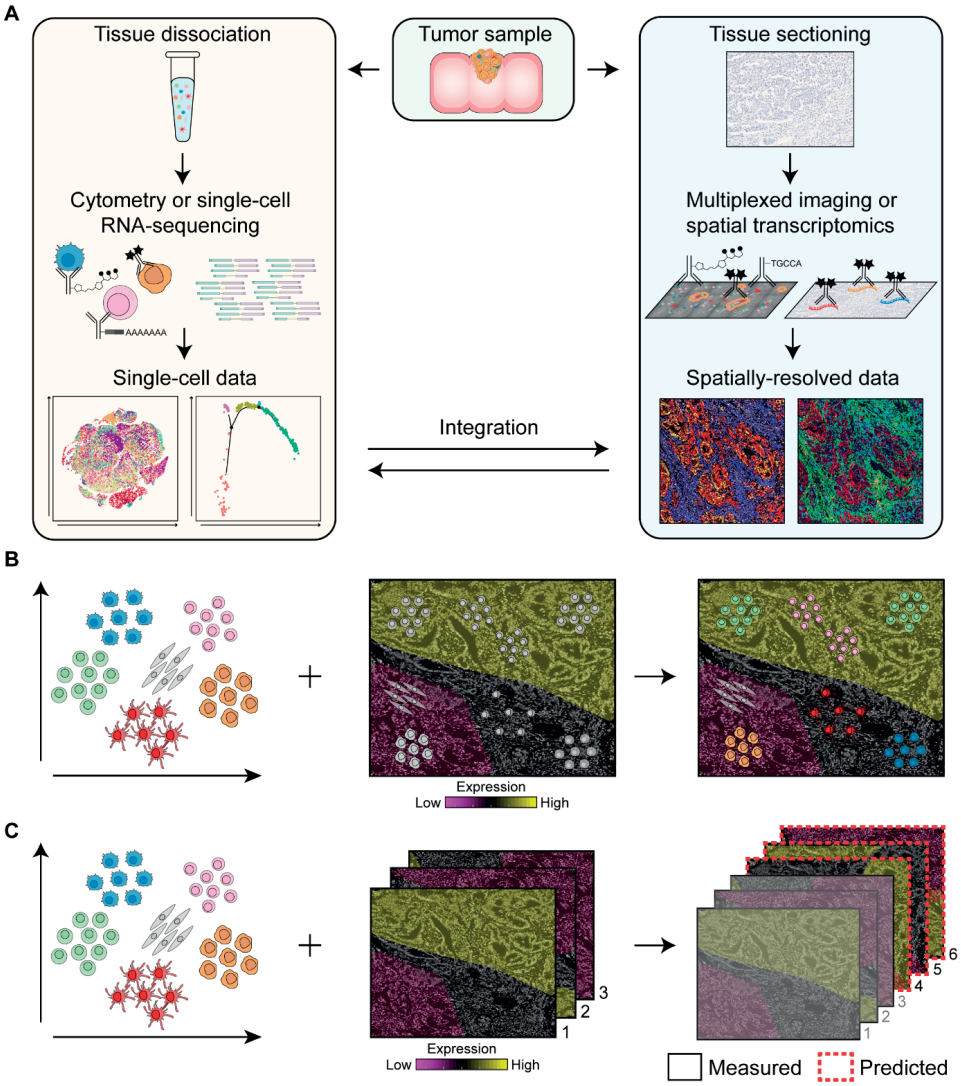


Figure 1. Overview of the pipeline for the integration of single-cell data of dissociated cells and spatially-resolved data.

A. Single-cell data can be obtained by flow and mass cytometry that make use of antibodies coupled to fluorochromes or heavy metal isotopes, respectively, for the immunodetection of dissociated cells. For single-cell RNA-sequencing, antibodies coupled to oligonucleotides can be used to simultaneously retrieve information on protein and RNA expression of single cells. Spatially-resolved data can be obtained by multiplexed imaging or spatial transcriptomics by immunodetection of tissue sections with antibodies coupled to fluorochromes, heavy metal isotopes or oligonucleotides. Integration of single-cell data of dissociated cells with spatially-resolved data will reveal the full cellular landscape of the cancer microenvironment. **B and C.** Integration approaches for single-cell data of dissociated cells and spatially-resolved data. Single-cell data of dissociated cells can be used as guide for cell type identification in spatial data and, vice versa, spatially-resolved data can be used to predict the location of dissociated cells based on the similarity of their expression profiles to spatially-mapped data (**B**). In addition, single-cell data can be used to predict the spatial profile of genes or proteins in the samples that have not been measured to expand the coverage of spatial data (**C**). Based on samples that have been measured (i.e. sample 1, 2, and 3), the expression of genes or proteins in sample 4, 5, and 6 can be predicted.

The first single-cell RNA-sequencing (scRNA-seq) experiment was published in 2009 by Tang and colleagues who profiled the transcriptome of a single cell from early embryonic development.[40] Rapid technological advances resulted in an exponential increase in the number of cells that can be studied by scRNA-seq analyses.[41] Just eight years later, 10x Genomics published a scRNA-seq dataset of more than one million individual cells from embryonic mice brains.[42] There are many different scRNA-seq library preparation platforms, which can be categorized into plate-based, droplet-based, and microwell-based.[41] The selection of the method depends on the research question, the number of input cells, the sequencing depth, the need for full-length coverage of transcriptomes, and costs, among others (reviewed by [43,44]). ScRNA-seq has demonstrated to be a powerful technique to decipher cancer biology. In 2012, Ramskold and colleagues applied scRNA-seq to study circulating tumor cells in melanoma, and could identify potential biomarkers for melanoma as well as SNPs and mutations in this relatively rare circulating tumor cell population.[45] Thereafter, scRNA-seq has been used to study the microenvironment of several cancer types including prostate cancer[46], breast cancer[47], glioma[48-50], renal cancer[51], lung cancer[52], melanoma[53-56], colorectal cancer[57-59], pancreatic ductal adenocarcinoma[60], liver cancer[61], head and neck cancer[62], leukemia[63], and glioma[64]. A pioneering study that applied scRNA-seq to primary glioblastomas uncovered inherent variability in oncogenic signaling, proliferation, immune responses, and regulators of stemness across cells sorted from five tumors.[48] However, this study was restricted to cancer cells and did not further investigate other cell types of the cancer microenvironment. Subsequently, another scRNA-seq study examined distinct genotypic and phenotypic states of malignant, immune, stromal, and endothelial cells of melanomas from 19 patients.[53] They identified cell states linked to resistance to targeted therapy, interactions between stromal factors and immune cell abundance, and potential biomarkers for distinguishing dysfunctional and cytotoxic T cells. A recent study in colorectal cancer broadened such scRNA-seq analysis by including a comparison of primary tumors to matched normal mucosa samples.[58] By projecting their scRNA-seq data to a large reference panel, the authors identified distinct subtypes of cancer-associated fibroblasts and new expression signatures that were predictive of prognosis in colorectal cancer. Further, scRNA-seq has been applied to investigate changes in the tumor microenvironment of cancer patients treated with immune checkpoint blockade to find signatures associated with positive responses to this therapy.[65,66]

Currently scRNA-seq can be combined with sequencing of T cell receptor and immunoglobulin repertoires thereby allowing to connect information of B and T cell specificity and phenotype. High-throughput single-cell B cell receptor sequencing of more than 250,000 B cells from different species has recently been pioneered to obtain paired antibody heavy- and light chain information at the single-cell level, and revealed a rapid discovery of antigen-reactive antibody candidates.[67] By a novel approach called RAGE-seq (Repertoire and Gene Expression by Sequencing), gene expression profiles can be paired with targeted full-length mRNA transcripts providing BCR and TCR sequences.[68] This method has been applied to study cells from the primary tumor and tumor-associated lymph node of a breast cancer patient and demonstrated the ability to track clonally related lymphocytes across tissues and link TCR and BCR clonotypes with gene-expression features.[68] A limitation of scRNA-seq is that RNA

levels are not fully representative of protein amounts. The advent of CITE-seq, REAP-seq, and Abseq overcame this limitation by enabling simultaneous detection of gene expression and protein levels in single cells by combining oligonucleotide-labeled antibodies against cell surface proteins with transcriptome profiling of thousands of single cells in parallel.[69-71] scRNA-seq, when employed in a discovery setting, can inform on the best markers to be used for the study of specific populations by complementary technologies such as flow or mass cytometry. However, aspects of sample preparation and handling have been shown to induce significant alterations in the transcriptome.[72] Furthermore, throughput is limited by cost, protocol complexity, available sequencing depth, and dropout events. Together, this can affect the downstream analysis pipeline such as clustering of cell populations and the inference of cellular relationships.

Computational analysis of scRNA-seq data is challenging and involves multiple steps, e.g. quality control, normalization, clustering, and identification of differentially expressed genes and/or trajectory inferences. Multiple unsupervised clustering analyses are available to identify putative cell types, of which graph-based clustering is most widely used.[73] For each of these steps, numerous computational tools are available, but in addition software packages have implemented the entire clustering workflow such as Seurat[16], scanpy[74], and SINCERA[75].

Single-cell epigenetic characterization

Although most high-throughput profiling studies to date have focused on DNA, RNA, and protein expression, recent progress in studying the epigenetic regulation of gene expression, at single-cell level, has been made. Over the last decades, methods have been developed including ATAC-seq to measure chromatin accessibility[76], bisulfite sequencing to measure DNA methylation[77], ChIC-sequencing to measure histone modifications[78], and chromosome conformation capture (3C) to analyze the spatial organization of chromatin in a cell[79]. Several studies revealed epigenetic programs that regulate T cell differentiation and dysfunction in tumors. Analysis of chromatin accessibility by ATAC-seq revealed that $CD8^+$ T cell dysfunction is accompanied with a broad remodeling of the enhancer landscape and transcription factor binding as compared to functional $CD8^+$ T cells in tumors.[80-83] Also, an increased chromatin accessibility at the enhancer site of the *PDCD1* gene (encoding for PD-1) has been found in the context of dysfunctional $CD8^+$ T cells.[82] In addition, studies have applied epigenetics to determine mechanisms of resistance to cancer immunotherapies by characterizing chromatin regulators of intratumoral T cell dysfunction before and after PD-1, PD-L1, or CTLA-4 blockade therapy.[84,85] Lastly, DNA hypermethylation may result in the inactivation of genes, such as mismatch repair gene *MLH1* associated with microsatellite instability in colorectal cancer.[86] Until recently, studies on epigenetic modifications depended on correlations between bulk cell populations. Since 2013, with the development of single-cell technologies, epigenomic techniques have been modified for application to single cells to study cell-to-cell variability in for instance chromatin organization in hundreds or thousands of single cells simultaneously. Several single-cell epigenomic techniques have been reported on recently, including measurements of DNA methylation patterns (scRRBS, scBS-seq, scWHBS)[87-89], chromatin accessibility (scATAC-seq)[90], chromosomal conformations

(scHi-C)[91], and histone modifications (scChIC-seq)[92]. A recent study applied scATAC-seq to characterize chromatin profiles of more than 200,000 single cells in peripheral blood and basal cell carcinoma. By analyzing tumor biopsies before and after PD-1 blockade therapy, Satpathy *et al.* could identify chromatin regulators of therapy-responsive T cell subsets at the level of individual genes and regulatory DNA elements in single cells.[93] Interestingly, variability in histone modification patterns in single cells have also been studied by mass cytometry, which was denominated EpiTOF[94]. In this way, Cheung *et al.* identified a variety of different cell-type and lineage-specific profiles of chromatin marks that could predict the identity of immune cells in humans. Lastly, scATAC-seq has been combined with scRNA-seq and CITE-seq analyses to find distinct and shared molecular mechanisms of leukemia.[95] These single-cell strategies will allow to further understand how the epigenome drives differentiation at the single-cell level and unravel drivers of epigenetic states that could be used as target for the treatment of cancer. Additionally, these methods may be used to measure genome structure in single cells to define the 3D structure of the genome. However, for many of these single-cell epigenetic techniques, disadvantages are the low coverage of regulatory regions such as enhancers (scRRBS), low coverage of sequencing reads (scChiP-seq, scATAC-seq), and low sequencing resolution (scHi-C).[96,97]

Single-cell protein measurements

Flow cytometry has been, in the past decades, the method of choice for high-throughput analysis of protein expression in single cells. The number of markers that can be simultaneously assayed was limited to approximately 14 markers due to the broad emission spectra of the fluorescent dyes. Recent developments with spectral flow cytometer machines enable the detection of up to 34 markers in a single experiment by measuring the full spectra from each cell, which are unmixed by reference spectra of the fluorescent dyes and the autofluorescence spectrum.[98] Fluorescence emission is registered by detectors consisting of avalanche photodiodes instead of photomultiplier tubes used in conventional flow cytometry. A variety of cellular features can be detected by flow cytometry including DNA and RNA content, cell cycle stage, detailed immunophenotypes, apoptotic states, activation of signaling pathways, and others (reviewed by [99]). This technique has thus been paramount in characterizing cell types, revealing the existence of previously unrecognized cell subsets, and for the isolation of functionally distinct cell subsets for the characterization of tumors. However, the design of multiparameter flow cytometry antibody panels is a challenging and laborious task, and most flow cytometry studies have so far focused on the in-depth analysis of specific cellular lineages, instead of a broad and system-wide approach.

In 2009, the advent of a new cytometry methodology, mass cytometry (CyTOF, cytometry by time-of-flight), overcame the limitation of spectral overlap by using metal-isotope-conjugated antibodies to detect antigens.[100] The metal isotopes attached to each cell are distinguished by mass and quantified in a quadrupole time-of-flight mass spectrometer. A mass cytometer is theoretically capable of detecting over 100 parameters per cell, but current chemical methods limit its use to approximately 40-50 parameters, simultaneously. Mass cytometry has expanded the breadth of single-cell data in each experiment, making it highly suitable for systems-

level analyses such as immunophenotyping of cancer microenvironments. By allowing the examination of large datasets at single-cell resolution, mass cytometry can be applied for the discovery of novel cell subsets as well as for the detection and identification of rare cells. Further advantages of mass cytometry are the irrelevance of autofluorescence, the low biological background as heavy metals are not naturally present in biological systems, and limited signal spillover between heavy metals, thereby reducing the complexity of panel design. Conversely, as compared to flow cytometry, mass cytometry suffers from a higher cell loss during acquisition, is more expensive, and is low-throughput, with a flow rate of up to 500 cells per second as compared to thousands of cells per second in flow cytometry. In addition, cells cannot be sorted for further analysis and forward and side scattered light is not detected.

Several studies have applied mass cytometry to further characterize immune cell profiles in peripheral blood or tissues from patients with breast cancer[101], renal cancer[102], melanoma[55,56,103-105], lung cancer[52,106,107], glioma[49,50], colorectal cancer[57,106,108,109], liver cancer[61,110], ovarian cancer[111], and myeloma[112-115], among others. In addition to characterizing immune cell profiles of different tissue types, mass cytometry has also been used to characterize immunophenotypes in tumors and monitor changes during immunotherapy.[56,103-105,114] In this way, factors that influence response to immunotherapy can be investigated and mechanisms at play during treatment can be characterized. This information can be used to understand and facilitate the identification and classification of responder *versus* non-responders to cancer immunotherapy. Most of the studies so far have focused on the CTLA-4 and PD-1/PD-L1 axis of cancer immunotherapy, but novel immunotherapeutic targets such as co-inhibitory molecules LAG-3 or TIM-3 or co-stimulatory molecules such as OX40 and GITR are currently being explored in mice models and clinical trials.[116] Moreover, mass cytometry has been employed to study antigen-specific T cells with a multiplex MHC class I tetramer staining approach, which has led to the identification of phenotypes associated with tumor antigen-specific T cells.[106] Most studies applied mass cytometry for measuring cell surface or intracellular markers, but it can also be used to evaluate cell signaling processes relying on the analysis of protein phosphorylation.[117] Altogether, these studies showed that immune responses in cancer are extremely diverse, within tumors from individual patients as well as between patients with equivalent tumor types. Hence, finding clinically-relevant characteristics based on overall differences can be challenging because of inter-patient variability; differences between cancer patients can be so large that they compromise the discovery of biomarkers.

Because the number of potential phenotypes (resulting from the combination of different markers) increases exponentially with the rise in number of antibodies being measured simultaneously, computational tools for the analysis and visualization of multidimensional data have become key in this field. Traditional workflows for analyzing flow cytometry datasets by manual gating are not efficient to capture the phenotypic differences in mass cytometry and complex flow cytometry data and suffer from individual user bias. In addition, flow and mass cytometry datasets can easily contain millions of cells, illustrating the need for scalable clustering algorithms for efficient analysis. Current single-cell computational tools employed for complex flow cytometry and mass cytometry datasets include unsupervised

clustering-based algorithms such as SPADE[118], Phenograph[119], and FlowSOM[120]. However, these clustering-based tools do not provide single-cell resolution of the data. On the other hand, non-linear dimensionality-reduction-based algorithms such as t-SNE[121] are widely used tools but limited by the number of cells that they can analyze simultaneously, resulting in down-sampled datasets and non-classified cells. Recently, a hierarchical approach of the t-SNE dimensionality-reduction-based technique, HSNE, was described to be scalable to tens of millions of cells.[122,123] In addition, a novel algorithm has recently been implemented in the single-cell analysis field as a dimensionality reduction tool, called uniform manifold approximation and projection (UMAP).[124]

Spatially-resolved data

Most of the multidimensional single-cell techniques such as flow cytometry, mass cytometry, and scRNA-seq require cellular dissociation to obtain cell suspensions prior to measuring the individual cells. Different dissociation methods are used, both mechanical and enzymatic, and may result in the loss of certain cell types and affect the expression of specific cell surface markers. Moreover, tissue specimens are often contaminated with blood or other tissues that are processed along with the tissue of interest. As such, not all subsets identified in single-cell data may be representative of the sample of interest. Another key limitation is the lack of information on spatial localization and cellular interactions within a tissue. Analysis of tissue sections by traditional IHC- and immunofluorescence-based methods are useful in providing spatial information[125], but are severely limited in the number of markers that can be measured simultaneously. Recent technological advances have greatly expanded the number of markers that can be captured on tissue slides. For instance, by applying the principles of secondary ion mass spectrometry to image antibodies conjugated to heavy metal isotopes in tissue sections with imaging mass cytometry (IMC)[126] and multiplexed ion beam imaging by time-of-flight (MIBI-TOF)[127]. In both imaging approaches, conventional IHC workflows are used but with metal-isotope-conjugated antibodies that are detected through a time-of-flight mass spectrometer. In IMC, a pulsed laser is used to ablate a tissue section by rasterizing over a selected region of interest. The liberated antibody-bound ions are subsequently introduced into the inductively coupled plasma time-of-flight mass spectrometer. IMC can currently image up to 40 proteins with a subcellular resolution of 1 μm. The principle of MIBI-TOF is similar, but it makes use of a time-of-flight mass spectrometer equipped with a duoplasmatron primary oxygen ion beam rather than a laser. It currently enables simultaneous imaging of 36 proteins at resolutions down to 260 nm.[128] Both techniques are, however, low-throughput due to the relatively long imaging time of two hours per field of 1 mm^2 in IMC and 1 hour 12 minutes per field of 1 mm^2 in MIBI-TOF.[129] IMC has been applied to study tumor heterogeneity in several types of cancers, such as pancreatic cancer[130], biliary tract cancer[131], breast cancer[126,132,133], and colorectal cancer[108,134]. MIBI-TOF has been used to study the tumor-immune microenvironment of breast cancer[127,128,135,136] and the metabolic state of T cells in colorectal cancer[109]. These spatially-resolved, single-cell analyses have great potential to characterize the spatial inter- and intratumoral phenotypic heterogeneity, which can guide cancer diagnosis, prognosis and the selection of treatment. A recent study was able to extend IMC data by integration

with genomic characterization of breast tumors and could, in this way, investigate the effect of genomic alterations on multidimensional tumor phenotypes of breast cancer.[137]

Other multiplexed imaging techniques such as the Digital Spatial Profiling (DSP) system from NanoString and co-detection by indexing (CODEX) make use of DNA oligonucleotides. In DSP, antibodies or probes are tagged with unique ultraviolet-photocleavable DNA oligos that are released after ultraviolet exposure in specific ROIs and quantified.[138] It enables simultaneous detection of up to 40 proteins or over 90 RNA targets from a tissue section and theoretically allows unlimited multiplexing using the NGS readout, but is time-consuming, does not allow for a reconstruction of the image, and has a lower resolution (10 µm).[129] In CODEX, antibodies conjugated to unique oligonucleotide sequences are detected in a cyclic manner by sequential primer extension with fluorescent dye-labeled nucleotides. CODEX currently allows the detection of over 50 markers with an automated fluidic setup platform including a three-color fluorescence microscope.[139] Of note, throughput is limited by sequential detection of antibody binding. A disadvantage of CODEX, but also of IMC, is the lack of signal amplification which hampers the detection of lowly abundant antigens. A novel imaging technique, called Immuno-SABER, overcame this limitation by implementing a signal amplification step using primer exchange reactions. Immuno-SABER makes use of multiple DNA-barcoded primary antibodies that are hybridized to orthogonal single-stranded DNA concatemers, generated via primer exchange reactions.[140] These primer exchange reactions allow multiplexed signal amplification with rapid exchange cycles of fluorophore-bearing imager strands. The Nanostring DSP platform has been used to study the tumor microenvironment and the outcome of various clinical trials of combination therapy for melanoma[141-144], interactions between macrophages and lymphocytes in metastatic uveal melanoma[145], immune cell subsets in lung cancer[129,143], and tumor microenvironments of different metastases in prostate cancer[146]. CODEX has been applied to study the immune tumor microenvironment of colorectal cancer patients with 56 protein markers simultaneously.[147]

These multiplexed imaging techniques can be applied to snap-frozen as well as FFPE samples that are usually stored in clinical repositories. However, they raise new analysis challenges such as the visualization of 40 markers simultaneously, the image segmentation for cell determination, and algorithms for image-based expression profiles. To understand the tissue architecture, it is necessary to have prior knowledge on which cell types can be present and what their physical relationship to one another could be. Several computational approaches have been developed to enable data analysis of spatially-resolved multiplexed tissue measurements including HistoCAT[148] and ImaCytE[149]. These approaches apply cell segmentation masks (using a combination of Ilastik[150] and CellProfiler[151]) to extract single-cell data from each image, which allow for deep characterization using multidimensional reduction tools such as t-SNE combined with the assessment of spatial localization and cellular interactions. In addition to cell-based analysis such imaging technologies also allow the employment of pixel-based analysis that do not depend on cell segmentation.

Integration of single cell transcriptome profiles with their spatial position in tissue context can

be achieved by labeling of DNA, RNA, or probes using *in situ* hybridization (ISH). Traditional ISH techniques have been improved to allow the detection of single molecules, named single-molecule fluorescence ISH (smFISH) that can be used to quantitate RNA transcripts at single-cell resolution within a tissue context.[152,153] However, only a small number of genes can be measured simultaneously and a main limitation is the lack of cellular resolution to hundreds of micrometers. To improve the throughput, several highly multiplex methods of *in situ* RNA visualization have been developed such as osmFISH[154], sequential FISH (seqFISH[155] and seqFISH+[156]) and error-robust FISH (MERFISH[157]). These allow the subcellular detection of up to a few hundred to ten thousand transcripts simultaneously in single cells *in situ* by using sequential rounds of hybridization with temporal barcodes for each transcript, but require a high number of probes and are time-consuming. Furthermore, ISH can suffer from probe-specific noise due to sequence specificity and background binding. Another approach which may be more applicable for tumors is *in situ* RNA sequencing on tissue sections. STARmap[158] and FISSEQ[159] can profile a few hundreds to thousands of transcripts by using enzymatic amplification methods, but at lower resolution and sensitivity compared to seqFISH and MERFISH. Spatial Transcriptomics[160] and Slide-seq[161] can profile whole transcriptomes by using spatially barcoded oligo-dT microarrays. The spatial transcriptomics method has been used to study and visualize the distribution of mRNAs within tissue sections of breast cancer[160,162], metastatic melanoma[56,163], prostate cancer[164], and pancreatic cancer[165]. These studies highlight the potential of gene expression profiling of cancer tissue sections to reveal the complex transcriptional landscape in its spatial context to gain insight into tumor progression and therapy outcome.

Integration of transcriptomic, (epi)genomic, proteomic, and spatially-resolved single-cell data

Traditionally, each type of single-cell data has been considered independently to investigate a biological system. However, cancer is a spatially-organized system composed of many distinct cell types (**Figure 2A**). These different cell types including immune cells, stromal cells, and malignant cells can be visualized and investigated in an interactive manner (**Figure 2B**). By applying multi-omics to individual cells in the cancer microenvironment, the molecular landscape of every cell[44] can be defined with its proteome (proteins), transcriptome (RNA sequence), genome (DNA sequence), epigenome (DNA methylation, chromatin accessibility), and spatial localization (x, y, z-coordinates) (**Figure 2C**). Integrating these different molecular layers for each cell will allow a detailed profiling of cancer as a complex biological system (**Figure 2D**). Data integration approaches have classically been categorized in three groups: early (concatenation-based), intermediate (transformation-based), and late (model-based) stage integration.[166] Early or intermediate stage integration approaches are more powerful than late stage integration since they can capture interactions between different molecular data-types. However, such approaches are also more challenging methodologically given the different data distributions across data types.

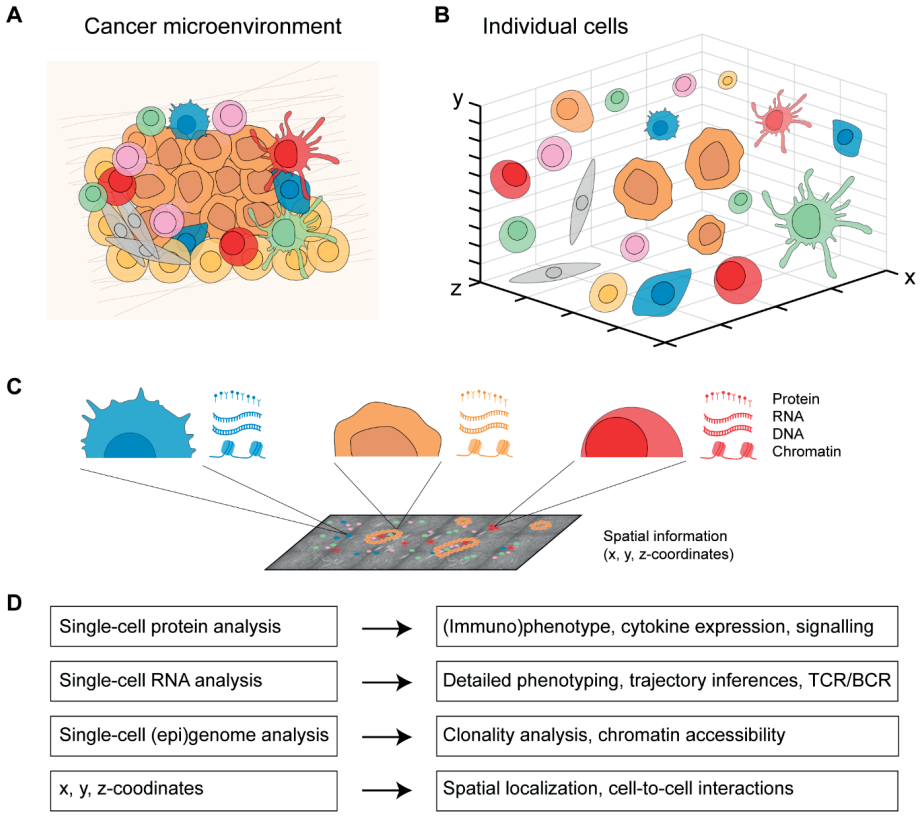


Figure 2. An integrated multicellular model of cancer.
A and B. From (**A**) cells in a spatially-organized cancer microenvironment to (**B**) a three-dimensional view of individual cells. **C.** From each individual cell in the cancer microenvironment, protein expression can be measured by single-cell protein analysis, RNA expression by single-cell RNA analysis, DNA and chromatin expression by single-cell (epi)genome analysis, and the x, y, z-coordinates with spatially-resolved analysis. **D.** Integrating all four molecular layers for each cell will allow a detailed profiling from individual cell-to-cell interactions to whole tissue context.

A number of studies have used complementary forms of multidimensional analysis on the same sample type in the context of cancer. We have performed a search strategy in PubMed, Web of Science, and Embase databases to find studies that have used mass cytometry in concert with scRNA-seq in the context of human cancer (**Table S1**). An overview of the eight relevant studies that applied mass cytometry together with scRNA-seq to study human cancer and their integration stage is shown in **Table 1**. In addition, we performed a search strategy in PubMed, Web of Science, and Embase databases on studies that applied single-cell mass cytometry in concert with spatially-resolved data obtained by IMC or MIBI-TOF in human cancer (**Table S1**). An overview of the two relevant studies and their integration stage is shown in **Table 2**. Notably, all different multidimensional datasets in these studies were analyzed separately and follow a late (model-based) stage integration. Only Goveia and colleagues applied an integration of clustered mass cytometry and scRNA-seq data.[107]

They merged scaled average gene expression data for each scRNA-seq cluster with scaled average protein expression data for each CyTOF cluster, an approach based on a recently described method from Giordani and colleagues.[167] As they integrated the data only after clustering each modality separately, it is still considered late stage integration.

Table 1. Overview of studies applying mass cytometry together with single-cell RNA-sequencing to study human cancer heterogeneity.

Reference	Methods for single-cell profiling	Cancer type	Integration stage
Lavin *et al*. 2017[52]	Mass cytometry and scRNA-seq	Lung cancer	Late
De Vries *et al*. 2019[57]	Mass cytometry and scRNA-seq	Colorectal cancer	Late
Zhang *et al.* 2019[61]	Mass cytometry and scRNA-seq	Liver cancer	Late
Sankowski *et al*. 2019[49]	Mass cytometry and scRNA-seq	Glioma	Late
Halaby *et al*. 2019[55]	Mass cytometry and scRNA-seq	Melanoma	Late
Goswami *et al*. 2020[50]	Mass cytometry and scRNA-seq	Glioblastoma	Late
Goveia *et al.* 2020[107]	Mass cytometry and scRNA-seq	Lung cancer	Late
Helmink *et al*. 2020[56]	Mass cytometry and scRNA-seq	Melanoma	Late

Abbreviations: scRNA-seq; single-cell RNA-sequencing.

Table 2. Overview of studies applying mass cytometry together with imaging mass cytometry or MIBI-TOF to study human cancer heterogeneity.

Reference	Methods for single-cell profiling	Cancer type	Integration stage
Zhang *et al.* 2019[108]	Mass cytometry and IMC	Colon cancer	Late
Hartmann *et al.* 2020[109]	Mass cytometry and MIBI-TOF	Colorectal cancer	Late

Abbreviations: IMC; imaging mass cytometry, MIBI-TOF; multiplexed ion beam imaging by time-of-flight.

Integrating multiple single-cell datasets is a challenging task because of the inherently high levels of noise and the large amount of missing data. Furthermore, the ever-expanding scale of single-cell experiments to millions of cells poses additional challenges. Several methods have been proposed to integrate multimodal single-cell data. State-of-the-art methods focus on embedding both spatial and standard datasets into a latent space using dimensionality reduction, such as Seurat[16], LIGER[17], and Harmony[168], or by employing factor analysis, such as MOFA[169], MOFA+[170], scMerge[171], and scCoGAPS[172]. In addition, a recent study introduced gimVI as a model for integrating spatial transcriptomics data with scRNA-seq data to impute missing gene expression measurements.[15] Of note, most of the methods so far follow an intermediate or late integration approach.[166] As such, these methods overcome challenges due to the different data distributions across data types, but they are less powerful in capturing interactions between different molecular data types.

Several methodologies have been developed to simultaneously acquire multiple measurements from the same cell (**Box 1**). Although obtaining simultaneous measurements from the same cell is becoming more feasible, it is still more common to perform subsequent measurements from the same sample (different sets of cells). Integrating spatial-based assays

with mRNA or protein expression measurements can be beneficial for several reasons. For instance, spatial measurements are often limited in terms of the number of features they can assess simultaneously, although the latest generations of MERFISH and seqFISH(+) can measure around 10,000 transcripts per cell. By integrating these imaging techniques with scRNA-seq, the amount of biologically-relevant information can be enhanced. Moncada and colleagues presented an integration of scRNA-seq with the spatial transcriptomics method generated from the same sample to study pancreatic cancer.[165] A clear challenge when integrating spatial protein (e.g. IMC, MIBI-TOF, CODEX) with scRNA-seq data is the need to model relationships between mRNA and protein expression levels, thus adding an extra layer of complexity. The advent of CITE-seq, combining antibody-based detection of protein markers with transcriptome profiling, could be used to bridge this gap since it allows simultaneous measurement of both mRNA and surface protein marker expression. We foresee an important role for CITE-seq data in the integration of IMC, MIBI-TOF, and CODEX spatial data with scRNA-seq data. Recently, the integration of CITE-seq with CODEX as well as with IMC has been pioneered by Govek *et al.*[173].

Box 1. Methods for the integration of transcriptomic, (epi)genomic, and proteomic single-cell data.

The analysis of protein expression has been extended to include transcript measurements at the single-cell level. CITE-seq[69], REAP-seq[174], and PLAYR[175] can be used to detect mRNA and protein levels simultaneously in single-cells. In CITE-seq and REAP-seq, oligonucleotide-labeled antibodies are used to integrate cellular protein and transcriptome measurements. In PLAYR, mass spectrometry is used to simultaneously analyze the transcriptome and protein expression levels. The analysis of mRNA expression and methylation status in single cells can be achieved by scM&Tseq[176]. In addition, mRNA expression and chromatin accessibility of single cells can be analyzed by sci-CAR[177], SNARE-seq[178], and Paired-seq[179]. Chromatin organization and DNA methylation from a single nucleus can jointly be profiled by snm3C-seq[180]. DR-seq[181] and G&T-seq[182] can assay genomic DNA and mRNA expression simultaneously in single cells, allowing correlations between genomic aberrations and transcriptional levels. Moreover, recent studies have reported on the development of single-cell triple-omics sequencing techniques, such as scTrio-seq[183] and scNMT-seq[184]. In scTrio-seq, the transcriptome, genome, and DNA methylome of individual cells are jointly captured. Lastly, scNMT-seq jointly profiles transcription, DNA methylation, and chromatin accessibility, allowing for a thorough investigation of different epigenomic layers with transcriptional status.

Potential avenues of how the integrated data will help to shed light on the complex role of the microenvironment in cancer

Cancer heterogeneity has long been recognized as a factor complicating the study and treatment of cancer but, until recently, it was difficult to account for in cancer research. The advent of multidimensional single-cell technologies has shed light on the tremendous cellular diversity that exists in cancer tissues and heterogeneity across patients. Moving forward, it will be important to work on the integration of available (spectral) flow cytometry, mass cytometry, scRNA-seq, and spatially-resolved datasets to investigate commonalities and differences in cellular landscapes between cancer tissues. Multiple flow and mass cytometry datasets can be matched if they include a shared marker set between panels, thereby extending the number

of markers per cell and allowing meta-analysis of different mass cytometry datasets with a common core of markers.[185] In addition, cell-type references from different single-cell datasets can improve the functional characterization of cells.[186] Such a system-wide approach will improve insights into how different components of the cancer microenvironment interact in a tissue context. This requires an extensive collaboration between multi-disciplinary research fields such as oncology, immunology, pathology, and bioinformatics.

Nevertheless, the development and widespread use of innovative methodologies also implies the development of analytical tools for the interpretation of complex datasets and their standardization across laboratories. Furthermore, systems-level analyses challenge a researcher's capacity to reconnect findings to their biological relevance. Studies should focus on the removal of unwanted variation and experimental noise in high-throughput single-cell technologies as well as the development of cell-type references, such as the Human Cell Atlas[187] and the Allen Brain Atlas[188] principles. We need to further develop algorithms to integrate data from different imaging and non-imaging single-cell technologies. Alternatively, technological developments should allow the acquisition of molecular profiles from single cells without the need of dissociating them from their tissue context. Lastly, it would be of great value to correlate multi-omics techniques with cell-to-cell signaling networks such as CellPhoneDB[189] and NicheNet[190]. We expect that this integrated and comprehensive data can be used to create a multicellular model of cancer, from single cells to its tissue context, to understand and exploit cancer heterogeneity for improved precision medicine for cancer patients.

How will such system-wide approaches contribute toward more effective therapies for the treatment of cancer? With the advent of targeted therapy and immunotherapy, remarkable advances have been made that changed the management of oncologic treatment for a significant number of patients. However, still only a minority of cancer patients benefit from these therapies, and resistance to treatment remains a major complication in the clinical management of advanced cancer patients. Integrated multi-omics data can help to improve our understanding of the variability in treatment response and resistance mechanisms. By linking detailed molecular and immunological profiles of cells in the cancer microenvironment with sensitivity to specific therapies, potential targets for cancer treatments and associated biomarkers can be identified. This would also support a rational selection of patients that are most likely to benefit from specific treatments. Furthermore, integrated multi-omics data has the potential to guide the development of alternative therapies, for instance through the identification of resistance mechanisms. We expect that such system-wide approaches, with technologies that include spatial information, will become standard methodologies in cancer research in the coming years.

ACKNOWLEDGEMENTS

We thank J.W. Schoones from Walaeus library of Leiden University Medical Center for his help with developing the literature search strategies and M.E. Ijsselsteijn from the dept. of Pathology of Leiden University Medical Center for providing imaging mass cytometry images of colorectal cancer.

AUTHOR CONTRIBUTIONS

NV and AM performed the bibliographic research for the manuscript and designed the figures. NV, AM, FK, and NM jointly wrote the manuscript.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1 Boveri, T. *Zur frage der entstehung maligner tumoren*. (Gustav Fischer, 1914).

2 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* 144, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).

3 Schreiber, R. D., Old, L. J. & Smyth, M. J. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science (New York, N.Y.)* 331, 1565-1570, doi:10.1126/science.1203486 (2011).

4 Vladoiu, M. C. *et al.* Childhood cerebellar tumors mirror conserved fetal transcriptional programs. *Nature* 572, 67-73, doi:10.1038/s41586-019-1158-7 (2019).

5 Tirosh, I. *et al.* Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* 539, 309-313, doi:10.1038/nature20123 (2016).

6 Fridman, W. H., Pages, F., Sautes-Fridman, C. & Galon, J. The immune contexture in human tumors: impact on clinical outcome. *Nat Rev Cancer* 12, 298-306, doi:10.1038/nrc3245 (2012).

7 Binnewies, M. *et al.* Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med* 24, 541-550, doi:10.1038/s41591-018-0014-x (2018).

8 Fridman, W. H., Zitvogel, L., Sautes-Fridman, C. & Kroemer, G. The immune contexture in cancer prognosis and treatment. *Nature reviews. Clinical oncology* 14, 717-734, doi:10.1038/nrclinonc.2017.101 (2017).

9 Angelova, M. *et al.* Evolution of Metastases in Space and Time under Immune Selection. *Cell* 175, 751-765.e716, doi:10.1016/j.cell.2018.09.018 (2018).

10 Paget, S. The distribution of secondary growths in cancer of the breast. *Lancet*, 571-573, doi:https://doi.org/10.1016/S0140-6736(00)49915-0 (1889).

11 Qian, X. *et al.* Probabilistic cell typing enables fine mapping of closely related cell types in situ. *Nature methods*, 10.1038/s41592-41019-40631-41594, doi:10.1038/s41592-019-0631-4 (2019).

12 Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nature biotechnology* 33, 495-502, doi:10.1038/nbt.3192 (2015).

13 Achim, K. *et al.* High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nature biotechnology* 33, 503-509, doi:10.1038/nbt.3209 (2015).

14 Nitzan, M., Karaiskos, N., Friedman, N. & Rajewsky, N. Gene expression cartography. *Nature*, 10.1038/s41586-41019-41773-41583, doi:10.1038/s41586-019-1773-3 (2019).

15 Lopez;, R. *et al.* A joint model of unpaired data from scRNA-seq and spatial transcriptomics for imputing missing gene expression measurements. *arXiv*, doi:1905.02269 (2019).

16 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888-1902.e1821, doi:10.1016/j.cell.2019.05.031 (2019).

17 Welch, J. D. *et al.* Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. *Cell* 177, 1873-1887.e1817, doi:10.1016/j.cell.2019.05.006 (2019).

18 Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nature methods* 12, 453-457, doi:10.1038/nmeth.3337 (2015).

19 Finotello, F. *et al.* Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome medicine* 11, 34-34, doi:10.1186/s13073-019-0638-6 (2019).

20 Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 366, 883-892, doi:10.1056/NEJMoa1113205 (2012).

21 Thorsson, V. *et al.* The Immune Landscape of Cancer. *Immunity* 48, 812-830.e814, doi:10.1016/j.immuni.2018.03.023 (2018).

22 Dean, F. B., Nelson, J. R., Giesler, T. L. & Lasken, R. S. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res* 11, 1095-1099, doi:10.1101/gr.180501 (2001).

23 Zong, C., Lu, S., Chapman, A. R. & Xie, X. S. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science (New York, N.Y.)* 338, 1622-1626, doi:10.1126/science.1229164 (2012).

24 Telenius, H. *et al.* Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics* 13, 718-725, doi:10.1016/0888-7543(92)90147-k (1992).

25 Navin, N. *et al.* Tumor evolution inferred by single-cell sequencing. *Nature* 472, 90-94, doi:10.1038/nature09807 (2011).

26 Xu, X. *et al.* Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* 148, 886-895, doi:10.1016/j.cell.2012.02.025 (2012).

27 Francis, J. M. *et al.* EGFR variant heterogeneity in glioblastoma resolved through single-nucleus sequencing. *Cancer discovery* 4, 956-971, doi:10.1158/2159-8290.CD-13-0879 (2014).

28 Gawad, C., Koh, W. & Quake, S. R. Dissecting the clonal origins of childhood acute lymphoblastic

leukemia by single-cell genomics. *Proceedings of the National Academy of Sciences of the United States of America* 111, 17947-17952, doi:10.1073/pnas.1420822111 (2014).

29 Casasent, A. K. *et al.* Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. *Cell* 172, 205-217.e212, doi:10.1016/j.cell.2017.12.007 (2018).

30 Yu, C. *et al.* Discovery of biclonal origin and a novel oncogene SLC12A5 in colon cancer by single-cell sequencing. *Cell Res* 24, 701-712, doi:10.1038/cr.2014.43 (2014).

31 Hughes, A. E. O. *et al.* Clonal architecture of secondary acute myeloid leukemia defined by single-cell sequencing. *PLoS Genet* 10, e1004462-e1004462, doi:10.1371/journal.pgen.1004462 (2014).

32 Jan, M. *et al.* Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Science translational medicine* 4, 149ra118-149ra118, doi:10.1126/scitranslmed.3004315 (2012).

33 Leung, M. L. *et al.* Single-cell DNA sequencing reveals a late-dissemination model in metastatic colorectal cancer. *Genome Res* 27, 1287-1299, doi:10.1101/gr.209973.116 (2017).

34 Heitzer, E. *et al.* Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res* 73, 2965-2975, doi:10.1158/0008-5472.Can-12-4140 (2013).

35 Lohr, J. G. *et al.* Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. *Nat Biotechnol* 32, 479-484, doi:10.1038/nbt.2892 (2014).

36 Ni, X. *et al.* Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. 110, 21083-21088, doi:10.1073/pnas.1320659110 %J Proceedings of the National Academy of Sciences (2013).

37 Wang, Y. *et al.* Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* 512, 155-160, doi:10.1038/nature13600 (2014).

38 Suzuki, A. *et al.* Single-cell analysis of lung adenocarcinoma cell lines reveals diverse expression patterns of individual cells invoked by a molecular target drug treatment. *Genome Biol* 16, 66, doi:10.1186/s13059-015-0636-y (2015).

39 Wang, Y. & Navin, N. E. Advances and applications of single-cell sequencing technologies. *Mol Cell* 58, 598-609, doi:10.1016/j.molcel.2015.05.005 (2015).

40 Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nature methods* 6, 377-382, doi:10.1038/nmeth.1315 (2009).

41 Svensson, V., Vento-Tormo, R. & Teichmann, S. A. Exponential scaling of single-cell RNA-seq in the past decade. *Nature Protocols* 13, 599-604, doi:10.1038/nprot.2017.149 (2018).

42 Genomics, X. *10X Genomics single cell gene expression datasets from https://support.10xgenomics.com/single-cell-gene-expression/datasets*, 2017).

43 Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Experimental & molecular medicine* 50, 96, doi:10.1038/s12276-018-0071-8 (2018).

44 Wagner, A., Regev, A. & Yosef, N. Revealing the vectors of cellular identity with single-cell genomics. *Nature biotechnology* 34, 1145-1160, doi:10.1038/nbt.3711 (2016).

45 Ramsköld, D. *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nature biotechnology* 30, 777-782, doi:10.1038/nbt.2282 (2012).

46 Miyamoto, D. T. *et al.* RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. *Science (New York, N.Y.)* 349, 1351-1356, doi:10.1126/science.aab0917 (2015).

47 Brady, S. W. *et al.* Combating subclonal evolution of resistant cancer phenotypes. *Nature communications* 8, 1231-1231, doi:10.1038/s41467-017-01174-3 (2017).

48 Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science (New York, N.Y.)* 344, 1396-1401, doi:10.1126/science.1254257 (2014).

49 Sankowski, R. *et al.* Mapping microglia states in the human brain through the integration of high-dimensional techniques. *Nat Neurosci* 22, 2098-2110, doi:10.1038/s41593-019-0532-y (2019).

50 Goswami, S. *et al.* Immune profiling of human tumors identifies CD73 as a combinatorial target in glioblastoma. *Nature medicine* 26, 39-46, doi:10.1038/s41591-019-0694-x (2020).

51 Kim, K.-T. *et al.* Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. *Genome biology* 17, 80-80, doi:10.1186/s13059-016-0945-9 (2016).

52 Lavin, Y. *et al.* Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. *Cell* 169, 750-765 e717, doi:10.1016/j.cell.2017.04.014 (2017).

53 Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science (New York, N.Y.)* 352, 189-196, doi:10.1126/science.aad0501 (2016).

54 Li, H. *et al.* Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell* 176, 775-789.e718, doi:10.1016/j.cell.2018.11.043 (2019).

55 Halaby, M. J. *et al.* GCN2 drives macrophage

and MDSC function and immunosuppression in the tumor microenvironment. *Sci Immunol* 4, eaax8189, doi:10.1126/sciimmunol.aax8189 (2019).

56 Helmink, B. A. *et al.* B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* 577, 549-555, doi:10.1038/s41586-019-1922-8 (2020).

57 de Vries, N. L. *et al.* High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of antitumour immunity. *Gut*, doi:10.1136/gutjnl-2019-318672 (2019).

58 Li, H. *et al.* Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet* 49, 708-718, doi:10.1038/ng.3818 (2017).

59 Zhang, Y. *et al.* Deep single-cell RNA sequencing data of individual T cells from treatment-naïve colorectal cancer patients. *Sci Data* 6, 131-131, doi:10.1038/s41597-019-0131-5 (2019).

60 Elyada, E. *et al.* Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts. *Cancer discovery* 9, 1102-1123, doi:10.1158/2159-8290.CD-19-0094 (2019).

61 Zhang, Q. *et al.* Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas. *Gut* 68, 2019-2031, doi:10.1136/gutjnl-2019-318912 (2019).

62 Puram, S. V. *et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* 171, 1611-1624.e1624, doi:10.1016/j.cell.2017.10.044 (2017).

63 Giustacchini, A. *et al.* Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nature medicine* 23, 692-702, doi:10.1038/nm.4336 (2017).

64 Filbin, M. G. *et al.* Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science (New York, N.Y.)* 360, 331-335, doi:10.1126/science.aao4750 (2018).

65 House, I. G. *et al.* Macrophage-Derived CXCL9 and CXCL10 Are Required for Antitumor Immune Responses Following Immune Checkpoint Blockade. *Clinical cancer research : an official journal of the American Association for Cancer Research* 26, 487-504, doi:10.1158/1078-0432.CCR-19-1868 (2020).

66 Jerby-Arnon, L. *et al.* A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. *Cell* 175, 984-997.e924, doi:10.1016/j.cell.2018.09.006 (2018).

67 Goldstein, L. D. *et al.* Massively parallel single-cell B-cell receptor sequencing enables rapid discovery of diverse antigen-reactive antibodies. *Communications biology* 2, 304, doi:10.1038/s42003-019-0551-y (2019).

68 Singh, M. *et al.* High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. *Nat Commun* 10, 3120, doi:10.1038/s41467-019-11049-4 (2019).

69 Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nature methods* 14, 865-868, doi:10.1038/nmeth.4380 (2017).

70 Peterson, V. M. *et al.* Multiplexed quantification of proteins and transcripts in single cells. *Nat Biotechnol*, doi:10.1038/nbt.3973 (2017).

71 Shahi, P., Kim, S. C., Haliburton, J. R., Gartner, Z. J. & Abate, A. R. Abseq: Ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding. *Sci Rep* 7, 44447, doi:10.1038/srep44447 (2017).

72 van den Brink, S. C. *et al.* Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nature methods* 14, 935-936, doi:10.1038/nmeth.4437 (2017).

73 Kiselev, V. Y., Andrews, T. S. & Hemberg, M. Challenges in unsupervised clustering of single-cell RNA-seq data. *Nat Rev Genet* 20, 273-282, doi:10.1038/s41576-018-0088-9 (2019).

74 Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome biology* 19, 15-15, doi:10.1186/s13059-017-1382-0 (2018).

75 Guo, M., Wang, H., Potter, S. S., Whitsett, J. A. & Xu, Y. SINCERA: A Pipeline for Single-Cell RNA-Seq Profiling Analysis. *PLoS computational biology* 11, e1004575-e1004575, doi:10.1371/journal.pcbi.1004575 (2015).

76 Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature methods* 10, 1213-1218, doi:10.1038/nmeth.2688 (2013).

77 Frommer, M. *et al.* A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America* 89, 1827-1831, doi:10.1073/pnas.89.5.1827 (1992).

78 Schmid, M., Durussel, T. & Laemmli, U. K. ChIC and ChEC; genomic mapping of chromatin proteins. *Mol Cell* 16, 147-157, doi:10.1016/j.molcel.2004.09.007 (2004).

79 Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science (New York, N.Y.)* 295, 1306-1311, doi:10.1126/science.1067799 (2002).

80 Mognol, G. P. *et al.* Exhaustion-associated regulatory regions in CD8(+) tumor-infiltrating T cells. *Proc Natl Acad Sci U S A* 114, E2776-e2785,

doi:10.1073/pnas.1620498114 (2017).

81 Philip, M. *et al.* Chromatin states define tumor-specific T cell dysfunction and reprogramming. *Nature* 545, 452-456, doi:10.1038/nature22367 (2017).

82 Sen, D. R. *et al.* The epigenetic landscape of T cell exhaustion. *Science* 354, 1165-1169, doi:10.1126/science.aae0491 (2016).

83 Pauken, K. E. *et al.* Epigenetic stability of exhausted T cells limits durability of reinvigoration by PD-1 blockade. *Science* 354, 1160-1165, doi:10.1126/science.aaf2807 (2016).

84 Pan, D. *et al.* A major chromatin regulator determines resistance of tumor cells to T cell-mediated killing. *Science (New York, N.Y.)* 359, 770-775, doi:10.1126/science.aao1710 (2018).

85 Miao, D. *et al.* Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science (New York, N.Y.)* 359, 801-806, doi:10.1126/science.aan5951 (2018).

86 Deng, G. *et al.* Methylation of hMLH1 promoter correlates with the gene silencing with a region-specific manner in colorectal cancer. *Br J Cancer* 86, 574-579, doi:10.1038/sj.bjc.6600148 (2002).

87 Guo, H. *et al.* Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res* 23, 2126-2135, doi:10.1101/gr.161679.113 (2013).

88 Smallwood, S. A. *et al.* Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nature methods* 11, 817-820, doi:10.1038/nmeth.3035 (2014).

89 Farlik, M. *et al.* Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. *Cell reports* 10, 1386-1397, doi:10.1016/j.celrep.2015.02.001 (2015).

90 Buenrostro, J. D. *et al.* Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523, 486-490, doi:10.1038/nature14590 (2015).

91 Nagano, T. *et al.* Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502, 59-64, doi:10.1038/nature12593 (2013).

92 Ku, W. L. *et al.* Single-cell chromatin immunocleavage sequencing (scChIC-seq) to profile histone modification. *Nature methods* 16, 323-325, doi:10.1038/s41592-019-0361-7 (2019).

93 Satpathy, A. T. *et al.* Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat Biotechnol* 37, 925-936, doi:10.1038/s41587-019-0206-z (2019).

94 Cheung, P. *et al.* Single-Cell Chromatin Modification Profiling Reveals Increased Epigenetic Variations with Aging. *Cell* 173, 1385-1397. e1314, doi:10.1016/j.cell.2018.03.079 (2018).

95 Granja, J. M. *et al.* Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nat Biotechnol* 37, 1458-1465, doi:10.1038/s41587-019-0332-7 (2019).

96 Lo, P. K. & Zhou, Q. Emerging techniques in single-cell epigenomics and their applications to cancer research. *Journal of clinical genomics* 1, doi:10.4172/jcg.1000103 (2018).

97 Clark, S. J., Lee, H. J., Smallwood, S. A., Kelsey, G. & Reik, W. Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. *Genome Biol* 17, 72, doi:10.1186/s13059-016-0944-x (2016).

98 Futamura, K. *et al.* Novel full-spectral flow cytometry with multiple spectrally-adjacent fluorescent proteins and fluorochromes and visualization of in vivo cellular movement. *Cytometry. Part A : the journal of the International Society for Analytical Cytology* 87, 830-842, doi:10.1002/cyto.a.22725 (2015).

99 Irish, J. M. & Doxie, D. B. High-dimensional single-cell cancer biology. *Current topics in microbiology and immunology* 377, 1-21, doi:10.1007/82_2014_367 (2014).

100 Bandura, D. R. *et al.* Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal Chem* 81, 6813-6822, doi:10.1021/ac901049w (2009).

101 Wagner, J. *et al.* A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. *Cell* 177, 1330-1345.e1318, doi:10.1016/j.cell.2019.03.005 (2019).

102 Chevrier, S. *et al.* An Immune Atlas of Clear Cell Renal Cell Carcinoma. *Cell* 169, 736-749 e718, doi:10.1016/j.cell.2017.04.016 (2017).

103 Wei, S. C. *et al.* Distinct Cellular Mechanisms Underlie Anti-CTLA-4 and Anti-PD-1 Checkpoint Blockade. *Cell* 170, 1120-1133.e1117, doi:10.1016/j.cell.2017.07.024 (2017).

104 Krieg, C. *et al.* High-dimensional single-cell analysis predicts response to anti-PD-1 immunotherapy. *Nature medicine* 24, 144-153, doi:10.1038/nm.4466 (2018).

105 Spitzer, M. H. *et al.* Systemic Immunity Is Required for Effective Cancer Immunotherapy. *Cell* 168, 487-502 e415, doi:10.1016/j.cell.2016.12.022 (2017).

106 Simoni, Y. *et al.* Bystander CD8(+) T cells are abundant and phenotypically distinct in human tumor infiltrates. *Nature* 557, 575-579, doi:10.1038/s41586-018-0130-2 (2018).

107 Goveia, J. *et al.* An Integrated Gene Expression Landscape Profiling Approach to Identify Lung Tumor Endothelial Cell Heterogeneity and Angiogenic Candidates. *Cancer cell* 37, 21-36.

e13, doi:10.1016/j.ccell.2019.12.001 (2020).

108 Zhang, T. *et al.* Immunocyte Profiling Using Single-Cell Mass Cytometry Reveals EpCAM(+) CD4(+) T Cells Abnormal in Colon Cancer. *Frontiers in immunology* 10, 1571-1571, doi:10.3389/fimmu.2019.01571 (2019).

109 Hartmann, F. J. *et al.* Multiplexed Single-cell Metabolic Profiles Organize the Spectrum of Cytotoxic Human T Cells. 2020.2001.2017.909796, doi:10.1101/2020.01.17.909796 %J bioRxiv (2020).

110 Chew, V. *et al.* Delineation of an immunosuppressive gradient in hepatocellular carcinoma using high-dimensional proteomic and transcriptomic analyses. *Proceedings of the National Academy of Sciences of the United States of America* 114, E5900-E5909, doi:10.1073/pnas.1706559114 (2017).

111 Gonzalez, V. D. *et al.* Commonly Occurring Cell Subsets in High-Grade Serous Ovarian Tumors Identified by Single-Cell Mass Cytometry. *Cell reports* 22, 1875-1888, doi:10.1016/j.celrep.2018.01.053 (2018).

112 Baughn, L. B. *et al.* Phenotypic and functional characterization of a bortezomib-resistant multiple myeloma cell line by flow and mass cytometry. *Leuk Lymphoma* 58, 1931-1940, doi:10.1080/10428194.2016.1266621 (2017).

113 Hansmann, L. *et al.* Mass cytometry analysis shows that a novel memory phenotype B cell is expanded in multiple myeloma. *Cancer immunology research* 3, 650-660, doi:10.1158/2326-6066.CIR-14-0236-T (2015).

114 Adams, H. C., 3rd *et al.* High-Parameter Mass Cytometry Evaluation of Relapsed/Refractory Multiple Myeloma Patients Treated with Daratumumab Demonstrates Immune Modulation as a Novel Mechanism of Action. *Cytometry. Part A : the journal of the International Society for Analytical Cytology* 95, 279-289, doi:10.1002/cyto.a.23693 (2019).

115 Marsh-Wakefield, F. *et al.* Mass Cytometry Discovers Two Discrete Subsets of CD39(-) Treg Which Discriminate MGUS From Multiple Myeloma. *Front Immunol* 10, 1596, doi:10.3389/fimmu.2019.01596 (2019).

116 Marin-Acevedo, J. A. *et al.* Next generation of immune checkpoint therapy in cancer: new developments and challenges. *J Hematol Oncol* 11, 39-39, doi:10.1186/s13045-018-0582-8 (2018).

117 Bodenmiller, B. *et al.* Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nature biotechnology* 30, 858-867, doi:10.1038/nbt.2317 (2012).

118 Qiu, P. *et al.* Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nature biotechnology* 29, 886-891, doi:10.1038/nbt.1991 (2011).

119 Levine, J. H. *et al.* Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* 162, 184-197, doi:10.1016/j.cell.2015.05.047 (2015).

120 Van Gassen, S. *et al.* FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry. Part A : the journal of the International Society for Analytical Cytology* 87, 636-645, doi:10.1002/cyto.a.22625 (2015).

121 van der Maaten, L. J. P. & Hinton, G. E. Visualizing high-dimensional data using t-SNE. *J. Mach. Learn. Res.* 9, 2579-2605 (2008).

122 Pezzotti, N., Höllt, T., Lelieveldt, B., Eisemann, E. & Vilanova, A. Hierarchical Stochastic Neighbor Embedding. *Computer Graphics Forum* 35, 21-30, doi:10.1111/cgf.12878 (2016).

123 van Unen, V. *et al.* Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types. *Nat Commun* 8, 1740, doi:10.1038/s41467-017-01689-9 (2017).

124 Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nature biotechnology*, 10.1038/nbt.4314, doi:10.1038/nbt.4314 (2018).

125 Ijsselsteijn, M. E. *et al.* Cancer immunophenotyping by seven-color multispectral imaging without tyramide signal amplification. *The journal of pathology. Clinical research* 5, 3-11, doi:10.1002/cjp2.113 (2019).

126 Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat Methods* 11, 417-422, doi:10.1038/nmeth.2869 (2014).

127 Angelo, M. *et al.* Multiplexed ion beam imaging of human breast tumors. *Nature medicine* 20, 436-442, doi:10.1038/nm.3488 (2014).

128 Keren, L. *et al.* MIBI-TOF: A multiplexed imaging platform relates cellular phenotypes and tissue structure. *Sci Adv* 5, eaax5851-eaax5851, doi:10.1126/sciadv.aax5851 (2019).

129 Decalf, J., Albert, M. L. & Ziai, J. New tools for pathology: a user's review of a highly multiplexed method for in situ analysis of protein and RNA expression in tissue. *J Pathol* 247, 650-661, doi:10.1002/path.5223 (2019).

130 Chang, Q. *et al.* Imaging Mass Cytometry. *Cytometry A* 91, 160-169, doi:10.1002/cyto.a.23053 (2017).

131 Umemoto, K. *et al.* The potential application of PD-1 blockade therapy for early-stage biliary tract cancer. *International immunology*, dxz080, doi:10.1093/intimm/dxz080 (2019).

132 Jackson, H. W. *et al.* The single-cell pathology landscape of breast cancer. *Nature*, 10.1038/s41586-41019-41876-x, doi:10.1038/s41586-

019-1876-x (2020).

133 Schulz, D. *et al.* Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry. *Cell Syst* 6, 25-36 e25, doi:10.1016/j.cels.2017.12.001 (2018).

134 Ijsselsteijn, M. E., van der Breggen, R., Farina Sarasqueta, A., Koning, F. & de Miranda, N. F. C. C. A 40-Marker Panel for High Dimensional Characterization of Cancer Immune Microenvironments by Imaging Mass Cytometry. *Frontiers in immunology* 10, 2534-2534, doi:10.3389/fimmu.2019.02534 (2019).

135 Keren, L. *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* 174, 1373-1387.e1319, doi:10.1016/j.cell.2018.08.039 (2018).

136 Rost, S. *et al.* Multiplexed ion beam imaging analysis for quantitation of protein expression in cancer tissue sections. *Laboratory investigation; a journal of technical methods and pathology* 97, 992-1003, doi:10.1038/labinvest.2017.50 (2017).

137 Ali, H. R. *et al.* Imaging mass cytometry and multiplatform genomics define the phenogenomic landscape of breast cancer. *Nature Cancer* 1, 163-175, doi:10.1038/s43018-020-0026-6 (2020).

138 Merritt, C. R. *et al.* High multiplex, digital spatial profiling of proteins and RNA in fixed tissue using genomic detection methods. 559021, doi:10.1101/559021 %J bioRxiv (2019).

139 Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* 174, 968-981.e915, doi:10.1016/j.cell.2018.07.010 (2018).

140 Saka, S. K. *et al.* Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nature biotechnology* 37, 1080-1090, doi:10.1038/s41587-019-0207-y (2019).

141 Cabrita, R. *et al.* Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* 577, 561-565, doi:10.1038/s41586-019-1914-8 (2020).

142 Amaria, R. N. *et al.* Neoadjuvant immune checkpoint blockade in high-risk resectable melanoma. *Nature medicine* 24, 1649-1654, doi:10.1038/s41591-018-0197-1 (2018).

143 Toki, M. I. *et al.* High-Plex Predictive Marker Discovery for Melanoma Immunotherapy-Treated Patients Using Digital Spatial Profiling. *Clinical cancer research : an official journal of the American Association for Cancer Research* 25, 5503-5512, doi:10.1158/1078-0432.CCR-19-0104 (2019).

144 Blank, C. U. *et al.* Neoadjuvant versus adjuvant ipilimumab plus nivolumab in macroscopic stage III melanoma. *Nature medicine* 24, 1655-1661, doi:10.1038/s41591-018-0198-0 (2018).

145 Figueiredo, C. R. *et al.* Loss of BAP1 expression is associated with an immunosuppressive microenvironment in uveal melanoma, with implications for immunotherapy development. *J Pathol*, 10.1002/path.5384, doi:10.1002/path.5384 (2020).

146 Ihle, C. L. *et al.* Distinct tumor microenvironments of lytic and blastic bone metastases in prostate cancer patients. *Journal for immunotherapy of cancer* 7, 293-293, doi:10.1186/s40425-019-0753-3 (2019).

147 Schürch, C. M. *et al.* Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front *Cell* (August 21, 2019), doi: CELL-D-19-02119.

148 Schapiro, D. *et al.* histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nature methods* 14, 873-876, doi:10.1038/nmeth.4391 (2017).

149 Somarakis, A., Van Unen, V., Koning, F., Lelieveldt, B. P. F. & Hollt, T. ImaCytE: Visual Exploration of Cellular Microenvironments for Imaging Mass Cytometry Data. *IEEE transactions on visualization and computer graphics*, 10.1109/TVCG.2019.2931299, doi:10.1109/TVCG.2019.2931299 (2019).

150 Sommer, C., Straehle, C., Köthe, U. & Hamprecht, F. A. in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro.* 230-233.

151 Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* 7, R100, doi:10.1186/gb-2006-7-10-r100 (2006).

152 Femino, A. M., Fay, F. S., Fogarty, K. & Singer, R. H. Visualization of single RNA transcripts in situ. *Science (New York, N.Y.)* 280, 585-590, doi:10.1126/science.280.5363.585 (1998).

153 Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nature methods* 5, 877-879, doi:10.1038/nmeth.1253 (2008).

154 Codeluppi, S. *et al.* Spatial organization of the somatosensory cortex revealed by osmFISH. *Nature Methods* 15, 932-935, doi:10.1038/s41592-018-0175-z (2018).

155 Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M. & Cai, L. Single-cell in situ RNA profiling by sequential hybridization. *Nature methods* 11, 360-361, doi:10.1038/nmeth.2892 (2014).

156 Eng, C.-H. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* 568, 235-239, doi:10.1038/s41586-019-1049-y (2019).

157 Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells.

Science (New York, N.Y.) 348, aaa6090-aaa6090, doi:10.1126/science.aaa6090 (2015).

158 Wang, X. *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science (New York, N.Y.)* 361, eaat5691, doi:10.1126/science.aat5691 (2018).

159 Lee, J. H. *et al.* Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nature protocols* 10, 442-458, doi:10.1038/nprot.2014.191 (2015).

160 Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science (New York, N.Y.)* 353, 78-82, doi:10.1126/science.aaf2403 (2016).

161 Rodriques, S. G. *et al.* Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science (New York, N.Y.)* 363, 1463-1467, doi:10.1126/science.aaw1219 (2019).

162 Vickovic, S. *et al.* High-definition spatial transcriptomics for in situ tissue profiling. *Nat Methods* 16, 987-990, doi:10.1038/s41592-019-0548-y (2019).

163 Thrane, K., Eriksson, H., Maaskola, J., Hansson, J. & Lundeberg, J. Spatially Resolved Transcriptomics Enables Dissection of Genetic Heterogeneity in Stage III Cutaneous Malignant Melanoma. *Cancer research* 78, 5970-5979, doi:10.1158/0008-5472.CAN-18-0747 (2018).

164 Berglund, E. *et al.* Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nature communications* 9, 2419-2419, doi:10.1038/s41467-018-04724-5 (2018).

165 Moncada, R. *et al.* Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nature biotechnology*, 10.1038/s41587-41019-40392-41588, doi:10.1038/s41587-019-0392-8 (2020).

166 Ritchie, M. D., Holzinger, E. R., Li, R., Pendergrass, S. A. & Kim, D. Methods of integrating data to uncover genotype-phenotype interactions. *Nat Rev Genet* 16, 85-97, doi:10.1038/nrg3868 (2015).

167 Giordani, L. *et al.* High-Dimensional Single-Cell Cartography Reveals Novel Skeletal Muscle-Resident Cell Populations. *Mol Cell* 74, 609-621.e606, doi:10.1016/j.molcel.2019.02.026 (2019).

168 Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature methods* 16, 1289-1296, doi:10.1038/s41592-019-0619-0 (2019).

169 Argelaguet, R. *et al.* Multi-Omics Factor Analysis-a framework for unsupervised integration of multi-omics data sets. *Mol Syst Biol* 14, e8124-e8124, doi:10.15252/msb.20178124 (2018).

170 Argelaguet, R. *et al.* MOFA+: a probabilistic framework for comprehensive integration of structured single-cell data. 837104, doi:10.1101/837104 %J bioRxiv (2019).

171 Lin, Y. *et al.* scMerge leverages factor analysis, stable expression, and pseudoreplication to merge multiple single-cell RNA-seq datasets. *Proceedings of the National Academy of Sciences of the United States of America* 116, 9775-9784, doi:10.1073/pnas.1820006116 (2019).

172 Stein-O'Brien, G. L. *et al.* Decomposing Cell Identity for Transfer Learning across Cellular Measurements, Platforms, Tissues, and Species. *Cell systems* 8, 395-411.e398, doi:10.1016/j.cels.2019.04.004 (2019).

173 Govek, K. W., Troisi, E. C., Miao, Z., Woodhouse, S. & Camara, P. G. Single-Cell Transcriptomic Analysis of mIHC Images via Antigen Mapping. 672501, doi:10.1101/672501 %J bioRxiv (2020).

174 Peterson, V. M. *et al.* Multiplexed quantification of proteins and transcripts in single cells. *Nature biotechnology* 35, 936-939, doi:10.1038/nbt.3973 (2017).

175 Frei, A. P. *et al.* Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nature methods* 13, 269-275, doi:10.1038/nmeth.3742 (2016).

176 Angermueller, C. *et al.* Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nature methods* 13, 229-232, doi:10.1038/nmeth.3728 (2016).

177 Cao, J. *et al.* Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science (New York, N.Y.)* 361, 1380-1385, doi:10.1126/science.aau0730 (2018).

178 Chen, S., Lake, B. B. & Zhang, K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nature biotechnology*, 10.1038/s41587-41019-40290-41580, doi:10.1038/s41587-019-0290-0 (2019).

179 Zhu, C. *et al.* An ultra high-throughput method for single-cell joint analysis of open chromatin and transcriptome. *Nat Struct Mol Biol* 26, 1063-1070, doi:10.1038/s41594-019-0323-x (2019).

180 Lee, D.-S. *et al.* Simultaneous profiling of 3D genome structure and DNA methylation in single human cells. *Nature methods* 16, 999-1006, doi:10.1038/s41592-019-0547-z (2019).

181 Dey, S. S., Kester, L., Spanjaard, B., Bienko, M. & van Oudenaarden, A. Integrated genome and transcriptome sequencing of the same cell. *Nature biotechnology* 33, 285-289, doi:10.1038/nbt.3129 (2015).

182 Macaulay, I. C. *et al.* G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nature methods* 12, 519-522, doi:10.1038/nmeth.3370 (2015).

183 Hou, Y. *et al.* Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell*

Res 26, 304-319, doi:10.1038/cr.2016.23 (2016).

184 Clark, S. J. *et al.* scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nature communications* 9, 781-781, doi:10.1038/s41467-018-03149-4 (2018).

185 Abdelaal, T. *et al.* CyTOFmerge: Integrating mass cytometry data across multiple panels. *Bioinformatics (Oxford, England)*, doi:10.1093/bioinformatics/btz180 (2019).

186 Gomes, T., Teichmann, S. A. & Talavera-López, C. Immunology Driven by Large-Scale Single-Cell Sequencing. *Trends in immunology* 40, 1011-1021, doi:10.1016/j.it.2019.09.004 (2019).

187 Regev, A. *et al.* The Human Cell Atlas. *Elife* 6, e27041, doi:10.7554/eLife.27041 (2017).

188 Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168-176, doi:10.1038/nature05453 (2007).

189 Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB v2.0: Inferring cell-cell communication from combined expression of multi-subunit receptor-ligand complexes. 680926, doi:10.1101/680926 %J bioRxiv (2019).

190 Bonnardel, J. *et al.* Stellate Cells, Hepatocytes, and Endothelial Cells Imprint the Kupffer Cell Identity on Monocytes Colonizing the Liver Macrophage Niche. *Immunity* 51, 638-654.e639, doi:10.1016/j.immuni.2019.08.017 (2019).

SUPPLEMENTAL TABLES

The supplementary table is available in the online version of this manuscript at https://doi.org/10.3389/fonc.2020.01254 and in the appendix to this thesis at the repository of Leiden University (https://hdl.handle.net/1887/3439882).

Table S1: Literature search strategy for PubMed, Web of Science, and Embase databases conducted on 27-02-2020.

High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of anti-tumor immunity

3

Natasja L. de Vries[1,2], Vincent van Unen[2#], Marieke E. Ijsselsteijn[1#], Tamim Abdelaal[3,4#], Ruud van der Breggen[1], Arantza Farina Sarasqueta[1], Ahmed Mahfouz[3,4], Koen C.M.J. Peeters[5], Thomas Höllt[4,6], Boudewijn P.F. Lelieveldt[3,7], Frits Koning[2†], Noel F.C.C. de Miranda[1†].

[1]Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands. [2]Department of Immunology, Leiden University Medical Center, Leiden, the Netherlands. [3]Department of Pattern Recognition and Bioinformatics, Delft University of Technology, Delft, the Netherlands. [4]Leiden Computational Biology Center, Leiden University Medical Center, Leiden, the Netherlands. [5]Department of Surgery, Leiden University Medical Center, Leiden, the Netherlands. [6]Computer Graphics and Visualization, Delft University of Technology, Delft, the Netherlands. [7]Department of LKEB Radiology, Leiden University Medical Center, Leiden, the Netherlands. [#]Equal contribution to this study. [†]Equal responsible authors.

Gut. 2020; 69: 691-703.
doi: 10.1136/gutjnl-2019-318672

ABSTRACT

Objective

A comprehensive understanding of anti-cancer immune responses is paramount for the optimal application and development of cancer immunotherapies. We unraveled local and systemic immune profiles in colorectal cancer (CRC) patients by high-dimensional analysis to provide an unbiased characterization of the immune contexture of CRC.

Design

Thirty-six immune cell markers were simultaneously assessed at the single-cell level by mass cytometry in 35 CRC tissues, 26 tumor-associated lymph nodes, 17 colorectal healthy mucosa, and 19 peripheral blood samples from 31 CRC patients. Additionally, functional, transcriptional, and spatial analyses of tumor-infiltrating lymphocytes were performed by flow cytometry, single-cell RNA-sequencing, and multispectral immunofluorescence.

Results

We discovered that a previously unappreciated innate lymphocyte population ($Lin^-CD7^+CD127^-CD56^+CD45RO^+$) was enriched in CRC tissues and displayed cytotoxic activity. This subset demonstrated a tissue-resident ($CD103^+CD69^+$) phenotype, and was most abundant in immunogenic mismatch repair (MMR)-deficient CRCs. Their presence in tumors was correlated with the infiltration of tumor-resident cytotoxic, helper, and γδ T cells with highly similar activated ($HLA\text{-}DR^+CD38^+PD\text{-}1^+$) phenotypes. Remarkably, activated γδ T cells were almost exclusively found in MMR-deficient cancers. Non-activated counterparts of tumor-resident cytotoxic and γδ T cells were present in CRC and healthy mucosa tissues, but not in lymph nodes, with the exception of tumor-positive lymph nodes.

Conclusion

This work provides a blueprint for the understanding of the heterogeneous and intricate immune landscape of CRC, including the identification of previously unappreciated immune cell subsets. The concomitant presence of tumor-resident innate and adaptive immune cell populations suggests a multi-targeted exploitation of their anti-tumor properties in a therapeutic setting.

INTRODUCTION

T cell checkpoint blockade immunotherapies have revolutionized cancer treatment following the clinical success achieved with therapeutic antibodies targeting CTLA-4 and the PD-1/PD-L1 axis in cancer patients. These strategies reinvigorate anti-tumor T cell responses, and are particularly effective in cancers with high mutation burden like melanomas, non-small cell lung cancers, and DNA mismatch repair (MMR)-deficient cancers.[1-5] MMR deficiency occurs in approximately 15-20% of colorectal cancers (CRCs) and leads to the widespread accumulation of somatic mutations in tumors, including insertions and deletions at DNA microsatellite sequences.[6,7] Such a theoretically immunogenic profile is corroborated by the presence of numerous intraepithelial lymphocytes in these cancers, in contrast to MMR-proficient cancers.[8,9] Nevertheless, not all MMR-deficient CRCs respond to immune checkpoint blockade, while MMR-proficient CRCs are insensitive to this therapy.

To understand the mechanisms that determine responses to current immunotherapies and for the design of alternative approaches, it is crucial to characterize the cancer microenvironment with multidimensional approaches that allow the simultaneous identification and characterization of immune cell populations across multiple lineages.[10,11] Mass cytometry allows a detailed single-cell characterization of adaptive and innate immune landscapes, thereby providing a unique platform to discriminate immune cell subsets that can be exploited in an immunotherapeutic setting.

We performed an in-depth characterization of immune landscapes across CRC tissues, tumor-associated lymph nodes, colorectal healthy mucosa, and peripheral blood samples from 31 CRC patients by high-dimensional single-cell mass cytometry. We revealed tumor tissue-specific immune signatures across the adaptive and innate compartments, and discovered a previously unappreciated innate immune cell population implicated in anti-tumor immunity that strongly differentiated immunogenic (MMR-deficient) from non-immunogenic (MMR-proficient) CRCs.

RESULTS

Tumor-resident immune cell populations derive from multiple lineages

Mass cytometric analysis of 36 immune cell markers was performed on single-cell suspensions isolated from cancer and healthy tissues of CRC patients (**Table S1, Table S2**). To decipher their immune composition, we performed HSNE analysis in Cytosplore on all acquired $CD45^+$ cells of the discovery cohort ($8.9*10^6$ cells in total) (**Figure 1A, Figure S1**). Based on the density features of the HSNE-embedded landmarks, 7 major immune cell clusters were identified by unsupervised GMS clustering, which corresponded to naive and memory (based on CD45RO and CCR7 expression) $CD4^+$ and $CD8^+/\gamma\delta$ T cells, B cells, Lin^-CD7^+ ILCs, and myeloid cells (**Figure 1A,B**). Memory $CD4^+$ and $CD8^+$ T cells, as well as myeloid cells, were dominant immune lineages in the tumor microenvironment, while B cells, Lin^-CD7^+

ILCs, and naive CD4+ and CD8+ T cells were present at a lower extent (**Figure S2**). The HSNE analysis also unveiled the presence of several tumor tissue-specific, phenotypically distinct landmarks within the memory CD4+ T cell, CD8+/γδ T cell, Lin−CD7+ ILC, and myeloid cell compartments (**Figure 1A**).

All 7 major immune lineages were analyzed in detail by hierarchical exploration of the data in HSNE. As an example, the embedding of the memory CD8+/γδ T cell compartment is shown in **Figure 1C**. Altogether, analysis of these 7 major immune lineages yielded 220 distinct immune cell clusters, of which 2 consisted of less than 100 cells and were excluded from further analysis. All acquired CD45+ cells of the validation cohort ($6.6*10^6$ cells in total) were subsequently classified into these pre-identified immune cell clusters based on their phenotype (see **Methods**).

The mass cytometric analysis was accompanied by single-cell RNA-sequencing of CD45+ cells from 7 CRC tissues. Seven immune cell clusters could be detected based on transcriptomic profiles (**Figure 1D**), corresponding to B cells, CD8+ and CD4+ T cells, ILCs, myeloid cells, proliferating cells, and plasma B cells (**Figure 1D,E**).

Activated CD8+ and γδ T cells are tumor tissue-specific and enriched in mismatch repair-deficient colorectal cancers

Hierarchical clustering analysis revealed that memory CD8+/γδ T cell phenotypes clustered in a tissue-specific manner (**Figure 2A**). Two CD8+CD103+PD-1+ populations (#60 and 96), distinguished by CD161 expression, were present in tumor tissues (constituting up to 28.2% of CD45+ cells) and infrequent in all other samples (**Figure 2B,C**), with the exception of one lymph node sample that was found to be infiltrated by tumor cells upon histological examination (data not shown). These CD8+CD103+PD-1+ cells were further characterized by the co-expression of CD69, FAS, HLA-DR, and CD38 (**Figure 2B**). Interestingly, the CD161− counterpart of CD8+CD103+PD-1+ T cells (#60) was particularly abundant in MMR-deficient tumors as compared to MMR-proficient tumors (**Figure S3**). Within the CD8+CD103+PD-1+CD38+ subset, we observed co-expression of CD39 (**Figure S3**), a marker that has recently been found to identify tumor-reactive CD8+ T cells.[12,13] Next to these tumor-resident cells, a cluster (#61) with a similar phenotype but lacking HLA-DR, PD-1, FAS, and possessing a lower expression of CD38 was present in both tumor and healthy colorectal samples (**Figure 2B,C**), and may represent a non-activated counterpart. Single-cell RNA-sequencing revealed that CD8+ T cells in colorectal tumors expressed cytolytic molecules (e.g. *GZMA*, *GZMB*, *GZMH, PRF1*) (**Figure 2D**). Furthermore, they displayed expression of the immune checkpoint molecule *LAG3* (**Figure 2D**).

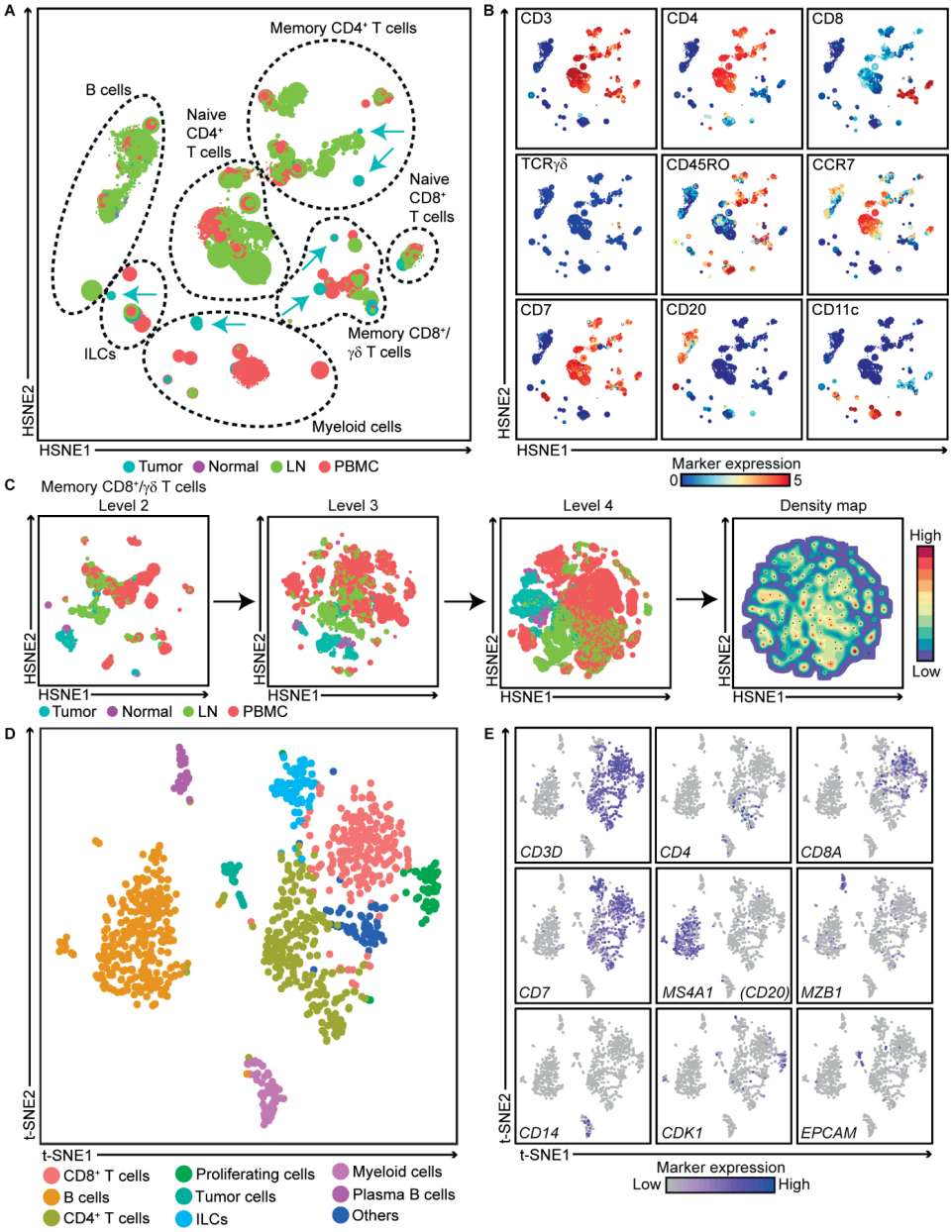


Figure 1. Tumor-resident immune cell populations derive from multiple lineages.
A and B. HSNE embedding showing $7.5*10^4$ landmarks representing immune cells ($8.9*10^6$ cells) isolated from CRC tissues (N=19), tumor-associated lymph nodes (N=17), colorectal healthy mucosa (N=4), and peripheral blood (N=9) samples from the discovery cohort. Colors represent the different tissue types (**A**) and the relative expression of indicated immune lineage markers (**B**). Arrows indicate the HSNE location of phenotypically distinct tumor-resident immune cell populations. **C.** Example of an HSNE analysis of $7.4*10^2$ landmarks representing $1.1*10^6$ cells from the memory CD8⁺/γδ T cell compartment as identified in (**A**). All landmarks are selected and embedded at the next, more detailed levels showing a finer granularity of structures with $5.0x10^3$ landmarks at level 2, to $3.0*10^4$ landmarks at level 3, and $1.6*10^5$ landmarks at level 4. Phenotypically distinct immune cell clusters were identified by unsupervised GMS clustering based on the density features. Black dots indicate the centroids of the identified clusters. **D and E.** t-SNE

embedding showing 1,079 cells from CRC tissues (N=7) analyzed by single-cell RNA-sequencing. Colors represent the different clusters (**D**) and the log-transformed expression levels of indicated immune lineage markers (**E**). Each dot represents a single cell.

Strikingly, a TCRγδ$^+$CD103$^+$PD-1$^+$ population (#99) was almost exclusively found in MMR-deficient tumors, constituting up to 8.4% of CD45$^+$ cells (**Figure 2B,C**). These γδ T cells had a phenotype similar to the CD8$^+$CD103$^+$PD-1$^+$ cells, as defined by co-expression of CD69, FAS, CD38, and HLA-DR (**Figure 2B**). An HLA-DR$^-$PD-1$^-$ counterpart of these cells (#97 and 101) was also observed in colorectal healthy mucosa and MMR-proficient tumors, and may represent a non-activated form of the CD103$^+$PD-1$^+$ γδ T cells in the tumor microenvironment (**Figure 2B,C**). We analyzed the cytotoxic potential of the tumor-resident γδ T cells by flow cytometry and determined that these were capable of producing IFN-γ and granzyme B/perforin upon stimulation with PMA/ionomycin (**Figure S4**).

3

ICOS$^+$ and activated CD4$^+$ T cells are dominant, tumor tissue-specific T cell populations in both mismatch repair-deficient and -proficient colorectal cancers

Next, we determined the cell surface phenotype of memory CD4$^+$ T cells in CRC patients. Memory CD4$^+$ T cells also distributed in a tissue-specific manner (**Figure 3A**). Here, a large population of CD4$^+$ICOS$^+$CD27$^-$ cells (#20 and 58) constituted up to 21.1% of CD45$^+$ cells in CRCs, while being absent in all other tissues with the exception of tumor-positive lymph node samples (**Figure 3B,C**). Part of this population co-expressed CD161 and PD-1 (#58), whereas the other part was negative for these markers but expressed high levels of CD25 (#20), indicative of a regulatory-like phenotype (**Figure 3B**). Flow cytometry analysis confirmed the expression of FOXP3 in 91-98% of ICOS$^+$ CD4$^+$CD45RO$^+$CD25$^+$CD127^{low} T cells in colorectal tumors (**Figure S5**). Interestingly, the ICOS$^+$ CD4$^+$ T cells were present in MMR-deficient as well as MMR-proficient tumors to a similar extent (**Figure 3B,C**).

In addition, CD4$^+$CD103$^+$PD-1$^+$ cells (#85 and 86), which constituted up to 23.8% of CD45$^+$ cells, were also enriched in tumor tissues (**Figure 3B,C**). Strikingly, several features of these cells mirrored our observations in the CD8$^+$/γδ compartment, including a tissue-resident phenotype defined by co-expression of CD69, FAS, CD38, and HLA-DR (**Figure 3B**). Moreover, expression of CD161 also subdivided CD4$^+$CD103$^+$PD-1$^+$ T cells into a positive (#85) and negative (#86) population, where CD161$^-$ cells were more abundant in MMR-deficient as compared to MMR-proficient tumors (**Figure S3**). In contrast to the tumor-resident CD8$^+$ and γδ T cells, a non-activated counterpart could not be detected for these cells.

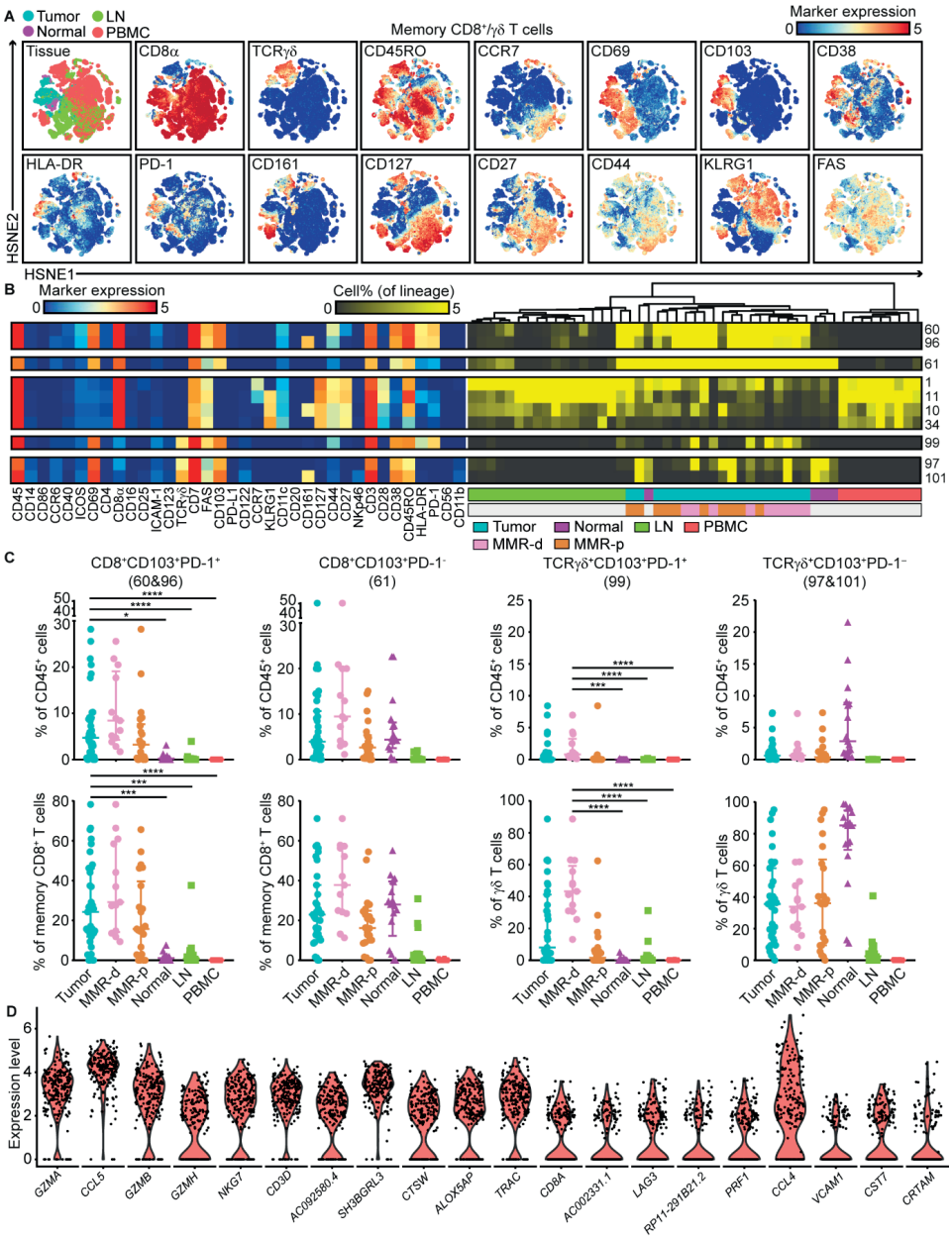


Figure 2. Activated CD8⁺ and γδ T cells are tumor tissue-specific and enriched in mismatch repair-deficient colorectal cancers.

A. HSNE embedding of $1.6*10^5$ landmarks representing the memory CD8⁺/γδ T cell compartment ($1.1*10^6$ cells) from the discovery cohort of CRC patients colored by tissue type (first plot) and relative expression of indicated markers. **B.** A heatmap showing median marker expression values (left) and a heatmap showing frequencies of selected memory CD8⁺/γδ T cell clusters (right). Hierarchical clustering was performed on cluster frequencies using Spearman's rank correlation. Color bars indicate tissue type. **C.** Frequencies of selected memory CD8⁺/γδ T cell clusters among CRC patients' tissues (N=35, further subdivided into MMR-deficient (N=13) and MMR-proficient (N=22)), colorectal healthy mucosa (N=17), tumor-associated lymph nodes (N=26), and peripheral blood (N=19) as percentage of total CD45⁺ cells

(upper panel) and memory CD8$^+$ or γδ T cells (lower panel). Cluster IDs correspond to the ones in (**B**). Bars indicate median ± IQR. Each dot represents an individual sample. Data from 22 independent experiments with mass cytometry. *P<0.05, ***P<0.001, ****P<0.0001 by Kruskal-Wallis test with Dunn's test for multiple comparisons. **D.** Violin plot showing log-transformed expression levels of the top 20 differentially expressed genes within CD8$^+$ T cells (N=217) analyzed by single-cell RNA-sequencing on CD45$^+$ cells from 7 tumors (**Figure 1D**). Each dot represents a single cell.

While ICOS$^+$ regulatory T cells (Tregs) were tumor tissue-specific, ICOS$^-$CD25$^+$CD127$^-$ Tregs (#13-73) were found in both tumor-associated lymph nodes and CRC tissues (**Figure 3B,C**). Lastly, immune cell populations such as CD4$^+$CD27$^+$CD127$^+$ central memory (CCR7$^+$CD45RO$^+$) cells (#1-37) were more abundant in peripheral blood and lymph nodes (**Figure 3B,C**). The expression of *ICOS* on CD4$^+$ T cells was confirmed by single-cell RNA-sequencing, which also revealed the expression of *TNFRSF4 (OX40R)* and *TNFRSF18 (GITR)* (**Figure 3D**). t-SNE analysis revealed the co-expression of all three immunotherapeutic targets by CD4$^+$ T cells (**Figure S6**).

CD127$^-$CD56$^+$CD45RO$^+$ ILCs are the prevalent ILC population in mismatch repair-deficient colorectal tumors

Mass cytometric profiles of the innate lymphoid compartment revealed the presence of three distinct Lin$^-$CD7$^+$ cell clusters: CD127$^-$CD56$^+$CD45RO$^-$ natural killer (NK) cells (90.4%), CD127$^+$ ILCs (3.4%), and a cluster of CD127$^-$CD56$^+$CD45RO$^+$ cells (6.2%) (**Figure 4A**). Analysis of cluster frequencies demonstrated that CD56^{dim}CD16^{bright} NK cells (#33-4) were present in high frequencies in peripheral blood, whereas CD56^{bright}CD16^{dim} NK cells (#10-82) were the dominant NK-type in lymph node samples (**Figure 4B,C**). CD127$^+$ ILCs (#6-9) were more abundant in healthy mucosa, lymph nodes and MMR-proficient tumors, and displayed a KLRG1$^-$ phenotype, characteristic of ILC3 cells (**Figure 4B,C**). Strikingly, the CD127$^-$CD56$^+$CD45RO$^+$ ILCs (#87,95,92,97) were enriched in tumor tissues, accounting for up to 80% of the innate lymphoid compartment (**Figure 4B,C**). Moreover, they were particularly abundant in MMR-deficient tumors, especially CD161$^-$ populations (#95, 92 and 97) (**Figure 4B,C**). The CD127$^-$CD56$^+$CD45RO$^+$ ILC population has recently been identified in human fetal intestine as intermediate-ILCs.[14] Consistent with that work, hierarchical clustering positioned the CD127$^-$CD56$^+$CD45RO$^+$ ILCs in between NK cells and CD127$^+$ ILCs (**Figure 4B**). We observed co-expression of CD69 and CD103 on all CD127$^-$CD56$^+$CD45RO$^+$ ILCs, but differential expression of CD16, ICAM-1, FAS, CD11c, CD161, CD44 and HLA-DR, indicative of further heterogeneity within this cell cluster (**Figure 4B**).

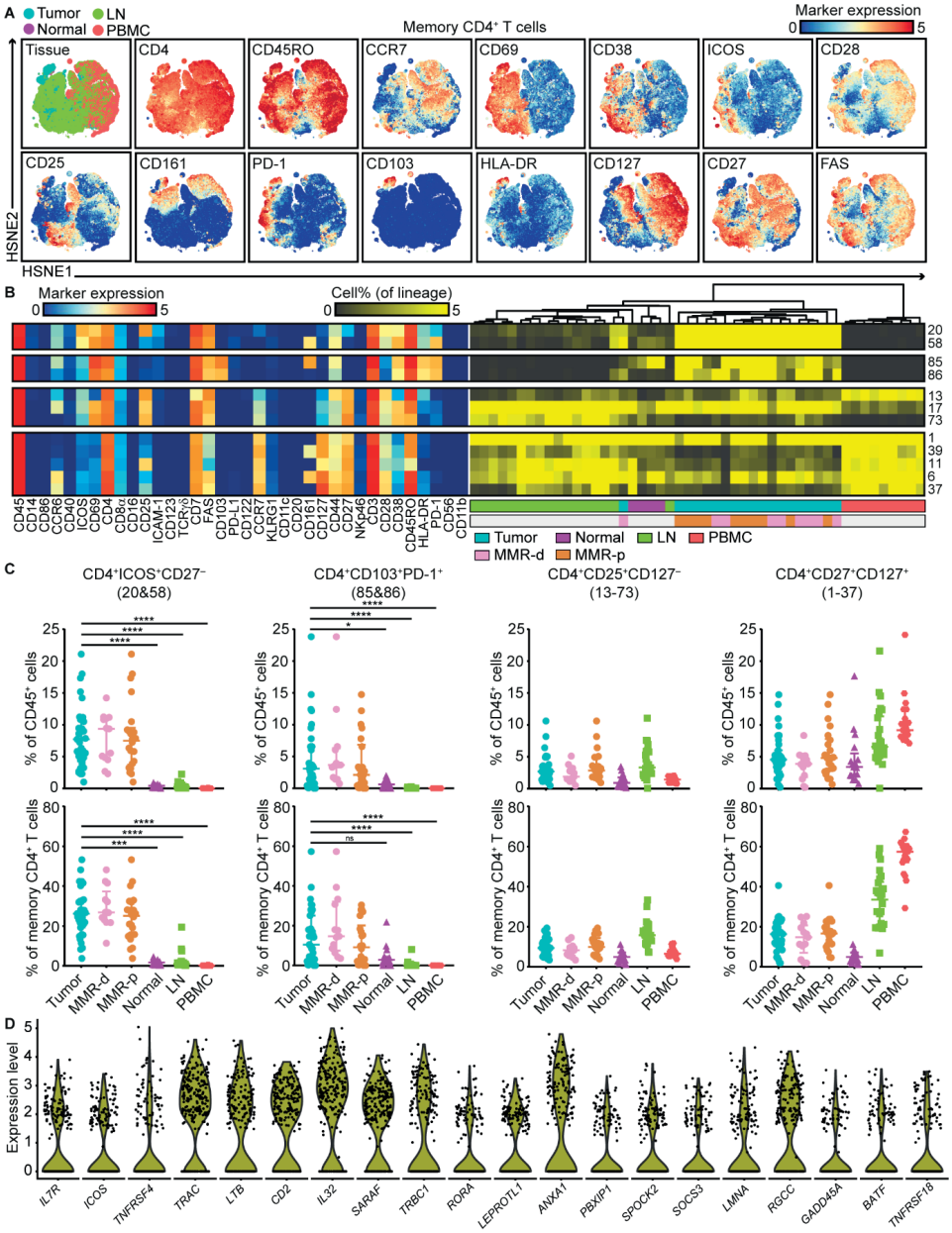


Figure 3. ICOS⁺ and activated CD4⁺ T cells are dominant, tumor tissue-specific T cell populations in both mismatch repair-deficient and -proficient colorectal cancers.

A. HSNE embedding of $3.1*10^5$ landmarks representing the memory CD4⁺ T cell compartment ($2.0*10^6$ cells) from the discovery cohort of CRC patients colored by tissue type (first plot) and relative expression of indicated markers. **B.** A heatmap showing median marker expression values (left) and a heatmap showing frequencies of selected memory CD4⁺ T cell clusters (right). Hierarchical clustering was performed on cluster frequencies using Spearman's rank correlation. Color bars indicate tissue type. **C.** Frequencies of selected memory CD4⁺ T cell clusters among CRC patients' tissues (N=35, further subdivided into MMR-deficient (N=13) and MMR-proficient (N=22)), colorectal healthy mucosa (N=17), tumor-associated lymph nodes (N=26), and peripheral blood (N=19) as percentage of total CD45⁺ cells

(upper panel) and memory CD4⁺ T cells (lower panel). Cluster IDs correspond to the ones in (**B**). Bars indicate median ± IQR. Each dot represents an individual sample. Data from 22 independent experiments with mass cytometry. *P<0.05, ***P<0.001, ****P<0.0001 by Kruskal-Wallis test with Dunn's test for multiple comparisons. **D.** Violin plot showing log-transformed expression levels of the top 20 differentially expressed genes within CD4⁺ T cells (N=245) analyzed by single-cell RNA-sequencing on CD45⁺ cells from 7 tumors (**Figure 1D**). Each dot represents a single cell.

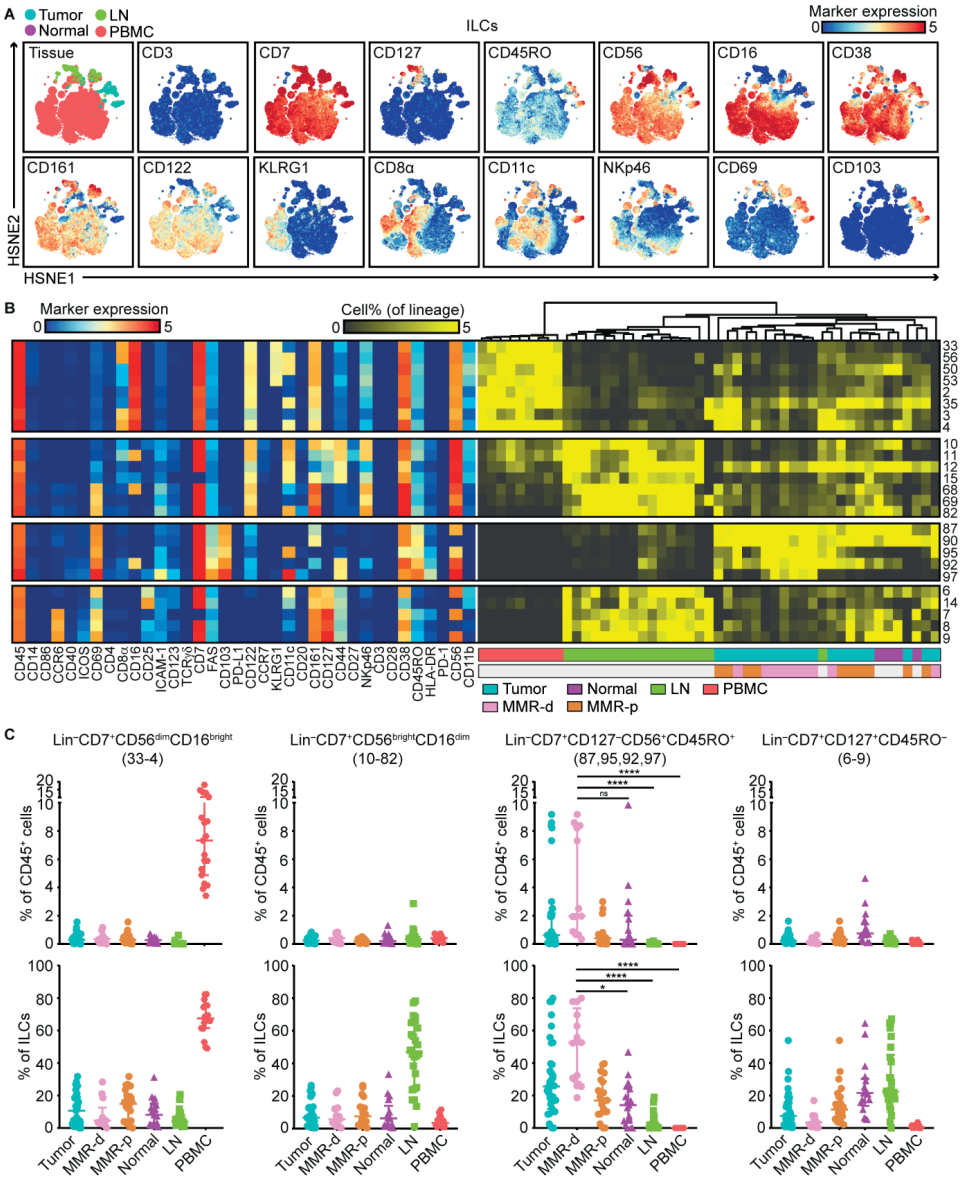


Figure 4. CD127⁻CD56⁺CD45RO⁺ ILCs are the prevalent ILC population in mismatch repair-deficient colorectal tumors.

A. HSNE embedding of 5.5*10⁴ landmarks representing the innate lymphoid compartment (0.4*10⁶ cells) from the discovery cohort of CRC patients colored by tissue type (first plot) and relative expression of indicated markers. **B.** A heatmap showing median marker expression values (left) and a heatmap showing frequencies of selected ILC

clusters (right). Hierarchical clustering was performed on cluster frequencies using Spearman's rank correlation. Color bars indicate tissue type. **C.** Frequencies of selected innate lymphoid clusters among CRC patients' tissues (N=35, further subdivided into MMR-deficient (N=13) and MMR-proficient (N=22)), colorectal healthy mucosa (N=17), tumor-associated lymph nodes (N=26), and peripheral blood (N=19) as percentage of total CD45$^+$ cells (upper panel) and ILCs (lower panel). Cluster IDs correspond to the ones in (**B**). Bars indicate median ± IQR. Each dot represents an individual sample. Data from 22 independent experiments with mass cytometry. NS, not significant, *P<0.05, ****P<0.0001 by Kruskal-Wallis test with Dunn's test for multiple comparisons.

Tumor-resident ILCs are involved in the anti-tumor immune response

Single-cell RNA-sequencing unveiled high expression levels of cytotoxic molecules (e.g. *GNLY*, *PRF1*, *GZMA*, *GZMB*) in the ILC cluster (**Figure 5A**). In addition, we observed the presence of transcripts for a member of the killer-cell immunoglobulin-like receptor (KIR) family, *KIR2DL4* (**Figure 5A**). We performed additional single-cell RNA-sequencing on CD45$^+$ cells from one MMR-deficient tumor with high numbers of Lin$^-$CD7$^+$CD127$^-$CD56$^+$CD45RO$^+$ ILCs (70% of the ILC cluster), as revealed by mass cytometry data. Here, we also observed high expression levels of cytotoxic molecules (e.g. *GNLY*, *PRF1*, *GZMA*) as well as the expression of *KIR2DL4* and *KIR3DL2* in the ILC cluster (**Figure S7**). Cell surface expression of KIRs was confirmed by flow cytometry in Lin$^-$CD7$^+$CD127$^-$CD56$^+$CD45RO$^+$ ILCs from this tumor (**Figure S7**).

To further investigate functional properties of tumor-resident lymphocytes, we designed a flow cytometry antibody panel to analyze the cytotoxic potential of Lin$^-$CD7$^+$CD127$^-$CD56$^+$CD45RO$^+$ ILCs, Lin$^-$CD7$^+$CD127$^-$CD56$^+$CD45RA$^+$ NK cells, and memory CD8$^+$ T cells in CRC tissues (**Table S3**). Strikingly, up to 82.3% of unstimulated CD127$^-$CD56$^+$CD45RO$^+$ ILCs displayed granzyme B/perforin expression in the tumor tissues (**Figure 5B**). Granzyme B/perforin expression by the ILCs was most abundant in MMR-deficient cancers as compared to MMR-proficient cancers (**Figure 5C**). Interestingly, the cytotoxic capacity of CD127$^-$CD56$^+$CD45RO$^+$ ILCs was accompanied by similar profiles in CD127$^-$CD56$^+$CD45RA$^+$ NK cells and memory CD8$^+$ T cells across samples (**Figure 5C**), suggesting a coordinated cytotoxic innate and adaptive immune response in CRC tissues.

To investigate the spatial localization of the ILCs in CRCs, we applied 6-color multispectral immunofluorescence to frozen tissue sections of 4 MMR-deficient and 4 MMR-proficient CRCs (**Table S4**). We simultaneously detected CD3, TCRαβ, CD127, CD7, CD45RO, and DAPI. We identified CD3$^-$TCRαβ$^-$CD127$^-$CD7$^+$CD45RO$^+$ ILCs in the tumors (**Figure 6A,B**), and observed an increased presence of these cells in MMR-deficient as compared to MMR-proficient CRCs (**Figure 6C**). Interestingly, the CD3$^-$TCRαβ$^-$CD127$^-$CD7$^+$CD45RO$^+$ ILCs frequently displayed an intraepithelial localization in agreement with their CD103$^+$CD69$^+$ tissue-resident phenotype (**Figure 6A**).

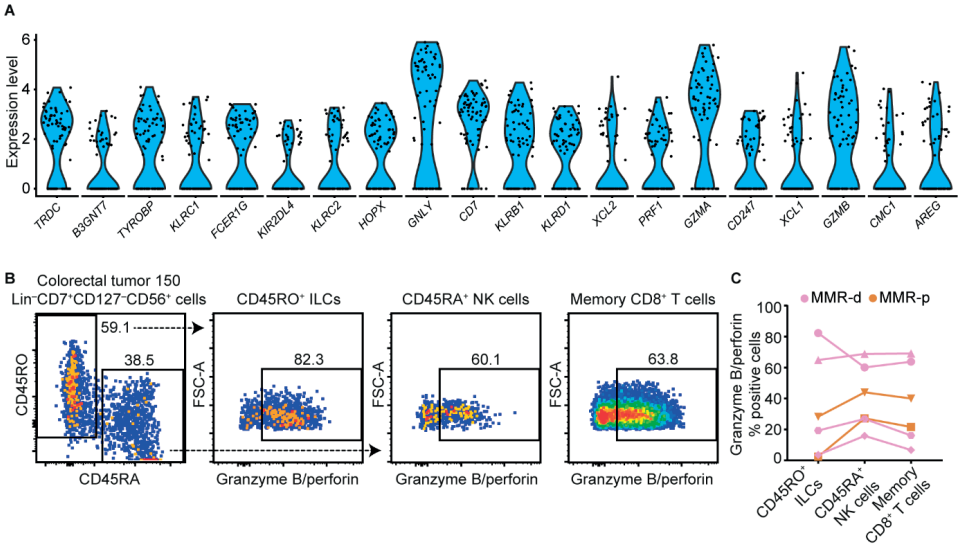


Figure 5. Tumor-resident ILCs are involved in the anti-tumor immune response.
A. Violin plot showing log-transformed expression levels of the top 20 differentially expressed genes within ILCs (N=74) analyzed by single-cell RNA-sequencing on CD45$^+$ cells from 7 tumors (**Figure 1D**). Each dot represents a single cell. **B.** Representative plots of a MMR-deficient tumor sample analyzed by flow cytometry without stimulation showing the distinction between CD45RO$^+$ ILCs and CD45RA$^+$ NK cells within Lin$^-$CD7$^+$CD127$^-$CD56$^+$ cells (first plot), and their expression of cytotoxic molecules. **C.** Granzyme B/perforin expression in different immune cell populations of CRC tissues (N=6, of which 4 MMR-deficient and 2 MMR-proficient). Dot shape indicates similar tumor samples. Data from three independent experiments with flow cytometry.

Immune-system-wide analysis reveals correlations between innate and adaptive immune cell subsets in colorectal cancer

Lastly, we integrated the identified immune cell clusters across all major immune lineages (N=218) in one immune-system-wide analysis to characterize the samples according to tissue type, MMR status, and available clinico-pathological parameters. The integrated t-SNE analysis confirmed the unique immune composition in the different tissue types, and visualized the top ten ranked immune cell clusters contributing to the distinctive clustering patterns of the samples (**Figure S8**). No association was observed with clinical stage while differences related to tumor location and HLA class I expression can be attributed to features that distinguish MMR-deficient and -proficient CRCs (**Figure S8**).

Spearman's rank correlation analysis performed on the top ten ranked unique immune cell clusters of each tissue type revealed strong correlations between the presence of CD127$^-$CD56$^+$CD45RO$^+$ ILCs (ILC97,92,95) and the presence of CD103$^+$PD-1$^+$ cytotoxic (CD8memory60,96), helper (CD4memory85,86), and γδ (TCR γδ99) T cell populations in MMR-deficient CRCs (**Figure 7**, **Table S5**). In contrast, MMR-proficient tumors were characterized by the presence of several myeloid populations (**Figure 7**, **Table S5**).

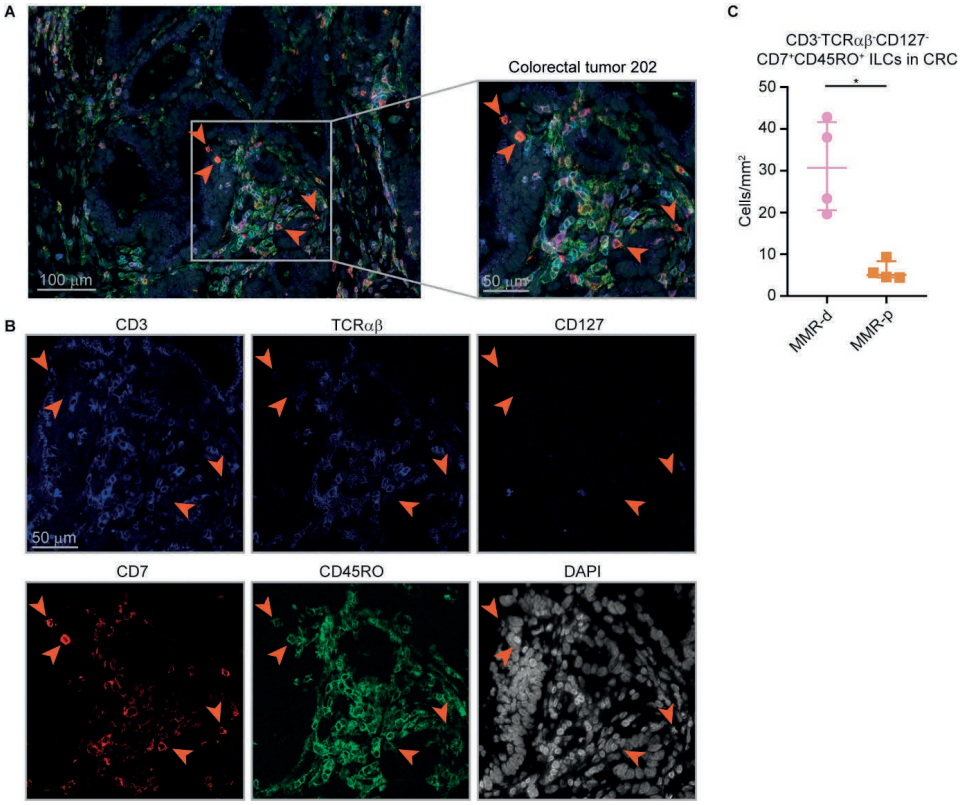


Figure 6. Higher cell density of CD127⁻CD45RO⁺ ILCs in mismatch repair-deficient colorectal cancers.
A and B. Representative image of the immunofluorescence microscopic detection of CD3⁻TCRαβ⁻CD127⁻CD7⁺CD45RO⁺ ILCs in a MMR-deficient tumor, showing CD3 (colored in blue), TCRαβ (colored in blue), CD127 (colored in blue), CD7 (colored in red), CD45RO (colored in green), and DAPI (colored in grey) as nuclear counterstain. **C.** Frequencies of CD3⁻TCRαβ⁻CD127⁻CD7⁺CD45RO⁺ ILCs in 4 MMR-deficient and 4 MMR-proficient CRCs. *P<0.05 by Mann-Whitney U test.

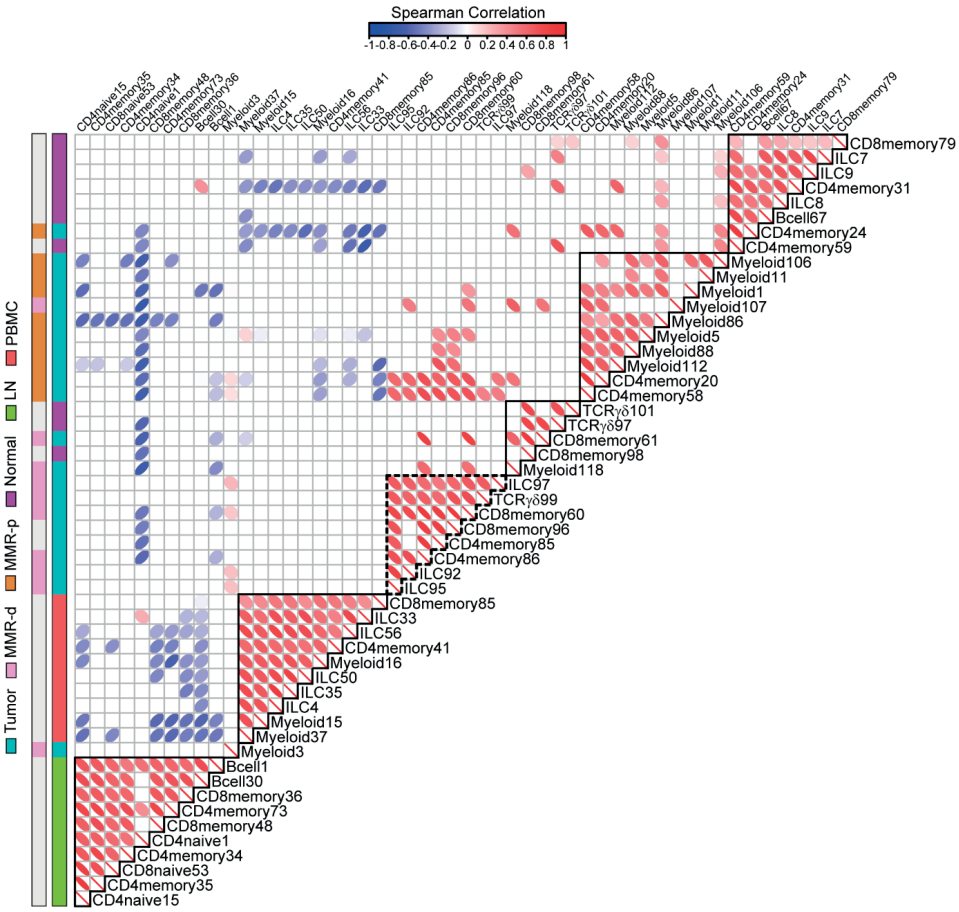


Figure 7. Immune-system-wide analysis reveals correlations between innate and adaptive immune cell subsets in colorectal cancer.

Matrix showing correlations (Spearman's ρ, see **Supplemental methods**) between unique top ten ranked immune cell clusters for each tissue type (shown in **Figure S8**) based on cell percentages (of total CD45$^+$ cells) corresponding to 97 samples from 31 CRC patients. Color and shape of the ellipses in the heatmap indicate the strength of the correlation. Only significant correlation coefficients are shown. Color bars indicate tissue type. Coefficient and *P*-values of correlations for CRC tissues are shown in **Table S5**. *P*-values were adjusted for multiple testing using the Benjamini-Hochberg procedure. Data from 22 independent experiments with mass cytometry.

DISCUSSION

We applied mass cytometry to comprehensively analyze the immune landscape of CRCs at single-cell level in tumor and healthy tissues. Our analysis revealed tumor tissue-specific immune signatures across the innate and adaptive immune compartments of CRC. Immunohistochemistry, flow cytometry, and recent transcriptomic approaches have provided insight into the complexity of tumor immune landscapes.[15-19] However, the number of markers that can be simultaneously assayed in immunohistochemistry or flow cytometry is limited, and bulk transcriptomic studies do not allow for discrimination of phenotypes at

the cellular level.[20,21] In mass cytometry over 40 markers can be simultaneously analyzed at single-cell level, providing a unique opportunity to obtain a comprehensive overview of tumor-resident lymphocytes.[22,23] Here, we combined mass cytometry phenotypes with functional, transcriptional, and spatial analyses of tumor-resident immune cell populations in CRC.

Within the innate compartment, we observed that a previously unappreciated innate lymphoid population, Lin$^-$CD7$^+$CD127$^-$CD56$^+$CD45RO$^+$ ILCs, is enriched in MMR-deficient tumors and displayed cytotoxic activity. *In-situ* detection of the ILCs confirmed a higher cell density in MMR-deficient CRCs, and showed a frequent intraepithelial localization. This is in line with their tissue-resident phenotype (CD103$^+$CD69$^+$), and supports an active role for these cells in the anti-tumor immune response. The ILCs resemble previous descriptions of TCR$^-$CD103$^+$ cells in mice that were found to express granzyme B.[24] Additionally, a unique subset of NK cells has been found in several human tissues and was described as NKp44$^+$CD103$^+$ intraepithelial ILC1-like.[25,26] In contrast to NKp44$^+$CD103$^+$ ILC1, the CD127$^-$CD56$^+$CD45RO$^+$ ILCs identified here lacked CD122 and NKp46 expression (**Figure 4B**), and showed low levels of NKp44 (data not shown). These variable marker expression patterns most likely represent additional levels of plasticity and heterogeneity within ILC subsets. Single-cell RNA-sequencing revealed the presence of transcripts for *KIR2DL4* and *KIR3DL2* in the ILC cluster, which hints toward potential activation mechanisms.[27] Common ligands of KIRs include HLA class I molecules,[28,29] and loss of HLA class I expression has been described to occur in the majority of MMR-deficient CRCs.[30-32] It is tempting to speculate that CD127$^-$CD56$^+$CD45RO$^+$ ILC-mediated cytotoxicity toward such HLA-loss variants may contribute to the anti-tumor response in MMR-deficient CRCs, a link that requires further investigation.

The presence of CD127$^-$CD56$^+$CD45RO$^+$ ILCs strongly correlated with tissue-resident CD103$^+$CD69$^+$ γδ T cells co-expressing activation markers HLA-DR, CD38, and PD-1 in MMR-deficient CRCs. It has been shown that human peripheral blood γδ T cells can express PD-1 and exhibit natural killer-like activity.[33] The expression of PD-1, in conjunction with their cytotoxic potential, suggest an active role of tumor-resident γδ T cells in the anti-tumor immune response and potentially as targets for PD-1 checkpoint blockade. This will be subject of further studies.

Within the adaptive compartment, we found dominant, tumor tissue-specific CD8$^+$ and CD4$^+$ T cell populations that displayed a highly similar activated tissue-resident phenotype. Such CD8$^+$ T cell populations have been described in ovarian cancer,[34,35] lung cancer,[36] and recently in melanoma,[37] cervical carcinoma[38] and CRC,[12,13] and their presence was associated with an improved clinical prognosis. Single-cell RNA-sequencing revealed that CD8$^+$ T cells in colorectal tumors showed a cytotoxic profile, indicative of potential anti-tumor reactivity. In addition, we found a dominant tumor tissue-specific population of ICOS$^+$ CD4$^+$ T cells. ICOS belongs to the CD28/CTLA-4 family and serves as a co-stimulatory molecule for T cell activation.[39] Activation of ICOS by agonists has been proposed for anti-cancer treatment.[40] Here, we identified a CD161$^+$PD-1$^+$ as well as a CD25$^+$ population of tumor-resident ICOS$^+$ CD4$^+$ T cells. The latter corresponds to a regulatory T cell subset displaying high levels of

FOXP3 expression, that, interestingly, expressed higher levels of ICOS as compared to the CD161^{+}PD-1^{+} counterpart. The use of ICOS agonists may, therefore, also result in activation of ICOS^{+} T cells with suppressive and regulatory properties in the tumor microenvironment. In contrast to the tumor-resident CD8^{+} T cells, ICOS^{+} CD4^{+} T cells were present in both MMR-deficient and MMR-proficient tumors to a similar extent.

We observed CD161^{+} and CD161^{-} counterparts of tumor-resident cytotoxic and helper T cells, and CD127^{-}CD56^{+}CD45RO^{+} ILCs. CD161 has been shown to mark a subset of tissue-resident memory CD8^{+} T cells with enhanced effector function and cytokine production.[41,42] In our study, the CD161^{-} counterpart of the tumor-resident T cell and ILC populations was particularly enriched in MMR-deficient CRCs as compared to MMR-proficient CRCs. The functional relevance of this observation will be subject of future studies. Nevertheless, we observed increased CD161 expression in PD-1 high cells as compared to PD-1 intermediate/negative cells for tumor-resident CD8^{+} and CD4^{+} T cell populations (**Figure S9**). As PD-1 high cells in human cancer have been associated with a state of T cell dysfunction,[43-45] CD161 expression could be an additional marker for this functional state.

Interestingly, we identified what could be the non-activated counterparts of the CD103^{+}PD-1^{+} cytotoxic and γδ T cells in both tumor and healthy colorectal tissues. Mobilization and activation of these cells from the colorectal healthy mucosa to the tumor tissue may be beneficial for immunotherapy in CRC. Strikingly, while lymph nodes are traditionally viewed as key players of anti-tumor immune responses, we did not detect non-activated precursors of tumor-resident immune cell populations in the lymph node samples, with the exception of tumor-positive lymph nodes. Furthermore, we observed that lymph nodes harboured a large population of CD4^{+}CD25^{+}CD127^{-} Tregs, suggesting they might be a primary source of Tregs in the cancer microenvironment. The tumor-resident immune cell populations were also not mirrored in peripheral blood, although the in-depth investigation of their presence in these tissues with complementary approaches should be conducted.

It should be noted that the mass cytometry antibody panel was primarily developed to characterize T cell, γδ T cell and ILC compartments, and in future studies additional efforts are required to further explore the myeloid and B cell compartment. Furthermore, the number and pattern of infiltrating lymphocytes can be influenced by various tumor characteristics. In this study we have shown profound differences in lymphocytic infiltration that distinguish MMR-deficient from MMR-proficient CRCs. Other factors not investigated in this study that can influence the infiltration of lymphocytes in tumors include for instance occurrence of somatic mutations (neoantigens) and the co-occurrence of inflammatory bowel disease. Although the results are of preliminary nature, they point to the involvement of additional subsets than T cells in immune responses to CRC, particularly ILCs and γδ T cells. This is especially relevant in the context of responses to checkpoint blockade therapy in absence of HLA class I expression.[46] Future approaches might opt for an in-depth investigation of these specific lineages for a detailed characterization of phenotypes that complement the markers used in this study. The next step will be to investigate the involvement of these subsets in

the clinical setting of patients treated by checkpoint blockade.

In conclusion, we identified a previously unappreciated innate immune cell population that was specifically enriched in CRC tissues, displayed cytotoxic activity, and strongly contributed to a data-driven distinction between immunogenic (MMR-deficient) and non-immunogenic (MMR-proficient) tumors. Furthermore, we revealed strong correlations between the presence of these innate cells and tumor-resident $CD8^+$, $CD4^+$, and $\gamma\delta$ T cells with an activated phenotype in MMR-deficient tumors that together may play a critical role in tumor control.

METHODS

Human samples

Primary CRC tissues (N=35, of which 22 MMR-proficient and 13 MMR-deficient) with matched tumor-associated lymph nodes (N=26), colorectal healthy mucosa (N=17), and pre-surgical peripheral blood samples (N=19) from 31 CRC patients were processed for this study (**Table S1**). All patients were treatment-naïve except five rectal cancer patients which received neo-adjuvant therapy (**Table S1**). One patient was diagnosed with multiple primary colorectal tumors (N=5) at different locations, all of which were included in the study (**Table S1**). No patient with a previous history of inflammatory bowel disease was studied. To account for tumor heterogeneity, macroscopic sectioning from the lumen to the most invasive area of the tumor was performed for further processing. This study was approved by the Medical Ethical Committee of the Leiden University Medical Center (protocol P15.282), and patients provided written informed consent. All specimens were anonymized and handled according to the ethical guidelines described in the Code for Proper Secondary Use of Human Tissue in the Netherlands of the Dutch Federation of Medical Scientific Societies.

Tissue processing and mass cytometry antibody staining

Details on tissue processing and mass cytometry antibody staining are available in **Supplemental methods** and **Table S2**.

Mass cytometry data analysis

Mass cytometry experiments were performed with a discovery and validation cohort of CRC patients. The discovery cohort consisted of 19 CRC tissues, 17 tumor-associated lymph nodes, 4 colorectal healthy mucosa, and 9 peripheral blood samples. Single, live $CD45^+$ cells were gated in Cytobank[47] (**Figure S1**). $CD45^+$ cells were sample-tagged, hyperbolic ArcSinh transformed with a cofactor of 5, and subjected to dimensionality reduction analysis in Cytosplore.[48] Of the 39 antibodies included in the panel, 36 showed clear discrimination between positive and negative cells (**Figure S1**). Major immune lineages (**Figure 1A,B**) were identified at the overview level of a 5-level Hierarchical Stochastic Neighbor Embedding (HSNE) analysis[49,50] on $CD45^+$ data from all samples ($8.9*10^6$ cells) with default perplexity and iterations (30 and 1,000, respectively). Naive and memory $CD4^+$ and $CD8^+/\gamma\delta$ T cell, B cell, Lin^-CD7^+ innate lymphoid cell (ILC), and myeloid cell lineages were analyzed in a data-driven manner up to a maximum number of $0.5*10^6$ landmarks.[50] Clustering of the data was performed by Gaussian Mean Shift (GMS) clustering in Cytosplore, and an algorithm was run that merged clusters showing high similarity in ArcSinh5-transformed median expression of all markers (<1). Hierarchical clustering on cell frequencies was performed in Matlab using Spearman's rank correlation.

The validation cohort consisted of 16 CRC tissues, 9 tumor-associated lymph nodes, 13 colorectal healthy mucosa, and 10 peripheral blood samples. Single, live $CD45^+$ cells were hyperbolic ArcSinh transformed with a cofactor of 5, and classified into the pre-identified immune cell clusters of the discovery cohort based on similarity in marker expression. To obtain consistent cell clusters across both cohorts, a Linear Discriminant Analysis classifier was trained using the cell clusters of the discovery cohort and was used to automatically predict the cluster label for each cell in the validation cohort.[51] To account for technical variation, a peripheral blood mononuclear cell (PBMC) reference sample was included in every mass cytometry experiment. ComBat was applied to align the PBMC reference samples and corresponding patient samples to correct for batch effects.[52]

Single-cell RNA-sequencing

$CD45^+$ cells from 7 tumors (4 MMR-deficient and 3 MMR-proficient) were MACS-sorted with anti-CD45-PE antibodies (clone 2D1, Thermo Fisher Scientific) and anti-PE microbeads (Miltenyi Biotec). Single-cell RNA-sequencing libraries were prepared using the Chromium Single Cell 3' Reagent Kit, Version 2

Chemistry (10x Genomics) according to the manufacturer's protocol. Libraries were sequenced on a NovaSeq6000 using paired-end 2x150bp sequencing (Illumina). Downstream analysis was performed using the Seurat R package according to the author's instructions.[53] Briefly, cells with fewer than 200 expressed genes, and genes that were expressed in less than 3 cells were excluded. Furthermore, cells with outlying percentages of differentially expressed mitochondrial genes (>0.20) and cells with outlying numbers of expressed genes (>5000) were excluded. This resulted in a final dataset of 1,079 cells expressing a total of 1,972 variable genes. Cells were pre-processed using principal component analysis, clustered using graph-based community detection,[54] and visualized by t-distributed Stochastic Neighbor Embedding (t-SNE)[55]. Differentially expressed genes were identified for each cell cluster and visualized in violin plots. In addition, CD45$^+$ cells from one MMR-deficient tumor with high numbers of Lin$^-$CD7$^+$CD127$^-$CD56$^+$CD45RO$^+$ ILCs were sorted on a FACS Aria II sorter (BD Biosciences) (**Table S3**). A similar single-cell RNA-sequencing analysis pipeline was performed while sequencing was performed on a HiSeq4000 (Illumina). Cut-offs for outlying percentages of differentially expressed mitochondrial genes (>0.05) and cells with outlying numbers of expressed genes (>5500) were used. Here, a final dataset of 795 cells expressing a total of 1,814 variable genes was obtained.

Flow cytometry

Single-cell suspensions of CRC tissues (N=8, of which 5 MMR-deficient and 3 MMR-proficient) were stimulated in IMDM/L-glutamine medium (Lonza) complemented with 10% human serum with 20 ng/mL PMA (Sigma-Aldrich) and 1 µg/mL ionomycin (Sigma-Alrdich) for 6 hr at 37°C. Ten µg/mL brefeldin A (Sigma-Aldrich) was added for the last 4 hours. A flow cytometry antibody panel was designed to detect granzyme B/perforin, IFN-γ, and TNF-α production by ILC, T cell, and γδ T cell populations (**Table S3**). In addition, FOXP3 expression by ICOS$^+$ regulatory T cells was assessed in single-cell suspensions of CRC tissues (N=4, of which 1 MMR-deficient and 3 MMR-proficient). Details on flow cytometry antibody staining are available in **Supplemental methods**.

Immunohistochemical staining

Details on immunohistochemical detection of MMR proteins and human leukocyte antigen (HLA) class I expression of CRC tissues are available in **Supplemental methods**.

Multispectral immunofluorescence

A six-marker immunofluorescence panel was applied to 5-µm frozen tissue sections of 4 MMR-deficient and 4 MMR-proficient colorectal tumors, as described previously[56]. Details on immunofluorescence antibody staining are available in **Supplemental methods** and **Table S4**. For each tumor, five different tissue sections were imaged at 20x magnification with the Vectra 3.0 Automated Quantitative Pathology Imaging System (Perkin Elmer). InForm Cell Analysis software (Perkin Elmer) was used for image analysis and spectral separation of dyes, by using spectral libraries defined with single-marker immunofluorescence detection. Tissue segmentation was trained manually with DAPI to segment images into tissue and 'no tissue' areas. All images were visually inspected for the number of CD3$^-$TCRαβ$^-$CD127$^-$CD7$^+$CD45RO$^+$ ILCs and cell counts were normalized by tissue area (number of cells per mm^2).

Statistical analysis

Data were presented as median ± interquartile range. Group comparisons were performed with Mann-Whitney U test, Kruskal-Wallis test with Dunn's test for multiple comparisons, or Friedman test with Dunn's test for multiple comparisons (GraphPad Prism version 7), as indicated. In the correlation analysis, *P*-values were adjusted for multiple testing using the Benjamini-Hochberg procedure. *P*-values < 0.05 were considered statistically significant.

ACKNOWLEDGEMENTS

We thank M.G. Kallenberg-Lantrua, A.M.E.G. Voet-van den Brink, and F.A. Holman for their help in collecting and providing samples from CRC patients; J. van den Bulk for the isolation of PBMCs from

CRC patients; W.E. Corver for flow cytometric cell sorting, R.J. McLaughlin for providing help with mass cytometry experiments, the Leiden Genome Technology Center for their help with single-cell RNA-sequencing, and J. Oosting for help with statistical analysis.

AUTHOR CONTRIBUTIONS

N.L. de Vries conceived the study, performed experiments, and wrote the manuscript. N.L. de Vries, V. van Unen, M.E. Ijsselsteijn, T. Abdelaal, and A. Mahfouz analyzed the data. A. Farina Sarasqueta and K.C.M.J. Peeters provided samples from patients. M.E. Ijsselsteijn, R. van der Breggen, and N.F. de Miranda processed the samples and performed experiments. T. Höllt, V. van Unen, and B.P.F. Lelieveldt developed Cytosplore and HSNE applications. F. Koning and N.F. de Miranda initiated and led the project, and wrote the manuscript. All authors discussed the results and commented on the manuscript.

COMPETING INTERESTS

None to declare.

REFERENCES

1 Hodi, F. S. *et al.* Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 363, 711-723, doi:10.1056/NEJMoa1003466 (2010).

2 Topalian, S. L. *et al.* Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med* 366, 2443-2454, doi:10.1056/NEJMoa1200690 (2012).

3 Rizvi, N. A. *et al.* Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124-128, doi:10.1126/science.aaa1348 (2015).

4 Kelderman, S., Schumacher, T. N. & Kvistborg, P. Mismatch Repair-Deficient Cancers Are Targets for Anti-PD-1 Therapy. *Cancer Cell* 28, 11-13, doi:10.1016/j.ccell.2015.06.012 (2015).

5 Le, D. T. *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409-413, doi:10.1126/science.aan6733 (2017).

6 Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D. & Perucho, M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363, 558-561, doi:10.1038/363558a0 (1993).

7 Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487, 330-337, doi:10.1038/nature11252 (2012).

8 Smyrk, T. C., Watson, P., Kaul, K. & Lynch, H. T. Tumor-infiltrating lymphocytes are a marker for microsatellite instability in colorectal carcinoma. *Cancer* 91, 2417-2422 (2001).

9 de Miranda, N. F. *et al.* Infiltration of Lynch colorectal cancers by activated immune cells associates with early staging of the primary tumor and absence of lymph node metastases. *Clin Cancer Res* 18, 1237-1245, doi:10.1158/1078-0432.CCR-11-1997 (2012).

10 Chevrier, S. *et al.* An Immune Atlas of Clear Cell Renal Cell Carcinoma. *Cell* 169, 736-749 e718, doi:10.1016/j.cell.2017.04.016 (2017).

11 Zheng, C. *et al.* Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* 169, 1342-1356 e1316, doi:10.1016/j.cell.2017.05.035 (2017).

12 Duhen, T. *et al.* Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat Commun* 9, 2724, doi:10.1038/s41467-018-05072-0 (2018).

13 Simoni, Y. *et al.* Bystander CD8(+) T cells are abundant and phenotypically distinct in human tumor infiltrates. *Nature* 557, 575-579, doi:10.1038/s41586-018-0130-2 (2018).

14 Li, N. *et al.* Mass cytometry reveals innate lymphoid cell differentiation pathways in the human fetal intestine. *J Exp Med*, doi:10.1084/jem.20171934 (2018).

15 Menon, A. G. *et al.* Immune system and prognosis in colorectal cancer: a detailed immunohistochemical analysis. *Lab Invest* 84, 493-501, doi:10.1038/labinvest.3700055 (2004).

16 Galon, J. *et al.* Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 313, 1960-1964, doi:10.1126/science.1129139 (2006).

17 Angelova, M. *et al.* Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. *Genome Biol* 16, 64, doi:10.1186/s13059-015-0620-6 (2015).

18 Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat Med* 21, 1350-1356, doi:10.1038/nm.3967 (2015).

19 Li, H. *et al.* Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet* 49, 708-718, doi:10.1038/ng.3818 (2017).

20 Peterson, V. M. *et al.* Multiplexed quantification of proteins and transcripts in single cells. *Nat Biotechnol*, doi:10.1038/nbt.3973 (2017).

21 Schulz, D. *et al.* Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry. *Cell Syst* 6, 25-36 e25, doi:10.1016/j.cels.2017.12.001 (2018).

22 Bandura, D. R. *et al.* Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal Chem* 81, 6813-6822, doi:10.1021/ac901049w (2009).

23 Bendall, S. C., Nolan, G. P., Roederer, M. & Chattopadhyay, P. K. A deep profiler's guide to cytometry. *Trends in immunology* 33, 323-332, doi:10.1016/j.it.2012.02.010 (2012).

24 Dadi, S. *et al.* Cancer Immunosurveillance by Tissue-Resident Innate Lymphoid Cells and Innate-like T Cells. *Cell* 164, 365-377, doi:10.1016/j.cell.2016.01.002 (2016).

25 Fuchs, A. *et al.* Intraepithelial type 1 innate lymphoid cells are a unique subset of IL-12- and IL-15-responsive IFN-gamma-producing cells. *Immunity* 38, 769-781, doi:10.1016/j.immuni.2013.02.010 (2013).

26 Simoni, Y. *et al.* Human Innate Lymphoid Cell Subsets Possess Tissue-Type Based Heterogeneity in Phenotype and Frequency. *Immunity* 46, 148-161, doi:10.1016/j.immuni.2016.11.005 (2017).

27 Wagtmann, N. *et al.* Molecular clones of the p58 NK cell receptor reveal immunoglobulin-related molecules with diversity in both the extra- and intracellular domains. *Immunity* 2, 439-449 (1995).

28 Rajagopalan, S. & Long, E. O. KIR2DL4 (CD158d): An activation receptor for HLA-G. *Front Immunol* 3, 258, doi:10.3389/fimmu.2012.00258 (2012).

29 Moretta, A. *et al.* P58 molecules as putative receptors for major histocompatibility complex (MHC) class I molecules in human natural killer (NK) cells. Anti-p58 antibodies reconstitute lysis of MHC class I-protected cells in NK clones displaying different specificities. *J Exp Med* 178, 597-604 (1993).

30 Dierssen, J. W. *et al.* HNPCC versus sporadic microsatellite-unstable colon cancers follow different routes toward loss of HLA class I expression. *BMC Cancer* 7, 33, doi:10.1186/1471-2407-7-33 (2007).

31 Kloor, M. *et al.* Immunoselective pressure and human leukocyte antigen class I antigen machinery defects in microsatellite unstable colorectal cancers. *Cancer Res* 65, 6418-6424, doi:10.1158/0008-5472.can-05-0044 (2005).

32 Ijsselsteijn, M. E. *et al.* Revisiting immune escape in colorectal cancer in the era of immunotherapy. *Br J Cancer* 120, 815-818, doi:10.1038/s41416-019-0421-x (2019).

33 Iwasaki, M. *et al.* Expression and function of PD-1 in human gammadelta T cells that recognize phosphoantigens. *Eur J Immunol* 41, 345-355, doi:10.1002/eji.201040959 (2011).

34 Webb, J. R., Milne, K., Watson, P., Deleeuw, R. J. & Nelson, B. H. Tumor-infiltrating lymphocytes expressing the tissue resident memory marker CD103 are associated with increased survival in high-grade serous ovarian cancer. *Clin Cancer Res* 20, 434-444, doi:10.1158/1078-0432.CCR-13-1877 (2014).

35 Webb, J. R., Milne, K. & Nelson, B. H. PD-1 and CD103 Are Widely Coexpressed on Prognostically Favorable Intraepithelial CD8 T Cells in Human Ovarian Cancer. *Cancer Immunol Res* 3, 926-935, doi:10.1158/2326-6066.CIR-14-0239 (2015).

36 Djenidi, F. *et al.* CD8+CD103+ tumor-infiltrating lymphocytes are tumor-specific tissue-resident memory T cells and a prognostic factor for survival in lung cancer patients. *J Immunol* 194, 3475-3486, doi:10.4049/jimmunol.1402711 (2015).

37 Edwards, J. *et al.* CD103+ tumor-resident CD8+ T cells are associated with improved survival in immunotherapy naive melanoma patients and expand significantly during anti-PD1 treatment. *Clin Cancer Res*, doi:10.1158/1078-0432.CCR-17-2257 (2018).

38 Santegoets, S. J. *et al.* The Anatomical Location Shapes the Immune Infiltrate in Tumors of Same Etiology and Affects Survival. *Clin Cancer Res* 25, 240-252, doi:10.1158/1078-0432.ccr-18-1749 (2019).

39 Hutloff, A. *et al.* ICOS is an inducible T-cell co-stimulator structurally and functionally related to CD28. *Nature* 397, 263-266, doi:10.1038/16717 (1999).

40 Burris, H. A. *et al.* Phase 1 safety of ICOS agonist antibody JTX-2011 alone and with nivolumab (nivo) in advanced solid tumors; predicted vs observed pharmacokinetics (PK) in ICONIC. *Journal of Clinical Oncology* 35, 3033-3033, doi:10.1200/JCO.2017.35.15_suppl.3033 (2017).

41 Fergusson, J. R. *et al.* CD161 defines a transcriptional and functional phenotype across distinct human T cell lineages. *Cell Rep* 9, 1075-1088, doi:10.1016/j.celrep.2014.09.045 (2014).

42 Fergusson, J. R. *et al.* CD161(int)CD8+ T cells: a novel population of highly functional, memory CD8+ T cells enriched within the gut. *Mucosal Immunol* 9, 401-413, doi:10.1038/mi.2015.69 (2016).

43 Thommen, D. S. *et al.* A transcriptionally and functionally distinct PD-1(+) CD8(+) T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. *Nat Med*, doi:10.1038/s41591-018-0057-z (2018).

44 Kansy, B. A. *et al.* PD-1 Status in CD8(+) T Cells Associates with Survival and Anti-PD-1 Therapeutic Outcomes in Head and Neck Cancer. *Cancer Res* 77, 6353-6364, doi:10.1158/0008-5472.can-16-3167 (2017).

45 Kim, H. D. *et al.* Association Between Expression Level of PD1 by Tumor-Infiltrating CD8(+) T Cells and Features of Hepatocellular Carcinoma. *Gastroenterology* 155, 1936-1950.e1917, doi:10.1053/j.gastro.2018.08.030 (2018).

46 Middha, S. *et al.* Majority of B2M-Mutant and -Deficient Colorectal Carcinomas Achieve Clinical Benefit From Immune Checkpoint Inhibitor Therapy and Are Microsatellite Instability-High. *JCO precision oncology* 3, doi:10.1200/po.18.00321 (2019).

47 Kotecha, N., Krutzik, P. O. & Irish, J. M. Web-based analysis and publication of flow cytometry experiments. *Curr Protoc Cytom* Chapter 10, Unit10 17, doi:10.1002/0471142956.cy1017s53 (2010).

48 Hollt, T. *et al.* Cytosplore: Interactive Immune Cell Phenotyping for Large Single-Cell Datasets. *EuroVis* 35 (2016).

49 Pezzotti, N., Höllt, T., Lelieveldt, B., Eisemann, E. & Vilanova, A. Hierarchical Stochastic Neighbor Embedding. *Computer Graphics Forum* 35, 21-30, doi:10.1111/cgf.12878 (2016).

50 van Unen, V. *et al.* Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types. *Nat Commun* 8,

1740, doi:10.1038/s41467-017-01689-9 (2017).

51 Abdelaal, T. *et al.* Predicting cell populations in single cell mass cytometry data. *Cytometry A*, doi:10.1002/cyto.a.23738 (2019).

52 Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8, 118-127, doi:10.1093/biostatistics/kxj037 (2006).

53 Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*, doi:10.1038/nbt.4096 (2018).

54 Waltman, L. & van Eck, N. J. A smart local moving algorithm for large-scale modularity-based community detection. *The European Physical Journal B* 86, 471, doi:10.1140/epjb/e2013-40829-0 (2013).

55 van der Maaten, L. J. P. & Hinton, G. E. Visualizing high-dimensional data using t-SNE. *J. Mach. Learn. Res.* 9, 2579-2605 (2008).

56 Ijsselsteijn, M. E. *et al.* Cancer immunophenotyping by seven-color multispectral imaging without tyramide signal amplification. *The journal of pathology. Clinical research* 5, 3-11, doi:10.1002/cjp2.113 (2019).

SUPPLEMENTAL METHODS

Tissue digestion and cell isolation

Tumor tissues, tumor-associated lymph nodes, and colorectal healthy mucosa were collected in IMDM+Glutamax medium (Gibco) complemented with 20% fetal calf serum (FCS) (Greiner Bio-One) after surgical resection. Samples were cut into small fragments in a petri dish, and enzymatically digested with 1 mg/mL collagenase D (Roche Diagnostics) and 50 µg/mL DNase I (Roche Diagnostics) in 5 mL of IMDM+Glutamax medium for 30 min at 37°C in gentleMACS C tubes (Miltenyi Biotec). During and after incubation, cell suspensions were mechanically dissociated on the gentleMACS Dissociator (Miltenyi Biotec). Cell suspensions were filtered through a 70-µm cell strainer (Corning) and washed in IMDM+Glutamax medium. Cell number and viability was determined with the Muse Count & Viability Kit (Merck) on the Muse Cell Analyzer (Merck). Cells were cryopreserved based on the number of viable cells in liquid nitrogen until time of analysis in 50% FCS and 10% dimethyl sulfoxide (DMSO) (Merck). Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll-Paque (provided by apothecary LUMC) density-gradient centrifugation and cryopreserved in liquid nitrogen until time of analysis in 20% FCS and 10% DMSO.

Mass cytometry antibodies

Heavy metal isotope-tagged monoclonal antibodies are listed in **Table S2**. Purified antibodies were conjugated with heavy metal reporters in-house using the MaxPar X8 Antibody Labeling Kit (Fluidigm) according to the manufacturer's instructions. All antibodies were titrated to determine the optimal labelling concentration.

Mass cytometry antibody staining and data acquisition

Percoll (GE Healthcare) density-gradient centrifugation was performed to isolate immune cells from tumor tissues and colorectal healthy mucosa. Cells were washed in Maxpar Cell Staining Buffer (CSB, Fluidigm), and incubated with 1 mL CSB containing 1 µM Cell-ID intercalator-^{103}Rh (Fluidigm) for 15 min at room temperature (rT). Cells were washed, incubated with human Fc receptor block (BioLegend) for 10 min at rT, and stained with cell surface antibodies for 45 min at rT in a final volume of 100 uL. After washing, cells were incubated with 1 mL Maxpar Fix and Perm buffer (Fluidigm Sciences) containing 0.125 µM Cell-ID intercalator-Ir (Fluidigm) overnight at 4°C. Cells were acquired on a Helios-upgraded CyTOF2 and Helios mass cytometer (Fluidigm) at an event rate of <500 events/sec in de-ionized water containing 10x diluted EQ Four Element Calibration Beads (Fluidigm). Data were normalized with the normalization passport EQ-P13H2302_ver2 for each experiment.

Mass cytometry data analysis

Correlation analysis of the presence of immune cell clusters across samples (**Figure 7**) was performed using Spearman's rank correlation in the Corrplot R package. PD-1 subsets in CRCs (**Figure S9**) were gated based on a healthy control PBMC sample into PD-1 negative cells (no PD-1 expression), PD-1 intermediate cells (PD-1 expression levels to a similar extent as healthy control PBMCs), and PD-1 high cells (PD-1 expression levels surpassing that of healthy control PBMCs) as described in a previous study on functional diversity between PD-1 subsets.[43] Tumors with less than 100 cells in one of the PD-1 subsets were excluded from this analysis. Sample and immune cell cluster t-distributed Stochastic Neighbor Embedding (t-SNE) maps (**Figure S8**) were computed as described previously.[57]

Flow cytometry antibody staining

Cells were first incubated with human Fc receptor block and cell surface antibodies for 45 min at 4°C. After washing, cells were stained for intracellular proteases and cytokines using Fixation Buffer and Intracellular Staining Perm Wash Buffer (BioLegend) or for FOXP3 expression using FOXP3 Transcription Factor Staining Buffer Set (eBioscience). Compensation was carried out with CompBeads

(BD Biosciences) and ArC reactive beads (Life Technologies). Cells were acquired on a LSR II and LSR Fortessa flow cytometer (BD Biosciences) running FACSDiva software version 8.0 (BD Biosciences). Data were analyzed with FlowJo software version 10.2 (Tree Star Inc). Antibody details are listed in **Table S3**.

Immunohistochemical staining

FFPE blocks from colorectal cancers (CRCs) were obtained from the department of Pathology at the Leiden University Medical Center (Leiden, The Netherlands). Tumor mismatch repair (MMR) status (MMR-proficient or MMR-deficient) was determined by immunohistochemical detection of PMS2 (anti-PMS2 antibodies; clone EP51, DAKO) and MSH6 (anti-MSH6 antibodies; clone EPR3945, Abcam) proteins.[58] MMR-deficiency was determined by lack of expression of at least one of the MMR-proteins in the presence of an internal positive control. Immunohistochemical detection of human leukocyte antigen (HLA) class I expression was performed with HCA2 and HC10 monoclonal antibodies (Nordic-MUbio), and classified as HLA class I positive, weak, or loss as described previously[32].

Multispectral immunofluorescence detection

Five-µm frozen tissue sections were cut on adhesive immunohistochemistry slides. Following fixation in ice-cold 100% methanol for 5 min and wash in PBS, the tissues were incubated with Superblock buffer (Thermo Fisher Scientific) for 30 min at rT. For the detection of CD127, TSA signal amplification was performed with Opal520 from the Opal 7-color manual IHC kit (Perkin Elmer), according to manufacturer's instructions. Thereafter, the tissues were incubated with primary TCR V beta F1 and CD45RO antibodies to be detected indirectly overnight at rT, followed by incubation with corresponding fluorescent secondary antibodies CF680 and CF633, respectively, for 1h at rT. The tissues were then incubated with directly conjugated primary CD3-AF594 and CD7-AF647 antibodies for 5h at rT. CD3 was conjugated to AF594 using the Alexa Fluor 594 antibody labelling kit (Thermo Fisher Scientific). Lastly, the tissues were incubated with 1 µM DAPI for nuclear counterstaining, and mounted with Prolong® Gold Antifade Reagent (Cell Signaling Technology). Antibody details are listed in **Table S4**.

REFERENCES

57 van Unen, V. *et al.* Mass Cytometry of the Human Mucosal Immune System Identifies Tissue- and Disease-Associated Immune Subsets. *Immunity* 44, 1227-1239, doi:10.1016/j.immuni.2016.04.014 (2016).

58 Hall, G. *et al.* Immunohistochemistry for PMS2 and MSH6 alone can replace a four antibody panel for mismatch repair deficiency screening in colorectal adenocarcinoma. *Pathology* 42, 409-413, doi:10.3109/00313025.2010.493871 (2010).

SUPPLEMENTAL FIGURES

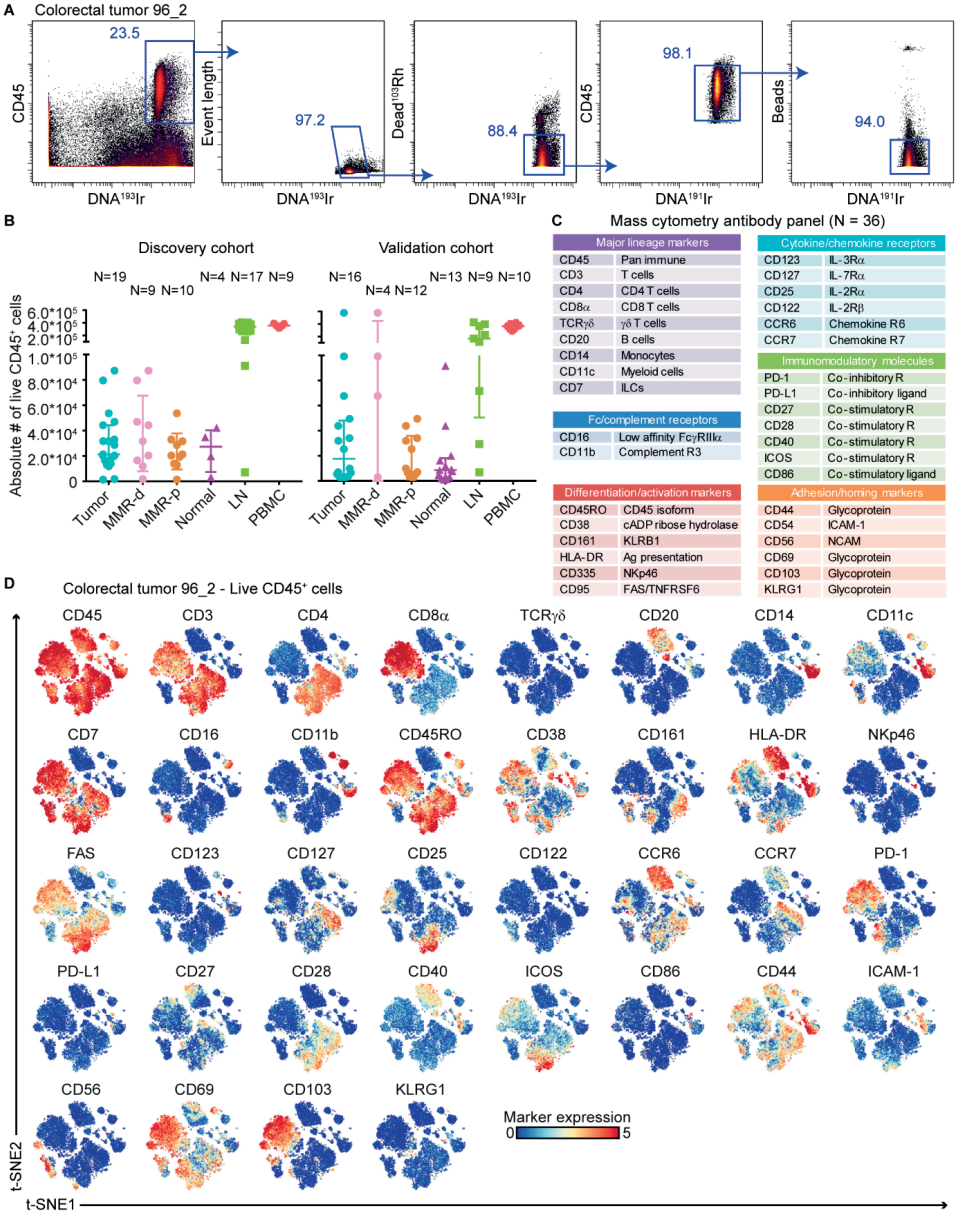


Figure S1. Mass cytometry gating strategy and antibody expression patterns.
A. Mass cytometry gating strategy for single, live CD45⁺ cells of a representative colorectal tumor sample showing sequential gates with percentages. **B.** Absolute number of live CD45⁺ cells of CRC tissues, colorectal healthy mucosa, tumor-associated lymph nodes, and peripheral blood samples of the discovery and validation cohort of CRC patients. Bars indicate median ± IQR. Each dot represents an individual sample. Data from 22 independent experiments with mass cytometry. **C.** Markers used to characterize immune cell phenotypes by mass cytometry. **D.** t-SNE embedding showing marker expression patterns of each antibody on single, live CD45⁺ cells ($2.0*10^4$) from the same tumor sample as shown in (**A**). Each dot represents a single cell. All markers are shown with an expression range of 0-5, with the exception of CD86 (0-3) due to lower sensitivity of the metal (^{115}In).

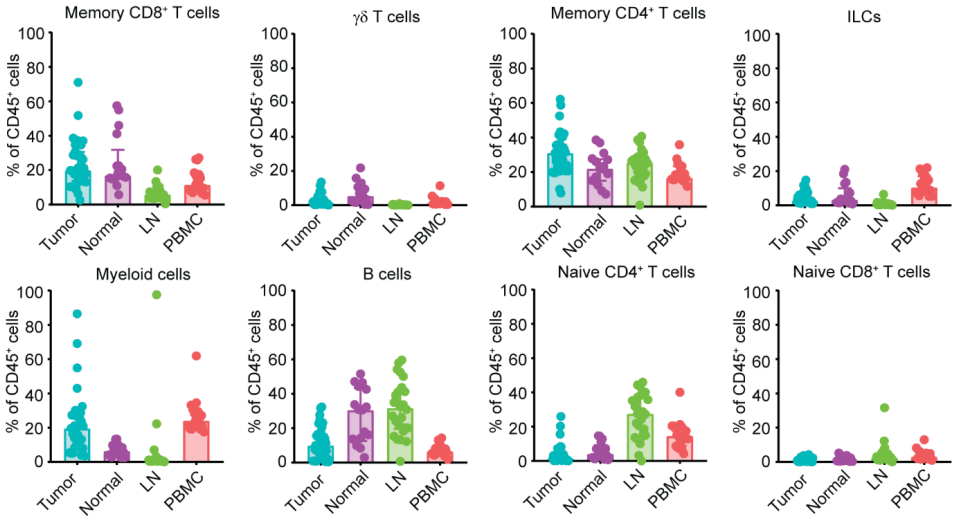


Figure S2. Major immune lineage frequencies in different tissue types of colorectal cancer patients.
Frequencies of major immune lineages across CRC tissues (N=35), colorectal healthy mucosa (N=17), tumor-associated lymph nodes (N=26), and peripheral blood (N=19) as percentage of total CD45⁺ cells. Bars indicate median ± IQR. Each dot represents an individual sample. Data from 22 independent experiments with mass cytometry.

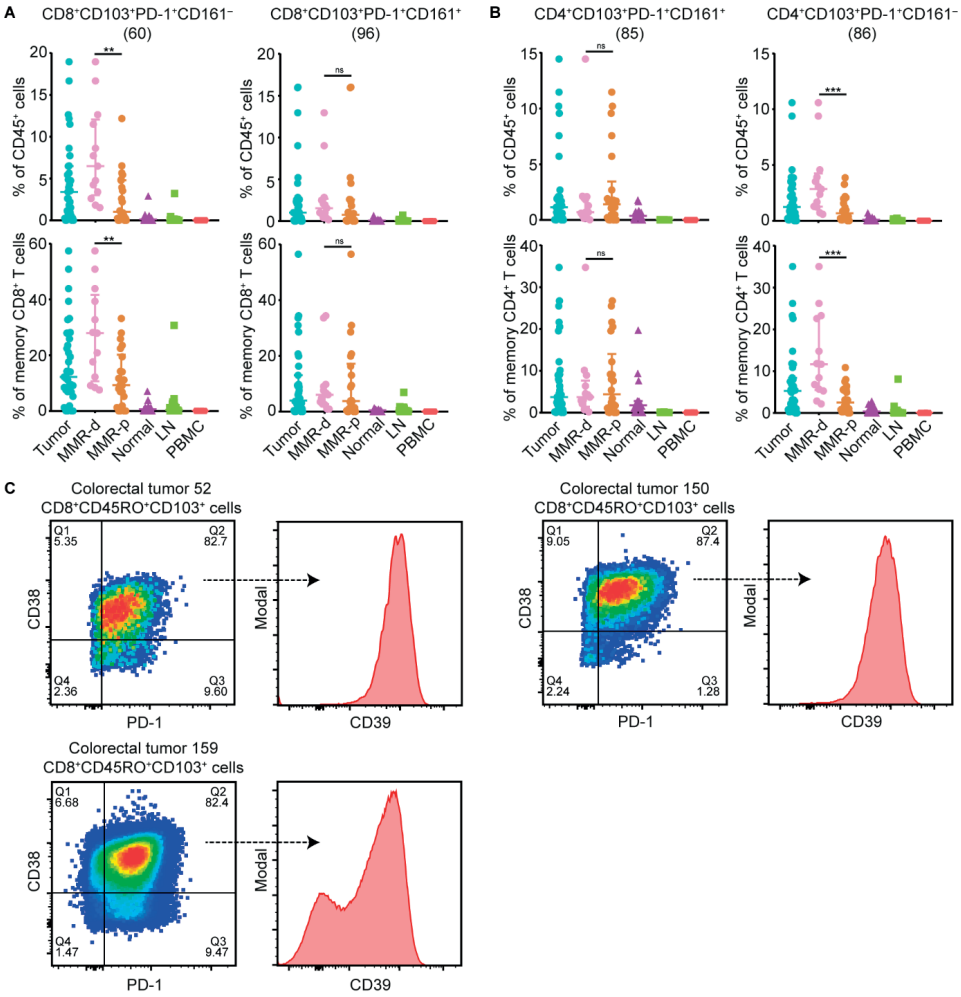


Figure S3. Characterization of tumor tissue-specific immune cell clusters corresponding to Figure 2 and 3.
A and B. Frequencies of CD103+PD-1+CD161− and CD103+PD-1+CD161+ memory CD8+ T cells (**A**) and CD4+ T cells (**B**) among CRC tissues (N=35, further subdivided into MMR-deficient (N=13) and MMR-proficient (N=22)), colorectal healthy mucosa (N=17), tumor-associated lymph nodes (N=26), and peripheral blood (N=19) as percentage of total CD45+ cells (upper panel) and memory CD8+ or CD4+ T cells (lower panel). Cluster IDs correspond to the ones in **Figure 2B** and **Figure 3B**. Bars indicate median ± IQR. Each dot represents an individual sample. Data from 22 independent experiments with mass cytometry. NS, not significant, ***P*<0.01, ****P*<0.001 by Mann-Whitney U test. **C.** Flow cytometry plots of colorectal tumors (N=3) showing the expression of CD39 within CD8+CD45RO+CD103+PD-1+CD38+ cells.

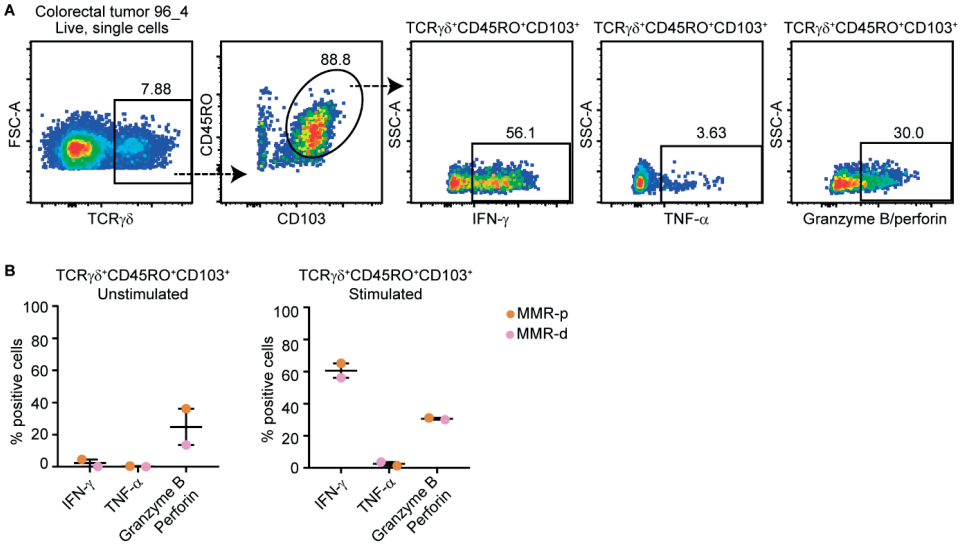


Figure S4. Tumor-resident γδ T cells are capable of producing cytokines and cytotoxic molecules upon stimulation.

A. Flow cytometry plots of a MMR-deficient tumor sample showing the expression of cytokines and cytotoxic molecules by TCRγδ+CD45RO+CD103+ cells upon stimulation with PMA/ionomycin. **B.** IFN-γ, TNF-α, and granzyme B/perforin expression by TCRγδ+CD45RO+CD103+ cells from a MMR-deficient and MMR-proficient CRC with and without stimulation with PMA/ionomycin. Bars indicate median ± IQR. Each dot represents an individual sample.

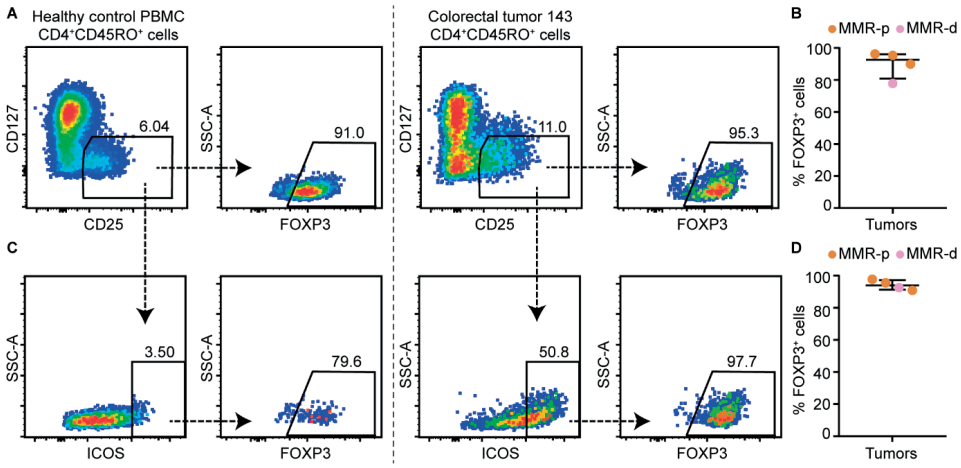


Figure S5. Expression of FOXP3 by ICOS+ regulatory T cells in colorectal tumors.

A. Representative plots of a healthy control PBMC sample and a MMR-proficient tumor sample analyzed by flow cytometry showing the expression of FOXP3 by regulatory T cells ($CD25^+CD127^{low}$). **B.** FOXP3 expression in regulatory T cells ($CD25^+CD127^{low}$) from CRC tissues (N=4, of which 1 MMR-deficient and 3 MMR-proficient). Bars indicate median ± IQR. Each dot represents an individual sample. Data from two independent experiments with flow cytometry. **C.** Representative plots of a healthy control PBMC sample and a MMR-proficient tumor sample analyzed by flow cytometry showing the expression of FOXP3 by $ICOS^+$ regulatory T cells ($CD25^+CD127^{low}$). **D.** FOXP3 expression in $ICOS^+$ regulatory T cells ($CD25^+CD127^{low}$) from CRC tissues (N=4, of which 1 MMR-deficient and 3 MMR-proficient). Bars indicate median ± IQR. Each dot represents an individual sample. Data from two independent experiments with flow cytometry.

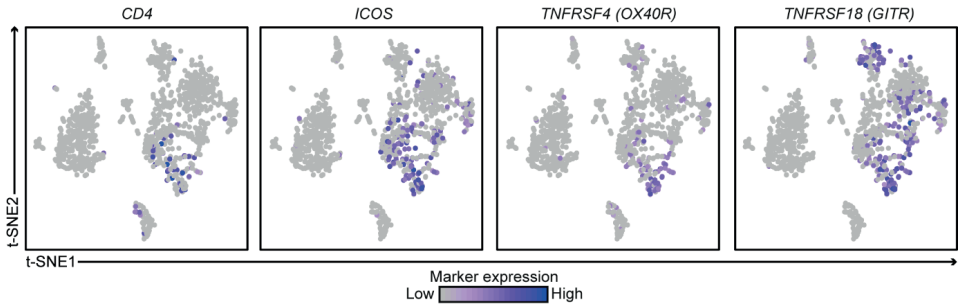


Figure S6. Co-expression of *ICOS*, *TNFRSF4 (OX40R)*, and *TNFRSF18 (GITR)* on CD4⁺ T cells in colorectal cancers.
t-SNE embedding showing 1,079 cells from CRC tissues (N=7) analyzed by single-cell RNA-sequencing. Colors represent the log-transformed expression levels of indicated markers. Each dot represents a single cell.

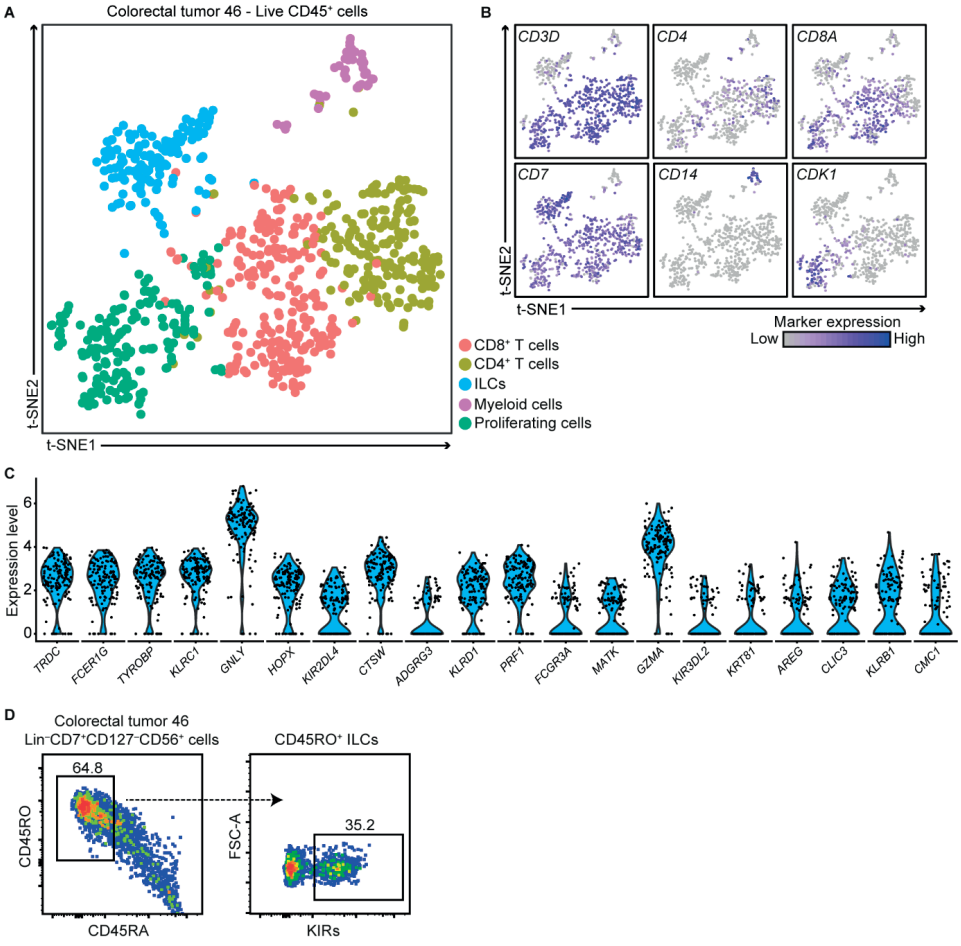


Figure S7. Expression of cytotoxic molecules and KIRs by tumor-resident ILCs.
A and B. t-SNE embedding of single-cell RNA-sequencing data showing 795 cells from one MMR-deficient tumor that was selected for its high numbers of Lin⁻CD7⁺CD127⁻CD56⁺CD45RO⁺ ILCs (70% of the ILC cluster) based on mass cytometry data. Colors represent the different clusters (**A**) and the log-transformed expression levels of indicated

markers (**B**). Each dot represents a single cell. **C.** Violin plot showing log-transformed expression levels of the top 20 differentially expressed genes within ILCs (N=137) as identified in (**A**). Each dot represents a single cell. **D.** Flow cytometry plots showing the cell surface expression of KIRs in Lin$^-$CD7$^+$CD127$^-$CD56$^+$CD45RO$^+$ ILCs from the same tumor as in (**A-C**).

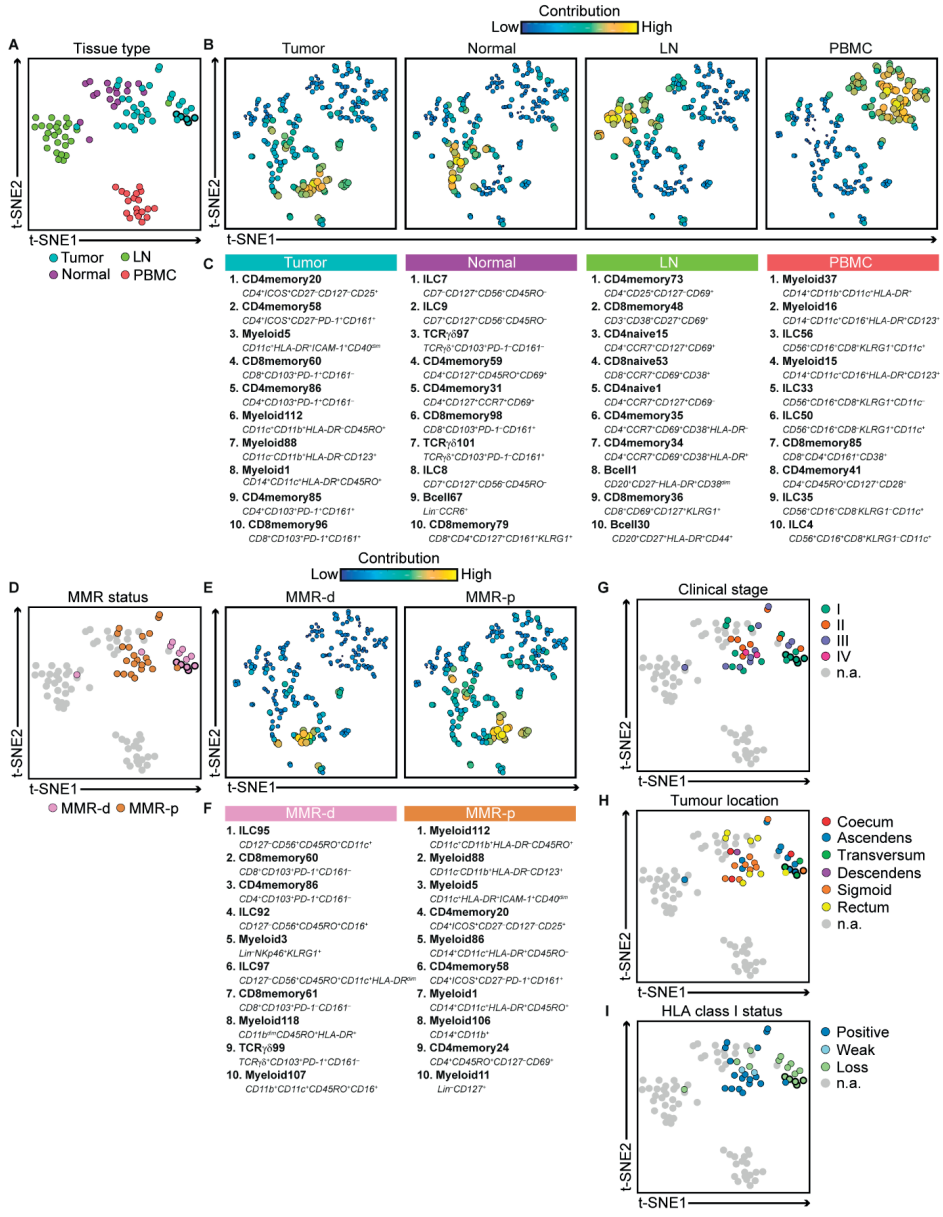


Figure S8. Integrated analysis of the immune composition in different tissue types of colorectal cancer patients.

A. Collective t-SNE analysis showing the clustering of 97 samples based on cell percentage data (of CD45$^+$ cells) of 218 immune cell clusters. Every dot represents a sample colored by tissue type. Five primary tumors at different locations from the same patient are highlighted. One lymph node sample clustered within the tumor samples, and was found to be infiltrated by tumor cells upon histological examination. One tumor sample clustered within the lymph

node samples, and was found to contain large populations of naive CD4+ T cells and B cells, which are enriched in lymph nodes. Histological examination of the tumor confirmed the presence of lymphoid aggregates with germinal centers, a Crohn-like lymphoid reaction that can be a feature of MMR-deficient tumors.[59] **B.** Collective t-SNE analysis showing the clustering of 218 immune cell clusters based on cell percentage data (of CD45+ cells) of 97 samples. Every dot represents an immune cell cluster. Dot color and size indicate the contribution of the immune cell cluster to the respective t-SNE sample signatures as shown in (**A**). **C.** Top ten ranked immune cell clusters contributing to the t-SNE sample signatures as shown in (**A**). Unique cluster IDs and a short description of their phenotype are displayed. **D.** Collective t-SNE analysis of (**A**) colored by MMR status of the tumor samples. **E.** Collective t-SNE analysis of (**B**) showing the contribution of the immune cell clusters to the respective t-SNE sample signatures as shown in (**D**). **F.** Top ten ranked immune cell clusters contributing to the t-SNE sample signatures as shown in (**D**). Unique cluster IDs and a short description of their phenotype are displayed. **G-I.** Collective t-SNE analysis of (**A**) colored by clinical stage (**G**), tumor location (**H**), and HLA class I status (**I**).

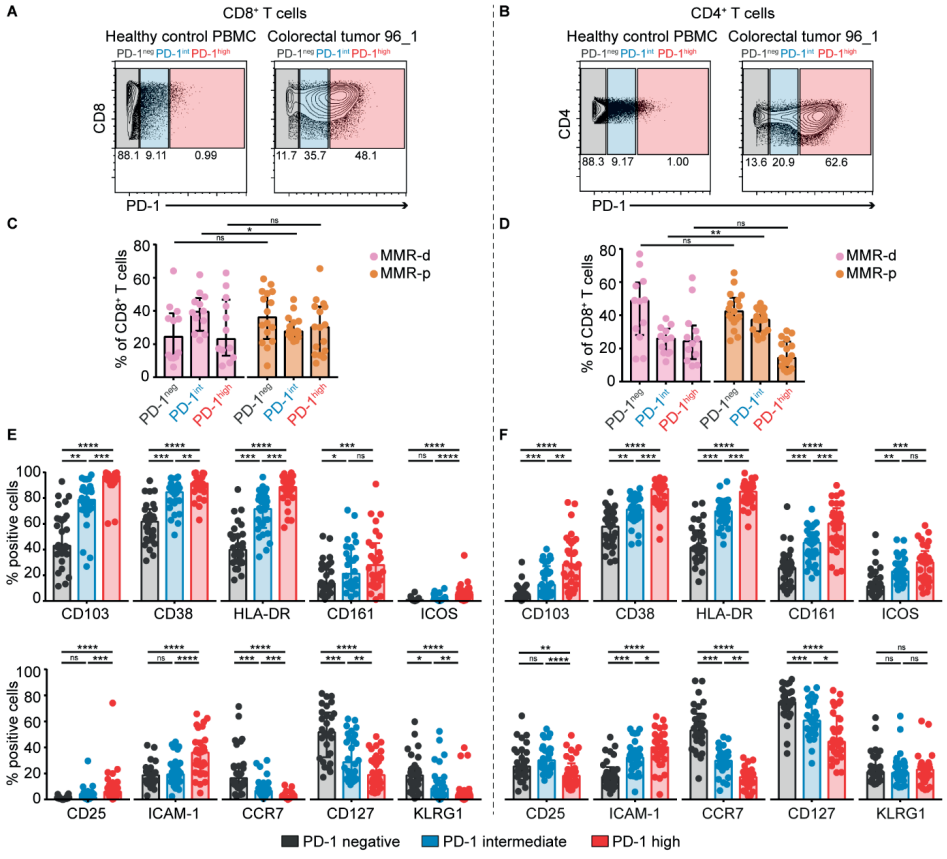


Figure S9. PD-1 expression level of CD8+ and CD4+ T cells correlates with distinct states of activation and differentiation in colorectal tumors.

A and B. Representative plots showing the gating strategy for PD-1 negative, intermediate and high CD8+ T cells (**A**) and CD4+ T cells (**B**) in healthy control PBMC and colorectal tumor tissues (see **Supplemental methods**). **C**. Frequencies of PD-1 negative, intermediate and high CD8+ T cells in MMR-deficient (N=12) and MMR-proficient (N=16) tumors. **D**. Frequencies of PD-1 negative, intermediate and high CD4+ T cells in MMR-deficient (N=12) and MMR-proficient (N=17) tumors. **C**,**D**. Bars indicate median ± IQR. Data from 22 independent experiments with mass cytometry. NS, not significant, **P*<0.05, ***P*<0.01 by Mann-Whitney U-test. **E**. Frequencies of selected immune cell markers expressed by PD-1 negative, intermediate and high CD8+ T cells in CRCs (N=28). **F**. Frequencies of selected immune cell markers expressed by PD-1 negative, intermediate and high CD4+ T cells in CRCs (N=29).

E,**F**. Bars indicate median ± IQR. Each dot represents an individual sample. Data from 22 independent experiments with mass cytometry. NS, not significant, **P*<0.05, ***P*<0.01, ****P*<0.001, *****P*<0.0001 by Friedman test with Dunn's test for multiple comparisons.

REFERENCES

59 Graham, D. M. & Appelman, H. D. Crohn's-like lymphoid reaction and colorectal carcinoma: a potential histologic prognosticator. *Mod Pathol* 3, 332-335 (1990).

SUPPLEMENTAL TABLES

The supplementary tables are available in the online version of this manuscript at http://dx.doi.org/10.1136/gutjnl-2019-318672 and in the appendix to this thesis at the repository of Leiden University (https://hdl.handle.net/1887/3439882).

Table S1: Overview of clinical samples and patient characteristics.
Table S2: Antibodies used for mass cytometry experiments.
Table S3: Antibodies used for flow cytometry and cell sorting experiments.
Table S4: Antibodies used for multispectral immunofluorescence detection.
Table S5: Coefficient and *P*-values of immune cell correlations observed in colorectal cancer tissues (corresponding to Figure 7).

4

γδ T cells are effectors of immune checkpoint blockade in mismatch repair-deficient colon cancers with antigen presentation defects

Natasja L. de Vries[1,2]*, Joris van de Haar[3,4,5]*, Vivien Veninga[3,4]*, Myriam Chalabi[3,6,7]*, Marieke E. Ijsselsteijn[1], Manon van der Ploeg[1], Jitske van den Bulk[1], Dina Ruano[1], Jose G. van den Berg[8], John B. Haanen[3,7], Laurien J. Zeverijn[3,4], Birgit S. Geurts[3,4], Gijs F. de Wit[3,4], Thomas W. Battaglia[3,4], Hans Gelderblom[9], Henk M.W. Verheul[10], Ton N. Schumacher[3,4,11], Lodewyk F.A. Wessels[4,5,12], Frits Koning[2†], Noel F.C.C. de Miranda[1†], Emile E. Voest[3,4†]

[1]Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands. [2]Department of Immunology, Leiden University Medical Center, Leiden, the Netherlands. [3]Department of Molecular Oncology and Immunology, Netherlands Cancer Institute, Amsterdam, The Netherlands. [4]Oncode Institute, Utrecht, the Netherlands. [5]Division of Molecular Carcinogenesis, Netherlands Cancer Institute, Amsterdam, the Netherlands. [6]Gastrointestinal Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands. [7]Medical Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands. [8]Department of Pathology, Netherlands Cancer Institute, Amsterdam, the Netherlands. [9]Department of Medical Oncology, Leiden University Medical Center, Leiden, the Netherlands. [10]Department of Medical Oncology, Radboud UMC, Nijmegen, the Netherlands. [11]Department of Hematology, Leiden University Medical Center, Leiden, the Netherlands. [12]Faculty of EEMCS, Delft University of Technology, Delft, the Netherlands. *Co-first authors. †Equal responsible authors.

Manuscript in revision (Nature)

ABSTRACT

DNA mismatch repair deficient (MMR-d) cancers present an abundance of neoantigens that likely underlies their exceptional responsiveness to immune checkpoint blockade (ICB).[1,2] In contrast to other cancer types,[3-5] we observed that 20 out of 21 (95%) MMR-d cancers with genomic inactivation of β2-microglobulin (*B2M*) retained responsiveness to ICB, suggesting the involvement of immune effector cells other than $CD8^+$ T cells in this context. We next identified a strong association between *B2M* inactivation and an increased infiltration by γδ T cells in MMR-d cancers. These γδ T cells were mainly composed of Vδ1 and Vδ3 subsets, and expressed high levels of PD-1, activation markers including cytotoxic molecules, and a broad repertoire of killer-cell immunoglobulin-like receptors (KIRs). *In vitro*, $PD\text{-}1^+$ γδ T cells, isolated from MMR-d colon cancers, exhibited a cytolytic response towards HLA class I-negative MMR-d colon cancer cell lines and *B2M*-knockout patient-derived tumor organoids (PDTOs), which was enhanced as compared to antigen presentation-proficient cells. By comparing paired tumor samples of MMR-d colon cancer patients obtained before and after dual PD-1 and CTLA-4 blockade, we found that ICB profoundly increased the intratumoral frequency of γδ T cells in β2m-deficient cancers. Taken together, these data indicate that γδ T cells contribute to the response to ICB in patients with HLA class I-negative, MMR-d colon cancers, and underline the potential of γδ T cells in cancer immunotherapy.

INTRODUCTION

Immune-checkpoint blockade (ICB) targeting the PD-1/PD-L1 and/or CTLA-4 axis provides durable clinical benefit to patients with DNA mismatch repair-deficient (MMR-d)/Microsatellite Instability-High (MSI-H) cancers.[6-9] The exceptional responses of MMR-d/MSI-H cancers to ICB are likely explained by their vast burden of putative neoantigens, which originate from the extensive accumulation of mutations in their genomes.[1,2] This is in line with the current view that PD-1 blockade mainly boosts endogenous antitumor immunity driven by $CD8^+$ T cells, which recognize Human Leukocyte Antigen (HLA) class I-bound neoepitopes on cancer cells.[10-12] However, MMR-d colon cancers frequently lose HLA class I-mediated antigen presentation due to silencing of HLA class I genes, inactivating mutations in β2-microglobulin (*B2M*), or other defects in the antigen processing machinery,[13-16] which may render these tumors resistant to $CD8^+$ T cell-mediated immunity.[3-5,17] Interestingly, early evidence has indicated that β2m-deficient, MMR-d cancers can obtain durable responses to PD-1 blockade,[18] suggesting that immune cell subsets other than $CD8^+$ T cells contribute to these responses.

HLA class I-unrestricted immune cell subsets, capable of tumor killing, include natural killer (NK) cells and γδ T cells. γδ T cells share many characteristics with their αβ T cell counterpart, such as cytotoxic effector functions, but express a distinct TCR composed of a γ and a δ chain. Different subsets of γδ T cells are defined by their TCR δ chain usage, of which those expressing Vδ1 and Vδ3 are primarily "tissue-resident" at mucosal sites, whereas those expressing Vδ2 are mainly found in blood.[19] Both adaptive and innate mechanisms of activation, e.g., through stimulation of their γδ TCR or innate receptors such as NKG2D, DNAM-1, NKp30 or NKp44, have been described for γδ T cells.[20] Killer-cell immunoglobulin-like receptors (KIRs) are expressed by γδ T cells and regulate their activity depending on HLA class I expression in target cells.[21] Furthermore, γδ T cells were found to express high levels of PD-1 in MMR-d colorectal cancers (CRCs), suggesting that these cells may be targeted by PD-1 blockade.[22]

Here, we applied a combination of transcriptomic and imaging approaches for an in-depth analysis of ICB-naïve and ICB-treated MMR-d colon cancers, as well as *in vitro* functional assays, and found evidence indicating that γδ T cells mediate responses to HLA class I-negative, MMR-d tumors during ICB treatment.

RESULTS

B2M-mutant MMR-d cancers retain responsiveness to PD-1 blockade

We evaluated responses to PD-1 blockade therapy in a cohort of 71 patients with MMR-d cancers from various anatomical sites treated in the Drug Rediscovery Protocol (DRUP)[23] in relation to their *B2M* status (**Figure 1A**, **Figure S1A-C**, **Table S1**). Clinical benefit (CB; defined as ≥4 months disease control; the primary outcome of the DRUP) was observed in 20 out of 21 (95%) of patients with $B2M^{MUT}$ tumors *vs* 31 out of 50 (62%) of patients

with *B2M*^{WT} tumors (Fisher's exact test-based two-sided P=0.0038; logistic regression-based P=0.022 and P=0.027, adjusted for tumor mutational burden [TMB], and TMB plus tumor type, respectively; **Figure 1B**). Among patients with *B2M*^{MUT} tumors, 12 out of 21 (57%) experienced a partial response (following RECIST1.1 criteria) and 3 (14%) a complete response. All 44 *B2M* alterations across 21 patients were clonal (**Methods**), consistent with earlier observations in MMR-d cancers[18]. Thirteen out of 21 (62%) patients with *B2M*^{MUT} tumors harbored bi-allelic *B2M* alterations, 4 (19%) potentially bi-allelic alterations, and 4 (19%) non bi-allelic alterations (**Figure 1C**, **Methods**). The latter have also been associated with complete loss of β2m protein expression in MMR-d tumors.[18] Thus, *B2M* alterations are associated with a high clinical benefit rate of PD-1 blockade in patients with MMR-d cancers.

γδ T cells are enriched in *B2M*-mutant MMR-d cancers

To gain insight into immune cell subsets involved in immune responses towards HLA class I-negative MMR-d cancers we used a large cohort of The Cancer Genome Atlas (TCGA) and studied the transcriptomic changes associated with genomic loss of *B2M* in three MMR-d cancer cohorts in colon adenocarcinoma (COAD; n=50 *B2M*^{WT}, n=7 *B2M*^{MUT}), stomach adenocarcinoma (STAD; n=48 *B2M*^{WT}, n=12 *B2M*^{MUT}), and endometrium carcinoma (UCEC; n=118 *B2M*^{WT}, n=4 *B2M*^{MUT}). We found that *B2M* was among the most significantly downregulated genes in *B2M*^{MUT} cancers (two-sided P=3.5 x10^{-4}, Benjamini-Hochberg corrected false discovery rate [FDR]=0.12; **Figure 1D**). Genes encoding components of the HLA class I antigen presentation machinery other than *B2M* were highly upregulated in *B2M*^{MUT} tumors, which may reflect reduced evolutionary pressure on somatic inactivation of these genes in the *B2M*^{MUT} context[18] (**Figure 1D**). Interestingly, we found *TRDV1* and *TRDV3*, which encode the variable regions of the δ1 and δ3 chains of the γδ T cell receptor (TCR), among the most significantly upregulated loci in *B2M*^{MUT} tumors (*TRDV1*: FDR=0.00090; *TRDV3*: FDR=0.0015; **Figure 1D**), regardless of allelic status of the *B2M* alteration (**Figure S1D**). In line with this, the expression level of *TRDV1* and *TRDV3* was higher in *B2M*^{MUT} compared to *B2M*^{WT} MMR-d cancers (Wilcoxon rank sum-based two-sided P=6.5x10^{-8} for all cohorts combined; linear regression-based two-sided P=4.7x10^{-6}, adjusted for tumor type; **Figure 1D-F**). In addition, *B2M*^{MUT} tumors showed overexpression of multiple KIRs (**Figure 1D**), which clustered together with *TRDV1* and *TRDV3* by hierarchical clustering (**Figure S1E**). The expression level of different KIRs (**Table S2**) was higher in *B2M*^{MUT} compared to *B2M*^{WT} MMR-d tumors (Wilcoxon rank sum-based two-sided P=4.4x10^{-6} for all cohorts combined; linear regression-based two-sided P=4.7x10^{-5}, adjusted for tumor type; **Figure 1D-F**). Together, these results suggest that ICB-naïve, *B2M*^{MUT} MMR-d cancers show increased levels of Vδ1 and Vδ3 T cells and (these or other) immune cells expressing KIRs, receptors implied in the recognition and killing of HLA class I-negative cells.

We used marker gene sets (modified from Danaher *et al.*[24]; **Methods**, **Table S2**) to estimate the abundance of a broad set of other immune cell types based on the RNA expression data of the TCGA cohorts. Hierarchical clustering identified a highly and a lowly infiltrated cluster in each of the three tumor types (**Figure 1E**). As compared to the Vδ1/Vδ3 T cell and KIR gene sets, the other marker gene sets showed no or only weak association between

expression level and *B2M* status, indicating that our findings were not solely driven by a generally more inflamed state of *B2M*^{MUT} tumors (**Figure 1E-F**, **Figure S1F**).

We next revisited the DRUP cohort and specifically applied the marker gene sets on RNA expression data. Despite the small patient numbers and high heterogeneity regarding tumor types and biopsy locations of this cohort, we could confirm an increased *TRDV1* and *TRDV3* expression in *B2M*^{MUT} tumors pan-cancer (linear regression-based two-sided P=0.017, adjusted for tumor type and biopsy site; **Figure 1G**, **Figure S1G**, **Methods**). KIR expression was only significantly associated with *B2M* status in CRC (**Figure 1G**). Taken together, *B2M* defects are positively associated with clinical benefit to ICB treatment, as well as infiltration by Vδ1/Vδ3 T cells and expression of KIRs.

Cytotoxic Vδ1 and Vδ3 T cells infiltrate MMR-d colon cancers

To investigate which γδ T cell subsets are present in MMR-d colon cancers and to determine their functional characteristics, we performed single-cell RNA-sequencing (scRNA-seq) on γδ T cells isolated from five MMR-d colon cancers (**Figure S2-3**, **Table S3**). Three distinct Vδ subsets were identified (**Figure 2A**), where Vδ1 T cells were the most prevalent (43% of γδ T cells), followed by Vδ2 (19%) and Vδ3 T cells (11%) (**Figure 2B**). *PDCD1* (encoding PD-1) was predominantly expressed by Vδ1 and Vδ3 T cells, while Vδ1 cells expressed high levels of genes encoding activation markers such as CD39 (*ENTPD1*) and CD38 (**Figure 2C**, **Figure S2B**). Furthermore, proliferating γδ T cells (expressing *MKI67*) were especially observed in the Vδ1 and Vδ3 subsets (**Figure 2C**). Other distinguishing features of Vδ1 and Vδ3 T cell subsets included the expression of genes encoding activating receptors NKp46 (*NCR1*), NKG2C (*KLRC2*), and NKG2D (*KLRK1*) (**Figure 2C**). Interestingly, the expression of several KIRs was also higher in the Vδ1 and Vδ3 subsets as compared to Vδ2 T cells (**Figure 2C**). Almost all γδ T cells displayed expression of genes encoding Granzyme B (*GZMB*), Perforin (*PRF1*), and Granulysin (*GNLY*) (**Figure 2C**). Altogether, these data support a role for γδ T cells in mediating natural cytotoxic antitumor responses in HLA class I-negative MMR-d colon cancers.

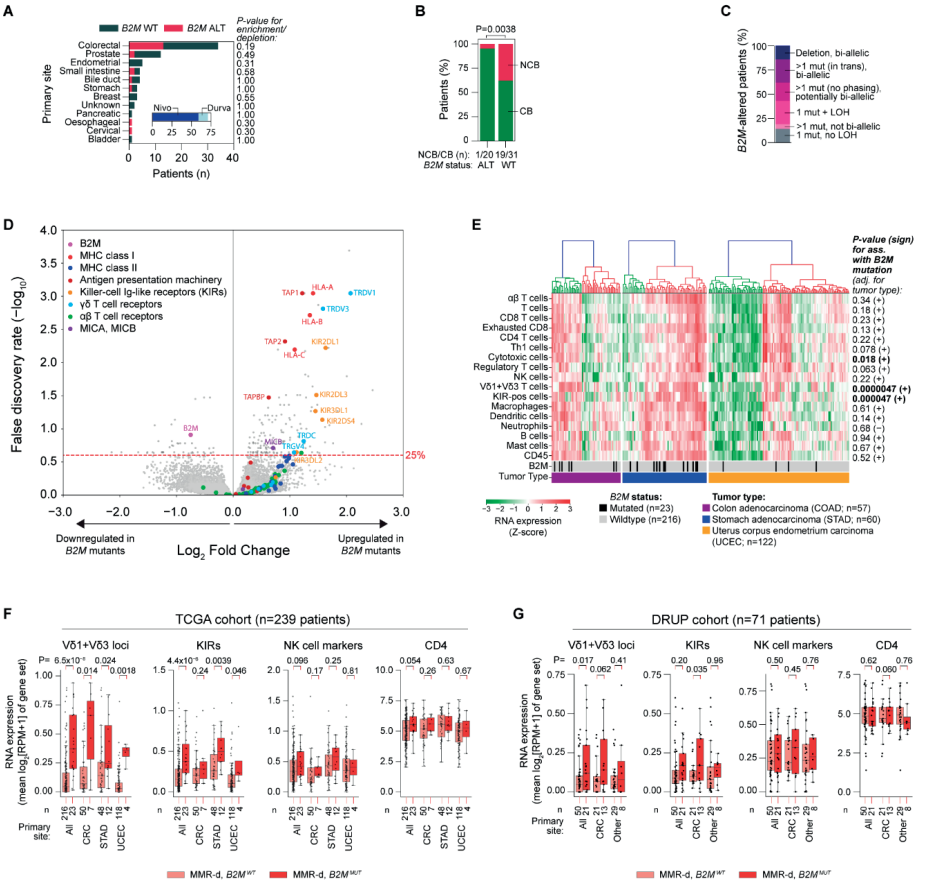


Figure 1. In MMR-d cancers, *B2M* defects are positively associated with clinical benefit to ICB treatment, as well as infiltration by Vδ1/3 T cells and killer-cell immunoglobulin-like receptor (KIR)-expressing cells.
A. The number of patients per primary tumor type for 71 patients in the DRUP with MMR-d cancers and available outcome of ICB therapy. Colors denote patients' *B2M* status (WT: wildtype, gray; ALT: altered, red). Fisher's exact test-based P-values for enrichment/depletion of *B2M* altered cases per primary site are shown. The inset denotes the type of ICB treatment (Nivo: nivolumab, dark blue; Durva: durvalumab, light blue). **B**. The clinical benefit of ICB treatment in the DRUP cohort *vs B2M* status, for the full DRUP cohort. Fisher's exact test-based P-value is shown. **C**. The allelic alteration status of *B2M* in the DRUP cohort. **D**. Volcano plot indicating differential gene expression between *B2M*-mutated *vs* -wildtype MMR-d cancers. The Benjamini Hochberg false discovery rate (FDR) significance threshold of 25% is indicated by the red dashed line. Results were obtained in a combined analysis of all MMR-d cancers of the TCGA COAD (colon adenocarcinoma; n=57 patients), STAD (stomach adenocarcinoma; n=60 patients) and UCEC (uterus corpus endometrial carcinoma; n=122 patients) cohorts, and were adjusted for tumor type. **E**. Heatmap of the expression (Z-score; see color bar) of gene sets whose expression marks infiltration of specific immune cell types in MMR-d cancers of the COAD, STAD and UCEC cohorts of TCGA. Cancers were ranked based on hierarchical clustering, as indicated by the dendrograms (top). The lower two bars indicate the *B2M* mutation status and cancer type. P-values and sign (+ for positive and – for negative) of associations of marker gene expression with *B2M* mutation status are shown on the right. P-values were obtained by ordinary least squares linear regression and adjusted for tumor type. Significant associations (nominal P<0.05) are in bold font. **F**. RNA expression of different immune marker gene sets MMR-d $B2M^{WT}$ (pink), and MMR-d $B2M^{MUT}$ (red) cancers. Results are obtained with the TCGA COAD, STAD and UCEC cohorts, and are shown for all cohorts combined (All), and for each cohort separately. Boxes, whiskers, and dots indicate quartiles, 1.5 interquartile ranges, and individual data points, respectively. Wilcoxon rank sum test-based P-values are shown for MMR-d $B2M^{WT}$ *vs* MMR-d $B2M^{MUT}$ cancers. **G**. As **F**, but for MMR-d cancers in the DRUP cohort. Results are shown for all cancers combined (All), only colorectal cancer (CRC), or all non-CRC cancers (Other).

Next, we applied imaging mass cytometry (IMC) on a cohort of 17 ICB-naïve MMR-d colon cancers (**Table S3**). High levels of γδ T cell infiltration were observed in cancers with β2m defects as compared to β2m-proficient cancers, albeit this difference was not significant (**Figure 2D**). Levels of other immune cells, including NK cells, CD4$^+$ T cells, and CD8$^+$ T cells, were similar between β2m-deficient and -proficient tumors (**Figure 2D**). In β2m-deficient cancers, γδ T cells showed frequent intraepithelial localization and expression of CD103 (tissue-residency), CD39 (activation), granzyme B (cytotoxicity), and Ki-67 (proliferation), as well as PD-1 (**Figure 2D-F**, **Figure S4A**), in line with the scRNA-seq data. Of note, γδ T cells in β2m-deficient cancers showed co-expression of CD103 and CD39 (**Figure S4B**), reported to identify tumor-reactive CD8$^+$ αβ T cells in a variety of cancers[25].

PD-1$^+$ γδ T cells are cytotoxic towards HLA class I-negative colon cancer cells

We next sought to determine whether tumor-infiltrating γδ T cells can recognize and kill CRC cells. We isolated and expanded PD-1$^-$ and PD-1$^+$ γδ T cells from five MMR-d colon cancers (**Figure S5A-C**, **Table S3**). In line with the scRNA-seq data, expanded PD-1$^+$ γδ T cell populations were devoid of Vδ2$^+$ cells and comprised of Vδ1$^+$ or Vδ3$^+$ subsets, whereas PD-1$^-$ fractions contained Vδ2$^+$ or a mixture of Vδ1/Vδ2/Vδ3$^+$ populations (**Figure 3A**, **Figure S5D**). Detailed immunophenotyping of the expanded γδ T cells (**Figure 3A**, **Figure S6A**) showed that all subsets expressed the activating receptor NKG2D, while the surface expression of KIRs was most frequent on PD-1$^+$ γδ T cells (Vδ1 or Vδ3$^+$), in line with the scRNA-seq results of unexpanded populations.

We measured the reactivity of the expanded γδ T cell populations towards HLA class I-negative and HLA class I-positive cancer cell lines (**Figure 3B**, **Figure S6B**). Upon co-culture with the different cancer cell lines, expression of activation markers and secretion of IFNγ was mainly induced in PD-1$^+$ γδ T cells (Vδ1 or Vδ3$^+$) and cell reactivity was most pronounced against HLA class I-negative cell lines (**Figure 3C**, **Figure S6-7**). Reactivity of PD-1$^-$ (enriched in Vδ2$^+$) subsets towards colorectal cancer cell lines was not detected (**Figure 3C**, **Figure S6-7**). To quantify and visualize the differences in killing of CRC cell lines by PD-1$^+$ and PD-1$^-$ γδ T cells, we co-cultured the γδ T cell populations with three CRC cell lines (HCT-15, LoVo, HT-29) in the presence of a fluorescent cleaved-caspase-3/7 reporter to measure cancer cell apoptosis over time (**Figure 3D-E**). This showed pronounced cancer cell apoptosis upon co-culture with PD-1$^+$ γδ T cells (Vδ1 or Vδ3$^+$) as compared to PD-1$^-$ cells, with highest killing of HLA class I-negative HCT-15 cells (**Figure 3E**, **Movie 1-2**). Re-introduction of *B2M* in the *B2M*-deficient HCT-15 and LoVo cells diminished their killing by PD-1$^+$ γδ T cells (Vδ1 or Vδ3$^+$) cells (**Figure S8**).

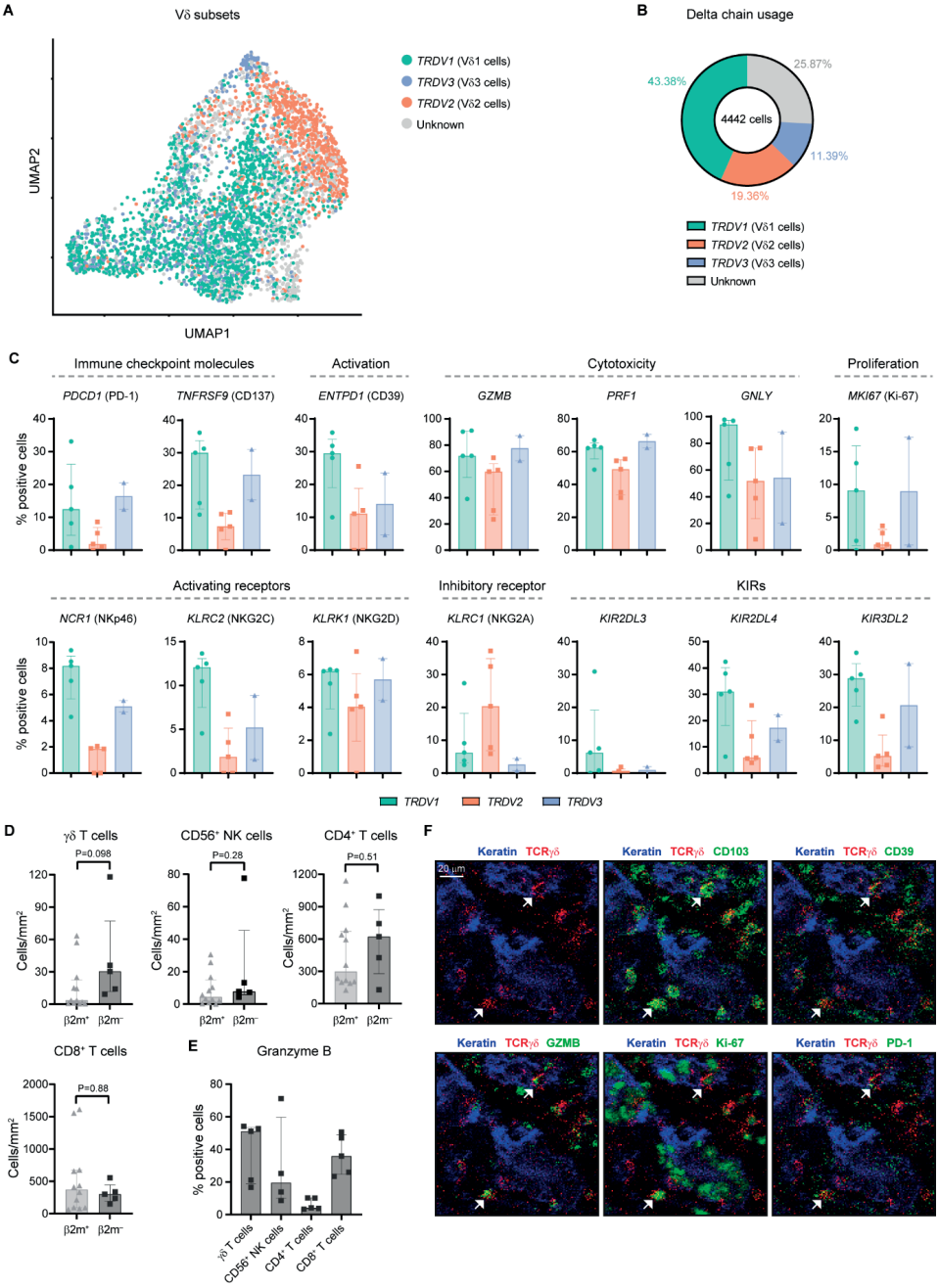


Figure 2. Tumor-infiltrating Vδ1 and Vδ3 T cell subsets display hallmarks of cytotoxic activity in MMR-d colon cancers.

A. UMAP embedding showing the clustering of γδ T cells (n=4442) isolated from MMR-d colon cancers (n=5) analyzed by single-cell RNA-sequencing. Colors represent the TCR Vδ chain usage. The functionally distinct γδ T cell clusters are shown in **Figure S3**. Each dot represents a single cell. **B**. Frequencies of the TCR Vδ chain usage of the γδ T cells (n=4442) analyzed by single-cell RNA-sequencing as a percentage of total γδ T cells. **C**. Frequencies of positive cells for

selected genes across Vδ1 (n=1927), Vδ2 (n=860), and Vδ3 (n=506) cells as percentage of total γδ T cells from each MMR-d colon tumor (n=5) analyzed by single-cell RNA-sequencing. Vδ3 cells were present in two out of five colon cancers. Bars indicate median ± IQR. Each dot represents an individual sample. **D**. Frequencies of γδ T cells, CD56$^+$ NK cells, CD4$^+$ T cells, and CD8$^+$ T cells in treatment-naïve β2m$^+$ (n=12) and β2m$^-$ (n=5) MMR-d colon cancers. Bars indicate median ± IQR. Each dot represents an individual sample. P-values were calculated by Wilcoxon rank sum test. **E**. Frequencies of granzyme B-positive γδ T cells, CD56$^+$ NK cells, CD4$^+$ T cells, and CD8$^+$ T cells in treatment-naïve β2m$^-$ (n=5) MMR-d colon cancers. CD56$^+$ NK cells were present in four out of five β2m$^-$ cancer samples. Bars indicate median ± IQR. Each dot represents an individual sample. **F**. Representative images of the detection of tissue-resident (CD103$^+$), activated (CD39$^+$), cytotoxic (granzyme B$^+$), proliferating (Ki-67$^+$), and PD-1$^+$ γδ T cells by imaging mass cytometry in an treatment-naïve, MMR-d colon cancer with β2m defect.

Next, we established two parental patient-derived tumor organoid lines (PDTOs; **Table S4**) of MMR-d CRC and generated isogenic *B2M^{KO}* lines using CRISPR. Genomic knockout of *B2M* effectively abrogated cell surface expression of HLA class I (**Figure S9**). We exposed two *B2M^{KO}* and their parental *B2M^{WT}* lines to the expanded γδ T cell subsets, and quantified γδ T cell activation by determination of IFNγ expression. Similarly to our cell line data, γδ T cells displayed increased reactivity towards *B2M^{KO}* PDTOs in comparison to the *B2M^{WT}* PDTOs (**Figure 3F-G**). Furthermore, γδ T cell reactivity towards *B2M^{KO}* tumor organoids was preferentially contained within the PD-1$^+$ population of γδ T cells (**Figure 3G**). Thus, lack of HLA class I antigen presentation in MMR-d tumor cells can be effectively sensed by γδ T cells and stimulates their antitumor response.

Expression of NKG2D on γδ T cells decreased during co-culture with target cells (**Figure S10A-B**), suggesting the involvement of the NKG2D receptor in γδ T cell activity. The NKG2D ligands MICA/B and ULBPs were expressed by the cancer cell lines (**Figure 3B**) and the MMR-d CRC PDTOs, irrespective of their *B2M* status (**Figure S9**). To explore which receptor-ligand interactions might regulate the activity of PD-1$^+$ γδ T cells, we performed blocking experiments focused on (i) NKG2D, (ii) DNAM-1, and (iii) γδ TCR signaling. Of these candidates, the only consistent inhibitory effect was observed for NKG2D ligand blocking on cancer cells, which decreased the activation and killing capacity of most PD-1$^+$ γδ T cells (**Figure 3H**, **Figure S10C-D**), confirming the mechanistic involvement of the NKG2D receptor in γδ T cell activation in this context. In addition, blocking NKG2D ligands on MMR-d CRC PDTOs reduced the PDTO-directed tumor reactivity of γδ T cells from CRC94 and CRC134 (**Figure 3I**). Together, these results show that γδ T cell reactivity towards MMR-d tumors is partly dependent on NKG2D/NKG2D-ligand interactions.

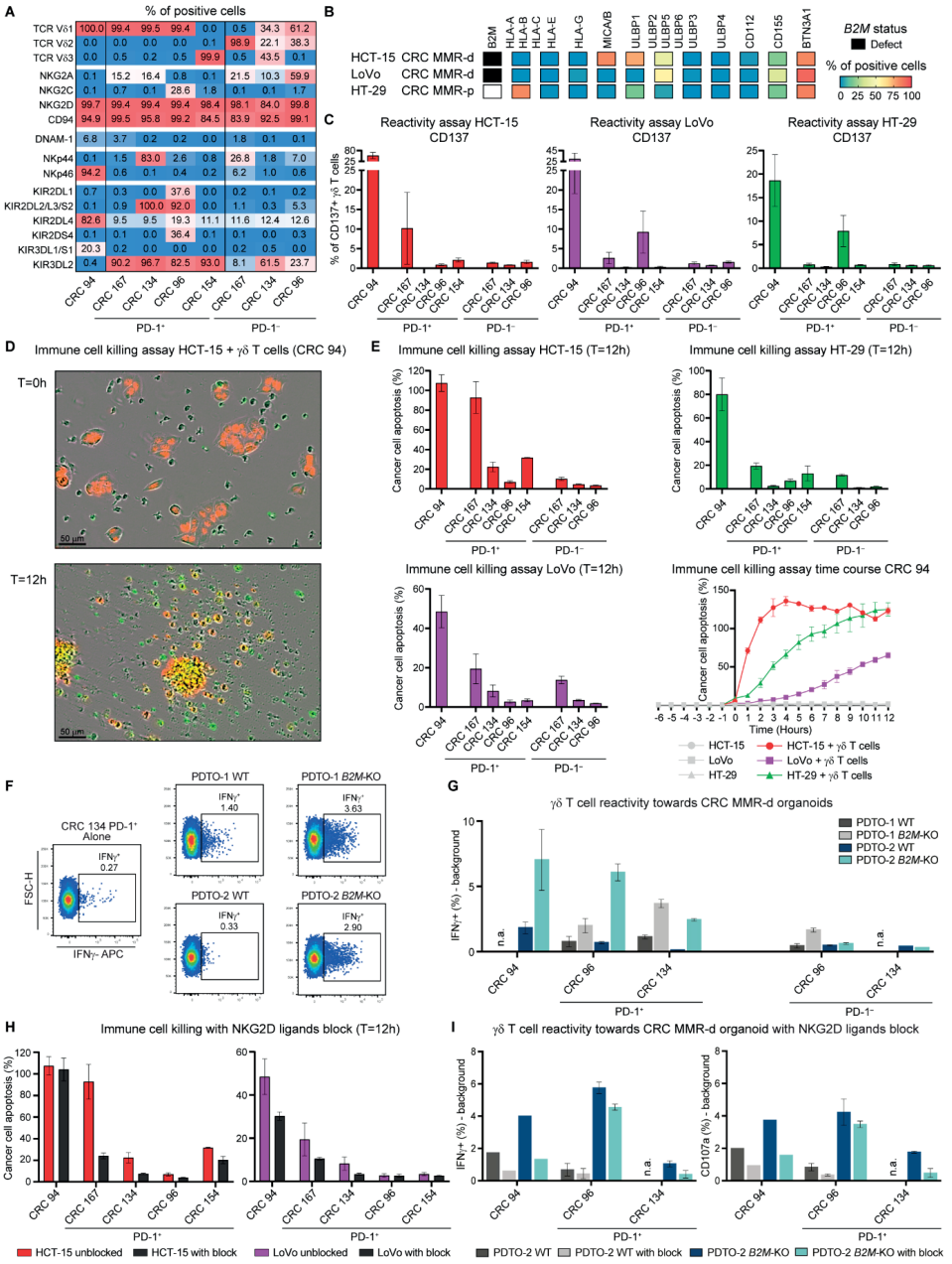


Figure 3. γδ T cells from MMR-d colon cancers show preferential reactivity towards HLA class I-negative cancer cell lines and organoids, which is regulated by NKG2D/NKG2D ligand interactions.

A. Table showing the percentage of positive cells for different TCR Vδ chains, innate immune receptors, and KIRs on expanded PD-1⁺ and PD-1⁻ γδ T cells sorted from MMR-d colon cancers (n=5) as percentage of total γδ T cells. **B**. Diagram showing the *B2M* mutational status and surface expression of HLA class I, NKG2D ligands, DNAM-1 ligands, and butyrophilin on CRC cell lines HCT-15, LoVo, and HT-29. **C**. Bar plots showing the percentage of CD137-positive γδ T cells after 18h co-culture of PD-1⁺ and PD-1⁻ γδ T cells from MMR-d colon cancers (n=5) with HCT-15, LoVo, and HT-29 cells. Medium as negative control and PMA/ionomycin as positive control are shown in **Figure S6**. Bars

indicate mean ± SEM. Data from four (CRC94), three (CRC167, CRC96), or two (CRC134, CRC154) independent experiments, depending on availability of γδ T cells. **D**. Representative images showing the killing of NucLight Red-transduced HCT-15 cells by γδ T cells (unlabeled) from CRC94 in the presence of a green fluorescent caspase-3/7 reagent in the IncuCyte S3. Images are taken immediately after the addition of γδ T cells (T=0) and 12h after. Cancer cell apoptosis is visualized in yellow. **E**. Bar plots showing the quantification of the killing of cancer cell lines by γδ T cells from MMR-d colon cancers (n=5) as in (**D**) after 12h co-culture. Bars indicate mean ± SEM of two wells with two images/well. At lower right, representative time course of cancer cell apoptosis in the presence or absence of γδ T cells derived from CRC94. **F**. Representative flow cytometry plots of PD-1$^+$ γδ T cells from CRC134 indicating IFN-γ expression in unstimulated condition (alone) and upon stimulation with two *B2M^{WT}* and *B2M^{KO}* CRC MMR-d organoids, as specified in the subplot titles. **G**. Histogram showing IFN-γ expression of γδ T cells from MMR-d colon cancers upon stimulation with two *B2M^{WT}* and *B2M^{KO}* CRC MMR-d organoids, as specified in the legend. Background IFN-γ signal of each unstimulated γδ T cell sample was subtracted from tumor organoid-stimulated IFN-γ signal. For all γδ T cell samples, data is shown for two biological replicates except for CRC134 PD-1$^-$ (n=1). Whiskers indicate SEM. **H**. Bar plots showing the quantification of killing of HCT-15 and LoVo cells by γδ T cells from MMR-d colon cancers (n=5) in the presence of blocking antibodies for NKG2D ligands as compared to the unblocked condition after 12h co-culture. Bars indicate mean ± SEM of two wells with two images/well. **I**. Histograms showing IFN-γ (left) and CD107a (right) expression in γδ T cells from MMR-d colon cancers upon stimulation with *B2M^{WT}* PDTO-2 (gray shades) or *B2M^{KO}* PDTO-2 (blue shades), with or without NKG2D ligand blocking (as indicated in the legend). For cultured γδ T cells, data is shown for two biological replicates (n=2) except for CRC94 (n=1). Whiskers indicate SEM.

Activated γδ T cells infiltrate ICB-treated *B2M*-mutant MMR-d colon cancers

We subsequently studied how ICB influences γδ T cell infiltration and activation in MMR-d colon cancers in a therapeutic context. For this purpose, we analysed pre- and post-treatment samples of the NICHE trial[9], in which colon cancer patients were treated with neoadjuvant PD-1 plus CTLA-4 blockade. In line with our observations in the DRUP cohort, 4 out of 5 (80%) *B2M^{MUT}* cancers in the NICHE trial showed a complete pathologic clinical response. IHC analysis confirmed tumor cell-specific loss of β2m protein expression in 4 out of 5 mutated cases. Whereas expression of immune marker gene sets in pre-treatment samples was similar between 5 *B2M^{MUT} vs* 13 *B2M^{WT}* cancers, ICB induced a clear immunological divergence between these two groups (**Figure 4A**). The *B2M^{MUT}* subgroup was most significantly associated with higher post-treatment expression of *TRDV1* and *TRDV3* (Wilcoxon rank sum-based two-sided P=0.0067; **Figure 4A**), followed by higher expression of the general immune cell marker CD45, NK cell-related markers, KIRs, and αβTCRs (Wilcoxon rank sum-based two-sided P=0.016, P=0.016, P=0.027, and P=0.043, respectively; **Figure 4A**, **Figure S11A**). The set of KIRs upregulated upon ICB in *B2M^{MUT}* cancers (**Figure S11B**) was consistent with the sets of KIRs upregulated in *B2M^{MUT}* MMR-d cancers in TCGA (**Figure 1E**), and those expressed by MMR-d tumor-infiltrating γδ T cells (**Figure 2C**). Pre- and post-ICB gene expression levels related to CD4 and CD8 infiltration were not associated with *B2M* status (**Figure 4A**, **Figure 11A**).

To quantify and investigate differences in immune profiles upon ICB treatment, we applied imaging mass cytometry in tissues derived from five *B2M^{MUT}* HLA class I-negative and five *B2M^{WT}* HLA class I-positive cancers pre- and post-ICB treatment. In the ICB-naïve setting, *B2M^{MUT}* MMR-d colon cancers showed higher γδ T cell infiltration as compared to *B2M^{WT}* MMR-d colon cancers (Wilcoxon rank sum-based two-sided P=0.032; **Figure 4B**, **Figure S11C**). No significant differences were observed in the infiltration of other immune cells, such as NK cells, CD4$^+$ T cells, and CD8$^+$ T cells in ICB-naïve *B2M^{MUT} vs B2M^{WT}* MMR-d colon cancers (**Figure 4B**). ICB treatment resulted in major pathologic clinical responses, and residual cancer

cells were absent in most post-ICB samples. All post-ICB tissues showed a profound infiltration of different types of immune cells (**Figure S11D**), of which γδ T cells were the only immune subset significantly higher in ICB-treated *B2M*^{MUT} as compared to *B2M*^{WT} MMR-d colon cancers (Wilcoxon rank sum-based two-sided P=0.016; **Figure 4B**, **Figure S11C**). In the sole *B2M*^{MUT} case that contained cancer cells, the majority of granzyme B^{+} immune cells infiltrating the tumor epithelium were γδ T cells (**Figure 4C**). These γδ T cells displayed co-expression of CD103, CD39, Ki-67, and PD-1 (**Figure S11E-G**). Taken together, these results show that ICB treatment of MMR-d colon cancer profoundly increases the intra-tumoral presence of activated, cytotoxic, and proliferating γδ T cells, especially when these cancers are β2m-deficient, highlighting γδ T cells as effectors of ICB treatment within this context.

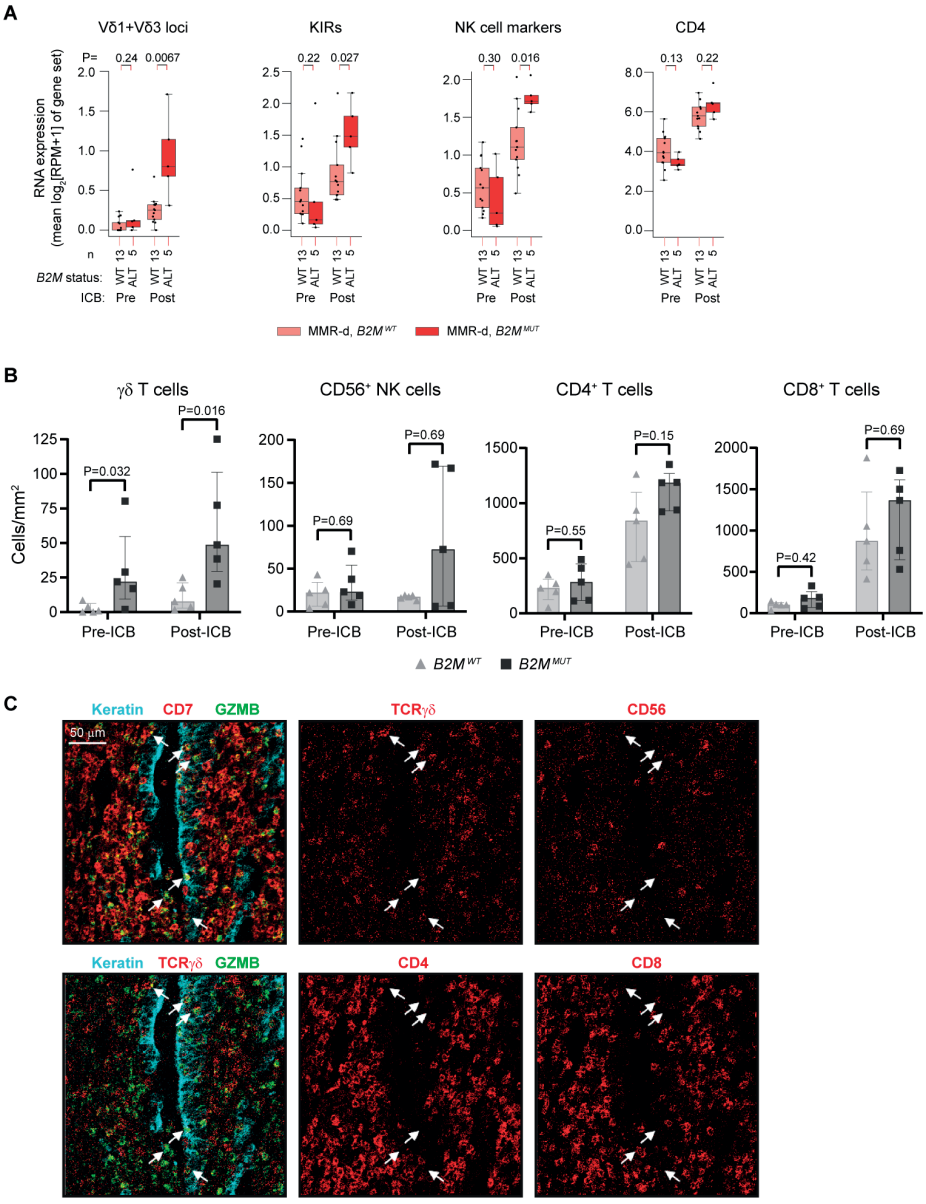


Figure 4. Immune checkpoint blockade (ICB) induces profound infiltration of γδ T cells into MMR-d colon cancers with antigen presentation defects.

A. The RNA expression of different immune marker gene sets in MMR-d *B2M*^{WT} (pink), and MMR-d *B2M*^{MUT} (red) cancers, before (left) and after (right) neoadjuvant ICB in the NICHE study. Boxes, whiskers, and dots indicate quartiles, 1.5 interquartile ranges, and individual data points, respectively. Wilcoxon rank sum test-based P-values are shown for MMR-d *B2M*^{WT} *vs* MMR-d *B2M*^{MUT} cancers. **B**. Frequencies of γδ T cells, CD56^{+} NK cells, CD4^{+} T cells, and CD8^{+} T cells in *B2M*^{WT} (n=5) and *B2M*^{MUT} (n=5) MMR-d colon cancers pre- and post-ICB treatment. Bars indicate median ± IQR. Each dot represents an individual sample. P-values were calculated by Wilcoxon rank sum test. **C**. Representative images of granzyme B-positive γδ T cells infiltrating the tumor epithelium in a *B2M*^{MUT} MMR-d colon cancer upon ICB treatment by imaging mass cytometry.

DISCUSSION

CD8$^+$ αβ T cells are major effectors of ICB[11,12,26] and rely on HLA class I antigen presentation of target cells. Paradoxically, we find that patients with HLA class I defects in MMR-d cancers retain their clinical benefit of ICB, suggesting the involvement of other immune effector cells compensating for the lack of conventional CD8$^+$ T cell immunity in this setting. We show that genomic inactivation of *B2M* in MMR-d colon cancers was associated with: (i) an elevated frequency of activated γδ T cells in ICB-naïve tumors, (ii) an increased presence of tumor-infiltrating γδ T cells upon ICB treatment, (iii) *in vitro* activation of tumor-infiltrating γδ T cells by colorectal cancer cell lines and PDTO, and iv) killing of these tumor cells by γδ T cells, in particular by Vδ1 and Vδ3 subsets expressing PD-1.

Different subsets of γδ T cells exhibit remarkably diverse functions which, in the context of cancer, ranges from tumor-promoting to tumoricidal effects.[27,28] Hence, it is of interest what defines antitumor reactivity of γδ T cells. This study is the first to isolate Vδ1/3-expressing PD-1$^+$ T cells as well as Vδ2-expressing PD-1$^-$ T cells from MMR-d tumor tissues. Our data suggest that especially tumor-infiltrating Vδ1 and Vδ3 T cells can recognize and kill HLA class I-negative MMR-d tumors, whereas Vγ9Vδ2 cells, the most studied and main subset of γδ T cells in the blood, appear to be less relevant within this context. This is in line with other studies showing that the cytotoxic ability of Vδ1 cells generally outperforms their Vδ2 counterparts.[29-33] Of note, cytotoxicity of tumor-infiltrating Vδ3 cells has, to our knowledge, not been reported before. Furthermore, the observation that PD-1$^+$ γδ T cells (Vδ1/Vδ3 phenotype) demonstrated clearly higher levels of antitumor reactivity as compared to their PD-1$^-$ counterparts (Vδ2 phenotype) suggests that, as for CD8$^+$ αβ T cells[34], PD-1 expression may be a marker of antitumor reactivity in γδ T cells.

The mechanisms of activation of γδ T cells are notoriously complex and diverse.[20] Specifically for Vδ1$^+$ cells, NKG2D has been described to be involved in tumor recognition, which is dependent on tumor cell expression of NKG2D ligands MICA/B and ULBPs.[35-37] In our study, MICA/B and ULBPs were highly expressed by the MMR-d CRC cell lines and tumor organoids, and blocking these ligands reduced γδ T cell activation and cytotoxicity. This suggests a role for the activation receptor NKG2D in γδ T cell reactivity towards HLA class I-negative MMR-d tumors. In addition, we detected expression of KIRs primarily on PD-1$^+$ γδ T cells (Vδ1 or Vδ3$^+$ subsets), whose antitumor reactivity and killing was clearly amplified when tumor cells lacked HLA class I.

Our findings have broad implications for cancer immunotherapy. First, our findings provide a basis for novel (combinatorial) immunotherapeutic approaches to further enhance γδ T cell-based antitumor immunity. Second, the presence or absence in tumors of specific γδ T cell subsets (e.g. Vδ1 or Vδ3) may help to define patients (un)responsive to ICB, especially in the case of MMR-d cancers and other malignancies with frequent HLA class I defects, like stomach adenocarcinoma[38] and Hodgkin lymphoma[39]. Third, our results suggest that MMR-d cancers and other tumors with HLA class I defects may be particularly attractive

targets for Vδ1 or Vδ3 T cell-based cellular therapies.

Although we have provided detailed and multidimensional analyses, it is likely that γδ T cells are not the only factor driving ICB responses in HLA class I-negative MMR-d CRC tumors. In this context, other HLA class I-independent immune subsets, like NK cells and (neoantigen-specific) $CD4^+$ T cells may also contribute. The latter were shown to play an important role in response to ICB (as reported in murine *B2M*-deficient MMR-d cancer models[40]), and may also support γδ T cell-driven responses. Of note, no equivalent subset to Vδ1 or Vδ3 T cells has been identified in mice which complicates their investigation in *in vivo* models. In conclusion, our results provide strong evidence that γδ T cells are cytotoxic effector cells of ICB treatment in HLA class I-negative MMR-d colon cancers, with implications for further exploitation of γδ T cells in cancer immunotherapy.

METHODS

TCGA data

RNA expression data (raw counts and Fragments Per Kilobase of transcript per Million mapped reads upper quartile FPKM-UQ) of the colon adenocarcinoma (COAD), stomach adenocarcinoma (STAD) and Uterus Corpus Endometrium Carcinoma (UCEC) cohorts of The Cancer Genome Atlas (TCGA) Research Network were downloaded via the GDC data portal (https://portal.gdc.cancer.gov) on April10th, 2019. Of these cohorts, mutation, copy number, purity and ploidy data were downloaded from GDC on November 11th, 2021, as the controlled access ABSOLUTE[41]-annotated MAF file (mutations), SNP6 whitelisted copy number segments file (copy numbers), and ABSOLUTE purity/ploidy file of the TCGA PanCanAtlas project[42]. Mismatch repair-deficiency status was obtained from Thorsson *et al.*[43] (TCGA Subtype = GI.HM-indel or UCEC.MSI).

DRUP data

A detailed description of the DRUP, including details on patient accrual, study design, oversight and endpoints has been published earlier in this journal[23]. Briefly, the DRUP is a national, non-randomized multi-drug and multi-tumor study in the Netherlands, in which patients receive off-label drugs registered for other treatment indications. We analyzed 71 patients with MMR-d cancers treated with PD-1 blockade in the DRUP[23] who had (i) clinical follow-up ≥16 weeks after start of PD-1 blockade treatment (i) WGS data passing standard quality controls (as defined before, including a sequencing-based tumor purity ≥20%[44]), (ii) available RNA-seq data (**Table S1**). MMR-d status was determined using routine diagnostics at the hospital of patient accrual and confirmed by WGS, based on an MSIseq[45] score >4, which represents a predefined threshold[44]. Clinical benefit was defined as disease control ≥16 weeks, consistent with the primary endpoint of the DRUP[23]. For genomics and transcriptomics analyses, fresh frozen tumor biopsies were obtained at baseline (i.e. prior to PD-1 blockade). WGS analysis (median depths ~100X and ~40X for tumor and normal, respectively) and bioinformatics were performed as previously described[23,44], with an optimized pipeline based on open source tools, which is freely available on GitHub (https://github.com/hartwigmedical/pipeline5). The tumor mutational burden per Mb (TMB) was determined by counting the genome-wide number of mutations (SNVs, MNVs, and indels) and dividing this number by the number of mega bases sequenced. For RNA-seq analysis, we extracted total RNA using the QIAGEN QIAsymphony RNA kit (cat nr 931636). Samples with approximately 100ng total RNA were prepared with KAPA RNA Hyper + RiboErase HMR (cat nr 8098131702) and RNA libraries were paired end sequenced on the Illumina NextSeq550 platform (2x75bp) or Illumina NovaSeq6000 platform (2x150bp). Raw RNA reads (FASTA files) were aligned to the human reference genome (GRCh38) with STAR software[46], version 2.7.7a, using default settings in two-pass mode.

NICHE study sequencing data

Raw RNA reads (FASTA files) of our recently published NICHE study[9] (ClinicalTrials.gov: NCT03026140) were generated as described in the original publication and aligned to the human reference genome (GRCh38) with STAR software[46], version 2.7.7a, using default settings in two-pass mode. For gene expression quantification, we used the gencode.v35.annotation.gtf annotation file. Somatic mutation data were obtained from DNA sequencing of pre-treatment tumor biopsies and matched germline DNA, as described in the original publication[9].

***B2M* status**

Consistent with the notion that both bi-allelic and mono-allelic non-synonymous *B2M* mutations are strongly associated with tumor-specific loss of β2m protein expression,[18] we considered all tumors with at least one somatic, non-synonymous *B2M* mutation to be "*B2M* mutant". As none of the *B2M* mutant tumors in TCGA harbored *B2M* copy number gains or losses, loss of heterozygosity (LOH) of *B2M* could be easily assessed by a simple calculation estimating the mutation's copy number:

$$Mut_{CN} = round(2 * \frac{VAF}{purity})$$

Here, Mut_{CN} represents the estimated mutation's copy number (rounded to an integer value), VAF represents the variant allele frequency of the mutation and purity equals the (ABSOLUTE[41]-based) tumor cell fraction of the sample.
A Mut_{CN} equal to 2 was considered consistent with LOH, as the most parsimonious explanation of such a result is the scenario in which all tumor-derived reads spanning the region of the *B2M* mutation harbor the mutation and none of the tumor-derived reads are wildtype.

In analyses of patients in the DRUP, LOH of *B2M* mutations was determined as an integrated functionality of PURPLE v2.34[47]. When multiple *B2M* mutations were present within a sample, we manually phased the mutations through inspection of the *B2M*-aligned reads using the Integrative Genomics Viewer (IGV)[48]. Here, mutations were only phased in case single reads were observed spanning the genomic locations of both mutations. We divided patients with multiple *B2M* mutations into three subgroups:

1. Bi-allelic, if:
    a. The multiple mutations were in trans AND the (integer) sum of the mutation copy numbers equaled (or exceeded) the integer copy number of the *B2M* gene (for mutations in cis, only one of these mutations was considered in the calculation); OR
    b. At least one of the mutations showed LOH.
2. Potentially bi-allelic, if:
    a. The multiple mutations affected too distant genomic locations and could therefore not be phased AND the (integer) sum of the mutation copy numbers equaled (or exceeded) the integer copy number of the *B2M* gene (for mutations in cis, only one of these mutations was considered in the calculation) AND none of the mutations showed LOH.
3. Not bi-allelic, if:
    a. The (integer) sum of the mutation copy numbers was smaller than the integer copy number of the *B2M* gene (for mutations in cis, only one of these mutations was considered in the calculation) AND none of the mutations showed LOH.

In these analyses, mutations were considered to be subclonal in case the probability of subclonality was >0.5 (the situation in which a mutation is more likely subclonal than clonal), as determined by PURPLE v2.34[47].

Association of *B2M* status with outcome and tumor characteristics
To test if somatic *B2M* alterations were associated with the clinical benefit rate of patients with MMR-d tumors treated with ICB in the DRUP, we used a Fisher's exact test (using the Python package Scipy[49]) for unadjusted analyses and logistic regression (as implemented by the Python package "Statsmodels" (https://pypi.org/project/statsmodels/), version 0.10.1) for analyses adjusted for the (continuous) tumor mutational burden per Mb (TMB) and/or the primary site of the tumor. The association of *B2M* status with TMB was tested using Scipy's Wilcoxon rank sum test. Associations of *B2M* status with the primary site of the tumor or the biopsy location was tested by Scipy's Fisher's exact test.

Association of TMB with ICB treatment outcome
For the DRUP cohort, the association of clinical benefit with TMB was tested using Statsmodels' Wilcoxon rank sum test.

Differential expression analysis

Differential RNA expression of genes was tested in R using EdgeR[50], Limma[51] and Voom[52]. Raw read counts were filtered by removing lowly expressed genes. Normalization factors were calculated using EdgeR, in order to transform the raw counts to $\log_2$ counts per million reads (CPM) and calculate residuals using Voom. Voom was then used to fit a smoothened curve to the √(residual standard deviation) by average gene expression, which was then plotted for visual inspection to confirm that the appropriate threshold was used for filtering of lowly expressed genes (defined as the minimal amount of filtering necessary to overcome a dipping mean-variance trend at low counts). Next, Limma was used to calculate differential expression of genes based on a linear model fit, considering the smoothened curve for sample weights, and empirical Bayes smoothing of standard errors. False discovery rates (FDRs) were calculated by Benjamini-Hochberg correction of the obtained p-values.

TCGA data

Using TCGA data, we calculated differential expression between tumors with and without high impact mutations in *B2M*, adjusting for tumor type and tumor mutational burden (TMB), using the following design formula: expression ~ Primary_Site + TMB + B2M_status (+ intercept by default), for which Primary_Site was a three-leveled factor (COAD, STAD, or UCEC), TMB was a continuous variable ($\log_{10}$[exome-wide number of mutations]) and B2M_status was a two-leveled factor (mutated, or wildtype).

NICHE study data

Using NICHE study data, we calculated differential expression between pre- and post-ICB treatment. In order to respect the paired nature of these data, we used the following design formula: expression ~ Patient + ICB (+ intercept by default), for which Patient was a factor for each individual patient and ICB was a two-leveled factor (ICB-treated yes/no).

Immune marker gene set expression analysis

To utilize RNA-seq data in order to obtain a relative estimate of the infiltration of specific immune cell types within tumors of TCGA, we summed the $\log_2$(FPKM-UQ+1) expression of genes that are specifically expressed in the immune cell types of interest. To this end, we used the marker gene sets published by Danaher *et al.*[24], and extended this by (i) *CD4* as $CD4^+$ T cell marker gene, (ii) *TRDV1* and *TRDV3* as γδ1/3T cell marker genes, and (iii) a killer-cell Ig-like receptor (KIR) gene set (comprised of all genes whose name starts with "KIR" and whose name contains "DL" or "DS". We excluded the gene set "NK CD56dim cells" of Danaher *et al*. (comprising IL21R, KIR2DL3, KIR3DL1, and KIR3DL2) from our analyses, as three out of four genes within this set were KIRs and hence this set showed high collinearity/redundancy to our full KIR gene set. As *XLC1* and *XLC2* are highly expressed by tumor-infiltrating γδ T cells, these genes were removed from the NK cell marker gene set and replaced by *KLRF1*, which encodes the well-established NK cell marker NKp80. The resulting gene set consisted of *NCR1* and *KLRF1*, encoding the well-established NK cell markers NKp46 and NKp80, respectively. Finally, we reduced the "cytotoxic cells" marker gene set of Danaher et al. to those genes in the set encoding cytotoxic molecules (GZMA, GZMB, GZMH, PRF1, GNLY, CTSW). The final collection of our marker gene sets can be found in **Table S2**.

Association of immune cell marker gene set expression with *B2M* alteration status (alteration yes/no) was calculated as follows:

1. For TCGA study-based analyses, we used (i) the Wilcoxon rank sum test (for unadjusted analyses) and (ii) ordinary least squares linear regression (for analyses adjusted for tumor type), using a similar design formula as for the differential gene expression analysis.
2. For DRUP cohort-based analyses, we used a linear mixed effects model (as implemented by the "lmer" function of the R package "lme4"), adjusting for tumor type and biopsy site as

random effects, using the following design formula: expression ~ B2M_status + (1ltumor_type) + (1lbiopsy_site) (+ intercept by default).

3. For NICHE study-based analyses, we used the Wilcoxon rank sum test (for unadjusted analyses).

Hierarchical clustering

Hierarchical clustering of expression profiles of individual genes or immune marker gene sets of TCGA cohorts was performed on Z-score-transformed $\log_2$(RPM+1) expression values, using the Python package Scipy[49], with Euclidean distance as distance metric and using the Ward variance minimization algorithm. Here, we used default settings with one exception: for visualization purposes, the color threshold was halved in the TCGA-based clustering of individual genes.

Patient samples

The DRUP was designed and conducted on behalf of the Center for Personalized Cancer Treatment (CPCT; clinicaltrials.gov: NCT02925234). This study was approved by the Medical Ethical Committee of the Netherlands Cancer Institute in Amsterdam, and was conducted in accordance with good clinical practice guidelines and the Declaration of Helsinki's ethical principles for medical research. Written informed consent was obtained from all study subjects. In addition, primary colon cancer tissues from a total of 17 patients with colon cancer who underwent surgical resection of their tumor at the Leiden University Medical Center (LUMC, the Netherlands; **Table S3**) were used for scRNA-seq, imaging mass cytometry, and functional assays. No patient with a previous history of inflammatory bowel disease was included. This study was approved by the Medical Ethical Committee of the Leiden University Medical Center (protocol P15.282), and patients provided written informed consent. In addition, primary colon cancer tissues from 10 patients with colon cancer included in the NICHE study (NCT03026140)[9] carried out at the Netherlands Cancer Institute (NKI, the Netherlands) were used for this study. All specimens were anonymized and handled according to the ethical guidelines described in the Code for Proper Secondary Use of Human Tissue in the Netherlands of the Dutch Federation of Medical Scientific Societies.

Processing of colorectal cancer tissues

Details on the processing of colorectal tumor tissues have been described previously[22]. In short, macroscopic sectioning from the lumen to the most invasive area of the tumor was performed. Tissues were collected in IMDM+Glutamax medium (Gibco) complemented with 20% fetal calf serum (FCS) (Sigma-Aldrich), 1% pen/strep (Gibco) and fungizone (Gibco), and 0.1% ciprofloxacin (provided by apothecary LUMC) and gentamicin (Invitrogen), and immediately cut into small fragments in a petri dish. Enzymatical digestion was performed with 1 mg/mL collagenase D (Roche Diagnostics) and 50 µg/mL DNase I (Roche Diagnostics) in 5 mL of IMDM+Glutamax medium for 30 min at 37°C in gentleMACS C tubes (Miltenyi Biotec). During and after incubation, cell suspensions were dissociated mechanically on the gentleMACS Dissociator (Miltenyi Biotec). Cell suspensions were filtered through a 70-µm cell strainer (Corning), washed in IMDM+Glutamax medium with 20% FCS, 1% pen/strep, and 0.1% fungizone, and cell count and viability were determined with the Muse Count & Viability Kit (Merck) on the Muse Cell Analyzer (Merck). Based on the number of viable cells, cells in IMDM+Glutamax medium were cryopreserved in liquid nitrogen until time of analysis complemented 1:1 with 80% FCS and 20% dimethyl sulfoxide (DMSO) (Merck).

Immunohistochemical detection of MMR, β2m, and HLA class I proteins

Tumor MMR status was determined by immunohistochemical detection of PMS2 (anti-PMS2 antibodies; clone EP51, DAKO) and MSH6 (anti-MSH6 antibodies; clone EPR3945, Abcam) proteins[53]. MMR-deficiency was defined as the lack of expression of at least one of the MMR-proteins in the presence of an internal positive control. Tumor β2m status was determined by immunohistochemical detection of β2m (anti-β2m antibodies; clone EP2978Y, Abcam). Immunohistochemical detection of HLA class I

expression on tumors was performed with HCA2 and HC10 monoclonal antibodies (Nordic-MUbio), and classified as HLA class I positive, weak, or loss as described previously[16]. For the tumor samples from the NICHE study, immunohistochemistry of the FFPE tissue was performed on a BenchMark Ultra autostainer (Ventana Medical Systems). Briefly, paraffin sections were cut at 3 µm, heated at 75°C for 28 min, and deparaffinized in the instrument with EZ prep solution (Ventana Medical Systems). Heat-induced antigen retrieval was carried out using Cell Conditioning 1 (CC1, Ventana Medical Systems) for 32 min at 95°C (HC10) or 64 min at 95°C (β2m and HCA2). HLA Class I Heavy Chain expression was detected using clone HCA2 (1/5000 dilution, 60 min at RT; Nordic-Mubio) and clone HC10 (dilution 1/20000, 32 min at 37°C; Nordic-Mubio). β2m was detected using clone D8P1H (dilution 1/1500, 60 min at RT; Cell Signaling). Bound antibody was detected using the OptiView DAB Detection Kit (Ventana Medical Systems). Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems). A PANNORAMIC® 1000 scanner from 3DHISTECH was used to scan the slides at a 40x magnification.

Sorting of γδ T cells from colon cancer and single-cell RNA-sequencing

scRNA-seq was performed on sorted γδ T cells from colon cancers (MMR-d) of five patients from the LUMC in the presence of hashtag oligo (HTOs) for sample ID and antibody-derived tags (ADTs) for CD45RA and CD45RO protein expression by CITE-seq[54]. Cells were thawed, rest at 37°C in IMDM (Lonza)/20% FCS for 1h, followed by incubation with human Fc receptor block (BioLegend) for 10 min at 4°C. Thereafter, cells were stained with cell surface antibodies (1:50 anti-CD3-PE [clone SK7, BD Biosciences], 1:160 anti-CD45-PerCP-Cy5.5 [clone 2D1, eBioscience], 1:200 anti-CD7-APC [clone 124-1D1, eBioscience], 1:60 anti-EPCAM-FITC [clone HEA-125, Miltenyi], 1:80 anti-TCRγδ-BV421 [clone 11F2, BD Biosciences], and a 1:1000 near-infrared viability dye [Life Technologies]), 1 µg of TotalSeq-C anti-CD45RA (clone HI100, BioLegend) and 1 µg of anti-CD45RO (clone UCHL1, BioLegend) antibodies, and 0.5 µg of a unique TotalSeq-C CD298/β2M hashtag antibody (clone LNH-94/2M2, BioLegend) for each sample (n=5) for 30 min at 4°C. Cells were washed three times in FACS buffer (PBS (Fresenius Kabi)/1% FCS) and kept cold and dark until cell sorting. Compensation was carried out with CompBeads (BD Biosciences) and ArC reactive beads (Life Technologies). Single, live $CD45^+$ $EPCAM^-$ $CD3^+$ $TCR\gamma\delta^+$ cells from five colorectal tumors (MMR-d) were sorted on a FACS Aria III 4L (BD Biosciences). After sorting, the samples were pooled.

scRNA-seq libraries were prepared using the Chromium Single Cell 5' Reagent Kit v1 chemistry (10X Genomics) following the manufacturer's instructions. The construction of 5' Gene Expression libraries allowed the identification of γδ T cell subsets according to Vδ and Vγ usage. Libraries were sequenced on a HiSeq X Ten using paired-end 2x150 bp sequencing (Illumina). Reads were aligned to the human reference genome (GRCh38) and quantified using Cell Ranger (version 3.1.0). Downstream analysis was performed using Seurat (version 3.1.5) according to the author's instructions[55]. Briefly, cells that had less than 200 detected genes and genes that were expressed in less than six cells were excluded. The resulting 5669 cells were demultiplexed based on HTO enrichment using the MULTIseqDemux algorithm[56]. Next, cells with a mitochondrial gene content greater than 10% and cells with outlying numbers of expressed genes (>3000) were filtered out from the analysis, resulting in a final dataset of 4442 cells. Data were normalized using the 'LogNormalize' function from Seurat with scale factor 10,000. Variable features were identified using the 'FindVariableFeatures' function from Seurat returning 2,000 features. We then applied the 'RunFastMNN' function from SeuratWrappers split by sample ID to adjust for potential batch-derived effects across samples[57]. Uniform manifold approximation (UMAP)[58] was used to visualize the cells in a two-dimensional space, followed by the 'FindNeighbors' and 'FindClusters' functions from Seurat. Data were scaled and heterogeneity associated with mitochondrial contamination was regressed out. Cell clusters were identified by performing differentially expressed gene analysis with the 'FindAllMarkers' function with min.pct and logfc.threshold at 0.25. Percentage of *TRDV1* (Vδ1), *TRDV2* (Vδ2), or *TRDV3* (Vδ3) positive cells was determined as the percentage of all cells with an expression level of >1, while <1 for the other TCR Vδ chains. CRC96, 134 and 167 had less than ten $TRDV3^+$ cells,

and were not included in the Vδ3 analysis. Transcripts of Vδ4 *(TRDV4),* Vδ5 *(TRDV5),* and Vδ8 *(TRDV8)* cells were not detected. Percentage of *TRGV1* (Vγ1) – *TRGV11* (Vγ11) positive cells was determined as the percentage of all cells with an expression level of >1, while <1 for the other TCR Vγ chains. Percentage of cells positive for a certain gene was determined as all cells with an expression level of >1.

Imaging mass cytometry staining and analysis

Imaging mass cytometry (IMC) was performed on ICB-naïve colon cancer tissues (MMR-d) of 17 patients from the LUMC, of which four HLA class I-positive, eight HLA class I-defect, and five β2m-defect (**Table S1**). In addition, IMC was performed on ICB-treated colon cancer tissues (MMR-d) of ten patients from the NKI, of which five *B2M*^{WT} and five *B2M*^{MUT}. Antibody conjugation and immunodetection were performed following the methodology published previously by Ijsselsteijn *et al.*[59]. Four-µm FFPE tissue were incubated with 41 antibodies in four steps. First, sections were incubated with anti-CD4 and anti-TCRδ overnight at RT, which were subsequently detected using metal-conjugated secondary antibodies (goat anti-rabbit IgG and goat anti-mouse IgG, respectively; Abcam). Second, sections were incubated with 20 antibodies (**Table S3**) for five hours at RT. Third, sections were incubated overnight at 4°C with the remaining 19 antibodies (**Table S3**). Fourth, sections were incubated with 0.125 µM Cell-ID intercalator-Ir (Fluidigm) to detect the DNA, and stored dry until measurement. For each sample, six 1000x1000µm regions were selected based on consecutive Haematoxylin and Eosin (H&E) stains and ablated using the Hyperion Imaging system (Fluidigm). Data was acquired with the CyTOF Software (version 7.0) and exported with MCD Viewer (version 1.0.5). Data was normalized using semi-automated background removal in ilastik[60], version 1.3.3, to control for variations in signal-to-noise between FFPE sections as described previously[61]. Thereafter, the phenotype data was normalized at pixel level. Cell segmentation masks were created for all cells in ilastik and CellProfiler[62], version 2.2.0. In ImaCytE[63], version 1.1.4, cell segmentation masks and normalized images were combined to generate single-cell FCS files containing the relative frequency of positive pixels for each marker per cell. Cells forming visual neighborhoods in a t-distributed Stochastic Neighbor Embedding (t-SNE)[64] embedding in Cytosplore[65], version 2.3.0, were grouped and exported as separate FCS files. The resulting subsets were imported back into ImaCyte and visualized on the segmentation masks. Expression of immunomodulatory markers was determined as all cells with a relative frequency of at least 0.2 positive pixels per cell. Differences in cells/mm^2 were calculated by Mann-Whitney tests in Graphpad Prism (version 9.0.1).

Sorting of γδ T cells from colon cancer and cell culturing

γδ T cells from colon cancers (MMR-d) of five patients from the LUMC were sorted for cell culture. Cells were thawed and rest at 37°C in IMDM (Lonza)/10% nHS for 1h. Thereafter, cells were incubated with human Fc receptor block (BioLegend) and stained with cell surface antibodies (1:20 anti-CD3-Am Cyan [clone SK7, BD Biosciences], 1:80 anti-TCRγδ-BV421 [clone 11F2, BD Biosciences], and 1:30 anti-PD-1-PE [clone MIH4, eBioscience] for 45 min at 4°C together with different additional antibodies for immunophenotyping (including 1:10 anti-CD103-FITC [clone Ber-ACT8, BD Biosciences], 1:200 anti-CD38-PE-Cy7 [clone HIT2, eBioscience], 1:60 anti-CD39-APC [clone A1, BioLegend], 1:20 anti-CD45RA-PE-Dazzle594 [clone HI100, Sony], 1:20 anti-CD45RO-PerCP-Cy5.5 [clone UCHL1, Sony], 1:40 anti-TCRαβ-PE-Cy7 [clone IP26, BioLegend], 1:50 anti-TCRVδ1-FITC [clone TS8.2, Invitrogen], or 1:200 anti-TCRVδ2-PerCP-Cy5.5 [clone B6, BioLegend]. A 1:1000 live/dead fixable near-infrared viability dye (Life Technologies) was included in each staining. Cell were washed three times in FACS buffer (PBS/1% FCS) and kept cold and dark until cell sorting. Compensation was carried out with CompBeads (BD Biosciences) and ArC reactive beads (Life Technologies). Single, live CD3$^+$ TCRγδ$^+$ PD-1$^+$ and PD-1$^-$ cells from five colorectal tumors (MMR-d) were sorted on a FACS Aria III 4L (BD Biosciences). For CRC94 all γδ T cells were sorted due to the low number of PD-1$^+$ cells. γδ T cells were sorted in medium containing feeder cells (1x10^6/mL), PHA (1 µg/mL; Thermo Fisher Scientific), IL-2 (1000 IU/mL; Novartis), IL-15 (10 ng/mL; R&D Systems), gentamicin (50 µg/mL), and fungizone (0.5 µg/mL). Sorted γδ T cells were expanded in the presence of 1000 IU/mL IL-2 and 10 ng/mL IL-15 for three-four weeks. Purity and phenotype of γδ T cells were assessed by flow cytometry. We obtained a >170,000-fold increase in 3-4

weeks of expansion of γδ T cells (**Figure S5C**).

Immunophenotyping of expanded γδ T cells by flow cytometry

Expanded γδ T cells from colon tumors were analyzed by flow cytometry for the expression of TCR Vδ chains, NKG2 receptors, NCRs, KIRs, tissue-residency/activation markers, cytotoxic molecules, immune checkpoint molecules, cytokine receptors, and Fc receptors. Briefly, cells were incubated with human Fc receptor block (BioLegend) and stained with cell surface antibodies (**Table S4**) for 45 min at 4°C, followed by three washing steps in FACS buffer (PBS/1% FCS). Granzyme B and perforin were detected intracellularly using Fixation Buffer and Intracellular Staining Permeabilization Wash Buffer (BioLegend). Compensation was carried out with CompBeads (BD Biosciences) and ArC reactive beads (Life Technologies). Cells were acquired on a FACS LSR Fortessa 4L (BD Biosciences) running FACSDiva software version 9.0 (BD Biosciences). Data were analyzed with FlowJo software version 10.6.1 (Tree Star Inc).

Cancer cell line models and culture

Human colorectal adenocarcinoma cell lines HCT-15 (MMR-d), LoVo (MMR-d), HT-29 (MMR-p), SW403 (MMR-p), and SK-CO-1 (MMR-p) as well as HLA class I deficient human leukemia cell line K-562 and Burkitt lymphoma cell line Daudi were used as targets for reactivity and immune cell killing assays. The cell lines were authenticated by STR profiling and tested for mycoplasma. HCT-15, LoVo, HT-29, K-562, and Daudi cells were maintained in RPMI (Gibco)/10% FCS. SW403 and SK-CO-1 were maintained in DMEM/F12 (Gibco)/10% FCS. All adherent cell lines were trypsinized before passaging. The *B2M*-knockin HCT-15 and LoVo cell lines were generated by using the *B2M* plasmid (pLV[Exp]-EF1A>hB2M[NM_004048.4](ns):T2A:Puro), produced in lentivirus according to standard methodology. Cells were selected using puromycin and afterwards FACS-sorted based on HLA-A/B/C expression using 1:100 anti-HLA-A/B/C-FITC [clone W6/32, eBioscience].

Organoid models and culture

Tumor organoids were derived from MMR-d CRC tumor of two patients via resection from the colon, tumor organoid 1, or peritoneal biopsy, tumor organoid 2 (**Table S2**). Establishment of the respective organoid lines from tumor material was performed as previously reported[66,67]. Briefly, tumor tissue was mechanically dissociated and digested with 1.5 mg/mL of collagenase II (Sigma-Aldrich), 10 µg/mL of hyaluronidase type IV (Sigma-Aldrich), and 10 µM Y-27632 (Sigma-Aldrich). Cells were embedded in Cultrex® RGF BME Type 2 (cat no. 3533-005-02, R&D systems) and placed in a 37°C incubator for 20 min. Human CRC organoids medium is composed of Ad-DF+++ (Advanced DMEM/F12 (GIBCO) supplemented with 2 mM Ultraglutamine I (Lonza), 10 mM HEPES (GIBCO), and 100/100 U/mL Penicillin/Streptomycin (GIBCO), 10% Noggin-conditioned medium, 20% R-spondin1- conditioned medium, 1x B27 supplement without vitamin A (GIBCO), 1.25 mM N- acetylcysteine (Sigma-Aldrich), 10 mM nicotinamide (Sigma-Aldrich), 50 ng/mL human recombinant EGF (Peprotech), 500 nM A83-01 (Tocris), 3 µM SB202190 (Cayman Chemicals) and 10 nM prostaglandin E2 (Cayman Chemicals). Organoids were passaged depending on growth every 1–2 weeks by incubating in TrypLE Express (Gibco) for 5–10 min followed by embedding in BME. Organoids were authenticated by SNP array or STR profile and regularly tested for Mycoplasma using Mycoplasma PCR43 and the MycoAlert Mycoplasma Detection Kit (cat no. LT07-318). In the first two weeks of organoid culture, 1x Primocin (Invivogen) was added to prevent microbial contamination. Procedures performed with patient specimens were approved by the Medical Ethical Committee of the Netherlands Cancer Institute – Antoni van Leeuwenhoek hospital (study NL48824.031.14) and written informed consent was obtained from all patients. Mismatch repair status was assessed by standard protocol for the Ventana automated immunostainer for MLH1 clone M1 (Roche), MSH2 clone G219-1129 (Roche), MSH6 clone EP49 (Abcam) and PMS2 clone EP51 (Agilant Technologies). The *B2M*^{KO} tumor organoid lines were generated by using sgRNA targeting *B2M* (GGCCGAGATGTCTCGCTCCG), cloned into LentiCRISPR v2 plasmid. The virus was produced by standard method.

Screening of cancer cell lines and tumor organoids by flow cytometry

The cancer cell lines used in the reactivity and killing assays were screened for the expression of β2m, HLA class I molecules, NKG2D ligands, DNAM-1 ligands, and butyrophilin by flow cytometry. Briefly, cells were incubated with human Fc receptor block (BioLegend) and stained with the different cell surface antibodies (1:10 anti-CD112-PE [clone R2.525, BD Biosciences], 1:10 anti-CD155-PE [clone 300907, R&D Systems], 1:50 anti-CD277/BTN3A1-PE [clone BT3.1, Miltenyi], 1:100 anti-β2M-PE [clone 2M2, BioLegend], 1:100 anti-HLA-A/B/C-FITC [clone W6/32, eBioscience], 1:160 anti-HLA-A/B/C-AF647 [clone W6/32, BioLegend], 1:20 anti-HLA-E-BV421 [clone 3D12, BioLegend], 1:20 anti-HLA-G-APC [clone 87G, BioLegend], 1:300 anti-MICA/B-PE [clone 6D4, BioLegend], 1:10 anti-ULBP1-PE [clone 170818, R&D Systems], 1:20 anti-ULBP2/5/6-PE [clone 165903, R&D Systems], 1:20 anti-ULBP3-PE [clone 166510, R&D Systems], or 1:20 anti-ULBP4-PE [clone 709116, R&D Systems] for 45 min at 4°C. A 1:1000 live/dead fixable near-infrared viability dye (Life Technologies) was included in each staining. Cells were washed three times in FACS buffer (PBS/1% FCS). Compensation was carried out with CompBeads (BD Biosciences) and ArC reactive beads (Life Technologies). Cells were acquired on a FACS Canto II 3L or FACS LSR Fortessa 4L (BD Biosciences) running FACSDiva software version 9.0 (BD Biosciences). Isotype or FMO controls were included to determine the percentage of positive cancer cells. Data were analyzed with FlowJo software version 10.6.1 (Tree Star Inc).

For organoid surface staining, tumor organoids were dissociated into single cells using TrypLE Express (Gibco) washed twice in cold FACS buffer (PBS, 5 mM EDTA, 1% bovine serum antigen) and stained with either 1:20 anti-HLA-A,B,C-PE (clone W6/32, BD Biosciences), 1:100 anti-β2m-FITC (clone 2M2, BioLegend), 1:200 anti-PD-L1 (clone MIH1, eBioscience) and 1:2000 near-infrared (NIR) viability dye (Life Technologies) or isotype controls (FITC, PE or APC) mouse IgG1 kappa (BD Biosciences). For NKG2D ligand expression analysis cells were stained with 1:300 anti-MICA/MICB, 1:10 anti-ULBP1, 1:20 anti-ULBP2/5/6, 1:20 anti-ULBP3, 1:20 anti-ULBP4, and 1:2000 near-infrared (NIR) viability dye (Life Technologies). Tumor cells were incubated for 30 min at 4°C in the dark and washed twice in FACS buffer. All samples were recorded at a Becton Dickinson Fortessa.

Reactivity assay γδ T cells

Reactivity of γδ T cells to the different cancer cell lines was assessed by a co-culture reactivity assay. γδ T cells were thawed and cultured in IMDM+Glutamax (Gibco)/8% nHS medium with pen (100 IU/mL)/strep (100 µg/mL) in the presence of low-dose IL-2 (25 IU/mL) and IL-15 (5 ng/mL) overnight at 37°C. Cancer cell lines were counted, adjusted to a concentration of $0.5x10^5$ cells/mL in IMDM+Glutamax/10% FCS medium with pen (100 IU/mL)/strep (100 µg/mL), and seeded (100 µL/well) in coated 96-well flat-bottom microplates (Greiner CellStar) (for 5,000 cells/well) overnight at 37°C. The next day, γδ T cells were harvested, counted, and adjusted to a concentration of $1.2x10^6$ cells/mL in IMDM+Glutamax/10% FCS medium. The γδ T cells were added in 50 µL (for 60,000 cells/well) and co-cultured (12:1 E:T ratio) at 37°C for 18h in biological triplicates. Medium (without cancer cells) was used as negative control and PMA (20 ng/mL)/ionomycin (1 µg/mL) as positive control. After co-culture, the supernatant was harvested to detect IFN-γ secretion by ELISA (Mabtech) following the manufacturer's instructions. Additionally, cells were harvested, incubated with human Fc receptor block (BioLegend), and stained with cell surface antibodies (1:100 anti-CD137-APC [clone 4B4-1, BD Biosciences], 1:150 anti-CD226/DNAM-1-BV510 [clone DX11, BD Biosciences], 1:400 anti-CD3-AF700 [clone UCHT1, BD Biosciences], 1:80 anti-CD39-APC [clone A1, BioLegend], 1:10 anti-CD40L-PE [clone TRAP1, BD Biosciences] or 1:30 anti-PD-1-PE [clone MIH4, eBioscience], 1:40 anti-TCRγδ-BV650 [clone 11F2, BD Biosciences], 1:300 anti-NKG2D-PE-Cy7 [clone 1D11, BD Biosciences], and 1:20 anti-OX40-FITC [clone ACT35, BioLegend] for 45 min at 4°C. A 1:1000 live/dead fixable near-infrared viability dye (Life Technologies) was included in each staining. Cells were washed three times in FACS buffer (PBS/1% FCS). Compensation was carried out with CompBeads (BD Biosciences) and ArC reactive beads (Life Technologies). Cells were acquired on a FACS LSR Fortessa X-20 4L (BD Biosciences) running FACSDiva software version 9.0 (BD Biosciences). Data were analyzed with FlowJo software version 10.6.1 (Tree Star Inc). All data are

representative of at least two independent experiments.

Immune cell killing assay γδ T cells

Killing of the different cancer cell lines by γδ T cells was visualized and quantified by a co-culture immune cell killing assay using the IncuCyte S3 Live-Cell Analysis System (Essen Bioscience). HCT-15, LoVo, and HT-29 cells were transduced with IncuCyte NucLight Red Lentivirus Reagent (EF-1α, Puro; Essen BioScience) providing a nuclear-restricted expression of a red (mKate2) fluorescent protein. In short, HCT-15, LoVo and HT-29 were seeded, transduced according to the manufacturer's instructions, and stable cell populations were generated using puromycin selection. The *B2M*-knockin cell lines were created under puromycin selection, hence, stable NucLight Red-expressing cell populations were generated by sorting mKate2 (the red fluorescent protein) in the PE Texas Red filter set instead. Cancer cell lines were counted, adjusted to a concentration of $1x10^5$ cells/mL in IMDM+Glutamax/10% FCS medium with pen (100 IU/mL)/strep (100 µg/mL), and seeded (100 µL/well) in 96-well flat-bottom clear microplates (Greiner CellStar) (for 10,000 cells/well). The target cell plate was placed in the IncuCyte system at 37°C to monitor for cell confluency for 3 days. On day 2, γδ T cells were thawed and cultured in IMDM+Glutamax/8% nHS medium with pen (100 IU/mL)/strep (100 µg/mL) in the presence of low-dose IL-2 (25 IU/mL) and IL-15 (5 ng/mL) overnight at 37°C. The next day, γδ T cells were harvested, counted, and adjusted to a concentration of $7.2x10^5$ cells/mL in IMDM+Glutamax/10% FCS medium. After aspiration of the medium of the target cell plate, 100 µL of new medium containing 3.75 µM IncuCyte Caspase-3/7 Green Apoptosis Reagent (Essen BioScience) (1.5x final assay concentration of 2.5 µM) was added together with 50 µL of γδ T cells (for 36,000 cells/well). They were co-cultured (4:1 E:T ratio) in the IncuCyte system at 37°C in biological duplicates. Cancer cells alone and cancer cells alone with Caspase-3/7 were used as negative controls. Images (2 images/well) were captured every hour at 20x magnification with the phase, green, and red channels for up to 4 days.

Analysis was performed in the IncuCyte software (version 2020B) for each cancer cell line separately. The following analysis definitions were applied: 1) for HCT-15 cells in the phase channel a minimum area of 200 μm^2, in the green channel a threshold of 2 GCU, and in the red channel a threshold of 2 RCU, 2) for LoVo and HT-29 cells in the phase channel a minimum area of 200 μm^2, in the green channel a threshold of 4 GCU, and in the red channel a threshold of 2 RCU. Cancer cell apoptosis was then quantified in the IncuCyte software by counting the total number of Green + Red objects per image normalized (by division) to the total number of Red objects per image after 12h co-culture and displayed as a percentage (mean ± SEM) of two wells with two images/well. For the comparison of the killing of *B2M*-knockin HCT-15 and LoVo cell lines *vs* the wildtype cell lines, Caspase-3/7 Red Apoptosis Reagent (Essen BioScience) was used. The transfection of the target reporter was not as successful in combination with the *B2M*-knockin. Hence, apoptosis was quantified by dividing the Red area by the Phase area and displayed as a percentage (mean ± SEM) of two wells with two images/well. The following analysis definitions were applied: a minimum phase area of 100 μm^2 and a RCU of 0.5 (for HCT-15) and 0.75 (for LoVo).

Tumor organoid recognition assay

For evaluation of tumor reactivity toward $B2M^{WT}$ and $B2M^{KO}$ organoids and NKG2D ligand blocking conditions, tumor organoids and γδ T cells were prepared as described previously.[9,66,67] Two days prior to the experiment organoids were isolated from BME by incubation in 2 mg/mL type II dispase (Sigma-Aldrich) for 15 min before addition of 5 mM ethylendiaminetetraacetic acid (EDTA) and washed with PBS before resuspended in CRC organoid medium with 10 µM Y-27632 (Sigma-Aldrich). Organoids were stimulated with 200 ng/mL IFN-γ (Peprotech) 24 hours before the experiment. For the recognition assay and intra-cellular staining tumor organoids were dissociated into single cells and plated in anti-CD28 (clone CD28.2 eBioscience) coated 96-well U-bottom plates with γδ T cells at a 1:1 target:effector ratio in the presence of 20 µg/mL anti-PD-1 (Merus). As positive control γδ T cells were stimulated with 50 ng/mL of phorbol 12-myristate 13 -acetate (Sigma-Aldrich) and 1 µg/mL of ionomycin (Sigma-Aldrich). After

1h of incubation at 37°C, GolgiSTOP (BD Biosciences, 1:1500) and GolgiPlug (BD Biosciences, 1:1000) were added. After 4h of incubation at 37°C, γδ T cells were washed twice in cold FACS buffer (PBS, 5 mM EDTA, 1% bovine serum antigen) and stained with 1:20 anti-CD3-PerCP-Cy5.5 (BD Biosciences), 1:20 anti-TCRγδ-PE (BD Bioscience), 1:20 anti-CD4-FITC (BD Bioscience) (not added in experiments with NKG2D ligand blocking), 1:200 anti-CD8-BV421 (BD Biosciences) and 1:2000 near-infrared (NIR) viability dye (Life Technologies) for 30 min at 4°C. Cells were washed, fixed and stained with 1:40 anti-IFN-γ-APC (BD Biosciences) for 30 min at 4°C, using the Cytofix/Cytoperm Kit (BD Biosciences). After two washing steps, cells were resuspended in FACS buffer and recorded at a BD LSRFortessa™ Cell Analyzer SORP flow cytometer with FACSDiVa 8.0.2 (BD Biosciences) software.

Blocking experiments with cancer cell lines and tumor organoids

Reactivity of and killing by the γδ T cells was examined in the presence of different blocking antibodies to investigate which receptor-ligand interactions are involved. For DNAM-1 blocking, γδ T cells were incubated with 3 µg/mL purified anti-DNAM-1 (clone DX11, BD Biosciences) for 1h at 37°C. For γδ TCR blocking, γδ T cells were incubated with 3 µg/mL purified anti-TCRγδ (clone 5A6.E9, Invitrogen) for 1h at 37°C, of which the clone we used was tested to be best for use in γδ TCR blocking assays[68]. NKG2D ligands were blocked on the cancer cell lines and single cells of tumor organoids by incubating the target cells with 12 µg/mL anti-MICA/B (clone 6D4, BioLegend), 1 µg/mL anti-ULBP1 (clone 170818, R&D Systems), 3 µg/mL anti-ULBP2/5/6 (clone 165903, R&D Systems), and 6 µg/mL anti-ULBP3 (clone 166510, R&D Systems) for 1h at 37°C prior to plating with γδ T cells. After incubation with the blocking antibodies, the γδ T cells were added to cancer cell lines HCT-15, LoVo, and HT-29 as described above with a minimum of two biological replicates per blocking condition. For organoid experiments, 1:50 anti-CD107a-FITC (clone H4A3, BioLegend) was added during incubation.

As a control for Fc-mediated antibody effector functions, γδ T cells alone were incubated with the blocking antibodies in the presence of 2.5 µM IncuCyte Caspase-3/7 Green Apoptosis Reagent (Essen BioScience) in the IncuCyte system at 37°C, and the number of apoptotic γδ T cells was quantified over time.

Data availability

The used TCGA data is publicly available via the National Cancer Institute GDC Data Portal (https://portal.gdc.cancer.gov; cohorts COAD, STAD and UCEC). Of DRUP study subjects included in this preliminary analysis across all (complete and incomplete) cohorts of the study, we included all clinical data, genomics data on *B2M* status and RNA-expression data of marker gene sets to this manuscript in **Table S1**. As mentioned in the original publication, NICHE study RNA- and DNA-sequencing data is deposited into the European Genome–Phenome Archive under accession no. EGAS00001004160 and will be made available on reasonable request for academic use and within the limitations of the provided informed consent. The single-cell RNA-sequencing data are available from the corresponding author upon request. All other data are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENTS

We thank K.C.M.J. Peeters, M.G. Kallenberg-Lantrua, D. Berends-van der Meer, and F.A. Holman for their help in collecting and providing samples from patients with colon cancer; the Flow cytometry Core Facility of the Leiden University Medical Center for their help with cell sorting; the Leiden Genome Technology Center for their help with single-cell RNA-sequencing; and M. Ganesh for help with cell culturing. We thank D. Thommen for fruitful discussions, I. S. Rodriguez for the establishment of a *B2M*-knockout organoid line, L. Hoes for initial clinical findings, the Flow Cytometry Core Facility at the Netherlands Cancer Institute for their support; X. Kong for providing the lentiCRISPR plasmid for *B2M*-knockout; Merus for providing anti-PD-1 for organoid experiments; The Cancer Genome Atlas (TCGA) for providing data used in this manuscript.

AUTHOR CONTRIBUTIONS

N.L.d.V., J.v.d.H. and V.V. conceived the study and performed experiments. J.v.d.H. performed bulk transcriptomic analyses of ICB-naïve as well as ICB-treated MMR-d cancers. N.L.d.V. performed single-cell RNA-sequencing and cell culturing experiments. V.V. performed organoid experiments. N.L.d.V., M.v.d.P. and J.v.d.B. performed cell line and IncuCyte experiments. N.L.d.V., V.V. and M.v.d.P. performed blocking experiments. M.C. provided tissue sections of patients in the NICHE study, which was designed and coordinated by M.C. under joint supervision of T.N.S., E.E.V. and J.B.H. M.E.I. carried out imaging mass cytometry experiments. M.E.I., N.F.C.C.d.M., N.L.d.V. and D.R. analyzed the imaging mass cytometry data. N.L.d.V. and D.R. analyzed the single-cell RNA-sequencing data. J.G.v.d.B. evaluated histological and immunohistochemical analyses. L.J.Z., B.S.G., G.F.d.W., T.W.B., H.G. and H.M.W.V. designed, coordinated, and analyzed data from the DRUP study. General scientific coordination by J.v.d.H. L.F.A.W., F.K., N.F.C.C.d.M. and E.E.V. supervised the study, with an advisory role of T.N.S. The manuscript was written by N.L.d.V., J.v.d.H. and V.V. in collaboration with all co-authors. All authors commented on and approved the manuscript.

COMPETING INTERESTS

The authors have no competing interests to declare.

REFERENCES

1 Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D. & Perucho, M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363, 558-561, doi:10.1038/363558a0 (1993).

2 Germano, G. *et al.* Inactivation of DNA repair triggers neoantigen generation and impairs tumour growth. *Nature* 552, 116-120, doi:10.1038/nature24673 (2017).

3 Zaretsky, J. M. *et al.* Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *N Engl J Med* 375, 819-829, doi:10.1056/NEJMoa1604958 (2016).

4 Gettinger, S. *et al.* Impaired HLA Class I Antigen Processing and Presentation as a Mechanism of Acquired Resistance to Immune Checkpoint Inhibitors in Lung Cancer. *Cancer Discov* 7, 1420-1435, doi:10.1158/2159-8290.Cd-17-0593 (2017).

5 Sade-Feldman, M. *et al.* Resistance to checkpoint blockade therapy through inactivation of antigen presentation. *Nat Commun* 8, 1136, doi:10.1038/s41467-017-01062-w (2017).

6 Le, D. T. *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409-413, doi:10.1126/science.aan6733 (2017).

7 Overman, M. J. *et al.* Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *The Lancet. Oncology* 18, 1182-1191, doi:10.1016/s1470-2045(17)30422-9 (2017).

8 Overman, M. J. *et al.* Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 36, 773-779, doi:10.1200/jco.2017.76.9901 (2018).

9 Chalabi, M. *et al.* Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat Med* 26, 566-576, doi:10.1038/s41591-020-0805-8 (2020).

10 Dolcetti, R. *et al.* High prevalence of activated intraepithelial cytotoxic T lymphocytes and increased neoplastic cell apoptosis in colorectal carcinomas with microsatellite instability. *Am J Pathol* 154, 1805-1813, doi:10.1016/s0002-9440(10)65436-3 (1999).

11 Tumeh, P. C. *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 515, 568-571, doi:10.1038/nature13954 (2014).

12 Taube, J. M. *et al.* Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res* 20, 5064-5074, doi:10.1158/1078-0432.Ccr-13-3271 (2014).

13 Bicknell, D. C., Kaklamanis, L., Hampson, R., Bodmer, W. F. & Karran, P. Selection for beta 2-microglobulin mutation in mismatch repair-defective colorectal carcinomas. *Current biology : CB* 6, 1695-1697, doi:10.1016/s0960-9822(02)70795-1 (1996).

14 Kloor, M. *et al.* Immunoselective pressure and human leukocyte antigen class I antigen machinery defects in microsatellite unstable colorectal cancers. *Cancer Res* 65, 6418-6424, doi:10.1158/0008-5472.can-05-0044 (2005).

15 Dierssen, J. W. *et al.* HNPCC versus sporadic microsatellite-unstable colon cancers follow different routes toward loss of HLA class I expression. *BMC Cancer* 7, 33, doi:10.1186/1471-2407-7-33 (2007).

16 Ijsselsteijn, M. E. *et al.* Revisiting immune escape in colorectal cancer in the era of immunotherapy. *Br J Cancer* 120, 815-818, doi:10.1038/s41416-019-0421-x (2019).

17 Hughes, E. A., Hammond, C. & Cresswell, P. Misfolded major histocompatibility complex class I heavy chains are translocated into the cytoplasm and degraded by the proteasome. *Proc Natl Acad Sci U S A* 94, 1896-1901, doi:10.1073/pnas.94.5.1896 (1997).

18 Middha, S. *et al.* Majority of B2M-Mutant and -Deficient Colorectal Carcinomas Achieve Clinical Benefit From Immune Checkpoint Inhibitor Therapy and Are Microsatellite Instability-High. *JCO precision oncology* 3, doi:10.1200/po.18.00321 (2019).

19 Groh, V. *et al.* Human lymphocytes bearing T cell receptor gamma/delta are phenotypically diverse and evenly distributed throughout the lymphoid system. *J Exp Med* 169, 1277-1294, doi:10.1084/jem.169.4.1277 (1989).

20 Silva-Santos, B., Serre, K. & Norell, H. γδ T cells in cancer. *Nat Rev Immunol* 15, 683-691, doi:10.1038/nri3904 (2015).

21 Halary, F. *et al.* Control of self-reactive cytotoxic T lymphocytes expressing gamma delta T cell receptors by natural killer inhibitory receptors. *Eur J Immunol* 27, 2812-2821, doi:10.1002/eji.1830271111 (1997).

22 de Vries, N. L. *et al.* High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of antitumour immunity. *Gut* 69, 691-703, doi:10.1136/gutjnl-2019-318672 (2020).

23 van der Velden, D. L. *et al.* The Drug Rediscovery protocol facilitates the expanded use of existing anticancer drugs. *Nature* 574, 127-131, doi:10.1038/s41586-019-1600-x (2019).

24 Danaher, P. *et al.* Gene expression markers of Tumor Infiltrating Leukocytes. *J Immunother Cancer* 5, 18, doi:10.1186/s40425-017-0215-8 (2017).

25 Duhen, T. *et al.* Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat Commun* 9, 2724, doi:10.1038/s41467-018-05072-0 (2018).

26 Kwon, M. *et al.* Determinants of Response and Intrinsic Resistance to PD-1 Blockade in Microsatellite Instability-High Gastric Cancer. *Cancer Discov*, doi:10.1158/2159-8290.Cd-21-0219 (2021).

27 Wu, P. *et al.* γδT17 cells promote the accumulation and expansion of myeloid-derived suppressor cells in human colorectal cancer. *Immunity* 40, 785-800, doi:10.1016/j.immuni.2014.03.013 (2014).

28 Lo Presti, E., Dieli, F. & Meraviglia, S. Tumor-Infiltrating γδ T Lymphocytes: Pathogenic Role, Clinical Significance, and Differential Programing in the Tumor Microenvironment. *Front Immunol* 5, 607, doi:10.3389/fimmu.2014.00607 (2014).

29 Maeurer, M. J. *et al.* Human intestinal Vdelta1+ lymphocytes recognize tumor cells of epithelial origin. *J Exp Med* 183, 1681-1696, doi:10.1084/jem.183.4.1681 (1996).

30 Siegers, G. M., Ribot, E. J., Keating, A. & Foster, P. J. Extensive expansion of primary human gamma delta T cells generates cytotoxic effector memory cells that can be labeled with Feraheme for cellular MRI. *Cancer Immunol Immunother* 62, 571-583, doi:10.1007/s00262-012-1353-y (2013).

31 Wu, D. *et al.* Ex vivo expanded human circulating Vδ1 γδT cells exhibit favorable therapeutic potential for colon cancer. *Oncoimmunology* 4, e992749, doi:10.4161/2162402x.2014.992749 (2015).

32 Almeida, A. R. *et al.* Delta One T Cells for Immunotherapy of Chronic Lymphocytic Leukemia: Clinical-Grade Expansion/Differentiation and Preclinical Proof of Concept. *Clin Cancer Res* 22, 5795-5804, doi:10.1158/1078-0432.Ccr-16-0597 (2016).

33 Mikulak, J. *et al.* NKp46-expressing human gut-resident intraepithelial Vδ1 T cell subpopulation exhibits high antitumor activity against colorectal cancer. *JCI insight* 4, doi:10.1172/jci.insight.125884 (2019).

34 van der Leun, A. M., Thommen, D. S. & Schumacher, T. N. CD8(+) T cell states in human cancer: insights from single-cell analysis. *Nat Rev Cancer* 20, 218-232, doi:10.1038/s41568-019-0235-4 (2020).

35 Groh, V., Steinle, A., Bauer, S. & Spies, T. Recognition of stress-induced MHC molecules by intestinal epithelial gammadelta T cells. *Science* 279, 1737-1740, doi:10.1126/science.279.5357.1737 (1998).

36 Groh, V. *et al.* Broad tumor-associated expression and recognition by tumor-derived gamma delta T cells of MICA and MICB. *Proc Natl Acad Sci U S A* 96, 6879-6884, doi:10.1073/pnas.96.12.6879 (1999).

37 Poggi, A. *et al.* Vdelta1 T lymphocytes from B-CLL patients recognize ULBP3 expressed on leukemic B cells and up-regulated by trans-retinoic acid. *Cancer Res* 64, 9172-9179, doi:10.1158/0008-5472.Can-04-2417 (2004).

38 Hause, R. J., Pritchard, C. C., Shendure, J. & Salipante, S. J. Classification and characterization of microsatellite instability across 18 cancer types. *Nat Med* 22, 1342-1350, doi:10.1038/nm.4191 (2016).

39 Cader, F. Z. *et al.* A peripheral immune signature of responsiveness to PD-1 blockade in patients with classical Hodgkin lymphoma. *Nat Med* 26, 1468-1479, doi:10.1038/s41591-020-1006-1 (2020).

40 Germano, G. *et al.* CD4 T cell dependent rejection of beta 2 microglobulin null mismatch repair deficient tumors. *Cancer Discov*, doi:10.1158/2159-8290.Cd-20-0987 (2021).

41 Carter, S. L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol* 30, 413-421, doi:10.1038/nbt.2203 (2012).

42 Taylor, A. M. *et al.* Genomic and Functional Approaches to Understanding Cancer Aneuploidy. *Cancer Cell* 33, 676-689 e673, doi:10.1016/j.ccell.2018.03.007 (2018).

43 Thorsson, V. *et al.* The Immune Landscape of Cancer. *Immunity* 51, 411-412, doi:10.1016/j.immuni.2019.08.004 (2019).

44 Priestley, P. *et al.* Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* 575, 210-216, doi:10.1038/s41586-019-1689-y (2019).

45 Huang, M. N. *et al.* MSIseq: Software for Assessing Microsatellite Instability from Catalogs of Somatic Mutations. *Sci Rep* 5, 13321, doi:10.1038/srep13321 (2015).

46 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* 29, 15-21, doi:10.1093/bioinformatics/bts635 (2013).

47 Cameron, D. L. *et al.* GRIDSS, PURPLE, LINX: Unscrambling the tumor genome via integrated analysis of structural variation and copy number. *bioRxiv*, 781013, doi:10.1101/781013 (2019).

48 Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* 29, 24-26, doi:10.1038/nbt.1754 (2011).

49 Virtanen, P. *et al.* SciPy 1.0: fundamental

algorithms for scientific computing in Python. *Nat Methods* 17, 261-272, doi:10.1038/s41592-019-0686-2 (2020).

50 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* 26, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

51 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43, e47, doi:10.1093/nar/gkv007 (2015).

52 Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* 15, R29, doi:10.1186/gb-2014-15-2-r29 (2014).

53 Hall, G. *et al.* Immunohistochemistry for PMS2 and MSH6 alone can replace a four antibody panel for mismatch repair deficiency screening in colorectal adenocarcinoma. *Pathology* 42, 409-413, doi:10.3109/00313025.2010.493871 (2010).

54 Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* 14, 865-868, doi:10.1038/nmeth.4380 (2017).

55 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888-1902 e1821, doi:10.1016/j.cell.2019.05.031 (2019).

56 McGinnis, C. S. *et al.* MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat Methods* 16, 619-626, doi:10.1038/s41592-019-0433-8 (2019).

57 Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* 36, 421-427, doi:10.1038/nbt.4091 (2018).

58 McInnes, L., Healy, J. & Melville, J. J. a. p. a. Umap: Uniform manifold approximation and projection for dimension reduction. (2018).

59 Ijsselsteijn, M. E., van der Breggen, R., Farina Sarasqueta, A., Koning, F. & de Miranda, N. F. C. C. A 40-Marker Panel for High Dimensional Characterization of Cancer Immune Microenvironments by Imaging Mass Cytometry. *Frontiers in immunology* 10, 2534-2534, doi:10.3389/fimmu.2019.02534 (2019).

60 Berg, S. *et al.* ilastik: interactive machine learning for (bio)image analysis. *Nat Methods* 16, 1226-1232, doi:10.1038/s41592-019-0582-9 (2019).

61 Ijsselsteijn, M. E., Somarakis, A., Lelieveldt, B. P. F., Höllt, T. & de Miranda, N. Semi-automated background removal limits data loss and normalises imaging mass cytometry data. *Cytometry A*, doi:10.1002/cyto.a.24480 (2021).

62 Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* 7, R100, doi:10.1186/gb-2006-7-10-r100 (2006).

63 Somarakis, A., Van Unen, V., Koning, F., Lelieveldt, B. P. F. & Hollt, T. ImaCytE: Visual Exploration of Cellular Microenvironments for Imaging Mass Cytometry Data. *IEEE transactions on visualization and computer graphics*, 10.1109/TVCG.2019.2931299, doi:10.1109/TVCG.2019.2931299 (2019).

64 van der Maaten, L. J. P. & Hinton, G. E. Visualizing high-dimensional data using t-SNE. *J. Mach. Learn. Res.* 9, 2579-2605 (2008).

65 Höllt, T. *et al.* Cytosplore: Interactive Immune Cell Phenotyping for Large Single-Cell Datasets. 35, 171-180, doi:https://doi.org/10.1111/cgf.12893 (2016).

66 Dijkstra, K. K. *et al.* Generation of Tumor-Reactive T Cells by Co-culture of Peripheral Blood Lymphocytes and Tumor Organoids. *Cell* 174, 1586-1598.e1512, doi:10.1016/j.cell.2018.07.009 (2018).

67 Cattaneo, C. M. *et al.* Tumor organoid-T-cell coculture systems. *Nat Protoc* 15, 15-39, doi:10.1038/s41596-019-0232-9 (2020).

68 Dutta, I., Postovit, L. M. & Siegers, G. M. Apoptosis Induced via Gamma Delta T Cell Antigen Receptor "Blocking" Antibodies: A Cautionary Tale. *Front Immunol* 8, 776, doi:10.3389/fimmu.2017.00776 (2017).

SUPPLEMENTAL FIGURES

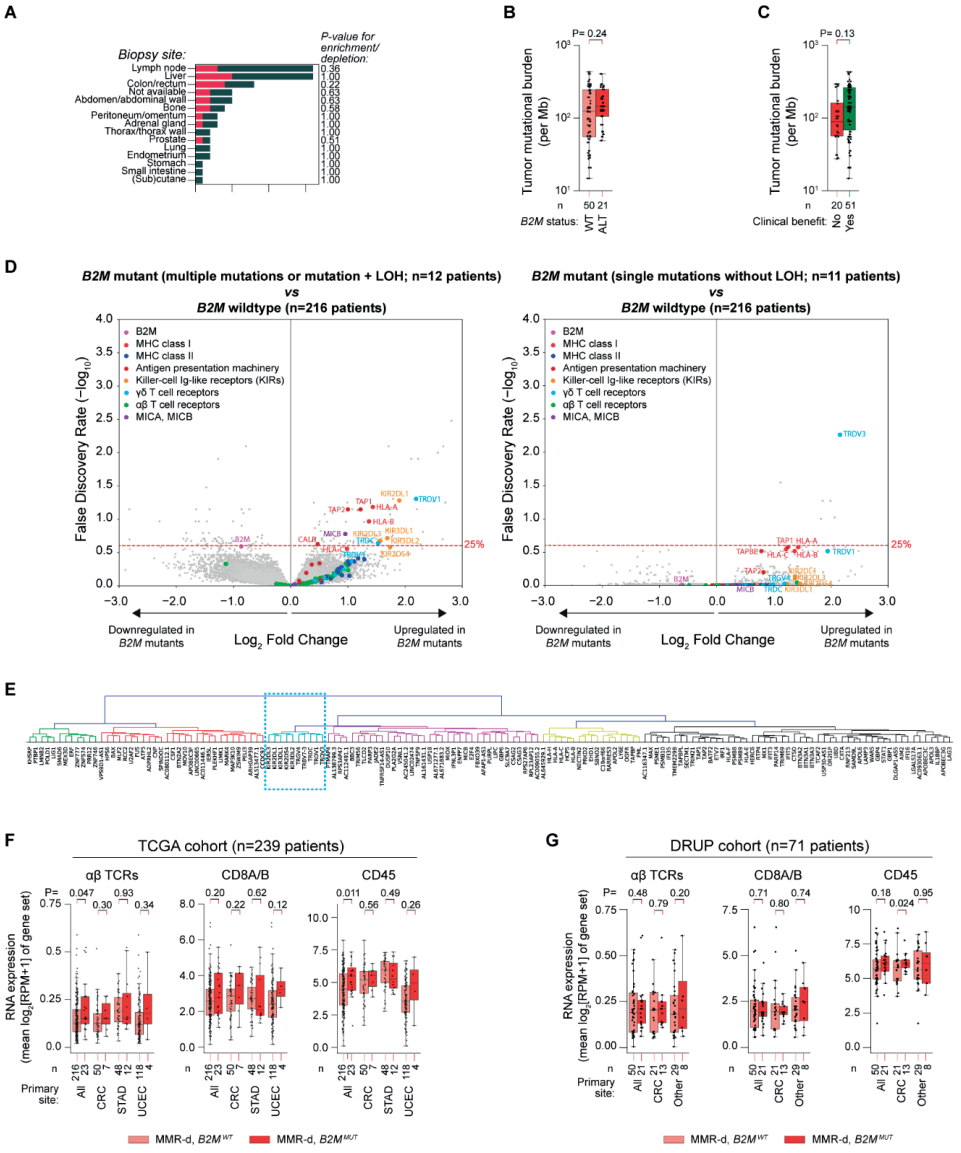


Figure S1. Association of *B2M* status to clinical, genomic, and transcriptomic characteristics of MMR-d tumors.

A. The number of patients per biopsy location for 71 patients in the DRUP with MMR-d cancers and available outcome of ICB therapy. Colors denote patients' *B2M* status (WT: wildtype, gray; ALT: altered, red). Fisher's exact test-based P-values for enrichment/depletion of *B2M* altered cases per primary site are shown. **B**. The tumor mutational burden per Mb *vs B2M* status. Wilcoxon rank sum test-based P-value is shown. **C.** The tumor mutational burden per Mb *vs* clinical benefit of immune checkpoint blockade (ICB). Wilcoxon rank sum test-based P-value is shown. **D.** Two volcano plots indicating differential gene expression between MMR-d cancers with wildtype *B2M vs* MMR-d cancers with multiple *B2M* mutations and/or *B2M* mutations + loss of heterozygosity (LOH; left plot), or *vs* MMR-d cancers with single *B2M* mutations without LOH (right plot). The Benjamini Hochberg false discovery rate (FDR) significance

threshold of 25% is indicated by the red dashed line. Results were obtained in a combined analysis of all MMR-d cancers of the TCGA COAD (colon adenocarcinoma; n=57 patients), STAD (stomach adenocarcinoma; n=60 patients) and UCEC (uterus corpus endometrial carcinoma; n=122 patients) cohorts, and were adjusted for tumor type. **E.** Dendrogram representing the hierarchical clustering result of gene expression profiles across all MMR-d cancers in the TCGA COAD (n=57 patients), STAD (n=60 patients) and UCEC (n=122 patients) cohorts. The genes included are those significantly (FDR <25%) upregulated in MMR-d *B2M*^{MUT} *vs* MMR-d *B2M*^{WT} cancers, after adjustment for tumor type. The blue dashed rectangle denotes the Vd1/3 T cell cluster. **F.** The RNA expression of different immune marker gene sets MMR-d *B2M*^{WT} (pink), and MMR-d *B2M*^{MUT} (red) cancers. Results are obtained with the TCGA COAD, STAD and UCEC cohorts, and are shown for all cohorts combined (All), and for each cohort separately. Boxes, whiskers, and dots indicate quartiles, 1.5 interquartile ranges, and individual data points, respectively. Wilcoxon rank sum test-based P-values are shown for MMR-d *B2M*^{WT} *vs* MMR-d *B2M*^{MUT} cancers. **G.** As **F**, but for MMR-d cancers in the DRUP cohort. Results are shown for all cancers combined (All), only colorectal cancer (CRC), or all non-CRC cancers (Other).

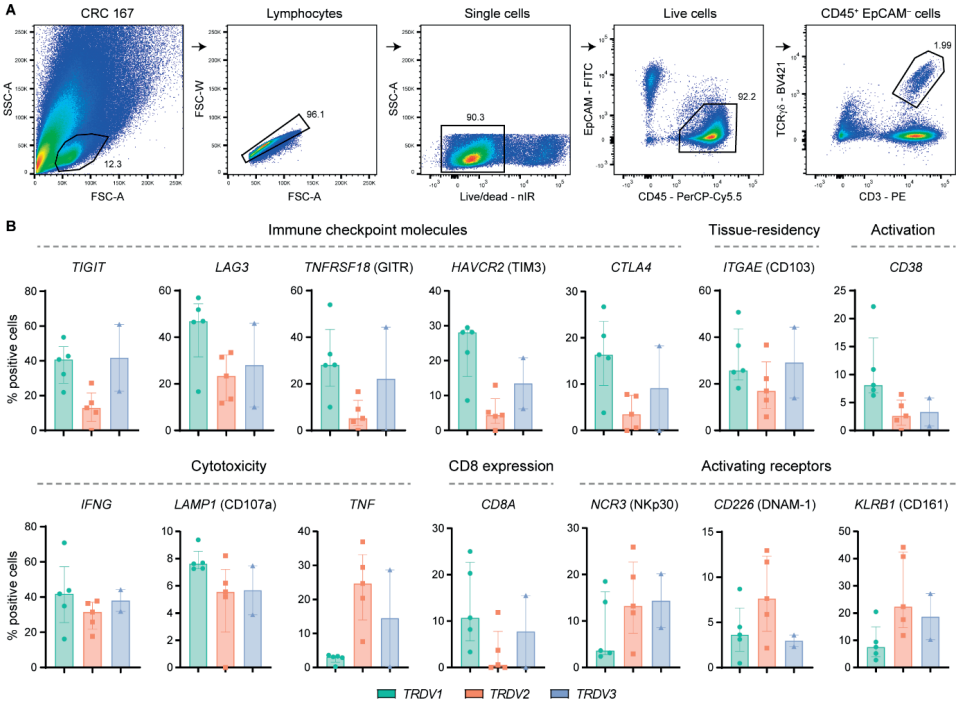


Figure S2. Characterization of γδ T cells from MMR-d colon cancers by single-cell RNA-sequencing.
A. FACS gating strategy for single, live CD45⁺ EpCAM⁻ CD3⁺ TCRγδ⁺ cells of a representative MMR-d colon cancer sample showing sequential gates with percentages. **B**. Frequencies of positive cells for selected genes across Vδ1 (n=1927), Vδ2 (n=860), and Vδ3 (n=506) cells as percentage of total γδ T cells from each MMR-d colon tumor (n=5) analyzed by single-cell RNA-sequencing. Vδ3 cells were present in two out of five colon cancers. Bars indicate median ± IQR. Each dot represents an individual sample.

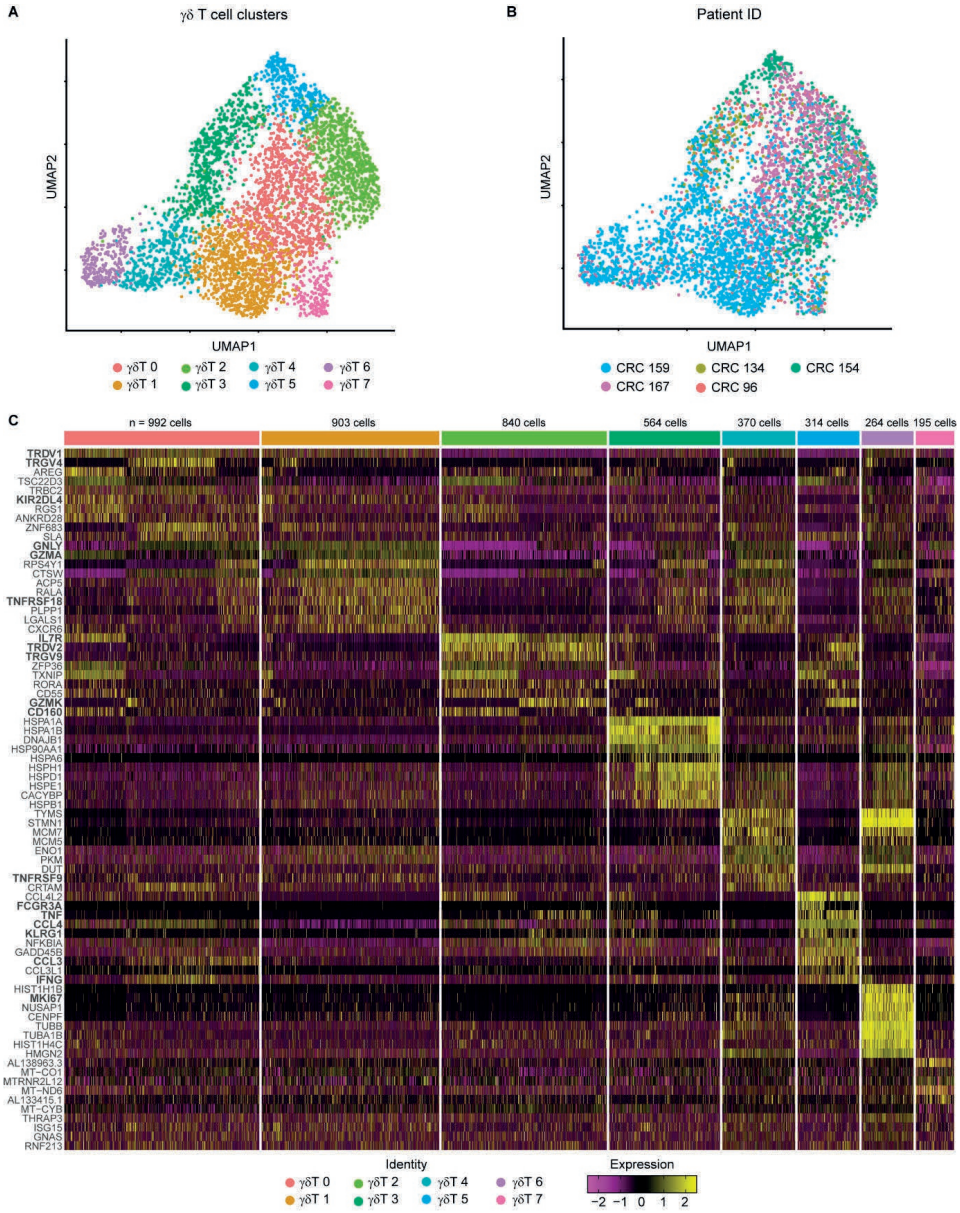


Figure S3. Distinct clusters of γδ T cells from MMR-d colon cancers by single-cell RNA-sequencing.
A. UMAP embedding showing γδ T cells (n=4442) isolated from MMR-d colon cancers (n=5) analyzed by single-cell RNA-sequencing. Colors represent the functionally different γδ T cell clusters identified by graph-based clustering and non-linear dimensional reduction. Each dot represents a single cell. **B**. UMAP embedding of (**A**) colored by patient ID. Each dot represents a single cell. **C**. Heatmap showing the normalized single-cell gene expression value (z-score, purple-to-yellow scale) for the top 10 differentially expressed genes in each identified γδ T cell cluster.

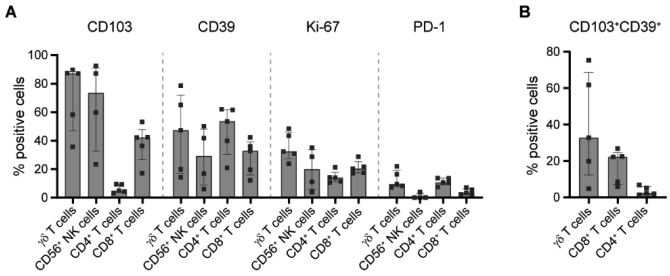
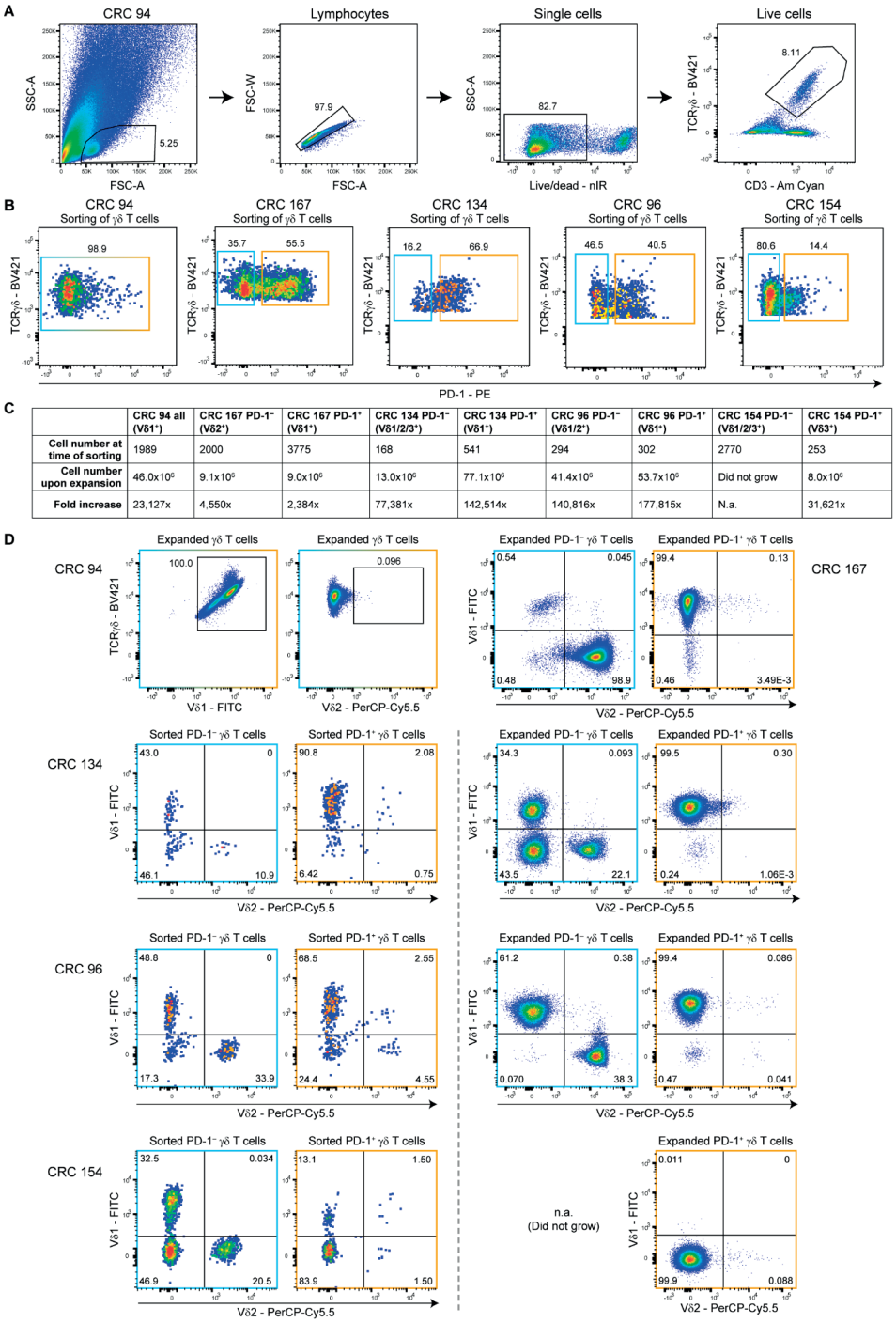


Figure S4. Phenotype of immune cell populations in β2m-positive and -negative MMR-d colon cancers.
A. Frequencies of marker-positive γδ T cells, $CD56^+$ NK cells, $CD4^+$ T cells, and $CD8^+$ T cells in treatment-naïve $\beta2m^-$ (n=5) MMR-d colon cancers. $CD56^+$ NK cells were present in four out of five $\beta2m^-$ cancer samples. Bars indicate median ± IQR. Each dot represents an individual sample. **B.** Frequencies of $CD103^+CD39^+$ γδ T cells, $CD8^+$ T cells, and $CD4^+$ T cells in treatment-naïve $\beta2m^-$ (n=5) MMR-d colon cancers. Bars indicate median ± IQR. Each dot represents an individual sample.



	CRC 94 all (Vδ1⁺)	CRC 167 PD-1⁻ (Vδ2⁺)	CRC 167 PD-1⁺ (Vδ1⁺)	CRC 134 PD-1⁻ (Vδ1/2/3⁺)	CRC 134 PD-1⁺ (Vδ1⁺)	CRC 96 PD-1⁻ (Vδ1/2⁺)	CRC 96 PD-1⁺ (Vδ1⁺)	CRC 154 PD-1⁻ (Vδ1/2/3⁺)	CRC 154 PD-1⁺ (Vδ3⁺)
Cell number at time of sorting	1989	2000	3775	168	541	294	302	2770	253
Cell number upon expansion	46.0x10⁶	9.1x10⁶	9.0x10⁶	13.0x10⁶	77.1x10⁶	41.4x10⁶	53.7x10⁶	Did not grow	8.0x10⁶
Fold increase	23,127x	4,550x	2,384x	77,381x	142,514x	140,816x	177,815x	N.a.	31,621x

Figure S5. Sorting of PD-1⁺ and PD-1⁻ γδ T cells from MMR-d colon cancers by FACS and their TCR Vδ chain usage.

A. FACS gating strategy for single, live CD3⁺ TCRγδ⁺ cells of a representative MMR-d colon cancer sample showing

sequential gates with percentages. **B**. Sorting of all γδ T cells from CRC94 (due to the low number of PD-1⁺ cells), and of PD-1⁻ (blue squares) and PD-1⁺ (orange squares) γδ T cells from CRC167, CRC134, CRC96, and CRC154. Each dot is a single cell. **C**. Table showing the number of γδ T cells isolated from colon cancers at the time of sorting *versus* 3-4 weeks after expansion, and the fold increase thereof. **D**. TCR Vδ chain usage after expansion of γδ T cells from CRC94 and CRC167 (first row), and at the time of sorting (left panel) as well as after expansion (right panel) of γδ T cells from CRC134, CRC96 and CRC154. From CRC154, the PD-1⁻ γδ T cells did not expand in culture. Each dot is a single cell.

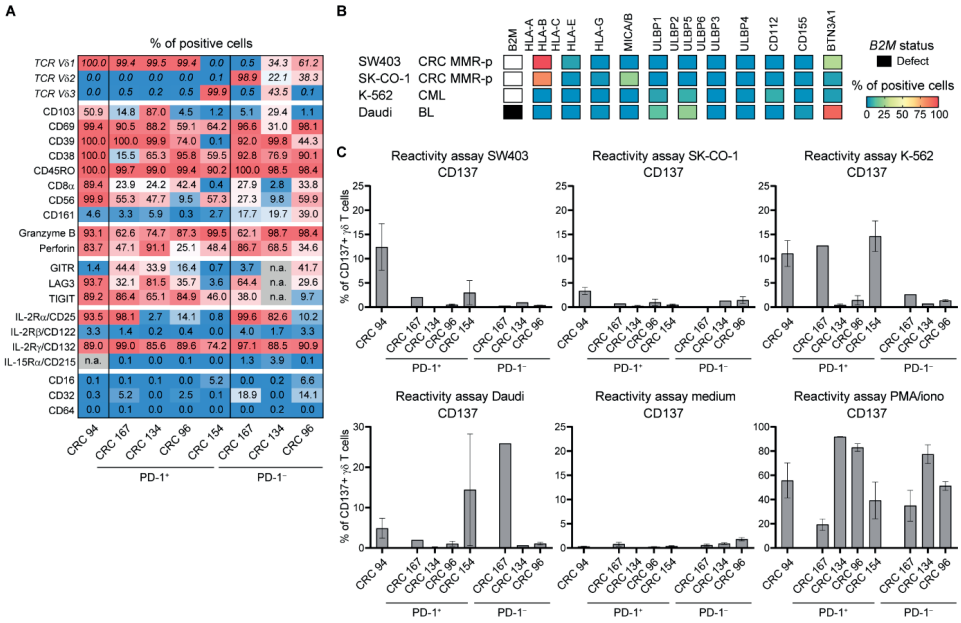


Figure S6. Phenotype and reactivity of γδ T cells from MMR-d colon cancers towards cancer cell lines.
A. Table showing the percentage of positive cells for different TCR Vδ chains (as in **Figure 3A**), tissue-residency/activation markers, cytotoxic molecules, immune checkpoint molecules, cytokine receptors, and Fc receptors on expanded PD-1⁺ and PD-1⁻ γδ T cells from MMR-d colon cancers (n=5) as percentage of total γδ T cells. **B**. Heatmap showing the *B2M* mutational status and surface expression of HLA class I, NKG2D ligands, DNAM-1 ligands, and butyrophilin on SW403, SK-CO-1, K-562, and Daudi cells. **C**. Bar plots showing the percentage of CD137-positive γδ T cells after 18h co-culture of PD-1⁺ and PD-1⁻ γδ T cells from MMR-d colon cancers (n=5) with SW403, SK-CO-1, K-562, and Daudi cells. Medium was used as negative control and PMA/ionomycin as positive control. Bars indicate mean ± SEM. Data from two independent experiments (CRC94, CRC134, CRC154, CRC96), depending on availability of γδ T cells.

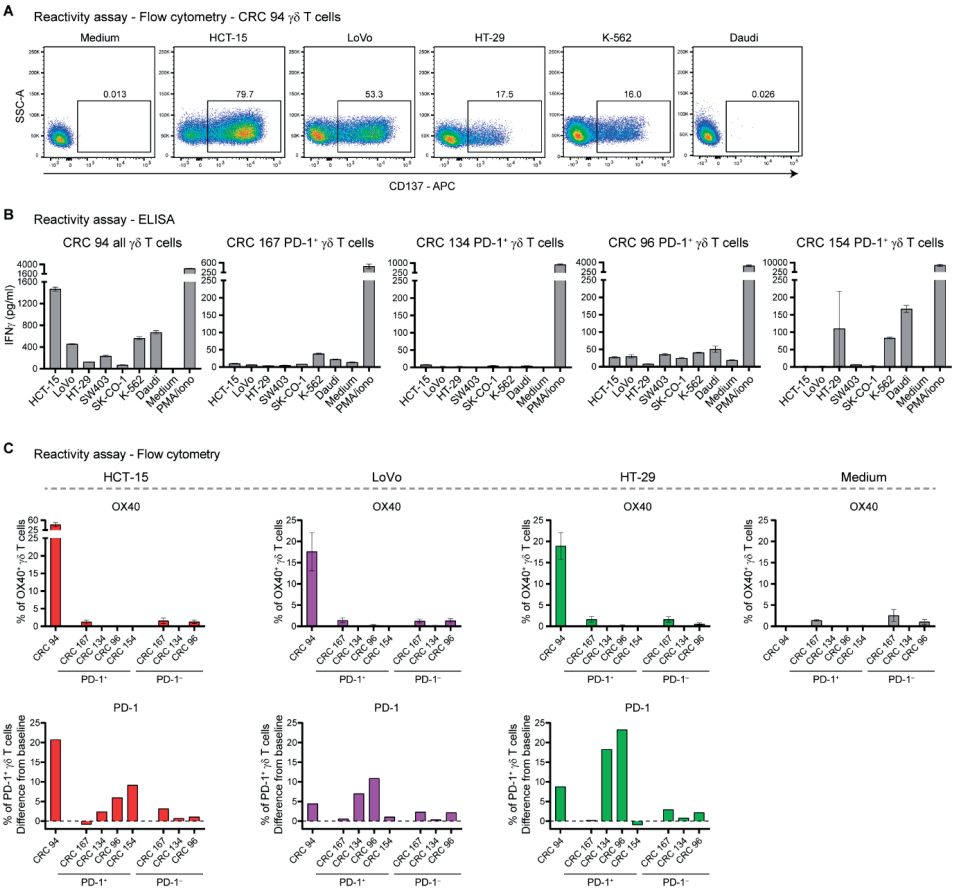


Figure S7. Reactivity of γδ T cells from MMR-d colon cancers towards cancer cell lines.
A. Representative flow cytometry plots showing the expression of CD137 on γδ T cells from CRC94 after 18h co-culture with HCT-15, LoVo, HT-29, K-562, and Daudi cells as compared to medium only. Gates indicate percentage of positive γδ T cells. **B**. Bar plots showing the presence of IFNg in the supernatant after 18h co-culture of PD-1$^+$ γδ T cells with the cancer cell lines. Medium as negative control and PMA/ionomycin as positive control are included. Bars indicate mean ± SEM of triplicates. **C**. Bar plots showing the percentage of OX40-positive (first row) and PD-1-positive (second row) γδ T cells after 18h co-culture of PD-1$^+$ and PD-1$^-$ γδ T cells from MMR-d colon cancers (n=5) with HCT-15, LoVo, and HT-29 cells. PD-1 expression is shown as difference from baseline (medium) condition. Bars indicate mean ± SEM. Data from four (CRC94), three (CRC167, CRC96), or two (CRC134, CRC154) independent experiments for OX40, depending on availability of γδ T cells.

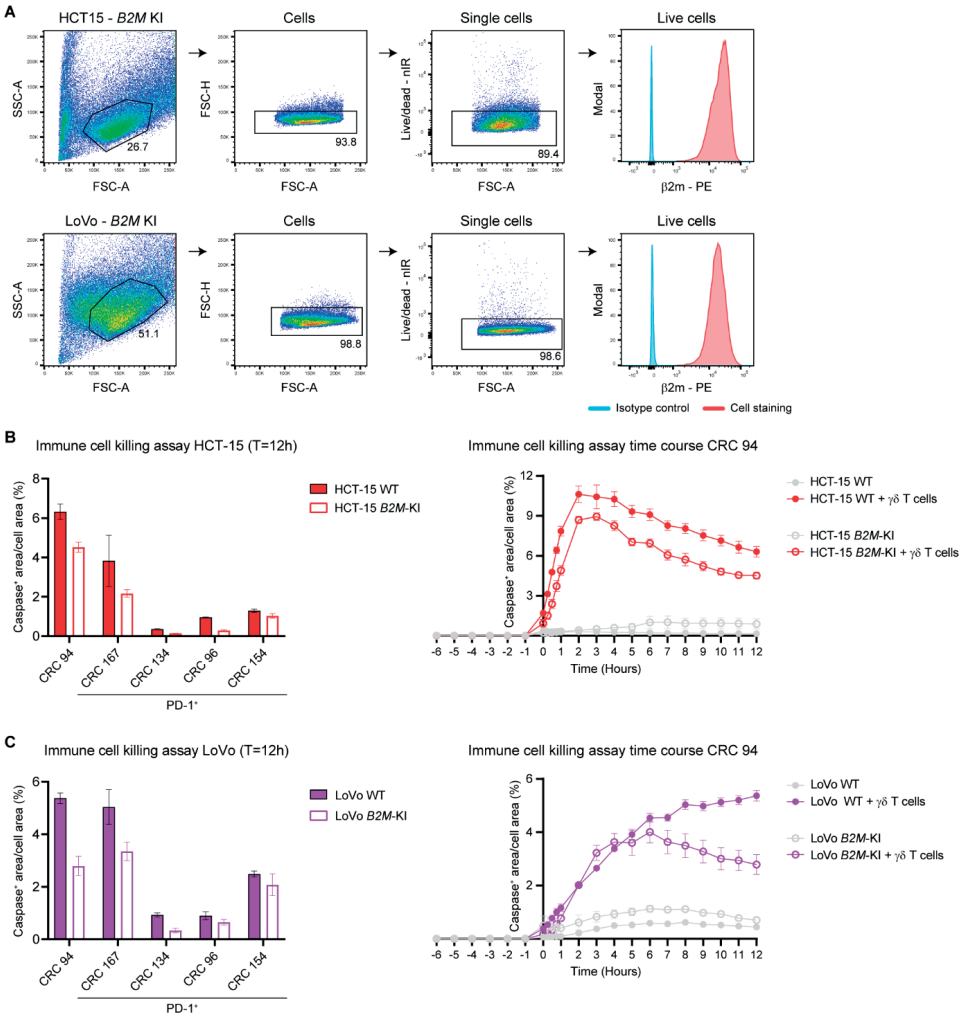


Figure S8. Characterization of *B2M*-knockin cancer cell lines and reactivity of γδ T cells towards *B2M*-knockin vs -wildtype cancer cell lines.

A. Flow cytometry gating strategy to validate β2m expression on HCT-15 and LoVo *B2M*-knockin (*B2M*-KI) cell lines. Isotype controls were included as negative control. **B**. Bar plots showing the quantification of the killing of HCT-15 *B2M*-KI *vs* wildtype (WT) cells by γδ T cells from MMR-d colon cancers (n=5) in the presence of a red fluorescent caspase-3/7 reagent after 12h co-culture. Bars indicate mean ± SEM of two wells with two images/well. Right panel shows representative time course of apoptosis (caspase-3/7) in the presence or absence of γδ T cells derived from CRC94. **C**. As **B**, but for LoVo *B2M*-KI *vs* WT cells.

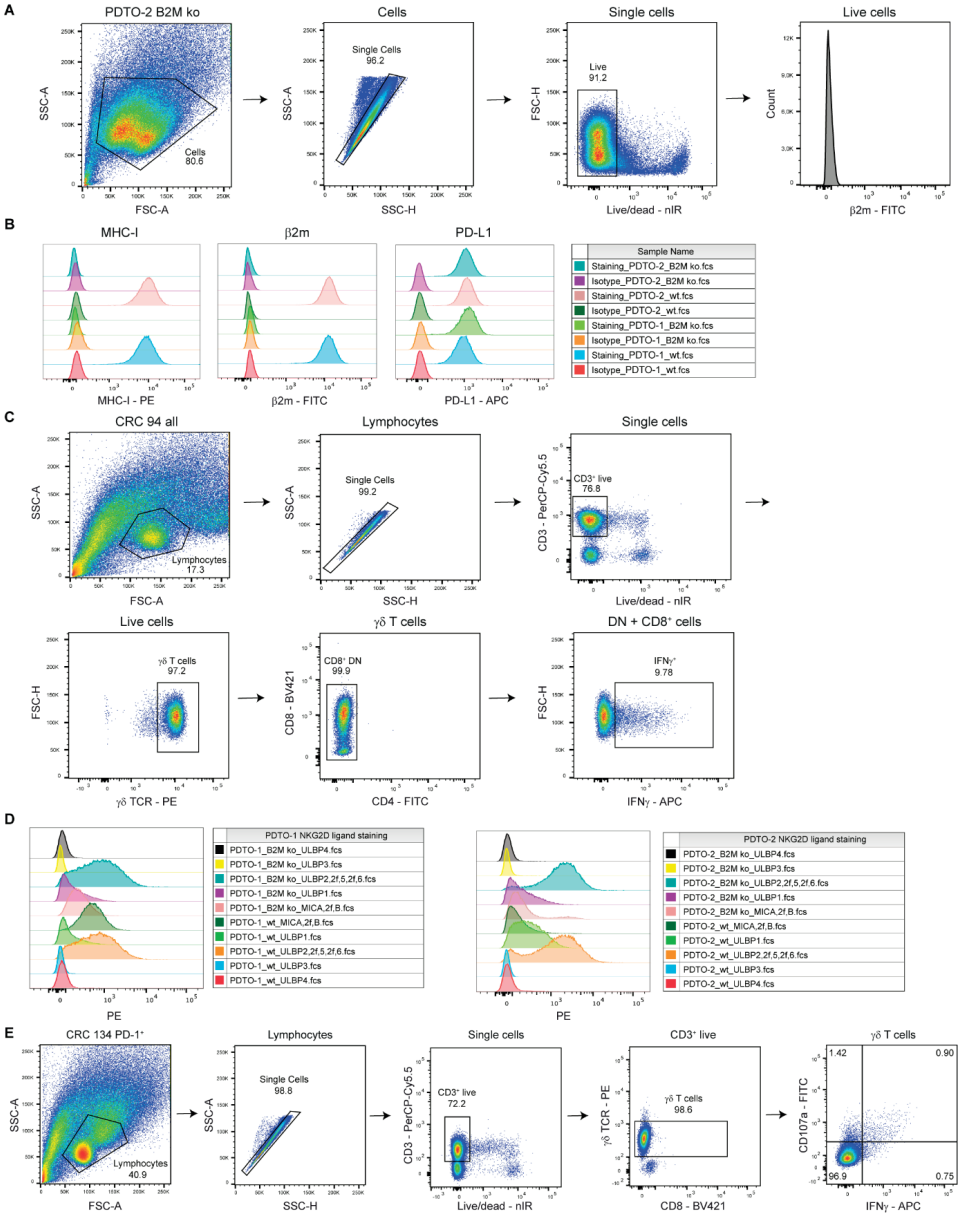


Figure S9. Reactivity of γδ T cells from MMR-d colon cancers towards cancer cell lines.
A. Flow cytometry gating strategy on PDTO cells for analysis of surface staining. Selected cells were gated on single, live cells before quantification of staining signal. **B**. Histogram representation and count for surface staining of MHC-I, PD-L1, and β2m expression on two PDTO lines *B2M*^{WT} and *B2M*^{KO} after IFN-γ pre-stimulation. Staining with isotype antibodies for each fluorochrome (PE, APC and FITC) were included as negative control. **C**. Flow cytometry gating strategy on γδ T cell samples for analysis of intracellular staining to test anti-tumor reactivity upon PDTO stimulation. Lymphocyte population was further gated on single cells, live and CD3^{+} cells, γδ TCR^{+} cells and CD8^{+} as well as CD8^{-} CD4^{-} cells. Reactivity of the sample was based on IFN-γ^{+} cells of the selected population. **D**. Histogram representation and count for surface staining of NKG2D ligands MICA/B, ULBP1, ULBP2/5/6, ULBP3, and ULBP4 on two PDTO lines *B2M*^{WT} and *B2M*^{KO} after IFN-γ pre-stimulation. **E**. Flow cytometry gating strategy on γδ T cell samples for analysis of intracellular staining after stimulation with PDTOs in the presence of NKG2D ligand blocking. Lymphocyte population

was further gated on single cells, live and CD3⁺ cells, followed by γδ TCR⁺ and CD8⁺ as well as CD8⁻ cells. Reactivity of final population was based on IFN-γ⁺ or CD107a⁺ cells.

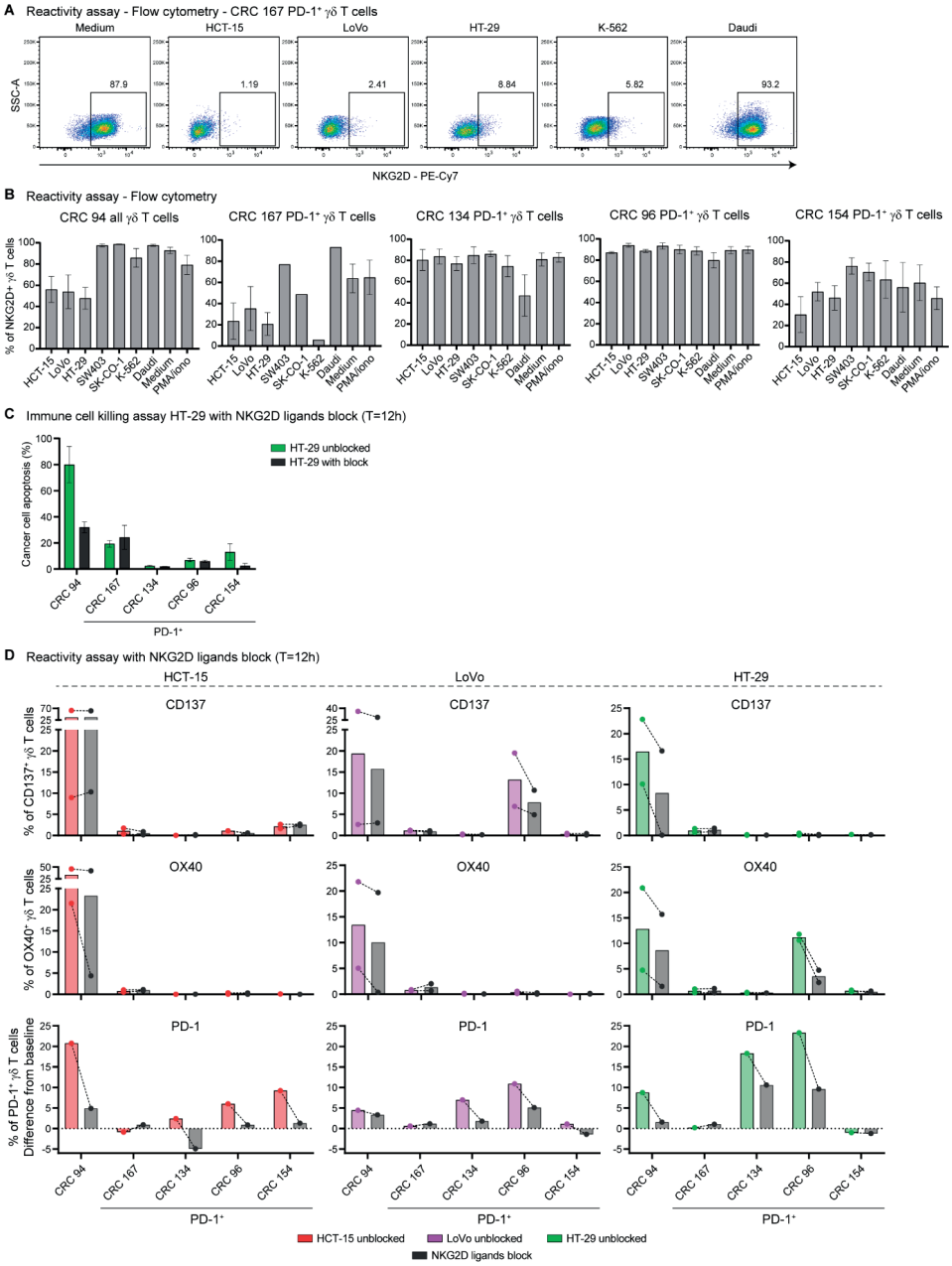


Figure S10. Reactivity towards and killing of cancer cell lines by PD-1⁺ γδ T cells from MMR-d colon cancers in the presence of NKG2D ligand blocking.

A. Representative flow cytometry plots showing the expression of NKG2D on PD-1⁺ γδ T cells from CRC167 after 18h co-culture with HCT-15, LoVo, HT-29, K-562, and Daudi cells as compared to medium only. Gates indicate percentage of positive γδ T cells. **B**. Bar plots showing the expression of NKG2D on PD-1⁺ γδ T cells from MMR-d colon cancers

(n=5) after 18h co-culture of PD-1⁺ γδ T cells with the cancer cell lines. Medium as negative control and PMA/ionomycin as positive control are included. Bars indicate mean ± SEM. Data from four (CRC94), three (CRC167, CRC96), or two (CRC134, CRC154) independent experiments, depending on availability of γδ T cells. **C**. Bar plots showing the quantification of killing of HT-29 cells by γδ d T cells from MMR-d colon cancers (n=5) in the presence of blocking antibodies for NKG2D ligands as compared to the unblocked condition after 12h co-culture. Bars indicate mean ± SEM of two wells with two images/well. **D**. Bar plots showing the percentage of CD137-positive (first row), OX40-positive (second row), and PD-1-positive (third row) γδ T cells after 18h co-culture of PD-1⁺ γδ T cells from MMR-d colon cancers (n=5) with HCT-15, LoVo, and HT-29 cells in the presence of blocking antibodies for NKG2D ligands. PD-1 expression is shown as difference from baseline (medium) condition. Bars indicate the mean and lines indicate similar experiments. Data from two independent experiments for CD137 and OX40.

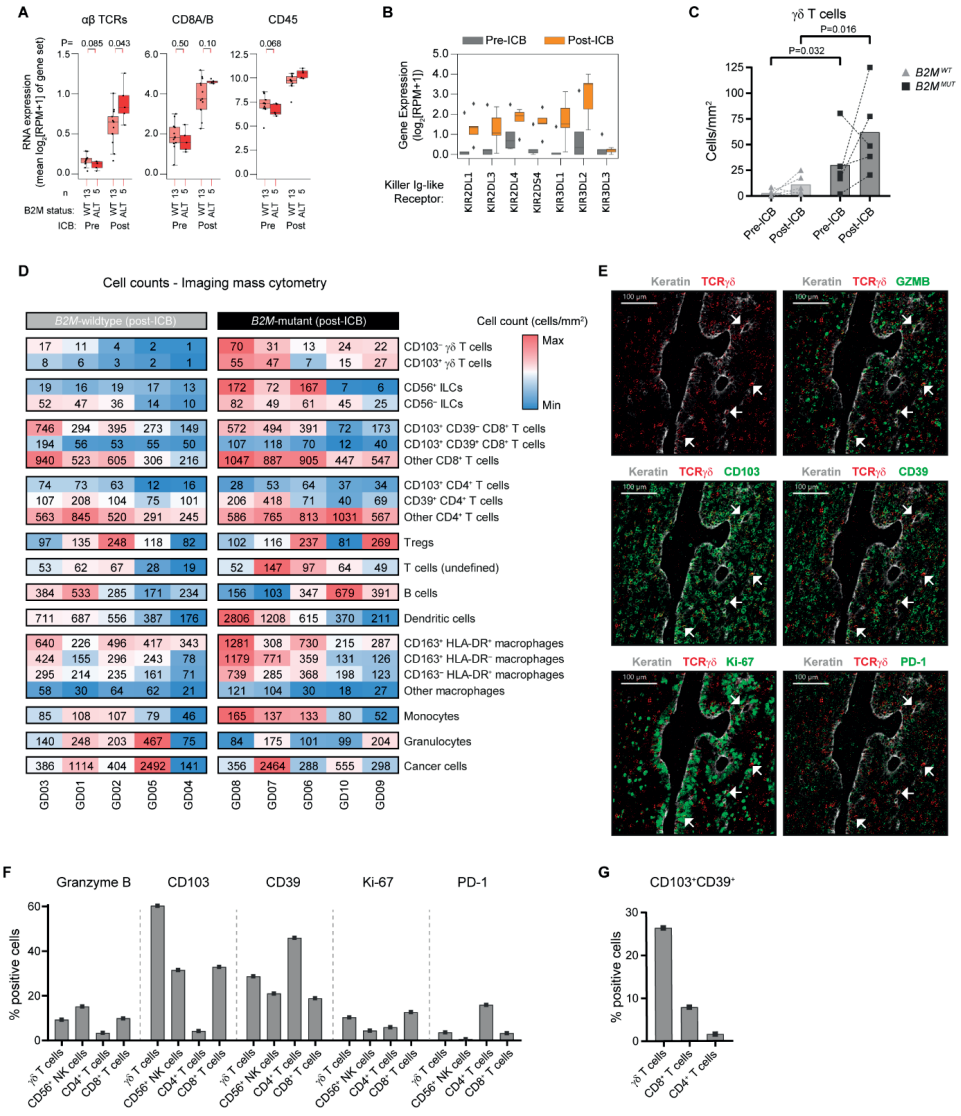


Figure S11. Distribution of immune cell populations in *B2M*-wildtype and *B2M*-mutant colon cancers upon immune checkpoint blockade (ICB) by imaging mass cytometry.

A. The RNA expression of different immune marker gene sets in MMR-d *B2M*^{WT} (pink), and MMR-d *B2M*^{MUT} (red)

cancers, before (left) and after (right) neoadjuvant ICB in the NICHE study. Boxes, whiskers, and dots indicate quartiles, 1.5 interquartile ranges, and individual data points, respectively. Wilcoxon rank sum test-based P-values are shown for MMR-d *B2M*^{WT} *vs* MMR-d *B2M*^{MUT} cancers. **B**. Boxplot showing the pre- (gray) and post-ICB (orange) RNA expression of killer-cell Ig-like receptors (KIRs) in MMR-d cancers in the NICHE study with high impact (inactivating) mutations in *B2M*. Boxes, whiskers, and dots indicate quartiles, 1.5 interquartile ranges, and outliers, respectively. P-values were calculated by Wilcoxon rank sum test. **C**. Frequencies of γδ T cells in *B2M*^{WT} (n=5) and *B2M*^{MUT} (n=5) MMR-d colon cancers pre- and post-ICB treatment corresponding to **Figure 4B** with lines indicating paired samples. Each dot represents an individual sample. P-values were calculated by Wilcoxon rank sum test. **D**. Table showing cell counts (cells/mm^2) of different (immune) cell types from the imaging mass cytometric detection of *B2M*^{WT} (n=5) and *B2M*^{MUT} (n=5) MMR-d colon cancers upon ICB treatment. Hierarchical clustering was performed on the samples within the two groups. Color bar is scaled per major immune lineage. **E**. Representative images of the detection of cytotoxic (granzyme B$^+$), tissue-resident (CD103$^+$), activated (CD39$^+$), proliferating (Ki-67$^+$), and PD-1$^+$ γδ T cells by imaging mass cytometry in a *B2M*^{MUT} MMR-d colon cancer upon ICB treatment. **F**. Frequencies of marker-positive γδ T cells, CD56$^+$ NK cells, CD4$^+$ T cells, and CD8$^+$ T cells in the sole *B2M*^{MUT} MMR-d colon cancer that contained cancer cells upon ICB treatment. **G**. Frequencies of CD103$^+$CD39$^+$ γδ T cells, CD8$^+$ T cells, and CD4$^+$ T cells in the sole *B2M*^{MUT} MMR-d colon cancer that contained cancer cells upon ICB treatment.

SUPPLEMENTAL TABLES AND MOVIES

The supplementary tables and movies are available in the appendix to this thesis at the repository of Leiden University (https://hdl.handle.net/1887/3439882) and can be requested from the author.

Table S1: Characteristics of clinical samples from 71 patients with MMR-d cancers from the DRUP study.
Table S2: Immune marker gene sets for bulk RNA-seq analyses.
Table S3: Characteristics of clinical samples from 17 patients with ICB-naïve MMR-d colon cancer.
Table S4: Characteristics of patient-derived organoids from MMR-d colorectal cancer.
Table S5: Antibodies used for imaging mass cytometry of colon cancers.
Table S6: Antibodies used for immunophenotyping of γδ T cells by flow cytometry.

Movie 1: Killing of HCT-15 cells by γδ T cells (Vδ1$^+$) from a MMR-d colon cancer.
Movie 2: Killing of HCT-15 cells by PD-1$^+$ (Vδ1$^+$) as compared to PD-1$^-$ (Vδ2$^+$) γδ T cells from a MMR-d colon cancer.

5

The ILC compartment in mismatch repair-deficient colorectal cancers is dominated by CD127-negative ILC1-like cells

Natasja L. de Vries[1,2], Li Zheng[1], Jessica Roelands[1], Manon van der Ploeg[1], Marieke E. Ijsselsteijn[1], Ruud van der Breggen[1], Juliette Krop[2], Michael Eikmans[2], Dina Ruano[1], Frits Koning[2†], Noel F.C.C. de Miranda[1†]

[1]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands.
[2]Department of Immunology, Leiden University Medical Center, Leiden, The Netherlands.
[†]Equal responsible authors.

Manuscript in preparation

ABSTRACT

Background

The colorectal cancer (CRC) immune microenvironment is an important determinant of the clinical prognosis of patients. We previously discovered that an unappreciated innate lymphocyte population ($Lin^{-}CD7^{+}CD127^{-}CD103^{+}CD45RO^{+}$) was enriched and displayed hallmarks of cytotoxic activity in mismatch repair (MMR)-deficient CRCs. Little is known about the role of innate lymphoid cells (ILCs) in CRC, and often their analysis has focused on the $CD127^{+}$ compartment. Here, we aimed to characterize the entire $CD7^{+}CD3^{-}CD127^{+/-}$ ILC compartment in MMR-deficient CRCs.

Methods

We performed an unbiased characterization of sorted $CD7^{+}CD3^{-}$ ILCs, independently of CD127 expression, from five patients with MMR-deficient CRC by single-cell RNA-sequencing. In concordance, mass and flow cytometric examination of immunophenotypic markers and transcription factor expression was performed to investigate the full spectrum of ILC subsets and their functional differences.

Results

The majority of ILCs in MMR-deficient CRCs consisted of $CD127^{-}CD103^{+}$ ILC1-like cells, followed by $CD127^{-}CD103^{-}$ NK-like cells, and a minor population of $CD127^{+}CD103^{-}$ conventional ILCs. The ILC1-like cells in CRC tissues specifically expressed CD39, CD45RO, and transcription factor *TBX21* (T-bet). They distinguished from NK-like cells and conventional ILCs by high expression of genes encoding KIRs, co-inhibitory and co-stimulatory molecules, NKG2A, and HLA class II. The ILC1-like cells were the only ILC population displaying hallmarks of proliferation in MMR-deficient CRCs.

Conclusions

Our work provides a comprehensive description of heterogeneity of the ILC population in MMR-deficient CRCs. Validation of our findings in a large CRC dataset confirmed that ILC1-like cells are the dominant ILC subset, which is particularly frequent in MMR-deficient samples. These findings suggest an active involvement of ILCs in the recognition of and cytotoxicity against MMR-deficient cancers, and holds potential for exploiting ILC1-like cells as targets for cancer immunotherapy.

INTRODUCTION

Colorectal cancer (CRC) is distinguished by an extensive and complex tumor microenvironment containing immune cells that impact carcinogenesis and patient prognosis.[1] The advent of immune checkpoint blockade (ICB) targeting the inhibitory PD-1/PD-L1 and/or CTLA-4 axis boosted the anti-tumor potential of cytotoxic T cells, resulting in enhanced immunity via recognition of human leukocyte antigen (HLA) class I-bound (neo)epitopes on cancer cells.[2-4] ICB has provided durable clinical responses in patients diagnosed with mismatch repair (MMR)-deficient CRCs that present with high mutation burden accompanied by an excess of frameshifted neoepitopes.[5-10] However, the large majority of these MMR-deficient CRCs lack expression of HLA class I, due to inactivating mutations in β2-microglobulin (*B2M*), the HLA class I genes themselves, or other components of the antigen processing pathway.[11-14] Hence, these tumors may evade $CD8^+$ T cell-mediated immunity. Interestingly, the majority of HLA class I-negative MMR-deficient CRCs have shown durable responses to ICB,[15] suggesting the involvement of other immune cell subsets than $CD8^+$ T cells.

We previously discovered that an unappreciated ILC1-like population ($Lin^-CD7^+CD127^-CD103^+CD45RO^+$) was enriched in MMR-deficient CRCs and showed high cytotoxic potential.[16] This subset displayed an intraepithelial localization in line with their tissue-resident ($CD103^+CD69^+$) phenotype.[16] ILCs represent a heterogeneous population of innate lymphocytes that reside in mucosal tissues such as the intestine where they have important effector and regulatory functions in infection, inflammation, and tissue remodeling.[17] ILCs lack the expression of rearranged T/B cell receptors and can be activated by cytokines and/or through activating and inhibitory innate immune receptors.[18] They include conventional NK cells as well as ILC1, ILC2, and ILC3 cells that display functional characteristics analogous to the main subsets of helper T cells. ILC1s express the transcription factor *TBX21* (T-bet) and can produce Th1-related cytokines such as IFN-γ, like conventional NK cells. ILC2s are *GATA3*-positive and can secrete Th2-related cytokines (IL-4, IL-5, IL-13). Last, ILC3s express the transcription factor *RORC* (RORγt) and have been described to be involved in IL-17 and IL-22 production.[19] In contrast to conventional NK cells, these ILC subsets display a CD127 (IL-7Rα)-positive phenotype, with the exception of the intraepithelial ILC1 subset,[20] and do not possess cytotoxic activity.[21] Moreover, additional levels of plasticity[22,23] and heterogeneity[24,25] have been described in the ILC classification.

As the intraepithelial ILC1-like ($CD127^-CD103^+CD45RO^+$) population is, in theory, capable of killing cancer cells independently of HLA class I expression, they may be particularly important in mediating anti-tumor immune responses in MMR-deficient CRCs, also in the therapeutic setting with ICB therapy. Here, we performed single cell-level analysis on the entire $CD7^+CD3^-CD127^{+/-}$ ILC population derived from MMR-deficient CRCs, including single-cell RNA-sequencing, flow cytometry, and mass cytometry. We find that the ILC compartment in MMR-deficient CRCs is dominated by $CD127^-CD103^+$ ILC1-like cells that exhibit a transcriptional profile related to cancer immunity.

RESULTS

Majority of ILCs in MMR-deficient colorectal cancers are CD127$^-$CD103$^+$ ILC1-like cells
Characterization of the CD7$^+$CD3$^-$CD127$^{+/-}$ ILC population in MMR-deficient CRCs by mass cytometric analysis (**Figure S1**, **Table S1**, **Table S2**) showed that the majority of the ILCs that infiltrate these tumors lack expression of conventional ILC marker CD127 (IL-7Rα) (**Figure 1A**). The majority of the CD7$^+$CD3$^-$ cells consisted of CD127$^-$CD103$^+$ ILC1-like cells, followed by a CD127$^-$CD103$^-$CD56$^{dim/bright}$ NK-like population, and a minor population of CD127$^+$ conventional ILCs (**Figure 1A**). The majority of CD127$^-$CD103$^+$ ILC1-like cells also expressed ectonucleotidase CD39, CD56, and CD45RO, whereas a minor population of CD127$^-$CD103$^+$CD45RA^{dim} cells was observed (**Figure 1A,B**). Nevertheless, CD45RA expression was mainly detected in the CD127$^-$CD103$^-$CD56$^{dim/bright}$ NK-like population (**Figure 1A,B**). Last, the CD127$^+$ conventional ILCs showed a CD45RA$^{+/-}$ phenotype (**Figure 1A,B**). Across four patients with MMR-deficient CRC, the CD127$^-$CD103$^+$CD45RO$^+$ ILC1-like population was the most prevalent constituting up to 75% (average 55%) of the total CD7$^+$CD3$^-$CD127$^{+/-}$ ILC population (**Figure 1C**). In addition to the aforementioned immune cell markers, the CD103$^+$ ILC1-like cells also expressed high levels of CD38 and CD69 (**Figure 1D**).

The CD127$^+$ conventional ILC population showed co-expression of the transcription factor *GATA3*, but not of *TBX21* (T-bet), *EOMES* or *RORC* (RORγt), indicating an ILC2 phenotype (**Figure 1E**). The CD127$^-$CD103$^+$CD45RO$^+$ population expressed *TBX21*, in line with their ILC1-like phenotype (**Figure 1E**). Both CD127$^-$CD103$^+$CD45RA^{dim} and CD127$^-$CD103$^-$CD45RA$^+$ populations showed the co-expression of *TBX21* and *EOMES*, indicative of a conventional NK cell phenotype (**Figure 1E**). Cytotoxic molecules granzyme B and perforin were highly expressed by all ILC populations, with the exception of the CD127$^+$ ILC2s (**Figure 1F**). NKp44, a marker used to characterize intraepithelial ILC1s,[20] was expressed at variable levels by the ILC1-like cells (**Figure S2**). We next investigated the presence of receptors recognizing HLA class I molecules. CD94 and NKG2A, able to bind to the non-classical HLA-E protein,[26] were highly expressed by the ILC1-like cells, and to a lower extent by the NK cell-like and ILC2 populations (**Figure S2**). Killer-cell immunoglobulin-like receptors (KIRs), recognizing different HLA class I molecules, were expressed by the ILC1-like and NK cell-like populations at variable levels, and to a lower extent by the ILC2 cells (**Figure S2**).

Figure 1. Majority of ILCs in MMR-deficient colorectal cancers are CD127$^-$CD103$^+$ ILC1-like cells.
A. t-SNE embedding showing 6.0x10^4 CD7$^+$CD3$^-$ innate lymphoid cells (ILCs) isolated from MMR-deficient CRCs (N=4) analyzed by mass cytometry. Colors represent relative expression of indicated immune cell markers. **B**. t-SNE embedding of (**A**) colored by ILC subsets. **C**. Circular plot showing the relative frequency of ILC subsets as average percentage of total CD7$^+$CD3$^-$ ILCs from MMR-deficient CRCs (N=4). **D.** Expression (ArcSinh5-transformed) of surface markers of CD127$^-$CD103$^+$CD45RO$^+$ ILC1-like (orange), CD127$^-$CD103$^+$CD45RA^{dim} ILC1-like (pink), CD127$^-$CD103$^-$CD45RA$^+$ NK-like (purple), and CD127$^+$CD103$^-$CD45RA$^{+/-}$ ILC2 (green) cells from MMR-deficient CRCs (N=4). **E.** Expression (ArcSinh5-transformed) of intracellular transcription factors and cytotoxicity markers of CD127$^-$CD103$^+$CD45RO$^+$ ILC1-like (orange), CD127$^-$CD103$^+$CD45RA^{dim} ILC1-like (pink), CD127$^-$CD103$^-$CD45RA$^+$ NK-like (purple), and CD127$^+$CD103$^-$CD45RA$^{+/-}$ ILC2 (green) cells from MMR-deficient CRCs (N=4).

We explored the full composition of the ILC compartment in MMR-deficient CRCs, independently of CD127 expression, by performing single-cell RNA-sequencing (scRNA-seq) on *ex vivo* isolated CD7^{+}CD3^{-}CD127$^{+/-}$ cells (**Figure S3, Table S1**). As the CD45 isoforms CD45RA and CD45RO cannot be detected at the RNA level, we included two CITE-sequencing antibodies against these markers (**Figure S3**). Transcriptomes were obtained from a total of 4270 ILCs and used for further analysis (**Figure S3**). A patient-specific analysis approach was performed to avoid loss of biologically-relevant populations upon integrating the scRNA-seq data of the different patients. Three to six distinct ILC clusters were identified per patient and visualized by t-SNE (**Figure 2A**). In each patient, we observed the presence of an *IL7R*$^{+}$*ITGAE*$^{-}$ (CD127^{+}CD103^{-}) conventional ILC cluster (except for CRC159), *IL7R*$^{-}$*ITGAE*$^{+}$ (CD127^{-}CD103^{+}) ILC1-like clusters, and *IL7R*$^{-}$*ITGAE*$^{-}$ (CD127^{-}CD103^{-}) NK-like clusters (**Figure 2A,B**). The ILC1-like clusters showed co-expression of CD45RO protein (antibody-derived tag, ADT), whereas CD45RA (ADT) protein expression was mainly found on the conventional ILC and/or NK-like clusters (**Figure 2A,B**). In line with the mass and flow cytometric data, the majority of the cells consisted of CD103^{+} ILC1-like cells, followed by NK-like cells, and a minor population of conventional ILCs (**Figure 2C**). Also in line with the protein data, the CD127^{+} conventional ILCs expressed *GATA3*, confirming their ILC2 phenotype (**Figure S4**).

CD127^{-}CD103^{+} ILC1-like cells exhibit a transcriptional profile related to cancer immunity

We next examined what functionally distinguished the different ILC clusters identified using differential expression analysis for each patient. Across patients, clusters with similar significantly differential expressed genes could be detected (**Figure 3A**, **Figure S5**). Differentially expressed genes for the CD103^{+} ILC1-like clusters across patients included genes encoding co-inhibitory and co-stimulatory molecules, KIRs, cytotoxicity markers, and HLA class II molecules (**Figure 3A**, **Figure S5**). As compared to the other populations, the CD103^{+} ILC1-like cells showed high expression of *ENTPD1* (CD39), activating receptor *KLRB1* (CD161), inhibitory receptor *KLRC1* (NKG2A), and activating receptor *NCR2* (NKp44) (**Figure 3B**). Genes encoding co-inhibitory (TIGIT, TIM3, LAG3) and co-stimulatory molecules (4-1BB) were also expressed mainly by the CD103^{+} ILC1-like cells (**Figure 3B**). CD103^{+} ILC1-like cells highly expressed *KIRs* as compared to the other populations (**Figure 3B**). Across all patients, the most commonly expressed KIR was *KIR2DL4*, which can bind to HLA-G molecules.[27] Of all ILC populations, the CD103^{+} ILC1-like cells were the only ones displaying hallmarks of proliferation (*MKI67*$^{+}$) in MMR-deficient CRCs (**Figure 3B**).These *MKI67*+ proliferating ILC1-like clusters commonly co-expressed genes encoding co-inhibitory and co-stimulatory molecules, KIRs, cytotoxicity markers, and HLA class II molecules (**Figure 3B**, **Figure S5**). The most frequently expressed interleukin receptors in the CD103^{+} ILC1-like population were *IL2RB*, and to a lower extent *IL2RG*, *IL4R*, *IL10RA*, and *IL21R* (**Figure S6**). Abundant expressed interleukin transcripts were *IL32*, and to a lower extent *IL16* (**Figure S6**).

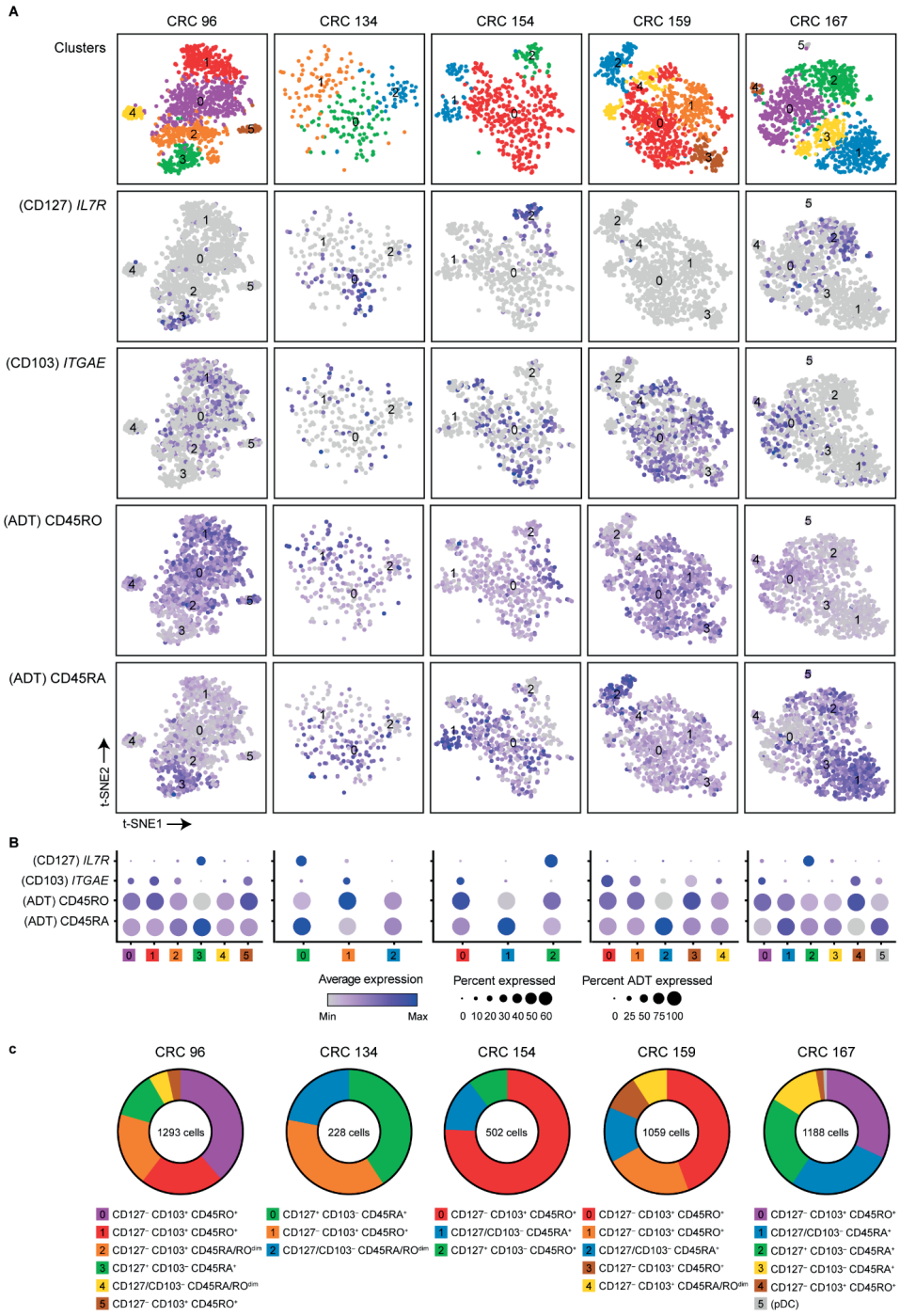


Figure 2. Unbiased, single-cell characterization of ILC populations in MMR-deficient colorectal cancers.
A. t-SNE embedding showing the clustering of CD7^{+}CD3^{-} ILCs isolated from MMR-deficient CRCs (N=5) analyzed by single-cell RNA-sequencing per patient. Colors represent functionally distinct ILC clusters (top row) and relative expression of indicated genes and proteins (remainder). Each dot represents a single cell. ADT; antibody-derived tag. **B**.

Dot plots showing the relative expression of indicated markers in the clusters identified in (**A**). Color intensity indicates average expression and size of the dot indicates the percentage of cells expressing the gene or protein. ADT; antibody-derived tag. **C**. Circular plots showing the frequencies of the distinct ILC clusters identified in (**A**) as a percentage of total $CD7^{+}CD3^{-}$ ILCs per patient.

Differentially expressed genes for the NK-like clusters across patients included genes encoding cytotoxicity markers and chemokines (**Figure 3A**, **Figure S5**). NK-like cells showed high expression of *FCG3RA* (encoding CD16), *NKG7* (NK cell granule protein), *KLRD1* (CD94), and activating receptors *KLRF1* (NKp80), *NCR1* (NKp46) and *NCR3* (NKp30) (**Figure 3B**). Both ILC1-like and NK-like clusters highly expressed cytotoxic molecules *GNLY*, *GZMB*, and *PRF1* as well as proinflammatory cytokine *IFNG* (**Figure 3B**). In opposition to $CD103^{+}$ ILC1-like cells, the NK-like cells generally showed lower expression of genes encoding co-inhibitory and co-stimulatory receptors and KIRs, and lacked expression of proliferation marker *MKI67* (**Figure 3B**). The NK-like clusters further distinguished from the other populations by high expression of chemokines CCL3 and/or CCL4 (**Figure 3B**). Interleukin receptor *IL2B* was expressed by the NK-like clusters, but to a lower extent than the $CD103^{+}$ ILC1-like cells (**Figure S6**). Last, differentially expressed genes for the ILC2 clusters across patients included *FOS* (transcriptional regulation), *XCL1* (lymphotactin), *KIT* (c-kit), *LTB* (regulation of inflammation), *SELL* (CD62L), and *NFKBIA* (IκBα) (**Figure 3A**, **Figure S5**). The ILC2 cells generally lacked expression of co-inhibitory and co-stimulatory molecules, KIRs, and cytotoxic molecules, with some exceptions (**Figure 3B**). Further, generally no expression of genes encoding chemokines CCL3, CCL4 or CCL5, HLA class II molecules, and proliferation marker Ki-67 were found in the ILC2s (**Figure 3B**). Importantly, the majority of the entire ILC population in MMR-deficient CRCs lacked expression of co-inhibitory receptors and immune checkpoint molecules *PDCD1* (PD-1) and *CTLA4*, both at the RNA and protein level (**Figure 3B**, **Figure S7**). Taken together, as compared to NK-like cells, the ILC1-like cells specifically express genes encoding CD103, CD39, CD161 and NKG2A, more frequently express genes encoding co-inhibitory and co-stimulatory molecules, KIRs and HLA class II, and display hallmarks of proliferation in MMR-deficient CRCs.

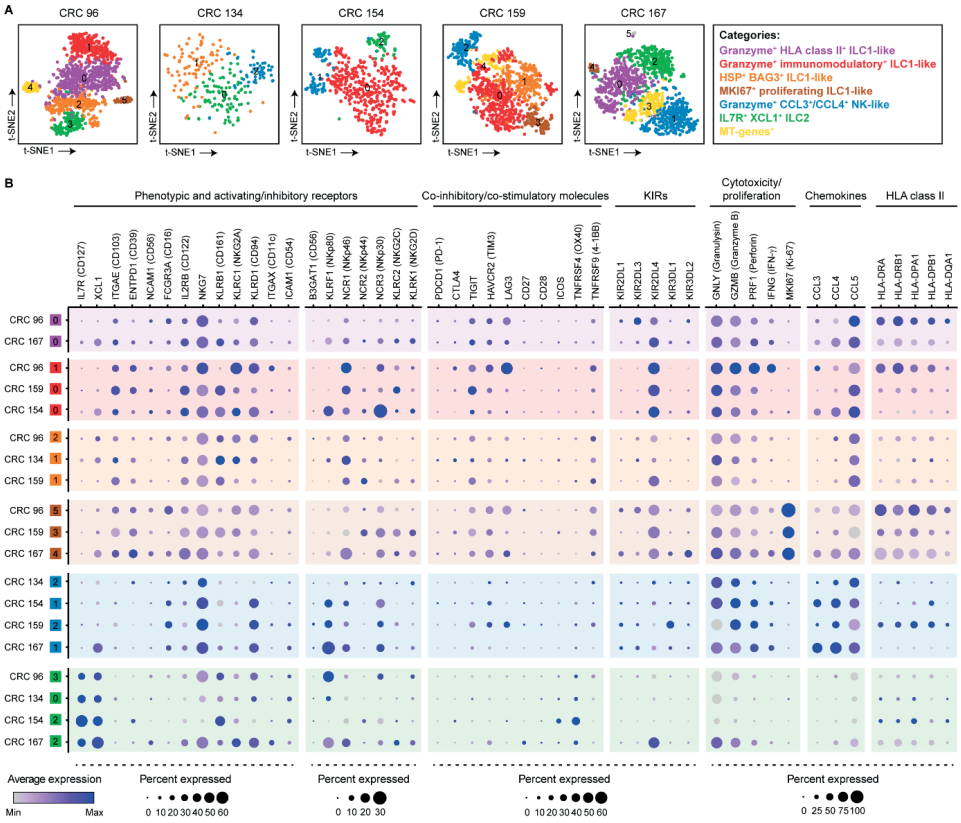


Figure 3. CD127⁻CD103⁺ ILC1-like cells exhibit a transcriptional profile related to cancer immunity.
A. t-SNE embedding showing the clustering of CD7⁺CD3⁻ ILCs from MMR-deficient CRC 96 (N=1293), CRC 134 (N=288), CRC 154 (N=502), CRC 159 (N=1059), and CRC 167 (N=1188) analyzed by single-cell RNA-sequencing. The top ten differentially expressed genes of the distinct ILC clusters are shown in **Figure S5**. Colors indicate functionally distinct categories of ILC subsets. Each dot represents a single cell. **B**. Dot plots showing the relative expression of indicated genes in the clusters identified in (**A**). Color intensity indicates average expression and size of the dot indicates the percentage of cells expressing the gene. Background colors indicate functionally distinct categories of ILC subsets as in (**A**).

Last, we validated our findings in a publicly available scRNA-seq cohort[28] consisting of MMR-deficient CRCs, MMR-proficient CRCs, and adjacent healthy tissues. This dataset confirmed that the majority of the ILC population consisted of ILC1-like cells that were particularly abundant in MMR-deficient CRCs, in line with our observations (**Figure 4A-C**).

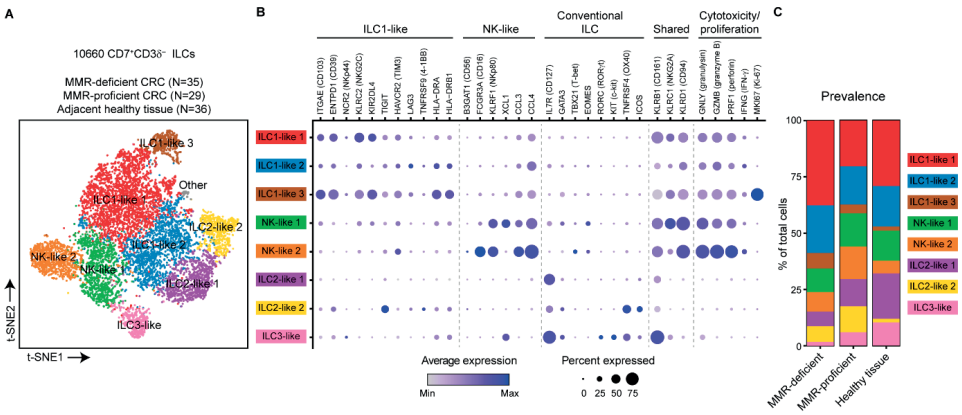


Figure 4. Validation of the abundance and transcriptional profile of CD127^{-}CD103^{+} ILC1-like cells in an independent CRC cohort.
A. t-SNE embedding showing the clustering of CD7^{+}CD3δ^{-} ILCs (N=10660) from MMR-deficient CRCs (N=35), MMR-proficient CRCs (N=29), and adjacent healthy tissues (N=36) obtained from the single-cell RNA-sequencing dataset of Pelka *et al.*[28]. The top ten differentially expressed genes of the distinct ILC clusters are shown in **Figure S9**. Each dot represents a single cell. **B**. Dot plots showing the relative expression of indicated genes in the clusters identified in (**A**). Color intensity indicates average expression and size of the dot indicates the percentage of cells expressing the gene. **C**. Prevalence of the clusters identified in (**A**) among MMR-deficient CRCs, MMR-proficient CRCs, and adjacent healthy tissues shown as percentage of total cells.

5

DISCUSSION

In this work, we provide an unbiased, single-cell characterization of CD7^{+}CD3^{-}CD127$^{+/-}$ ILCs from MMR-deficient CRCs. We defined biological heterogeneity within the ILC population through scRNA-seq, flow cytometry, and mass cytometry on thousands of individual ILCs, independently of CD127 expression. Unsupervised clustering of ILCs distinguished ILC1-like (CD127^{-}CD103^{+}), NK-like (CD127^{-}CD103^{-}), and conventional ILC (CD127^{+}CD103^{-}) subsets across different patients, and we defined the similarities and differences between these ILC subsets. Specifically, we showed that i) the major ILC population in MMR-deficient CRCs consisted of CD127^{-}CD103^{+} ILC1-like cells, ii) these ILC1-like cells specifically expressed CD39, CD45RO, and transcription factor T-bet, iii) high expression of genes encoding KIRs, co-inhibitory and co-stimulatory molecules, NKG2A, and HLA class II distinguished the ILC1-like cells from NK-like cells and conventional ILCs, and iv) the ILC1-like cells are actively proliferating in MMR-deficient CRCs.

The role of ILCs in colorectal cancer immunity and immunotherapy remains largely unclear. Recent single-cell transcriptome characterizations specifically focused on the Lin^{-}CD127^{+} compartment in healthy and malignant colorectal tissue.[29,30] Here, we aimed to characterize the entire CD7^{+}CD3^{-} ILC compartment, independently of CD127 expression, in MMR-deficient CRCs. Our results revealed that the CD127^{+} conventional ILC population is a relative small subset of the total ILC population in CRC tissues, whereas CD127^{-}CD103^{+}

ILC1-like cells and CD127$^-$CD103$^-$ NK-like cells are present in larger frequencies. Such intraepithelial ILC1-like cells have been described to be found in solid tumors, resembling the ones described in the mass cytometry study of lung and colorectal tumors of Simoni *et al.*[31] In line with their study, NKp44, a marker used to characterize intraepithelial ILC1 cells,[20] was not or infrequently observed on the ILC1-like cells in CRC tissues (**Figure S2**). We found that the CD103$^+$ ILC1-like cells could express *IL2RB* and, to a lower extent, *IL2RG*, known to transduce signals by IL-15 in addition to IL-2. IL-15 is a common γ-chain cytokine that can be expressed by epithelial cells and is critical for the maintenance of intestinal intraepithelial lymphocytes,[32] in line with their intraepithelial localization. In contrast to the study of Simoni *et al.* that reported a low expression of cytotoxic granule components,[31] we did observe that the CD103$^+$ ILC1-like cells exhibited a high expression of cytotoxic molecules and IFN-γ. Cytotoxicity and tumor-specific expansion of CD127$^-$CD103$^+$ ILC1-like cells have been shown in murine cancer models, which contrasted the lack of cytotoxicity in the distinct population of CD127$^+$CD103$^-$ conventional ILC1.[33] We found consistent co-expression of CD103, CD39 and CD45RO on the ILC1-like population, implying that these markers can be used to identify the cells. In our study, the CD103$^+$ ILC1-like populations appeared to be transcriptionally distinct but related to conventional NK cells, as has been described before,[31,34] with regard to expression of activating and/or inhibitory innate immune receptors and cytotoxicity markers. In addition, the CD103$^+$ ILC1-like cells showed characteristics of tissue-resident cytotoxic T cells in terms of expression of tissue-residency markers, co-inhibitory and co-stimulatory molecules, cytotoxicity markers, and the expression of transcripts encoding variable regions of the TCR. We found gene expression of CD3ε as well as TRCδ, TRCα, and TCRβ by the ILCs (**Figure S8**), in line with what others have reported.[35-37] TCR rearrangement and clonality of the ILCs will be subject of future studies.

Although previous studies in murine models reported on PD-1 and/or CTLA-4 expression on NK cells and ILC2s in other cancer types and their subsequent involvement in mediating responses to PD-1/PD-L1 blockade,[38-40] the majority of CD7$^+$CD3$^-$ ILCs in MMR-deficient CRCs lacked expression of immune checkpoint molecules PD-1 and CTLA-4. This highlights the need to further investigate the contribution of ILCs in response to PD-1/PD-L1 blockade immunotherapy in human cancer and the involvement of other immune cells such as CD4$^+$ T cells. We found that the ILC1-like cells express high levels of HLA class I-binding receptors KIRs, CD94, and NKG2A, which may sense the lack of HLA class I expression on the cancer cells. KIR2DL4 was the most frequently expressed KIR across all five patients, and has been shown to function as an activating receptor involved in cytokine and chemokine secretion of which the sole ligand known is HLA-G[27]. CD94/NK2G2A is an inhibitory receptor able to bind to the non-classical HLA-E protein,[26,41] suggesting that blocking NKG2A may result in an increased effector function of the CD103$^+$ ILC1-like cells, which will be tested in future functional studies. Last, the expression of HLA class II by CD103$^+$ ILC1-like cells may suggest that these cells have antigen-presenting properties, an interesting direction for further exploration.

Our study is limited with respect to the unraveling of heterogeneity within the CD127$^+$ conventional ILC population, as these cells were present at low frequencies (average 9.7%

of the total CD7^{+}CD3^{-} population; **Figure 1C**). The CD127^{+} ILCs were positive for GATA3, indicating an ILC2 phenotype, in line with other studies on ILC2s in CRC.[29] No RORγt^{+} ILC3s could be detected in our cohort of MMR-deficient CRCs, while we detected an ILC3 cluster in the dataset of Pelka *et al.*[28] that was mainly derived from adjacent healthy tissues. Previous studies reported on remarkably decreased ILC3 populations in colorectal tumor tissues,[29,42-44] albeit being present at high frequencies in healthy colon tissues.[30,31,42] Studies have shown plasticity of ILC3s into ILC1-like cells driven by the presence of cytokines such as IL-12, IL-23, TGF-β in the microenvironment of CRCs.[22,45] In addition, ILC1-like phenotypes could emerge from NK cell differentiation driven by TGF-β signaling in murine cancer models.[46] It remains to be determined what the precursor of the CD103^{+} ILC1-like cells is in colorectal tumors and whether they can further differentiate into other ILC populations depending on environmental signals. In general, the conventional CD127^{+} ILCs in our study did not show expression of cytotoxic molecules both at the transcriptome and protein level, in line with what others found.[17] We will examine the localization and interacting cells of the different ILC populations with imaging mass cytometry. The expression of intracellular CD3ε and TCRs complicates the detection of these cells in tissues. We found that the ILCs lacked expression of (intracellular) CD3δ, a marker we are currently employing to pinpoint the cells by imaging mass cytometry in colorectal tumors with and without HLA class I expression, accompanied by functional studies.

In conclusion, this study provides a comprehensive description of heterogeneity of the ILC population in MMR-deficient CRCs. Our results point toward an active involvement of ILCs in the recognition of and cytotoxicity against MMR-deficient cancers. Our findings may facilitate the exploitation of ILC1-like cells as targets for cancer immunotherapy, and thereby broadening the potential of tumor-resident immune cell populations.

METHODS

Patient samples

Primary CRC tissues (N=10) were from a total of nine patients with CRC undergoing surgical resection of their tumor at the Leiden University Medical Center (LUMC, the Netherlands) (**Table S1**). All patients were treatment-naïve, and did not have a previous history of inflammatory bowel disease. One patient was diagnosed with multiple primary colorectal tumors at different locations, two of which were included in this study. All patients included in this study provided written informed consent. The study was approved by the Medical Ethical Committee of the Leiden University Medical Center (protocol P15.282). All specimens were anonymized and handled according to the ethical guidelines described in the Code for Proper Secondary Use of Human Tissue in the Netherlands of the Dutch Federation of Medical Scientific Societies.

Processing of colorectal cancer tissues

Details on the processing of primary CRC tissues have been described previously.[16] Briefly, macroscopic sectioning from the lumen to the most invasive area of the tumor was performed. Tissue specimens were collected in IMDM+Glutamax medium (Gibco) with 20% fetal calf serum (FCS) (Sigma-Aldrich), 1% pen/strep (Gibco) and fungizone (Gibco), and 0.1% ciprofloxacin (provided by apothecary LUMC) and gentamicin (Invitrogen). Tissues were minced into small fragments in a petri dish, followed by enzymatical digestion with 1 mg/mL collagenase D (Roche Diagnostics) and 50 µg/mL DNase I (Roche Diagnostics) in 5 mL IMDM+Glutamax medium for 30 min at 37°C in gentleMACS C tubes (Miltenyi Biotec). Cell suspensions were dissociated mechanically during and after incubation using the gentleMACS Dissociator (Miltenyi Biotec). Thereafter, cell suspensions were filtered through a 70-µm cell strainer (Corning) and washed in IMDM+Glutamax medium with 20% FCS, 1% pen/strep, and 0.1% fungizone. Cell count and viability were determined with the Muse Count & Viability Kit (Merck) on the Muse Cell Analyzer (Merck), and cells were cryopreserved in liquid nitrogen based on the number of viable cells in IMDM+Glutamax medium complemented 1:1 with 80% FCS and 20% dimethyl sulfoxide (DMSO) (Merck).

Immunohistochemical detection of MMR, β2m, and HLA class I

MMR status of CRC tissues was examined by immunohistochemical staining of PMS2 (anti-PMS2 antibodies; clone EP51, DAKO) and MSH6 (anti-MSH6 antibodies; clone EPR3945, Abcam) proteins.[47] Tumor samples showing the lack of expression of at least one of the MMR-proteins in the presence of an internal positive control were defined as MMR-deficient. Further, β2m expression on CRC tissues was determined by immunohistochemical staining of β2m (anti-β2m antibodies; clone EP2978Y, Abcam). Last, immunohistochemical staining of HLA class I expression on CRC tissues was performed with HCA2 and HC10 monoclonal antibodies (Nordic-MUbio), and classified as HLA class I positive, weak, or loss as described previously.[14]

Mass cytometry staining and data analysis

Mass cytometric analysis was performed on four MMR-deficient CRC tissues with an antibody panel of 44 immune cell markers covering immune lineage markers, differentiation/activation markers, co-inhibitory and co-stimulatory molecules, adhesion molecules, transcription factors, and cytotoxicity markers (**Table S2**). Briefly, single-cell suspensions were thawed and Percoll (GE Healthcare) density-gradient centrifugation was performed to isolate immune cells. Cells were washed in Maxpar Cell Staining Buffer complemented with EDTA (CSB, Fluidigm) and counted. Up to 3 million cells of each sample were incubated with 1 mL CSB containing 1 µM Cell-ID intercalator-103Rh (Fluidigm) for 15 min at room temperature (rT) to discriminate live from dead cells. Cells were washed in CSB, incubated with human Fc receptor block (BioLegend) for 10 min at rT, and stained with a cell-surface antibody cocktail for 45 min at rT in a final volume of 100 µL (**Table S2**). Thereafter, intracellular staining was performed

with Fixation Buffer and Intracellular Staining Permeabilization Wash Buffer (BioLegend) and stained with an intracellular antibody cocktail for 30 min at rT. Cells were washed once in Perm buffer and twice in CSB, followed by incubation with 1 mL Maxpar Fix and Perm buffer (Fluidigm Sciences) containing 0.125 μM Cell-ID intercalator-Ir (Fluidigm) overnight at 4°C to discriminate singlets from doublets. The next day, cells were washed three times in CSB, and one time in de-ionized water immediately prior to data acquisition. Cells were acquired on a Helios mass cytometer (Fluidigm) at an event rate of <500 events/sec in de-ionized water containing 10x diluted EQ Four Element Calibration Beads (Fluidigm). Data were normalized with the normalization passport EQ-P13H2302_ver2 for each experiment.
Normalized FCS files were analyzed in FlowJo (version 10.6.1, Tree Star Inc). Data were checked for quality of staining and gated for live, single, $CD45^+$ cells using 191/193Ir DNA intercalator, CD45, event length, center, width, residual, 103Rh DNA intercalator, and 140Ce bead channels (gating strategy shown in **Figure S1**). Next, we specifically exported the $CD7^+CD3^-$ gate for each sample as FCS file for downstream analysis (**Figure S1**). $CD7^+CD3^-$ cells were hyberbolic ArcSinh transformed with a cofactor of 5, sample-tagged, and subjected to dimensionality reduction analysis in the Cytosplore software[48].

Sorting of ILCs from colorectal cancers for immunophenotyping and cell culturing
$CD7^+CD3^-CD127^{+/-}$ ILCs were sorted from MMR-deficient CRC tissues of six patients by FACS for immunophenotyping and cell culture (gating strategy shown in **Figure S2**). Cells were thawed, rest at 37°C in IMDM+L-glutamine (Lonza)/10% nHS for 1h, and washed in FACS buffer (PBS/1% FCS). Thereafter, cells were incubated with human Fc receptor block (BioLegend) and stained with the following cell surface antibodies: 1:20 anti-CD7-V450 [clone M-T701, BD Biosciences], 1:25 anti-CD3-Am Cyan [clone SK7, BD Biosciences] or 1:50 anti-CD3-PE [clone SK7, BD Biosciences], 1:20 anti-CD127-FITC [clone A019D5, Sony] or 1:150 anti-CD127-PE-Cy7 [clone A01905, BioLegend], 1:20 anti-CD45RO-PerCP-Cy5.5 [clone UCHL1, Sony], 1:20 anti-CD45RA-PE/Dazzle [clone HI100, Sony] or 1:30 anti-CD45RA-FITC [clone L48, BD Biosciences], and 1:150 anti-CD56-APC-R700 [clone NCAM16.2, BD Biosciences] for 45 min at 4°C together with different additional antibodies for immunophenotyping (including 1:200 anti-CD94-BV605 [clone HP-3D9, BD Biosciences], 1:20 anti-NKp44-PE [clone 253415, R&D Systems], 1:60 anti-CD39-APC [clone A1, BioLegend], 1:30 anti-NKG2A-APC [clone z199, Beckman Coulter], 1:150 anti-DNAM-1-BV510 [clone DX11, BD Biosciences], 1:10 anti-KIR3DL2-PE [clone #539304, R&D Systems], 1:10 anti-KIR2DL4-PE [clone #181703, R&D Systems], 1:50 anti-KIR2DL1/DS1-PE [clone EB6, Beckman Coulter], 1:40 anti-KIR2DL3-PE [clone GL183, Beckman Coulter], 1:20 anti-KIR3DL1/DS1-PE [clone Z27, Beckman Coulter], and 1:40 anti-KIR2DS4-PE [clone FES172, Beckman Coulter]. To discriminate live from dead cells, a live/dead fixable near-infrared viability dye (1:1000, Life Technologies) was included in each staining. Cells were washed three times in FACS buffer before sorting the cells on a FACS Aria III 4L (BD Biosciences). CompBeads (BD Biosciences) and ArC reactive beads (Life Technologies) were used for compensation controls. The ILCs were sorted in IMDM+L-glutamine medium containing feeder cells ($1x10^6$/mL), PHA (1 μg/mL; Thermo Fisher Scientific), gentamicin (50 μg/mL), and fungizone (0.5 μg/mL), and different cytokine mixes with: IL-2 (100 IU/mL and 500 IU/mL; Novartis), IL-15 (10 ng/mL; R&D Systems), IL-7 (10 ng/mL; Peprotech), IL-21 (10 ng/mL; Gibco), TNF-α (10 ng/mL; Peprotech), SCF (25 ng/mL; Miltenyi). The cytokine mixes tested included i) SCF (25 ng/mL), IL-2 (10 CU/mL), IL-7 (25 ng/mL) with feeder cells and PHA (1 μg/mL), ii) OP9-DL1 cells with SCF (25 ng/mL), IL-2 (10 CU/mL), IL-7 (25 ng/mL), iii) IL-2 (10 CU/mL) and IL-15 (10 ng/mL) with feeder cells and PHA (1 μg/mL), iv) TCGF (10%), IL-7 (10 ng/mL), IL-15 (10 ng/mL), IL-21 (10 ng/mL), TNF-α (10 ng/mL) with feeder cells/PHA (1 μg/mL) and EBV, and v) TCGF (10%), IL-7 (10 ng/mL), IL-15 (10 ng/mL), IL-21 (10 ng/mL), TNF-α (10 ng/mL) with feeder cells and PHA (1 μg/mL). During and after 3-4 weeks expansion, the purity and phenotype of the ILCs were assessed by flow cytometry.

Sorting of ILCs from colorectal cancers for single-cell RNA-sequencing
$CD7^+CD3^-CD127^{+/-}$ ILCs were sorted from MMR-deficient CRC tissues of five patients by FACS and processed for single-cell RNA-sequencing (scRNA-seq). Hashtag oligo (HTOs) antibodies were included

for sample ID, and antibody-derived tags (ADTs) for CD45RA and CD45RO protein expression using CITE-seq[49] (**Figure S3**), as described previously.[50] In short, cell suspensions were thawed, rest at 37°C in IMDM (Lonza)/20% FCS for 1h, and washed in FACS buffer (PBS (Fresenius Kabi)/1% FCS). Cells were subsequently incubated with human Fc receptor block (BioLegend) for 10 min at 4°C. Thereafter, cells were stained with the following cell surface antibodies: 1:160 anti-CD45-PerCP-Cy5.5 [clone 2D1, eBioscience], 1:50 anti-CD3-PE [clone SK7, BD Biosciences], 1:200 anti-CD7-APC [clone 124-1D1, eBioscience], 1:60 anti-EPCAM-FITC [clone HEA-125, Miltenyi], and 1:80 anti-TCRγδ-BV421 [clone 11F2, BD Biosciences]. In parallel, the following ADTs for CD45RA and CD45RO were added in the antibody mix: 1 µg of TotalSeq-C anti-CD45RA (clone HI100, BioLegend) and 1 µg of anti-CD45RO (clone UCHL1, BioLegend), together with the following HTOs: 0.5 µg of a unique TotalSeq-C CD298/β2M hashtag antibody (clone LNH-94/2M2, BioLegend) for each sample (N=5) for 30 min at 4°C. To discriminate live from dead cells, a live/dead fixable near-infrared viability dye (1:1000, Life Technologies) was included in each staining. Cells were washed three times in FACS buffer before sorting the cells. CompBeads (BD Biosciences) and ArC reactive beads (Life Technologies) were used for compensation controls. Single, live $EPCAM^{-}CD45^{+}CD7^{+}CD3^{-}$ cells were sorted on a FACS Aria III 4L (BD Biosciences). After sorting, the samples were pooled in FACS buffer, washed, and resuspended in PBS with 0.04% BSA at a concentration of 1,000 cells/µL for downstream application.

Single-cell RNA-sequencing data analysis

Libraries were prepared using the Chromium Single Cell 5' Reagent Kit v1 chemistry (10X Genomics), following the manufacturer's instructions, and sequenced on a HiSeq X Ten using paired-end 2x150 bp sequencing (Illumina) as described previously.[50] Briefly, reads were aligned to the human reference genome (GRCh38) and quantified using Cell Ranger (version 3.1.0). Seurat (version 4.0.4)[51] was used for downstream analysis. Cells with less than 200 detected genes and genes expressed in less than six cells were excluded, resulting in a dataset of 5748 cells. These were demultiplexed based on HTO enrichment using the MULTIseqDemux algorithm[52], resulting in 4755 cells excluding doublet and negative HTOs. Next, a patient-specific analysis approach was performed were cells with a mitochondrial gene content greater than 10% and cells with outlying numbers of expressed genes (>3000) were filtered out for each patient. Transcriptomes from 1293 cells (HTO1), 228 cells (HTO6), 502 cells (HTO7), 1059 cells (HTO8), and 1188 cells (HTO9) were used for further analysis (**Figure S3**). Data normalization was performed using the 'LogNormalize' function with scale factor 10,000. The 'FindVariableFeatures' function was used to identify 2,000 variable features in the data. Data were scaled and heterogeneity associated with mitochondrial contamination was regressed out. Next, PCA was run on the scaled data and the optimal number of dimensions was determined for each patient individually. Clustering was performed with the 'FindNeighbors' and 'FindClusters' functions, where the optimal resolution parameter was set for each patient individually. Non-linear dimensional reduction with t-SNE[53] and uniform manifold approximation (UMAP)[54] were used to visualize the cells. Differentially expressed genes were identified with the 'FindAllMarkers' function with min.pct and logfc.threshold at 0.25.

Validation in a publicly available single-cell RNA-sequencing CRC dataset

We validated our findings in the publicly available scRNA-seq CRC dataset of Pelka *et al.*[28] consisting of MMR-deficient CRC (N=35), MMR-proficient CRC (N=29), and adjacent healthy tissue (N=36). Briefly, cells of main partition T/NK/ILC were extracted for all samples. In Seurat (version 4.0.3), $CD7^{+}CD3\delta^{-}$ cells were selected and re-clustered with resolution parameter 0.4. Non-linear dimensional reduction with t-SNE was used to visualize the clustered cells. Differentially expressed genes were identified with the 'FindAllMarkers' function with min.pct and logfc.threshold at 0.25.

ACKNOWLEDGEMENTS

We thank K.C.M.J. Peeters, M.G. Kallenberg-Lantrua, D. Berends-van der Meer, and F.A. Holman for their help in collecting and providing samples from patients with colorectal cancer; the Flow cytometry

Core Facility of the Leiden University Medical Center for their help with cell sorting; the Leiden Genome Technology Center for their help with single-cell RNA-sequencing; A. Mahfouz for his help with the scRNA-sequencing analysis.

AUTHOR CONTRIBUTIONS

N.L. de Vries conceived the study and performed experiments. N.L. de Vries, L. Zheng, J. Roelands, and D. Ruano performed the computational analyses presented in this paper. N.L. de Vries, L. Zheng, J. Roelands, D. Ruano, N.F. de Miranda, and F. Koning analyzed and interpreted the data. M. van der Ploeg, M.E. Ijsselsteijn, and R. van der Breggen assisted with experiments. J. Krop designed the mass cytometry antibody panel under supervision of M. Eikmans. F. Koning and N.F. de Miranda supervised the study. The manuscript was written by N.L. de Vries, F. Koning, and N.F. de Miranda in collaboration with all co-authors.

COMPETING INTERESTS

The authors declare no competing interests exist.

REFERENCES

1 Fridman, W. H., Zitvogel, L., Sautes-Fridman, C. & Kroemer, G. The immune contexture in cancer prognosis and treatment. *Nature reviews. Clinical oncology* 14, 717-734, doi:10.1038/nrclinonc.2017.101 (2017).

2 Dolcetti, R. *et al.* High prevalence of activated intraepithelial cytotoxic T lymphocytes and increased neoplastic cell apoptosis in colorectal carcinomas with microsatellite instability. *Am J Pathol* 154, 1805-1813, doi:10.1016/s0002-9440(10)65436-3 (1999).

3 Tumeh, P. C. *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 515, 568-571, doi:10.1038/nature13954 (2014).

4 Taube, J. M. *et al.* Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res* 20, 5064-5074, doi:10.1158/1078-0432.Ccr-13-3271 (2014).

5 Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D. & Perucho, M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363, 558-561, doi:10.1038/363558a0 (1993).

6 Germano, G. *et al.* Inactivation of DNA repair triggers neoantigen generation and impairs tumour growth. *Nature* 552, 116-120, doi:10.1038/nature24673 (2017).

7 Le, D. T. *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409-413, doi:10.1126/science.aan6733 (2017).

8 Overman, M. J. *et al.* Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *The Lancet. Oncology* 18, 1182-1191, doi:10.1016/s1470-2045(17)30422-9 (2017).

9 Overman, M. J. *et al.* Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 36, 773-779, doi:10.1200/jco.2017.76.9901 (2018).

10 Chalabi, M. *et al.* Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat Med* 26, 566-576, doi:10.1038/s41591-020-0805-8 (2020).

11 Bicknell, D. C., Kaklamanis, L., Hampson, R., Bodmer, W. F. & Karran, P. Selection for beta 2-microglobulin mutation in mismatch repair-defective colorectal carcinomas. *Current biology : CB* 6, 1695-1697, doi:10.1016/s0960-9822(02)70795-1 (1996).

12 Dierssen, J. W. *et al.* HNPCC versus sporadic microsatellite-unstable colon cancers follow different routes toward loss of HLA class I expression. *BMC Cancer* 7, 33, doi:10.1186/1471-2407-7-33 (2007).

13 Kloor, M. *et al.* Immunoselective pressure and human leukocyte antigen class I antigen machinery defects in microsatellite unstable colorectal cancers. *Cancer Res* 65, 6418-6424, doi:10.1158/0008-5472.can-05-0044 (2005).

14 Ijsselsteijn, M. E. *et al.* Revisiting immune escape in colorectal cancer in the era of immunotherapy. *Br J Cancer* 120, 815-818, doi:10.1038/s41416-019-0421-x (2019).

15 Middha, S. *et al.* Majority of B2M-Mutant and -Deficient Colorectal Carcinomas Achieve Clinical Benefit From Immune Checkpoint Inhibitor Therapy and Are Microsatellite Instability-High. *JCO precision oncology* 3, doi:10.1200/po.18.00321 (2019).

16 de Vries, N. L. *et al.* High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of antitumour immunity. *Gut* 69, 691-703, doi:10.1136/gutjnl-2019-318672 (2020).

17 Vivier, E. *et al.* Innate Lymphoid Cells: 10 Years On. *Cell* 174, 1054-1066, doi:10.1016/j.cell.2018.07.017 (2018).

18 Glatzer, T. *et al.* RORγt$^+$ innate lymphoid cells acquire a proinflammatory program upon engagement of the activating receptor NKp44. *Immunity* 38, 1223-1235, doi:10.1016/j.immuni.2013.05.013 (2013).

19 Spits, H. *et al.* Innate lymphoid cells--a proposal for uniform nomenclature. *Nat Rev Immunol* 13, 145-149, doi:10.1038/nri3365 (2013).

20 Fuchs, A. *et al.* Intraepithelial type 1 innate lymphoid cells are a unique subset of IL-12- and IL-15-responsive IFN-gamma-producing cells. *Immunity* 38, 769-781, doi:10.1016/j.immuni.2013.02.010 (2013).

21 Hazenberg, M. D. & Spits, H. Human innate lymphoid cells. *Blood* 124, 700-709, doi:10.1182/blood-2013-11-427781 (2014).

22 Cella, M. *et al.* Subsets of ILC3-ILC1-like cells generate a diversity spectrum of innate lymphoid cells in human mucosal tissues. *Nat Immunol* 20, 980-991, doi:10.1038/s41590-019-0425-y (2019).

23 Bernink, J. H. *et al.* Interleukin-12 and -23 Control Plasticity of CD127(+) Group 1 and Group 3 Innate Lymphoid Cells in the Intestinal Lamina Propria. *Immunity* 43, 146-160, doi:10.1016/j.

immuni.2015.06.019 (2015).

24 Crinier, A. *et al.* High-Dimensional Single-Cell Analysis Identifies Organ-Specific Signatures and Conserved NK Cell Subsets in Humans and Mice. *Immunity* 49, 971-986.e975, doi:10.1016/j.immuni.2018.09.009 (2018).

25 Krabbendam, L. *et al.* CD127+ CD94+ innate lymphoid cells expressing granulysin and perforin are expanded in patients with Crohn's disease. *Nat Commun* 12, 5841, doi:10.1038/s41467-021-26187-x (2021).

26 Lee, N. *et al.* HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A. *Proc Natl Acad Sci U S A* 95, 5199-5204, doi:10.1073/pnas.95.9.5199 (1998).

27 Rajagopalan, S. & Long, E. O. KIR2DL4 (CD158d): An activation receptor for HLA-G. *Front Immunol* 3, 258, doi:10.3389/fimmu.2012.00258 (2012).

28 Pelka, K. *et al.* Spatially organized multicellular immune hubs in human colorectal cancer. *Cell* 184, 4734-4752.e4720, doi:10.1016/j.cell.2021.08.003 (2021).

29 Qi, J. *et al.* Single-cell transcriptomic landscape reveals tumor specific innate lymphoid cells associated with colorectal cancer progression. *Cell Rep Med* 2, 100353, doi:10.1016/j.xcrm.2021.100353 (2021).

30 Mazzurana, L. *et al.* Tissue-specific transcriptional imprinting and heterogeneity in human innate lymphoid cells revealed by full-length single-cell RNA-sequencing. *Cell Res* 31, 554-568, doi:10.1038/s41422-020-00445-x (2021).

31 Simoni, Y. *et al.* Human Innate Lymphoid Cell Subsets Possess Tissue-Type Based Heterogeneity in Phenotype and Frequency. *Immunity* 46, 148-161, doi:10.1016/j.immuni.2016.11.005 (2017).

32 Sheridan, B. S. & Lefrançois, L. Intraepithelial lymphocytes: to serve and protect. *Curr Gastroenterol Rep* 12, 513-521, doi:10.1007/s11894-010-0148-6 (2010).

33 Dadi, S. *et al.* Cancer Immunosurveillance by Tissue-Resident Innate Lymphoid Cells and Innate-like T Cells. *Cell* 164, 365-377, doi:10.1016/j.cell.2016.01.002 (2016).

34 Flommersfeld, S. *et al.* Fate mapping of single NK cells identifies a type 1 innate lymphoid-like lineage that bridges innate and adaptive recognition of viral infection. *Immunity* 54, 2288-2304.e2287, doi:10.1016/j.immuni.2021.08.002 (2021).

35 Roan, F. *et al.* CD4+ Group 1 Innate Lymphoid Cells (ILC) Form a Functionally Distinct ILC Subset That Is Increased in Systemic Sclerosis. *J Immunol* 196, 2051-2062, doi:10.4049/jimmunol.1501491 (2016).

36 Bekiaris, V. *et al.* Human CD4+CD3- innate-like T cells provide a source of TNF and lymphotoxin-αβ and are elevated in rheumatoid arthritis. *J Immunol* 191, 4611-4618, doi:10.4049/jimmunol.1301672 (2013).

37 Björklund Å, K. *et al.* The heterogeneity of human CD127(+) innate lymphoid cells revealed by single-cell RNA sequencing. *Nat Immunol* 17, 451-460, doi:10.1038/ni.3368 (2016).

38 Moral, J. A. *et al.* ILC2s amplify PD-1 blockade by activating tissue-specific cancer immunity. *Nature* 579, 130-135, doi:10.1038/s41586-020-2015-4 (2020).

39 Hsu, J. *et al.* Contribution of NK cells to immunotherapy mediated by PD-1/PD-L1 blockade. *J Clin Invest* 128, 4654-4668, doi:10.1172/jci99317 (2018).

40 Davis-Marcisak, E. F. *et al.* Transfer learning between preclinical models and human tumors identifies a conserved NK cell activation signature in anti-CTLA-4 responsive tumors. *Genome Med* 13, 129, doi:10.1186/s13073-021-00944-5 (2021).

41 Kamiya, T., Seow, S. V., Wong, D., Robinson, M. & Campana, D. Blocking expression of inhibitory receptor NKG2A overcomes tumor resistance to NK cells. *J Clin Invest* 129, 2094-2106, doi:10.1172/jci123955 (2019).

42 Carrega, P. *et al.* Characterisation of innate lymphoid cell subsets infiltrating colorectal carcinoma. *Gut* 69, 2261-2263, doi:10.1136/gutjnl-2020-320908 (2020).

43 Wang, S. *et al.* Transdifferentiation of tumor infiltrating innate lymphoid cells during progression of colorectal cancer. *Cell Res* 30, 610-622, doi:10.1038/s41422-020-0312-y (2020).

44 Ikeda, A. *et al.* Human NKp44(+) Group 3 Innate Lymphoid Cells Associate with Tumor-Associated Tertiary Lymphoid Structures in Colorectal Cancer. *Cancer Immunol Res* 8, 724-731, doi:10.1158/2326-6066.Cir-19-0775 (2020).

45 Bernink, J. H. *et al.* Human type 1 innate lymphoid cells accumulate in inflamed mucosal tissues. *Nat Immunol* 14, 221-229, doi:10.1038/ni.2534 (2013).

46 Gao, Y. *et al.* Tumor immunoevasion by the conversion of effector NK cells into type 1 innate lymphoid cells. *Nat Immunol* 18, 1004-1015, doi:10.1038/ni.3800 (2017).

47 Hall, G. *et al.* Immunohistochemistry for PMS2 and MSH6 alone can replace a four antibody panel for mismatch repair deficiency screening in colorectal adenocarcinoma. *Pathology* 42, 409-413, doi:10.3109/00313025.2010.493871 (2010).

48 Höllt, T. *et al.* Cytosplore: Interactive Immune Cell Phenotyping for Large Single-Cell Datasets. 35, 171-180, doi:https://doi.org/10.1111/cgf.12893 (2016).

49 Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* 14, 865-868, doi:10.1038/nmeth.4380

(2017).

50 de Vries, N. L. *et al.* γδ T cells are effectors of immune checkpoint blockade in mismatch repair-deficient colon cancers with antigen presentation defects. *bioRxiv*, 2021.2010.2014.464229, doi:10.1101/2021.10.14.464229 (2021).

51 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888-1902 e1821, doi:10.1016/j.cell.2019.05.031 (2019).

52 McGinnis, C. S. *et al.* MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat Methods* 16, 619-626, doi:10.1038/s41592-019-0433-8 (2019).

53 van der Maaten, L. J. P. & Hinton, G. E. Visualizing high-dimensional data using t-SNE. *J. Mach. Learn. Res.* 9, 2579-2605 (2008).

54 McInnes, L., Healy, J. & Melville, J. J. a. p. a. Umap: Uniform manifold approximation and projection for dimension reduction. (2018).

SUPPLEMENTAL FIGURES

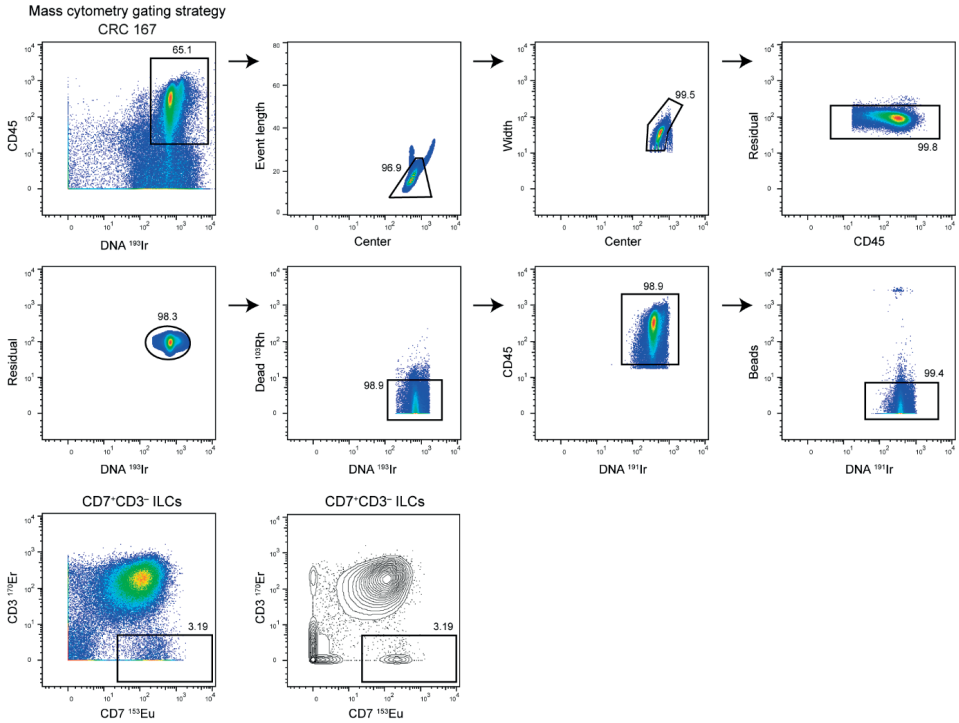


Figure S1. Mass cytometry gating strategy for the ILC population from MMR-deficient colorectal cancers. Mass cytometry gating strategy for single, live CD45$^+$CD7$^+$CD3$^-$ ILCs of a representative MMR-deficient colorectal cancer sample showing sequential gates with percentages.

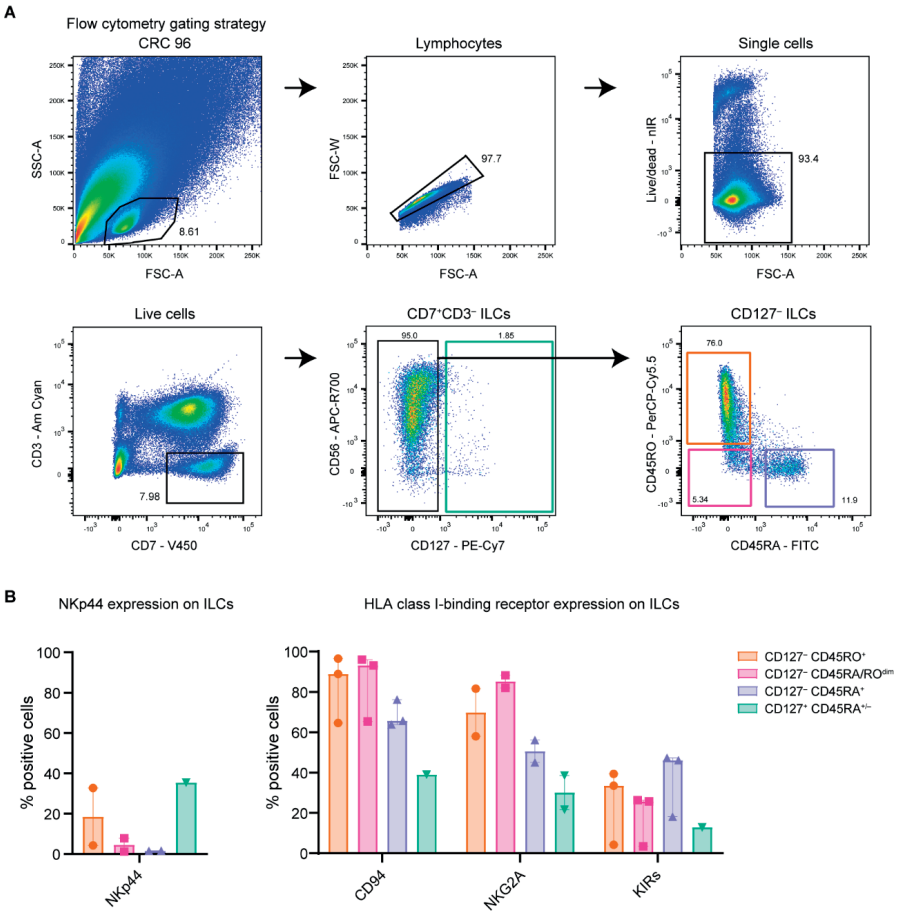
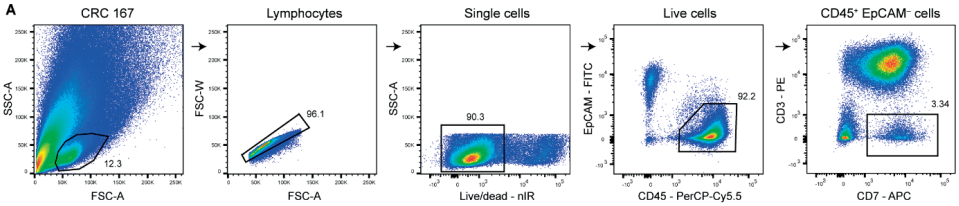
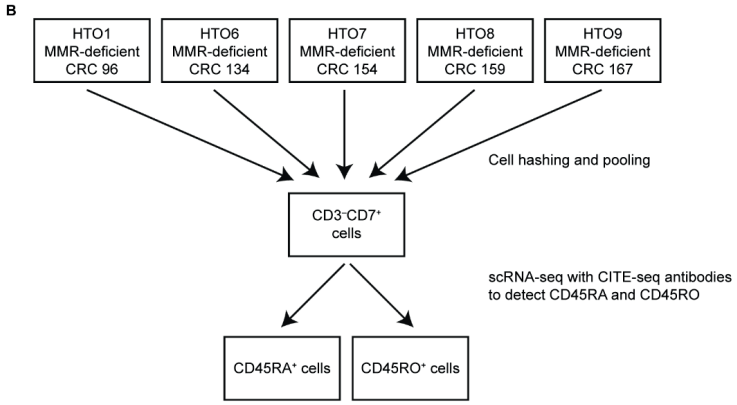


Figure S2. Flow cytometry gating strategy and frequencies of NKp44 and HLA class I-binding receptors on ILC subsets from MMR-deficient colorectal cancers.

A. Flow cytometry gating strategy for CD127$^-$CD45RO$^+$ ILC1-like (orange), CD127$^-$CD45RA/RO^{dim} ILC1-like (pink), CD127$^-$CD45RA$^+$ NK-like (purple), and CD127$^+$CD45RA$^{+/-}$ ILC2 (green) cells showing sequential gates with percentages. **B**. Frequencies of NKp44 and HLA class I-binding receptors among CD127$^-$CD45RO$^+$ ILC1-like (orange), CD127$^-$CD45RA/RO^{dim} ILC1-like (pink), CD127$^-$CD45RA$^+$ NK-like (purple), and CD127$^+$CD45RA$^{+/-}$ ILC2 (green) cells as percentage of total ILCs from MMR-deficient CRCs *ex vivo* (N=2-3) by flow cytometry. Marker frequencies for ILC2s (green) were only included when >100 cells were present. Bars indicate the median with IQR. Data from five independent experiments.





C

Patient ID	HTO1 CRC 96	HTO6 CRC 134	HTO7 CRC 154	HTO8 CRC 159	HTO9 CRC 167
Cell number at FACS sorting:	6662 cells	3307 cells	3226 cells	6140 cells	6656 cells
Cell number at scRNA-seq analysis:	1293 cells	228 cells	502 cells	1059 cells	1188 cells
Fraction in:	19.4%	6.9%	15.6%	17.2%	17.8%

Figure S3. Gating strategy and experimental design for single-cell RNA-sequencing of the ILC population from MMR-deficient colorectal cancers.

A. FACS gating strategy for single, live CD45$^+$EpCAM$^-$CD7$^+$CD3$^-$ ILCs of a representative MMR-deficient colorectal cancer sample showing sequential gates with percentages. **B**. Experimental design of the single-cell RNA-sequencing study with HTO- and CITE-sequencing antibodies performed on CD7$^+$CD3$^-$ cells isolated from MMR-deficient CRCs (N=5). **C**. Table showing the number of CD7$^+$CD3$^-$ cells isolated from MMR-deficient CRCs at the time of sorting *versus* the number of cells in the single-cell RNA-sequencing analysis, and the fraction thereof.

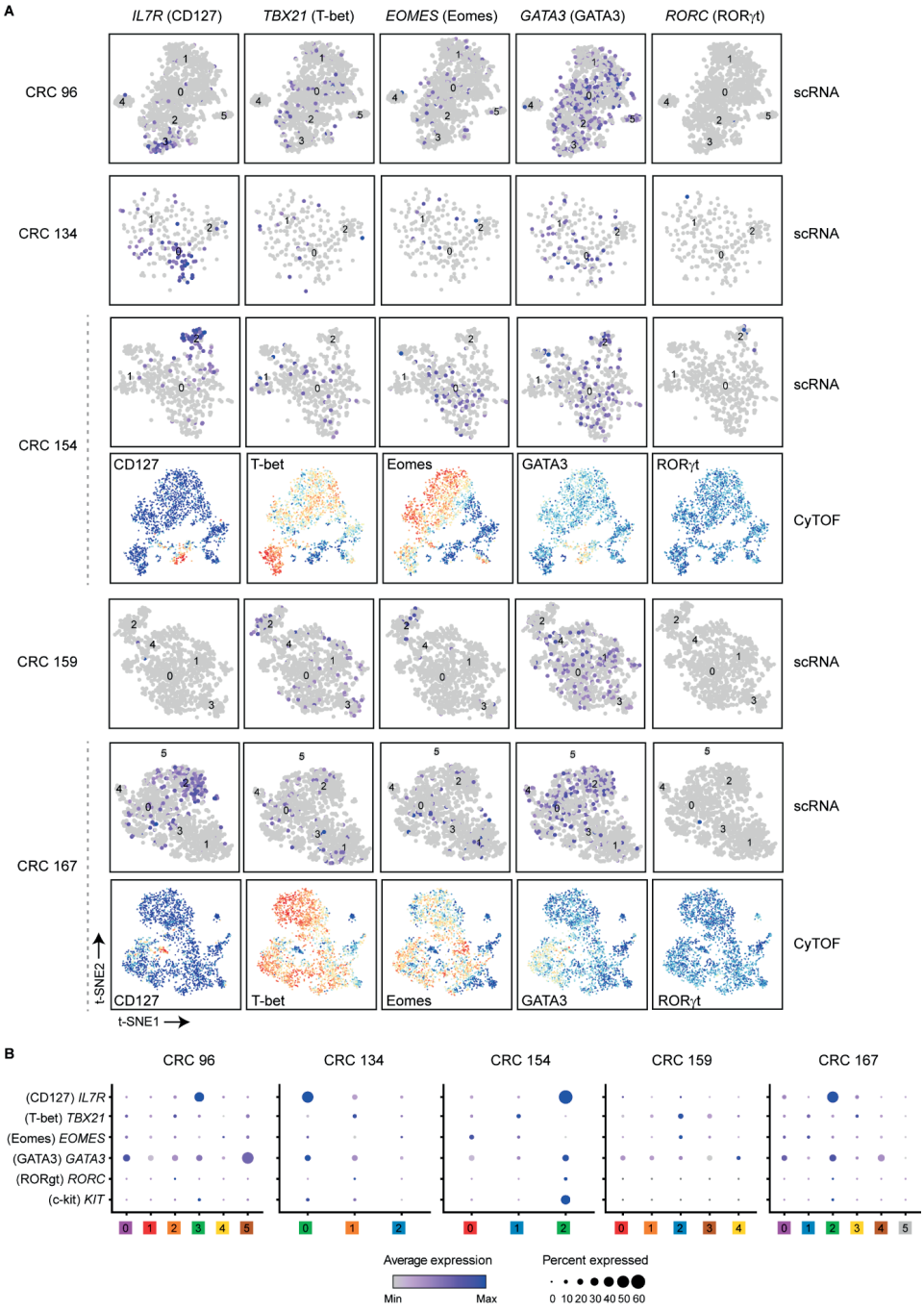
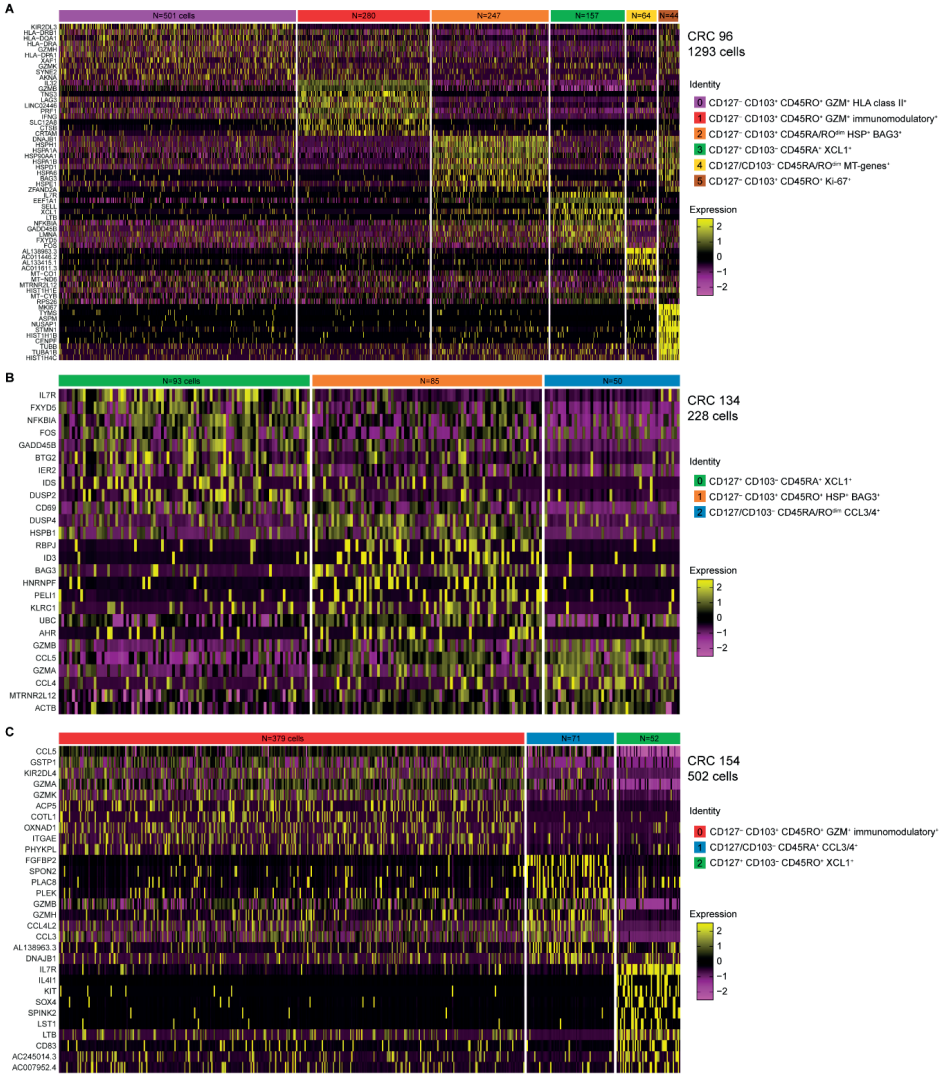


Figure S4. Transcription factor analysis of ILCs from MMR-deficient colorectal cancers by single-cell RNA-sequencing and mass cytometry.

A. t-SNE embedding showing the relative expression of indicated markers on CD7^{+}CD3^{-} ILCs isolated from MMR-deficient CRCs (N=5) analyzed by single-cell RNA-sequencing (scRNA; gene expression) per patient, and as compared

to mass cytometry (CyTOF; protein expression) for two paired samples. Colors represent relative expression of indicated genes (scRNA) and proteins (CyTOF). Each dot represents a single cell. **B**. Dot plots showing the relative expression of indicated genes in the clusters identified in **Figure 2**. Color intensity indicates average expression and size of the dot indicates the percentage of cells expressing the gene.



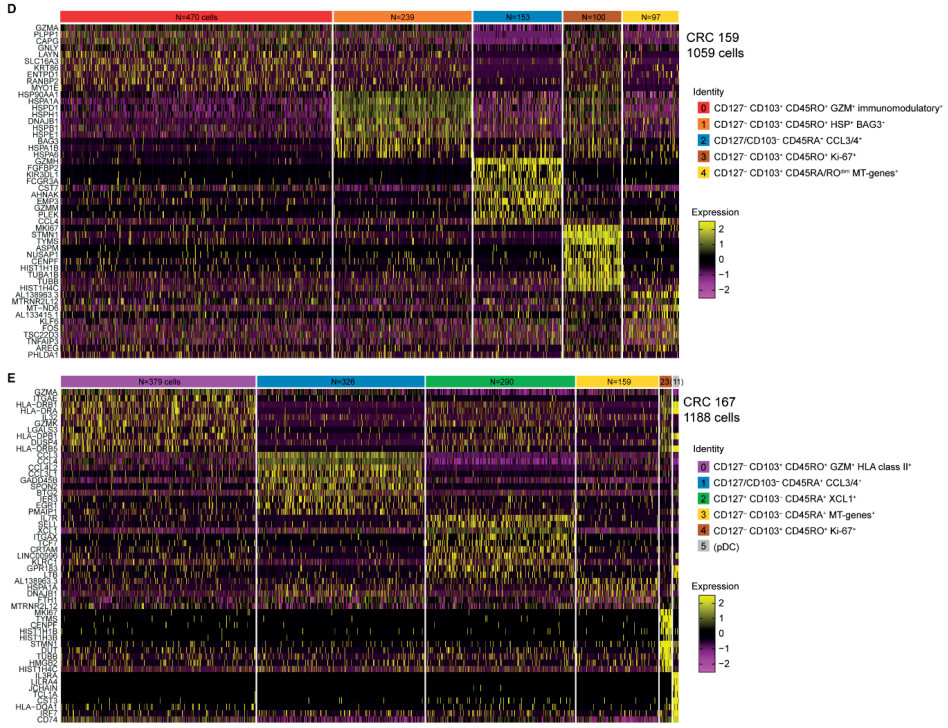


Figure S5. Distinct ILC clusters in MMR-deficient colorectal cancers by single-cell RNA-sequencing.
A. Heatmap showing the normalized single-cell gene expression value (z-score, purple-to-yellow scale) for the top 10 differentially expressed genes in each identified CD7^{+}CD3^{-} ILC cluster for MMR-deficient CRC 96 (N=1293 cells). Color bar at the top represents functionally distinct categories of ILC subsets. **B**. As (**A**), but for MMR-deficient CRC 134 (N=228 cells). **C**. As (**A**), but for MMR-deficient CRC 154 (N=502 cells). **D**. As (**A**), but for MMR-deficient CRC 159 (N=1059 cells). **E**. As (**A**), but for MMR-deficient CRC 167 (N=1188 cells).

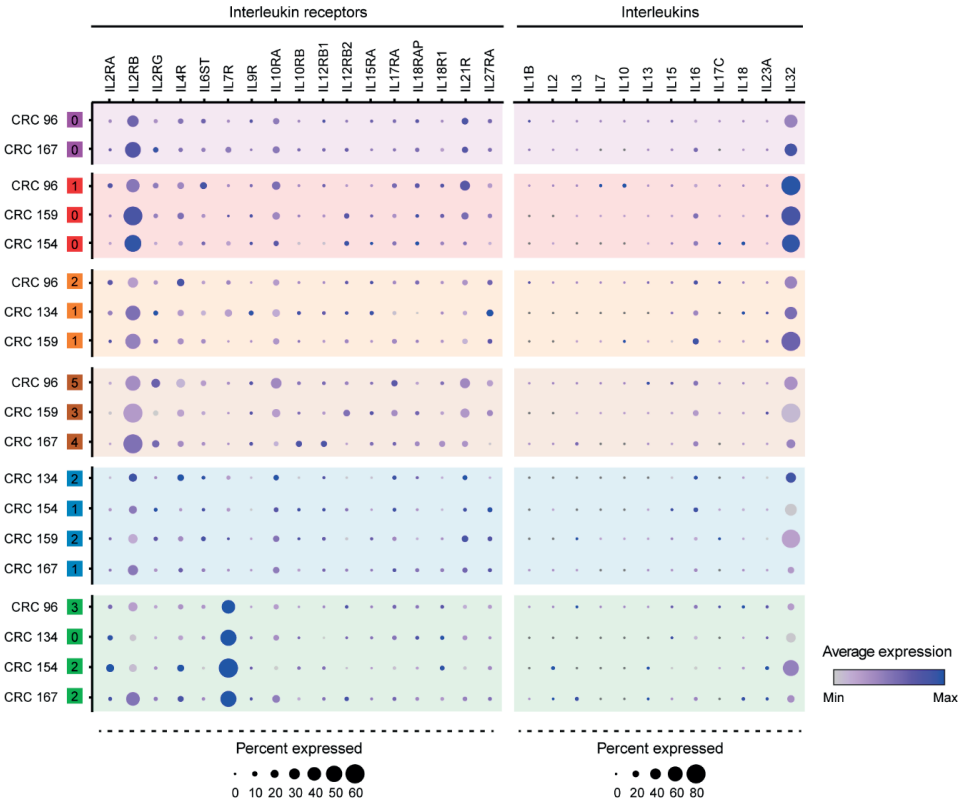
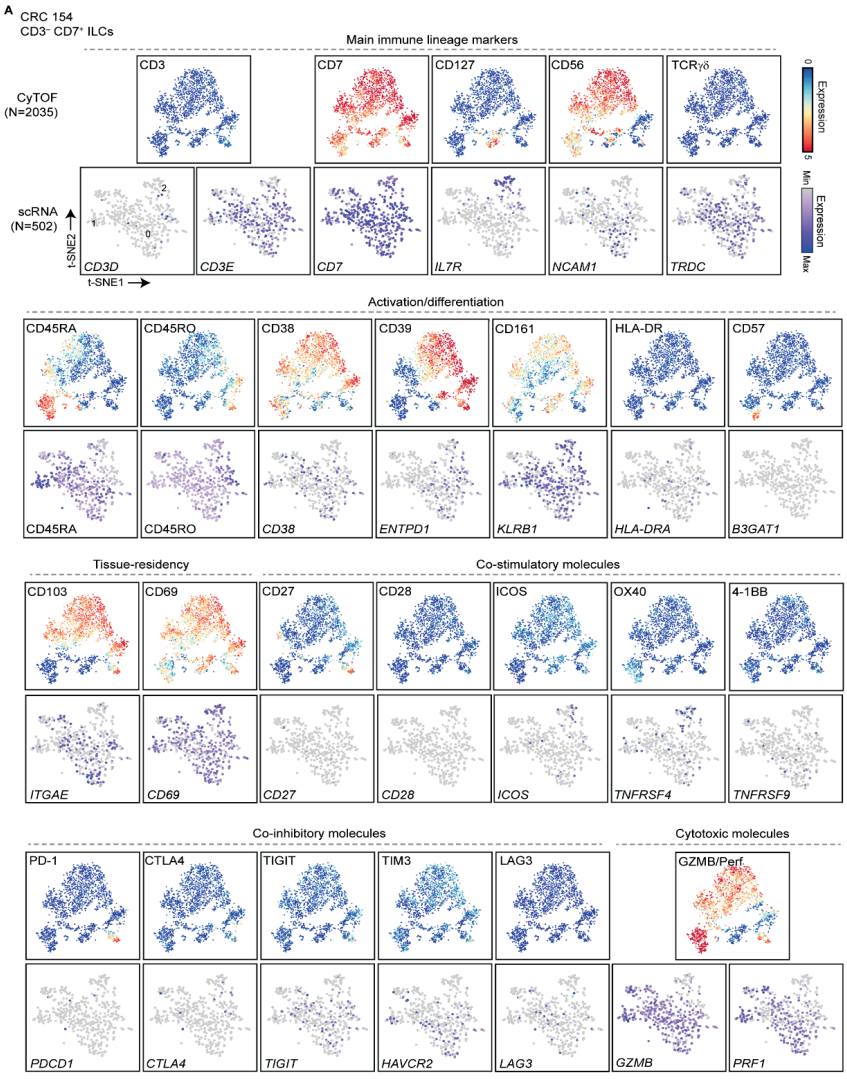


Figure S6. Transcriptional profiles of interleukin receptors and interleukins by ILCs from MMR-deficient colorectal cancers.

Dot plots showing the relative expression of indicated genes in the clusters identified in (**Figure 3A**). Color intensity indicates average expression and size of the dot indicates the percentage of cells expressing the gene. Background colors indicate functionally distinct categories of ILC subsets as in (**Figure 3A**). Interleukin receptor and interleukin transcripts with a median percentage of at least 1% of positive ILCs (N=4270) were included.


A
CRC 154
CD3⁻ CD7⁺ ILCs
Main immune lineage markers
CyTOF
(N=2035)
CD3
CD7
CD127
CD56
TCRγδ
Expression
0
5
scRNA
(N=502)
t-SNE2
t-SNE1
CD3D
CD3E
CD7
IL7R
NCAM1
TRDC
Min
Max
Activation/differentiation
CD45RA
CD45RO
CD38
CD39
CD161
HLA-DR
CD57
CD45RA
CD45RO
CD38
ENTPD1
KLRB1
HLA-DRA
B3GAT1
Tissue-residency
Co-stimulatory molecules
CD103
CD69
CD27
CD28
ICOS
OX40
4-1BB
ITGAE
CD69
CD27
CD28
ICOS
TNFRSF4
TNFRSF9
Co-inhibitory molecules
Cytotoxic molecules
PD-1
CTLA4
TIGIT
TIM3
LAG3
GZMB/Perf
PDCD1
CTLA4
TIGIT
HAVCR2
LAG3
GZMB
PRF1

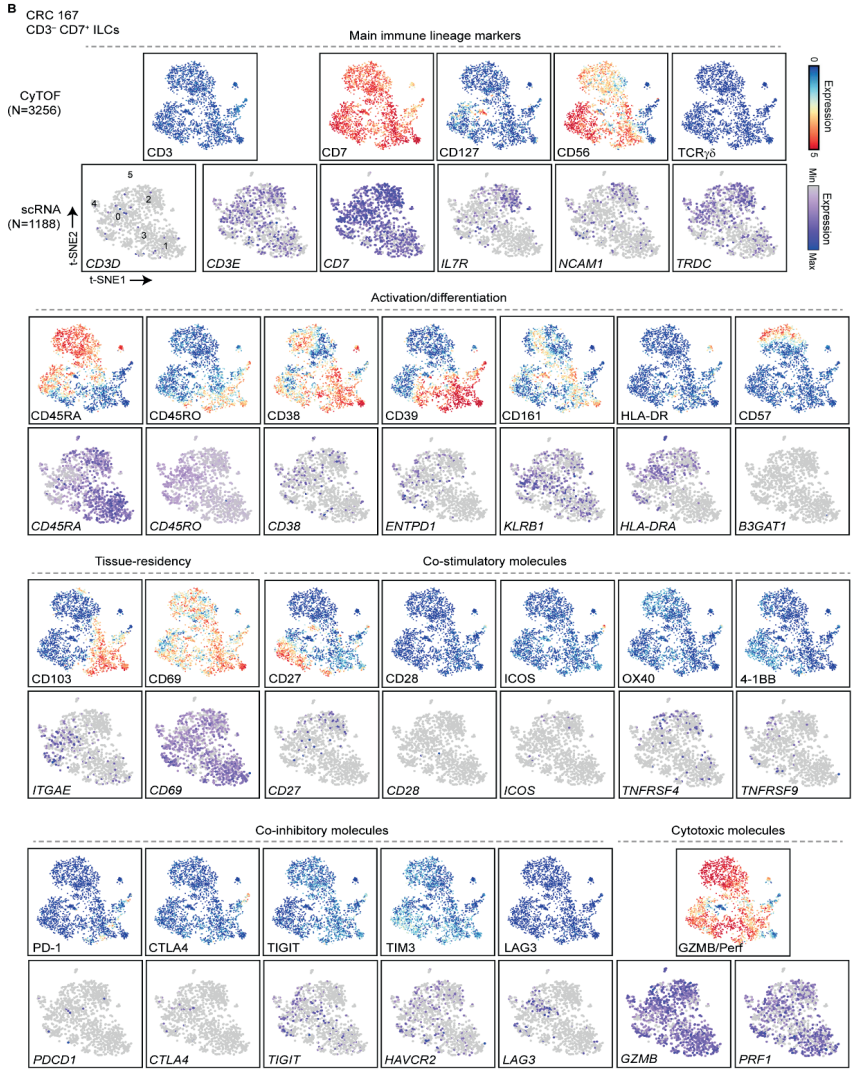


Figure S7. Mass cytometric *versus* transcriptional profiles of immune cell markers on ILCs in MMR-deficient colorectal cancers.

A. t-SNE embedding showing the relative expression of indicated immune cell markers on CD7⁺CD3⁻ ILCs analyzed by mass cytometry (CyTOF, N=2035) as compared to CD7⁺CD3⁻ ILCs analyzed by single-cell RNA-sequencing (scRNA, N=502) isolated from MMR-deficient CRC 154. Colors represent relative expression of indicated genes (scRNA) and proteins (CyTOF). Each dot represents a single cell. **B**. As (**A**), but for CD7⁺CD3⁻ ILCs analyzed by mass cytometry (CyTOF, N=3256) as compared to CD7⁺CD3⁻ ILCs analyzed by single-cell RNA-sequencing (scRNA, N=1188) isolated from MMR-deficient CRC 167.

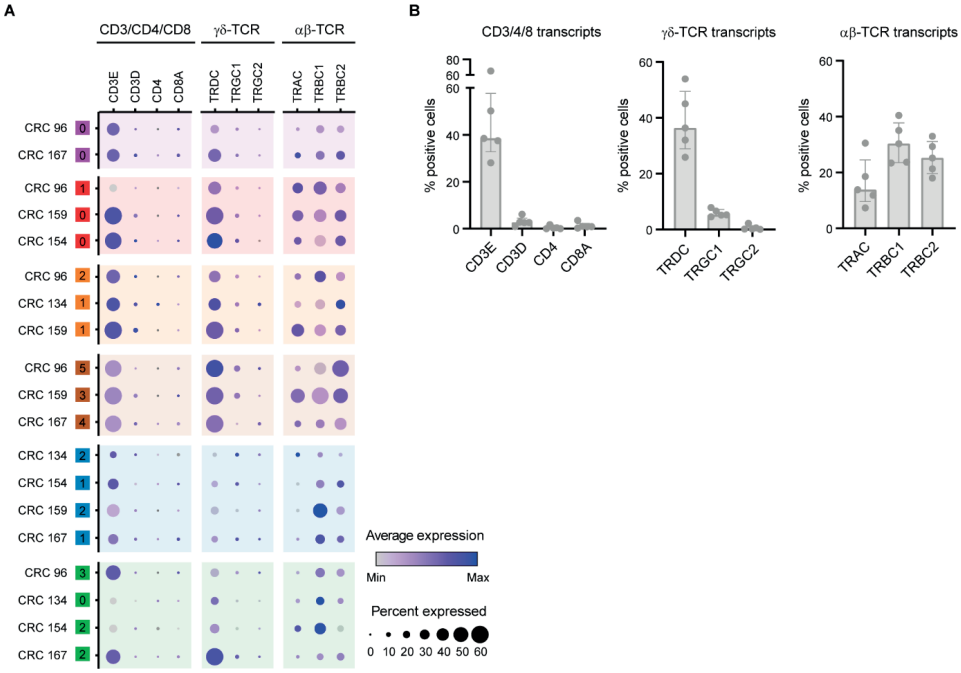


Figure S8. Transcriptional profiles of CD3 and T cell receptors by ILCs from MMR-deficient colorectal cancers. **A**. Dot plots showing the relative expression of indicated genes in the clusters identified in (**Figure 3A**). Color intensity indicates average expression and size of the dot indicates the percentage of cells expressing the gene. Background colors indicate functionally distinct categories of ILC subsets as in (**Figure 3A**). **B**. Frequencies of CD3, γδ-TCR, and αβ-TCR gene expression among all ILCs (N=4270) as percentage of positive cells from MMR-deficient CRCs (N=5). Bars indicate the median with IQR.

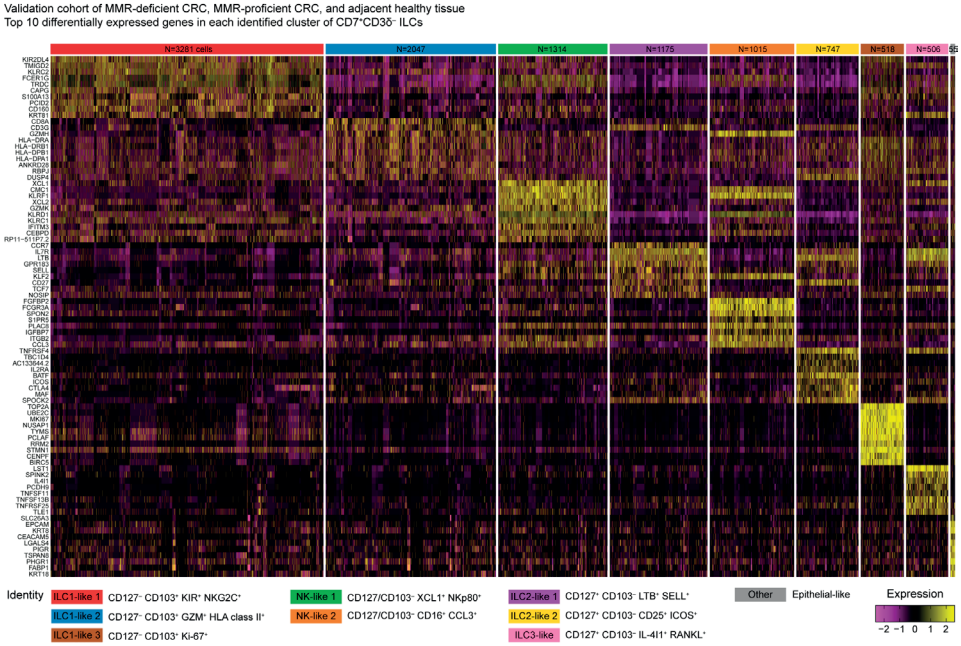


Figure S9. Distinct ILC clusters in MMR-deficient and -proficient colorectal cancers and adjacent healthy tissues by single-cell RNA-sequencing.

Heatmap showing the normalized single-cell gene expression value (z-score, purple-to-yellow scale) for the top 10 differentially expressed genes in each identified cluster from **Figure 4** of CD7^{+}CD3δ^{-} ILCs (N=10660) from MMR-deficient CRCs (N=35), MMR-proficient CRCs (N=29), and adjacent healthy tissues (N=36) obtained from the single-cell RNA-sequencing dataset of Pelka *et al.*[28].

SUPPLEMENTAL TABLES

The supplementary tables are available in the appendix to this thesis at the repository of Leiden University (https://hdl.handle.net/1887/3439882) and can be requested from the author.

Table S1: Characteristics of clinical samples from patients with MMR-deficient colorectal cancer included in the study.
Table S2: Antibodies used for mass cytometry experiments.

T cells with tumor-specific phenotypes largely originate from pericolic lymph nodes in colorectal cancer

6

Natasja L. de Vries[1,2], Joana Bras Gomes Nunes[1], Ruud van der Breggen[1], Frits Koning[2], Noel F.C.C. de Miranda[1]

[1]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands.
[2]Department of Immunology, Leiden University Medical Center, Leiden, The Netherlands.

Manuscript in preparation

ABSTRACT

Background

Cancer immunotherapies aiming at reactivating anti-tumor T cell responses have yielded significant clinical benefits. To further improve the success rate of such T cell-based immunotherapies, it is important to characterize and understand the origin and dynamics of anti-tumor T cell responses. Previously, we showed that colorectal cancer (CRC) tissues contained specific populations of tissue-resident activated $CD8^+$ and $CD4^+$ T cells. Here, we aimed to identify their T cell receptor (TCR) repertoire, clonal frequencies, and clonal overlap with paired non-malignant tissues of the same patients.

Methods

Distinct $CD8^+$ and $CD4^+$ T cell populations, defined by the expression of markers of activation (CD38, PD-1) and tissue-residency (CD103), along with their non-activated and non-tissue-resident counterparts, were isolated from seven patients with CRC and processed for bulk TCRβ sequencing. The TCR repertoire in colorectal tumors was subsequently compared with that of matched samples from adjacent healthy mucosa, pericolic lymph nodes, and peripheral blood.

Results

We found clonal enrichment in tissue-resident, activated $CD8^+$ and $CD4^+$ T cells from colorectal tumors, with the most frequent T cell clone presenting frequencies ranging between 10-29% for $CD8^+$ T cells and between 8-39% for $CD4^+$ T cells in the different tumors. TCR clonotypes of tissue-resident, activated $CD8^+$ and $CD4^+$ T cells from colorectal tumors showed most overlap with matched pericolic lymph node samples, while their non-activated counterparts showed frequent TCRβ clonal overlap with T cells from adjacent healthy mucosa.

Conclusion

The TCRβ clonal overlap of tissue-resident, activated T cells with pericolic lymph node samples suggests that lymph nodes are an important source of tumor-reactive $CD8^+$ and $CD4^+$ T cells. These findings should be extended to a larger cohort of patients with CRC.

INTRODUCTION

T cell-mediated adaptive immune responses play an important role in cancer immunosurveillance and have predictive value with respect to natural disease progression[1,2] and response to immunotherapy[3-5]. Cancer cells express antigens presented by HLA class I molecules that may activate anti-tumor CD8+ T cell responses.[6] Intratumoral CD8+ T cells are particularly abundant in patients with DNA mismatch repair (MMR)-deficient cancers bearing a high mutational load and, therefore, a vast burden of antigens derived from somatic mutations (neoantigens).[7,8] Cancer immunotherapies targeting the T cell repertoire, such as T cell checkpoint blockade with antibodies against PD-1/PD-L1 and/or CTLA-4 are proving to be effective at restoring T cell-mediated immune responses.[5,9-11] However, durable clinical responses are observed in only a minority of patients, highlighting the need to better understand the origin and dynamics of anti-tumor T cell responses.

In colorectal cancer (CRC), the presence of intratumoral CD8+ T cells is correlated with improved prognosis.[12,13] We previously identified tumor tissue-specific CD8+ T cell clusters in colorectal tumors.[14] These T cell populations were characterized by the co-expression of tissue-residency (CD103) and activation markers (CD38, CD39, and PD-1), and were infrequent in colorectal healthy mucosa, pericolic lymph nodes, and peripheral blood samples. Interestingly, co-expression of CD39 and CD103 was reported to identify tumor-reactive CD8+ T cells in different types of cancer, including CRC.[15-17] Of note, tumor tissue-specific CD4+ T cells with a similar activated tissue-resident phenotype were found in colorectal tumor samples.[14] Understanding the clonal relationships between such distinct CD8+ and CD4+ T cell populations and their tissue of origin might shed light on the dynamics of anti-tumor T cell responses. In αβ T cells, most receptor diversity is contained within the third complementarity-determining region (CDR3) of the T cell receptor (TCR) alpha and beta chains. Targeted sequencing of the CDR3 region can identify T cell clones, clonal frequencies, and clonal expansion, a characteristic of antigen specific responses. Clonal enrichment of T cell populations has been shown to correlate with durable responses to PD-1 blockade.[3,18,19]

In this study, we performed DNA sequencing of the CDR3 region of the TCR beta-chain (TCRβ) to determine the TCR repertoires of tissue-resident, activated CD8+ and CD4+ T cells in colorectal tumors, and subsequently compare those with total CD8+ and CD4+ T cells from matched samples of adjacent healthy mucosa, pericolic lymph nodes, and peripheral blood from seven patients with CRC.

RESULTS

TCR repertoire analysis indicates clonal enrichment in tissue-resident, activated CD8+ and CD4+ T cells from colorectal tumors

TCR repertoire analysis was performed on DNA isolated from different FACS-sorted T cell populations from colorectal tumors defined as activated (CD38+ PD-1+) and non-activated

(CD38$^-$ PD-1$^-$), tissue-resident (CD103$^+$) and non-tissue-resident (CD103$^-$) (**Figure S1**). Those TCR sequences were compared to the CD8$^+$ and CD4$^+$ TCR repertoire of matched adjacent healthy mucosa, pericolic lymph nodes, and peripheral blood samples of the same patients (**Figure S1**). A total of 10,374 unique clonotypes were detected in the different T cell subsets isolated from colorectal tumors, 7,653 in adjacent healthy mucosa, 34,607 in pericolic lymph nodes, and 36,562 in peripheral blood samples (**Table 1**).

The majority of CD8$^+$ T cells in colorectal tumors contained a tissue-resident activated phenotype (CD103$^+$ CD38$^+$ PD-1$^+$), while in adjacent healthy mucosa non-activated cells were most abundant (**Figure 1A**). In pericolic lymph nodes and peripheral blood samples, almost all CD8$^+$ T cells consisted of CD103$^-$ CD38$^-$ PD-1$^-$ cells (**Figure 1A**). The Shannon diversity index was lower in intratumoral CD8$^+$ T cell populations as compared to CD8$^+$ T cells from adjacent healthy mucosa, pericolic lymph nodes, and peripheral blood samples of the patients, indicating a low CD8$^+$ TCR repertoire diversity in the tumors (**Figure 1B**). Within colorectal tumors, highest clonality was observed for tissue-resident activated (CD103$^+$ CD38$^+$ PD-1$^+$) CD8$^+$ T cell populations (**Figure 1C**; details are provided in Methods). Enrichment of clonotypes was detected in tissue-resident activated (CD103$^+$ CD38$^+$ PD-1$^+$) CD8$^+$ T cell populations from all patients, with the most frequent T cell clone presenting frequencies ranging between 10-29% in the different tumors (**Table 2**, **Table S1**).

Table 1. Overview of total numbers of cells, reads, and unique clonotypes detected in the clinical samples.

Sample	Colorectal tumor (N=7)			Adjacent healthy mucosa (N=7)**			Lymph node (N=7)**			PBMC (N=7)**		
Population	**All**	**CD8**	**CD4**	**All**	**CD8**	**CD4**	**All**	**CD8**	**CD4**	**All**	**CD8**	**CD4**
Total cells	149,975	58,841	91,134	102,137	45,878	56,259	1,076,387	106,207	970,180	2,866,060	921,506	1,944,554
Total reads	1,7596,118	7,127,674	10,468,444	8,319,137	3,740,426	3,073,885	15,847,118	5,508,073	5,944,784	16,531,234	5,586,063	7,593,642
Total unique clonotypes*	10,374	2057	8334	7653	1243	3009	34,607	3069	22,137	36,562	4826	26,875

*Defined as distinct CDR3β amino acid sequence. **For two samples, cells were not subdivided into $CD8^+$ and $CD4^+$ T cell populations as DNA was isolated directly from fresh-frozen tissue or EDTA blood.*

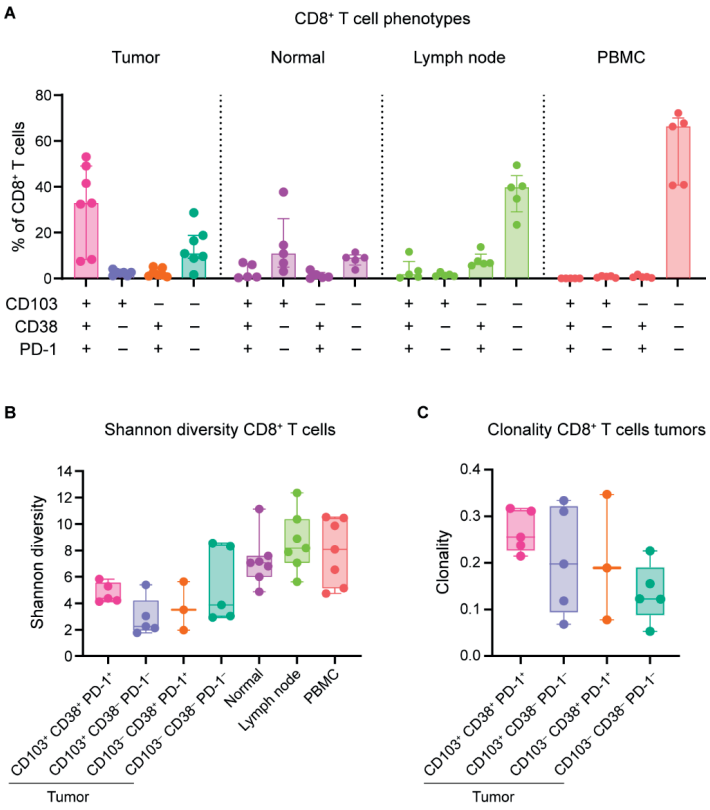


Figure 1. TCR repertoire analysis indicates clonal enrichment in tissue-resident, activated CD8⁺ T cells from colorectal tumors.

A. Frequencies of different CD8⁺ T cell populations among colorectal tumors, adjacent healthy mucosa ("normal"), pericolic lymph nodes, and peripheral blood samples from patients with CRC (N=7) as percentage of total CD8⁺ T cells by flow cytometry. Bars indicate median with IQR. Each dot represents an individual sample. **B**. Shannon diversity of different CD8⁺ T cell populations from colorectal tumors and CD8⁺ T cells from matched adjacent healthy mucosa ("normal"), pericolic lymph nodes, and peripheral blood samples from patients with CRC (N=7) by TCRβ sequencing. Box indicates median with IQR, whiskers min to max. Each dot represents an individual sample. **C**. Clonality of different CD8⁺ T cell populations from colorectal tumors from patients with CRC (N=7) by TCRβ sequencing. Details are provided in Methods. Box indicates median with IQR, whiskers min to max. Each dot represents an individual sample. Data from seven independent experiments.

Table 2. Top ten most frequent CDR3β sequences in tissue-resident, activated CD8⁺ T cells from colorectal tumors.

	CRC 164		CRC 181		CRC 220		CRC 102*		CRC 223*	
Rank	CDR3β AA sequence	Frequency	CDR3β AA sequence	Frequency	CDR3β AA sequence	Frequency	CDR3β AA sequence	Frequency	CDR3β AA sequence	Frequency
1	SARDPTSARQY	23.7%	ASSESTIGTQY	28.6%	ASSSRTSGGVGSYNEQF	26.9%	SARTSGGGETQY	10.4%	ASSLATGPNTEAF	13.8%
2	SAWDSPRGPRKLF	12.0%	ASSLREAGLAGGPFSYEQY	15.7%	ASRFLAGGHNEQF	14.5%	ASRKTLGGVRYGYT	7.6%	ASSLTVRYEQY	8.1%
3	ASSKGYGYT	8.4%	ASSATAGSNQPQH	10.5%	ASSWQGPNNEQF	5.5%	ASSLQSRDGYT	6.5%	ASSYTGSTYEQY	6.3%
4	ASRGQEAF	7.6%	ASSFLNEEEQY	6.0%	ASSQEDPYYEQY	3.1%	ASSRLAGRGEQF	6.4%	ASSLVGGPYEQY	6.1%
5	ASSTSGRGYEQY	6.1%	ASIITDSQETQY	3.3%	ASSSQGVEQY	3.0%	ASSYSGRGGDEQF	6.3%	ASSQDRGSYEQY	4.3%
6	ASSLIGSREQY	3.9%	ASIQQGNNQPQH	2.0%	SVGAGGTNEKLF	2.6%	ASSTGRSYGYT	5.4%	ASSFPSAAQY	4.2%
7	ASSLVSAHTNEQF	3.7%	ASSQEEGRQYGYT	2.0%	ASSMSGNYEQY	2.5%	SVKGGYEQY	4.3%	ASSLGWSYEQY	3.9%
8	ASSPTWYSGANVLT	3.5%	ASSQDRPLGAGTLPDTQY	2.0%	ASRGLAEKYNEQF	2.3%	ASSGGAYEQY	2.6%	ASSPDRDVYEQY	3.9%
9	ASSLTKRWGNNGETQY	3.1%	ASSNREGTEAF	1.6%	ASSPTTVGNQPQH	2.0%	ASQNGPGTEQF	2.5%	ASSQAGQPQH	3.3%
10	ASQPWASGAYEQY	2.5%	ASSPFNTEAF	1.6%	ASSMYADEQF	1.9%	ASSAGGVEQY	2.3%	ASSGGDLIANEQF	3.0%

**MMR-deficient colorectal tumors.*

In line with previous observations,[14] CD103$^+$ CD4$^+$ T cells were less prevalent in colorectal tumors as compared to their CD8 counterpart (**Figure 2A**). Tissue-resident activated (CD103$^+$ CD38$^+$ PD-1$^+$) CD4$^+$ T cells were exclusively found in colorectal tumor samples (**Figure 2A**). Within CD103$^-$ CD4$^+$ T cells, both activated (CD38$^+$ PD-1$^+$) and non-activated (CD38$^-$ PD-1$^-$) populations were abundant in CRC tissues (**Figure 2A**). In adjacent healthy mucosa, pericolic lymph nodes, and peripheral blood samples, CD4$^+$ T cells lacking CD103, CD38, and PD-1 were most prevalent (**Figure 2A**). Intratumoral CD4$^+$ T cells showed a low Shannon diversity as compared to matched samples from adjacent healthy mucosa, pericolic lymph nodes, and peripheral blood of the patients (**Figure 2B**). In line with the intratumoral CD8$^+$ T cells, tissue-resident activated (CD103$^+$ CD38$^+$ PD-1$^+$) CD4$^+$ T cells showed the highest clonality in colorectal tumors (**Figure 2C**; details are provided in Methods). Clonal enrichment was observed in tissue-resident activated (CD103$^+$ CD38$^+$ PD-1$^+$) CD4$^+$ T cell populations from all patients, with the most frequent T cell clone presenting frequencies ranging between 8-39% in the different tumors (**Table 3**, **Table S2**). Of note, CD103$^+$ CD38$^-$ PD-1$^-$ CD4$^+$ T cells were not detected in such frequencies to allow TCRβ sequencing.

We next examined whether homology could be observed within CDR3β sequences detected in the intratumoral T cells. For CD8$^+$ T cells, we found CDR3β sequences that were shared by two different colorectal tumors (**Table 4**). The first of those CDR3β amino acid sequences has been reported in dysfunctional CD8$^+$ T cells in hepatocellular carcinoma.[20] Interestingly, we detected a CDR3β sequence that was shared by CD4$^+$ T cells in three different colorectal tumors (**Table 5**). The CDR3β amino acid sequence had frequencies ranging from 0.11 – 2.26% in the tumors (**Table 5**). This specific sequence has been reported in effector (CD57$^+$CD28$^-$) CD4$^+$ T cells in fibrosing mediastinitis.[21]

Tissue-resident, activated CD8$^+$ and CD4$^+$ T cells in colorectal tumors show most overlapping TCR clonotypes with pericolic lymph nodes

The CDR3β sequences detected in intratumoral CD8$^+$ and CD4$^+$ T cell populations were subsequently compared to the CDR3β repertoires of CD8$^+$ and CD4$^+$ T cells from matched samples of adjacent healthy mucosa, pericolic lymph nodes, and peripheral blood to investigate TCRβ clonal overlap. An average of 19% of the T cell clones were shared between the intratumoral tissue-resident activated (CD103$^+$ CD38$^+$ PD-1$^+$) CD8$^+$ T cells and CD8$^+$ T cells from pericolic lymph node samples (**Figure 3A**, **Figure S2**). This indicates a strong clonal relationship between tissue-resident, activated CD8$^+$ T cells and CD8$^+$ T cells from the lymph nodes. In contrast, the non-activated counterpart (CD103$^+$ CD38$^-$ PD-1$^-$) showed frequent TCRβ clonal overlap with CD8$^+$ T cells from healthy mucosal tissue (**Figure 3A**, **Figure S2**), indicating that such T cells with intraepithelial phenotypes migrate from healthy mucosal tissue to the tumor tissue but do not become activated. For the non-tissue-resident, non-activated (CD103$^-$ CD38$^-$ PD-1$^-$) CD8$^+$ T cells, most TCRβ clonal overlap was found (shared) with CD8$^+$ T cells from blood samples (**Figure 3A**, **Figure S2**). Last, CD103$^-$ CD38$^+$ PD-1$^+$ CD8$^+$ T cells had an equivalent distribution of TCRβ clonotypes with CD8$^+$ T cells from adjacent mucosa, lymph node, and peripheral blood samples (**Figure 3A**, **Figure S2**).

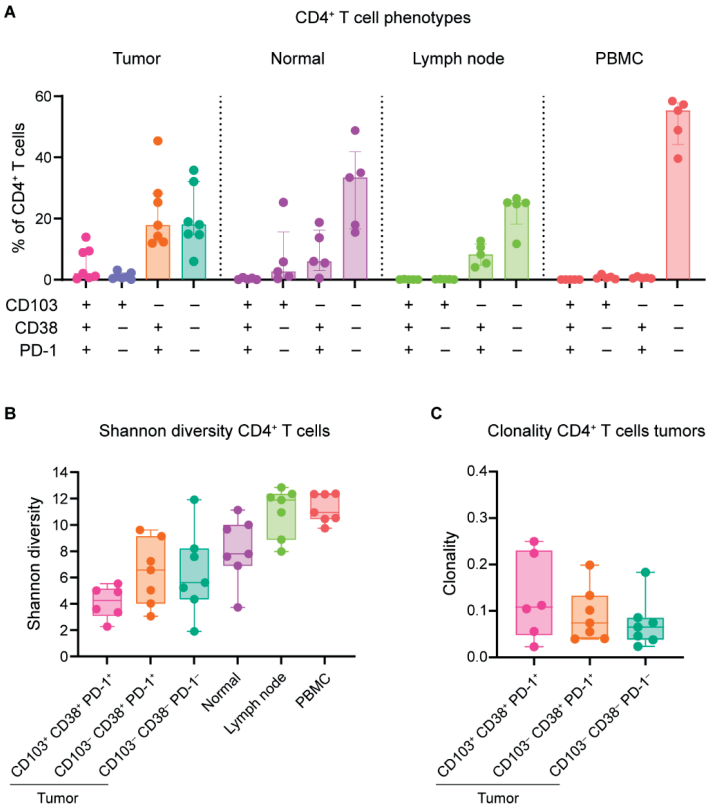


Figure 2. TCR repertoire analysis indicates clonal enrichment in tissue-resident, activated CD4⁺ T cells from colorectal tumors.

A. Frequencies of different CD4⁺ T cell populations among colorectal tumors, adjacent healthy mucosa ("normal"), pericolic lymph nodes, and peripheral blood samples from patients with CRC (N=7) as percentage of total CD4⁺ T cells by flow cytometry. Bars indicate median with IQR. Each dot represents an individual sample. **B**. Shannon diversity of different CD4⁺ T cell populations from colorectal tumors and CD4⁺ T cells from matched adjacent healthy mucosa ("normal"), pericolic lymph nodes, and peripheral blood samples from patients with CRC (N=7) by TCRβ sequencing. **C**. Clonality of different CD4⁺ T cell populations from colorectal tumors from patients with CRC (N=7) by TCRβ sequencing. Details are provided in **Methods**. **B,C**. Box indicates median with IQR, whiskers min to max. Each dot represents an individual sample. CD103⁺ CD38⁻ PD-1⁻ CD4⁺ T cells were not detected in such frequencies to allow TCRβ sequencing. Data from seven independent experiments.

Table 3. Top ten most frequent CDR3β sequences in tissue-resident, activated CD4⁺ T cells from colorectal tumors.

	CRC 44*		CRC 179**		CRC 181		CRC 220		CRC 102***		CRC 223***	
Rank	CDR3β AA sequence	Frequency	CDR3β AA sequence	Frequency	CDR3β AA sequence	Frequency	CDR3β AA sequence	Frequency	CDR3β AA sequence	Frequency	CDR3β AA sequence	Frequency
1	ASSLRQGGDSPLH	12.1%	ASSAGTGDYGYT	24.7%	SASGLAGSLPHEQY	12.6%	ASSPHGQGARVGYT	8.4%	ASSYSAYSNQPQH	17.1%	ASSFRETNTEAF	38.8%
2	ATSSPSADTQY	11.5%	SARSREVTYEQY	22.0%	ASSPGQGRPEAF	7.8%	ASSLNQGEQY	7.6%	ASSLGGTGSTDTQY	10.2%	ATSRALREGLKTQY	7.9%
3	ATSASQGYTDTQY	10.9%	ASSLGKERGNQPQH	21.8%	SASGAGQPQH	6.2%	ATSETGYTDTQY	6.7%	SLRQDLEGYT	8.7%	ASRFGTSGIYEQY	6.9%
4	ASSFGLAGGPTGTDTQY	10.2%	SARAPGVAGSYEQY	21.4%	ASGRGVGQETQY	5.7%	ASSDVWTANEKLF	6.2%	ASSLDSNQPQH	8.0%	ASSLDGRYEQY	5.9%
5	ASRSGNWGYT	9.0%	ASTGETEAF	10.2%	SEGLGDNTQY	4.4%	ASQSFGTSYNSPLH	3.2%	SAKQGADQPQH	3.7%	ASGGYGYT	5.5%
6	ASFESGANVLT	8.4%			SARGLLAGGTPATDTQY	4.2%	ASSTGTSDGELF	3.0%	ASSSEAGWGNSPLH	3.4%	ASMQSSGANVLT	4.3%
7	ASSSGGYGYT	7.4%			ASSVYRENIQY	3.3%	SAYPEGQSYEQY	2.5%	ASSLDKYGYT	3.3%	ASSEAGGGAYTGELF	4.0%
8	ASSYSTSGGVPETQY	7.4%			ASSSGGISTKNIQY	3.0%	SARMLAAPTDTQY	2.1%	SARVQTGRSYGYT	2.9%	ASSTPSGGRGNTIY	3.8%
9	ASIPGQGLFSYNSPLH	7.1%			ASSYGTGTGRYGYT	2.5%	ASSRDSNQPQH	2.0%	ASSQGGDSPLH	2.3%	ASSYLGTGSHEQF	3.6%
10	ASSLGTADTQY	7.1%			SPGRPGTEAF	2.5%	ASSLPGTYPPDTQY	2.0%	ASSLGRSYNSPLH	2.2%	ASSLWSSTDTQY	2.8%

**For CRC 44, a total of 14 clonotypes were detected. **For CRC 179, a total of five clonotypes were detected. ***MMR-deficient colorectal tumors.*

Table 4. Shared CDR3β sequences by CD8⁺ T cells from different colorectal tumors.

	CRC 220		CRC 102*			CRC 223*		
CDR3β AA sequence	Frequency	CDR3β Vβ gene	CDR3β AA sequence	Frequency	CDR3β Vβ gene	CDR3β AA sequence	Frequency	CDR3β Vβ gene
ASSLGETQY	0.27%	TRBV7-7				ASSLGETQY	0.21%	TRBV5-6
ASSPGYEQY	0.25%	TRBV7-6	ASSPGYEQY	0.003%	TRBV12-3			

**MMR-deficient colorectal tumors.*

Table 5. Shared CDR3β sequences by CD4⁺ T cells from different colorectal tumors.

CRC 179			CRC 220			CRC 102*		
CDR3β AA sequence	**Frequency**	**CDR3β Vβ gene**	**CDR3β AA sequence**	**Frequency**	**CDR3β Vβ gene**	**CDR3β AA sequence**	**Frequency**	**CDR3β Vβ gene**
ASSRDSNQPQH	2.26%	TRBV6-1	ASSRDSNQPQH	1.95%	TRBV27	ASSRDSNQPQH	0.11%	TRBV7-2
			ASSLGGGTYNEQF	0.27%	TRBV28	ASSLGGGTYNEQF	0.01%	TRBV7-9
			ASSLGQGYTEAF	0.15%	TRBV4-3	ASSLGQGYTEAF	0.03%	TRBV11-3
			ASSRTASYEQY	0.43%	TRBV19	ASSRTASYEQY	0.10%	TRBV6-6
ASSRTGVYGYT	0.88%	TRBV6-1				ASSRTGVYGYT	0.03%	TRBV28
			ASSSWTGSYGYT	0.11%	TRBV5-1	ASSSWTGSYGYT	0.04%	TRBV27
SVSTNTGELF	1.20%	TRBV29-1				SVSTNTGELF	0.01%	TRBV29-1

**MMR-deficient colorectal tumor.*

Analysis of the CD4⁺ TCR repertoire of colorectal tumors *versus* paired tissues revealed that, in line with their CD8 counterparts, the tissue-resident activated (CD103⁺ CD38⁺ PD-1⁺) CD4⁺ T cells showed most TCRβ clonal overlap with pericolic lymph node samples (an average of 17%; **Figure 3B**, **Figure S3**). Within the CD103⁻ CD4⁺ T cells, minimal overlap in CDR3β repertoire was observed with CD4⁺ T cells from adjacent healthy mucosa, pericolic lymph nodes, or peripheral blood samples (**Figure 3B**, **Figure S3**). Activated (CD38⁺ PD-1⁺) populations shared little TCRβ clonal overlap with CD4⁺ T cells from lymph node and blood samples, and non-activated (CD38⁻ PD-1⁻) populations with CD4⁺ T cells from peripheral blood (**Figure 3B**, **Figure S3**). Of note, the large majority of TCRβ clonotypes detected in tissue-resident activated (CD103⁺ CD38⁺ PD-1⁺) CD8⁺ and CD4⁺ T cells in colorectal tumors did not share overlap with matched healthy mucosal tissue (**Figure 3A,B**).

Last, with the emerging data on the value of investigating co-expression of CD39 and CD103 as markers for tumor-reactive T cells,[15-17] we examined the expression levels of CD39 within the different T cell populations and found that CD39 was highly expressed by tissue-resident activated (CD103⁺ CD38⁺ PD-1⁺) CD8⁺ and CD4⁺ T cells (**Figure S4**). In addition, CD38 or PD-1 single-positive CD103⁺ CD8⁺ and CD4⁺ T cells showed expression of CD39. In contrast to CD8⁺ T cells, CD39 expression could also be observed on CD38⁺ PD-1⁺ CD103⁻ CD4⁺ T cells and CD38 or PD-1 single-positive CD103⁻ CD4⁺ T cells (**Figure S4**). These observations suggest that CD38 or PD-1 are better surrogates for tumor-reactivity than CD103 in the case of CD4⁺ T cells.

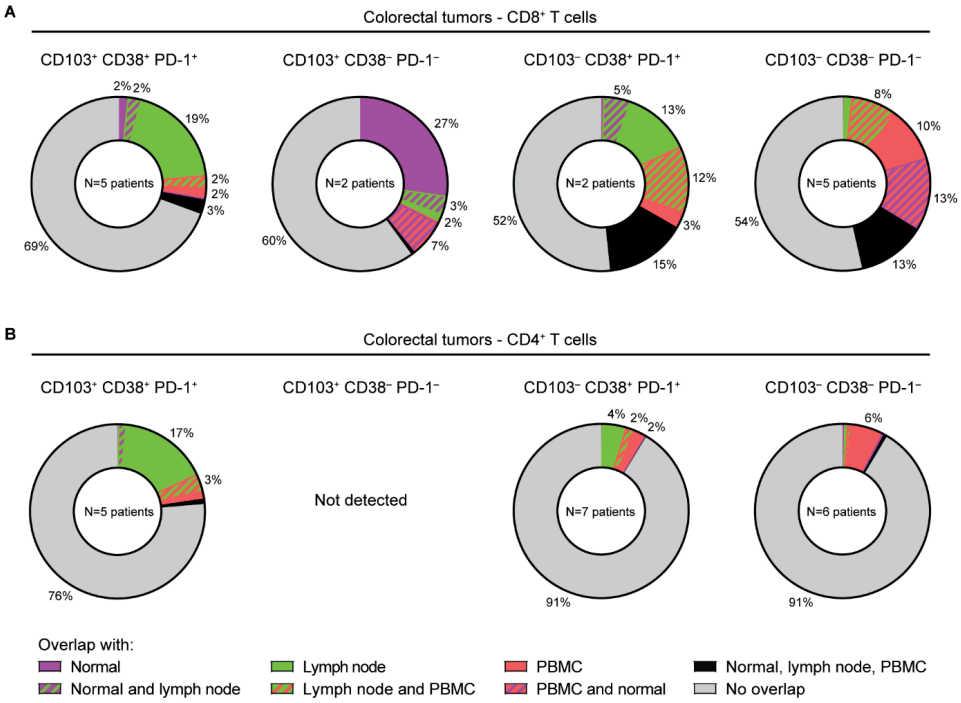


Figure 3. Tissue-resident, activated CD8⁺ and CD4⁺ T cells show most overlapping TCR clonotypes with pericolic lymph nodes.
A. Circular graphs showing the average frequency of overlapping TCR clonotypes of different CD8⁺ T cell populations from colorectal tumors with matched adjacent healthy mucosa ("normal"), pericolic lymph nodes, and peripheral blood samples from patients with CRC. **B**. Circular graphs showing the average frequency of overlapping TCR clonotypes of different CD4⁺ T cell populations from colorectal tumors with matched adjacent healthy mucosa ("normal"), pericolic lymph nodes, and peripheral blood samples from patients with CRC. CD103⁺ CD38⁻ PD-1⁻ CD4⁺ T cells were not detected in the tumor samples. **A,B**. Samples were only included if ≥ 10 clonotypes were detected. Data from seven independent experiments.

DISCUSSION

The field of cancer immunotherapy is rapidly developing and there is an emerging interest in (combinational) T cell checkpoint blockade immunotherapies designed to reactivate intratumoral T cells in cancers.[22,23] Unraveling the dynamic behavior of the intratumoral TCR repertoire and their capacity to recognize the tumor will be highly valuable for the understanding of anti-tumor T cell responses and ultimately for the development of novel T cell-mediated immunotherapeutic strategies. It is becoming increasingly evident that tissue-resident (CD103⁺) cells can be very abundant in CRC tissues, and have the capacity to recognize tumor cells.[24] However, it was shown that colorectal tumors can also contain so called 'bystander' T cells lacking tumor reactivity (with specificities for viruses such as EBV and CMV).[16] We and others previously identified tumor tissue-specific CD8⁺ and CD4⁺ T cell populations with an activated (CD38⁺, PD-1⁺, CD39⁺) tissue-resident (CD103⁺) phenotype in CRC.[14-16] These tissue-resident activated T cells were particularly enriched in MMR-deficient

CRCs.[14] Here, we specifically investigated the TCR repertoire, clonal frequencies, and clonal overlap with paired tissues of i) tissue-resident, activated (CD103$^+$ CD38$^+$ PD-1$^+$), ii) tissue-resident, non-activated (CD103$^+$ CD38$^-$ PD-1$^-$), iii) non-tissue-resident, activated (CD103$^-$ CD38$^+$ PD-1$^+$), and iv) non-tissue-resident, non-activated (CD103$^-$ CD38$^-$ PD-1$^-$) CD8$^+$ and CD4$^+$ T cells in colorectal tumors by bulk TCRβ sequencing. We showed that clonal enrichment was most frequently observed in the tissue-resident activated CD8$^+$ and CD4$^+$ T cell populations as compared to the other T cell populations, suggesting the existence of antigenic responses.

It has been proposed that the interaction of intratumoral T cells with tumor antigens in the tumor microenvironment can shape their phenotype toward an 'exhausted' cell state accompanied by the expression of different phenotypic markers depending on the cancer type.[25] Co-expression of CD39 and CD103 could function as markers for tumor-reactive CD8$^+$ T cells in different types of solid cancers including CRC.[15-17] In addition, expression of PD-1 and/or CD137 have been reported as markers for tumor-reactive T cells in melanoma and ovarian cancer.[25-29] However, the aforementioned markers not always separated tumor-reactive CD8$^+$ T cells from bystander T cells.[24] In our study, highest clonality was observed in the intratumoral CD103$^+$ CD38$^+$ PD-1$^+$ CD8$^+$ T cell populations, suggesting that tissue-residency and activation markers as CD103, CD38, PD-1 (and CD39) may serve as surrogates for tumor-reactive CD8$^+$ T cells. In contrast to CD8$^+$ T cells, phenotypic characterization of markers to identify tumor-reactive CD4$^+$ T cells are largely lacking. It was reported that CD39$^-$ CD4$^+$ T cells showed specificity for cancer-unrelated antigens (e.g. CMV epitopes).[30] However, no direct evidence has linked expression of CD39 to tumor-reactive CD4$^+$ T cells as of yet. In our data, clonal enrichment was mainly observed in the CD103$^+$ CD38$^+$ PD-1$^+$ CD4$^+$ T cell subset in colorectal tumors. This may suggest that tumor reactivity can be associated with expression of markers of tissue-residency (CD103) and activation/dysfunction (CD38, PD-1, CD39) for CD4$^+$ T cells in CRC. To test whether those CDR3β sequences of the clonally expanded CD103$^+$ CD38$^+$ PD-1$^+$ CD8$^+$ and CD4$^+$ T cells have the capacity to recognize the tumor, identified TCRs could be introduced into donor T cells and tested against the original tumor.

In our study, the large majority of TCRβ repertoires of intratumoral CD103$^+$ CD38$^+$ PD-1$^+$ CD8$^+$ and CD4$^+$ T cells did not share clonal overlap with T cells from matched healthy mucosal tissue. Instead, we found that CD103$^+$ CD38$^+$ PD-1$^+$ cells showed most overlapping TCR clonotypes with pericolic lymph nodes, for both CD8$^+$ and CD4$^+$ T cell populations, suggesting their potential origin from the lymph nodes. In the pericolic lymph nodes, T cells are primed by antigen presenting cells and subsequently migrate to the tumor where they recognize their cognate antigen and clonally expand. It has been proposed that upon TCR activation, intratumoral T cells may upregulate tissue-residency and dysfunctional markers that retain the cells within the tumor.[15,25] Our findings further support this hypothesis as CD103$^+$ CD38$^+$ PD-1$^+$ T cells were generally infrequent in the pericolic lymph node samples (**Figure 1**). We previously detected a high frequency of these cells in a tumor-positive lymph node sample.[14] TCRβ sequencing showed 86% overlap between TCR clonotypes of CD103$^+$ PD-1$^+$ CD4$^+$ T cells in the tumor sample with CD45RO$^+$ CD4$^+$ T cells in the pericolic lymph

node sample (data not shown). It would therefore be of interest to further study the TCR repertoires of tumor-positive lymph nodes to test for the presence of tumor-reactivity at those sites.

Due to the relatively small study size, it is important to extend our findings to a larger cohort of patients with CRC. Moreover, further studies could investigate whether the markers for tissue-residency and activation tested by us and others could be used to expand the Immunoscore[2,13], which may help to further segregate tumor-reactive T cells from bystander T cells, and add to its prognostic value for patients with CRC. Indeed, in a pilot study we observed that scoring $CD3^+CD103^+$ T cells further separated MMR-deficient from -proficient colorectal tumors as compared to scoring $CD3^+CD8^+$ T cells only (data not shown).

Altogether, our findings suggest that the addition of tissue-residency and activation markers as CD103, CD38, PD-1, and CD39 on intratumoral T cells may help in the search of finding enriched CDR3β sequences that correspond to tumor antigen-specific T cells. Immunotherapeutic strategies aiming at increasing the frequency of tumor antigen-specific TCRs (e.g. transduction of tumor antigen-specific TCRs on functional donor αβ T cells) may boost the clinical activity of T cell checkpoint blockade therapies for patients with CRC.

METHODS

Patient samples

Primary CRC tissues with matched adjacent healthy mucosa, pericolic lymph nodes, and peripheral blood samples were from patients with CRC (N=7) undergoing surgical resection of their tumor at the Leiden University Medical Center (LUMC, the Netherlands). All patients were treatment-naïve, and did not have a previous history of inflammatory bowel disease. Patients provided written informed consent, and the study was approved by the Medical Ethical Committee of the Leiden University Medical Center (protocol P15.282). All specimens were anonymized and handled according to the ethical guidelines described in the Code for Proper Secondary Use of Human Tissue in the Netherlands of the Dutch Federation of Medical Scientific Societies.

Processing of samples

Colorectal tumor tissues, adjacent healthy mucosa, and pericolic lymph nodes were processed as described previously.[14] Briefly, tissue specimens were minced into small fragments in a petri dish, followed by enzymatical digestion with 1 mg/mL collagenase D (Roche Diagnostics) and 50 µg/mL DNase I (Roche Diagnostics) in 5 mL of IMDM+Glutamax medium for 30 min at 37°C in gentleMACS C tubes (Miltenyi Biotec). Cell suspensions were mechanically dissociated during and after incubation with the gentleMACS Dissociator (Miltenyi Biotec). Thereafter, cell suspensions were filtered through a 70-µm cell strainer (Corning) and washed in IMDM+Glutamax medium. With regard to the peripheral blood samples, peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll-Paque (provided by apothecary LUMC) density-gradient centrifugation. Cell count and viability of the different samples were determined with the Muse Count & Viability Kit (Merck) on the Muse Cell Analyzer (Merck), and cells were cryopreserved in liquid nitrogen based on the number of viable cells until time of analysis.

FACS sorting

Cells from tumor digests, healthy mucosa, lymph nodes and PBMCs were thawed and rest at 37°C in IMDM+L-glutamine (Lonza) with 20% fetal calf serum (FCS) (Sigma-Aldrich) for 30min. After washing in FACS buffer (DPBS (Braun) with 1% FCS), cells were incubated with human Fc receptor block (BioLegend) and stained with the following cell surface antibodies: 1:100 anti-CD4-BV421 [clone RPA-T4, Sony], 1:200 anti-CD8a-BV650 [clone RPA-T8, BD Biosciences], 1:10 anti-CD103-FITC [clone Ber-ACT8, BD Biosciences], 1:160 anti-CD45-PerCP-Cy5.5 [clone 2D1, eBioscience], 1:30 anti-PD-1-PE [clone MIH4, eBioscience], 1:200 anti-CD38-PE-Cy7 [clone HIT2, eBioscience], and 1:60 anti-CD39-APC [clone A1, BioLegend] for 45 min at 4°C. To discriminate live from dead cells, a live/dead fixable near-infrared viability dye (1:1000, Life Technologies) was included in each staining. Thereafter, cells were washed three times in FACS buffer before cell sorting. Cells were resuspended in DPBS before sorting in 1.5 mL RNase-free Eppendorf tubes containing 80 mL of T1 buffer from the NucleoSpin Tissue XS kit. CompBeads (BD Biosciences) and ArC reactive beads (Life Technologies) were used for compensation controls. A gating strategy for single, live $CD45^+$ $CD8^+$ T cells as well as $CD4^+$ T cells was applied, from which the following T cell populations were sorted from colorectal tumor samples: i) $CD103^+$ $CD38^+$ $PD\text{-}1^+$, ii) $CD103^+$ $CD38^-$ $PD\text{-}1^-$, iii) $CD103^-$ $CD38^+$ $PD\text{-}1^+$, and iv) $CD103^-$ $CD38^-$ $PD\text{-}1^-$ (**Figure S1**). FMO controls for PE and PE-Cy7 were included on a PBMC samples used as positive control. In addition, $CD8^+$ and $CD4^+$ T cells were sorted from matched adjacent healthy mucosa, pericolic lymph nodes, and peripheral blood samples. Cells were sorted on a FACS Aria III 4L (BD Biosciences). Of note, for two patients (CRC44 and CRC102), no CD8/CD4 sorting was performed for healthy mucosa, lymph nodes, and peripheral blood samples, but instead fresh-frozen tissue slides (for healthy mucosa and lymph node samples) and EDTA blood (for peripheral blood samples) were used to isolate DNA.

DNA isolation

After sorting, cells were centrifuged for 5 min at 1,500 x *g* at 4°C and drop-wise resuspended in T1

buffer (80 mL for 10-100,000 cells and 200 mL for 100,000-10^7 cells). Samples were only included in downstream analysis if they contained at least 100 cells. Samples were transferred to -20°C until DNA isolation. DNA was isolated with the NucleoSpin® Tissue (XS) Kit (740952, Macherey-Nagel) according to the manufacturer's instructions. Upon thawing the samples, Proteinase K was added. Jurkat cells were included as control in a pilot experiment. DNA quality was assessed by running a qPCR on the *PPP2R1B* housekeeping gene on all samples before continuing with TCRβ sequencing.

TCRβ-chain sequencing

For bulk TCRβ repertoire profiling, TCRβ DNA libraries were constructed with the Ion AmpliSeq™ Library Kit Plus (4488990) and the Oncomine™ TCRβ-SR Assay (A39072) according to the user guide (MAN0017438). The individual libraries were quantified using the Ion Library TaqMan™ Quantitation Kit (4468802), and pooled to a total concentration of 25 pM. Final libraries were templated on ION 530™ Chips (A27764) using the Ion Chef™ System and sequenced on the Ion Genestudio™ S5 Series System with the Ion 510™ & Ion 520™ & Ion 530™ Kit - Chef (A34461) either in house or by GenomeScan BV (Leiden, the Netherlands) according to the user guide references (MAN0016854). All kits and equipment used, unless stated otherwise, are from ThermoFisher Scientific. The diversity analysis was performed with the IonReporter 5.14 software using the Oncomine TCRβ-SR - w1.2 - DNA - Single Sample (version 5.12) workflow with default settings.

Data analysis

Processed sequence data for each sample were analyzed with custom scripts in PostgreSQL (version 12.9). Number of clonotypes, CDR3β length, Shannon diversity, evenness, and the distribution and frequency of clones were processed for each sample. Clonality represented the distribution of clone sizes and was defined as 1 minus the normalized Shannon diversity of the TCRβ abundances (i.e., the evenness). Unique clonotypes were defined as TCRβ sequences with the same V-gene family, J-gene segment, and CDR3β amino acid and nucleotide sequences. For investigations of clonal relations between samples, the number of overlapping TCR clonotypes and their frequency were analyzed for the $CD8^+$ and $CD4^+$ TCR repertoires separately. Bars and boxes represent median with IQR.

ACKNOWLEDGEMENTS

We thank K.C.M.J. Peeters, M.G. Kallenberg-Lantrua, D. Berends-van der Meer, and F.A. Holman for their help in collecting and providing samples from patients with colorectal cancer; the Flow cytometry Core Facility of the Leiden University Medical Center for their help with cell sorting; Genome Scan for their help with TCR-sequencing; E.J. Heeres for his help with the TCR-sequencing analysis.

AUTHOR CONTRIBUTIONS

N.L. de Vries conceived the study. N.L. de Vries and J. Bras Gomes Nunes performed experiments. N.L. de Vries performed the computational analyses presented in this paper. N.L. de Vries, J. Bras Gomes Nunes, N.F. de Miranda, and F. Koning analyzed and interpreted the data. R. van der Breggen assisted with experiments. F. Koning and N.F. de Miranda supervised the study. The manuscript was written by N.L. de Vries, F. Koning, and N.F. de Miranda in collaboration with all co-authors.

COMPETING INTERESTS

The authors declare no competing interests exist.

REFERENCES

1 Pagès, F. *et al.* Effector memory T cells, early metastasis, and survival in colorectal cancer. *N Engl J Med* 353, 2654-2666, doi:10.1056/NEJMoa051424 (2005).

2 Galon, J. *et al.* Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 313, 1960-1964, doi:10.1126/science.1129139 (2006).

3 Tumeh, P. C. *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 515, 568-571, doi:10.1038/nature13954 (2014).

4 Huang, A. C. *et al.* A single dose of neoadjuvant PD-1 blockade predicts clinical outcomes in resectable melanoma. *Nat Med* 25, 454-461, doi:10.1038/s41591-019-0357-y (2019).

5 Chalabi, M. *et al.* Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat Med* 26, 566-576, doi:10.1038/s41591-020-0805-8 (2020).

6 Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* 348, 69-74, doi:10.1126/science.aaa4971 (2015).

7 Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D. & Perucho, M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363, 558-561, doi:10.1038/363558a0 (1993).

8 Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487, 330-337, doi:10.1038/nature11252 (2012).

9 Le, D. T. *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409-413, doi:10.1126/science.aan6733 (2017).

10 Overman, M. J. *et al.* Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *The Lancet. Oncology* 18, 1182-1191, doi:10.1016/s1470-2045(17)30422-9 (2017).

11 Overman, M. J. *et al.* Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 36, 773-779, doi:10.1200/jco.2017.76.9901 (2018).

12 Fridman, W. H., Zitvogel, L., Sautes-Fridman, C. & Kroemer, G. The immune contexture in cancer prognosis and treatment. *Nature reviews. Clinical oncology* 14, 717-734, doi:10.1038/nrclinonc.2017.101 (2017).

13 Pagès, F. *et al.* International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet* 391, 2128-2139, doi:10.1016/s0140-6736(18)30789-x (2018).

14 de Vries, N. L. *et al.* High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of antitumour immunity. *Gut* 69, 691-703, doi:10.1136/gutjnl-2019-318672 (2020).

15 Duhen, T. *et al.* Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat Commun* 9, 2724, doi:10.1038/s41467-018-05072-0 (2018).

16 Simoni, Y. *et al.* Bystander CD8(+) T cells are abundant and phenotypically distinct in human tumor infiltrates. *Nature* 557, 575-579, doi:10.1038/s41586-018-0130-2 (2018).

17 van den Bulk, J. *et al.* Neoantigen-specific immunity in low mutation burden colorectal cancers of the consensus molecular subtype 4. *Genome Med* 11, 87, doi:10.1186/s13073-019-0697-8 (2019).

18 Roh, W. *et al.* Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Science translational medicine* 9, doi:10.1126/scitranslmed.aah3560 (2017).

19 McGranahan, N. *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* 351, 1463-1469, doi:10.1126/science.aaf1490 (2016).

20 Zheng, C. *et al.* Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* 169, 1342-1356 e1316, doi:10.1016/j.cell.2017.05.035 (2017).

21 Allard-Chamard, H. *et al.* CD4(+)CTLs in Fibrosing Mediastinitis Linked to Histoplasma capsulatum. *J Immunol* 206, 524-530, doi:10.4049/jimmunol.2000433 (2021).

22 Sharma, P. & Allison, J. P. The future of immune checkpoint therapy. *Science* 348, 56-61, doi:10.1126/science.aaa8172 (2015).

23 Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer-immune set point. *Nature* 541, 321-330, doi:10.1038/nature21349 (2017).

24 Scheper, W. *et al.* Low and variable tumor reactivity of the intratumoral TCR repertoire in human cancers. *Nat Med* 25, 89-94, doi:10.1038/s41591-018-0266-5 (2019).

25 Oliveira, G. *et al.* Phenotype, specificity and avidity of antitumor CD8(+) T cells in melanoma. *Nature* 596, 119-125, doi:10.1038/s41586-021-03704-y (2021).

26 Gros, A. *et al.* PD-1 identifies the patient-specific $CD8^+$ tumor-reactive repertoire infiltrating human tumors. *J Clin Invest* 124, 2246-2259, doi:10.1172/jci73639 (2014).

27 Gros, A. *et al.* Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients. *Nat Med* 22, 433-438, doi:10.1038/nm.4051 (2016).

28 Ye, Q. *et al.* CD137 accurately identifies and enriches for naturally occurring tumor-reactive T cells in tumor. *Clin Cancer Res* 20, 44-55, doi:10.1158/1078-0432.Ccr-13-0945 (2014).

29 Eiva, M. A., Omran, D. K., Chacon, J. A. & Powell, D. J., Jr. Systematic analysis of CD39, CD103, CD137, and PD-1 as biomarkers for naturally occurring tumor antigen-specific TILs. *Eur J Immunol*, doi:10.1002/eji.202149329 (2021).

30 Simoni, Y. *et al.* Bystander CD4⁺ T cells infiltrate human tumors and are phenotypically distinct. *bioRxiv*, 2020.2007.2015.204172, doi:10.1101/2020.07.15.204172 (2020).

SUPPLEMENTAL FIGURES

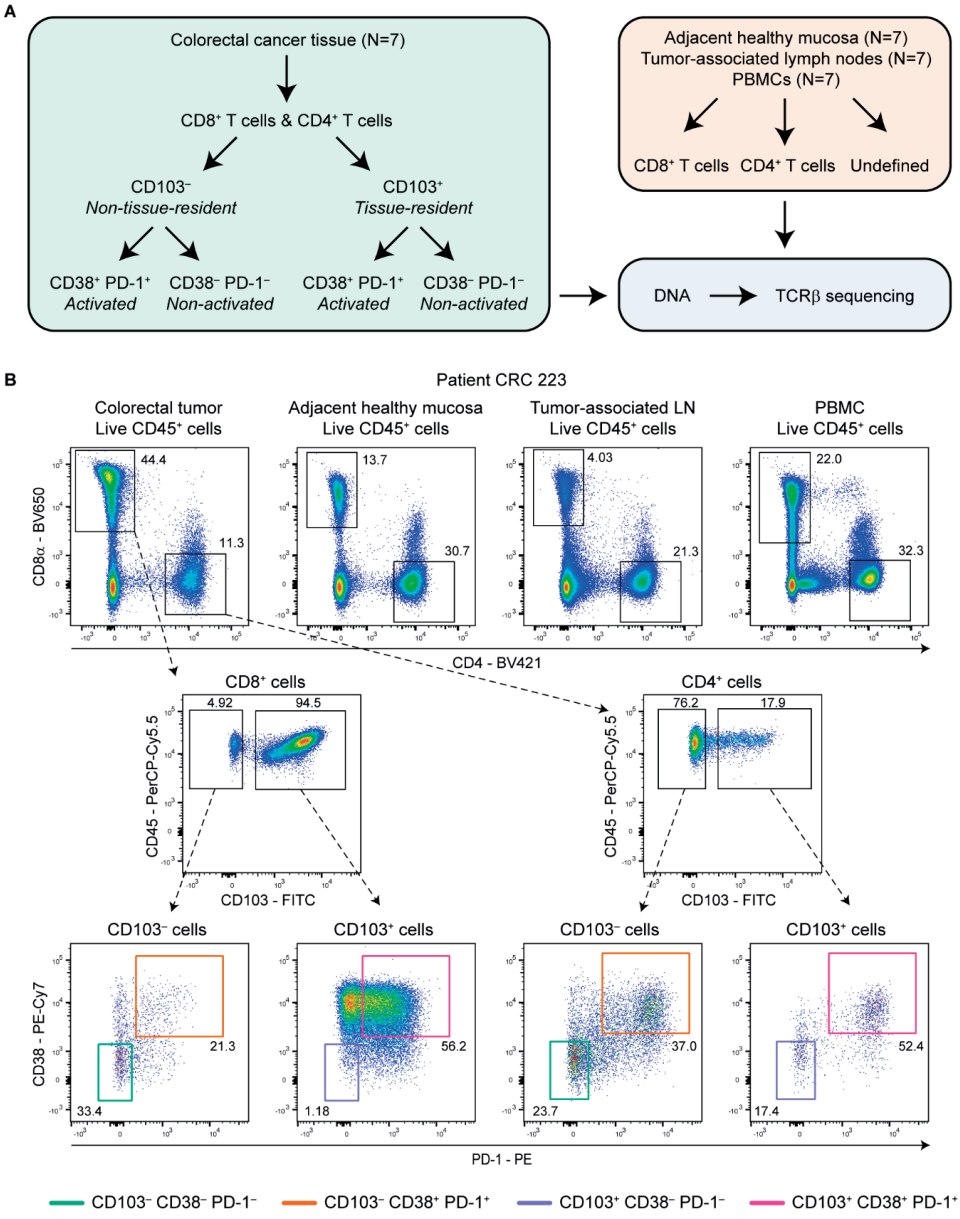


Figure S1. Gating strategy for the sorting of CD8⁺ and CD4⁺ T cell populations from colorectal tumors and matched non-malignant tissues.

A. Experimental design of TCRβ sequencing performed on DNA isolated from different T cell populations FACS-sorted from colorectal tumors and matched adjacent healthy mucosa, pericolic lymph nodes, and peripheral blood samples from patients with CRC. **B**. Flow cytometry gating strategy for the sorting of different CD8⁺ and CD4⁺ T cell populations from a representative MMR-deficient CRC sample with matched non-malignant tissues showing sequential gates with percentages.

Figure S2. Overlap in TCR clonotypes of different CD8+ T cell populations for each individual colorectal cancer patient.

Circular graphs showing the frequency of overlapping TCR clonotypes of different CD8+ T cell populations from colorectal tumors with matched adjacent healthy mucosa ("normal"), pericolic lymph nodes, and peripheral blood samples for each patient with CRC. Samples with < 10 clonotypes are colored in red. Data from seven independent experiments.

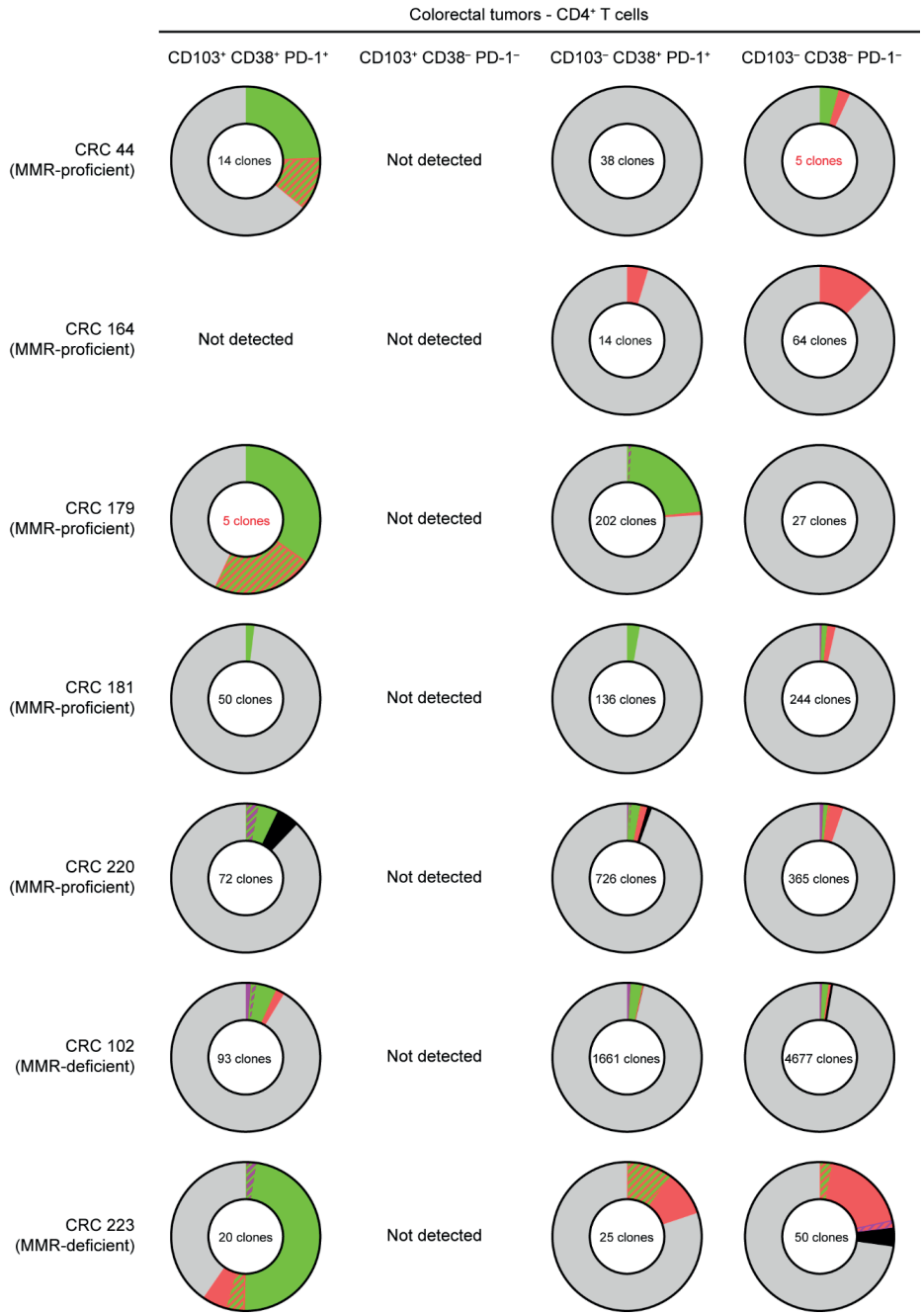


Figure S3. Overlap in TCR clonotypes of different CD4⁺ T cell populations for each individual colorectal cancer patient.

Circular graphs showing the frequency of overlapping TCR clonotypes of different CD4⁺ T cell populations from colorectal tumors with matched adjacent healthy mucosa ("normal"), pericolic lymph nodes, and peripheral blood samples for each patient with CRC. CD103⁺ CD38⁻ PD-1⁻ CD4⁺ T cells were not detected in the tumor samples. Samples with < 10 clonotypes are colored in red. Data from seven independent experiments.

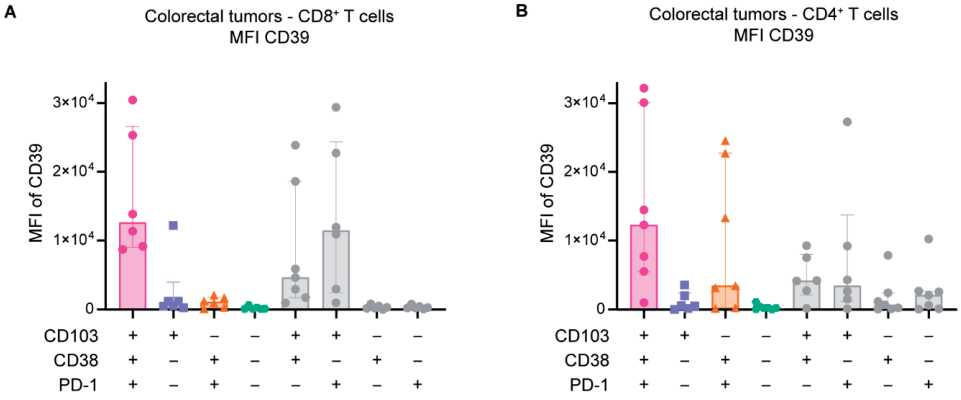


Figure S4. Tissue-resident, activated T cell populations show high expression levels of CD39.
A. MFI of CD39 in the different CD8⁺ T cell populations from colorectal tumors (N=6-7) analyzed by flow cytometry. Background MFI of CD39 of each unstained sample was subtracted. **B**. MFI of CD39 in the different CD4⁺ T cell populations from colorectal tumors (N=6-7) analyzed by flow cytometry. Background MFI of CD39 of each unstained sample was subtracted. **A,B**. Bars represent median with IQR. Samples were only included if ≥100 cells were present. Data from seven independent experiments.

SUPPLEMENTAL TABLES

The supplementary tables are available in the appendix to this thesis at the repository of Leiden University (https://hdl.handle.net/1887/3439882) and can be requested from the author.

Table S1: Overview of number of cells, number of clones, Shannon diversity, and clonality for $CD8^+$ T cells from all samples.
Table S2: Overview of number of cells, number of clones, Shannon diversity, and clonality for $CD4^+$ T cells from all samples.

Local and systemic immune profiles of human pancreatic ductal adenocarcinoma revealed by single-cell mass cytometry

7

Thomas P. Brouwer[1,2]*, **Natasja L. de Vries**[1,3]*, Tamim Abdelaal[4,5], Ricki T. Krog[1], Zheng Li[1], Dina Ruano[1], Arantza Fariña[6], Boudewijn P. F. Lelieveldt[5,7], Hans Morreau[1], Bert A. Bonsing[2], Alexander L. Vahrmeijer[2], Frits Koning[3], Noel F.C.C. de Miranda[1].

[1]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands. [2]Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands. [3]Department of Immunology, Leiden University Medical Center, Leiden, The Netherlands. [4]Pattern recognition and Bioinformatics, Delft University of Technology, Delft, The Netherlands. [5]Leiden computational biology center, Leiden University Medical Center, Leiden, The Netherlands. [6]Department of Pathology, Amsterdam UMC, University of Amsterdam, Cancer Center Amsterdam, Amsterdam, The Netherlands. [7]LKEB radiology, Leiden University Medical Center, Leiden, The Netherlands. *Co-first authors.

J Immunother Cancer. 2022; 10: e004638.
doi: 10.1136/jitc-2022-004638

ABSTRACT

Background

Pancreatic ductal adenocarcinoma (PDAC) is a highly lethal malignancy in need of effective (immuno)therapeutic treatment strategies. For the optimal application and development of cancer immunotherapies, a comprehensive understanding of local and systemic immune profiles in patients with PDAC is required. Here, our goal was to decipher the interplay between local and systemic immune profiles in treatment-naïve PDAC patients.

Methods

The immune composition of PDAC, matched non-malignant pancreatic tissue, regional lymph nodes, spleen, portal vein blood, and peripheral blood samples (collected before and after surgery) from 11 patients with PDAC was assessed by measuring 41 immune cell markers by single-cell mass cytometry. Furthermore, the activation potential of tumor-infiltrating lymphocytes as determined by their ability to produce cytokines was investigated by flow cytometry. In addition, the spatial localization of tumor-infiltrating innate lymphocytes in the tumor microenvironment was confirmed by multispectral immunofluorescence.

Results

We found that $CD103^+$ $CD8^+$ T cells with cytotoxic potential are infrequent in the PDAC immune microenvironment and lack the expression of activation markers and checkpoint blockade molecule PD-1. In contrast, PDAC tissues showed a remarkable increased relative frequency of B cells and regulatory T cells as compared to non-malignant pancreatic tissues. Besides, a previously unappreciated innate lymphocyte cell (ILC) population ($CD127^-$ $CD103^+CD39^+CD45RO^+$ ILC1-like) was discovered in PDAC tissues. Strikingly, the increased relative frequency of B cells and regulatory T cells in pancreatic cancer samples was reflected in matched portal vein blood samples but not in peripheral blood, suggesting a regional enrichment of immune cells that infiltrate the PDAC microenvironment. After surgery, decreased frequencies of myeloid dendritic cells were found in peripheral blood.

Conclusions

Our work demonstrates an immunosuppressive landscape in PDAC tissues, generally deprived of cytotoxic T cells and enriched in regulatory T cells and B cells. The antitumor potential of ILC1-like cells in PDAC may be exploited in a therapeutic setting. Importantly, immune profiles detected in blood isolated from the portal vein reflected the immune cell composition of the PDAC microenvironment, suggesting that this anatomical location could be a source of tumor-associated immune cell subsets.

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal cancer types in the industrialized world.[1] The outlook for patients with PDAC is bleak with a five-year survival rate of 10%, and without novel therapeutic strategies the survival rate will not improve during the coming years.[2-4] The development of cancer immunotherapy, using therapeutic antibodies targeting CTLA-4 and the PD-1/PD-L1 axis in particular, have yielded sustained clinical responses in traditionally incurable cancer types including non-small cell lung cancer and melanoma, amongst others.[5-8] These immunotherapeutic strategies reinvigorate antitumor T cell responses. PDAC, however, is non-immunogenic and characterized by a lack of naturally occurring immune responses due to a generally low mutational burden as well as a large stromal compartment consisting of few and mostly immunosuppressive immune cells compared to other cancer types,[9] thereby rendering the cancer insensitive to immunotherapy.[10-12]

Novel omics-technologies can aid in the improvement of patient selection for immunotherapy by giving a more profound insight into the tumorigenic and immunologic mechanisms at play in the tumor microenvironment. With the rise in T cell checkpoint blockade therapies, the T cell landscape has been predominantly looked into. In PDAC, studies have provided insights into the prevalence and distribution of T cells in PDAC tissues and revealed that the T cell landscape is highly complex and varies across patients.[13-16] Others, have also highlighted possible myeloid checkpoint targets as well as a possible role for B cells in the response to immune checkpoint blockade. Unfortunately, these new insights regarding the immune landscape of PDAC have not impacted clinical responses or survival, with the exception of mismatch repair (MMR)-deficient PDAC tumors.[17-21]

To improve our understanding of additional biological features that could play a role in the lack of response to current immunotherapies and for the development of alternative therapeutic approaches, it is pivotal to determine the immune processes mediating antitumor responses within the local and systemic environment. Here we aimed to decipher the interplay between the local and systemic immune landscapes across PDAC tumors and their matched non-malignant pancreatic tissue, regional lymph nodes, portal vein blood, and peripheral blood obtained before and after surgery from treatment-naïve PDAC patients. We discovered the presence of an activated, $CD103^+$ innate lymphoid cell (ILC)1-like population in PDAC tissues that was infrequent in matched non-malignant, lymph node, and blood samples. Furthermore, we found that PDAC tumors had notably reduced $CD103^+$ memory $CD8^+$ T cells in comparison to non-malignant pancreatic tissue, while memory B cells and regulatory T cells were significantly increased. The immune profiles in PDAC tissue were mimicked in portal vein blood but not peripheral blood samples of the patients.

RESULTS

The PDAC immune microenvironment is enriched for B cells and regulatory T cells

The immune composition of PDAC and non-malignant pancreatic tissues, matched regional lymph nodes, spleen, portal vein blood, and peripheral blood samples (collected before and after surgery) from 11 patients with PDAC was deciphered by measuring 41 immune cell markers by single-cell mass cytometry (**Figure S1**). HSNE analysis was performed on all acquired $CD45^+$ cells ($18.2x10^6$ cells in total), and showed the presence of eight major immune cell clusters (**Figure S2**). First, we specifically explored the immune landscape in PDAC and non-malignant pancreatic tissues by single-cell mass cytometric analysis of 11 PDAC tissues with 6 matched non-malignant pancreatic tissues. This analysis showed the presence of different clusters of $CD4^+$ T cells, $CD8^+$ T cells/γδ T cells, innate lymphoid cells (ILCs), B cells, and myeloid cells (**Figure 1A**, **Figure S3**). In non-malignant pancreatic tissue, memory $CD8^+$ T cells were the most prevalent subset, followed by myeloid cells, and memory $CD4^+$ T cells, albeit the absolute number of immune cells was generally low (**Figure 1B**, **Figure S3**). Compared to non-malignant tissues, PDAC tissues showed a higher relative frequency of B cells and $CD4^+CD25^+CD127^{low}ICOS^+$ cells ($P<0.01$ by Mann-Whitney tests; **Figure 1B**). By flow cytometry, we demonstrated that this population was mainly comprised of $FOXP3^+$, regulatory T cells (**Figure S4**), in line with our previous observations in colorectal cancer.[22] In contrast, PDAC tissues showed a remarkable decreased frequency of $CD8^+$ T cells with an effector memory phenotype as compared to non-malignant tissue ($P<0.01$ by Mann-Whitney test; **Figure 1B**).

PDAC tissues are largely deprived of infiltration by $CD8^+$ T cells with cytotoxic potential

We further investigated the phenotype of $CD8^+$ and $CD4^+$ T cells in PDAC and non-malignant pancreatic tissues (**Figure 2A**, **Figure S5**). Within the effector memory $CD8^+$ T cells, a reduced relative frequency in $CD103^+$ (tissue-residency marker) $CD8^+$ T cells was observed in PDAC tissues as compared to their matched non-malignant tissues ($P<0.01$ by Mann-Whitney test; **Figure 2B**), with the exception of one patient (ISPIC20; indicated by the dark blue dot). The majority of $CD8^+$ T cells contained in PDAC tissues did not show expression of activation markers such as CD38, HLA-DR, and CD39 (**Figure 2A**). The $CD103^+$ memory $CD8^+$ T cells contained in PDAC tissues also carried low amounts of cytolytic enzymes granzyme B/perforin and of pro-inflammatory cytokines TNF-α and IFN-γ (**Figure 2C**). However, the $CD103^+$ memory $CD8^+$ T cells still had the potential to become activated as determined by their ability to produce TNF-α and IFN-γ upon stimulation (**Figure 2C**). Co-expression of CD39 and CD103 has been proposed to identify tumor-reactive $CD8^+$ T cells in different solid malignancies.[23,24] Of the 11 patients studied here, we only found $CD39^+CD103^+CD8^+$ T cells in one PDAC tissue (ISPIC20), constituting 35% of all $CD45^+$ cells in this patient (**Figure 2B**). Interestingly, the pancreatic tumor of patient ISPIC20 showed a relatively higher number of coding, non-synonymous, somatic mutations (N=70) as compared to the rest of the cohort (average N=29, SD±16). Whole-exome sequencing of the tumor tissue of ISPIC20 revealed the presence of a homozygous, somatic mutation in *EXO1*, involved in DNA

mismatch repair and homologous recombination, which potentially underlies the higher mutational load observed in this sample. In addition to investigating the presence of T cells with tumor-reactive phenotypes, we determined whether T cells in PDAC tissues expressed the immune checkpoint blockade molecule PD-1. We found that PD-1^{+}CD8^{+} T cells were almost completely absent in the tumor tissues, with the exception of one patient (ISPIC22). In contrast to the absence of cytotoxic T cells in the pancreatic tumors, CD4^{+} T cells were relatively abundant and showed expression of activation markers (**Figure 2A**). CD4^{+} T cell phenotypes contained in PDAC tissues included PD-1^{+}CD4^{+} T cells, of which the relative frequency was significantly increased in pancreatic tumors as compared to the non-malignant tissues ($P<0.01$ by Mann-Whitney test; **Figure 2D**). A subset of these PD-1^{+}CD4^{+} T cells co-expressed ICOS and this population was almost exclusively found in PDAC tissues ($P<0.01$ by Mann-Whitney test; **Figure 2D**). In line with the CD8^{+} compartment, memory CD4^{+} T cells in PDAC tissues showed low levels of granzyme B/perforin, TNF-α, and IFN-γ *ex vivo* (**Figure S6**). Interestingly, the PDAC tissue of patient ISPIC20 showed the most CD103^{+} CD4^{+} T cells as well as CD39^{+}CD103^{+}CD4^{+} T cells, reflecting the characteristics of the CD8^{+} compartment (**Figure 2D**). Last, a population of CD25^{+} activated CD4^{+} T cells showed an increased relative frequency in PDAC as compared to non-malignant tissues ($P<0.01$ by Mann-Whitney test; **Figure 2D**). Altogether, these data demonstrate a lack of activated, CD103^{+} cytotoxic T cells in the immune microenvironment of PDAC, while an adequate profile for the generation of helper T cell responses may be present.

Activated, CD103^{+} ILC1-like cells are prevalent in pancreatic tumors

Detailed analysis of Lin^{-}CD7^{+} ILCs in PDAC and non-malignant pancreatic tissue revealed the presence of a largely tumor tissue-specific ILC population, characterized by the lack of conventional ILC marker CD127 and the expression of CD103, CD39 and CD45RO (**Figure 3A**, **Figure S5**). These CD127^{-}CD103^{+}CD39^{+}CD45RO^{+} ILCs constituted up to 16% of the immune cell infiltrate in pancreatic tumors (**Figure 3B**), and were present at a higher relative frequency than in non-malignant tissue ($P<0.05$ by Mann-Whitney test; **Figure 3B**). Interestingly, we previously found that such CD127^{-}CD103^{+}CD45RO^{+} ILCs are particularly enriched in MMR-deficient colorectal cancers, and have characteristics of intraepithelial ILC1-like cells.[22] In contrast to the exceptional high levels of cytotoxic enzymes present in this population in colorectal cancers, we found that only up to 10% of the CD45RO^{+} ILC1-like cells expressed granzyme B/perforin in PDAC tissues. When stimulated, low levels of TNF-α and IFN-γ were also produced by the CD45RO^{+} ILC1-like cells, up to 14 and 4%, respectively (**Figure 3C**, **Figure S7**). In non-malignant tissues, CD127^{-}CD103^{-}CD39^{-}CD45RA^{+} NK-like cells were the most frequent ILC population (**Figure 3B**, **Figure S5**).

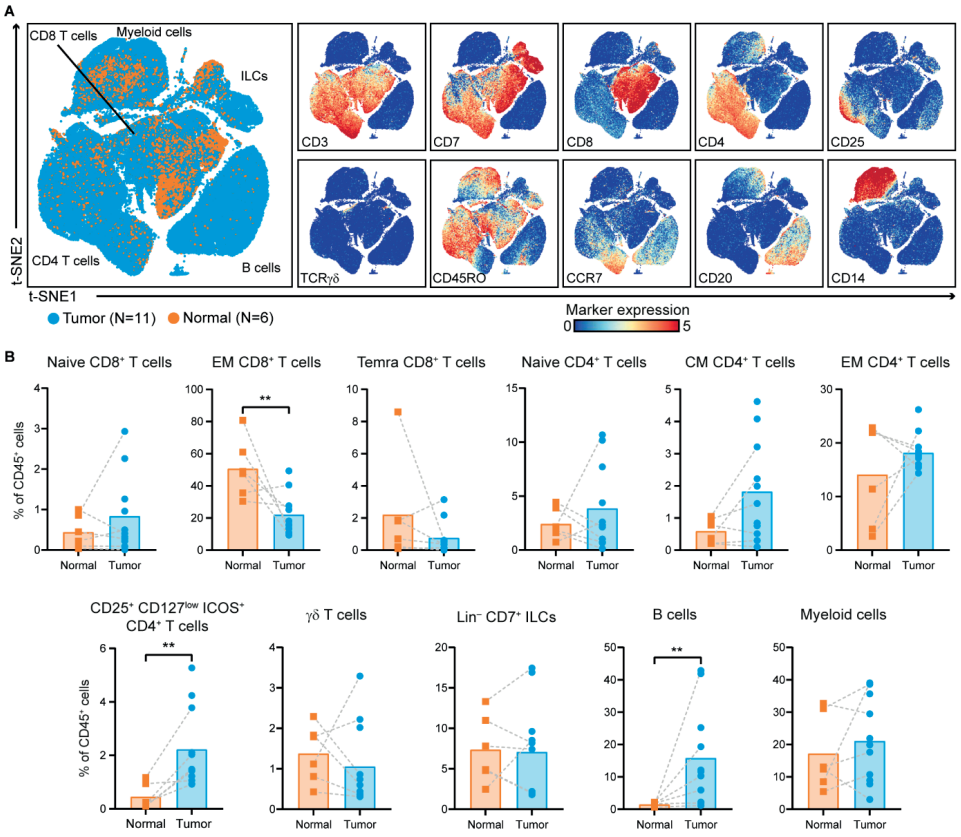


Figure 1. The PDAC immune microenvironment is enriched for B cells and regulatory T cells.
A. **t**-SNE embedding showing 155,433 immune cells isolated from PDAC (N=11) and non-malignant pancreatic tissues (N=6), clustered based on the expression of 41 immune cell markers by single-cell mass cytometry. In the right panels, the relative expression of indicated immune lineage markers is shown. **B.** Frequencies of the major immune lineages in PDAC tissues (N=11) as compared to non-malignant pancreatic tissues (N=6) shown as percentage of total CD45⁺ cells. Each dot represents an individual sample. Bars indicate the mean and lines indicate matched samples. Data from 11 independent experiments with mass cytometry. **P<0.01 by Mann-Whitney test.

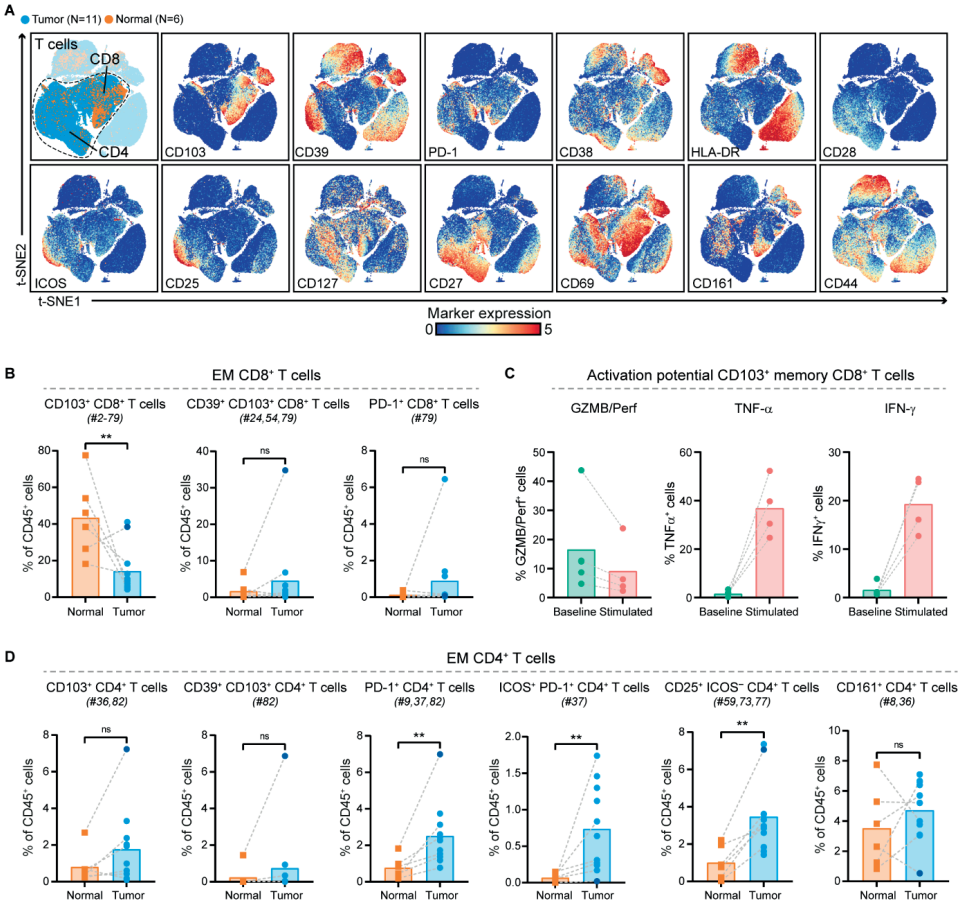


Figure 2. PDAC tissues are largely deprived of infiltration by CD8⁺ T cells with cytotoxic potential.
A. t-SNE embedding from **Figure 1A** highlighting the CD8⁺ and CD4⁺ T cell populations, colored by tissue type (first plot) and relative expression of indicated T cell markers (remainder). An overview of all T cell clusters identified is shown in **Figure S5**. **B.** Frequencies of selected effector memory CD8⁺ T cell clusters in PDAC (N=11) and non-malignant pancreatic (N=6) tissues shown as percentage of total CD45⁺ cells. **C.** Cytokine production by CD103⁺ (tissue-resident) memory CD8⁺ T cells from PDAC tissues *ex vivo* (N=5) and upon stimulation with PMA/ionomycin (N=4) measured by flow cytometry. Bars indicate the mean and lines indicate matched samples. **D.** Frequencies of selected effector memory CD4⁺ T cell clusters in PDAC (N=11) and non-malignant pancreatic (N=6) tissues shown as percentage of total CD45⁺ cells. **B,D.** Cluster IDs correspond to the ones in **Figure S5**. Each dot represents an individual sample (dark blue color represents patient ISPIC20). Bars indicate the mean and lines indicate matched samples. Data from 11 independent experiments with mass cytometry. **P<0.01 by Mann-Whitney test.

We subsequently examined whether the ILC1-like population was also present in regional lymph nodes, spleen, portal vein blood, or peripheral blood by running an HSNE analysis on all samples from the PDAC patients (**Figure 3D**). This analysis confirmed the specific localization within the tumor tissue of the $CD127^{-}CD103^{+}CD39^{+}CD45RO^{+}$ ILC1-like population, as the cells were not found in regional lymph nodes, spleen, portal vein blood, or peripheral blood (**Figure 3E**). To determine the spatial localization of the ILC1-like cells in PDAC tissues, a 5-color multispectral immunofluorescence panel was used on FFPEs tissue sections of the pancreatic tumors. We identified a high proportion of $CD7^{+}CD3^{-}CD45RO^{+}CD103^{+}$ ILC1-like cells in the immune microenvironment of the tumors, in line with the CyTOF data (**Figure 3F**). The ILC1-like cells frequently localized adjacent to the ductal cells, which is in agreement with their $CD103^{+}$ tissue-resident phenotype. Thus, we discovered an activated, $CD103^{+}$ ILC1-like population specifically enriched in pancreatic cancers.

Increased frequencies of $CD27^{+}$ memory B cells in PDAC

Immune marker profiles of the B cell compartment showed that $CD27^{-}$ naïve B cells were predominant in PDAC tissues (accounting for 67% of the B cell population), followed by $CD27^{+}$ memory B cells (accounting for 33% of the B cell population) (**Figure 4A,B**, **Figure S5**). Compared to non-malignant tissues, the relative frequencies of both B cell subsets were increased in PDAC tissues ($CD27^{+}$ memory B cells: $P<0.001$ by Mann-Whitney test; **Figure 4C**). Importantly, none of the tumor samples with higher B cell frequency showed signs of pancreatitis or contained tertiary lymphoid structures (TLS) as confirmed by histological examination of the PDAC tissues. Both $CD27^{-}$ naïve B cells and $CD27^{+}$ memory B cells in PDAC tissues expressed HLA-DR and CD40 (**Figure 4A**), the latter receptor involved in mediating T cell help. In contrast, expression of activation marker CD69 was found only on the $CD27^{+}$ memory B cell population (**Figure 4A**). Last, examination of the composition of the myeloid compartment in PDAC and non-malignant pancreatic tissues showed no significant differences in the frequency of $CD33^{+}HLA\text{-}DR^{low}$ myeloid cells, myeloid dendritic cells (DCs), and plasmacytoid DCs (**Figure S8**). Intriguingly, expression of ICAM-1 and CD16 on myeloid cells was mainly observed in PDAC tissues and not in non-malignant pancreatic tissue (**Figure S8**).

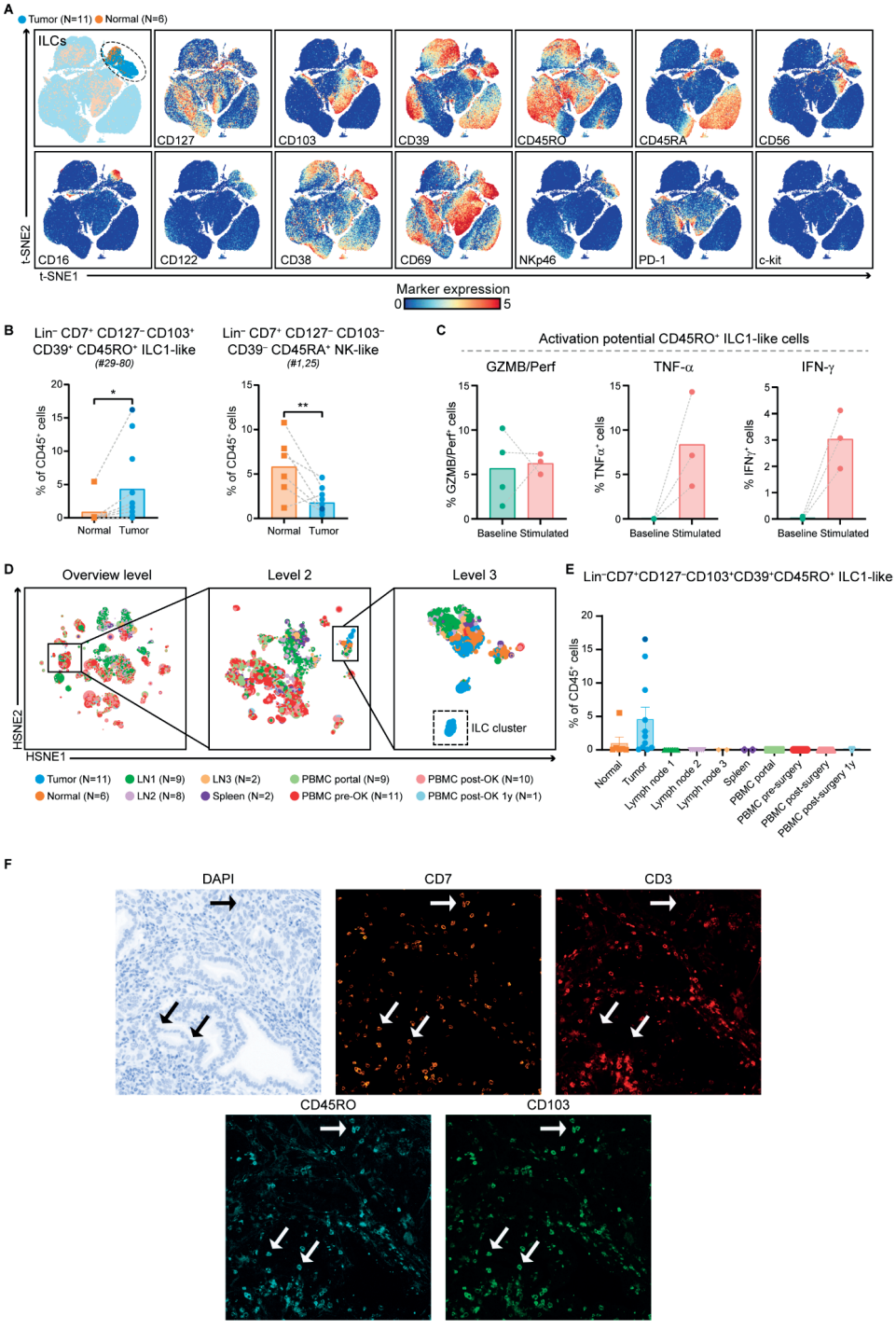


Figure 3. Activated, CD103⁺ ILC1-like cells are prevalent in pancreatic tumors.
A. t-SNE embedding from **Figure 1A** highlighting the Lin⁻CD7⁺ innate lymphoid cell (ILC) population, colored by tissue

type (first plot) and relative expression of indicated ILC markers (remainder). An overview of all ILC clusters identified is shown in **Figure S5**. **B.** Frequencies of selected ILC clusters in PDAC tissues (N=11) as compared to non-malignant pancreatic tissues (N=6) shown as percentage of total CD45$^+$ cells. Cluster IDs correspond to the ones in **Figure S5**. Each dot represents an individual sample (dark blue color represents patient ISPIC20). Bars indicate the mean and lines indicate matched samples. Data from 11 independent experiments with mass cytometry. *P<0.05, **P<0.01 by Mann-Whitney test. **C.** Cytokine production by CD45RO$^+$ ILC1-like cells from PDAC tissues *ex vivo* (N=4) and upon stimulation with PMA/ionomycin (N=3) measured by flow cytometry. Bars indicate the mean and lines indicate matched samples. **D.** HSNE analysis showing 17592 landmarks representing 18.2x10^6 immune cells isolated from PDAC and non-malignant pancreatic tissues, regional lymph nodes, spleen, portal vein blood, and peripheral blood obtained before and after surgery from 11 patients with PDAC clustered based on the expression of 41 immune cell markers by CyTOF. Specific landmarks were selected and embedded at the next, more detailed levels to identify the CD127$^-$CD103$^+$CD39$^+$CD45RO$^+$ ILC1-like cells and their distribution across the different tissue types. Colors represent the different tissue types. **E.** Frequencies of CD127$^-$CD103$^+$CD39$^+$CD45RO$^+$ ILC1-like cells among non-malignant pancreatic (N=6) and PDAC (N=11) tissues, regional lymph nodes (details below), spleen (N=2), portal vein blood (N=8), peripheral blood before surgery (N=11), directly after surgery (N=10), and one year after surgery (N=1) from 11 patients with PDAC shown as percentage of total CD45$^+$ cells. Each dot represents an individual sample. Bars indicate the mean with SEM. Data from 11 independent experiments with mass cytometry. Lymph node 1 derived from the common hepatic artery (N=9), lymph node 2 from the hepatoduodenal ligament (N=8), and lymph node 3 from the abdominal aorta (N=2). **F.** Representative images of the immunofluorescence microscopic detection of CD7$^+$CD3$^-$ CD103$^+$CD45RO$^+$ ILC1-like cells in a pancreatic tumor.

Portal vein blood can reflect the immune cell populations residing in pancreatic cancers

In addition to characterizing the immune infiltration in PDAC and non-malignant pancreatic tissues, we examined the systemic immune profiles of PDAC patients through analysis of different blood samples collected from i) portal vein blood, ii) peripheral blood before surgery, and iii) peripheral blood after surgery. HSNE analysis of all immune cells showed an equivalent distribution of the major immune lineages across the different blood sample types with large populations of myeloid cells and CD4$^+$ T cells (**Figure 5A**, **Figure S9**), except for an increase in the relative frequency of γδ T cells in portal vein blood as compared to the matched peripheral blood samples (P<0.05 by Wilcoxon test; **Figure S9**). The clustering of all blood samples based on their complete immune composition showed that inter-patient heterogeneity was greater than potential differences in immune cell profiles between the different compartments or time points, per patient (**Figure 5B,C**). As such, we performed an analysis per immune lineage to investigate phenotypic differences between the different types of blood samples from PDAC patients.

Strikingly, increased frequencies of the immune cell populations enriched in PDAC tissue were also found in portal vein blood as compared to peripheral blood. For example, in portal vein blood, an increased frequency of CD27$^+$ memory B cells was found as compared to peripheral blood (P<0.05 by Wilcoxon test; **Figure 5D**, **Figure S10**), similar to the increased presence of memory B cells observed in the PDAC tissues (**Figure 4C**). The large majority of the CD27$^+$ memory B cells in portal blood also expressed CD40 (**Figure 5D**, **Figure S10**). Besides, portal vein blood contained higher frequencies of HLA-DR$^+$CD11c$^+$CD16$^+$ myeloid DCs as compared to pre-surgery peripheral blood (P<0.05 by Wilcoxon test; **Figure 5D**). Moreover, portal vein blood showed a higher relative frequency of CD4$^+$CD25$^+$CD127^{low}ICOS$^+$ regulatory T cells as compared to pre-surgery peripheral blood (P<0.05 by Wilcoxon test; **Figure 5D**, **Figure S11**), in line with the increased presence of regulatory T cells found in the PDAC tissues (**Figure 2B**). Of note, the only patient that showed infiltration of PD-1$^+$CD8$^+$ T

cells in the tumor (ISPIC22), showed the highest levels of PD-1^{+}CD8^{+} T cells in portal blood (17% of CD8 T cells; **Figure 5D**), suggesting that local immune profiles in PDAC tissues can also be detected in the blood circulation close to the tumor.

When comparing peripheral blood collected before and after surgery, we observed a decreased frequency of CD16^{-} myeloid DCs in post- versus pre-surgery peripheral blood (P<0.01 by Wilcoxon test; **Figure 5E**, **Figure S10**). This was accompanied by an increase in CD14^{+}CD33^{+}HLA-DR^{-} myeloid cells in peripheral blood after surgery as compared to before (P<0.05 by Wilcoxon test; **Figure 5E**). A higher proportion of MDSCs relative to the myeloid DCs has been previously described and suggested to play a role in the clinical recovery of patients following surgical procedure.[25] Last, a population of CD14^{-}HLA-DR^{-} CD123^{+} basophils was substantially lower in peripheral blood after surgery compared to before (P<0.05 by Wilcoxon test; **Figure 5E**). Analysis of the CD4^{+} T cell compartment in blood showed an increase in CD4^{+} T cells with a central memory phenotype in peripheral blood after surgery as compared to before (P<0.05 by Wilcoxon test; **Figure 5E**), while no differences were observed in the frequencies of naïve and effector memory CD4^{+} T cells (**Figure S11**). Last, peripheral blood obtained after surgery showed a reduction in the number of CD161^{+}CD4^{+} T cells (P<0.05 by Wilcoxon test; **Figure 5E**), a subset of CD4^{+} T cells described to possess enhanced innate-like cytokine production[26].

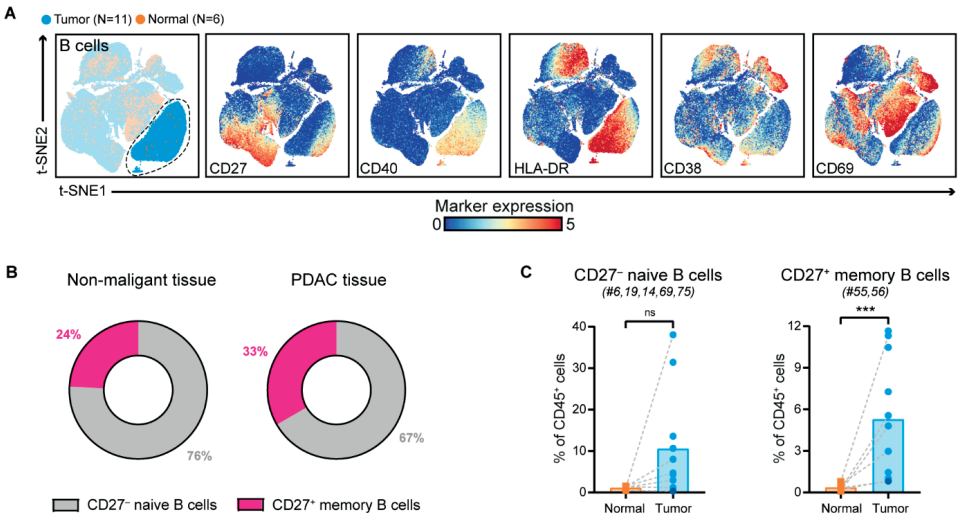


Figure 4. Increased frequencies of CD27^{+} memory B cells in PDAC.
A. t-SNE embedding from **Figure 1A** highlighting the CD20^{+} B cells, colored by tissue type (first plot) and relative expression of indicated B cell markers (remainder). An overview of all B cell clusters identified is shown in **Figure S5**. **B.** Circular plots showing the relative frequency of CD27^{-} naïve B cells and CD27^{+} memory B cells as average percentage of total B cells in non-malignant (N=6) and PDAC tissues (N=11). **C.** Frequencies of CD27^{-} naïve B cells and CD27^{+} memory B cells in PDAC tissues as compared to non-malignant pancreatic tissues shown as percentage of total CD45^{+} cells. Cluster IDs correspond to the ones in **Figure S5**. Each dot represents an individual sample (dark blue color represents patient ISPIC20). Bars indicate the mean and lines indicate matched samples. Data from 11 independent experiments with mass cytometry. ***P<0.001 by Mann-Whitney test.

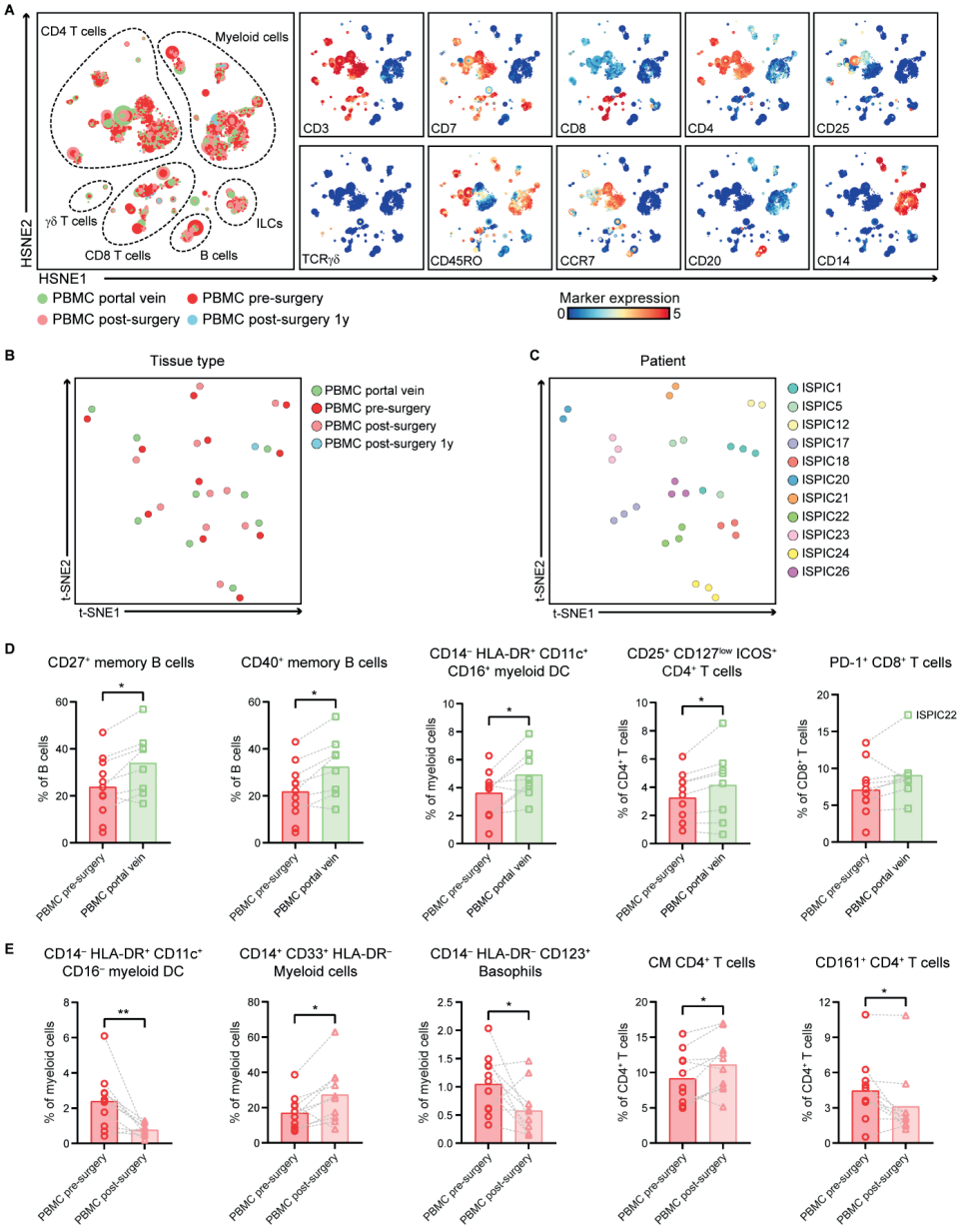


Figure 5. Portal vein blood can reflect the immune cell populations residing in pancreatic cancers.
A. HSNE analysis showing 9921 landmarks representing 10.9x10^6 immune cells isolated from portal vein blood (N=8), peripheral blood before surgery (N=11), peripheral blood directly after surgery (N=10) as well as one year after surgery (N=1) from 11 patients with PDAC clustered based on the expression of 41 immune cell markers by single-cell mass cytometry. In the right panel, the relative expression of indicated immune lineage markers is shown. **B.** Collective t-SNE analysis showing the clustering of all blood samples from the patients with PDAC (N=31) based on cell percentage data (of CD45$^+$ cells) of all immune cell clusters identified (N=417). Every dot represents a sample colored by blood type. **C.** Collective t-SNE analysis of (**B**) colored by patient ID. **D.** Frequencies of selected immune cell populations in portal vein blood (N=8) and peripheral blood pre-surgery (N=11) as percentage of the indicated major immune lineage. **E.** Frequencies of selected immune cell populations in peripheral blood pre-surgery (N=11) and post-surgery (N=10) as

percentage of the indicated major immune lineage. **D,E.** Each dot represents an individual sample. Bars indicate the mean and lines indicate matched samples. Data from 11 independent experiments with mass cytometry. *P<0.05, **P<0.01 by Wilcoxon test.

DISCUSSION

By studying the innate and adaptive immune composition of PDAC and matched non-malignant pancreatic tissues, regional lymph nodes, portal vein blood, and peripheral blood before and after surgery from treatment-naïve PDAC patients, our work provides an in-depth immune characterization of local and systemic immune profiles in patients with PDAC.

Although it has been shown that $CD8^+$ T cells in proximity to PDAC cells correlate with increased survival,[27] we found that memory $CD8^+$ T cells were less frequent in pancreatic cancers in comparison to non-malignant pancreatic tissue. Moreover, the few $CD8^+$ T cells contained in the tumors showed a lack of tissue-residency and activation markers, cytolytic enzymes, and immune checkpoint molecules, hallmarks associated with enhanced efficacy of immune checkpoint blockade therapy. In contrast to Steele *et al.* that reported heterogenous expression of immune checkpoints in PDAC tissues[15], we showed that the majority of the $CD8^+$ T cells lacked expression of immune checkpoint molecules ICOS, CD28, CD27, and PD-1. The low amount of infiltrating PD-1 expressing immune cells in PDAC tissues was recently confirmed by conventional IHC.[9] Moreover, $CD8^+$ and $CD4^+$ T cells co-expressing CD39 and CD103 were only observed in one patient that showed a higher mutational load than the remaining samples, suggesting an absence of tumor-reactive T cells in the majority of pancreatic tumors. Finally, and in line with other studies, we found that γδ T cells comprised a small proportion of all T cells in PDAC.[9,15]

B cells and $CD4^+CD25^+CD127^{low}ICOS^+$ cells, identified as being mainly comprised of regulatory T cells,[22] were more frequent in PDAC tissues as compared to non-malignant pancreatic tissues. The role of B cells in PDAC remains controversial. Murine studies have reported on tumor-promoting roles for B cells,[28-30] whereas human PDAC data indicated that tumor-infiltrating B cells are a positive prognostic factor.[31,32] Especially the location of B cells in TLS was correlated with a longer survival.[31] In our cohort of pancreatic tumors, the increase in B cells was not related to the presence of TLS. The high relative frequency of regulatory T cells in PDAC is consistent with previous observations in mouse models and human, and in agreement with the notion that regulatory T cells, together with myeloid cells, are key immunosuppressive components in PDAC.[33-36] More importantly, the regulatory T cells in pancreatic tumors expressed high levels of ICOS, which highlights the potential use of ICOS antagonists to inhibit Treg interactions with ICOSL (such as NCT03829501).

Within the innate compartment, we discovered a previously unappreciated ILC1-like population, $CD127^-CD103^+CD39^+CD45RO^+$, that was specifically found in pancreatic cancers and not in matched regional lymph nodes, spleen, portal vein blood, or peripheral blood. Interestingly, these cells resembled ILC1-like cells found in colorectal cancer expressing

high levels of granzyme B and perforin.[22] Here, we found that the ILC1-like cells produced lower amounts of cytolytic enzymes, possibly related to the highly immunosuppressive microenvironment in PDAC. Importantly, these ILCs did not express the immune checkpoint molecule PD-1, in contrast to a recent study reporting on the existence of PD-1$^+$ ILC2 in mouse models of PDAC.[37] As opposed to the CD103$^+$CD39$^+$CD45RO$^+$ ILC1-like cells in PDAC tissues, ILCs in the matched non-malignant pancreatic tissue demonstrated a CD103$^-$CD39$^-$ CD45RA$^+$ NK-like phenotype. Further research needs to elucidate the specific mechanism by which these cells infiltrate or differentiate in PDAC; whether the ILCs acquire the CD103$^+$CD39$^+$CD45RO$^+$ phenotype locally in PDAC tissues differentiating from CD103$^-$ CD39$^-$CD45RA$^+$ NK cells, or whether they are already CD45RO$^+$ but acquire markers of tumor-residence and activation (CD103/CD39) later.

A unique feature of this study is the sample collection of portal vein blood in addition to peripheral blood before and after surgery. Intriguingly, we found that immune profiles of portal vein blood mirrors the increased frequency of memory B cells and regulatory T cells in pancreatic cancers, suggesting a regional enrichment of immune cells involved in the anti-tumor immune response. Furthermore, the only patient that harbored PD-1$^+$CD8$^+$ T cells in the tumor, also showed the highest relative frequency of PD-1$^+$CD8$^+$ T cells in the matched portal blood sample. This highlights a possible novel source of T cells with tumor-reactive phenotypes. It would be of interest to study if the PD-1$^+$CD8$^+$ T cells found in portal vein blood include tumor-reactive T cells.

In conclusion, our results confirm the unconventional nature of the immune microenvironment in PDAC, highlighting the involvement of ILC1-like cells and B cells. Larger patient cohorts are required to determine the functional role of ILC1-like cells and B cells in pancreatic cancers and their possible contribution to immunotherapeutic response. Furthermore, the regional enrichment in portal vein blood of immune cells infiltrating the PDAC microenvironment might provide an interesting source of tumor-associated immune cells that could be exploited in the context of adoptive T cell transfer.

METHODS

Human samples

A total of 11 PDAC patients, of which none received neoadjuvant chemo(radio)therapy, were recruited at the Leiden University Medical Center through informed consent following approval by the Medical Ethical Committee of the Leiden University Medical Center (protocol P17.047 and B18.049). All specimens were anonymized and handled according to the ethical guidelines described in the Code for Proper Secondary Use of Human Tissue in the Netherlands of the Dutch Federation of Medical Scientific Societies. Fresh unfixed primary PDAC tissues (N=11) with matched regional lymph nodes adjacent to the common hepatic artery, hepatoduodenal ligament or abdominal aorta (N=19), adjacent non-malignant pancreatic tissue (N=6), spleen (N=2), pre- and post-surgical peripheral blood samples (N=11 and N=10, respectively, with post-surgical blood withdrawn between three and six days post operatively), as well as portal vein blood samples (N=9) and one-year post-operative blood (N=1) were processed for this study (**Table 1**, **Figure S1**). Guided by specialized pathologists (AF and HM) macroscopic sectioning was performed to define the tumorous areas for specimen retrieval and subsequent further processing. After sectioning, the chosen tumor samples retrieved were partially snap frozen. The fresh frozen sections were cut and stained with hematoxylin and eosin (H&E) to determine type of tumor, tumor percentage as well as other factors that might play a role in the immune cell composition (**Table 2**).

Table 1. Different types of samples collected from respective patients.

Sample	Tumor	Normal	Lymph node 1	Lymph node 2	Lymph node 3	Spleen	PBMC portal	PBMC pre-surgery	PBMC post-surgery	PBMC post-surgery 1y
ISPIC1	+		+	+			+	+	+	+
ISPIC5	+	+	+	+	+		+	+	+	
ISPIC12	+	+				+		+	+	
ISPIC17	+		+	+	+		+	+	+	
ISPIC18	+	+	+	+			+	+	+	
ISPIC20	+	+	+	+			+	+		
ISPIC21	+					+		+	+	
ISPIC22	+	+		+	+		+	+	+	
ISPIC23	+		+		+		+	+	+	
ISPIC24	+		+				+	+	+	
ISPIC26	+	+	+		+		+	+	+	

Lymph node 1, arteria hepatica; lymph node 2, ligamentum hepatoduodenale; lymph node 3, aortacaval.

Table 2. Pathological characteristics of resected specimens from the respective patients.

Sample	Gender	Age (years)	Stage	Location	TNM	Type of specimen	Neoadjuvant (chemo) therapy	History of malignancy	Type of malignancy	Mucine	Pancreatitis in tumor specimen	Dirty necrosis	TLS
ISPIC1	Male	52	IIB	Head	pT1N1M0	PPPD/Whipple	No			+	–	–	+
ISPIC5	Female	49	IIB	Head	pT3N1M0	PPPD/Whipple	No			–	–	+	–
ISPIC12	Male	64	IIB	Tail	pT2N1M0	Lap.distal pancreatectomy	No	Yes	Colon	–	+	–	–
ISPIC17	Male	68	IA	Head	pT1cN0M0	PPPD/Whipple	No			+	+	+	+
ISPIC18	Female	75	IIB	Head	pT1cN2M0	PPPD/Whipple	No			–	–	+	–
ISPIC20	Male	73	IB	Head	pT2N0M0	PPPD/Whipple	No			–	–	–	–
ISPIC21	Female	68	IIB	Tail	pT2N2M0	Open distal pancreatectomy	No			+	–	–	–
ISPIC22	Male	77	IV	Head	pT2N2M1	PPPD/Whipple	No	Yes	Prostate	–	–	–	–
ISPIC23	Male	54	IIB	Head	pT2N2M0	PPPD/Whipple	No			–	–	+	+
ISPIC24	Male	71	IIB	Head	pT2N1M0	PPPD/Whipple	No			–	–	+	–
ISPIC26	Male	54	IIB	Head	pT2N2M0	PPPD/Whipple	No			–	–	+	–

TLS, tertiary lymphoid structure; TNM, tumor node metastasis; PPPD, pylorus-preserving pancreaticoduodenectomy.

Tissue processing

As mentioned in the previous section, a fraction of the tumor samples was snap-frozen, another part was cut into small fragments and digested using 1 mg/mL collagenase D (Roche, Basel, Switzerland) and 50 µg/mL DNAse I (Roche) in IMDM medium (Lonza BioWhittaker, Breda, The Netherlands) supplemented with 2 mM Glutamax (Thermo Fisher Scientific, Waltham, MA, US), 20% Fetal Bovine Serum (Sigma-Aldrich, Saint Louis, MO, US), 1% penicillin/streptomycin (Thermo Fisher Scientific), 1% Fungizone (Thermo Fisher Scientific), 0.1% Ciprofloxacin (provided by the LUMC pharmacy), and 0.1% Gentamicin (Sigma-Aldrich). Tissue fragments were incubated for 30 minutes at 37°C interrupted by three mechanical dissociations on a gentleMACS Dissociator (Miltenyi Biotec, Bergisch Gladback, Germany) in gentleMACS C tubes (Miltenyi Biotec), and subsequently processed through a 70 µm strainer (Miltenyi Biotec). In parallel, the lymph node tissue and spleen samples were processed in the same way. Single cell digests and remaining tumor fragments were cryopreserved for analysis and culturing at later stages.

Blood samples were obtained at their respective timepoints. Peripheral blood mononuclear cells (PBMCs) were isolated from patients' heparinized venous blood by Ficoll-Amidotrizoate (provided by the LUMC pharmacy) gradient centrifugation.

Mass cytometry antibody staining and data acquisition

To decipher the immune composition in the different samples from patients with PDAC, mass cytometric analysis was performed with 41 immune cell markers covering immune lineage markers, Fc/complement receptors, differentiation/activation markers, cytokine/chemokine receptors, immunomodulatory molecules, and adhesion/homing molecules (**Figure S1**). Briefly, single-cell suspensions were thawed, after which Percoll (GE Healthcare) density-gradient centrifugation was performed to isolate immune cells from PDAC and non-malignant pancreatic tissue. Cells were washed in Maxpar Cell Staining Buffer (CSB, Fluidigm) and counted. Up to 3 million cells of each sample were incubated with 1 mL CSB containing 1 µM Cell-ID intercalator-103Rh (Fluidigm) for 15 min at room temperature (rT) to discriminate dead from live cells. Cells were washed in CSB, incubated with human Fc receptor block (BioLegend) for 10 min at rT, and stained with a cell-surface antibody cocktail for 45 min at rT in a final volume of 100 µL. The antibodies are listed in **Table S1**. Purified antibodies were conjugated in-house with heavy metal reporters using the MaxPar X8 Antibody Labeling Kit (Fluidigm) according to the manufacturer's instructions. All antibodies were titrated to determine the optimal concentration. After three washing steps in CSB, cells were incubated with 1 mL Maxpar Fix and Perm buffer (Fluidigm Sciences) containing 0.125 µM Cell-ID intercalator-Ir (Fluidigm) overnight at 4°C to discriminate singlets from doublets. The next day, cells were washed three times in CSB, and one time in de-ionized water immediately prior to data acquisition. Cells were acquired on a Helios mass cytometer (Fluidigm) at an event rate of <500 events/sec in de-ionized water containing 10x diluted EQ Four Element Calibration Beads (Fluidigm). Data were normalized with the normalization passport EQ-P13H2302_ver2 for each experiment.

Mass cytometry data analysis

Normalized FCS files were analyzed in the FlowJo software version 10.6.1 (Tree Star Inc). Data were checked for quality of staining and gated for live, single, $CD45^+$ cells using the 191/193Ir DNA intercalator, CD45, residual, center, width, event length, 103Rh DNA intercalator, and 140Ce bead channels (gating strategy shown in **Figure S1**). The final $CD45^+$ gate was exported for each sample as FCS file for downstream analysis (**Figure S1**). To account for technical variation, a PBMC reference sample was included in every mass cytometry experiment (batch). ComBat was applied to correct for batch effects by aligning the PBMC reference samples and corresponding patient samples across all batches.[38] $CD45^+$ cells were hyberbolic ArcSinh transformed with a cofactor of 5, sample-tagged, and subjected to dimensionality reduction analysis in the Cytosplore software[39]. All 41 antibodies in the panel showed consistent staining over time and were included in the analysis.

First, CD45$^+$ data from all samples (18.2x10^6 cells) were subjected to a 5-level hierarchical stochastic neighbor embedding (HSNE) analysis[40,41] with default perplexity and iterations. One portal vein blood sample with an outlying number of CD45$^+$ cells (3941; ISPIC5) was excluded from further data analysis. Due to the large number of input cells, we first identified global clusters of naïve CD8 T cells, CD8 memory T cells, naïve CD4 T cells, memory CD4 T cells, γδ T cells, Lin$^-$CD7$^+$ ILCs, B cells, and myeloid cells (**Figure S2**). These clusters were separately analyzed in a data-driven manner up to a maximum number of 0.5x10^6 cells or landmarks. Clustering of the cells was performed with the Gaussian mean shift (GMS) algorithm, and clusters that showed high similarity in ArcSinh5-transformed median expression of all markers were merged. Second, CD45$^+$ cells from PDAC and non-malignant pancreatic tissue (N=11 and N=6, respectively) were subjected to a t-distributed stochastic neighbor embedding (t-SNE) analysis[42] with a total of 155,433 cells analyzed with default perplexity (30) and iterations (1000). Clustering of the cells with a sigma value of 20 resulted in the identification of different immune subsets across the major immune lineages of CD4$^+$ T cells (N=18), CD8$^+$ T cells (N=16), CD25$^+$CD127^{low}ICOS$^+$ CD4$^+$ T cells (N=1), γδ T cells (N=1), Lin$^-$CD7$^+$ innate lymphoid cells (ILCs, N=9), B cells (N=7), and myeloid cells (N=16). Absolute cell counts of these immune cell clusters are provided in **Figure S3**. Third, CD45$^+$ data from all blood samples (10.9x10^6 cells) were subjected to a 5-level HSNE analysis with default perplexity and iterations. Here, we also first identified global clusters of the major immune lineages, which were separately analyzed in a data-driven manner up to a maximum number of 0.5x10^6 cells or landmarks. Clusters that showed high similarity in ArcSinh5-transformed median expression of all markers were merged. To visualize phenotypic differences among the blood samples, we additionally performed t-SNE analyses on each major immune lineage where we downsampled to a maximum number of 0.5x10^6 cells. For all analyses, frequencies of the clusters were calculated and hierarchical clustering thereof was performed using Spearman's rank correlation in Matlab version R2016a 9.0.

A dual t-SNE analysis was performed to quantify the similarity between individual samples based on the cluster composition.[43] The samples t-SNE map is constructed using a data matrix ($N_{samples} \times K_{clusters}$) containing the cluster frequencies of the individual samples, hence samples with similar cluster composition profiles end up close together in the map. The data matrix was normalized by centering to zero mean and scaling to unit variance. The normalized data matrix was transposed and used to generate the clusters t-SNE map, hence clusters with similar profiles across individual samples end up close together in the map.

Flow cytometry

Single-cell suspensions of PDAC tissues (N=5) were analyzed by flow cytometry for the expression of cytotoxic molecules and pro-inflammatory cytokines by ILC1-like cells and CD8$^+$ and CD8$^-$ T cells at baseline *versus* upon stimulation. Briefly, cells were thawed and rest at 37°C in IMDM/L-glutamine medium (Lonza) complemented with 10% human serum for 1h. Thereafter, cells were stimulated with 20 ng/mL PMA (Sigma-Aldrich) and 1 µg/mL ionomycin (Sigma-Aldrich) for 6 hours at 37°C. Ten µg/mL brefeldin A (Sigma-Aldrich) was added for the last 4 hours. Next, cells were washed in FACS buffer (PBS (Fresenius Kabi)/1% FCS) and incubated with human Fc receptor block (BioLegend) for 10 min at 4°C. Thereafter, cells were stained with cell surface antibodies (1:20 anti-CD7-V450 [clone M-T701, BD Biosciences], 1:20 anti-CD3-Am Cyan [clone SK7, BD Biosciences], 1:100 anti-CD8α-BV605 [clone SK1, BD Biosciences], 1:25 anti-CD127-BV711 [clone A019D5, BioLegend], 1:10 anti-CD103-FITC [clone Ber-ACT8], 1:20 anti-CD45RO-PerCP-Cy5.5 [clone UCHL1, Sony], 1:20 anti-CD45RA-PE/Dazzle [clone HI100, Sony], 1:150 anti-CD56-APC-R700 [clone NCAM16.2, BD Biosciences], and a 1:1000 nIR viability stain [Life Technologies]) for 45 min at 4°C. After two washing steps in FACS buffer (PBS/1% FCS), intracellular staining was performed using Fixation Buffer and Intracellular Staining Permeabilization Wash Buffer (BioLegend) with 1:20 anti-Perforin-PE [clone δG9, BD Biosciences], 1:50 anti-Granzyme B-PE [clone GB11, eBioscience], 1:300 anti-TNFα-PE-Cy7 [clone Mab11, eBioscience], and 1:50 anti-IFNγ-APC [clone 25723.11, BD Biosciences] for 20 min at rT, followed by three washing steps in permwash

buffer. In addition, single-cell suspensions of PDAC tissues (N=7) were analyzed by flow cytometry for the expression of FOXP3 by $CD25^{+}CD127^{low}$ and $CD25^{+}CD127^{low}ICOS^{+}$ $CD4^{+}$ T cells. After incubation with human Fc receptor block, cells were stained with the following cell surface antibodies for 45 min at 4°C: 1:100 anti-CD4-BV421 [clone RPA-T4, Sony], 1:20 anti-CD45RO-PerCP-Cy5.5 [UCHL1, Sony], 1:300 anti-ICOS-PE [clone C398.4A, BioLegend], 1:150 anti-CD127-PE-Cy7 [clone AO1905, BioLegend], 1:20 anti-CD25-APC [clone 2A3, BD Biosciences], and a 1:1000 nIR viability stain [Life Technologies]. After washing, cells were stained using the FOXP3 Transcription Factor Staining Buffer Set (eBioscience) with the following antibodies: 1:20 anti-FOXP3-FITC [clone PCH101, Thermo Fisher Scientific] and 1:20 anti-T-bet-BV605 [clone 4B10, BioLegend]. Compensation in flow cytometry experiments was carried out with CompBeads (BD Biosciences) and ArC reactive beads (Life Technologies). Cells were acquired on a FACS LSR Fortessa 4L (BD Biosciences) running FACSDiva software version 8.0 (BD Biosciences). Data were analyzed with FlowJo software version 10.6.1 (Tree Star Inc).

Multispectral immunofluorescence

Multispectral immunofluorescence (IF) analysis was performed on 4-µm FFPE tissue sections. The FFPE tissue sections were handled and stained as described previously.[44] For the ILC panel FFPE tissue sections were deparaffinized with xylene and washed in ethanol. Heat-induced antigen retrieval in citrate buffer (10 mM, pH 6) was performed and the slides were allowed to cool down to room temperature. Subsequently, the tissues were blocked with Superblock buffer (Thermo Fisher Scientific) and incubated overnight at 4 degrees with the following primary antibodies: anti-CD103 (1:100 dilution, EPR4166(2)), Abcam, Cambridge, UK), anti-CD45RO (1:50 dilution, UCHL1, Cell Signalling Technology, Danvers, MA, USA). After incubation, slides were washed in PBS and detection with the following fluorescent secondary antibodies was performed: Alexa Fluor 680-labelled anti-rabbit antibody (1:400 dilution, Thermo Fisher Scientific) and CF633-labelled goat-anti-mouse IgG2a antibody(1:400 dilution, Biotium, Fremont, CA, USA). After washing, tissues were incubated with directly conjugated primary antibodies: anti-CD7 (EPR4242, Abcam) labelled with Alexa Fluor 647 (1:100 dilution, Thermo Fisher Scientific) and anti-CD3 (D7A6E, Cell Signalling Technology) labelled with Alexa Fluor 594 (1:50 dilution, Thermo Fisher Scientific). Alexa Fluor labelling kits were employed to label anti-CD7 and anti-CD3 antibodies (Thermo Fisher Scientific, catalogue numbers A20186 and A10239, respectively).

Image acquisition and cell counting

For each tumor, 3 different regions of interest (selected by an experienced pathologist) within the tumor microenvironment were imaged at 20x magnification with the Vectra 3.0 Automated Quantitative Pathology Imaging System (Perkin Elmer). An analysis algorithm was trained manually for cell segmentation as well as immunophenotyping of cells. Cellular segmentation was performed using a counterstain-based approach with DAPI to segment nuclei and membrane markers (CD3, CD7) to detect cell contours. All images were then visually inspected for the number of $CD7^{+}CD3^{-}CD45RO^{+}CD103^{+}$ ILC1-like cells. For each case, cell counts were normalized by tissue area (number of cells/mm^2).

Whole-exome sequencing analysis

Exomes of PDAC and respective normal samples were captured using the Twist Human Core Exome kit (Twist Bioscience) and paired-end reads were sequenced on the NovaSeq 6000 system. Raw reads were then mapped with bwa-mem to the hg38 reference genome. Variant calling was done by muTect[45]. Finally, coding variants were functionally annotated using the ensembl Variant Effect Predictor (VEP)[46].

Statistical analysis

Cell frequencies were presented either as mean with lines indicating matched samples, as mean with standard error of the mean (SEM), or as median with interquartile range (IQR), as specified in the legend. Comparisons of cell frequencies between PDAC and non-malignant pancreatic tissue were performed by Mann-Whitney tests, as matched non-malignant samples were only available from 6 out of the total

of 11 patients. Comparisons of cell frequencies between the different blood samples were performed by Wilcoxon matched-pairs signed rank tests. Statistical tests were conducted in GraphPad Prism (version 9.0.1) and P values <0.05 were considered statistically significant.

ACKOWLEDGEMENTS
We thank the Flow cytometry Core Facility of the Leiden University Medical Center for their help with mass cytometry experiments.

AUTHOR CONTRIBUTIONS
All authors contributed to the study conception and design. Material withdrawal, preparation, and analysis was performed by T.B., N.L.d.V., A.F., H.M., B.B. and A.V. Data collection and analysis was performed by T.B., N.L.d.V., T.A., R.T.K., Z.L., D.R., B.L., F.K. and N.F.d.M. The manuscript was written by T.B. and N.L.d.V. under supervision of F.K. and N.F.d.M. with input of all authors.

COMPETING INTEREST
None to declare.

REFERENCES

1 Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68, 394-424, doi:10.3322/caac.21492 (2018).

2 Rahib, L. *et al.* Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* 74, 2913-2921, doi:10.1158/0008-5472.CAN-14-0155 (2014).

3 Ferlay, J., Partensky, C. & Bray, F. More deaths from pancreatic cancer than breast cancer in the EU by 2017. *Acta Oncol* 55, 1158-1160, doi:10.1080/0284186X.2016.1197419 (2016).

4 Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer Statistics, 2021. *CA Cancer J. Clin.* 71, 7-33, doi:https://doi.org/10.3322/caac.21654 (2021).

5 Antonia, S. J. *et al.* Four-year survival with nivolumab in patients with previously treated advanced non-small-cell lung cancer: a pooled analysis. *The Lancet. Oncology* 20, 1395-1408, doi:10.1016/S1470-2045(19)30407-3 (2019).

6 Schadendorf, D. *et al.* Pooled Analysis of Long-Term Survival Data From Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma. *J Clin Oncol* 33, 1889-1894, doi:10.1200/JCO.2014.56.2736 (2015).

7 von Pawel, J. *et al.* Long-term survival in patients with advanced non-small-cell lung cancer treated with atezolizumab versus docetaxel: Results from the randomised phase III OAK study. *Eur J Cancer* 107, 124-132, doi:10.1016/j.ejca.2018.11.020 (2019).

8 Rosenberg, S. A. & Restifo, N. P. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science* 348, 62-68, doi:10.1126/science.aaa4967 (2015).

9 Blando, J. *et al.* Comparison of immune infiltrates in melanoma and pancreatic cancer highlights VISTA as a potential target in pancreatic cancer. *Proc Natl Acad Sci U S A* 116, 1692-1697, doi:10.1073/pnas.1811067116 (2019).

10 Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* 500, 415-425, doi:10.1038/nature12477 (2013).

11 Feig, C. *et al.* The pancreas cancer microenvironment. *Clinical Cancer Research* 18, 4266-4276, doi:10.1158/1078-0432.CCR-11-3114.The (2012).

12 Ryan, D. P., Hong, T. S. & Bardeesy, N. Pancreatic Adenocarcinoma. *N Engl J Med* 371, 1039-1049, doi:10.1056/NEJMra1404198 (2014).

13 Stromnes, I. M., Hulbert, A., Pierce, R. H., Greenberg, P. D. & Hingorani, S. R. T-cell Localization, Activation, and Clonal Expansion in Human Pancreatic Ductal Adenocarcinoma. *Cancer Immunol Res* 5, 978-991, doi:10.1158/2326-6066.CIR-16-0322 (2017).

14 Bernard, V. *et al.* Single Cell Transcriptomics of Pancreatic Cancer Precursors Demonstrates Epithelial and Microenvironmental Heterogeneity as an Early Event in Neoplastic Progression. *Clinical Cancer Research*, doi:10.1158/1078-0432.ccr-18-1955 (2018).

15 Steele, N. G. *et al.* Multimodal Mapping of the Tumor and Peripheral Blood Immune Landscape in Human Pancreatic Cancer. *Nat Cancer* 1, 1097-1112, doi:10.1038/s43018-020-00121-4 (2020).

16 Sivakumar, S. *et al.* Activated Regulatory T-Cells, Dysfunctional and Senescent T-Cells Hinder the Immunity in Pancreatic Cancer. *Cancers (Basel)* 13, doi:10.3390/cancers13081776 (2021).

17 O'Reilly, E. M. *et al.* Durvalumab With or Without Tremelimumab for Patients With Metastatic Pancreatic Ductal Adenocarcinoma: A Phase 2 Randomized Clinical Trial. *JAMA Oncol*, doi:10.1001/jamaoncol.2019.1588 (2019).

18 Brahmer, J. R. *et al.* Safety and Activity of Anti-PD-L1 Antibody in Patients with Advanced Cancer. *N Engl J Med* 366, 2455-2465, doi:10.1056/NEJMoa1200694.Safety (2012).

19 Royal, R. E. *et al.* Phase 2 trial of single agent Ipilimumab (anti-CTLA-4) for locally advanced or metastatic pancreatic adenocarcinoma. *J Immunother* 33, 828-833, doi:10.1097/CJI.0b013e3181eec14c (2010).

20 Le, D. T. *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409-413, doi:10.1126/science.aan6733 (2017).

21 Le, D. T. *et al.* Results from a Phase IIb, Randomized, Multicenter Study of GVAX Pancreas and CRS-207 Compared with Chemotherapy in Adults with Previously Treated Metastatic Pancreatic Adenocarcinoma (ECLIPSE Study). *Clin Cancer Res* 25, 5493-5502, doi:10.1158/1078-0432.ccr-18-2992 (2019).

22 de Vries, N. L. *et al.* High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of antitumour immunity. *Gut* 69, 691-703, doi:10.1136/gutjnl-2019-318672 (2020).

23 Duhen, T. *et al.* Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat Commun* 9, 2724, doi:10.1038/s41467-018-05072-0 (2018).

24 Simoni, Y. *et al.* Bystander CD8(+) T cells are abundant and phenotypically distinct in human

tumour infiltrates. *Nature* 557, 575-579, doi:10.1038/s41586-018-0130-2 (2018).

25 Gaudillière, B. *et al.* Clinical recovery from surgery correlates with single-cell immune signatures. *Science translational medicine* 6, 255ra131, doi:10.1126/scitranslmed.3009701 (2014).

26 Fergusson, J. R. *et al.* CD161 defines a transcriptional and functional phenotype across distinct human T cell lineages. *Cell Rep* 9, 1075-1088, doi:10.1016/j.celrep.2014.09.045 (2014).

27 Carstens, J. L. *et al.* Spatial computation of intratumoral T cells correlates with survival of patients with pancreatic cancer. *Nat Commun* 8, 15095, doi:10.1038/ncomms15095 (2017).

28 Pylayeva-Gupta, Y. *et al.* IL35-Producing B Cells Promote the Development of Pancreatic Neoplasia. *Cancer Discov* 6, 247-255, doi:10.1158/2159-8290.Cd-15-0843 (2016).

29 Lee, K. E. *et al.* Hif1a Deletion Reveals Pro-Neoplastic Function of B Cells in Pancreatic Neoplasia. *Cancer Discov* 6, 256-269, doi:10.1158/2159-8290.Cd-15-0822 (2016).

30 Gunderson, A. J. *et al.* Bruton Tyrosine Kinase-Dependent Immune Cell Cross-talk Drives Pancreas Cancer. *Cancer Discov* 6, 270-285, doi:10.1158/2159-8290.Cd-15-0827 (2016).

31 Castino, G. F. *et al.* Spatial distribution of B cells predicts prognosis in human pancreatic adenocarcinoma. *Oncoimmunology* 5, e1085147, doi:10.1080/2162402x.2015.1085147 (2016).

32 Tewari, N. *et al.* The presence of tumour-associated lymphocytes confers a good prognosis in pancreatic ductal adenocarcinoma: an immunohistochemical study of tissue microarrays. *BMC Cancer* 13, 436, doi:10.1186/1471-2407-13-436 (2013).

33 Liyanage, U. K. *et al.* Prevalence of regulatory T cells is increased in peripheral blood and tumor microenvironment of patients with pancreas or breast adenocarcinoma. *J Immunol* 169, 2756-2761, doi:10.4049/jimmunol.169.5.2756 (2002).

34 Clark, C. E. *et al.* Dynamics of the immune reaction to pancreatic cancer from inception to invasion. *Cancer Res* 67, 9518-9527, doi:10.1158/0008-5472.Can-07-0175 (2007).

35 Zhang, Y. *et al.* CD4+ T lymphocyte ablation prevents pancreatic carcinogenesis in mice. *Cancer Immunol Res* 2, 423-435, doi:10.1158/2326-6066.Cir-14-0016-t (2014).

36 Zhang, Y. *et al.* Regulatory T-cell Depletion Alters the Tumor Microenvironment and Accelerates Pancreatic Carcinogenesis. *Cancer Discov* 10, 422-439, doi:10.1158/2159-8290.CD-19-0958 (2020).

37 Moral, J. A. *et al.* ILC2s amplify PD-1 blockade by activating tissue-specific cancer immunity. *Nature* 579, 130-135, doi:10.1038/s41586-020-2015-4 (2020).

38 Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8, 118-127, doi:10.1093/biostatistics/kxj037 (2007).

39 Höllt, T. *et al.* Cytosplore: Interactive Immune Cell Phenotyping for Large Single-Cell Datasets. *Computer Graphics Forum* 35, 171-180, doi:10.1111/cgf.12893 (2016).

40 Pezzotti, N., Höllt, T., Lelieveldt, B., Eisemann, E. & Vilanova, A. Hierarchical Stochastic Neighbor Embedding. *Computer Graphics Forum* 35, 21-30, doi:10.1111/cgf.12878 (2016).

41 van Unen, V. *et al.* Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types. *Nat Commun* 8, 1740, doi:10.1038/s41467-017-01689-9 (2017).

42 Van Der Maaten, L. J. P. & Hinton, G. E. Visualizing high-dimensional data using t-sne. *Journal of Machine Learning Research* 9, 2579-2605, doi:10.1007/s10479-011-0841-3 (2008).

43 van Unen, V. *et al.* Mass Cytometry of the Human Mucosal Immune System Identifies Tissue- and Disease-Associated Immune Subsets. *Immunity* 44, 1-13, doi:10.1016/j.immuni.2016.04.014 (2016).

44 Ijsselsteijn, M. E. *et al.* Cancer immunophenotyping by seven-colour multispectral imaging without tyramide signal amplification. *The journal of pathology. Clinical research* 5, 3-11, doi:10.1002/cjp2.113 (2019).

45 Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 31, 213-219, doi:10.1038/nbt.2514 (2013).

46 McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* 17, 122, doi:10.1186/s13059-016-0974-4 (2016).

SUPPLEMENTAL FIGURES

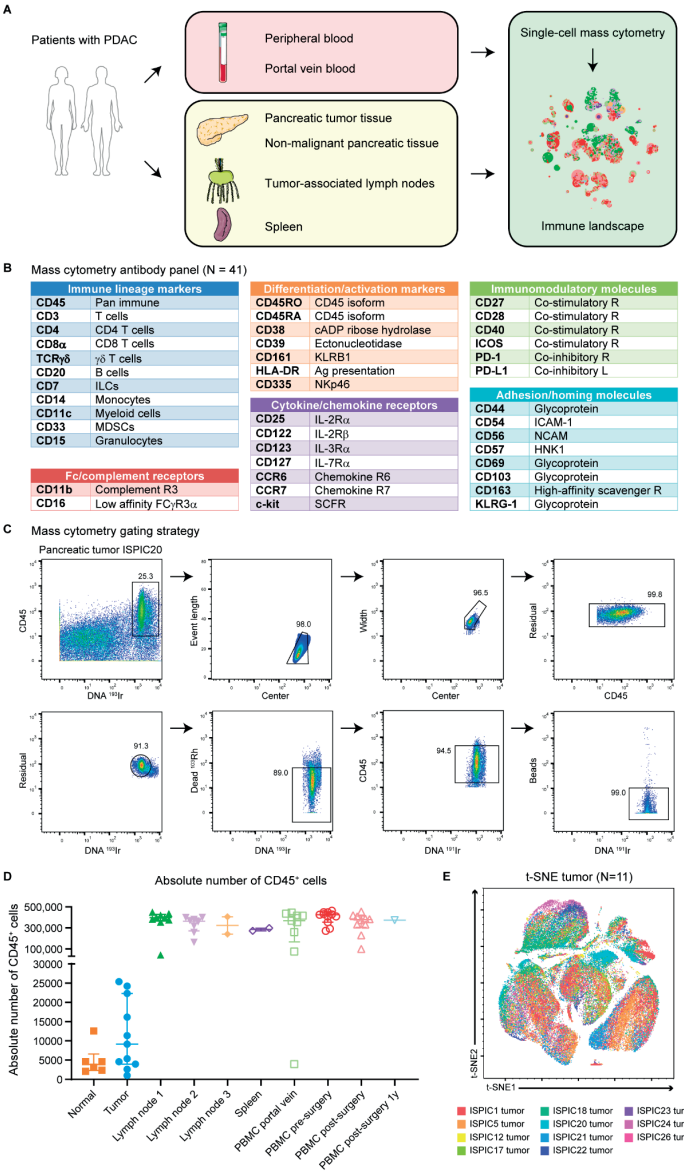


Figure S1. Overview of the study design, the mass cytometry antibody panel, and the gating strategy.
A. Schematic overview of the study design. **B.** Mass cytometry antibody panel composed of 41 immune cell markers. **C.** Mass cytometry gating strategy for single, live CD45⁺ cells of a representative pancreatic tumor sample showing sequential gates with percentages. **D.** Absolute number of CD45⁺ cells from non-malignant pancreatic (N=6) and PDAC (N=11) tissues, regional lymph nodes (details below), spleen (N=2), portal vein blood (N=9), peripheral blood before surgery (N=11), directly after surgery (N=10), and one year after surgery (N=1) from 11 patients with PDAC. Bars indicate the median with IQR. Each dot represents an individual sample. Data from 11 independent experiments with mass cytometry. Lymph node 1 derived from the common hepatic artery (N=9), lymph node 2 from the hepatoduodenal ligament (N=8), and lymph node 3 from the abdominal aorta (N=2). **E.** t-SNE embedding of **Figure 1A** showing the clustering of immune cells from PDAC tissues (N=11) colored by patient ID.

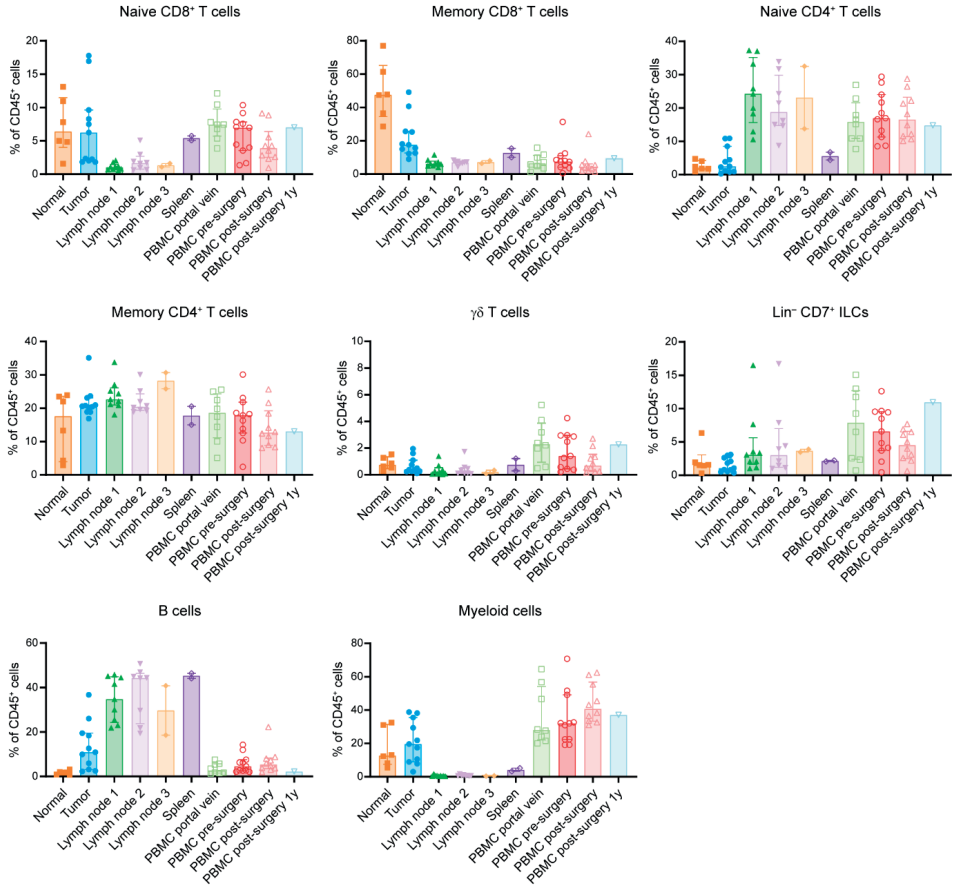


Figure S2. Frequencies of the major immune lineages in the different tissue types obtained from PDAC patients.

Frequencies of the major immune lineages across non-malignant pancreatic (N=6) and PDAC (N=11) tissues, regional lymph nodes (details below), spleen (N=2), portal vein blood (N=8), peripheral blood before surgery (N=11), directly after surgery (N=10), and one year after surgery (N=1) from 11 patients with PDAC. Frequencies were determined by HSNE analysis including all samples (details in methods), and are shown as percentage of total CD45⁺ cells. Bars indicate the median with IQR. Each dot represents an individual sample. Data from 11 independent experiments with mass cytometry. Lymph node 1 derived from the common hepatic artery (N=9), lymph node 2 from the hepatoduodenal ligament (N=8), and lymph node 3 from the abdominal aorta (N=2).

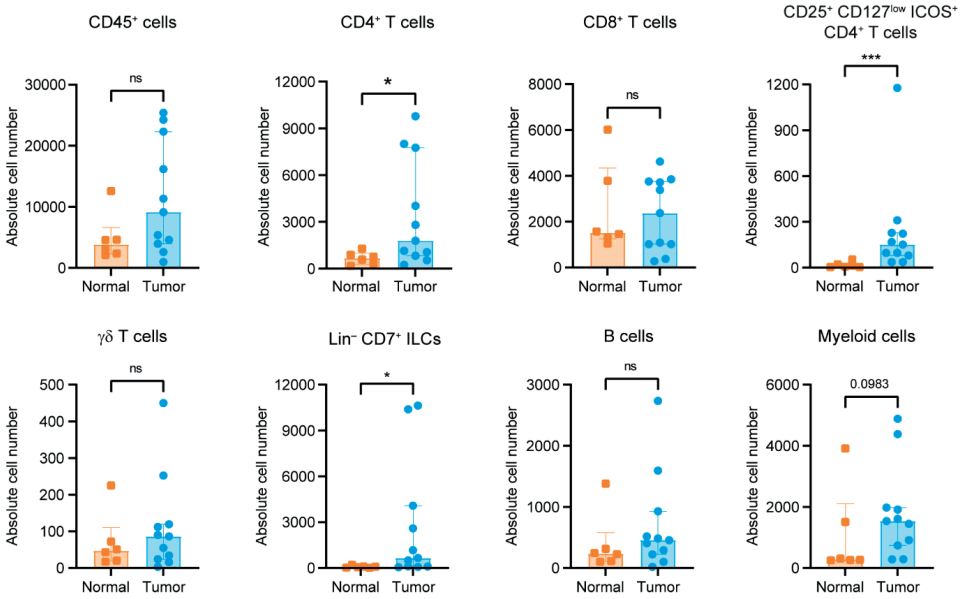


Figure S3. Absolute cell counts of the major immune lineages in PDAC and non-malignant pancreatic tissue. Absolute cell counts of the major immune lineages in non-malignant pancreatic (N=6) and PDAC (N=11) tissues as determined by mass cytometric analysis. Bars indicate the median with IQR. Data from 11 independent experiments with mass cytometry. *P<0.05, ***P<0.001 by Mann-Whitney test.

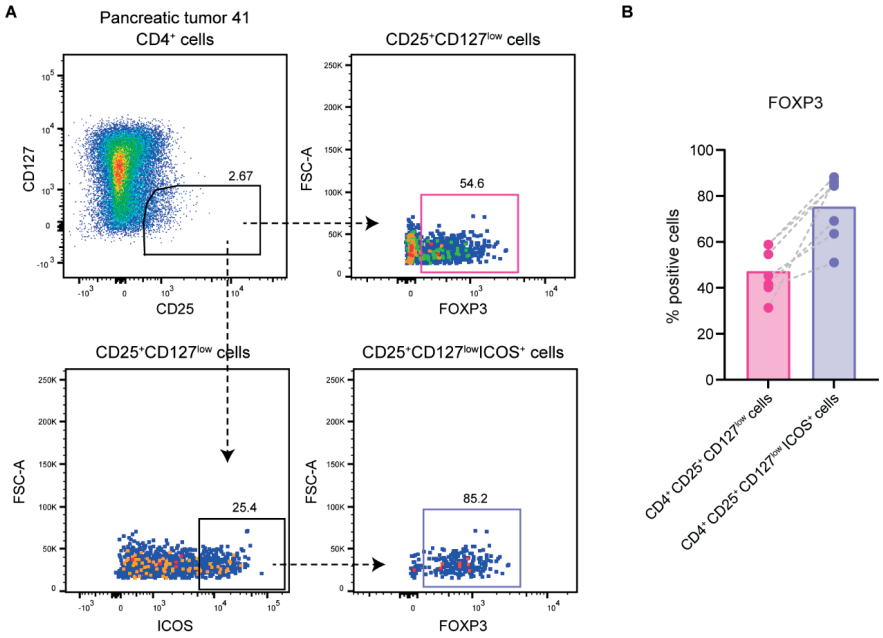


Figure S4. FOXP3 expression in CD25$^+$CD127^{low} and CD25$^+$CD127^{low}ICOS$^+$ CD4$^+$ T cells in pancreatic tumors. **A.** Flow cytometry gating strategy used to determine FOXP3-positive cells in a representative pancreatic tumor. **B.** Frequency of FOXP3-positive cells within CD25$^+$CD127^{low} CD4$^+$ T cells as well as within CD25$^+$CD127^{low}ICOS$^+$ CD4$^+$ T cells of pancreatic tumors (N=7).

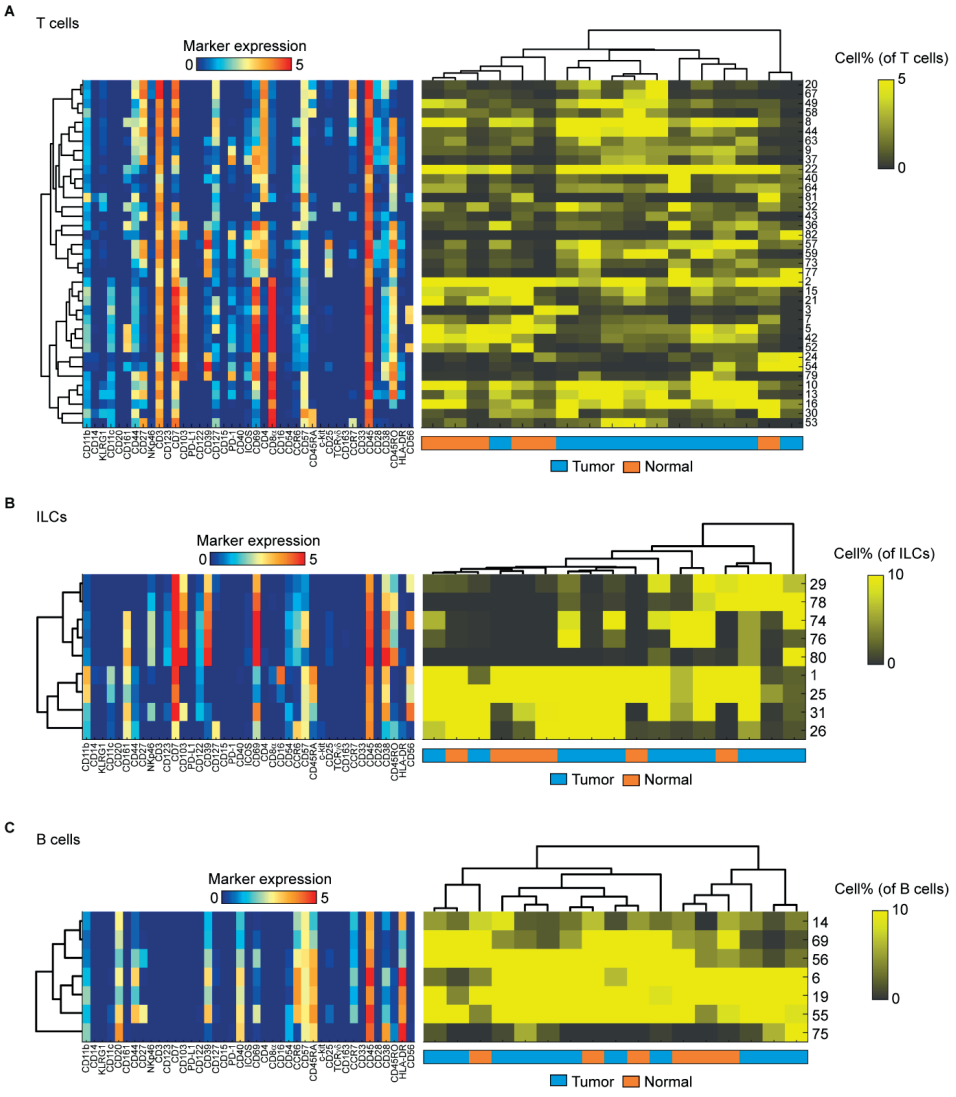


Figure S5. Marker expression and cell frequency of T cell, ILC, and B cell clusters identified in PDAC and non-malignant pancreatic tissues.

A. A heatmap showing median marker expression values (left) and frequencies of all identified T cell clusters among PDAC and non-malignant pancreatic tissues (right) belonging to **Figure 2**. **B.** A heatmap showing median marker expression values (left) and frequencies of all identified ILC clusters among PDAC and non-malignant pancreatic tissues (right) belonging to **Figure 3**. **C.** A heatmap showing median marker expression values (left) and frequencies of all identified B cell clusters among PDAC and non-malignant pancreatic tissues (right) belonging to **Figure 4**. Hierarchical clustering was performed on cluster frequencies using Spearman's rank correlation.

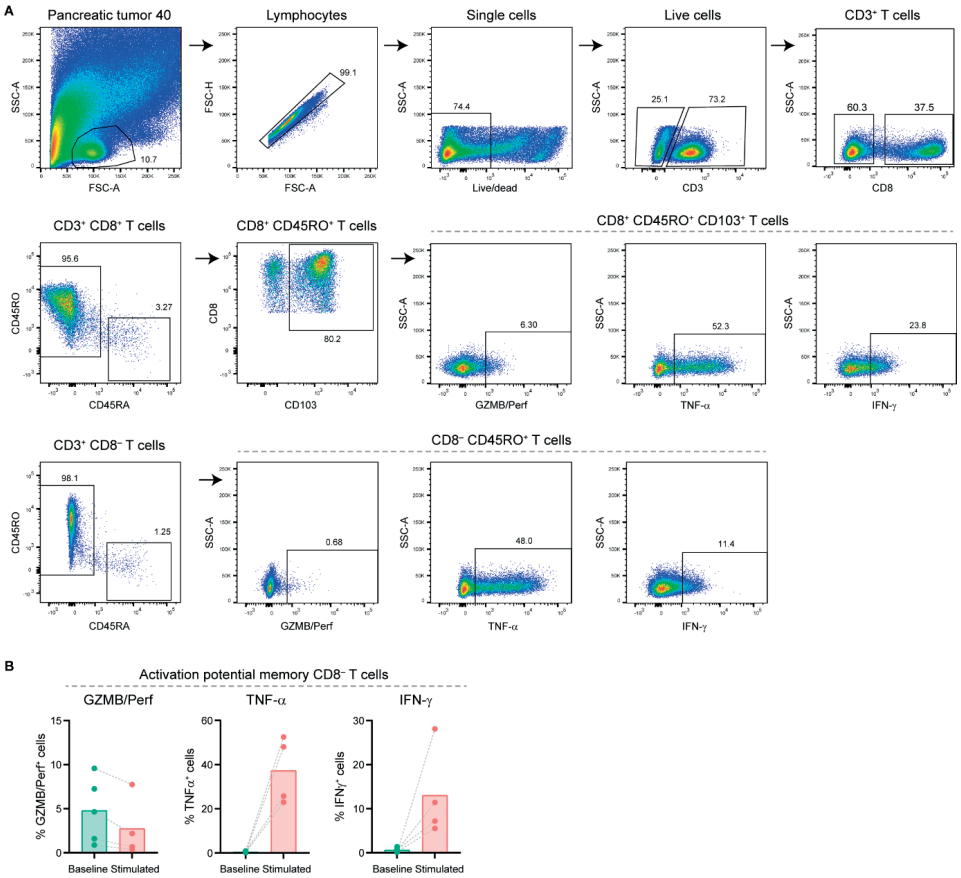


Figure S6. Gating strategy for the detection of cytotoxic molecules and cytokines by T cells.
A. Flow cytometry gating strategy to detect granzyme B/perforin, TNF-α, and IFN-γ in CD103⁺ (tissue-resident) memory CD8⁺ T cells and in memory CD8⁻ T cells of a representative pancreatic tumor sample (upon stimulation with PMA/ionomycin) showing sequential gates with percentages. **B.** Cytokine production by memory CD8⁻ T cells from PDAC tissues *ex vivo* (N=5) and upon stimulation with PMA/ionomycin (N=4) measured by flow cytometry. Bars indicate the mean and lines indicate matched samples.

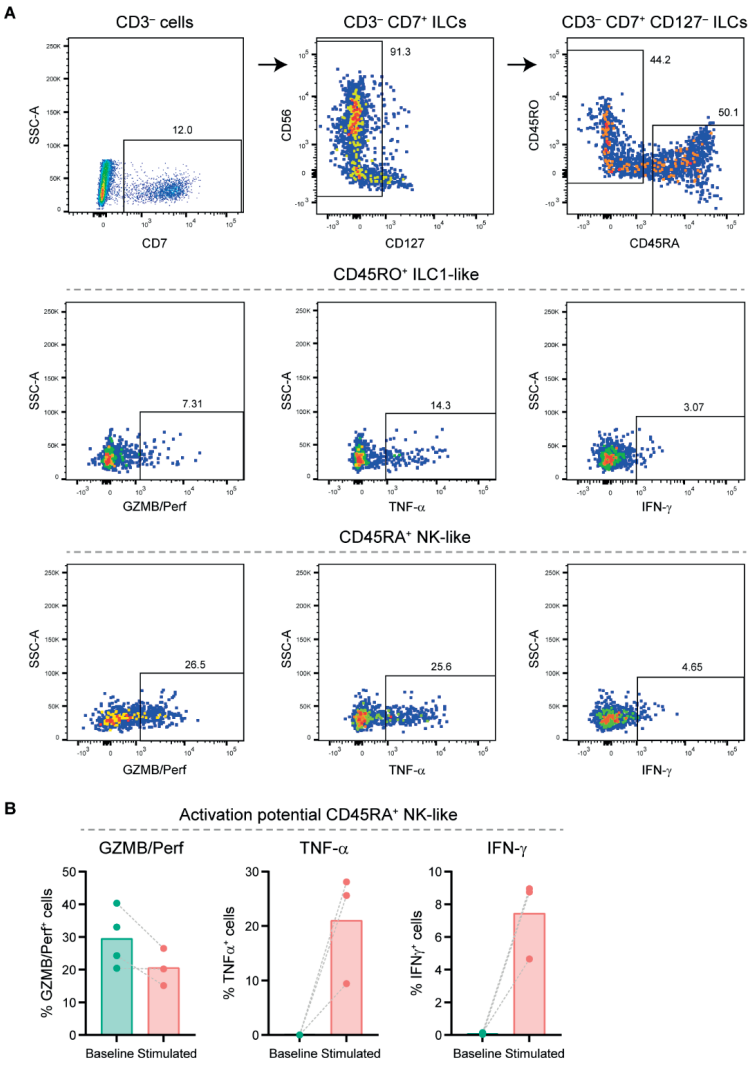


Figure S7. Gating strategy for the detection of cytotoxic molecules and cytokines by ILCs.
A. Flow cytometry gating strategy to detect granzyme B/perforin, TNF-α, and IFN-γ in CD45RO⁺ ILC1-like cells and CD45RA⁺ NK-like cells of a representative pancreatic tumor sample (upon stimulation with PMA/ionomycin) showing sequential gates with percentages, starting with the CD3⁻ gate from **Figure S6**. **B.** Cytokine production by CD45RA⁺ NK-like cells from PDAC tissues *ex vivo* (N=4) and upon stimulation with PMA/ionomycin (N=3) measured by flow cytometry. Bars indicate the mean and lines indicate matched samples.

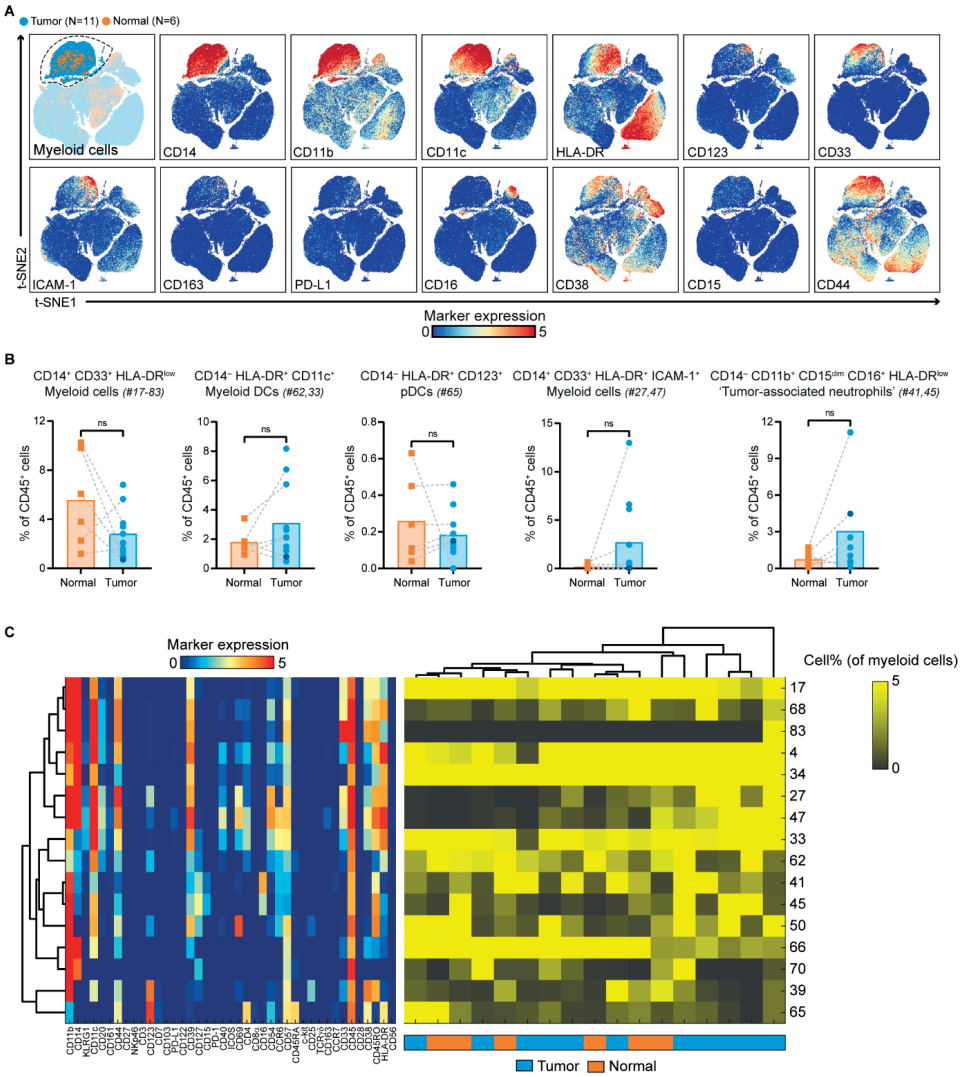


Figure S8. Myeloid cell phenotypes identified in PDAC and non-malignant pancreatic tissues.
A. t-SNE embedding from **Figure 1A** highlighting the myeloid cells, colored by tissue type (first plot) and relative expression of indicated myeloid cell markers (remainder). **B.** Frequencies of selected myeloid cell clusters in PDAC tissues (N=11) as compared to non-malignant pancreatic tissues (N=6) shown as percentage of total CD45+ cells. Cluster IDs correspond to the ones in (**C**). Each dot represents an individual sample (dark blue color represents patient ISPIC20). Bars indicate the mean and lines indicate matched samples. Data from 11 independent experiments with mass cytometry. **C.** A heatmap showing median marker expression values (left) and frequencies of all identified myeloid cell clusters among PDAC and non-malignant pancreatic tissues (right). Hierarchical clustering was performed on cluster frequencies using Spearman's rank correlation.

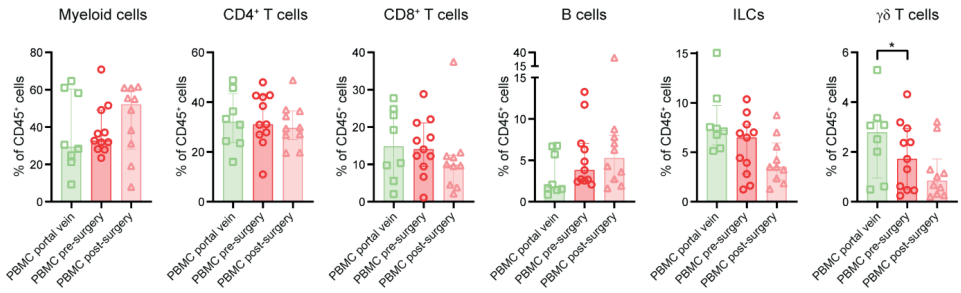


Figure S9. Frequency of the major immune lineages in portal vein and peripheral blood from patients with PDAC.

Frequencies of the major immune lineages in portal vein blood (N=8) and peripheral blood before (N=11) as well as after (N=10) surgery from 11 patients with PDAC shown as percentage of total CD45⁺ cells. Each dot represents an individual sample. Bars indicate the median with IQR. Data from 11 independent experiments with mass cytometry. *P<0.05 by Wilcoxon test.

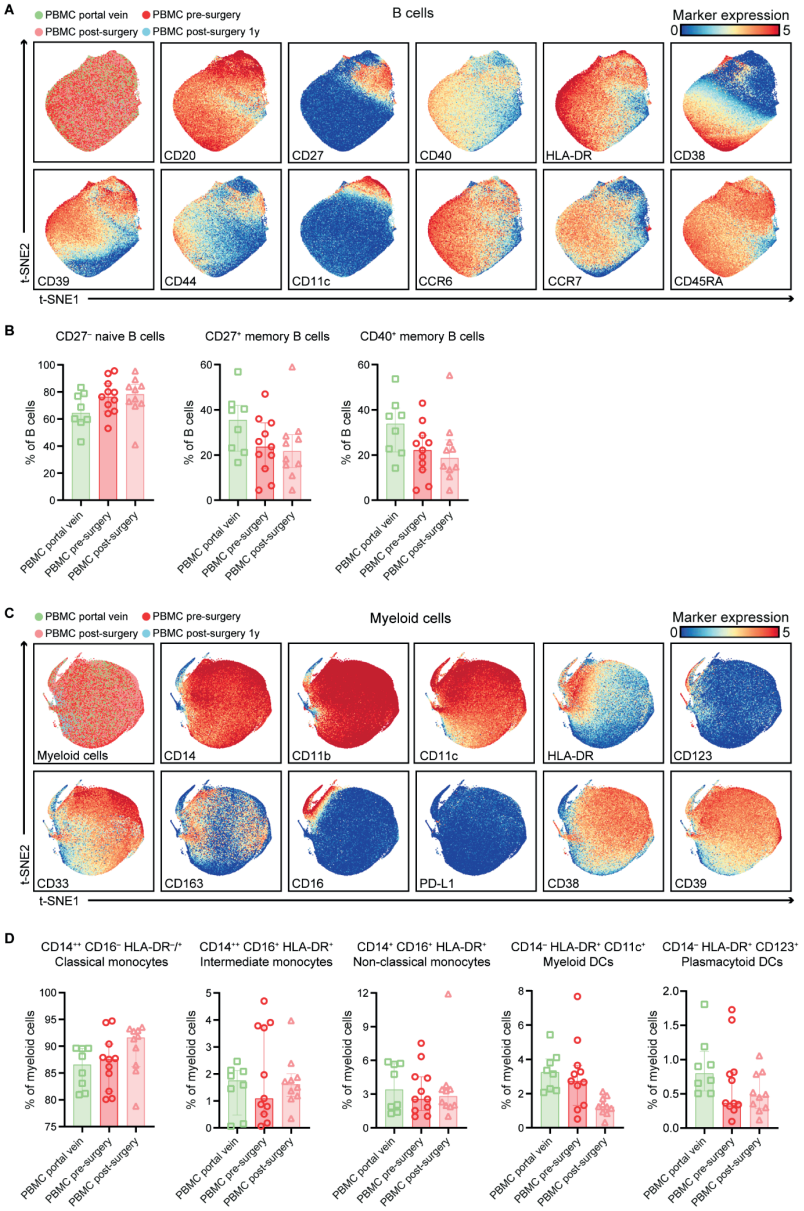


Figure S10. Overview of B cell and myeloid cell phenotypes in portal vein and peripheral blood from patients with PDAC.

A. t-SNE embedding showing 363,140 B cells isolated from portal vein blood (N=8), peripheral blood before surgery (N=11), peripheral blood directly after surgery (N=10) as well as one year after surgery (N=1) from 11 patients with PDAC clustered based on the expression of 41 immune cell markers by single-cell mass cytometry. The cells are colored by tissue type (first plot) and relative expression of indicated B cell markers (remainder). **B.** Frequencies of general B cell populations in the different blood samples as percentage of total B cells. **C.** t-SNE embedding showing 499,980 myeloid cells isolated from the same samples as in (**A**). The cells are colored by tissue type (first plot) and relative expression of indicated myeloid cell markers (remainder). **D.** Frequencies of general myeloid cell populations in the different blood samples as percentage of total myeloid cells. **B,D.** Each dot represents an individual sample. Bars indicate the median with IQR. Data from 11 independent experiments with mass cytometry.

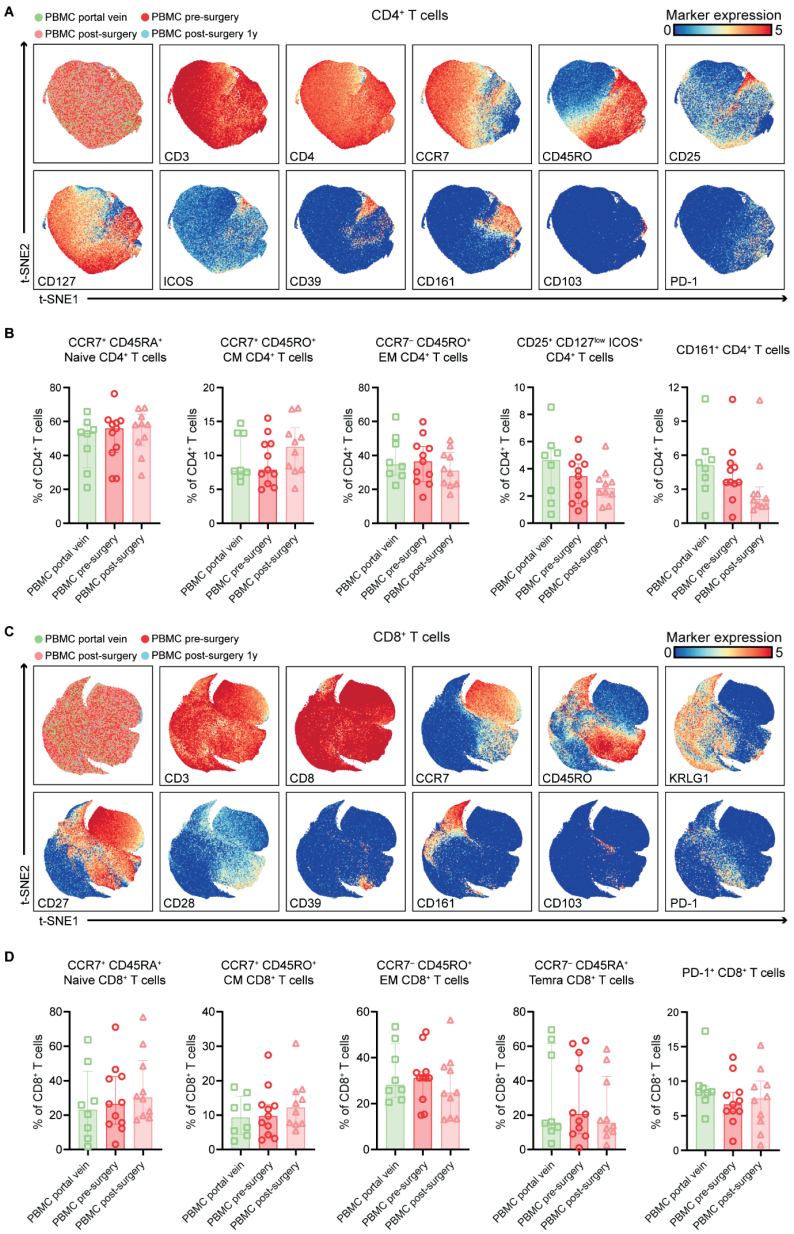


Figure S11. Overview of CD4⁺ and CD8⁺ T cell phenotypes in portal vein and peripheral blood from patients with PDAC.

A. t-SNE embedding showing 499,980 CD4⁺ T cells isolated from portal vein blood (N=8), peripheral blood before surgery (N=11), peripheral blood directly after surgery (N=10) as well as one year after surgery (N=1) from 11 patients with PDAC clustered based on the expression of 41 immune cell markers by single-cell mass cytometry. The cells are colored by tissue type (first plot) and relative expression of indicated CD4⁺ T cell markers (remainder). **B.** Frequencies of selected CD4⁺ T cell clusters in different blood samples as percentage of total CD4⁺ T cells. **C.** t-SNE embedding showing 445,521 CD8⁺ T cells isolated from the same samples as in (**A**). The cells are colored by tissue type (first plot) and relative expression of indicated CD8⁺ T cell markers (remainder). **D.** Frequencies of selected CD8⁺ T cell clusters in different blood samples as percentage of total CD8⁺ T cells. **B,D.** Each dot represents an individual sample. Bars indicate the median with IQR. Data from 11 independent experiments with mass cytometry.

SUPPLEMENTAL TABLES

The supplementary table is available in the online version of this manuscript at http://dx.doi.org/10.1136/jitc-2022-004638 and in the appendix to this thesis at the repository of Leiden University (https://hdl.handle.net/1887/3439882).

Table S1: Antibodies used for mass cytometry experiments.

General discussion and future perspectives

8

GENERAL DISCUSSION AND FUTURE PERSPECTIVES

Cancer immunotherapy has established itself as a novel pillar of cancer care in recent years, despite its long history with the first experiments already performed in the late 19th century.[1] T cells, and their capacity for antigen recognition and cytotoxicity, have become a central focus in anti-tumor immunity. Remarkable progress in fundamental research elucidating the molecular and cellular biology of T cells have resulted in successful cancer immunotherapeutic strategies, such as immune checkpoint blockade therapy and chimeric antigen receptor (CAR) T cell therapy. In contrast to chemo- or radiotherapeutic approaches, these immunotherapies have shown curative potential in advanced cancers.[2] However, immune checkpoint blockade therapy and CAR T cell therapy are not yet applicable to a large proportion of cancer patients. Although T cells have long served as the cellular underpinnings of cancer immunotherapies, T cell-based immunotherapies might not be a solution for all cancer patients. Therefore, characterization of other immune cell subsets with anti-tumor potential in the cancer microenvironment is important and may pave the way toward exploiting previously unappreciated immune cells in an immunotherapeutic setting. Furthermore, it has recently been shown that cancer immunotherapies employed at earlier stages of cancer, in the neo-adjuvant setting, appear to be more successful than at advanced tumor stages.[3] This indicates that different mechanisms are at play in the primary tumor *versus* the metastatic immune microenvironment influencing the effect of immunotherapies. Moving forward to improving outcomes of patients with advanced cancer, it would be crucial to understand how to specifically target immune evasion mechanisms in the metastatic setting. Taken together, to progress cancer immunology research, we need to focus on: i) broadening the scope of anti-tumor immune mediators in primary tumors, before immune evasion can occur, and ii) improving our understanding of the metastatic immune microenvironment and how it can be targeted.

Broadening the scope of anti-tumor immune mediators in primary tumors

The work in this thesis focused on the characterization of the colorectal cancer (CRC) and pancreatic ductal adenocarcinoma (PDAC) immune microenvironments for a comprehensive understanding of anti-cancer immune responses across the innate and adaptive immune compartments. Most cancer immunology research studied the role of cytotoxic T cells in both cancer types, while a comprehensive analysis of both innate and adaptive components of cancer immunity was largely lacking. With such an approach, we demonstrated an important involvement of understudied unconventional (γδ T cells) and innate (innate lymphoid cells (ILCs)) immune effector cells in anti-tumor immunity. Based on the findings of us and others, we foresee emerging roles for other immune effector cells in the coming decades such as Vδ1/Vδ3 T cells and ILC1-like cells in anti-tumor immunity and cancer immunotherapies. In addition, the recent findings that B cells, tertiary lymphoid structures, and the presence of antibodies to cancer antigens are associated with a favorable prognosis in several types of cancer[4-9] highlight further investigations into the potential of naturally-generated antibodies in the context of cancer. Research should prioritize the elucidation of the functions that these understudied immune effector cells perform and how they are

involved in anti-tumor immune responses. Exploitation of their anti-tumor reactivity may provide an alternative immunotherapeutic approach and/or may complement current T cell-mediated immunotherapies to enhance cancer immunotherapeutic success.

In this thesis, we observed that mismatch repair (MMR)-deficient colorectal tumors contained tumor tissue-specific PD-1$^+$ γδ T cells that were infrequent in colorectal healthy tissue, tumor-associated lymph nodes, and peripheral blood of the same patients. Furthermore, PD-1$^+$ γδ T cells were generally not observed in MMR-proficient colorectal tumors. γδ T cells are one of the least understood immune cell types at the interface of innate and adaptive immunity. The precise mechanisms of recognition by and activation of intratumoral γδ T cells remain largely unknown. Our findings of i) elevated frequencies of γδ T cells in MMR-deficient tumors with human leukocyte antigen (HLA) class I defects, ii) *in vitro* activation of intratumoral γδ T cells by CRC cell lines and tumor organoids, iii) killing of these tumor cells by γδ T cells, mainly by Vδ1 and Vδ3 subsets expressing PD-1, iv) increased presence of intratumoral γδ T cells upon immune checkpoint blockade therapy, and v) the localization of γδ T cells next to apoptotic cancer cells in these tumors indicate an active role for these cells in this context. Recent work has revealed an important involvement of butyrophilins binding to the T cell receptor (TCR) in Vδ2 T cell recognition and activation.[10,11] However, different subsets of γδ T cells have remarkably diverse functions. For Vδ1 and Vδ3 cells, it is not yet established whether tumor recognition is established through their γδ TCR, via innate immune receptors such as NKG2D or DNAM-1, or both. Like ILCs/NK cells, γδ T cells express important activating and inhibitory innate immune receptors that potentially are involved in tumor recognition including NKG2D and KIRs (**Figure 1**). For a precise understanding of γδ TCR – ligand interactions, it would be necessary to screen for (novel) γδ TCR ligands involved in the activation of γδ TCRs in combination with single-cell RNA/VDJ-sequencing to dissect antigenic specificities of intratumoral γδ T cells. This might help to elucidate the mechanism behind the specific enrichment and cytotoxicity of Vδ1 and Vδ3 T cells in CRC tissues, while Vδ2 cells are the main γδ T cell population in the circulation. Furthermore, such mechanistic studies might shed light on the observation that γδ T cells were virtually absent in PDAC tissues.

Furthermore, we detected tumor tissue-specific ILC1-like populations in both the CRC and PDAC microenvironment. In CRC, this population was particularly frequent in MMR-deficient tumors that show a relatively high number of somatic mutations and commonly lose HLA class I expression. In these tumors, the ILC1-like cells had a frequent intraepithelial localization and showed hallmarks of activation, cytotoxicity, and proliferation. In line, the ILC1-like population was most abundant in a PDAC tumor with a DNA repair defect, potentially underlying the increased mutational load and resulting inflammatory response observed in this tumor. However, the ILC1-like cells exhibited lower cytotoxicity in the PDAC microenvironment as compared to the CRC microenvironment. Possible reasons for this observation may be the largely immunosuppressive microenvironment in PDAC, the lack of immune evasion through loss of HLA class I-mediated antigen presentation in PDAC,[12] or the lack of expression of NKG2D ligands on PDAC cells. Interestingly, intratumoral NK cells are

scarce in both cancer types,[13,14] while these cells constitute a major immune cell population in the circulation. It would be of interest to study the possible conversion of peripheral NK cells to ILC1-like cells in the cancer microenvironment, as in murine cancer models ILC1-like phenotypes could emerge from NK cell differentiation driven by transforming growth factor-β (TGF-β) signaling.[15] TGF-β is a pleiotropic cytokine that, apart from its immunosuppressive activity, can drive the differentiation of immune cells.[16,17] In support of this, decidual stromal medium with TGF-β can convert peripheral NK cells into decidual NK cells exhibiting an ILC1-like phenotype.[18] Whether ILC1-like cells reflect a truly distinct lineage or acquired this phenotype upon encountering the cancer microenvironment remains elusive and requires further functional studies. In both cancer types, the presence of the ILC1-like population correlated with the presence of $CD8^+$ T cells with tumor-reactive phenotypes ($CD103^+CD39^+$[19-22]), raising questions about whether and how the functions of ILC1-like cells and $CD8^+$ T cells are related during anti-tumor immune responses. It is interesting that the ILC1-like and $CD8^+$ T cell populations shared such highly similar immunophenotypes, localization, and expression of cytotoxic molecules in different cancer types. Like cytotoxic T cells, intraepithelial ILC1-like cells have been shown to be cytotoxic against cancer cells.[23,24] However, they differ greatly in the way they sense cancer cells and in their kinetics of action. ILCs do not depend on the expression of antigen-specific receptors for their activation and respond rapidly, whereas T cell responses require clonal expansion and differentiation. Nevertheless the anti-tumor immune functions of these two diverse immune subsets appear related. From an evolutionary perspective, shared immunological features by subsets of ILCs and T cells[25] may account for robust immunity in the face of continuous cancer immune evasion.[25] The study of ILC1-like cells in CRC is currently being continued in which we examine i) the localization and interacting cells of ILC subsets in colorectal tumors with and without HLA class I expression by imaging mass cytometry, and ii) the cytotoxicity of ILC1-like cells toward CRC cell lines with and without HLA class I expression by *in vitro* functional assays. These follow-up studies may facilitate the exploitation of ILC1-like cells as targets for cancer immunotherapy. Of note, both the ILC1-like and Vδ1/Vδ3 T cell populations showed expression of HLA class II, suggesting that these cells might have antigen-presenting properties, another interesting direction for further research (**Figure 1**).

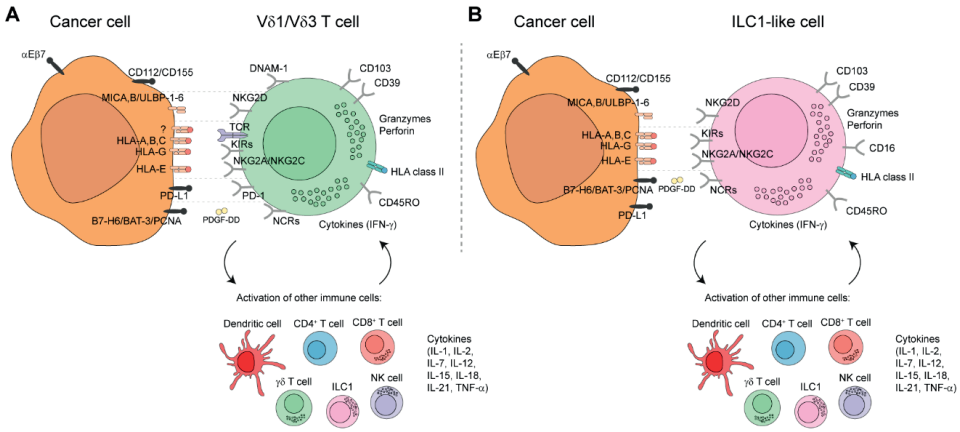


Figure 1. γδ T cells and ILCs are endowed with various recognition mechanisms to sense cancer cells and to initiate anti-tumor immune responses.

Schematic overview of possible receptor – ligand interactions for Vδ1/Vδ3 T cells (**A**) and for ILC1-like cells (**B**) in the context of cancer. NCRs; natural cytotoxicity receptors (NKp30, NKp44, NKp46), PDGF-DD; platelet-derived growth factor-DD.

Awaiting further insights into the function of these previously unappreciated unconventional and innate immune cells, we foresee multiple ways forward for the translation of Vδ1/Vδ3 T cells and ILC1-like cells into cancer immunotherapies. For γδ T cells, immunotherapeutic strategies could include: i) the induction of their activation by PD-1 blockade, ii) the upregulation of NKG2D (ligand) expression to enhance their anti-tumor functions (e.g. via cytokines), iii) the adoptive cell transfer of *ex vivo* expanded Vδ1/Vδ3 T cells, and iv) the transduction of tumor-reactive Vδ1/Vδ3 TCRs on donor αβ T cells. For ILC1-like cells, immunotherapeutic strategies could include i) the induction of their activation by NKG2A blockade, ii) the upregulation of NKG2D (ligand) expression to enhance their anti-tumor functions (e.g. via cytokines), iii) the adoptive cell transfer of *ex vivo* expanded ILC1-like cells, and iv) the engineering of ILC1-like cells to express CARs directed against surface molecules on cancer cells (CAR-ILC1-like cells) providing an "off-the-shelf" allogeneic product. Vδ1/Vδ3 and ILC1-like cancer immunotherapies could have broad implications for MMR-deficient cancers and other malignancies with common HLA class I defects. The first immunotherapies with γδ T cells (mainly Vδ2 subsets) and conventional NK cells are currently being developed for hematological malignancies.[26-28] These clinical trials have not yet reached the stage of conventional T cell-based therapies, but we foresee that we will see such products within the coming years. While it needs to be considered that the Vδ1/Vδ3 T cells and ILC1-like cells might also display alloreactivity, the application of these cells isolated and expanded from a universal donor rather than the autologous setting might be a reasonable approach for future application. Vδ1/Vδ3 T cells and/or ILC1-like cells from a universal donor with the highest cytotoxicity could, in theory, be distributed to many different patients. Harnessing such unconventional and innate immune cells may have the potential to lead to a cheaper, faster, and more universal form of cancer immunotherapy.

Improving our understanding of the metastatic immune microenvironment and how it can be targeted

The work in this thesis focused on the study of primary tumors, while the majority of cancer patients die from metastatic disease. Moving forward to improving outcomes of patients with cancer, it would be important to also study the role of innate and adaptive immune cells in metastatic lesions, how the metastatic immune microenvironment differs from the primary tumor, and how it could be targeted. The structure and composition of the cancer microenvironment vary between the primary tumor and metastasis, and between different metastatic niches in individual patients. During dissemination, cancer cells from primary tumors are exposed to different types of stromal cells, immune cells, platelets, and metabolic stress, and have to adapt to the new tissue microenvironment. In CRC, HLA class I expression can be a determining factor for metastatic behavior. Liver metastases originating from HLA class I-negative/β*2-microglobulin* (*B2M*)-mutated MMR-deficient CRCs are rare and HLA class I-positive.[29-31] This could be related to NK cell-mediated elimination of metastatic cancer cells that lack HLA class I expression in the circulation and/or the high abundance of NK cells (30-50% of intrahepatic lymphocytes) and, to a lower extent, γδ T cells (3-5% of intrahepatic lymphocytes) in the liver[32]. However, metastases could be derived through lymph nodes where NK cells and γδ T cells are virtually absent, or altered metastatic HLA phenotypes could be acquired during the process of dissemination. Analysis of such immune evasion mechanisms of metastases may provide important information required to determine to what type of cancer immunotherapy the metastatic lesions respond. It would be interesting to examine whether we can find similar tumor-resident immune cell populations in metastases as in the primary tumor. In PDAC, the majority of patients already present with advanced disease at the time of diagnosis. Hence, dissecting tumor-immune cell interactions in the metastatic lesions would be necessary to find out whether the approach to treating the primary tumor and metastases needs to vary.

Taken together, cancers are complex ecosystems deregulated at multiple levels. A "one-size fits all" approach will not suffice to achieve effective immunotherapy responses for the majority of cancer patients. We need to determine for individual tumors what immune evasion mechanisms are at play, and target these specifically to get an effective anti-tumor immune response. Multi-omics technologies that connect the immunophenotype, gene expression, and spatial landscape have the potential to accelerate personalized immunotherapy based on each individual's tumor and metastatic microenvironment (**Figure 2**). In parallel, genomic profiling of the primary tumor as well as metastatic lesions as standard diagnostics for patients is needed to gain insight into genomic differences between primary tumor and metastases, and to select (targeted) treatment based on their unique tumor DNA profile. Examining immune phenotypes in primary tumors and metastatic lesions of immunotherapy-responsive patients could provide new avenues on how we could introduce or adapt these tumor-reactive immune cells in non-responsive patients. These analyses could for instance indicate whether cancer immunotherapies with Vδ1/Vδ3 T cells and/or ILC1-like cells may also be effective in the metastatic setting. We can extract enormous amounts of relevant information from multi-omics data, but the challenge is to reconnect findings to their biological relevance.

Therefore, multi-omics technologies must always go hand in hand with laboratory research such as functional studies to properly interpret what we see. This requires an extensive collaboration between multi-disciplinary researchers including bioinformaticians, oncologists, and immunologists. The findings in this thesis will facilitate further mechanistical studies of γδ T cells and ILCs in anti-tumor immunity, and how such cells can be harnessed as novel cancer immunotherapies. We expect that these understudied immune effector cells hold promise for increasing the success rate of immunotherapy across a wide range of cancer types.

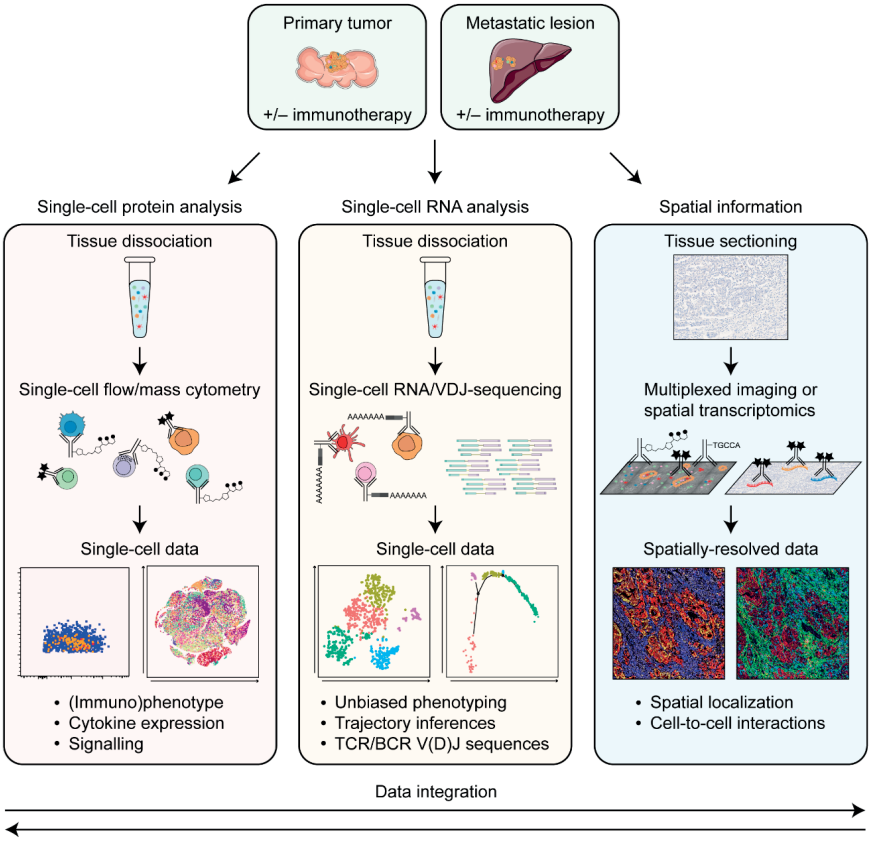


Figure 2. Toward understanding primary tumor and metastatic microenvironments by integrative single-cell technologies.
Schematic overview of the integration of single-cell data of dissociated cells and spatially-resolved data from primary tumors and metastatic lesions, which has the potential to reveal the full cellular landscape of the cancer microenvironment of the primary tumor and metastatic lesions. Adapted from de Vries *et al.* (2020)[33].

REFERENCES

1 Coley, W. B. II. Contribution to the Knowledge of Sarcoma. *Ann Surg* 14, 199-220, doi:10.1097/00000658-189112000-00015 (1891).

2 Melenhorst, J. J. *et al.* Decade-long leukaemia remissions with persistence of CD4(+) CAR T cells. *Nature*, doi:10.1038/s41586-021-04390-6 (2022).

3 Chalabi, M. *et al.* Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat Med* 26, 566-576, doi:10.1038/s41591-020-0805-8 (2020).

4 Helmink, B. A. *et al.* B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* 577, 549-555, doi:10.1038/s41586-019-1922-8 (2020).

5 Petitprez, F. *et al.* B cells are associated with survival and immunotherapy response in sarcoma. *Nature* 577, 556-560, doi:10.1038/s41586-019-1906-8 (2020).

6 Cabrita, R. *et al.* Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* 577, 561-565, doi:10.1038/s41586-019-1914-8 (2020).

7 Meylan, M. *et al.* Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. *Immunity*, doi:https://doi.org/10.1016/j.immuni.2022.02.001 (2022).

8 Ijsselsteijn, M. *et al.* Multidimensional spatial profiling of immune landscapes in colorectal cancer across Consensus Molecular Subtypes. *Manuscript in preparation* (2022).

9 Patil, N. S. *et al.* Intratumoral plasma cells predict outcomes to PD-L1 blockade in non-small cell lung cancer. *Cancer Cell*, doi:10.1016/j.ccell.2022.02.002 (2022).

10 Harly, C. *et al.* Key implication of CD277/butyrophilin-3 (BTN3A) in cellular stress sensing by a major human γδ T-cell subset. *Blood* 120, 2269-2279, doi:10.1182/blood-2012-05-430470 (2012).

11 Rigau, M. *et al.* Butyrophilin 2A1 is essential for phosphoantigen reactivity by γδ T cells. *Science* 367, doi:10.1126/science.aay5516 (2020).

12 Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* 32, 185-203. e113, doi:10.1016/j.ccell.2017.07.007 (2017).

13 Halama, N. *et al.* Natural killer cells are scarce in colorectal carcinoma tissue despite high levels of chemokines and cytokines. *Clin Cancer Res* 17, 678-689, doi:10.1158/1078-0432.Ccr-10-2173 (2011).

14 Lim, S. A. *et al.* Defective Localization With Impaired Tumor Cytotoxicity Contributes to the Immune Escape of NK Cells in Pancreatic Cancer Patients. *Front Immunol* 10, 496, doi:10.3389/fimmu.2019.00496 (2019).

15 Gao, Y. *et al.* Tumor immunoevasion by the conversion of effector NK cells into type 1 innate lymphoid cells. *Nat Immunol* 18, 1004-1015, doi:10.1038/ni.3800 (2017).

16 van den Bulk, J., de Miranda, N. & Ten Dijke, P. Therapeutic targeting of TGF-β in cancer: hacking a master switch of immune suppression. *Clin Sci (Lond)* 135, 35-52, doi:10.1042/cs20201236 (2021).

17 Li, M. O., Wan, Y. Y., Sanjabi, S., Robertson, A. K. & Flavell, R. A. Transforming growth factor-beta regulation of immune responses. *Annu Rev Immunol* 24, 99-146, doi:10.1146/annurev.immunol.24.021605.090737 (2006).

18 Kopcow, H. D. *et al.* Human decidual NK cells form immature activating synapses and are not cytotoxic. *Proc Natl Acad Sci U S A* 102, 15563-15568, doi:10.1073/pnas.0507835102 (2005).

19 Duhen, T. *et al.* Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat Commun* 9, 2724, doi:10.1038/s41467-018-05072-0 (2018).

20 Simoni, Y. *et al.* Bystander CD8(+) T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature* 557, 575-579, doi:10.1038/s41586-018-0130-2 (2018).

21 van den Bulk, J. *et al.* Neoantigen-specific immunity in low mutation burden colorectal cancers of the consensus molecular subtype 4. *Genome Med* 11, 87, doi:10.1186/s13073-019-0697-8 (2019).

22 Simoni, Y. *et al.* Bystander CD4⁺ T cells infiltrate human tumors and are phenotypically distinct. *bioRxiv*, 2020.2007.2015.204172, doi:10.1101/2020.07.15.204172 (2020).

23 Moreno-Nieves, U. Y. *et al.* Landscape of innate lymphoid cells in human head and neck cancer reveals divergent NK cell states in the tumor microenvironment. *Proc Natl Acad Sci U S A* 118, doi:10.1073/pnas.2101169118 (2021).

24 Nixon, B. G. *et al.* Cytotoxic granzyme C-expressing ILC1s contribute to antitumor immunity and neonatal autoimmunity. *Sci Immunol* 7, eabi8642, doi:10.1126/sciimmunol.abi8642 (2022).

25 Vivier, E., van de Pavert, S. A., Cooper, M. D. & Belz, G. T. The evolution of innate lymphoid cells. *Nat Immunol* 17, 790-794, doi:10.1038/ni.3459 (2016).

26 Almeida, A. R. *et al.* Delta One T Cells for Immunotherapy of Chronic Lymphocytic Leukemia: Clinical-Grade Expansion/Differentiation and Preclinical Proof of Concept. *Clin Cancer Res* 22, 5795-5804, doi:10.1158/1078-0432.Ccr-16-0597 (2016).

27 Kabelitz, D., Serrano, R., Kouakanou, L., Peters, C. & Kalyan, S. Cancer immunotherapy with γδ T cells: many paths ahead of us. *Cell Mol Immunol* 17, 925-939, doi:10.1038/s41423-020-0504-x (2020).

28 Shimasaki, N., Jain, A. & Campana, D. NK cells for cancer immunotherapy. *Nat Rev Drug Discov* 19, 200-218, doi:10.1038/s41573-019-0052-1 (2020).

29 Ijsselsteijn, M. E. *et al.* Revisiting immune escape in colorectal cancer in the era of immunotherapy. *Br J Cancer* 120, 815-818, doi:10.1038/s41416-019-0421-x (2019).

30 Kloor, M. *et al.* Beta2-microglobulin mutations in microsatellite unstable colorectal tumors. *Int J Cancer* 121, 454-458, doi:10.1002/ijc.22691 (2007).

31 Menon, A. G. *et al.* Immune system and prognosis in colorectal cancer: a detailed immunohistochemical analysis. *Lab Invest* 84, 493-501, doi:10.1038/labinvest.3700055 (2004).

32 Gao, B., Jeong, W. I. & Tian, Z. Liver: An organ with predominant innate immunity. *Hepatology* 47, 729-736, doi:10.1002/hep.22034 (2008).

33 de Vries, N. L., Mahfouz, A., Koning, F. & de Miranda, N. Unraveling the Complexity of the Cancer Microenvironment With Multidimensional Genomic and Cytometric Technologies. *Front Oncol* 10, 1254, doi:10.3389/fonc.2020.01254 (2020).

Summary in English
Nederlandse samenvatting
Curriculum Vitae & List of Publications
Dankwoord

A

SUMMARY IN ENGLISH

Overcoming challenges in cancer immunology with multidimensional technologies
Cancer immunotherapy treatment currently benefits only a minority of cancer patients, despite the successful application across a broad range of tumors. Cancer is a dynamic disease characterized by extensive heterogeneity, within a primary tumor, between lesions, and across patients (**Chapter 2**). Major challenges in the field of cancer immunology research are the identification of predictive biomarkers to select cancer patients that are likely to respond to specific immunotherapy treatments, the detection of resistance mechanisms to cancer immunotherapies, and the development of novel immunotherapeutic strategies to improve cancer survival. The advent of multidimensional single-cell technologies such as spectral flow cytometry, mass cytometry, single-cell RNA-sequencing, and imaging mass cytometry may play a crucial role to address the former challenges. As cancer is a complex biological system, we highlighted the need of integrating single-cell data derived from multidimensional technologies to i) improve our understanding of the variability in response to immunotherapy treatment and the resistance mechanisms that are at play, and ii) guide the development of alternative immunotherapeutic strategies (**Chapter 2**). As of now, only a handful of studies, including **Chapter 3** of this thesis, have used complementary forms of multidimensional single-cell analyses on the same tissue to study cancer.

Novel mediators of anti-tumor immunity in colorectal cancer
In **Chapter 3** of this thesis, we charted the complexity of colorectal cancer (CRC) immunity by single-cell mass cytometry complemented with single-cell RNA-sequencing. We identified tumor-resident immune cell populations across the innate (ILCs, γδ T cells) and adaptive (cytotoxic and helper T cells) immune compartments. The tumor tissue-specific T cell subsets showed a highly similar activated (HLA-DR$^+$CD38$^+$PD-1$^+$), tissue-resident (CD103$^+$CD69$^+$), memory (CD45RO$^+$) phenotype, and were infrequent in colorectal healthy tissues, tumor-associated lymph nodes, and peripheral blood. We found that two immune cell populations were particularly enriched in mismatch repair (MMR)-deficient CRCs, namely PD-1$^+$ γδ T cells and a previously unappreciated innate lymphoid cells (ILC) population (Lin$^-$CD7$^+$CD127$^-$ CD103$^+$CD45RO$^+$ ILC1-like).

With regard to the first immune cell population, the expression of PD-1 by the γδ T cells in conjunction with their cytotoxic potential suggested an active role for these cells in the anti-tumor immune response and potentially as effector cells of PD-1 immune checkpoint blockade therapy. These research questions were further addressed in **Chapter 4**. We found that intratumoral PD-1$^+$ γδ T cells, isolated from MMR-deficient colon cancers, were composed of Vδ1 and Vδ3 T cells, and showed potent cytotoxicity toward human leukocyte antigen (HLA) class I-negative/*β2-microglobulin* (*B2M*)-defect MMR-deficient CRC cell lines and patient-derived tumor organoids. This response was partly reduced upon blocking the interaction between NKG2D and DNAM-1, and their respective ligands on the cancer cells. As a next step, the relevance of γδ T cells in the clinical setting of patients treated with immune checkpoint blockade was investigated. We addressed the paradoxical observation

that MMR-deficient cancers show exceptional clinical responses to immune checkpoint blockade therapy, even when they lost the expression of HLA class I, an essential component for antigen presentation to $CD8^+$ T cells. This observation contradicts the current view that PD-1 blockade boosts anti-tumor immunity through $CD8^+$ T cells, but suggests the involvement of other immune effector cells than $CD8^+$ T cells. Our work on the in-depth analysis of MMR-deficient colon cancer samples before and after immune checkpoint blockade (**Chapter 4**) was the first to provide evidence that γδ T cells are cytotoxic effector cells of immune checkpoint blockade therapy in cancers with HLA class I defects. Overall, HLA class I deficiency was associated with an increased frequency of activated γδ T cells in treatment-naïve tumors. Furthermore, immune checkpoint blockade profoundly increased the intratumoral frequency of γδ T cells in HLA class I-negative/*B2M*-mutant cancers. In combination with the *in vitro* functional data, our observations illustrate the potential of γδ T cells, in particular Vδ1 and Vδ3 subsets, as novel targets for the development of cancer immunotherapies.

The second immune cell subset prominently enriched in MMR-deficient CRCs, $Lin^-CD7^+CD127^-CD103^+CD45RO^+$ ILC1-like cells, accounted for up to 80% of the innate lymphoid compartment in these tumors. The $CD127^-CD103^+CD45RO^+$ ILC1-like cells positioned, upon hierarchical clustering based on their immune cell profiles, in between conventional NK cells ($CD127^-CD56^+CD45RO^-$) and $CD127^+$ conventional ILCs. This positioning hints toward the possibility that the ILC1-like population can be an intermediate subset in between conventional NK cells and $CD127^+$ conventional ILCs. We attempted to isolate the ILC1-like cells from MMR-deficient CRC tissues for cell culture and functional studies to gain more insight into where the cells come from and whether they could further differentiate, however, this proved difficult. The ILC1-like population had a frequent intraepithelial localization and showed hallmarks of activation, cytotoxicity, and proliferation in MMR-deficient CRC tissues *ex vivo* (**Chapter 3 and Chapter 5**). We performed an unbiased characterization of sorted $CD7^+CD3^-$ ILCs, independently of CD127 expression, by a single-cell RNA-sequencing approach, and showed that high expression of genes encoding killer-cell immunoglobulin-like receptors (KIRs), NKG2A, immunomodulatory molecules, and HLA class II further distinguished the ILC1-like cells from conventional NK cells and conventional ILCs (**Chapter 5**). The ILC1-like cells are an attractive immune cell population to study in light of the expression of KIRs, implied in the recognition of HLA class I-negative cells, and the frequent loss of HLA class I expression observed in MMR-deficient cancers.

The origin of tissue-resident $CD8^+$ and $CD4^+$ T cell populations in colorectal cancer

We identified tumor-specific populations of tissue-resident ($CD103^+$), activated ($CD38^+PD\text{-}1^+CD39^+$) $CD8^+$ and $CD4^+$ T cells in CRC tissues (**Chapter 3**). The non-activated counterparts ($CD103^+CD38^-PD\text{-}1^-$) of the tumor tissue-specific $CD8^+$ and $CD4^+$ T cells were present in tumors and adjacent healthy tissue, but could not be found in pericolic lymph nodes (with the exception of tumor-positive lymph nodes). Co-expression of CD103 and CD39 has been reported to identify tumor-reactive $CD8^+$ T cells in different types of cancer, including CRC. Understanding the clonal relationships between such distinct $CD8^+$ and $CD4^+$ T cell

populations and their tissue of origin might shed light on the dynamics of anti-tumor T cell responses. In **Chapter 6**, we aimed to address the question where the tissue-resident, activated T cells in CRC come from and whether they are clonally expanded. Our TCRβ sequencing data of $CD8^+$ and $CD4^+$ T cell subsets in colorectal tumors and matched non-malignant tissue samples showed that the tissue-resident, activated ($CD103^+CD38^+PD\text{-}1^+$) $CD8^+$ and $CD4^+$ T cells largely originate from pericolic lymph nodes, while their non-activated counterparts ($CD103^+CD38^-PD\text{-}1^-$) showed frequent TCRβ clonal overlap with T cells from adjacent healthy tissue. Within colorectal tumors, highest clonality was observed for $CD103^+CD38^+PD\text{-}1^+$ $CD8^+$ and $CD4^+$ T cell populations, which co-expressed CD39, suggesting that tissue-residency and activation markers as CD103, CD38, PD-1, and CD39 may serve as surrogates for tumor-reactive $CD8^+$ T cells. With regard to $CD4^+$ T cells, phenotypic characterization of markers to identify tumor-reactive $CD4^+$ T cells are largely lacking, and it could be tested whether tumor reactivity is indeed associated with markers as CD103, CD38, PD-1, and CD39 in CRC. The observation that both activated ($CD38^+PD\text{-}1^+$) and non-activated ($CD38^-PD\text{-}1^-$) tissue-resident ($CD103^+$) T cells were infrequent in pericolic lymph node samples (**Chapter 6**) supports the current view that T cells are primed by antigen presenting cells in the lymph nodes, subsequently migrate to the tumor where they recognize their cognate antigen and clonally expand. Upon this TCR activation, the intratumoral T cells start upregulating tissue-residency and activation markers that retain the cells within the tumor.

Novel mediators of anti-tumor immunity in pancreatic ductal adenocarcinoma

In contrast to the immune cell-rich microenvironment of CRCs (**Chapter 3**), the immune microenvironment of pancreatic ductal adenocarcinoma (PDAC) was found to be immunosuppressive and deprived of infiltration by T cells with cytotoxic potential in the majority of pancreatic tumors (**Chapter 7**). The intratumoral $CD8^+$ T cells showed a lack of expression of activation markers as well as immune checkpoint protein PD-1, which might underly the lack of clinical responses to immune checkpoint blockade as of yet. As compared to non-malignant pancreatic tissues, B cells and regulatory T cells were remarkably increased in PDAC tissues. Interestingly, we identified a tumor tissue-specific ILC1-like population ($CD127^-CD103^+CD39^+CD45RO^+$) that resembled the ILC1-like cells found in CRCs, but showed lower expression of cytotoxic molecules. It would be of interest to study how the cytotoxicity of these ILC1-like cells in the PDAC microenvironment could be enhanced, and whether these cells exhibit a cytolytic response to PDAC cells. A unique feature of this study was the sample collection of portal vein blood, in addition to peripheral blood, enabling a characterization of immune cell profiles in portal vein blood in comparison to peripheral blood from PDAC patients. We showed that the immune composition in peripheral blood is extremely diverse, with large differences between patients, which troubles the discovery of disease-related features. Intriguingly, immune cells infiltrating the PDAC microenvironment could, to some extent, be detected in portal vein blood, but not in peripheral blood, suggesting a regional enrichment of immune cells involved in the anti-tumor immune response. For instance, the only patient that harbored $PD\text{-}1^+$ $CD8^+$ T cells in the pancreatic tumor, showed high frequencies of $PD\text{-}1^+$ $CD8^+$ T cells in the matched portal blood sample.

This observation may indicate that portal vein blood could be a novel source of T cells with tumor-reactive phenotypes that could be further exploited for PDAC patients.

NEDERLANDSE SAMENVATTING

De geschiedenis van kanker

Het woord kanker stamt af van het Griekse 'καρκίνος' (karkinos) en het Latijn 'cancer', wat beide krab betekent. Deze naam werd voor het eerst gebruikt door de Griekse arts Hippocrates (ca. 460 v.Chr. – 370 v.Chr.) toen hij zag dat tumoren met hun vele bloedvaten die alle kanten op groeien, lijken op de ledematen van een krab. Wereldwijd staat de krab nog steeds symbool voor de ziekte kanker. Hoewel kanker wordt gezien als een moderne ziekte, kwam het door de eeuwen heen voor. Lang was er geen verklaring of behandeling van kanker voorhanden. In de negentiende eeuw zorgden de opkomst van de microscoop en de herkenning van kankercellen voor veel vooruitgang in de wetenschap over kanker. Door de kankercellen beter te bestuderen, kon men zien dat de ziekte ontstond in een primaire tumor, en zich vervolgens verspreidde naar andere delen in het lichaam. In die tijd was de enige behandeling het operatief verwijderen van de tumor, wat door de slechte hygiëne vaak een risicovolle ingreep was. De opkomst van kankerverwekkende stoffen zoals vervuiling door industrie en tabaksrook, maar ook de stijging in levensverwachting, leidde tot een gestage toename van het aantal kankergevallen. Toen in 1895 door de Duits natuurkundige Röntgen radioactieve röntgenstralen werden ontdekt, werd bestraling met radiotherapie al snel ingezet als effectieve behandeling van kanker mede dankzij baanbrekend werk van de Pools-Frans schei- en natuurkundige Marie Curie (1867 – 1934), de eerste vrouwelijke Nobelprijs winnaar. Tijdens de Tweede Wereldoorlog werd met onderzoek naar de gevolgen van mosterdgas per toeval ontdekt dat mensen die ermee in aanraking kwamen een daling in het aantal witte bloedcellen lieten zien. De Amerikaanse arts Rhoads en farmacologen Goodman en Gilman speelden hier een belangrijke rol. Uit mosterdgas werd het allereerste chemotherapeuticum ontwikkeld voor de behandeling van bloedkanker. Hoewel radiotherapie en chemotherapie zich ondertussen veel verder hebben ontwikkeld, vormen ze tot op heden samen met chirurgie belangrijke pijlers in de behandeling van kanker. Met de opkomst van nieuwe technologieën zijn er steeds meer verklaringen gevonden voor het ontstaan van kanker, en zijn er nieuwe behandelmethoden ontwikkeld die de vooruitzichten voor kankerpatiënten aanzienlijk hebben verbeterd. Desalniettemin blijven kankerpreventie en -bestrijding cruciaal. Momenteel krijgen in Nederland jaarlijks ongeveer 124.000 mensen de diagnose kanker, waarvan 65% vijf jaar na de diagnose nog in leven is. Elk jaar overlijden ongeveer 46.000 mensen aan kanker, waarmee kanker in Nederland de meest voorkomende doodsoorzaak is. Naast de gevolgen voor de gezondheid van de patiënt, heeft kanker ook een grote maatschappelijke en economische impact.

Het ontstaan van kanker

Kanker is een ziekte die wordt gekenmerkt door ongeremde en ongecontroleerde celgroei. Kanker ontwikkelt zich geleidelijk langs evolutionaire principes, waarbij er in cellen in ons lichaam in de loop van de tijd fouten (mutaties) in het DNA ontstaan. Deze mutaties treden meestal sporadisch op, maar kunnen ook erfelijk zijn. Alleen cellen waarin toevalligerwijs een opeenstapeling van specifieke mutaties zijn opgetreden, en die daardoor een selectief groeivoordeel krijgen, kunnen uitgroeien en het functioneren van de gezonde cellen in het

lichaam verstoren. Met name mutaties in genen die van belang zijn voor de celdeling staan centraal in het ontstaan van kanker. Onder deze typen genen vallen proto-oncogenen, tumorsuppressorgenen, genen die de apoptose (geprogrammeerde celdood) regelen, en genen die de DNA-repair regelen. De genetische basis van kanker werd in 1914 erkend door de Duits zoöloog Theodor Boveri. Lang dacht men dat zulke genetische mutaties voldoende waren om kanker te krijgen, maar inmiddels weten we dat andere cellen in de omgeving van kankercellen, zoals bepaalde typen bindweefselcellen en afweercellen, ook wel immuuncellen genoemd, de ontwikkeling en kwaadaardige progressie van kanker kunnen beïnvloeden.

De rol van het immuunsysteem bij kanker

De afgelopen twintig jaar is er veel vooruitgang geboekt in het wetenschappelijk onderzoek naar de rol van het immuunsysteem in het voorkomen en bestrijden van kanker, ondanks dat de eerste experimenten hierover al meer dan honderd jaar geleden werden uitgevoerd. Het belang van het immuunsysteem in het voorkomen en bestrijden van kanker werd met name erkend door observaties van een verhoogde incidentie van kanker bij transplantatie patiënten die immuunsuppressiva gebruiken, patiënten met immuundeficiënties, en muismodellen waarin essentiële componenten van het immuunsysteem werden uitgeschakeld. Het immuunsysteem helpt het lichaam om kanker te bestrijden door ontspoorde eigen lichaamscellen op te ruimen. Om dat te kunnen doen, is het essentieel dat immuuncellen de kankercellen herkennen en geactiveerd worden.

We hebben een breed scala aan immuuncellen met diverse functies, die onderverdeeld kunnen worden in verschillende verdedigingslinies. Het aangeboren immuunsysteem kan een snelle, algemene immuunreactie tegen kankercellen in gang zetten. Belangrijke (cellulaire) spelers in de aangeboren immuniteit zijn lymfoïde cellen zoals 'innate lymphoid cells' (ILCs) en 'natural killer' (NK)-cellen (zie **Figuur 1, Hoofstuk 1**). ILCs en NK-cellen kunnen de kankercellen herkennen, bijvoorbeeld aan de hand van stressmoleculen op het oppervlak van kankercellen, en een immuunreactie in gang zetten. Daarnaast kunnen zij de afwezigheid van speciale oppervlakte moleculen op kankercellen herkennen als lichaamsvreemd. Bijna alle cellen in ons lichaam bevatten moleculen op de celmembraan waarmee delen van eiwitten afkomstig uit de binnenkant van de cel worden gepresenteerd. Cellen presenteren hiermee als het ware een stukje van 'zichzelf' aan hun omgeving in gespecialiseerde moleculen die humaan leukocyten antigenen (HLA) worden genoemd. Myeloïde cellen zijn ook onderdeel van het aangeboren immuunsysteem, en bestaan onder andere uit dendritische cellen, macrofagen en monocyten (zie **Figuur 1, Hoofstuk 1**). Myeloïde cellen kunnen lichaamseigen of lichaamsvreemde eiwitdelen (antigenen) fagocyteren en presenteren aan het adaptieve immuunsysteem (antigeen-presenterende cellen). Ze zetten een immuunreactie in gang en werken intensief samen met het adaptieve immuunsysteem om de kankercellen te bestrijden. In het adaptieve immuunsysteem spelen B-lymfocyten en T-lymfocyten (helper $CD4^+$ T-cellen en cytotoxische $CD8^+$ T-cellen) een belangrijke rol (zie **Figuur 1, Hoofdstuk 1**). In tegenstelling tot aangeboren immuuncellen, werken B- en T-cellen met specifieke antigeenreceptoren, het immunoglobuline en de

T-celreceptor, respectievelijk. Hierdoor zijn ze in staat om vrijwel elk antigeen te herkennen. Het adaptieve immuunsysteem geeft dus een specifieke immuunreactie, maar heeft meer tijd nodig dan het aangeboren immuunsysteem. Daarentegen kunnen cellen van het adaptieve immuunsysteem lang bijdragen aan het immunologische geheugen voor kanker. Tot slot zijn er nog immuuncellen die eigenschappen bezitten van zowel het aangeboren als het adaptieve immuunsysteem, zoals gamma delta (γδ) T-cellen (zie **Figuur 1, Hoofdstuk 1**).
In een continu proces, genaamd immuunsurveillance, zijn onze immuuncellen op zoek naar antigenen die cellen in ons lichaam (zoals kankercellen) op hun oppervlak laten zien. Antigeen-presenterende cellen, bijvoorbeeld dendritische cellen, kunnen verdachte tumor-antigenen op kankercellen herkennen. Ze nemen deze antigenen op, en brengen ze naar de lymfeklieren, waar de antigenen gepresenteerd worden aan naïve T-cellen. T-cellen die specifiek dat verdachte tumor-antigeen herkennen met hun T-celreceptor, worden geactiveerd. Deze geactiveerde T-cellen vermeerderen zich en migreren naar de tumor, waar cytotoxische ($CD8^+$) T-cellen de kankercellen die hetzelfde antigeen laten zien kunnen herkennen en vernietigen. Helper ($CD4^+$) T-cellen geven verschillende soorten cytokinen af die de werking van $CD8^+$ T-cellen stimuleren. De aanmaak van T-cellen tegen tumor-specifieke antigenen met de hulp van dendritische cellen is schematisch weergegeven in **Figuur 2** van **Hoofdstuk 1**. T-celresponsen worden in toom gehouden door de aanwezigheid van immuun checkpoints, die ervoor zorgen dat de immuunreactie zo min mogelijk schade aanricht aan gezonde weefsels in het lichaam. Cytotoxische immuuncellen van zowel het aangeboren als het adaptieve immuunsysteem (bijvoorbeeld NK-cellen en $CD8^+$ T-cellen) kunnen na herkenning de kankercellen vernietigen via twee belangrijke routes: i) het afscheiden van granules die cytotoxische moleculen bevatten (perforine, granzymen), en ii) expressie van moleculen van de 'cell death pathway' (Fas-ligand, TRAIL-ligand).

Immuun ontsnappingsmechanismen van kanker

Het immuunsysteem speelt een cruciale rol bij het tegengaan van kanker, waarbij er een belangrijke rol is weggelegd voor $CD8^+$ T-cellen. Wanneer $CD8^+$ T-cellen de verdachte tumor-antigenen op kankercellen herkennen, kunnen ze de kankercellen vernietigen. Uit onderzoek blijkt dat patiënten die veel van deze cytotoxische T-cellen in hun tumor hebben, een betere overleving hebben dan patiënten bij wie dat niet zo is. Door de opeenstapeling van specifieke mutaties in kankercellen kunnen er echter varianten ontstaan die de herkenning en vernietiging van het immuunsysteem omzeilen. Kankercellen kunnen zich dus aanpassen en zo ontsnappen aan de immuunsurveillance. Een van de meest onderzochte ontsnappingsmechanismen van tumoren is het afschakelen van de expressie van HLA-moleculen op het oppervlak van kankercellen. Hierdoor is er geen antigeen presentatie meer, en kunnen de T-cellen de kankercellen niet meer herkennen en vernietigen. Daarnaast zijn er ook immuuncellen die de kankercellen kunnen helpen overleven door de immuunreactie te onderdrukken. Voorbeelden hiervan zijn een subtype van $CD4^+$ T-cellen die regulatoire T-cellen worden genoemd, alsook subtypen van myeloïde cellen. Ten slotte kunnen kankercellen zich beschermen tegen T-celresponsen door moleculen tot expressie te brengen die kunnen binden aan inhiberende immuun checkpoints op T-cellen, wat een remmend signaal aan de T-cellen geeft (zie **Figuur 3, Hoofdstuk 1**). De interactie tussen

immuuncellen en kankercellen is dus zeer complex.

Immuun checkpoint blokkade: doorbraak in kankertherapie

Onderzoek naar de rol van T-cellen in de immuunrespons tegen kanker heeft geleid tot de ontwikkeling van nieuwe immuuntherapieën die erop gericht zijn om de T-cellen te helpen en bij te sturen. Waar chemotherapie en radiotherapie erop gericht zijn om de kankercellen zelf te vernietigen of hun celdeling te remmen, is het doel van immuuntherapie om het eigen immuunsysteem te activeren of te versterken om de kankercellen op te ruimen. Een grote doorbraak in de behandeling van kanker is immuuntherapie met checkpointremmers. Deze vorm van immuuntherapie richt zich op het blokkeren van het remmende signaal van de immuun checkpoint moleculen PD-1 en/of CTLA-4 op de T-cellen met hun liganden PD-L1/PD-L2 en B7.1/B7.2, respectievelijk, op de kankercellen. Dit wordt gedaan met therapeutische antistoffen (anti-PD-1, anti-PD-L1, anti-CTLA-4) die specifiek de eiwitten PD-1, PD-L1 of CTLA-4 herkennen (zie **Figuur 3, Hoofdstuk 1**). Door deze specifieke antilichamen toe te dienen die de blokkades opheffen, worden de T-cellen geactiveerd en kunnen ze de kankercellen vernietigen. Voorbeelden van checkpointremmers zijn nivolumab, pembrolizumab, atezolizumab, durvalumab en ipilimumab. Deze doorbraak leverde kankeronderzoekers James P. Allison en Tasuku Honjo de Nobelprijs voor Fysiologie of Geneeskunde op in 2018. Immuuntherapie met checkpointremmers kan zeer effectief zijn bij de behandeling van verschillende typen kanker zoals huidkanker, longkanker, en bepaalde soorten darmkanker. Daarnaast zijn er andere soorten immuuntherapie in opkomst, zoals CAR-T-therapie en therapeutische vaccins. Bij CAR-T-therapie worden de eigen T-cellen genetisch aangepast om een speciale receptor (CAR; chimerische antigeenreceptor) tot expressie te brengen, waarmee de T-cellen de kankercellen kunnen herkennen en vernietigen. CAR-T-therapie kan zeer effectief zijn bij verschillende soorten bloedkankers. Therapeutische vaccins hebben als doel om T-cel responsen tegen tumor-specifieke antigenen te stimuleren. Er bestaan verschillende soorten therapeutische vaccins tegen kanker, waaronder vaccins met eiwitten of eiwitdelen, vaccins met DNA- of RNA-moleculen, en vaccins met dendritische cellen. Het is nog grotendeels een experimentele therapie voor de behandeling van kanker.

Uitdagingen op het gebied van immuuntherapie

Immuuntherapie wordt sinds 2011 ingezet als nieuwe behandeling van kanker. Het laat zeer effectieve responsen zien bij verschillende typen kanker. Op dit moment heeft echter maar een kleine groep van de patiënten baat bij immuuntherapie. Bovendien kan er resistentie tegen de immuuntherapie optreden. Het is daarom van belang om meer inzicht te krijgen in de factoren die de respons van een patiënt op immuuntherapie bepalen, en om nieuwe behandelingsmogelijkheden te ontwikkelen. Om dat te kunnen doen, is het cruciaal om te weten welke rol het immuunsysteem precies speelt bij patiënten met kanker, zodat de immuuntherapie afgestemd kan worden op de tumor. T-cellen hebben een centrale rol als targets voor immuuntherapie, en er is veel onderzoek verricht naar de biologie en functie van T-cellen in tumoren. Echter zijn er, naast T-cellen, nog vele andere soorten immuuncellen in tumoren waar nog niet veel onderzoek naar gedaan is. Over bijvoorbeeld de rol van ILCs, NK-cellen en γδ T-cellen in de anti-tumor immuunreactie en in respons op immuuntherapie is

nog weinig bekend. De afgelopen jaren zijn er nieuwe, 'single-cell' technologieën ontwikkeld die het mogelijk maken om het eiwit en genetisch profiel van individuele cellen te bepalen. Voorbeelden van die nieuwe technologieën zijn massa cytometrie en 'single-cell RNA-sequencing'. Met deze nieuwe, 'single-cell' technologieën kunnen we de verschillende typen immuuncellen en hun functie in tumoren in detail in kaart brengen (zie **Figuur 4, Hoofdstuk 1**). Dit geeft ons een uniek platform om de anti-tumor immuunreacties van cellen van zowel het aangeboren als het adaptieve immuunsysteem tegelijk te analyseren. Hiermee kunnen we onderzoeken welke immuuncellen nog meer potentie hebben als targets voor immuuntherapie.

Dit proefschrift

In dit proefschrift hebben we een uitgebreide karakterisatie van het immuuncel-infiltraat in colorectale (dikkedarm) kanker en alvleesklierkanker uitgevoerd. Colorectale kanker is de derde meest voorkomende vorm van kanker wereldwijd. Momenteel komen er per jaar 1,9 miljoen nieuwe gevallen van colorectale kanker bij, met naar schatting 935.000 sterfgevallen. Colorectale kanker is een heterogene ziekte waarbij verschillende subtypen worden onderscheiden aan de hand van hun genetische eigenschappen. Ongeveer 15-20% van de patiënten heeft het subtype microsatelliet instabiele (MSI) colorectale kanker. Dit subtype wordt gekarakteriseerd door vele honderden mutaties in de tumor door een defect in het DNA mismatch-repairsysteem (MMR-deficiënt). Mismatch-repair is een systeem in onze cellen waarmee fouten in het DNA hersteld kunnen worden. Bij een defect mismatch-repair systeem zullen er grote hoeveelheden mutaties ontstaan. Door deze grote hoeveelheid mutaties worden er waarschijnlijk veel antigenen aan het immuunsysteem gepresenteerd, wat leidt tot veel tumor-infiltrerende immuuncellen. Dit tumortype wordt ook wel 'immunogeen' genoemd. Daarentegen heeft het merendeel van de patiënten met colorectale kanker een microsatelliet stabiele (MSS) tumor met een normaal werkend DNA mismatch-repairsysteem (MMR-proficiënt). In vergelijking met MMR-deficiënte tumoren, vertoont dit subtype over het algemeen een lager aantal mutaties en een minder grote infiltratie van immuuncellen in de tumor. Het type colorectale kanker is gerelateerd aan de klinische prognose van de patiënten en de respons op immuuntherapie met checkpointremmers. MMR-deficiënte tumoren hebben over het algemeen een betere prognose dan MMR-proficiënte tumoren. Het is bekend dat immuuntherapie met checkpointremmers weinig of niet aanslaat bij MMR-proficiënte tumoren, maar wel effectief kan zijn bij een deel van de MMR-deficiënte tumoren. Een veel voorkomend ontsnappingsmechanisme van MMR-deficiënte colorectale kanker is de afschakeling van HLA-expressie. Hierdoor zouden MMR-deficiënte tumoren, in theorie, ongevoelig moeten zijn voor immuuntherapie met checkpointremmers. Deze vorm van immuuntherapie grijpt namelijk in op T-cellen in tumoren, die in dit geval de HLA-negatieve kankercellen niet (zouden moeten) kunnen herkennen. Verrassend genoeg reageert de meerderheid van HLA-negatieve, MMR-deficiënte colorectale kanker goed op immuuntherapie met checkpointremmers. Dit suggereert dat andere soorten immuuncellen bijdragen aan de respons op de immuuntherapie in deze tumoren. Hier zou een belangrijke rol weggelegd kunnen zijn voor ILCs, NK-cellen en γδ T-cellen, die kankercellen zonder HLA-expressie kunnen herkennen en elimineren.

Alvleesklierkanker is een van de meest dodelijke vormen van kanker. Het wordt vaak pas in een laat stadium ontdekt, waarbij slechts ongeveer 20% van de patiënten in aanmerking komt voor operatie. Vanwege de slechte prognose heeft alvleesklierkanker ongeveer evenveel gevallen (496.000) als sterfgevallen (466.000) per jaar wereldwijd. Alvleesklierkanker ontstaat door de stapsgewijze opeenstapeling van genetische veranderingen via niet-invasieve voorloperafwijkingen. Er zijn verschillende typen alvleesklierkanker. De meest voorkomende vorm van alvleesklierkanker ontstaat uit cellen van de alvleeskliergang. Dit type kanker heet ductaal adenocarcinoom, en hebben we in dit proefschrift bestudeerd. Alvleesklierkanker is niet-immunogeen en wordt gekenmerkt door een laag aantal mutaties, en weinig en vooral immuun-onderdrukkende immuuncellen. Defecten in DNA-repairsystemen zoals MMR-deficiëntie komen zelden voor in alvleesklierkanker (ongeveer in 1% van de gevallen). Tot op heden reageert alvleesklierkanker niet op behandeling met immuun checkpointremmers. Momenteel worden nieuwe vormen van immuuntherapie en combinaties van therapieën onderzocht om te proberen om de immunogeniteit van alvleesklierkanker te versterken, en daarmee de respons op immuuntherapie te verbeteren.

In **Hoofdstuk 2** beschrijven we hoe onze kennis over de rol van het immuunsysteem in kanker vergroot is door de opkomst van nieuwe, 'single-cell' technologieën zoals massa cytometrie en 'single-cell RNA-sequencing'. Deze technieken hebben wij toegepast in de experimenten die in dit proefschrift zijn beschreven om zo een beter beeld te krijgen van de immuuncellen die zich in de omgeving van de tumor bevinden.

In **Hoofdstuk 3** hebben we een gedetailleerde karakterisatie van het immuuncel-infiltraat in colorectale kanker, gezond colorectaal weefsel, lymfeklieren en bloed samples van patiënten uitgevoerd met massa cytometrie. We vonden dat cellen van zowel het aangeboren als het adaptieve immuunsysteem een belangrijke rol spelen bij de anti-tumor immuunreactie. Daarnaast werden nieuwe immuuncel-subsets gevonden die verhoogd aanwezig waren in MMR-deficiënte tumoren. Deze nieuwe subsets bestaan uit PD-1$^+$ γδ T-cellen en CD103$^+$ ILC1-achtige cellen, en laten cytotoxische activiteit zien in de tumoren. Ten slotte vonden we weefsel-residente, geactiveerde CD8$^+$ en CD4$^+$ T-cel populaties die specifiek aanwezig waren in tumoren en niet in gezonde weefsels van dezelfde patiënten. Deze observaties vormden de basis voor nieuwe onderzoeksvragen die we in **Hoofdstuk 4**, **Hoofdstuk 5**, en **Hoofdstuk 6** bestuderen.

In **Hoofdstuk 4** hebben we verder onderzocht hoe de PD-1$^+$ γδ T-cellen bijdragen aan de anti-tumor immuunrespons en aan het effect van immuuntherapie met checkpointremmers in MMR-deficiënte colorectale kanker. Tumoren voor en na behandeling met checkpointremmers werden geanalyseerd met een combinatie van genexpressie, massa cytometrie en functionele experimenten. We vonden dat in tumoren waarin de HLA-expressie afgeschakeld is, waardoor ze ongevoelig (zouden moeten) zijn voor CD8$^+$ T-cellen, γδ T-cellen een belangrijke rol spelen in de vernietiging van de kankercellen.

In **Hoofdstuk 5** analyseerden we het volledige spectrum van ILC subsets en hun functionele verschillen in MMR-deficiënte colorectale kanker door middel van 'single-cell RNA-sequencing' in combinatie met analyse van eiwitexpressies en transcriptiefactoren. We laten zien dat de meerderheid van de ILC-populatie in colorectale tumoren bestaat uit cytotoxische CD103$^+$ ILC1-achtige cellen, die een actieve rol lijken te spelen in de anti-tumor immuunrespons.

In **Hoofdstuk 6** hebben we de DNA-sequenties van de T-celreceptor van weefsel-residente, geactiveerde CD8$^+$ en CD4$^+$ T-cellen uit colorectale tumoren bepaald en vergeleken met die van T-cellen uit gezond colorectaal weefsel, lymfeklieren en bloed samples van dezelfde patiënten. Daarmee kregen we inzicht in waar de weefsel-residente, geactiveerde T-cellen in colorectale kanker vandaan komen, en of ze klonale expansie laten zien, een kernmerk van antigeen-specifieke T-celresponsen.

In **Hoofdstuk 7** onderzochten we immuuncel-populaties in alvleesklierkanker, gezond alvleesklierweefsel, lymfeklieren, milt en bloed samples van patiënten met massa cytometrie. We beschrijven een immuun-onderdrukkende omgeving rondom de tumor met weinig cytotoxische T-cellen en veel B cellen en regulatoire T-cellen ten opzichte van gezond alvleesklierweefsel. Daarnaast vonden we een CD103$^+$ ILC1-achtige populatie in alvleeskliertumoren die lijkt op de populatie die we gevonden hebben in colorectale tumoren (**Hoofdstuk 3** en **Hoofdstuk 5**). Een unieke invalshoek van deze studie was de inclusie van immuuncellen geïsoleerd uit de poortader, waar we een verhoogde aanwezigheid van tumor-geassocieerde immuuncellen zagen.

In **Hoofdstuk 8** bespreken we de belangrijkste bevindingen van dit proefschrift in een breder perspectief en geven we implicaties voor verder onderzoek en de klinische praktijk.

Conclusie en toekomstperspectief

In dit proefschrift hebben we de complexiteit van het immuunsysteem in colorectale kanker en alvleesklierkanker blootgelegd met massa cytometrie en 'single-cell RNA-sequencing'. Terwijl de meeste studies zich focussen op de rol van conventionele T-cellen in de ontwikkeling en progressie van kanker, hebben wij een gedetailleerde karakterisatie van immuuncellen van zowel het aangeboren als het adaptieve immuunsysteem uitgevoerd in tumoren en gezonde weefsels van patiënten. We vonden dat naast conventionele T-cellen, γδ T-cellen (in colorectale kanker) en ILCs (in colorectale kanker en alvleesklierkanker) een belangrijke rol spelen bij de anti-tumor immuunreactie. In colorectale kanker waren deze immuuncel-subsets verhoogd aanwezig in MMR-deficiënte tumoren die vaak afschakeling van HLA-expressie laten zien, en daardoor ongevoelig (zouden moeten) zijn voor herkenning en vernietiging door CD8$^+$ T cellen. Aan de hand van het eiwit en genetisch profiel van de immuuncellen, en met behulp van functionele experimenten, zagen we dat ze een actieve rol spelen in de herkenning en vernietiging van de kankercellen (zie **Figuur 1, Hoofdstuk 8**). Op basis van onze eigen bevindingen, maar ook de bevindingen van anderen, voorzien we een belangrijke rol voor γδ T-cellen en ILCs als effector cellen voor immuuntherapie. Verdere

studies zullen zich richten op de werking en potentie van deze subsets als immuuntherapie bij kanker. Deze resultaten zijn belangrijk, omdat momenteel maar een kleine groep van de patiënten met kanker baat heeft bij immuuntherapie. Het benutten van de anti-tumor reactiviteit van γδ T-cellen en ILCs zou het succes van immuuntherapie bij kanker kunnen verbeteren.

In de toekomst is het belangrijk om niet alleen de rol van het immuunsysteem in primaire tumoren te bestuderen, zoals beschreven in dit proefschrift, maar ook de rol van het immuunsysteem bij uitzaaiingen van kanker. De meerderheid van de patiënten met kanker overlijdt aan de gevolgen van uitzaaiingen. Om de vooruitzichten van patiënten met gevorderde kanker te verbeteren, is het belangrijk om de rol van immuuncellen van het aangeboren en adaptieve immuunsysteem bij uitzaaiingen te onderzoeken, en om in kaart te brengen hoe de omgeving van uitzaaiingen verschilt van die van de primaire tumor (zie **Figuur 2, Hoofdstuk 8**). Tijdens de verspreiding van kankercellen via het bloed of lymfe, worden de kankercellen blootgesteld aan verschillende soorten bindweefselcellen en immuuncellen. Vervolgens moeten de kankercellen zich aanpassen aan de nieuwe omgeving op de plek van de uitzaaiing. De aanwezigheid van andere soorten immuuncellen in de omgeving van de uitzaaiing en wellicht andere immuun ontsnappingsmechanismen ten opzichte van de primaire tumor kan van invloed zijn op het effect van immuuntherapie. Het is daarom belangrijk om meer inzicht te krijgen in de immuuncellen rondom uitzaaiingen, en hoe deze kunnen worden aangepakt met immuuntherapie.

De resultaten van dit proefschrift zijn in verschillende wetenschappelijke tijdschriften gepubliceerd. Dit proefschrift is tot stand gekomen dankzij samenwerkingen tussen verschillende onderzoeksgroepen onder leiding van Frits Koning (Immunologie, LUMC), Noel de Miranda (Pathologie, LUMC), Boudewijn Lelieveldt (Radiologie, LUMC en TU Delft), Ahmed Mahfouz (Humane Genetica, LUMC en TU Delft) en Emile Voest (Moleculaire Oncologie en Immunologie, NKI).

CURRICULUM VITAE AND LIST OF PUBLICATIONS

Curriculum vitae

Natasja Lieuwke de Vries was born on October 19th, 1992 in Leiden. She grew up in the rural area of Vlist with her parents (Bote and Christine), two younger sisters (Carmen and Simone), and a younger brother (Bas). In 2010, she graduated from the Coornhert Gynmasium in Gouda, during which she spent 1.5 years (2007-2008) at the Vespucci College in Julianadorp, Curaçao. She started her studies Biomedical Sciences at the Leiden University Medical Center (LUMC) in Leiden, and obtained her Bachelor's degree in 2013 with a minor in Science Based Business. She then attended the Master Biomedical Sciences at the LUMC, and followed the specialization Cancer Pathogenesis and Therapy. As part of the Master, she joined the laboratory of Prof. Marianne Hokland at the department of Biomedicine of Aarhus University in Denmark, for a 10-month internship in the field of cancer immunology. In collaboration with the department of Surgery at the LUMC under supervision of Dr. Peter Kuppen, she set up a project to characterize the phenotype of T cells, NK cells, and NKT cells in peripheral blood of patients with colorectal cancer using multiparameter flow cytometry. She obtained her Master's degree *cum laude* in 2016. Next, she initiated her PhD research in the laboratories of Prof. Frits Koning and Dr. Noel de Miranda at the department of Immunology and the department of Pathology of the LUMC. As described in this thesis, her research focused on dissecting immune cell populations across cancer and healthy tissues from patients with colorectal cancer and patients with pancreatic cancer by single-cell mass cytometry, single-cell RNA-sequencing, functional assays, and spatial analyses. During her PhD studies, she visited 11 conferences, of which she attended four with an oral presentation and four with a poster presentation. She was awarded the best poster prize at a European Association for Cancer Research (EACR) conference in Barcelona, Spain. Further, she was an invited speaker at the Single-Cell Analysis course of the Medical Genetics Center (MGC) at the Erasmus MC in Rotterdam. She supervised the internships of a student Biomedical Sciences and a student Life, Science and Technology. In 2021, she was a reviewer for the scientific journals *Frontiers in Immunology* and *Journal for ImmunoTherapy of Cancer*. Since May 2022, Natasja is working as a Scientist at Genmab in Utrecht, an international biotech company specialized in the creation and development of antibody therapeutics for the treatment of cancer and autoimmune diseases. She is living in Amsterdam with her partner (Erik).

List of publications

Published

Brouwer, T.P.*, **de Vries, N.L.***, Abdelaal, T., Ruano, D., Farina Sarasqueta, A., Lelieveldt, B.P.F., Morreau, H., Bonsing, B.A., Vahrmeijer, A.L., Koning, F., de Miranda, N.F.C.C., Local and systemic immune profiles of human pancreatic ductal adenocarcinoma revealed by single-cell mass cytometry, **J Immunother Cancer.** 10: e004638 (2022)

Krijgsman, D., **de Vries, N.L.**, Andersen, M.N., Skovbo, A., Tollenaar, R.A.E.M., Bastiaannet, E., Kuppen, P.J.K.†, Hokland, M.†, The effects of tumor resection and adjuvant therapy on the peripheral blood immune cell profile in patients with colon carcinoma, **Cancer Immunol Immunother.** 69: 2009-2020 (2020)

Krijgsman, D., **de Vries, N.L.**, Andersen, M.N., Skovbo, A., Tollenaar, R.A.E.M., Møller, H.J., Hokland, M.†, Kuppen, P.J.K.†, CD163 as a Biomarker in Colorectal Cancer: The expression on circulating monocytes and tumor-associated macrophages, and the soluble form in the blood, **Int J Mol Sci.** 21: 5925 (2020)

de Vries, N.L., Mahfouz, A., Koning, F., de Miranda, N.F.C.C., Unraveling the complexity of the cancer microenvironment with multidimensional genomic and cytometric technologies, **Front Oncol.** 10: 1254 (2020)

de Vries, N.L., van Unen, V.#, Ijsselsteijn, M.E.#, Abdelaal, T.#, van der Breggen, R., Farina Sarasqueta, A., Mahfouz, A., Peeters, K.C.M.J., Höllt, T., Lelieveldt, B.P.F., Koning, F.†, de Miranda, N.F.C.C.†, High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of antitumour immunity, **Gut.** 69: 691-703 (2020)

van den Bulk, J., Verdegaal, E.M.E., Ruano, D., Ijsselsteijn, M.E., Visser, M., van der Breggen, R., Duhen, T., van der Ploeg, M., **de Vries, N.L.**, Oosting, J., Peeters, K.C.M.J., Weinberg, A.D., Farina-Sarasqueta, A., van der Burg, S.H., de Miranda, N.F.C.C., Neoantigen-specific immunity in low mutation burden colorectal cancers of the consensus molecular subtype 4, **Genome Med.** 11: 87 (2019)

Krijgsman, D.*, **de Vries, N.L.***, Skovbo, A., Andersen, M.N., Swets, M., Bastiaannet, E., Vahrmeijer, A.L., van de Velde, C.J.H., Heemskerk, M.H.M., Hokland, M.†, Kuppen, P.J.K.†, Characterization of circulating T-, NK-, and NKT cell subsets in patients with colorectal cancer: the peripheral blood immune cell profile, **Cancer Immunol Immunother.** 68: 1011-1024 (2019)

de Vries, N.L., Swets, M., Vahrmeijer, A.L., Hokland, M.†, Kuppen, P.J.†, The immunogenicity of colorectal cancer in relation to tumor development and treatment, **Int J Mol Sci.** 17: 1030 (2016)

In revision

de Vries, N.L.*, van de Haar, J.*, Veninga, V.*, Chalabi, M.*, Ijsselsteijn, M.E., van der

Ploeg, M., van den Bulk, J., Ruano, D., van den Berg, J.G., Haanen, J.B., Zeverijn, L.J., Geurts, B.S., de Wit, G.F., Battaglia, T.W., Gelderblom, H., Verheul, H.M.W., Schumacher, T.N., Wessels, L.F.A., Koning, F.†, de Miranda, N.F.C.C.†, Voest, E.E.†, γδ T cells are effectors of immune checkpoint blockade in mismatch repair-deficient colon cancers with antigen presentation defects, **Nature**

In preparation

de Vries, N.L., Zheng, L., Roelands, J., van der Ploeg, M., Ijsselsteijn, M.E., van der Breggen, R., Krop, J., Eikmans, M., Ruano, D., Koning, F.†, de Miranda, N.F.C.C.†, The ILC compartment in mismatch repair-deficient colorectal cancers is dominated by CD127-negative ILC1-like cells

de Vries, N.L., Bras Gomes Nunes, J., van der Breggen, R., Koning, F., de Miranda, N.F.C.C., T cells with tumor-specific phenotypes largely originate from pericolic lymph nodes in colorectal cancer

* Co-first author, # Co-second author, † Co-last author

ACKNOWLEDGEMENTS / DANKWOORD

The work in this thesis would not have been possible without the people that have helped, guided, and supported me through this journey. Here, I would like to thank a few people that have been particularly important:

Mijn promotor, Frits Koning: van het begin af aan gaf je me de vrijheid en het vertrouwen om zelf uit te vinden hoe de immunologie in elkaar steekt. Tegelijkertijd overzag jij altijd het grote plaatje in de enorme hoeveelheid complexe data, en stuurde je bij waar nodig. Ik ben je erg dankbaar voor de goede adviezen, de eerlijke visies en de mooie resultaten die we samen bereikt hebben.

My co-promotor, Noel de Miranda: you, and the rest of the group, immediately made me feel welcome and gave me the opportunity to join the discussions in immunogenomics and pathology. Thank you for guiding me, always being critical, and being a great inspiration during this early stage of my career. It was a pleasure to discuss science, and life in general, with you and Frits, and I hope we will continue working together in the future.

The King group: Yvonne, dank dat je altijd klaarstond voor vragen en hulp. Vincent and Li Na, thank you for your advice with CyTOF experiments and the great time at conferences. Zhou Lin, thank you for your help with the PDAC project. Tessa en Laura, dank dat jullie er altijd waren en zijn om lief en leed te delen. Nannan, Sanne, Jia Li, Fernanda, Juliette, Munisha, Mette, Emma, Qinyue, and Ciska, thank you for your scientific input and the fun times during walks outside, cheese making, canoeing, and dinners.

The Immunogenomics group: Marieke, altijd sta je klaar om te helpen met projecten van anderen. Dank voor je enorme inzet en bijdrage aan dit proefschrift. Jitske, dank voor je luisterend oor en de fijne theepauzes. Manon en Ruud, ik ben jullie erg dankbaar voor alle hulp met experimenten. Dina, Joana, Jessica, thank you for your help with experiments and data analyses. Thomas, dank voor de goede samenwerking in het PDAC project. Li Zheng, I am grateful that you will be continuing the projects and I wish you the best of luck. Arantza, Eva, Katja, Ziena, Debora, Ricki, and Siddh, thank you for your scientific input and the great time during group activities and dinners.

Mijn studenten Roelof en Dennis: dank voor jullie inzet en enthousiasme tijdens de stage projecten in het lab.

Guillaume, my paranymph: thank you for traveling the world together for conferences; from skiing in Colorado to running in Stockholm. We met many inspiring researchers and science from all over the world and at the same time enjoyed our shared passion for mountains, nature, and sports.

Boudewijn Lelieveldt and Thomas Höllt: thank you for your support with Cytosplore. Ahmed Mahfouz: I enjoyed writing the review together, and I am grateful for your help with analyses. Tamim: thank you for all the bioinformatics analyses and support, and the good time at conferences.

The CyTOF team, FCF team, and Wim Corver: thank you for your experimental advice and support. Jan Oosting: bedankt dat ik altijd bij je binnen kon lopen voor vragen over statistiek. Colleagues from the NKI, Emile Voest, Joris, Vivien, and Myriam: thank you for the highly productive collaboration on the gamma delta T cell project. I hope we can soon celebrate a publication in Nature.

Lieve Rekemcrew, Jip, Laura, Fleur, Stephanie, Nienke en Kate: het is een voorrecht om met zulke fijne, talentvolle en ambitieuze vriendinnen lief en leed te kunnen delen. Dank dat jullie altijd voor me klaarstaan. Lieve Dorien: dank voor je steun, betrokkenheid en de bijzondere gesprekken. Lieve Annelotte, Maaike en Melissa: dank voor jullie altijd belangstellende vragen over mijn onderzoek en de fijne borrelavonden met elkaar. Lieve Naomi, Annarein en Josine: oud-huisgenoten uit Den Haag, ik heb genoten van de leuke tijd op de WBR 205 en hoop dat er nog vele diners volgen. Andere jaarclubgenoten van Olle, oud-huisgenoten van de Vreewijk 16, en oud-teamgenoten van HGC Dames 3: ik kijk met veel plezier terug op de mooie tijd samen in Leiden en op het hockeyveld in Den Haag. Lieve Jette, Erik Stigter, Anne, Erik Schouwstra, Frederieke en Luuk: dank voor jullie luisterend oor en de afleiding. Ik kijk uit naar een mooie tijd in Utrecht (omgeving) met elkaar.

Lieve familie, schoonfamilie, Hanneke, Heidi en Désirée: ik ben jullie erg dankbaar voor de vele rondjes vogelplas – met een lach en een traan, de steun en afleiding, de fietstochten en vlaaien in Limburg, de heerlijke (fiets)vakanties, en natuurlijk de belangstelling in mijn onderzoek.

Lieve broertje en zusjes, Carmen, Bas en Simone: wat hebben wij een sterke en bijzondere band samen. Carmen, we hebben een gedeelde passie voor hoe het menselijk lichaam werkt. Jij aan de kant van de patiënten, en ik aan de kant van de cellen. Ik ben ongelooflijk trots op jou en vind het heel bijzonder dat je vandaag naast me staat als paranimf. Saksham, dank dat je oneindig spelletjes met ons wilt spelen. Bas, ik ben enorm trots op hoe je je door de pieken en dalen van je studie en stage hebt heen geslagen. Wat ben jij een doorzetter. Sabrina, dank dat je altijd voor ons klaarstaat in goede en minder goede tijden. Simone, ik vond het erg gezellig om te kunnen helpen met jouw onderzoek en scriptie in de lunchpauzes in het LUMC. Ook jij bent een enorme doorzetter en ik ben zo trots op jou.

Lieve papa: je wijze raad en adviezen zijn mij erg waardevol. Dank voor het vertrouwen en de vrijheid die je me hebt meegegeven. Mijn interesse in de biologie komt toch een beetje van jou vandaan. Hopelijk worden we ooit net zo'n goed (zeil)team op de Losbol. Lieve Kirsten, ik vond het erg leuk om de groepsuitjes op de boerderij te mogen organiseren. Dank voor je inzet, enthousiasme en gastvrijheid.

Lieve mama: dank voor je onvoorwaardelijke liefde, zorg en belangstelling. Jouw interesse in en aandacht voor de mensen om je heen zijn een groot voorbeeld voor mij. Dankzij jou heb ik geleerd om te relativeren en te genieten van het leven en de natuur. In gedachten zit je hier vooraan vandaag. Zoals je zelf zei: "Onderzoek alles en behoud het goede! Carpe diem!" Lieve Erik de Kooter, we hebben elkaar in korte tijd goed leren kennen. Dank dat je altijd voor ons klaarstaat.

Tot slot, lieve Erik: het leven is zoveel leuker met jou. Dank voor jouw oneindige liefde, steun en positiviteit, en de tolerantie van mijn drukte en stress. Je staat altijd klaar om mij te helpen, zo ook met analyses in dit proefschrift en het ontwerp van de cover. Ik kijk met veel plezier uit naar alle avonturen en reizen die we samen nog mogen beleven. Wie weet steken we ooit samen de oceaan over op onze zeilboot!