Role of bacteria-reactive T lymphocytes in immune responses against human colorectal cancer

A doctoral dissertation presented by
Julija Djordjevic

Under the supervision of
Prof. Giandomenica Iezzi

Submitted to the
Faculty of Biomedical Sciences
Università della Svizzera italiana

For the degree of
Ph.D. in Biomedical Sciences

April 2024

Note to the Reader

During my PhD studies, I conducted an extensive investigation into the role of the gut microbiota-adaptive immunity axis in human colorectal cancer (CRC). The primary objective was to deepen our understanding of the interplay between the intricate gut bacteria community and immune responses developing in the host, for the future development of novel patient treatment strategies.

This doctoral dissertation comprises several key sections. The **introduction** emphasizes the clinical relevance of the "immune contexture" in CRC and offers a comprehensive overview of the determinants in the CRC microenvironment, including gut microbiota and their influence on immune cells. The **hypothesis** and **aims** of this study are outlined, followed by a detailed description of the **materials** and **methods** used in our research. The **results** section is presented in a chronological format, reporting most of the results obtained throughout my PhD studies, as well as preliminary results of our most recent experiments. In the **discussion**, I provide an in-depth analysis of the major findings and their broader implications. These data have been included in a manuscript, which is currently under preparation.

In addition to this primary study, during my PhD, I was also involved in additional research projects. One study focusing on the potential benefits of directly triggering Toll-like receptors in colorectal cancer-associated stromal cells led to the submission of a manuscript, which is currently under evaluation upon revision, and is included in the **appendix**. Furthermore, I actively contributed to another study focusing on the identification of combined immune-microbial signatures predicting response to neoadjuvant chemoradiotherapy in patients with locally advanced rectal cancer. The results of this study are also currently submitted for publication. Finally, I was involved in a study evaluating the impact of gut microbiota composition on immune responses against SARS-CoV-2. The results of this study have been included in a manuscript, which is currently under preparation.

Table of Contents

Summary

Colorectal cancer (CRC) is the third most diagnosed cancer globally and the second leading cause of cancer-related mortality. Many studies have convincingly demonstrated that CRC infiltration by T cells is unequivocally associated with prolonged patient survival, although this spontaneously occurs in a minority of cases, and underlying mechanisms remain unclear. Paradoxically, immunotherapies based on immunological checkpoint inhibitors (ICIs) have proven beneficial only in a small subset of CRCs (up to 15%), characterized by microsatellite instability (so-called MSI phenotype), high tumor mutational burden (TMB) and high immunogenicity and increased T cell infiltration. Instead, most CRC cases exhibit a microsatellite stable (MSS) phenotype, low mutational burden, and poor response to ICIs. In these CRCs, it has been recognized that only a minor fraction of tumor-infiltrating T cells (TILs) is specific for tumor-derived neoantigens, while a large abundance of TILs is identified as “bystander” cells, with antigenic specificities remaining unclear.

CRC originates within an environment teeming with trillions of gut bacteria. During the oncogenic process, changes in gut barrier integrity result in increased permeability of the gut mucosa, thus allowing gut bacteria to actively translocate into the submucosa and directly engage with tumor cells and resident immune cells.

Previous studies from our group have revealed significant correlations between the abundance of specific bacterial species within the tumor tissue and the extent of T cell infiltration. Based on this background, in this study, we followed the hypothesis that at least a subset of T cells infiltrating CRC tissues may be specific to previously identified CRC-associated bacteria, and following activation within tumor tissues may exert anti-tumor effects.

We aimed to address the following specific points:

1. Testing of the ability of previously identified bacterial strains to induce T cell activation;
2. Characterization of the phenotypic and functional profiles of bacteria-reactive T cells;
3. Evaluation of anti-tumor potential of bacteria-reactive T cells;
4. Identification of bacteria-reactive T cells among TILs in CRC tissue.

We found that CRC-associated bacteria induced the expansion of peripheral blood T cells both in healthy donors and CRC patients, to similar extents. Specifically, bacteria-induced expansion was derived from the memory T cell compartment. Surprisingly, the predominant population of proliferating cells was identified as CD4-CD8- (double negative, DN) T cells induced by all three tested bacteria. We showed that bacteria-induced DN T cells possess cytotoxic innate-like properties. Furthermore, we demonstrated their T cell-mediated and cytokine-mediated killing capacity towards tumor cells. Additionally, we detected abundant DN T cell clusters in the analysis of the primary CRC infiltrates. We then employed single-cell RNA sequencing coupled with TCR sequencing of TILs from primary CRCs of five patients and autologous peripheral blood bacteria-reactive T cells, to comparatively assess their functional identities and clonal repertoire. Whole genome transcriptomic analysis revealed high identity matching between DN T cells and CD8+ cells, and high expression of innate receptors NKG2D and DNAM-I. Interestingly, large proportions of the bacteria-expanded cells were identified as unconventional TCRγδ or mucosa-associated invariant T (MAIT) cells in these five patients. Further TCR sequencing analysis revealed several dominant clonotypes induced by CRC-associated bacteria, of which some appeared to be cross-reactive to different bacteria. Ultimately, we identified bacteria-specific T cell clonotypes within the TILs.

Our findings cumulatively suggest that defined species of CRC-infiltrating bacteria are capable of inducing unconventional DN T cells endowed with cytotoxic anti-tumor properties. These findings may pave the way towards the development of novel therapeutic concepts aimed at boosting the intratumoral expansion of DN T cells for enhancing anti-tumor immune responses, even in poorly immunogenic CRCs.

Keywords: colorectal cancer, gut microbiota, double-negative T cells, single-cell RNA sequencing

Abbreviations

AMP	Antimicrobial peptide
anti-EGFR	Anti-epidermal growth factor receptor
anti-VEGF	Anti-vascular endothelial growth factor
APC	Antigen-presenting cell
AvrA	Anti-virulence factor A
B.f.	Bacteroides fragilis
BRAF	B-raf proto-oncogene, serine/threonine kinase
BSA	Bovine serum albumin
BTN	Butyrophilin
CCL17	Chemokine (C-C motif) ligand 17
CCL22	Chemokine (C-C motif) ligand 22
CCL5	Chemokine (C-C motif) ligand 5
CCR6	Chemokine receptor 6
CEA	Carcinoembryonic antigen
CEACAM8	Carcinoembryonic antigen-related cell adhesion molecule 8
CFSE	Carboxyfluorescein succinimidyl ester
CFU	Colony-forming units
CIMP	CpG island methylation pathway
CIN	Chromosomal instability
CMS	Consensus molecular subtypes
CRC	Colorectal cancer
CTL	Cytotoxic T lymphocyte
CTLA-4	Cytotoxic T-lymphocyte associated protein 4
CX3CR1	CX3C motif chemokine receptor 1
CXCL10	Chemokine (C-X-C motif) ligand 10
CXCL12	Chemokine (C-X-C motif) ligand 12

CXCL9	Chemokine (C-X-C motif) ligand 9
CXCR3	C-X-C chemokine receptor type 3
CXCR5	C-X-C chemokine receptor type 5
DFS	Disease-free survival
DMSO	Dimethyl sulfoxide
DN T	Double-negative T cell
DNAM-I	DNAX Accessory Molecule-1
ELISA	Enzyme-linked immunosorbent assay
EOMES	Eomesodermin
ER	Endoplasmatic reticulum
ETBF	Enterotoxigenic Bacteroides fragilis
F.n.	Fusobacterium nucleatum
FAP	Familial adenomatous polyposis
FBS	Fetal bovine serum
Foxp3	Forkhead box P3
GALT	Gut-associated lymphoid tissue
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GNLY	Granulysin
GZMA	Granzyme A
GZMB	Granzyme B
GZMK	Granzyme K
HDI	Human development index
HLA	Human leukocyte antigen
IBD	Inflammatory bowel disease
ICIs	Immune checkpoint inhibitors
IgA	Immunoglobulin A
IL	Interleukin
IFN-γ	Interferon gamma
IRF	Interferon regulatory factor

KLRG1	Killer cell lectin-like receptor subfamily G member 1
KRAS	Kirsten rat sarcoma virus
LAG3	Lymphocyte-activation gene 3
MAIT	Mucosal-associated invariant T cell
MAMP	Microbial-associated molecular pattern
MDSC	Myeloid-derived suppressor cell
MHC	Major Histocompatibility Complex
MICA	MHC class I polypeptide–related sequence A
MICB	MHC class I polypeptide-related sequence B
MLH1	MutL homolog 1
MMR	Mismatch repair system
MPO	Myeloperoxidase
MSH2	MutS homolog 2
MSH6	MutS homolog 6
MSI	Microsatellite instability
MSS	Microsatellite stability
MUC	Mucin
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NK cell	Natural killer cell
NKG2D	Natural killer group 2D
NKT	Natural killer T cell
NLR	NOD-like receptor
OD	Optical density
OS	Overall survival
OTU	Operational Taxonomic Unit
PAMP	Pathogen-associated molecular pattern
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate-buffered saline
PD-1	Programmed cell death protein 1

PD-L1	Programmed death-ligand 1
PHA	Phytohaemagglutinin
PMS2	PMS homolog 2
PRF1	Perforin
PRR	Pattern recognition receptor
PVR	Poliovirus receptor
RLR	Retinoid acid-inducible gene-I-like receptor
SCFA	Short-chain fatty acid
STAT3	Signal transducer and activator of transcription 3
T-bet or TBX21	T-box transcription factor
TAM	Tumor-associated macrophage
TAN	Tumor-associated neutrophil
TAP	Transporter Associated with Antigen Processing
Tcm	Central memory T cell
TCR	T cell receptor
TCRαβ	Alpha-beta T cell
TCRγδ or γδ T	Gamma-delta T cell
Tem	Effector memory T cell
TEMRA	Terminally differentiated effector memory T cell re-expressing CD45RA
Tex	Exhausted T cell
Tfh	T follicular helper cells
TGFβ	Transforming growth factor β
Th1	T helper 1 cell
Th17	T helper 17 cell
Th2	T helper 2 cell
Th9	T helper 9 cell
TIGIT	T cell immunoreceptor with Ig and ITIM domains
TIL	Tumor-infiltrating lymphocyte
TLR	Toll-like receptor

TMB	Tumor mutational burden
TNF-α	Tumor necrosis factor α
TP53	Tumor protein p53
Tpex	Precursor-exhausted T cell
Treg	Regulatory T cell
TRIF	TIR-domain-containing adapter-inducing interferon-β
TSB	Tryptic soy broth
Wnt	Wingless-related integration site
WT	Whole transcriptome
5-FU	5-fluorouracil

Introduction

1. Pathogenesis of human colorectal cancer

1.1. Epidemiology, risk factors, and symptoms

Colorectal cancer (CRC) is a prevalent malignancy affecting the colon (large intestine) and/or rectum. According to the World Health Organization's 2020 statistics, colorectal cancer is the third most commonly diagnosed cancer globally and the second leading cause of cancer-related mortality (Figure 1). However, the incidence of colorectal cancer exhibits notable regional and national disparities. Developed countries, including North America, Europe, and Australia, record the highest CRC incidence rates, potentially influenced by dietary and lifestyle factors[1] (Figure 2). In Switzerland, the frequency is relatively high, with approximately 6,000 new cases diagnosed annually, according to data from the Federal Statistical Office of Switzerland in 2018. Contrastingly, many developing countries, particularly those in Asia and South America, have witnessed a rising trend in CRC incidence. This phenomenon can be attributed, in part, to a shift towards more Westernized diets and lifestyles [1].

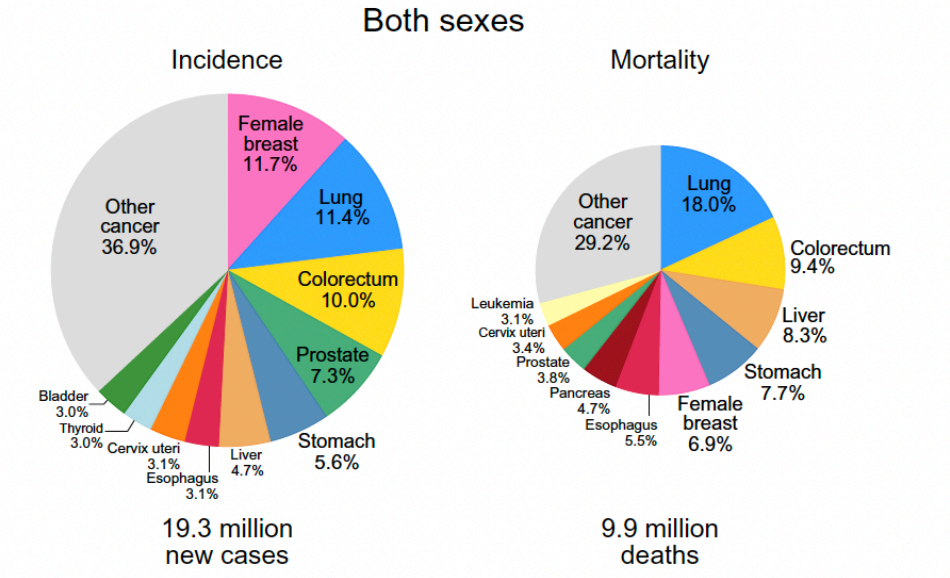


Figure 1: Distribution of Cases and Deaths for the Top 10 Most Common Cancers Worldwide in 2020 [1]

Despite the availability of advanced diagnostic tools and high-quality healthcare services in countries with higher Human Development Index (HDI), it remains a major concern that the highest proportion of CRC cases in these regions is detected at advanced stages, substantially diminishing the prospects for successful treatment [2]. This predicament may lie more in the lack of awareness or participation in screening programs. In recognition of this challenge, the American Cancer Society took proactive steps in 2018 by lowering the recommended age for initiating screening in individuals with a predisposition from 50 to 45 years [1]. This strategic move aimed to enhance early detection and intervention, thereby addressing the evolving landscape of CRC incidence, and facilitating timely and potentially life-saving treatments.

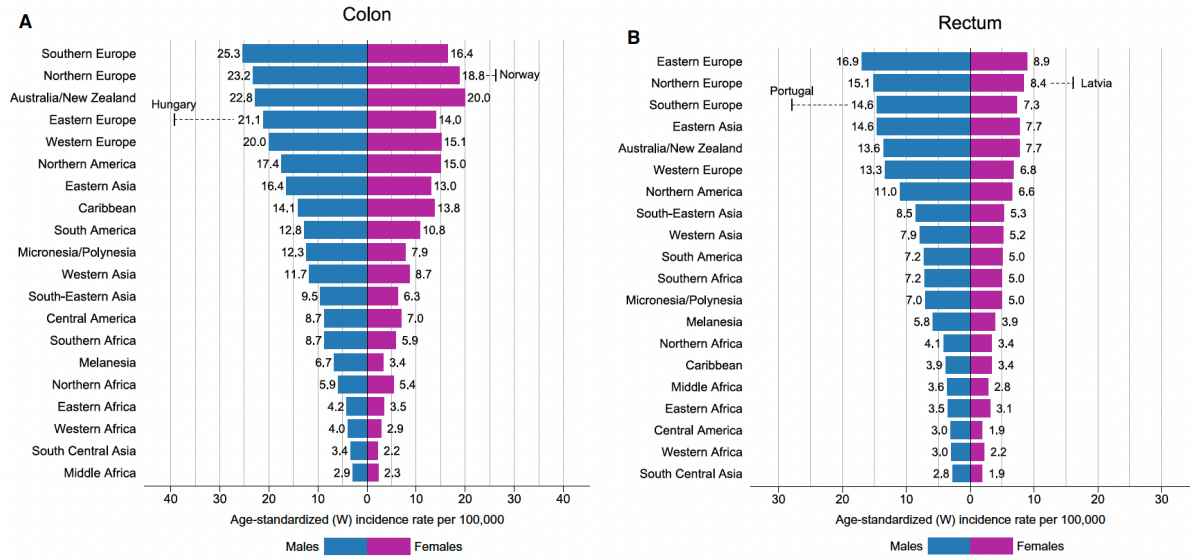


Figure 2: Region-Specific Incidence Age-Standardized Rates by Sex for Cancers of the (A) Colon and (B) Rectum in 2020 [1]

In a broader context, the risk of developing CRC increases with advancing age, with most cases diagnosed in individuals over the age of 50. Other reported risk factors encompass a family history of the disease, a diet rich in red or processed meats, smoking, excessive alcohol consumption, and obesity [3–5].

The symptoms of CRC can be varied and may not always be present during the early stages of the disease. Nevertheless, certain common symptoms include changes in bowel habits, such as diarrhea or constipation, occurrences of rectal bleeding or blood in the stool, abdominal pain or discomfort, feelings of weakness or fatigue, and unintended weight loss [4,5].

1.2. Carcinogenesis and subtypes

Based on the cellular origin, the most common type of CRC, accounting for about 90% of all cases, is adenocarcinoma, which develops in the glandular epithelial cells that line the inner surface of the colon and rectum. Neuroendocrine, squamous cell, adenosquamous, spindle cell, and undifferentiated carcinomas are very rare forms of CRC and will not be discussed further[6].

The process of carcinogenesis in CRC generally proceeds slowly over the years and includes a series of alterations in the genetic and morphological background (Figure 3). This traditional adenoma-carcinoma sequence model was first introduced by Fearon and Vogelstein [7]. Consistently, genetic alterations entail morphological and phenotypical changes in the colonic mucosa, and they are defined throughout the processes of tumor *initiation*, *promotion*, and *progression*. Over time, normal epithelium transforms into benign adenomatous polyps, subsequently developing adenomas with high-grade dysplasia, and finally progressing to the carcinoma stage, followed by tumor invasion, and spreading [6]. Although it is difficult to precisely determine the duration of each developmental phase, some studies suggest that it may take anywhere from 10 to 15 years or longer for a polyp to progress to cancer and on average 9 years to form metastasis [8]. However, interestingly, recent full-genome sequencing data from primary cancers and distant metastasis revealed that new mutations were not required for a tumor cell to migrate to distant sites, suggesting that metastatic seeding could happen at the early adenoma level. Indeed, the study performed by Hu et al. confirmed that in more than 80% of patients, metastatic dissemination took place while the primary tumor was not yet clinically detectable [9], opposing the prevailing linear progression model. Although the small patient cohort is the limitation of this study, it's noteworthy that it provided an alternative model distinct from the traditional adenoma-carcinoma, recognized as a parallel progression model.

CRC tumors have different potential for propagation, depending on the presence and expansion of selected aggressive clones, their specific phenotypic characteristics, as well as interactions with the host microenvironment and immune system [10]. Upon local spreading of the primary tumor mass, there are several mutually non-exclusive pathways engaged in propagation. Primary tumors with metastatic potential normally spread through the lymphatic system or the bloodstream [11]. Cancer cells can enter the lymphatic vessels surrounding the primary tumor site, and then travel through the lymphatic system to nearby regional lymph nodes, and other

organs. As well, cancer cells can invade nearby blood vessels and travel through the bloodstream to distant organs such as the liver, lungs, peritoneum, and brain. The liver is the most common site of metastasis in CRC due to the direct blood flow from the colon and rectum into the portal vein system. Over the course of the disease, approximately 50% of CRC patients develop liver metastases, an event representing a major cause of death [12].

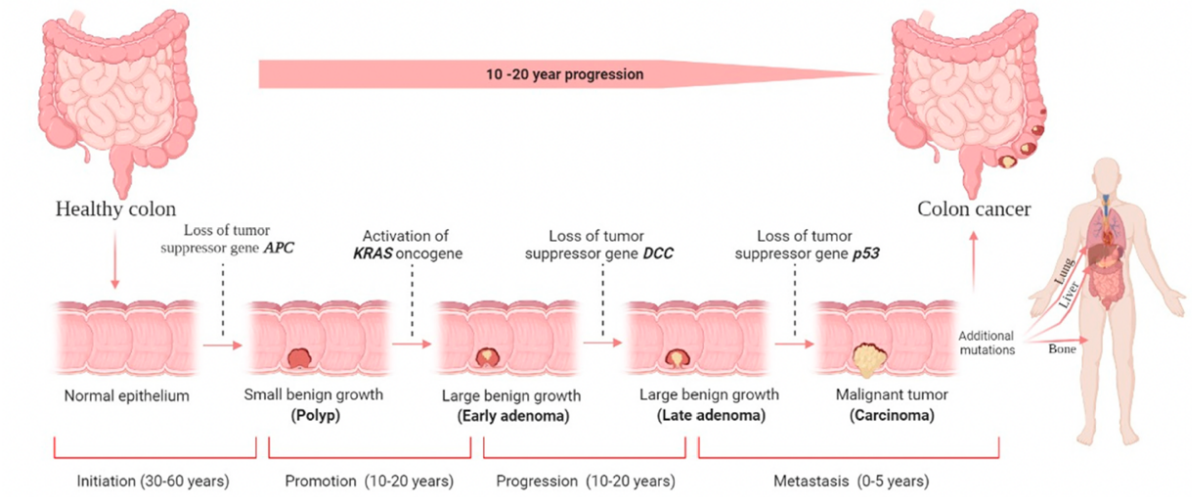


Figure 3: The development of CRC [8]

Considering that mutations in genes responsible for the initiation of CRC can be inherited or acquired, a distinction is made between hereditary CRC syndromes and sporadic CRC cases, respectively. There are several inherited colorectal cancer syndromes, including Familial adenomatous polyposis (FAP) and other polyposis syndromes, hereditary nonpolyposis colorectal cancer (also known as Lynch syndrome), altogether accounting for about 5% of all CRCs [13]. FAP arises from a germline mutation in the APC gene, while the genetic basis for Lynch syndrome lies in the mutations in one of the DNA mismatch repair genes (MLH1, MSH2, or MSH6) [14]. Overall, described pathologies are significantly less prevalent compared to sporadic CRCs which account for about 80-90% of all CRCs. These cancers have no family history and are typically caused by various mutations that develop spontaneously due to a combination of factors, such as aging, lifestyle choices (e.g., diet, physical activity, smoking), and environmental exposures [15].

There are three main underlying molecular pathways of CRC (Figure 4). The chromosomal instability pathway (CIN) is the most common type of genomic instability in CRC (around 70%), causing many alterations in chromosomal copy number and structure, resulting in activation of the Wnt pathway [16]. Fundamentally, these tumors are caused by oncogene mutational activation combined with tumor-suppressor gene mutational inactivation, and comprehensively by mutations in at least 4 to 5 genes [7]. Specifically, mutations of the APC, followed by KRAS and TP53 mutations were found to be the drivers of CIN conventional pathway, where the occurrence of each mutation usually follows the specific timeline during development. Nevertheless, the order of these mutations is not constant, and the accumulation of these mutations appears to be more essential than their order with respect to one another [7]. Another pathway leading to CRC is the CpG island methylation pathway (CIMP), which is characterized by a high level of DNA methylation in the CpG island and is often associated with microsatellite instability and BRAF mutations. The third pathway is microsatellite instability (MSI) which reflects the inability of the DNA mismatch-repair system to correct errors occurring during DNA replication. It is caused by mutations and/or epigenetic silencing of the MLH1, MSH2, MSH6, or PMS2 genes, and consequently, it includes an accumulation of short, repeated nucleotide sequences, called microsatellites, throughout the genome. These tumors are usually right-sided, poorly differentiated, and characterized by higher immune infiltrations [15]. It is important to note that the three molecular pathways are not mutually exclusive and mutational makeup alone does not allow precise and final tumor classification, nor prognosis of disease development [17].

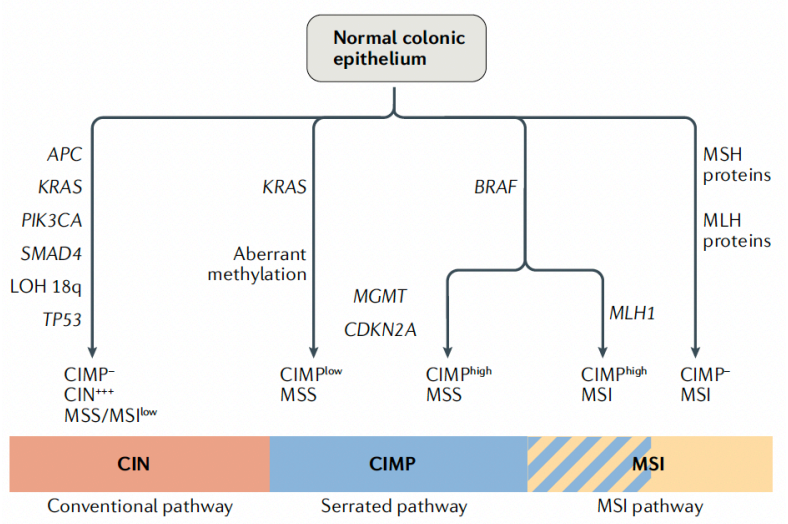


Figure 4: The three main molecular pathways of CRC [17]

In 2015 an international consortium of researchers established a new classification system for CRC called Consensus Molecular Subtypes (CMS), with the goal of developing a molecular classification system that would better reflect the heterogeneity of CRC. It is based on gene expression profiling from CRC samples obtained from multiple independent cohorts and is designed to identify robust molecular subtypes with clinical relevance. There are four identified CMSs with distinct clinical, pathological, and biological features [18,19]:

1. CMS1 (MSI immune): This subtype is characterized by microsatellite instability (MSI), high levels of immune cell infiltration (mainly Th1 and CTLs), and hypermutation. Patients with CMS1 tumors have a better prognosis than other subtypes and may benefit from immune checkpoint inhibitor therapy. It is represented in 14% of cases.
2. CMS2 (canonical): This subtype is characterized by the activation of canonical WNT and MYC signaling pathways and is associated with a high frequency of KRAS and APC mutations. Patients with CMS2 tumors have a relatively good prognosis, and they represent 37% of cases.
3. CMS3 (metabolic): This subtype is characterized by metabolic dysregulation and the overrepresentation of KRAS mutations. CMS3 tumors are associated with a worse prognosis, and they account for around 13%.
4. CMS4 (mesenchymal): This subtype is characterized by the prominent TGFβ signaling, and activation of stromal and angiogenic pathways, and is associated with high levels of cancer-associated fibroblasts and immune suppression. Patients with CMS4 tumors have the worst prognosis and may benefit from therapies that target the tumor microenvironment. It is represented in 23% of cases.

Due to its comprehensiveness, the CMS classification system holds promise for improving the diagnosis, prognosis, and treatment of CRC. However, it is not yet incorporated into routine clinical practice.

1.3. Staging and prognosis

Pathological staging of CRCs refers to the extent of the spread of cancer in the body [4]. The currently most widely used staging system for cancer worldwide is the TNM staging system, established by the Union for International Cancer Control (UICC) and the American Joint Committee on Cancer (AJCC). By this classification, all CRCs can be characterized for the presence of the primary tumor burden (T), involvement of draining and regional lymph nodes (N), and distant metastases (M) and assigned the relevant cancer stage[20] (Figure 5).

However, despite being a valuable tool for assessing the severity of cancer disease based on tumor invasion parameters, the TNM staging system comes with critical limitations. For instance, it cannot provide information regarding the tumor type and grade, tumor budding, genetic mutations, and overall patient health [21]. Also, it may not instruct on the best treatment options for individual cases. Importantly, the TNM staging system does not accurately predict patients' post-operative prognosis; i.e., there is substantial heterogeneity in clinical outcomes within the same TNM tumor stage [22–24]. Therefore, additional examinations such as histological subtyping, grading, analysis of genetic mutations and protein expression, mismatch-repair testing, and immunoscoring are increasingly being recognized in clinical practice [14,24].

Overall, a multidisciplinary approach that takes into account all determinants that contribute to the complexity and heterogeneity of CRC can help ensure that treatment decisions are individualized and based on the best available evidence [19].

T=primary tumour
TX=primary tumour cannot be assessed
T0=no evidence of primary tumour
Tis=carcinoma in situ: intraepithelial or invasion of lamina propria
T1=tumour invades submucosa
T2=tumour invades muscularis propria
T3=tumour invades through the muscularis propria into subserosa or into non-peritonealised pericolic or perirectal tissues
T4a=tumour penetrates the surface of the visceral peritoneum
T4b=tumour directly invades or is histologically adherent to other organs or structures

N=regional lymph nodes
NX=regional lymph nodes cannot be assessed
N0=no regional lymph node metastasis
N1a=metastasis in one regional lymph node
N1b=metastasis in two to three regional lymph nodes
N2a=metastasis in four to six regional lymph nodes
N2b=metastasis in seven or more regional lymph nodes

M=distant metastasis
MX=distant metastasis cannot be assessed
M0=no distant metastasis
M1a=distant metastasis to one site
M1b=distant metastasis to more than one site
Surveillance, Epidemiology and End Results Program data for 5-year stage-specific relative survival rates in colon cancers:[91,92]

- Stage I (T1, T2, N0): 97·1%
- Stage IIA (T3, N0): 87·5%
- Stage IIB (T4, N0): 71·5%
- Stage IIIA (T1, T2, N1): 87·7%
- Stage IIIB (T1, T2, N2): 75·0%
- Stage IIIB (T3, N1): 68·7%
- Stage IIIC (T3, N2): 47·3%
- Stage IIIC (T4, N1): 50·5%
- Stage IIIC (T4, N2): 27·1%

5-year stage-specific relative survival rates were similar for rectal cancer as compared with colon cancer.

Figure 5: CRC staging and correlation of TNM stages with prognosis [15]

1.4. Current treatments

Treatment options for individual colon or rectal cancer are determined by various factors, but generally, the decision is tailored based on the tumor staging. Surgery is considered a first-line treatment for tumors at stages I, II, and III. However, for rectal cancer, neoadjuvant treatment (chemoradiotherapy) is administered before surgery, aiming at the reduction of the large tumor mass [25]. Following surgery, patients at stage III and specific high-risk stage II patients undergo adjuvant chemotherapy as standard care, typically involving 5-FU and oxaliplatin or irinotecan[26]. Patients at stage IV with distant metastasis receive a combination of chemotherapeutic agents and in some cases targeted therapies such as anti-EGFR or anti-VEGF monoclonal antibodies potentially limiting the growth and spread of cancer cells and angiogenesis [27]. If the primary tumor and/or metastatic lesions are deemed resectable, surgery may also benefit some of these patients [25]. However, despite undergoing multiple rounds of treatment, established chemotherapy protocols have shown low response rates, leading to relapse in many patients [15].

As a novel approach to the treatment of various cancers, immunotherapy has brought tremendous advances. Immune checkpoint inhibitors directed towards CTLA-4, PD-1, and PD-L1 have demonstrated significant efficacy in the treatment of diverse malignancies [28]. In colon cancer therapy the key breakthrough occurred in 2017 with the approval of the first anti-PD1 antibody, pembrolizumab, for the treatment of metastatic patients with MSI tumors. Soon after, two more drugs targeting PD-1 (nivolumab) and CTLA-4 (ipilimumab) were approved for the same indication [29]. Multiple studies have demonstrated the remarkable efficacy of these therapies in the treatment of MSI colon cancers [30,31] and a very recent clinical trial has demonstrated complete response in 100% of patients with MSI locally advanced rectal cancers [32]. However, the use of immune checkpoint inhibitors is still limited only to patients with high mutational load tumors (15-20% are MSI CRCs), leaving all microsatellite stable (MSS) patients with a very limited choice of effective treatments [28].

2. Immune contexture in CRC

2.1. CRC-infiltrating immune cell populations

The concept of the immune contexture in cancer was first introduced by Jérôme Galon and his colleagues. It refers to the composition, functional orientation, and spatial organization of immune cells within the tumor microenvironment (Figure 6) [33].

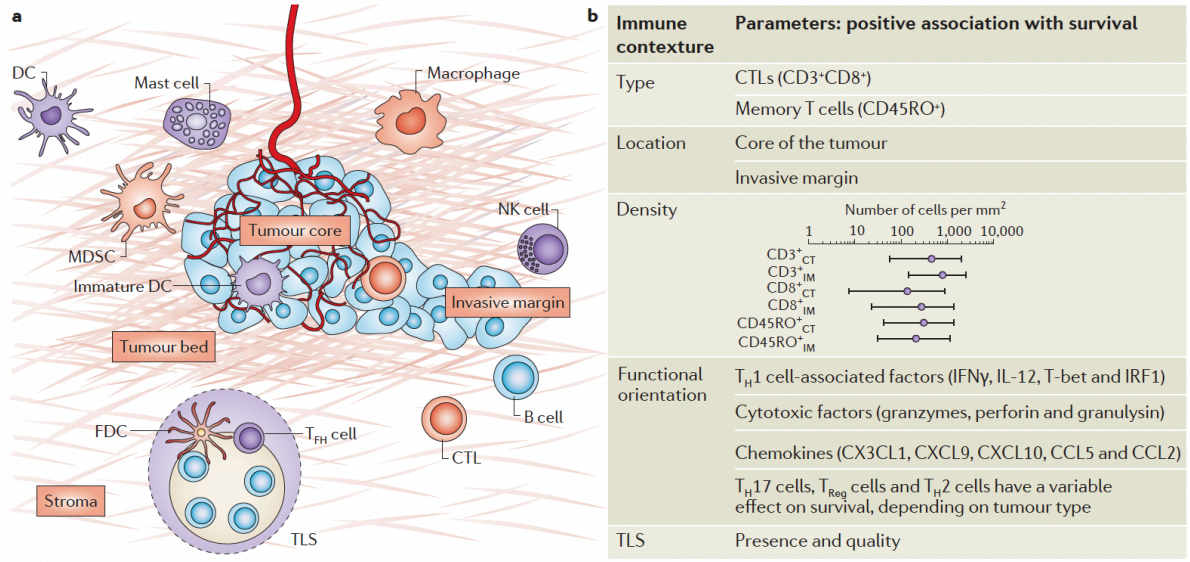


Figure 6: The immune contexture [34]

Accumulating histopathological examinations of solid tumors have provided evidence that all immune cell types, including macrophages, dendritic cells, neutrophils, NK cells, B cells, and various subsets of effector T cells may be found across the tumor mass [34]. Such immune infiltrates are variably scattered, exhibit vast heterogeneity between different tumor types, and can vary significantly among individual patients [24,33,34]. These observations suggested the different roles of immune cells in tumor control and prompted the investigation of their prognostic significance. To date, numerous studies have demonstrated the association between the level of specific immune infiltrates and clinical outcomes. E.g. tumor infiltration by Th1 and CD8+CD45RO+ T cells has been proven beneficial in various malignancies, including melanoma, breast, bladder, ovarian, lung, pancreatic, and colorectal cancers [34]. On the contrary, tumor infiltration by myeloid cells, and in particular, by myeloid-derived suppressor cells (MDSCs) has been linked to unfavorable prognostic outcomes and resistance to treatment in

CRC and many other malignancies [35]. MDSCs are known to promote tumor progression through various suppressive effects, such as the production of reactive oxygen and nitrogen species, secretion of IL-10 and TGFβ, inhibition of T cell activation and proliferation, reduction of NKG2D expression in NK cells, and polarization of macrophages toward a pro-tumoral M2 phenotype, thus facilitating immune evasion [36].

Tumor-associated macrophages (TAM) and tumor-associated neutrophils (TAN) are also important constituents of CRC immune infiltrates. Nevertheless, conflicting findings regarding their prognostic significance in CRC have been reported in several studies. This ambiguity can be attributed to the marked plasticity displayed by these cells or a lack of consensus on the method to identify them.

High infiltration of CD68+ TAMs predicted favorable long-term overall survival irrespective of their localization and correlated with less tumor budding, absence of lymph node metastasis [37], and the presence of T and B cell infiltration [38]. Surprisingly, an abundance of CD163+ TAMs, representing the M2 phenotype, has also shown an association with higher overall survival in CRC [37]. In contrast, Herrera et al. reported that individual expression of M2 macrophages, or their combination with cancer-associated fibroblast was negatively correlated with disease-free (DFS) and overall survival (OS) in a similar size patient cohort [39].

TANs constitute the predominant cell type in CRC tissue, with their infiltration evolving over stages I-III similarly to T cells but significantly decreasing at stage IV when the tumor enters the phase of immune escape [40]. High tissue densities of CD16+MPO+ [41,42] and CD66b+ TANs in invasive margins [43] and throughout whole tumor sections [40] have been found to significantly correlate with improved clinical outcomes in CRC. However, contrasting findings have indicated that neutrophil-rich primary CRCs were associated with higher tumor grade and poorer 5-year recurrence-free survival in both mismatch repair system (MMR)-proficient and MMR-deficient CRC patients [44] and infiltration by CEACAM8+ TANs has been identified as an independent factor of poor DFS [45].

Furthermore, the detection of CD83+ dendritic cells has been linked to a favorable prognosis in CRC [46], while certain immune cell subsets, such as NK cells, have not been found to be predictive of the clinical outcome [42].

2.2. CRC-infiltrating T lymphocytes: phenotype, functional orientation, and antigenic specificities

In the context of colorectal cancer, a systematic analysis of immune cell infiltrates through comprehensive gene expression and immunohistochemical investigations on extensive patient cohorts, as conducted by Galon and colleagues, has yielded crucial insights into the significance of the T cells within the immune contexture. Their collective findings have demonstrated that within CRC tissue, T cells are present both in the tumor core and invasive margins, albeit their density diminishes in advanced TNM stages [34,47]. Notably, high infiltration levels of specific T cell subpopulations, including CD8+CD45RO+ T cells [47,48], CXCR5+ Tfh cells [49], and INF-γ-producing Th1 cells [50,51], have been correlated with improved patient survival. These observations have been validated by two other studies, which confirmed the prognostic value of tumor-infiltrating lymphocytes or CD45RO+-cell densities in CRC, independently of clinical, pathological, and molecular characteristics [52,53].
Furthermore, the examination of cytotoxic memory T cell subsets within CRC tissues has revealed distinct prognostic implications. Masuda et al. employed single-cell analysis of TILs in treatment-naïve CRC patients to differentiate between the GZMK+KLRG1+ effector subset, enriched in patients with favorable clinical outcomes, and the GNLY+CD103+ resident subset, which showed no prognostic value [54]. However, when dissecting cytotoxic TILs in MSS tumors, functionally exhausted immunoscenescent CD8+CD28- T cell subset has been identified as dominating the T cell compartment. Crucially, this phenotype was observed across both early and advanced stages of CRC, indicating the early establishment of an immunosuppressive immune landscape during CRC development [55].

Surprisingly, the infiltration of Foxp3+ regulatory T cells (Treg), which typically portends poor prognosis in most malignancies, has been associated with improved clinical outcomes in CRC [50,56–58]. Single-cell analysis of CRC TILs has enabled the identification of two subsets of Foxp3+ Tregs with divergent prognostic implications: the Helios+ subset correlated with favorable prognosis, while the Helios- infiltrates, exhibiting elevated expression of CD38 and LAG3 molecules, correlated with poorer clinical outcomes [54].

Conversely, the role of interleukin-17-producing CD4+ T cells (Th17) remains a subject of debate. Tosolini et al. predicted a poor prognosis for patients with high expression of Th17

cluster genes [50], and in the same year, another study confirmed the detrimental impact of IL-17-producing cells, which were implicated in promoting CRC development through angiogenesis [59]. However, a more recent study involving over 1400 CRC patients found no association between IL-17+ infiltrates and clinical outcomes [60].

Foremost among these findings, Galon's study has demonstrated that immunological data can surpass the traditional TNM-based prognostic predictions for the first time in any cancer type. This study laid the foundation for the concept of the Immunoscore, wherein the combined enumeration of two lymphocyte populations (CD3 and CD8) in both the tumor core and invasive margin has been established as a clinically valuable biomarker in CRC, with prognostic significance superior to TNM predictions [61]. The Immunoscore ranges from I0 (absence of both cell subsets in both regions) to I4 (high T cell densities in both regions), with the former categorized as 'cold tumors' and the latter as 'hot tumors' [62]. Importantly, the Immunoscore has been shown to predict overall survival and identify patients who would derive the most benefit from chemotherapy based on their pre-existing immunity [63].

While higher T cell densities in CRC tissues are generally associated with the MSI phenotype, it is essential to recognize that some patients with MSS CRC also present tumors enriched with T cells. Galon et al. demonstrated that high T cell infiltration can confer benefits to both MSS and MSI CRC patients [47], thereby prompting inquiries into the antigenic specificities and factors that drive T cell accumulation in these distinct contexts, irrespective of the TMB status.

Several in vitro studies have provided evidence of CD8+ T cell recognition of autologous tumor-associated antigens, including but not limited to carcinoembryonic antigen (CEA), mucin, EGFR, and HER-2 [64,65]. However, it is important to note that only a minority of TILs have been unequivocally identified as tumor-specific cells [66–70]. Remarkably, recent reports have indicated that the proportions of neoantigen-specific CD8+ T cells within human cancers, including CRC, rarely exceed 0.5% of the total TIL population [66–69]. This observation indirectly underscores the notable clinical relevance of "bystander" T cells, which recognize antigens unrelated to the tumor [69]. Nevertheless, the precise mechanisms governing their disparate infiltration into CRC tissue and their prognostic implications remain subjects of extensive investigation.

3. CRC-associated microbiota

3.1. Physiology of colonic mucosa and gut microbiota

The intestinal barrier is a dynamic entity whose primary function is to maintain the homeostatic relationship with the large density of gut commensals by minimizing their translocation and dissemination across the body. Therefore, the preserved integrity of the gut barrier is crucial for good health [71,72].

The normal colonic mucosa is composed of several layers that collectively form the inner lining of the colon (Figure 7). The entire lumen of the intestine is lined with a thick mucus layer providing a primary shield. The major constituents of mucus are membrane-bound and secretory mucins (MUCs)- glycoproteins that provide lubrication and prevent microbial penetration [71,73]. Underneath mucus is a single-cell-layered epithelium, consisting of different types of epithelial cells, including enterocytes (specialized in nutrient absorption and electrolyte transport), goblet cells (specialized in mucus production), Paneth cells (specialized for production of antimicrobial peptides, AMP), and hormone-producing enteroendocrine cells [74]. The critical elements holding the epithelial barrier intact are intercellular junctions, and they are positioned between the apical and basal surfaces of enterocytes, tentatively allowing, or restricting the paracellular passage of molecules [72]. The epithelium is further supported by connective tissue called the lamina propria, infiltrated by a spectrum of immune cellular components, including T lymphocytes, dendritic cells, macrophages, plasma cells, and innate lymphoid cells. Lamina propria provides local host defense through a combination of innate and adaptive immune responses orchestrated by cells producing immunoglobulin A (IgA), cytokines, and other factors [72,74].

The human gastrointestinal tract hosts the largest collection of microbes in the body. Trillions of gut bacteria, collectively known as *'gut microbiota,'* have been recognized as an integral part of the human organism, shaping immunological, metabolic, structural, and neurological landscapes [75]. Microbial colonization of the gut begins right upon birth when neonates experience the first exposure to the microbes derived from the mother and environment. During early life, along with the seeding of the primordial microbial species, the foundations of the immune responses are established [76]. Nonetheless, during life, the composition of the gut microbiota is shaped by numerous factors such as diet, medications, and lifestyle. In healthy

individuals, the two most represented phyla are *Bacteroidetes* and *Firmicutes*, comprising more than 90% of the total gut population, which is predominantly anaerobic [75].

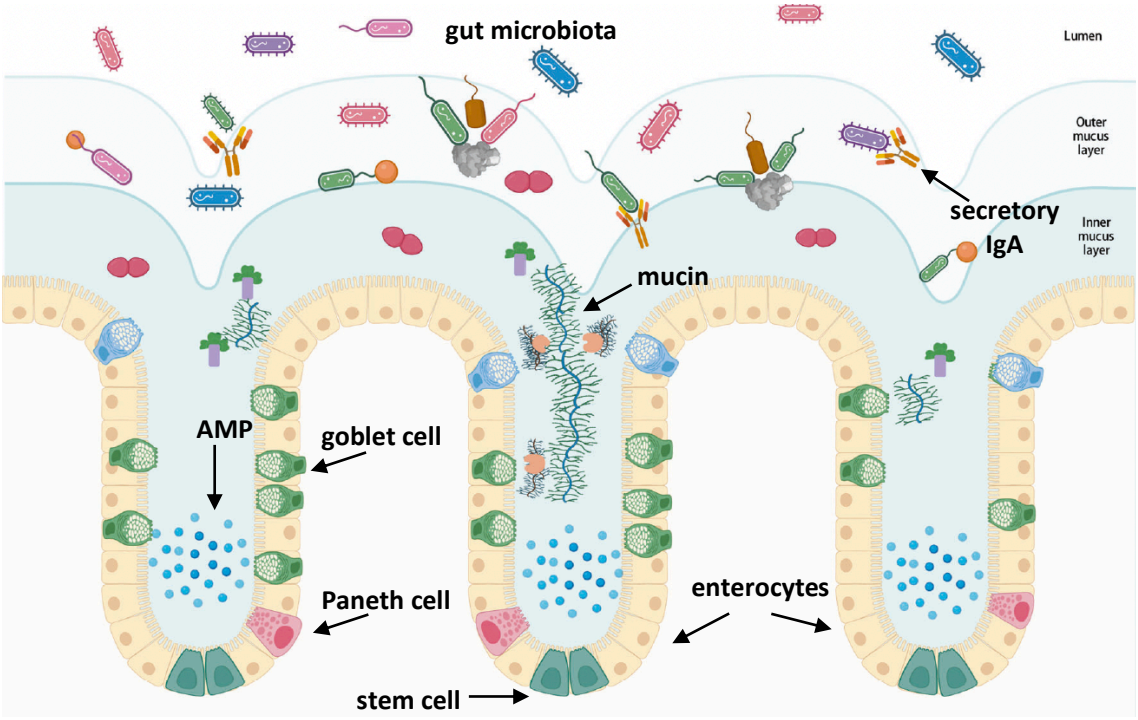


Figure 7: The main components of colonic mucosa (adapted from [77])

Homeostatic interactions between gut microbiota and the colonic mucosa involve a complex and mutually beneficial relationship that helps maintain the balance and function of both the microbiota and the host. These bacteria have a symbiotic relationship with the host, and they are referred to as commensal bacteria [73]. In this habitual exchange, gut commensals receive nutrients and an environment suitable for their survival, while they aid in digestion and absorption, produce vitamins, and prevent the colonization of harmful pathogens. They also contribute to the synthesis of short-chain fatty acids (SCFAs), which are the main energy source for colonocytes and beneficial modulators of intestinal health and immune functions [78].

Intestinal epithelial cells normally express several types of pattern recognition receptors (PRRs) responsible for sensing the surrounding microbes. These receptors belong to functionally distinct classes, including Toll-like receptors (TLRs), nucleotide-binding oligomerization domain (NOD)-like receptors (NLR), and RIG-I (retinoid acid-inducible gene-I)-like receptors (RLRs). In the colonic mucosa, microbial-associated molecular patterns

(MAMPs), such as microbial lipopolysaccharides, peptidoglycans, and flagellins, are recognized mainly via TLRs [74]. TLRs have different localization in the mucosa and depending on the nature of the MAMP induce a specific inflammatory response. Normally, TLR signaling results in several protective effects including the release of chemokines and antimicrobial peptides, IgA secretion, and epithelial cell proliferation upon injury [78]. Commensal bacteria are generally harmless, and they elicit homeostatic, controlled inflammatory, or tolerogenic responses. However, under certain circumstances, such as a weakened immune system or disruption of the microbial balance, known as gut dysbiosis, they can become opportunistic pathogens and cause infections and chronic inflammation, contributing to inflammatory bowel disease (IBD), metabolic disorders, and other pathologies [78].

3.2. The role of CRC-associated microbiota in CRC development and progression

In recent years, studies investigating perturbations in gut bacterial composition between healthy individuals and CRC patients have attracted a lot of attention. The comparative analysis of stool samples from healthy individuals, IBD patients, and CRC patients enabled the identification of unique bacterial strains that correlate with the pathology of the disease [79]. One such study conducted by Feng et al. addressed the profound shifts in the gut bacterial composition during the development of CRC; patients at advanced adenoma and carcinoma stages displayed higher levels of *Bacteroides*, *Alistipes,* and *Fusobacterium* species, while *Ruminococcus*, *Streptococcus*, and *Bifidobacterium* were significantly decreased [79]. Importantly, an increasing number of studies support this discrimination, emphasizing the fecal abundance of *Fusobacterium* and *Bacteroides* species in CRC patients [80,81].

Inflammation in the gut is largely associated with CRC. Chronic and poorly controlled gut inflammation may precede the development of CRC, although this is usually addressed as the cause in less than 5% of all CRCs [17]. In most sporadic cancers, loss of function in tumor-suppressor genes p53 or APC [17] has critical impacts on the differentiation of certain epithelial cell lineages causing the loss of intercellular junctional proteins and intestinal barrier function [82,83]. As a result of the impaired permeability, gut commensals may actively translocate from the lumen into the submucosa, populating the tumor stroma and enabling direct contact with residing immune cells. Thus, CRC represents a unique cancer type arising in a dynamic environment heavily populated by microbes. An increasing number of reports highlight the

significance of bacteria that infiltrate CRC and their involvement in the development of CRC through various mechanisms. To date, several bacterial species have been recognized to own genotoxic or oncogenic properties, and/or to drive specific pro-tumoral immune responses [17] (Figure 8).

Fusobacterium nucleatum (F.n.) is one of the most extensively studied gut species in the context of CRC. It is an invasive, adherent, Gram-negative anaerobic bacterium, found overrepresented in primary CRC tissues and distant metastasis [84]. Thus far, it has been mostly associated with pro-tumorigenic effects in CRC, by acting on both tumor cells and immune cells in the tumor microenvironment. Adhesive and invasive properties of F.n. have been described. In a study conducted by Rubinstein et al., F.n. was suggested to have a direct effect on the proliferation of E-cadherin-expressing CRC cells upon FadA-mediated binding to E-cadherin and activation of the Wnt signaling pathway [85]. Kostic et al. have demonstrated the ability of F.n. to promote intestinal tumorigenesis in the ApcMin/+ mouse model and the recruitment of myeloid-derived suppressor cells [86]. Gur et al. highlighted Fap2, another virulence factor of F.n. that prevents NK- and T cell-mediated cytotoxicity toward CRC via binding to TIGIT [87]. In a large cohort study performed by Mima et al., the abundance of F.n. in CRC tissue correlated with stage and shorter survival of CRC patients and has been proposed as a prognostic biomarker [88]. However, the presence of F.n. was negatively associated with TILs in MSI-high tumors, but positively associated with TILs in MSS tumors [89]. This observation suggests conflicting effects of F.n. relative to the MSI status, and proinflammatory properties of F.n. that might be exploited in MSS CRCs. Moreover, F.n. supplementation has been demonstrated to enhance the effectiveness of PD-L1 treatment and promote the accumulation of IFN-γ +CD8+ T cells in mice undergoing treatment, therefore contributing to favorable immune modulation [90].

Bacteroides fragilis (B.f.) is another anaerobe comprising two molecular subtypes - non-toxigenic, and enterotoxigenic B.f., together representing a relatively small proportion of the fecal microbiota, around 1% [91]. Several studies in mice have demonstrated that enterotoxigenic B.f. (ETBF) toxin can trigger colitis progressing to colon carcinogenesis via STAT-3 signaling and by inducing Th17 cell responses [92,93]. Metagenomic studies in humans have found a higher abundance of ETBF in CRC tissue, compared to healthy controls [91], especially in early-stage lesions [94], supporting the implications of this bacterium in CRC development.

Certain virulence factors have been identified for their ability to cause direct DNA damage, as seen with colibactin produced by *Escherichia coli* [95]. Others, like the anti-virulence factor A (AvrA) from *Salmonella enterica*, target the Wnt/β-catenin pathway, interfering with tumor-suppressor protein p53, and ultimately resulting in uncontrolled proliferation of CRC cells [96,97]. The relative abundance of these bacteria in stool samples has been proposed to be used as a CRC biomarker. However, larger cohort studies are needed to validate the diagnostic values of microbial signatures, which could prove invaluable in categorizing patients for best-personalized treatment.

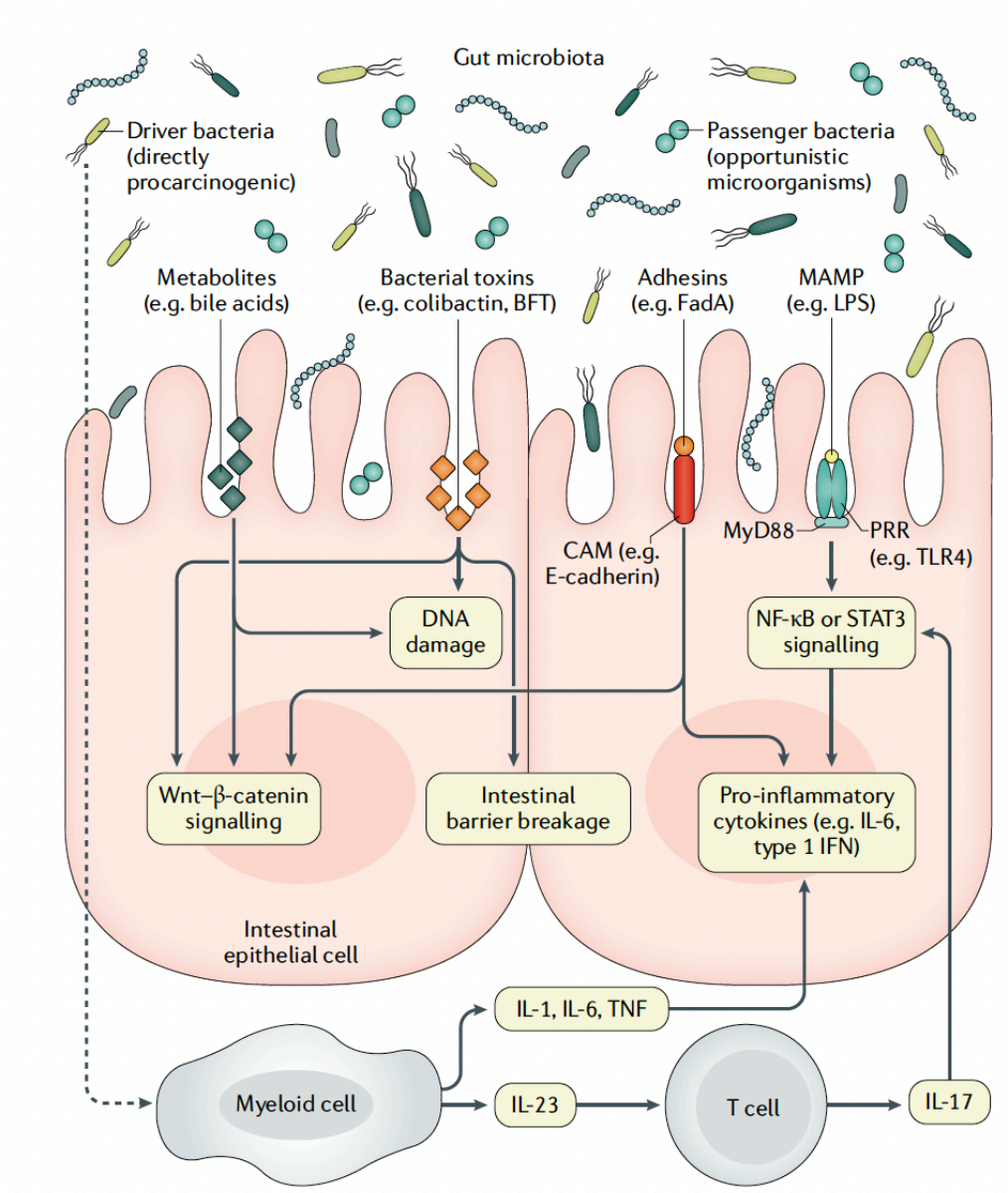


Figure 8: Microbiota-associated mechanisms involved in the pathogenesis of colorectal cancer [98]

3.3. The interplay between tumor-associated microbiota and immune contexture

Numerous studies have reported noteworthy associations between specific bacteria and immune infiltration, emphasizing the importance of differentiating between various bacterial species. Notably, *Blautia* and *Faecalibacterium* species have demonstrated correlations with increased tumor T cell infiltration [99]. Conversely, F.n. tumor infection has been associated with adverse survival outcomes and an abundance of exhausted CD8+ and FoxP3+ regulatory T cells within the tumor microenvironment [100]. However, in mucinous CRC, a high F.n. abundance has shown promise by promoting an increased proportion of CD4+ lymphocytes and M1 macrophages, potentially enhancing clinical outcomes [101].

Furthermore, our study conducted by Cremonesi et al. has significantly contributed to our comprehension of the microbiota's role in orchestrating T cell trafficking and influencing prognosis in colorectal cancer [102]. This study unveiled distinct abundance patterns of multiple bacterial species, correlated with varying degrees of T cell infiltration. In-depth analysis of chemokine gene signatures within CRC tumors has revealed their association with specific T cell subsets. For instance, chemokines like CCL5, CXCL9, and CXCL10 were associated with cytotoxic T lymphocytes and T-helper (Th)1 cells, while others like CCL17, CCL22, and CXCL12 were associated with Th1 and regulatory T cells [102]. Importantly, these findings highlighted the capacity of tumor cells to express these chemokines in response to gut bacteria, both in vitro and in vivo, underscoring the intricate interplay between gut microbiota and the immune landscape in CRC.

Crucially, recent studies have increasingly demonstrated the protective anti-tumor responses upon administration of certain bacterial species. Montalban et al. have shown that a combination of *Clostridiales* strains produces a potent anti-tumor effect in CRC via CD8+ T cells [103]. Similarly, Lu et al. found that fecal bacteria from CRC patients can enhance the degranulation and cytotoxicity of CD8+ T cells. This effect was reliant on antigen-presenting cells (APCs) and exhibited a higher proinflammatory profile compared to fecal bacteria sourced from individuals without CRC [104]. Additionally, Tanoue et al. identified a consortium of 11 bacterial strains that significantly enhanced the efficacy of anti-PD-1 treatment by eliciting CD8+-mediated anti-tumor response [105].

Given the complexity of the tumor microenvironment and the numerous factors involved, predicting the overall impact of individual bacterial species is challenging. Significant efforts have been made to understand the interactions between gut commensals and immune cells. However, our comprehension of these processes is still evolving. Specifically, we have yet to explore whether TILs could be reactive to tumor-infiltrating bacteria and potentially play a role in anti-tumor responses.

4. T cell responses to gut bacteria

4.1. Antigen-presenting cells-mediated responses

The innate immune system serves as the primary defense against microbial pathogens, providing immediate recognition and response to pathogen-associated molecular patterns (PAMPs) through PRRs. This rapid response is pivotal for controlling microbial infections and plays a central role in conveying structural information to the adaptive immune system [106]. Upon the entry of pathogens, the host's innate immune system coordinates a complex interplay between molecules derived from pathogens and host sensors. This dynamic network involves various innate immune cells, such as macrophages, dendritic cells, and neutrophils, which express a range of PRRs, TLRs, NLRs, C-type lectin receptors, and RLRs, which collectively are capable of sensing a wide spectrum of PAMPs. TLRs, in particular, are critical for recognizing bacteria [107]. Activation of PRRs initiates a cascade of signaling pathways that activate various adapter proteins, such as MyD88 and TRIF, and transcription factors, such as NF-κB and interferon regulatory factors (IRFs). This ultimately induces the secretion of antimicrobial peptides, proinflammatory cytokines and chemokines, and type I interferons, effectively combating infection through direct pathogen elimination or replication inhibition[108].

Simultaneously, structural information is transmitted to the adaptive immune system, which generates a robust antigen-specific response, ultimately resolving the infection [106]. Microbial epitopes form complexes with Major Histocompatibility Complex (MHC) molecules and are transported to the cell surface. MHC molecules can present processed peptides with substantial structural variations, enabling the recognition of even mutated foreign molecules, provided the corresponding TCR exists. This exclusive recognition mechanism, designed to avoid

detrimental immune responses against self-antigens, aligns well with the development of immune memory and rapid reactivation [109].

There are two major classes of MHC molecules: MHC class I and MHC class II, each playing distinct roles in antigen presentation and recognition. Cytosolic proteins, such as those from viruses or intracellular bacteria, are broken down into peptides within the host cell's cytoplasm by the proteasome. These peptides are then transported into the endoplasmic reticulum (ER) with the assistance of a transport protein called TAP (Transporter Associated with Antigen Processing). In the ER, the peptides are loaded onto MHC class I molecules, forming MHC class I-peptide complexes. These complexes are transported to the cell surface and presented to CD8+ T cells. When the specific TCR recognizes a foreign peptide presented on MHC class I with high affinity, it triggers a cascade of signaling events within the T cell, leading to its activation and response to the antigen [109].
In contrast, extracellular proteins, such as those from extracellular bacteria or parasites, are engulfed by APCs and digested by enzymes in endosomes. Peptides are loaded onto MHC class II molecules within the endosomes, creating MHC class II-peptide complexes. These complexes are also transported to the cell surface but are presented to CD4+ T cells. CD4+ (helper) T cells play a central role in coordinating the immune response, providing signals to other immune cells, such as B cells and cytotoxic T cells, to orchestrate an effective immune response [109].

It's important to note that bacterial antigen presentation extends beyond peptide antigens and includes non-peptide microbial metabolites or bacterial cell wall components, which are presented by MHC-like nonpolymorphic molecules. These components initiate adaptive responses through unconventional T cells, a topic that will be discussed later.

The relative importance of each antigen presentation pathway is likely dependent on the nature of the antigen, the cell type involved, and the uptake mechanism, which can direct the antigen to different processing compartments [109].

4.2. Conventional T cell-mediated responses

During the first three years of a human's life, the composition of the microbiota undergoes a gradual stabilization process, coinciding with a phase of expansion in the T cell repertoire [110]. In a recent study conducted by Zegarra-Ruiz et al., significant insights were gained into the thymic development of gut microbiota-specific T cells. Early trafficking of microbial antigens from the gut to the thymus results in the expansion of microbiota-specific T cells and the process is facilitated by CX3CR1+ dendritic cells, abundant in the intestinal sights [111]. Moreover, the commensal microbiota present in the colon plays a pivotal role in educating the immune system after thymic development. The adaptive immune response, with a specific focus on CD4+ T cell-mediated responses, is shaped to effectively distinguish between harmless intestinal stimuli and pathogens. Physiological CD4+ T cell differentiation at the mucosal barrier ensures the development of tolerance rather than the expansion of pathogenic effector responses, thus providing natural prevention against the onset of colitis [112,113].

The existing body of literature extensively discusses the antigen-specific recognition of commensal bacteria by conventional CD4+ T cells, exploring their distinct responses and their potential for cross-reactivity with other antigens [114]. Notably, in patients with inflammatory bowel disease, *Bacteroides* and *Bifidobacterium* species have been found to induce clonal expansion of CD4+TCRαβ+ T cells, which exhibit cross-reactivity with other *Enterobacteria* [115]. Furthermore, Hegazy, West, et al. have shed light on the robust peripheral immune responses to seven prevalent intestinal species, highlighting that gut commensal-specific T cells circulate systemically in healthy individuals. The expanded CD4+ memory T clones exhibit a diverse TCR repertoire and demonstrate functional heterogeneity, displaying characteristics of both Th1 and Th17 phenotypes [116]. It is worth noting that CCR6+CXCR3+ Th1/Th17 cells are emerging as highly enriched memory T cell subsets with broad reactivity to various commensals. These cells have been reported to undergo antigen-specific proliferation in response *to S. aureus, E. coli,* and *L. rhamnosus* [117], as well as other *Enterobacteriaceae* species [118] in healthy individuals. They produce cytokines such as IFN-γ, IL-17A, and IL-22, and exhibit cross-reactivity with different *Enterobacteriaceae* species presented by MHC-II molecules [118].

4.3. Unconventional bacteria-reactive T cells

A particular subset of innate-like T cells, poorly represented in the peripheral circulation but extensively enriched at the mucosal sights, comprises unconventional T cells. This subset includes natural killer T cells (NKT), mucosal-associated invariant T (MAIT) cells, and gamma-delta (γδ) T cells. These unconventional T cells possess unique antigen recognition properties, as they are constrained to recognize antigens presented by monomorphic CD1, MR1, and butyrophilin (BTN) molecules [106,119]. Importantly, they exhibit the ability to recognize non-peptide antigens, such as lipids, vitamin B metabolites, and phosphoantigens, respectively [120]. An integral aspect of unconventional T cells is their developmental trajectory. These cells acquire effector characteristics before thymic egress, resulting in their rapid cytokine release and the expression of chemokine receptors and integrins. This early differentiation facilitates their accumulation in tissues, where they promptly detect and respond to alarmins and environmental changes [121]. Furthermore, despite having a relatively limited diversity of TCRs, recent findings have unveiled the remarkable responsiveness of unconventional T cells to microbial-derived antigens. This semi-invariant nature of their TCRs shares similarities with innate immune receptors, narrowing their antigenic repertoire and aligning them with the concept of innate-like T cells [119].

MAIT cells are predominantly localized in mucosal tissues and comprise up to 10% of the total circulating T cell population [119]. Their tissue abundance is critically influenced by early-life microbiota colonization. MAIT cells appear as CD8+ or CD4-CD8- (double-negative, DN) cells, and exhibit both effector and memory characteristics, with rapid responses to microbial metabolites originating from riboflavin (vitamin B2) and folic acid (vitamin B9) presented by MR1 molecules [106]. Activation through their semi-invariant TCR, specifically composed of Vα7.2–Jα33 paired with Vβ2 or Vβ13 chains, initiates the production of inflammatory cytokines including IFN-γ, TNF-α, IL-17, and IL-22, alongside cytotoxic molecules like granzyme B and perforin. Moreover, MAIT cells enhance their tissue-homing capacity by upregulating chemokine receptors CCR5, CCR6, and CXCR6 [106].

MAIT cells play pivotal protective roles in the context of bacterial infections. Notably, CD8+ MAIT cells have demonstrated their capability to mount immunostimulatory responses against pathogens like *Mycobacterium tuberculosis* and *Escherichia coli* upon the recognition of MR1-

ligand complexes, resulting in the production of key cytokines such as IFN-γ and TNF-α [122]. Remarkably, specific MAIT cell clones have displayed cross-reactivity towards other bacterial species, including *Salmonella typhimurium* and *Staphylococcus aureus* [122]. Furthermore, MAIT cells have been instrumental in controlling invasive pneumococcal infections, with *Streptococcus pneumoniae*-infected dendritic cells inducing robust MAIT cell activation and IFN-γ production in an antigen-dependent manner [123].

MAIT cells have long been recognized for their distinctive specificity towards riboflavin-producing bacteria [124]. Experimental investigations involving 47 bacterial species from diverse phyla have shown that only those bacterial species harbouring the riboflavin pathway, primarily found in the *Bacteroidetes* and *Proteobacteria* phyla, were capable of stimulating MAIT-TCRs, triggering cytotoxic responses mediated by IFN-γ, TNF-α, and granzyme B [125]. Conversely, bacteria lacking the genes responsible for the riboflavin pathway, such as *Enterococcus faecalis*, failed to stimulate MAIT cells [125]. However, intriguingly, recent findings have expanded our understanding of MAIT cell reactivity. Notably, in healthy individuals, MR1-restricted MAIT cells producing IFN-γ have been identified, and they exhibit specificity towards *Fusobacterium nucleatum*, a microbe unrelated to riboflavin production [126]. This intriguing discovery highlights the versatility of the MR1 binding pocket and underscores the remarkable plasticity exhibited by these unconventional T cells.

Gamma-delta (γδ) T cells represent a distinctive subset of unconventional T cells, predominantly belonging to the double-negative subset, making up approximately 5% of the total T cell population in circulation [127]. The highest frequency of γδ T cells is found in the gut mucosa [128]. In the peripheral blood of healthy adults, the predominant population consists of Vγ9Vδ2+ γδ T cells, while Vδ1 subsets, coupled with diverse γ-chains are more prevalent in peripheral tissues [119]. These cells exhibit remarkable functional plasticity, responding to local microenvironmental cues by differentiating into Th1-like, Th2-like, Th9-like, Th17-like, and Treg-like cells. Consequently, γδ T cells contribute to a wide array of immunological responses, characterized by the production of an extensive spectrum of cytokines, including IFNγ, TNFα, IL-4, IL-10, IL-9, IL-17, and IL-22 [119].

Recent advances in our understanding have unveiled the pivotal role of butyrophilin-related proteins, particularly BTN3A1, in mediating the recognition of microbial-derived

phosphoantigens by Vγ9Vδ2+ γδ T cells [129]. These cells have also emerged as key players in the defense against transformed and infected cells [127]. They can produce Th1 cytokines, and antimicrobial peptides, and exhibit cytolytic activity, particularly in response to specific pathogens like mycobacteria [129]. Notably, their activation can occur independently of TCR stimulation, inducing production of IFN-γ and IL-17 and enabling swift responses to infections[128].

Vγ9Vδ2 T cells have been shown to undergo expansion when exposed to live Bacillus Calmette-Guérin (BCG) or mycobacterial lysates. These expanded Vγ9Vδ2 T cells possess the ability to produce Th1 cytokines, granulysin, and various antimicrobial peptides. Additionally, when co-cultured with mycobacteria-infected human monocytes, these expanded Vγ9Vδ2 T cells demonstrate cytolytic activity [130].

NKT cells have also been recognized for their reactivity towards bacteria. This reactivity is attributed to their invariant TCRs capable of binding to a range of lipid antigens, including microbial lipids, when presented by CD1d molecules [106]. In various pulmonary bacterial infection models, such as those involving *Streptococcus pneumoniae*, *Listeria monocytogenes*, and *Pseudomonas aeruginosa*, NKT cells have demonstrated their crucial role in protection in a CD1d-dependent manner [131]. Notably, also *Bacteroides fragilis* has been found to produce distinctive α-Galactosyl ceramides (α-GalCers), which are loaded onto CD1d molecules and serve as targets for recognition by NKT TCRs [106].

NKT cells actively contribute to immune responses against pathogens by producing pro-inflammatory cytokines, including IFN-γ, TNF-α, and IL-17, and engaging with diverse immune cell types such as macrophages, neutrophils, and cytotoxic T cells [106]. Intriguingly, the nature of APCs can influence the effector functions of NKT cells. While APCs of hematopoietic origin typically induce a robust Th1 signal in NKT cells, other CD1d-expressing cell types like colonic epithelial cells or adipocytes may induce IL-10 production in response to antigen stimulation [106].

4.4. Antigen mimicry

In recent years, research endeavors have yielded substantial insights into the phenomenon of antigen mimicry between commensal bacteria and tumor antigens, which holds the potential to elicit advantageous antitumor responses. These studies have, for the first time, elucidated the mechanistic aspects of bacteria-specific T cell expansion that exhibits cross-reactivity with model tumor neoantigens [132]. Such epitopes have been identified within the genomes of *Bifidobacterium breve* [132] and *Enterococcus hirae* [133], resulting in robust CD8+ T cell-mediated cytotoxic responses. Utilizing HLA peptidomics, Kalaora et al. successfully demonstrated the presence of bacteria-derived peptides presented on HLA-I and HLA-II molecules in melanoma [134]. Notably, 11 of these HLA-I-bound peptides originated from *Fusobacterium nucleatum*, *Staphylococcus aureus*, and *Staphylococcus capitis*. These peptides were capable of being presented by both APCs and melanoma cells, provoking the release of IFN-γ from tumor-infiltrating T cells [134]. Additionally, a recent and significant discovery by Naghavian et al. identified bacterial peptides presented by tumor-derived HLA-II molecules in glioblastoma patients, leading to the clonal expansion of specific tumor-infiltrating CD4+ T cells [135].

The concept of T cell cross-reactivity, often referred to as "molecular mimicry," between antigens derived from tumors and bacteria is currently receiving substantial attention due to its potential to enhance antitumor responses. Specifically, pathogens like *Fusobacterium nucleatum*, typically associated with adverse effects in CRC, may potentially assume a beneficial role.

Hypothesis and Aims of the Study

Hypothesis:

High levels of T cell infiltration in CRC tissues confer favorable outcomes across various disease stages, in contrast to low T cell infiltration, which predicts a dismal prognosis, even in early-stage cases. Importantly, this characteristic is not confined solely to MSI CRCs with high mutational burdens, but is also observed in MSS phenotypes, characterized by poor immunogenicity [47]. Consequently, the factors driving T cell infiltration into tumors and possibly their expansion within the tissue, remain obscure. Mechanisms underlying the correlation between T cell infiltration and improved prognosis also require elucidation.

Given that CRC originates within an environment teeming with trillions of gut bacteria, and that during the oncogenic process changes in gut barrier integrity result in increased permeability of the gut mucosa [82,83], gut bacteria actively translocate into the submucosa and directly engage with tumor cells and resident immune cells. Previous studies from our group have revealed significant correlations between the abundance of specific bacterial species within the tumor tissue and the extent of T cell infiltration [102].

In the present study, our central hypothesis postulates that at least a subset of T cells infiltrating CRC tissues may be reactive to previously identified tumor-infiltrating bacteria, and following activation within tumor tissues may exert anti-tumor effects.

Specific Aims:

We aimed to assess the ability of bacteria associated with T cell infiltration to elicit T cell-mediated anti-tumor immune responses. In particular, we have addressed the following specific points:

1. Testing the ability of previously identified bacterial strains to induce T cell activation;
2. Characterization of the phenotypic and functional profiles of bacteria-reactive T cells;
3. Evaluation of the anti-tumor potential of bacteria-reactive T cells;
4. Identification of bacteria-reactive T cells among TILs in CRC tissue.

Materials and Methods

1. Purification of immune cells from peripheral blood

Peripheral blood mononuclear cells (PBMCs) from healthy donors or CRC patients were isolated by Ficoll-Paque PLUS (Sigma-Aldrich) density gradient centrifugation. To remove red blood cells, which can interfere with our analysis, the samples were treated with a lysis buffer (Biolegend). For some experiments, T cells and monocytes were further isolated by positive selection using magnetic microbeads conjugated with anti-CD3 and anti-CD14 antibodies, respectively, according to the protocol (Miltenyi Biotech). In both cases the purity of isolated cells exceeded 99%.

2. Tissue processing

Clinical specimens were collected from consenting patients undergoing surgical resection for primary CRC at Ente Ospedaliero Cantonale, Ticino. The clinicopathological characteristics of patients involved in this study are listed in Table 1. Tumor and adjacent mucosa tissues were snap-frozen for RNA extraction or treated by enzymatic digestion to obtain single-cell suspensions. Briefly, tissues were minced and digested in phosphate-buffered saline (PBS) supplemented with 2 mg/ml collagenase D (Sigma-Aldrich), 0.2 mg/ml DNAse I (Sigma-Aldrich), and 0.4 mg/ml hyaluronidase type I-S (Sigma-Aldrich) for 1 hour on an orbital shaker at 37°C. Single-cell suspensions were then filtered through 70 μm cell strainers and washed twice with PBS supplemented with 2% human serum. The samples were subsequently used for flow cytometry characterization or preserved at -150°C in human serum with a 10% DMSO final concentration.

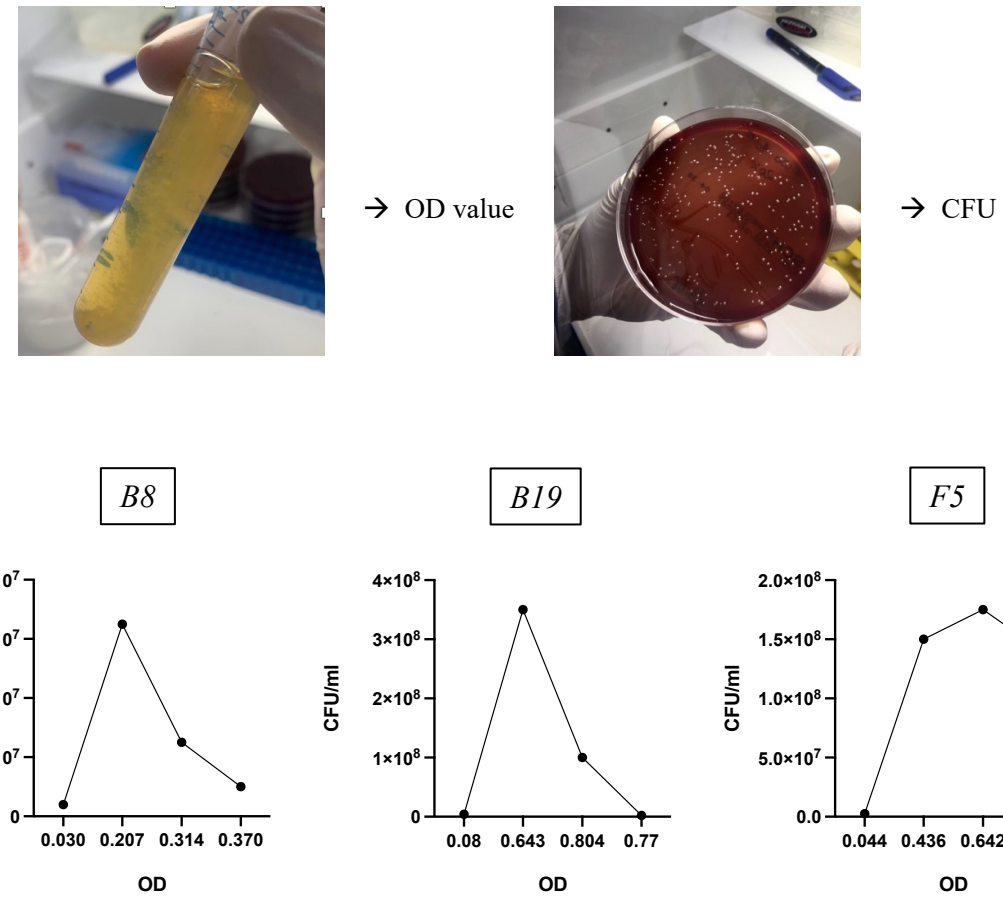
Table 1. Clinicopathological characteristics of CRC patients involved in the study (n=54).

Patients Characteristics *	**Frequency n (%)**
AGE (years, MEDIAN)	77 (44-90)
SEX	
male n (%)	29 (54)
female n (%)	25 (46)
TUMOR LOCATION	
Right colon n (%)	29 (54)
Transverse colon n (%)	5 (9)
Left colon n (%)	4 (7)
Sigmoid colon n (%)	16 (30)
GRADE	
Low (G1 and G2) n (%)	39 (74)
High (G3) n (%)	14 (26)
T STAGE	
T1 n (%)	6 (11)
T2 n (%)	6 (11)
T3 n (%)	30 (57)
T4 n (%)	11 (21)
N STAGE	
N0 n (%)	34 (63)
N1 n (%)	16 (30)
N2 n (%)	4 (7)
AJCC STAGE	
I n (%)	10 (19)
II n (%)	22 (41)
III n (%)	19 (36)
IV n (%)	2 (4)
MICROSATELLITE STABILITY STATUS	
MSS n (%)	33 (62)
MSI n (%)	20 (38)

* Grade, T stage, AJCC stage and MS status data for 1 patient were not available.

3. Bacterial cultures preparation

Fusobacterium OTU5 (F5), *Bacteroides* OTU19 (B19), and *Bacteroides* OTU8 (B8) were purchased from the Leibniz Institute DSMZ. Lyophilized cultures were resuspended in filtered Tryptic Soy Broth (TSB, Sigma-Aldrich) supplemented with vitamin K and hemin (VWR), and cultured under anaerobic conditions (5% CO2, 0.1% O2, 37°C). First inoculums were diluted in fresh TSB medium and seeded on Schaedler Anaerobe Agar plates (Thermo Scientific) during the expansion phase. To obtain growth curves, colony-forming units (CFU) on plates were correlated to the optical density (OD) values of liquid cultures, as assessed by NanoDrop One/One^{C} spectrophotometer.



Frozen stocks of fresh liquid bacterial cultures were prepared following a previously published protocol [136], which were subsequently used for experiments.

4. Co-culture experiments

PBMCs from healthy donors or CRC patients were labelled with carboxyfluorescein succinimidyl ester (CFSE, Invitrogen) for 10 minutes at 37°C. After labeling, cells were washed twice with PBS supplemented with 1% BSA and co-cultured with bacteria at 5:1 bacteria-to-cell ratio in 96-well U bottom plates, in RPMI 1640, supplemented with 5% human serum, 1% GlutaMAX-I, 1% non-essential amino acids, 1% sodium pyruvate, 100 U/mL penicillin and 100 μg/mL streptomycin (all components supplied from ThermoFisher). Bacteria-reactive T cells were identified by flow cytometry after 48h based on the upregulation of activation markers CD69 and CD25 and after 5 days as CD3+CFSE^{int}CD25^{high} cells. Phenotypes of bacteria-reactive T cells were characterized based on the expression of subset-specific markers.

To evaluate T cell responses to bacteria in culture with or without APCs, T cells were enriched using magnetic microbeads, CFSE-labelled as previously described, and co-cultured with bacteria in the presence or absence of autologous monocytes at a 1:2 monocyte-to-T cell ratio. The frequencies of bacteria-reactive T cells were determined by the upregulation of activation markers CD69 and CD25 and by the reduction in CFSE intensity after 5 days of culture.

To determine the requirement for direct T cell and APC contact, monocytes or CD3- fractions containing APCs were pulsed with bacteria overnight and subsequently co-cultured with CFSE-labelled total T cells or sorted CD4+, CD8+, and DN cells for 5 days, in normal conditions (U-bottom plate) or transwell plates (0.4 μm pore size), to prevent direct contact with T cells. T cell proliferation was evaluated by measuring CFSE dilution and CD25 upregulation.

In specific experiments, naïve and total memory T cells were further sorted from bead-isolated T cells by flow cytometry as CD3+CD45RA+CCR7+ (naïve) and CD3+CD45RA+CCR7-, CD3+CCR7+CD45RA-, CD3+CD45RA-CCR7- (memory), respectively. Sorted T cells were CFSE-labelled as previously described, and co-cultured with autologous monocytes and bacteria for 5 days. The proliferation of T cells was determined by measuring CFSE dilution and CD25 upregulation. Statistical significance of variance in percentages of activated and proliferating T cells in response to different bacteria, was tested by Friedman test with multiple

comparisons or paired Wilcoxon signed-rank test, as appropriate, and it is detailed in the Figures.

5. Flow cytometry and cell sorting

For the assessment of surface marker expression, cell cultures or single-cell suspensions from tissues were incubated with PBS containing 1% human serum and 2 mM EDTA and stained with fluorochrome-conjugated marker-specific antibodies for 20 minutes at 4°C. Cell viability was determined using the LIVE/DEAD™ Fixable Yellow Dead Cell Stain Kit (Invitrogen) to exclude non-viable cells from the analysis. For the analysis of intracellular markers, cells were fixed and permeabilized using the FOXP3 Fix/Perm Buffer Set (Biolegend) and stained with the specific antibodies for 30 minutes at room temperature (RT). Following staining, the cells were washed and analyzed using the FACSymphony A5 Cell Analyzer (BD Biosciences). Data were collected using FACSDiva Software. Compensation controls were set up using single-stained compensation beads or cells to correct for spectral overlap. Data analysis was performed using FlowJo software.

Statistical significance of differences in densities of specific immune cell subsets between tumor and adjacent mucosal tissues, as well as between MSS versus MSI tumors, was assessed by paired Wilcoxon test, as detailed in the figure legends.

For the isolation of specific cell subsets based on the expression of the specific markers, sorting was performed using the FACSymphony S6 Cell Sorter (BD Biosciences). The antibodies used in this study are listed in Table 2.

Table 2. List of antibodies used in this study.

Antigen	Clone	Provider
CD103	BER-ACT8	Biolegend
CD11B	ICRF44	BD Pharmingen
CD127	A019D5	Biolegend
CD161	DX12	Biolegend
CD183	1C6	BD Pharmingen
CD19	SJ25C1	BD Pharmingen
CD197	G043H7	Biolegend
CD25	BC96	Biolegend
CD3	SK7, OKT3	Biolegend
CD39	A1	Biolegend
CD4	SK3, RPA-T4	BD Pharmingen
CD45	HI30	BD Pharmingen
CD45RA	HI100	Biolegend
CD45RO	UCHL1	Biolegend
CD69	FN50	Biolegend
CD8	SK1, RPA-T8	BD Pharmingen
DNAM-I	11A8	Biolegend
EPCAM	9C4	Biolegend
FAS	DX2	Biolegend
FOXP3	150D	Biolegend
HLA-DR	G46-6	BD Pharmingen
ICOS	C398.4A	Biolegend
NKG2D	1D11	Biolegend
NKP46	9E2	Biolegend
PD-1	EH12.1	BD Pharmingen
TCRαβ	IP26	BD Pharmingen
TCRγδ	B1	Biolegend
TCRVα7.2	3C10	Biolegend
TIGIT	A15153G	Biolegend
TIM-3	F38-2E2	Biolegend

6. Cytokine detection

Culture supernatants were collected at different time points and stored at -20°C until analysis. Cytokine production was quantified from cell culture supernatants using enzyme-linked immunosorbent assay (ELISA) kits (for detection of IFNγ, Invitrogen) or Legendplex assay kits (Human CD8/NK panel, or T helper cytokine panel, Biolegend), following the manufacturer's instructions. Each sample was analyzed in duplicate or triplicate for accuracy, and the average value was recorded.

7. Bacteria-reactive T cell clone generation

The cloning of T cells was performed using the limiting dilution method, as previously described [137]. In summary, bacteria-reactive T cells were sorted from bacteria-stimulated PBMCs as CD3+CFSE^{int}CD25^{high} cells after a 5-day culture and seeded into 384-well plates at 1cell/well density, with irradiated allogeneic PBMCs (45Gy, 2,5x10^4/well) as feeder cells, in complete medium supplemented with IL-2 500 IU/ml, PHA 1 µg/ml (50 µl/well). After 12 days, single-cell-grown clones were transferred to 96-well U-bottom plates and further expanded in a complete medium supplemented with IL-2. At day 20, their specificity towards the respective bacteria was evaluated. Expanded T cell clones were restimulated with autologous monocytes, pulsed with bacteria overnight, or unpulsed as control, and after 24h cytokine release in culture supernatants was assessed by Legendplex.

8. Cell line culture

LS180 human CRC cell line was purchased from the European Collection of Cell Cultures and maintained in RPMI 1640 medium (GIBCO), supplemented with 10% fetal bovine serum (FBS), GlutaMAX-I, and kanamycin, at 37°C, 5% CO2. Cells were passaged when reaching 80% confluency, upon trypsinization. The medium was refreshed every 2-3 days, and cells were used for experiments at early passages.

9. RT-PCR and qPCR

Total RNA was extracted from preserved primary tumor and adjacent mucosa tissues or CRC cell lines using a NucleoSpin RNA kit (Macherey-Nagel) following the manufacturer's instructions. The RNA quality and quantity were assessed using a NanoDrop One/One^{C} spectrophotometer. For RT-PCR, complementary DNA (cDNA) was synthesized from the isolated RNA using Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT, Invitrogen) and stored at -20°C until qPCR analysis. For the detection of bacteria abundance in the primary CRC tissues, bacteria-specific primers were designed (see Table 3), and gene expression levels were quantified using SYBR Green (Invitrogen). Gene expression levels were normalized to 18s, used as a housekeeping gene. For other genes, expression levels were quantified using TaqMan (Invitrogen) and were normalized to GAPDH. All PCR reactions were performed on the QuantStudio real-time PCR system (Applied Biosystems). Primer sequences used in this study are listed in Table 3.

Table 3. List of primers used in this study.

Gene	Sequence	Provider
Cd4	Hs01058407_m1	Applied Biosystems
Cd8a	Hs00233520_m1	Applied Biosystems
Tbx21	Hs00203436_m1	Applied Biosystems
Cd44	Hs01075861_m1	Applied Biosystems
Eomes	Hs00172872_m1	Applied Biosystems
Gzma	Hs00196206_m1	Applied Biosystems
Gzmb	Hs00188051_m1	Applied Biosystems
Prf1	Hs00169473_m1	Applied Biosystems
Mica	Hs00741286_m1	Applied Biosystems
Micb	Hs00792952_m1	Applied Biosystems
Pvr	Hs00197846_m1	Applied Biosystems
Nectin2	Hs01071562_m1	Applied Biosystems
GAPDH	Hs02786624_g1	Applied Biosystems

18s	AGTCCCTGCCCTTTGTACACA (Forward) GATCCGAGGGCCTCACTAAAC (Reverse)	Microsynth
F5	GCCTCACAGCTAGGGACAAC (Forward) GAGTAAGGGCCGTGTCTCAG (Reverse)	Microsynth
B19	AGGCGGACGCTTAAGTCAGTT (Forward) CGCTACACCACGAATTCCGC (Reverse)	Microsynth
B8	TGAGGTAGGCGGAATTCGTG (Forward) GCCTTCGCAATCGGAGTTCT (Reverse)	Microsynth

10. Cytotoxicity assays

Previously generated T cell clones were plated at a density of $5x10^5$ cells per well in 0.5 ml complete medium and stimulated with anti-CD3/CD28 beads (Miltenyi Biotech) at a 4:1 bead-to-T cell ratio to promote cytokine production. Supernatants were collected after 3-5 days to evaluate their capacity to kill tumor cells.

LS180 cells were seeded in triplicate in 24-well plates at a density of $3,5x10^5$ cells/well in 0.5 ml RPMI 1640 medium. The following day, the culture medium was removed, and the cells were treated with 0.25 ml of fresh medium (control), or culture supernatant derived from stimulated T cell clones. Following a 24-hour treatment period, cells were harvested and stained with Live/Dead, and sequentially incubated with Annexin V, according to the protocol (both from Invitrogen). The percentage of CRC cell death was determined by flow cytometry and calculated as the sum of the proportions of Annexin V+ and Annexin V+Live/Dead+ cells.

To assess the direct T cell-mediated tumor killing, LS180 cells were seeded in a 48-well plate at a density of $5x10^4$ cells/well in 200 µl RPMI 1640 medium, and following overnight incubation, bacteria-reactive DN T cells were added at 5:1 effector-to-target ratio. After overnight incubation, CRC cells were harvested and stained with Live/Dead, and sequentially incubated with Annexin V, according to the protocol (both from Invitrogen). The percentage of CRC cell death was determined by flow cytometry and calculated as the sum of the proportions of Annexin V+ and Annexin V+Live/Dead+ cells. The statistical significance of differences detected in cell death rates was assessed by Kruskal-Wallis test.

11. Single-cell RNA sequencing

Bacteria-reactive T cells were induced as detailed above, from PBMCs of five patients with MSS-CRCs, and sorted as live CD3+CFSE^{int}CD25^{high} cells on day 5 of culture. Concurrently, TILs were sorted as live CD3+ cells from the autologous tumor cell suspensions. Sorted cells were fixed using a fixation buffer provided in the kit (PARSE Biosciences), according to the instruction manual, and stored in a cell buffer (PARSE Biosciences) at -80°C until sequencing.

After thawing, cells were counted to assess the final concentration. Three rounds of sample barcoding were carried out as outlined in the protocol from PARSE Bioscience. Quantification of cDNA was performed using the Qubit fluorometer and the Bioanalyzer system.

The preparation of Whole Transcriptome (WT) libraries and TCR Library was conducted according to the protocol from PARSE Bioscience. Evercode sequencing libraries were diluted at 4nM, then pooled and diluted to 800 pM with RSB buffer + tween according to the NextSeq 2000 (Illumina) instructions.

The sequencing was performed on NextSeq 2000 (Illumina) system, using 66-8-8-86 cycles, and NextSeq 1000/2000 P2 Reagents (100 Cycles) v3 (Illumina – catalog #20046811) for WT, and 216-8-8-86 cycles, and NextSeq 1000/2000 P1 Reagents (300 Cycles) (Illumina - catalog #20050264) for TCR Library.

For the data analysis, sequencing reads were mapped to the GRCh38 with the ParseBioscience processing pipeline (v0.9.6p), and the downstream steps were performed using the R package Seurat (v4.0.2). Cells with fewer than 300 or more than 8000 detected unique genes, or with more than 10% mitochondrial reads were excluded from the analysis. The resulting gene–cell matrix was normalized and scaled using Seurat's NormalizeData and ScaleData functions. The dimensionality reduction was performed using the principal component analysis with Seurat's RunPCA function. Cell clustering was performed using the FindNeighbors (30 dimensions of reduction) and FindClusters (resolution = 1.4) functions; for visualizing clusters, RunUMAP (30 dimensions) was used to project the cells into a 2D space.

To determine differentially expressed genes between clusters a Limma tests using the FindAllMarkers function (minimal fraction of 25% and log-transformed fold-change threshold of 0.25) was performed. The identity of cell clusters was determined by a mixed supervised and unsupervised approach. Top differentially expressed transcripts were cross-referenced with previous studies reporting on single-cell transcriptomes. Moreover, the identification of cell types was enhanced through integration with Celldex, a powerful automatic tool for cell type annotation based on known marker genes and signature expression profiles.

To further explore T cell states, data were further projected into a published reference atlas of T cell states, using the functionalities provided by the authors. (https://github.com/carmonalab/ProjecTILs).

The AddModuleScore function was employed to score the expression of the genes assigned to the TCRγδ T cells as reported by Pizzolato et al.[138].

Ambiguous cells with expression of distinct lineage markers were deemed to be likely multiplets and were excluded from the analysis.

Results

1. CRC-associated bacteria induce the expansion of peripheral blood T cells in healthy donors and CRC patients.

For this study, we selected three CRC-associated bacterial strains, belonging to genera mostly prevalent in human CRC tissues, as identified within the publicly available database (The Cancer Microbiome Atlas, Figure 1A), whose abundance was correlated with the extent of T cell infiltration in tumor samples, as previously reported [102] (Figure 1B). The presence of selected strains - hereafter referred to as *Fusobacterium* OTU5 (F5), *Bacteroides* OTU19 (B19), and *Bacteroides* OTU8 (B8), – in the tumor tissues of the patients in our cohort was confirmed by specific PCR, as shown in Figure 1C. We initially evaluated the capacity of the selected strains to activate peripheral blood T cells derived from healthy donors and CRC patients. T cell activation was assessed based on the expression of activation marker CD69 alone, and in combination with CD25, following 48 hours or 5 days of co-culture with each bacterial strain (Figure 2A). A significant increase in activation markers as compared to the unstimulated control was detected in healthy donors after both 48 hours and 5 days in response to all three bacteria, although with individual variability (Figure 2B,C). Moreover, after a five-day co-culture most of the activated T cells underwent proliferation as indicated by CFSE dilution, demonstrating their full commitment (Figure 2D).

Parallel assessments of T cells from CRC patients revealed reactivity to bacteria comparable to that of healthy donors, at both time points (Figure 2E,F). Remarkably, certain patient samples exhibited substantial T cell activation and proliferation, up to 25% and 20% of the total CD3+ cell population, respectively (Figure 2F,G).

Figure 1.

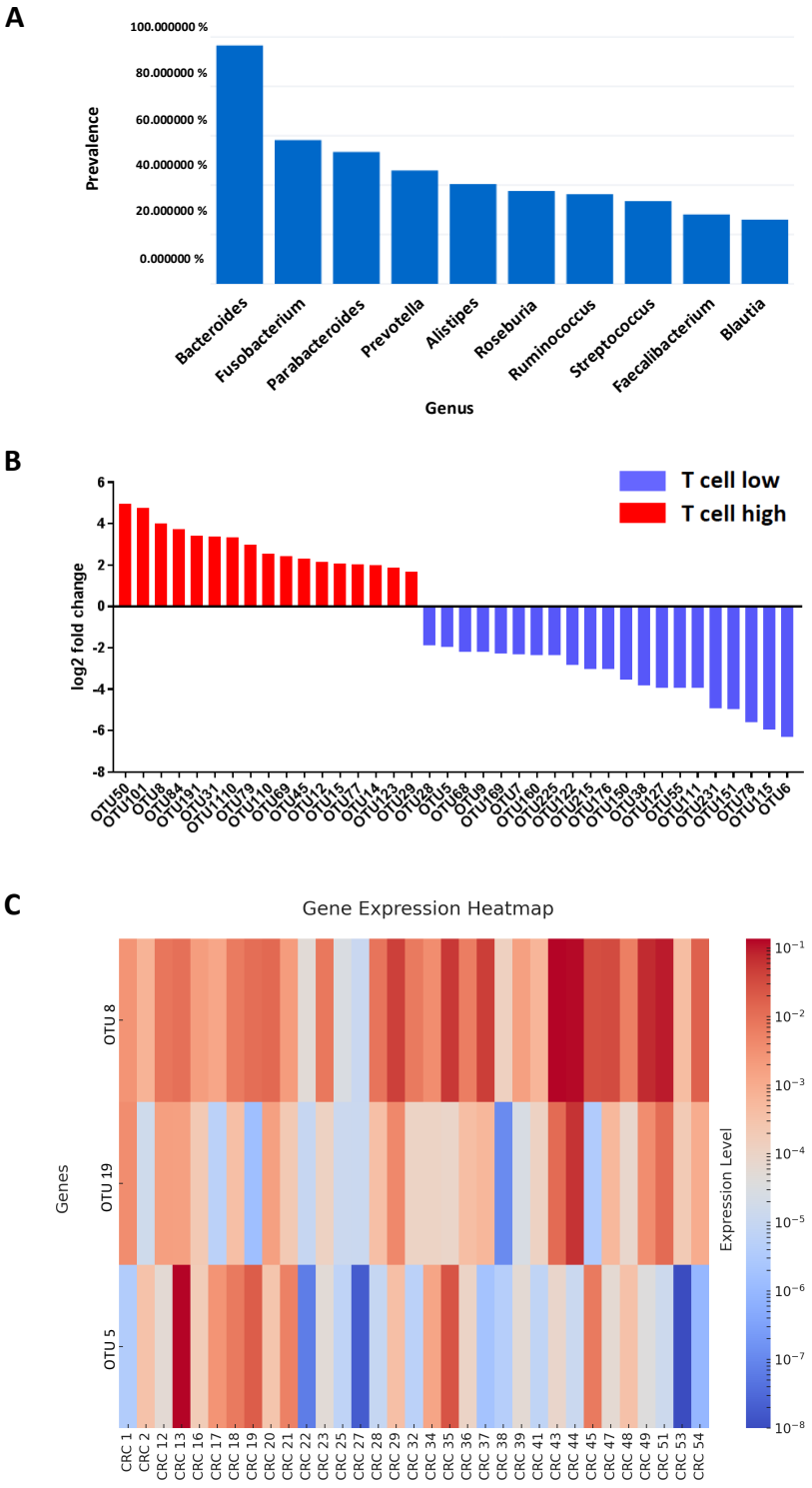


Figure 1: CRC-associated bacteria detected in primary CRC tissues. (A) Prevalence of bacterial genera in human CRC tissues (n=146) as reported in https://tcma.pratt.duke.edu; (B) Correlation between the relative abundance of specific bacterial species (operational taxonomic unit, OTU) and T cell infiltration in human primary CRC tissues. Graph refers to OTUs enriched in highly versus poorly infiltrated CRC, as reported in Cremonesi et al.[102]; (C) Gene expression heatmap displaying the relative abundance of the three bacterial strains in primary tumor tissues from CRC patients (n=33), as detected by specific PCR.

Figure 2.

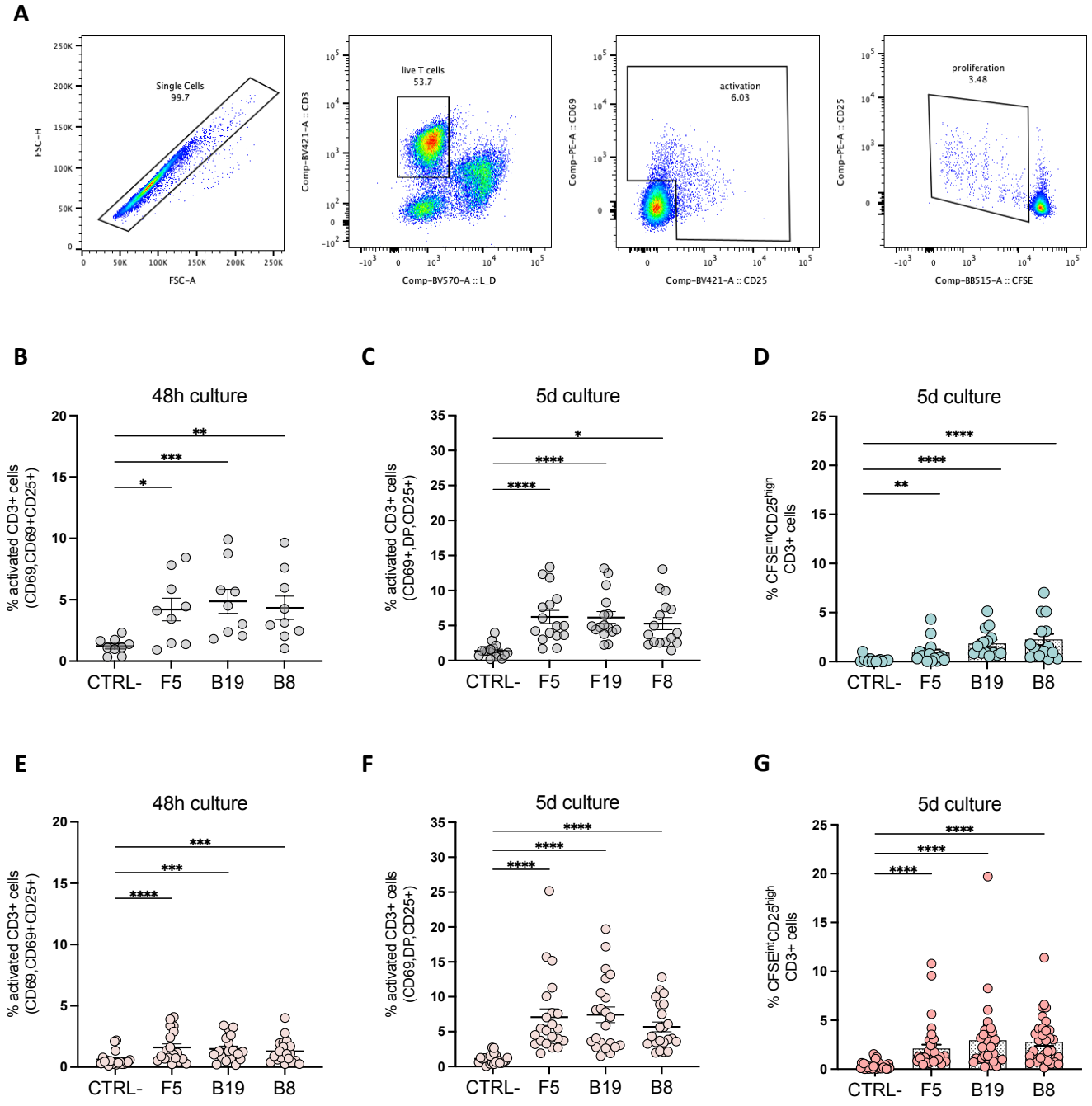


Figure 2: Peripheral blood T cell expansion induced by CRC-associated bacteria in healthy donors and CRC patients. PBMCs from healthy donors (A-D) and CRC patients (E-G) were labeled with CFSE and co-cultured with bacterial preparations for 48h, or 5 days to evaluate T cell responses. Bacteria:T cell ratio was 5:1. (A) Gating strategy and representative dot plots illustrating T cell activation and proliferation after 5 days of culture; Bacteria-reactive T cells were identified as CD3+CD69+CD25+/- cells upon 48h (B,E), or collectively as CD3+CD69+, CD3+CD69+CD25+, CD3+CD25+ cells upon 5 days of culture (C,F). Proliferating T cells were identified as CD3+CFSE^{int}CD25^{high} cells upon 5 days of culture (D,G). Each data point represents a single donor. Mean values ± SEM are shown. Statistical significance was determined using the Friedman test, multiple comparisons (****p<0.0001). DP=double positive, CD69+CD25+.

2. CRC-associated bacteria preferentially induce the expansion of TCRαβ CD4-CD8- memory T cells.

To characterize the functional phenotypes of T cells expanded following bacterial stimulation, we first isolated naïve and memory T cell populations and compared their responses to bacterial stimulation. This approach allowed us to confirm that the bacteria-induced expansion was predominantly derived from the memory T cell compartment, rather than from naïve T cells (Figure 3A).

We then characterized bacteria-reactive T cells by immunophenotyping. Unexpectedly, among healthy donors and in response to all three types of bacterial stimulation, the predominant population of proliferating cells was identified as double-negative (DN) CD4-CD8- T cells, accounting for up to 92% of proliferating cells, although the degree of their expansion varied across individual donors (Figure 3B). In contrast, in CRC patients, CD4+ and DN T cells expanded to almost equal extents (up to 97 and 89%, respectively) whereas CD8+ cells accounted for up to 55%, as depicted in Figure 3C. Regulatory T cells, identified as CD3+FoxP3+, constituted only a minor fraction of the proliferating cells in all donors, across all conditions (Figure 3B,C).

DN T cells usually account for only a small percentage of PBMCs [139]. We therefore grew curious to further elucidate the nature of bacteria-reactive DN T cells. By co-culturing sorted CD4+, CD8+, and DN cells in the presence of autologous APCs, we could confirm a higher response of DN T cell subset to bacteria compared to CD4+ or CD8+ cells (Figure 3D). Furthermore, this approach ruled out that DN T might derive from single-positive cells potentially downregulating CD4 or CD8 markers following bacteria stimulation. Consistently with findings on total CD3+ cells, expanded DN T cells were characterized as effector memory T cells (Figure 3E), indicating a robust, adaptive immune response to these commensals. We further sought to determine if these cells belong to the known unconventional T cell subsets. As Figure 3F illustrates, a notable proportion of bacteria-reactive DN T cells were indeed TCRγδ+ or MAIT cells, whereas no NKT-like cells were detected.

Figure 3.

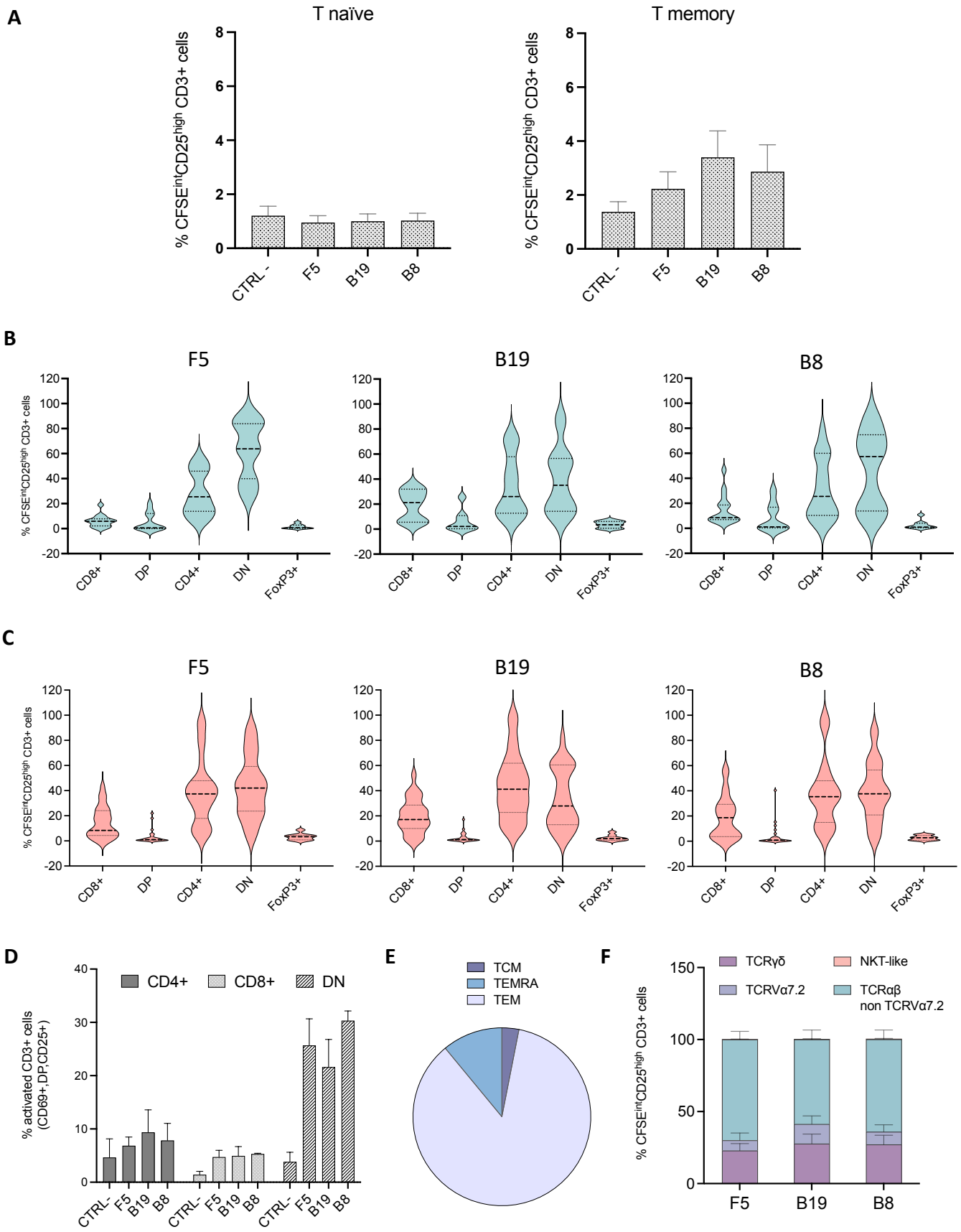

A
T naïve
T memory
% CFSE^intCD25^high CD3+ cells
CTRL -
F5
B19
B8
B
F5
B19
B8
CD8+
DP
CD4+
DN
FoxP3+
C
D
CD4+
CD8+
DN
% activated CD3+ cells (CD69+,DP,CD25+)
E
TCM
TEMRA
TEM
F
TCRγδ
NKT-like
TCRVα7.2
TCRαβ non TCRVα7.2

Figure 3: Characterization of bacteria-reactive T cell phenotypes in peripheral blood. (A) Naïve and memory T cells, sorted from healthy donors (n=6), were labeled with CFSE and co-cultured with autologous monocytes in the presence of bacteria. After 5 days, frequencies of bacteria-reactive T cells were determined; (B,C) PBMCs from HD and CRC patients were stimulated with bacteria as described in Figure 2. On day 5, expanding cells, identified as CD3+CFSE^{int}CD25^{high} were further characterized by flow cytometry based on the expression of subset-specific markers. Phenotypic distribution of bacteria-reactive T cells from healthy donors (B, n=15) and CRC patients (C, n=25) is depicted; (D) Pre-sorted CD4+, CD8+, and DN T cell populations from PBMCs of healthy donors (n=3), were co-cultured with autologous monocytes and bacteria for 5 days. Proportions of bacteria-reactive T cells are depicted; (E) Classification of DN T cells isolated from healthy donors' peripheral blood into central memory (TCM; CD3+CD45RA-CD45RO+CCR7+), terminally-differentiated effector memory (TEMRA; CD3+CCR7-CD127-CD45RA+), and effector memory (TEM; CD3+CD45RA-CD45RO+CCR7-) subsets, using marker expression profiles; (F) Identification of MAIT, γδ T cells, and non-MAIT αβ T cells among bacteria-expanded DN T cells from peripheral blood of CRC patients (n=20), based on specific TCR expression, as assessed by flow cytometry upon counterstaining with TCR-specific antibodies. The graph shows mean values ± SEM. DP=double positive, CD4+CD8+; DN=double negative, CD4-CD8-.

3. Bacteria-reactive T cells require direct contact with APCs for their functional response.

To elucidate whether the induction of T cell responses to CRC-associated bacteria requires the involvement of APCs, T cells were cultured with bacteria in the presence or absence of autologous monocytes. Our findings revealed a clear dependency on APCs for T cell expansion, as evidenced by the drastic reduction of T cell activation (Figure 4A) and proliferation (Figure 4B) in the absence of monocytes after a five-day culture. Furthermore, the production of cytokines, such as IFNγ, was exclusively observed in co-cultures containing APCs (Figure 4C). Collectively, this indicates that CRC-associated bacterial stimuli elicit T cell responses through interaction with APCs.

However, bacterial stimulation of APCs induces their rapid release of proinflammatory cytokines, such as IL-6, IL-1β, IL-18, and IL-12, which could potentially induce non-cognate T cell responses[140]. Therefore, we further investigated the necessity for physical interaction between T cells and APCs. Notably, T cell expansion was only observed when cell-to-cell contact with APCs was permitted (Figure 4D). Consistently, we found that DN T cells alone did not respond directly to bacterial exposure (Figure 4E). Instead, DN T cells, alongside CD4+ and CD8+ counterparts, require direct contact with APCs to undergo functional expansion (Figure 4F). This outcome supports the conclusion that bacteria-induced T cell activation is not merely driven by cytokine signaling from APCs but requires direct contact, implying antigen presentation as the primary mechanism.

To further investigate whether bacteria-reactive T cells are restricted by classical MHC-I and MHC-II, or non-classical monomorphic MHC molecules, we have first tested the capacity of MHC-I-, MHC-II-, MR1-, CD1b-, CD1c-, CD1d- blocking antibodies to reduce activation and/or proliferation of bacteria-reactive T cells. However, we failed to demonstrate a consistent impairment in bacteria-driven T cell responses (data not shown). Currently, we are testing the effect of MHC-deficient APCs, generated through CRISPR/Cas9-mediated genome editing, on the induction of bacteria-reactive T cells.

Figure 4.

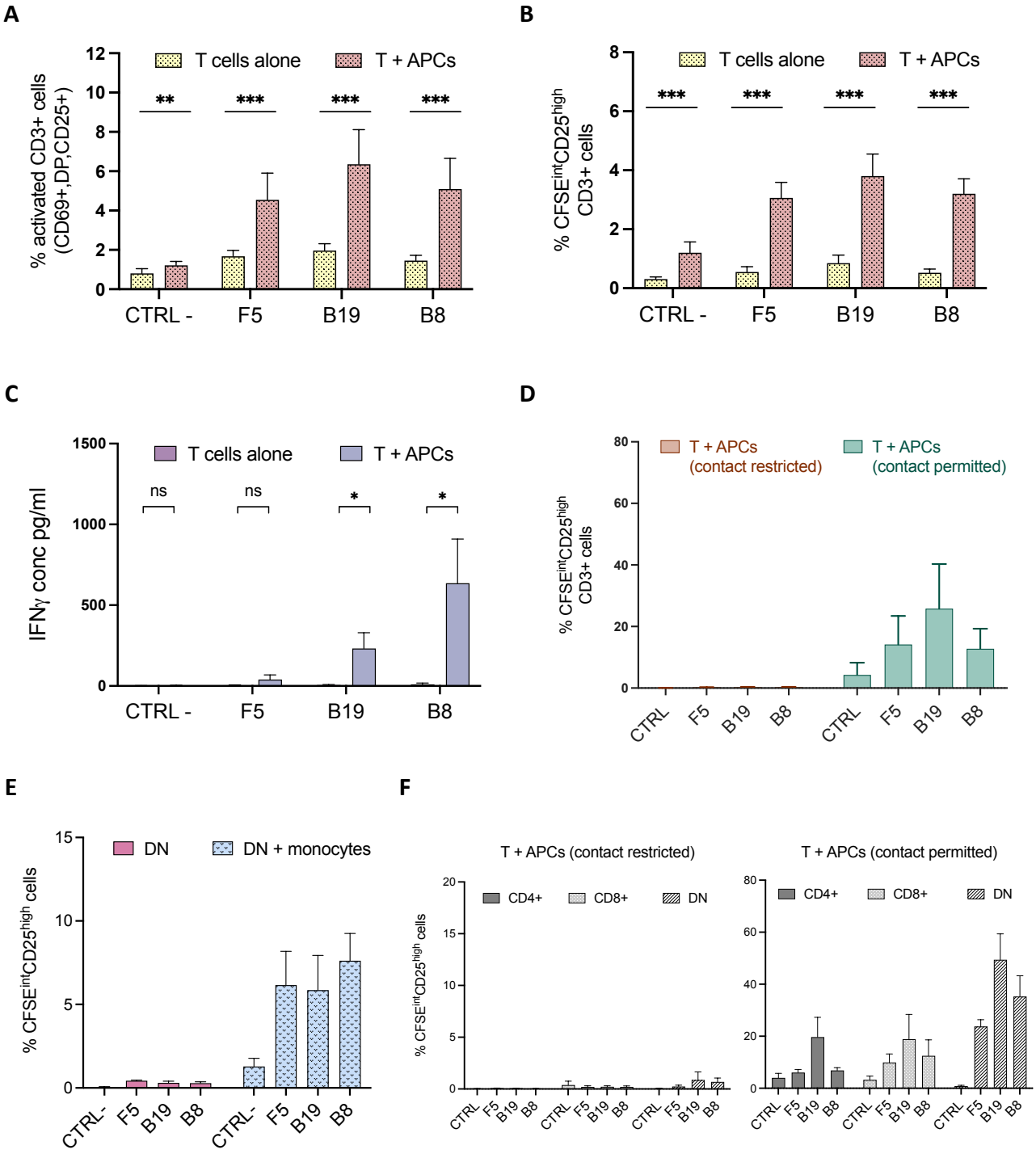


Figure 4: T cells require the presence of APCs for proliferation and cytokine production. Total CD3+ cells were isolated from PBMCs of healthy donors (n=12) and co-cultured with bacteria in the presence or absence of autologous monocytes at a 1:2 monocyte-to-T cell ratio. The frequencies of activated (A) and proliferating (B) T cells after a 5-day culture period are shown; (C) Five-day culture supernatants were assessed for IFNγ production; (D) CD3- fractions containing APCs were stimulated with bacteria overnight and subsequently co-cultured with CD3+ cells for 5 days, in conditions permitting (normal plates) or

restricting (transwell plate, pore size 0.4 μm) the direct APC-T cell contact. The frequencies of proliferating T cells are depicted (E) Proliferation rates of DN T cells from healthy donors (n=3), with or without co-culture with autologous monocytes; (F) Comparison of bacteria-induced expansion of pre-sorted CD4+, CD8+, and DN T cells from healthy donors (n=3), co-cultured with APCs pulsed overnight with bacteria, under conditions permitting (normal plates) or restricting (transwell plate, pore size 0.4 μm) the direct APC-T cell contact; The frequency of proliferating T cells was assessed after 5-days culture. Mean values ± SEM of three independent experiments are shown. Statistical significance was evaluated using the Wilcoxon test.

4. DN T cells exhibit innate-like cytotoxic features.

Our findings prompted us to further characterize the functional profiles of the highly expanded DN T cell subset. Comparative gene expression analysis between DN T cells and CD4+ and CD8+ cells revealed intermediate expression of T-bet, CD44, EOMES, granzyme A, granzyme B and perforin genes. Indeed, except for CD44, all genes were more highly expressed in DN T cells compared to CD4+ cells and less so than in CD8+ cells (Figure 5A). In addition, a significant fraction of the expanded DN T cell population expressed CD39, while a smaller subset expressed PD-1 (Figure 5B).

Further, assessment of culture supernatants derived from DN T cell cultures, as well as from expanded DN T cell clones, revealed a substantial secretion of cytotoxic molecules including TNFα, IFNγ, granzyme A, and granzyme B, with lower levels of granulysin (Figure 5C,D), in particular in cultures stimulated by B8. Conversely, IL-10 levels were low in all cultures (Figure 5C,D). The detected gene expression profiles, together with the panel of cytokines released, cumulatively indicated a fully functional state and cytotoxic nature of DN T cells, underscoring their potential role in immune defense mechanisms, particularly in anti-tumor responses.

In the context of anti-tumor immunity, DN T cells have been recognized as part of both innate and adaptive immune responses[139], and have shown potent cytotoxicity against various cancers, including acute myeloid leukemia and solid tumors like non-small cell lung and pancreatic cancers[141–145]. Notably, the mechanism behind DN T cell-mediated tumor lysis appears to involve the innate receptors NKG2D and DNAM-1, known to recognize a range of stress-induced ligands present on target cells, which leads to a positive feedback loop involving the release of IFNγ, TNFα, and perforin[142,145]. In line with these findings, our results showed that bacteria-stimulated DN T cells progressively upregulated NKG2D and DNAM-1, in contrast to non-reactive T cells, which exhibited no expression of these receptors (Figure 5E,F).

Figure 5.

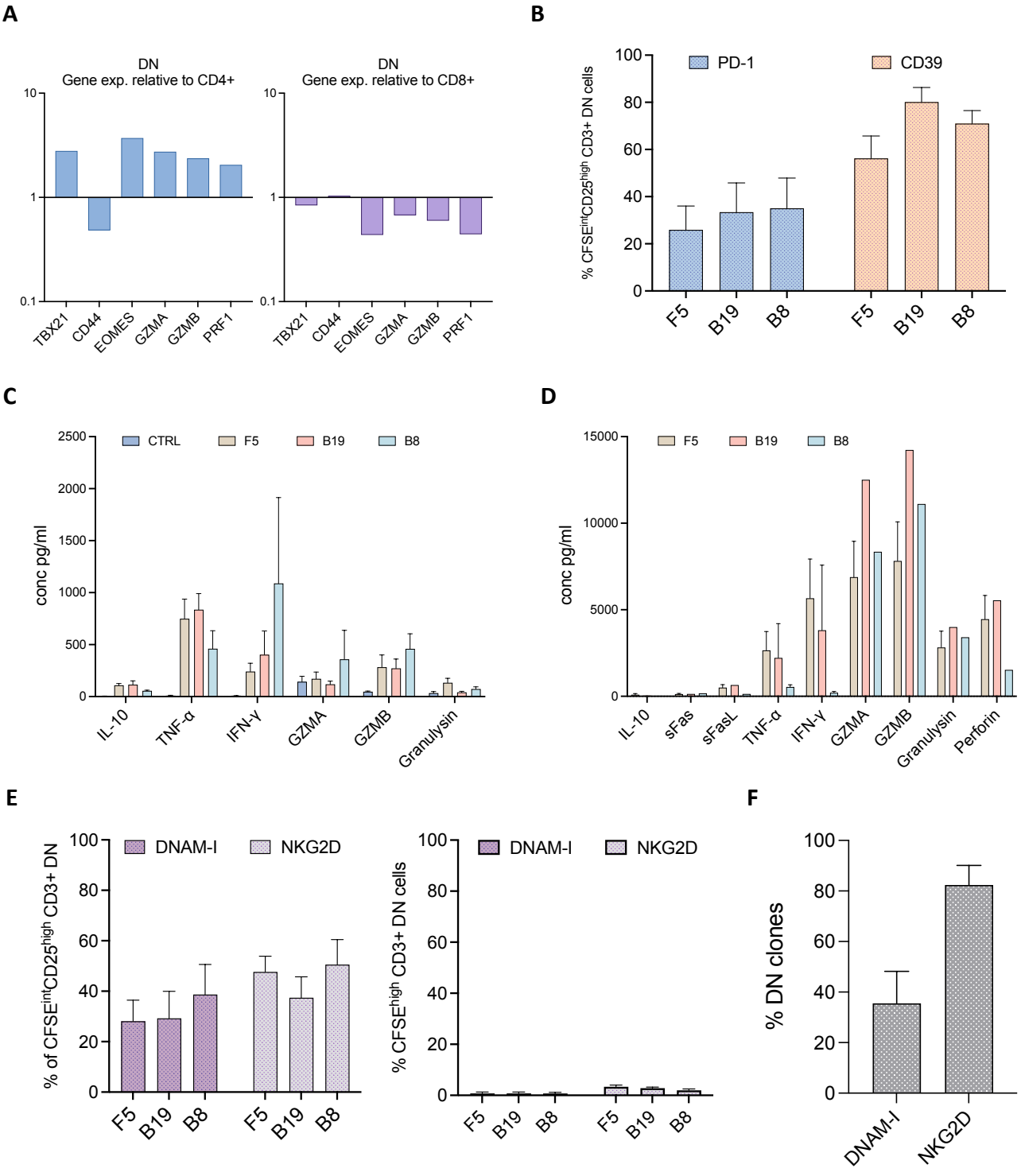
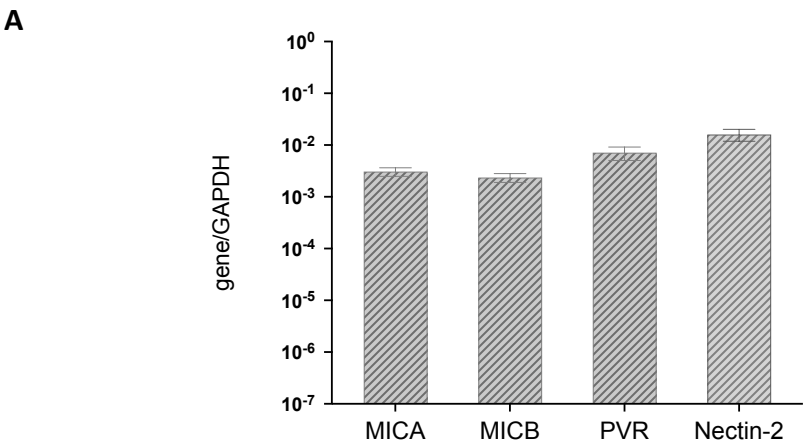

A
DN
Gene exp. relative to CD4+
DN
Gene exp. relative to CD8+
10
1
0.1
TBX21
CD44
EOMES
GZMA
GZMB
PRF1
B
PD-1
CD39
% CFSE^int CD25^high CD3+ DN cells
F5
B19
B8
C
CTRL
F5
B19
B8
conc pg/ml
IL-10
TNF-α
IFN-γ
GZMA
GZMB
Granulysin
D
F5
B19
B8
conc pg/ml
IL-10
sFas
sFasL
TNF-α
IFN-γ
GZMA
GZMB
Granulysin
Perforin
E
DNAM-I
NKG2D
% of CFSE^int CD25^high CD3+ DN
% CFSE^high CD3+ DN cells
F
% DN clones

Figure 5: Functional characterization of DN T cells induced by CRC-associated bacteria. (A) Relative gene expression analysis in DN T cells from healthy donor (n=1) compared to CD4+ and CD8+ T cells; (B) Proportions of CD39+ and PD1+ bacteria-responsive DN T cells in CRC patients (n=7); (C) Cytokine secretion profiles in 5-day culture supernatants of DN T cells from healthy donors (n=4) co-cultured with autologous APCs; (D) Single bacteria-reactive DN T clones (n=17) from CRC patients were stimulated with anti-CD3/CD28 beads and expanded for 3-5 days. Cytokine production in culture supernatants was quantified using a multiplex assay; Expression levels of DNAM-I and NKG2D in proliferating vs. non-proliferating DN T cell compartments from CRC patients (E, n=8) or in expanding DN T clones (F, n=5). Each donor was analyzed in four technical replicates, and the mean value ± SEM was plotted.

5. DN T cells induce CRC cell killing.

Encouraged by the previous findings, we further explored the anti-tumor potential of bacteria-expanded DN T cells. After verifying the expression of ligands for NKG2D (MICA, MICB) and DNAM-I (PVR, Nectin-2) on CRC cells (Figure 6A), we tested the capacity of bacteria-reactive DN T cells to kill CRC cells. Indeed, overnight incubation of CRC with DN T cells, resulted in remarkable tumor cell killing (up to 40%), as detected based on Annexin V /Live-Dead staining (Figure 6B,C). Furthermore, supernatants derived from stimulated DN T cell clones also induced apoptosis in CRC cell lines, as illustrated in Figure 6D.

Figure 6.



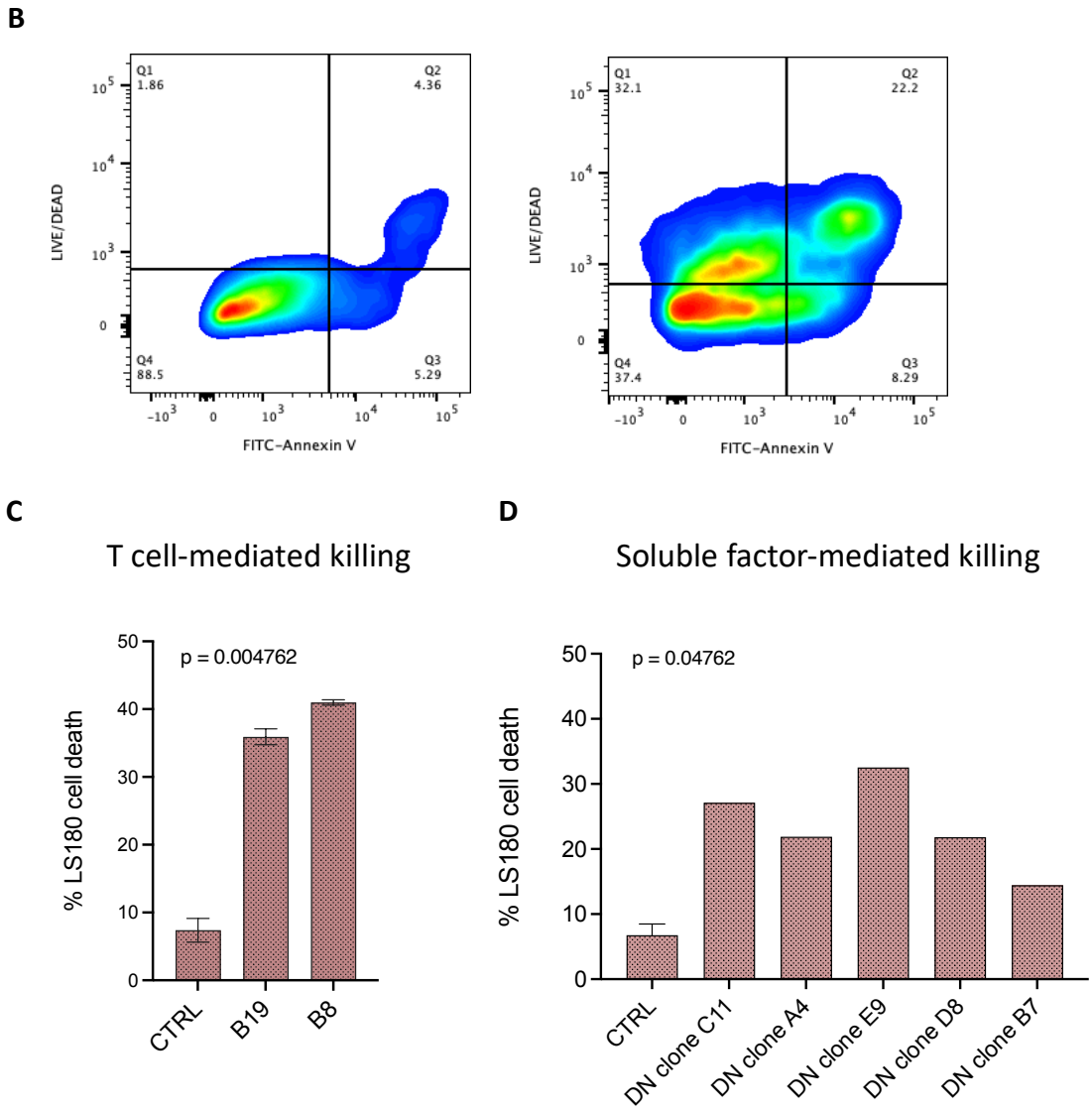


Figure 6: Anti-tumor potential of bacteria-expanded DN T cells. (A) Relative gene expression of NKG2D (MICA, MICB) and DNAM-I (PVR, Nectin-2) ligands on CRC cell line LS180; (B) Representative plots of DN T cell-mediated CRC killing - left: control, right: treatment with supernatant; (C) Expanded DN T cells were co-cultured overnight with LS180 CRC cell line at 5:1 effector:target ratio. Cell death was determined by flow cytometry, by summing the proportions of Annexin V+ and Annexin V+Live/Dead+ cells; (D) Single bacteria-reactive DN T clones (n=17) from CRC patients were stimulated with anti-CD3/CD28 beads and expanded for 3-5 days. Culture supernatants of selected DN T clones were used to assess cytotoxicity against LS180 CRC cell line upon 24h treatment. Cell death was determined by summing the proportions of Annexin V+ and Annexin V+Live/Dead+ cells. Data are presented as mean values ± SD; Statistical significance was evaluated using the Kruskal-Wallis test.

6. DN T cells are detected among TILs in fresh CRC tissue.

The above findings indicated that bacteria abundant within CRC tissues are capable of expanding DN T cells endowed with cytotoxic potential. This led us to investigate whether cytotoxic DN T cells are included within TILs.

We analyzed T cell infiltrates in fresh primary CRC tissues and adjacent tumor-free mucosa specimens obtained from patients undergoing surgical removal. In most samples, the tumor tissue consistently showed a higher degree of CD45+ and CD3+ cell infiltration compared to the adjacent mucosal tissue (Figure 7A). Interestingly, the frequency of infiltrating T cells did not significantly differ between the MSS and MSI CRCs (Figure 7B).

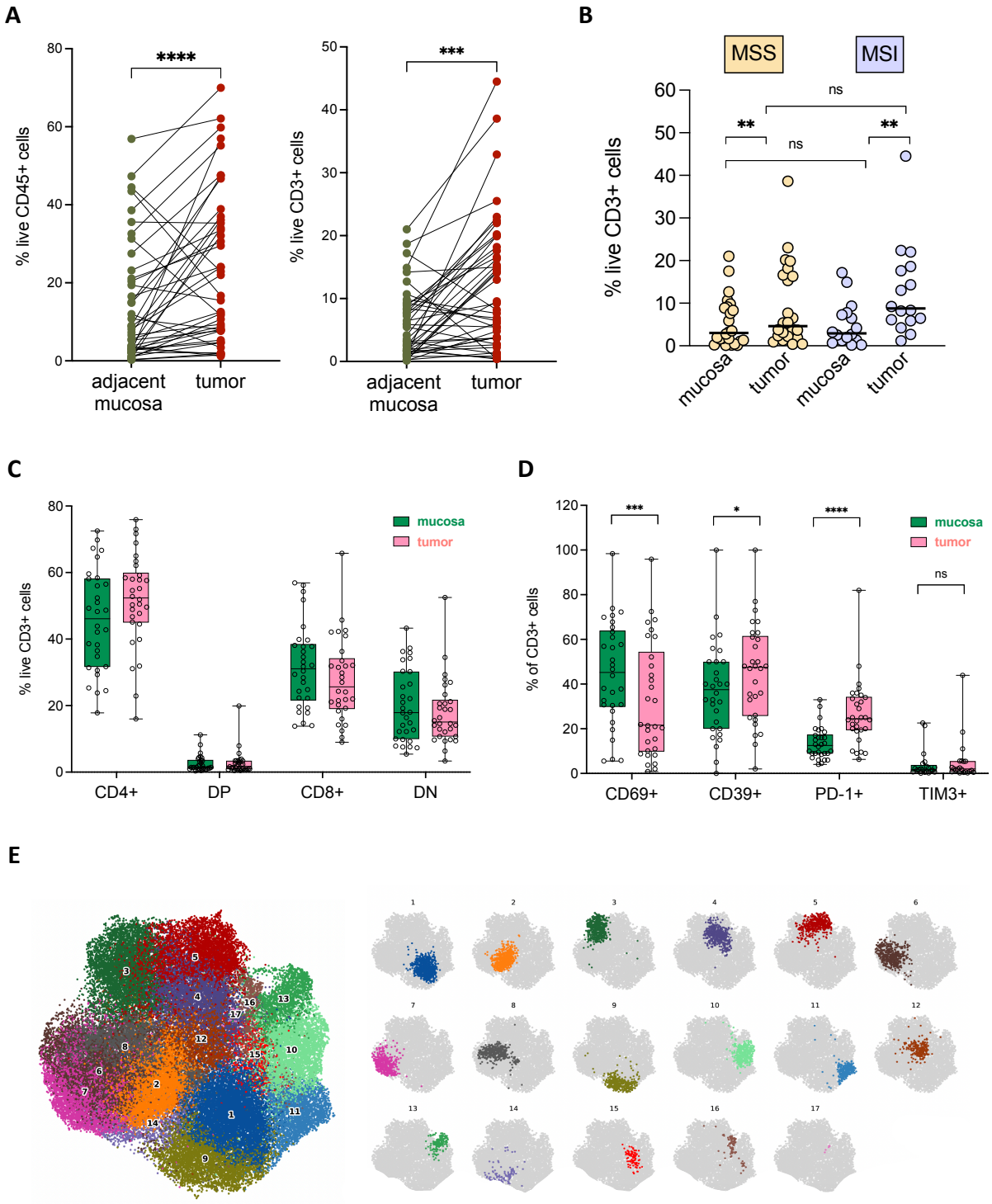
TILs mostly included CD4+ T cells and to lower extents, CD8+ T cells. Remarkably, DN T cells were also largely present in TILs, and in certain individuals, they represented the most abundant T cell subset (Figure 7C). In tumor-free areas, infiltrating T cells showed significantly higher expression of CD69 compared to those in the tumor (Figure 7D). In contrast, tumor-infiltrating T cells exhibited significantly higher levels of CD39 and PD-1 (Figure 7D), indicative of T cell exhaustion. This finding aligns with the known adaptive response of tumor-infiltrating lymphocytes to the immunosuppressive tumor environment.

To obtain a detailed understanding of the T cell landscape in CRC, we employed unsupervised clustering analysis of high-dimensional cytometry data using the publicly available CRUSTY method. This approach enabled us to identify 17 distinct clusters of tumor-infiltrating T cells, each with unique properties and varying frequencies among individual CRCs (Figure 7E). The prevalence of each cluster among individuals, as well as their distribution in MSS and MSI tumors, is shown in Figure 7F.

Further analysis allowed us to determine the specific marker expression for each cluster, providing a precise phenotypic characterization of the cells within each group, as depicted in Figure 7G. Several clusters including DN T cells were identified. Notably, Cluster 12, included DN T NKT-like cells characterized by high expression of exhaustion markers (Figure 7G). Clusters 8 and 16 corresponded to two distinct states of DN TCRγδ cells, whereas clusters 6, 7, and 14 were identified as DN MAIT cells (Figure 7G).

Furthermore, C2, identified as one of the most abundant T cell clusters in MSS tumors, consisted of CD69-expressing non-MAIT, non-TCRγδ, DN T cells. This raises the intriguing possibility that these DN T cells could be reactive to gut commensals and contribute to effective anti-tumor immunity.

Figure 7.



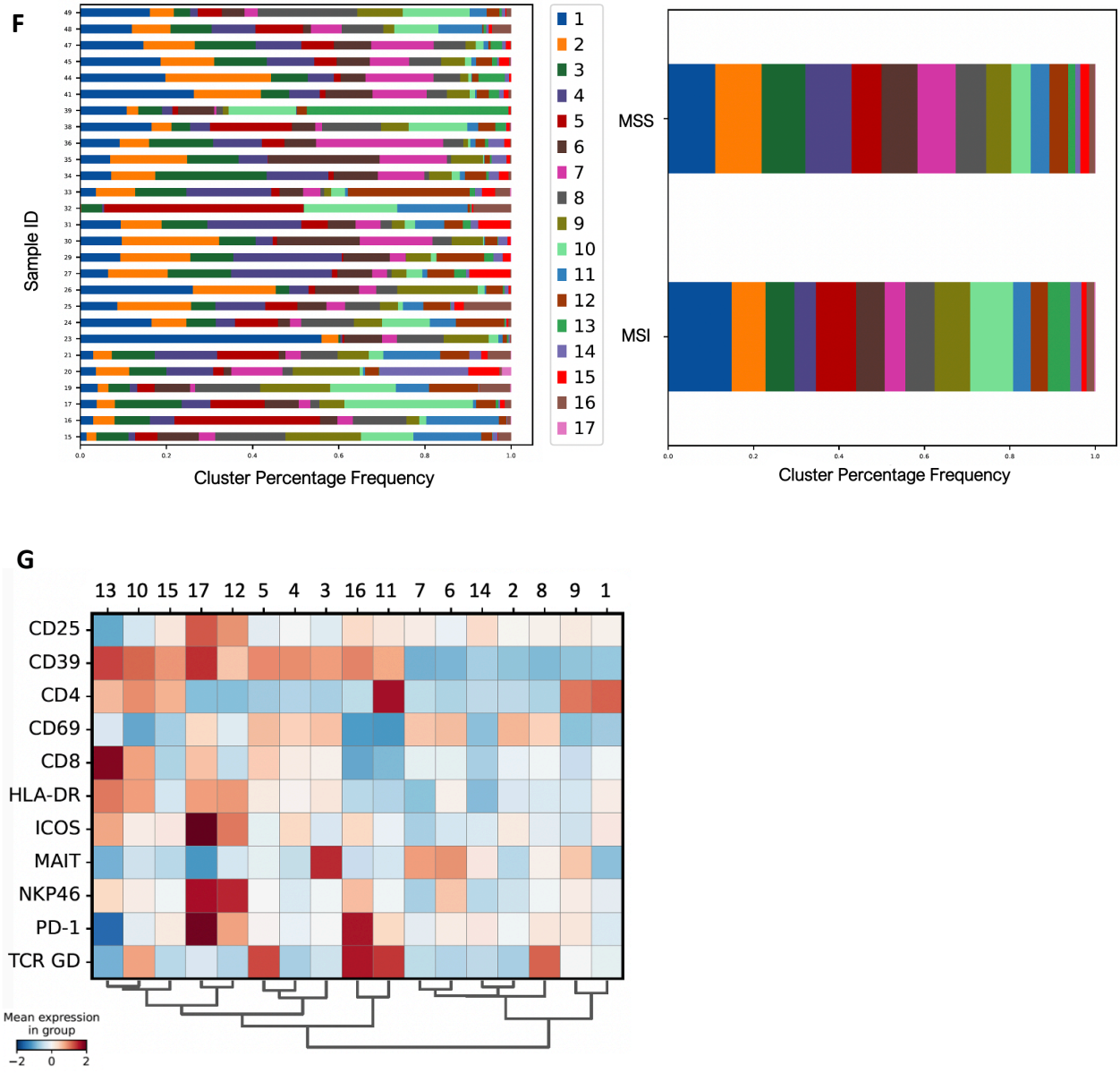


Figure 7: Assessment and analysis of T cell infiltrates in fresh CRC tissue. (A) Single-cell suspensions prepared from fresh primary tumor and adjacent mucosa tissues of CRC patients (n=43) were analyzed for the prevalence of CD45+ and CD3+ cells; (B) Comparative analysis of CD3+ cell infiltration in tumor versus adjacent mucosa and between MSS and MSI tumors; (C) Identification and quantification of major T cell subsets within TILs; (D) Assessment of T cell activation and exhaustion markers in TILs; (E) Unsupervised cluster analysis of concatenated TIL data (2,500 TILs per CRC patient) from 27 CRC patients, performed using the CRUSTY web tool, with each of the 17 identified clusters represented by a distinct color; (F) Distribution patterns of identified clusters within tumor samples of individual patients (left) and comparison between MSS and MSI tumor types (right); (G) Detailed phenotypic characterization of TILs from each of the 17 clusters, based on specific marker expression profiles. Each data point represents an individual patient. Statistical analysis was conducted using the Wilcoxon test. TCR GD=TCRγδ.

7. TILs include DN T cells sharing transcriptomic profiles with bacteria-reactive T cells.

Our final aim of the study was to assess the reactivity of freshly isolated tumor-infiltrating T cells to CRC-associated bacteria. However, determining this reactivity using our established experimental approach proved challenging. The markers of T cell activation in freshly isolated cells were already highly expressed, and by day 5, we observed proliferation in all experimental conditions, including the unstimulated control. This finding was somewhat expected, considering the array of stimuli T cells are exposed to within the tumor microenvironment. It underscores the inherently active state of these cells in situ, reflecting the dynamic and complex nature of the tumor immune landscape.

To further gain insights into the potential presence of T cells reactive to CRC-associated bacteria among TILs, we embarked on a comprehensive single-cell analysis. First, we performed a comparative analysis of the single-cell whole transcriptome of bacteria-reactive T cells expanded from the peripheral blood of five CRC patients (antigens F5, B19, B8) against those of their autologous TILs.

TILs ('Tumor') formed a unique cluster, distinctly separate from the peripheral blood T cells exposed to the bacterial antigens (F5, B19, B8), which implies a unique expression profile specific to the tumor environment. Conversely, the clustering of peripheral blood T cells reflects the broadly similar transcriptional responses to the three bacterial strains (Figure 8A). Cells responding to B8 clustered together, overlapping with B19-responsive cells, whereas T cells responding to the F5 were segregated from the former ones (Figure 8A).

Interestingly, a large majority of peripheral blood bacteria-reactive T cells, which were initially classified as DN T cells, demonstrated an absence or very poor expression of CD4 and CD8 also at the transcriptomic level (Figure 8B). However, based on their whole transcriptome, these cells were predominantly classified as cytotoxic cells, similar to CD8+ T cells, with smaller subsets corresponding to NKT cells and CD4+ cells (Figure 8C). Notably, TILs displayed a more balanced distribution across different T cell subsets, with a significant presence of CD4+ and regulatory T cells, in addition to DN T cells (Figure 8C).

Cumulative T cell phenotype distribution across four conditions (Tumor, F5, B19, B8) provided a more detailed classification of tumor-infiltrating T cells and a better understanding of T cell responses induced by the three CRC-associated bacteria (Figure 8D). Intriguingly, a large fraction of T cells expanded by all three bacteria showed transcriptomic profiles typical of unconventional T cells. Specifically, T cells reactive to B8 and B19 included large proportions of CD8+ MAIT and TCR-γδ T cells. Conversely, among F5-reactive cells, higher percentages of NKT cells were detected. Furthermore, as depicted in Figure 8D, F5 elicited an increased proportion of CD4+ cells in comparison to B8 and B19. In addition, a significant expansion of CD8+ cells, with traits indicative of exhaustion or a precursor to exhaustion (referred to as CD8+ TEX or CD8+ TPEX, respectively), was also noted. Furthermore, the analysis revealed that peripheral blood T cell responses were almost completely governed by the memory T cell compartment, consistent with our previous findings.

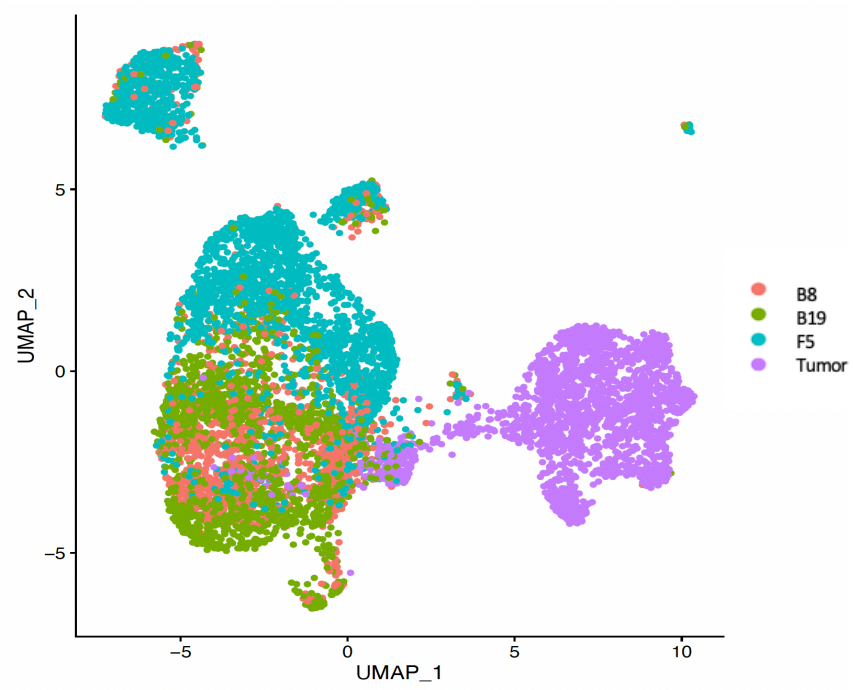
As expected, the phenotype distribution of T cells within the tumor differed markedly from those stimulated by bacterial antigens in the peripheral blood, particularly concerning the prominence of CD4+ cells and regulatory T cell function (Figure 8D). Notably, these tumors reflect notable infiltration by CD8+ effector memory cells, whereas fractions of MAIT and TCR-γδ T cells appeared to be reduced in comparison to bacteria-reactive T cells from peripheral blood.

Our analysis further explored the expression patterns of PD-1 and CD39 in our samples. We observed a pronounced expression of the CD39 marker, both in bacteria-induced T cells and TILs, aligning with our previous experimental results. Surprisingly, PD-1 expression was detected at very low levels in all samples (Figure 8E,F).

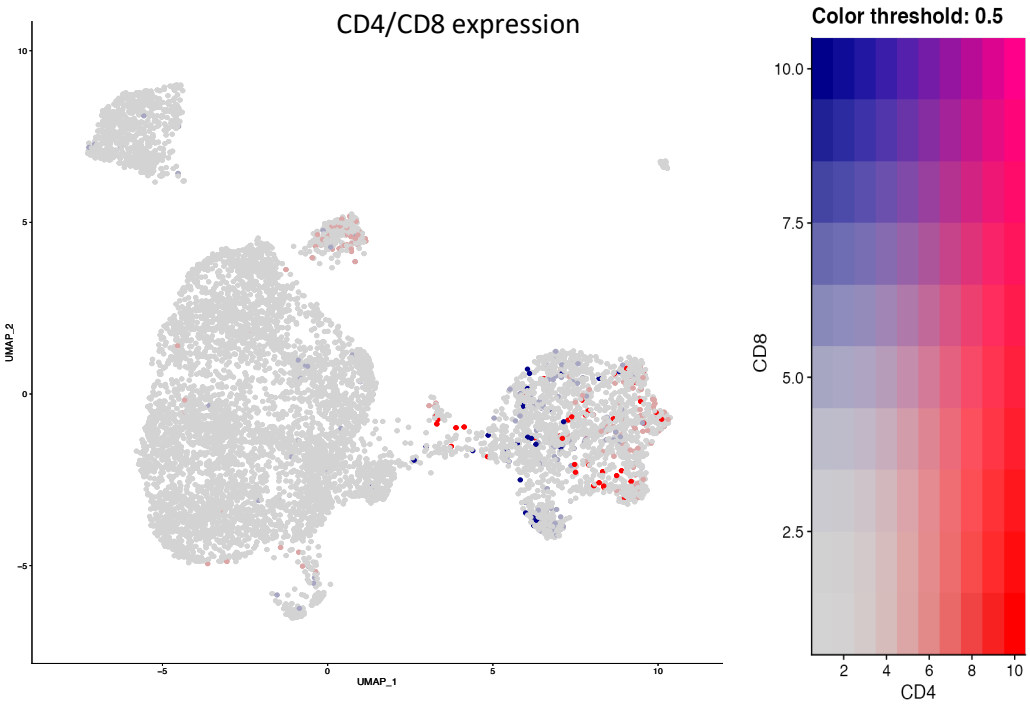
Moreover, we discovered that a substantial fraction of bacteria-expanded peripheral blood T cells exhibited high expression levels of DNAM-I and NKG2D receptors, which were also found to be expressed within TILs (Figure 8G,H). This finding further suggests the cytotoxic capabilities of CRC-associated bacteria-induced T cells and their possible role in anti-tumor responses.

Figure 8.

A


UMAP_2
UMAP_1
B8
B19
F5
Tumor

B


CD4/CD8 expression
Color threshold: 0.5
UMAP_2
UMAP_1
CD8
CD4

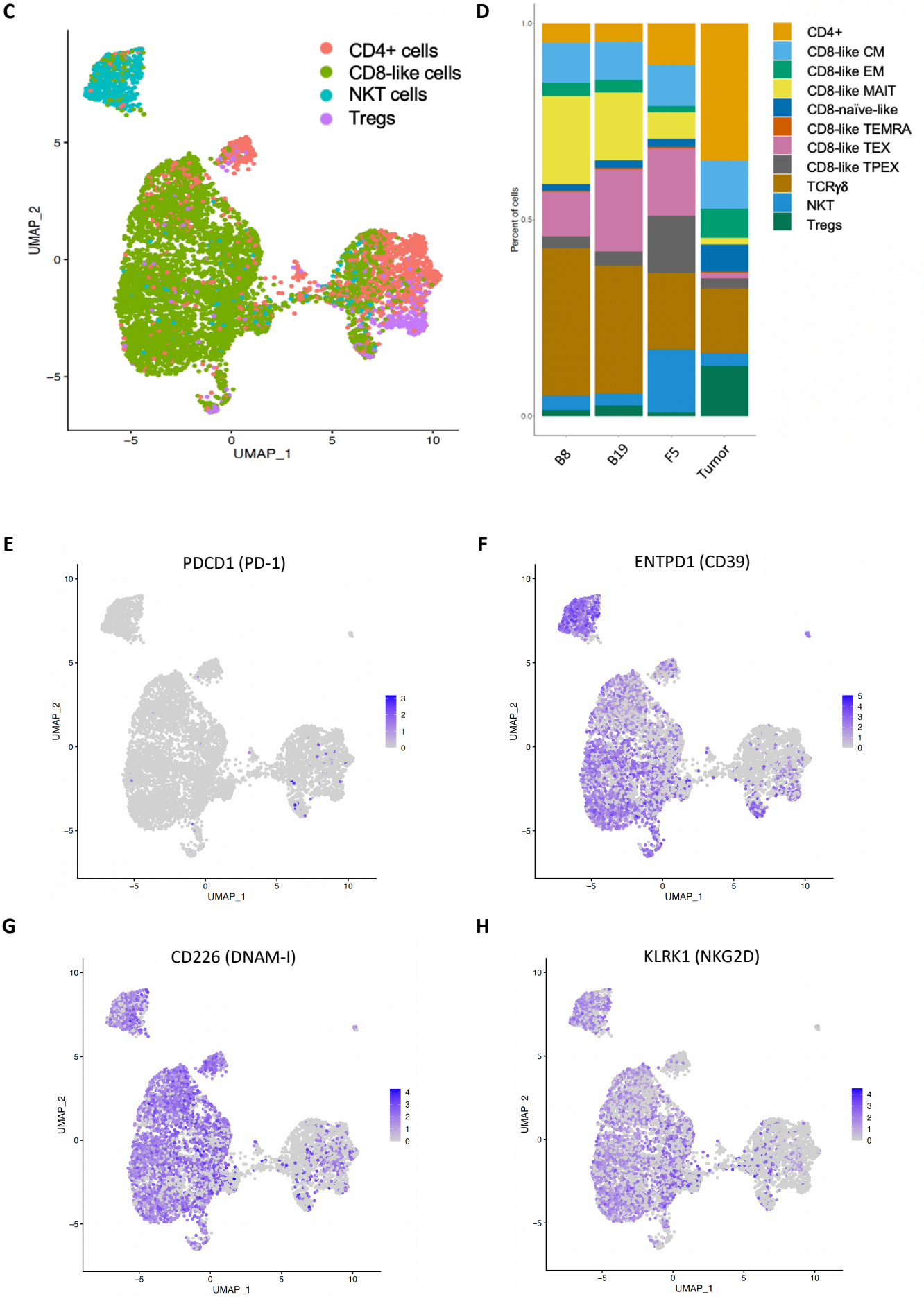
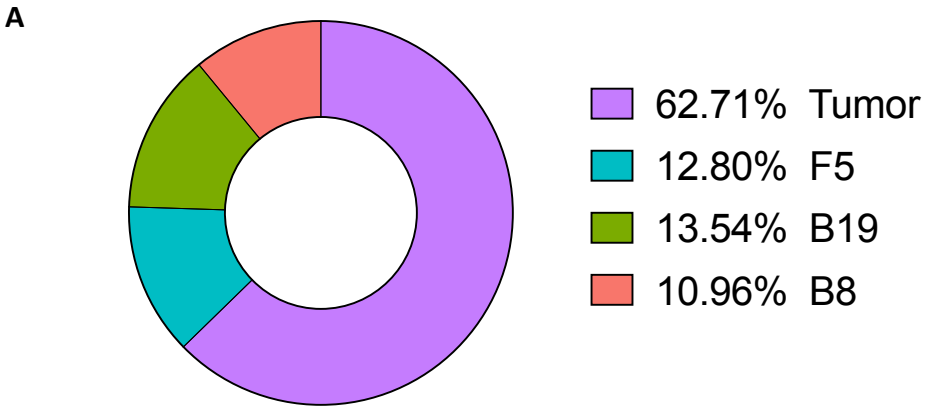

C
CD4+ cells
CD8-like cells
NKT cells
Tregs
UMAP_2
UMAP_1
D
CD4+
CD8-like CM
CD8-like EM
CD8-like MAIT
CD8-naïve-like
CD8-like TEMRA
CD8-like TEX
CD8-like TPEX
TCRγδ
NKT
Tregs
Percent of cells
B8
B19
F5
Tumor
E
PDCD1 (PD-1)
F
ENTPD1 (CD39)
G
CD226 (DNAM-I)
H
KLRK1 (NKG2D)

Figure 8: Comparative analysis of bacteria-expanded peripheral blood T cells and tumor-infiltrating T cells – whole transcriptome. Single-cell whole-transcriptome sequencing was performed on T cells isolated from MSS CRC patients (n=5). (A) UMAP representation of a total collection of 7,790 analyzed T cells, comprising purified T cells from primary CRC tissues (Tumor) and autologous peripheral blood T cells expanded by 3 bacteria (F5, B19, B8); (B) UMAP representation of CD4 and CD8 expressions in T cells; (C) UMAP representation of the main cell phenotypes annotated according to the whole transcriptome profiles; (D) Bar plot representing proportions of identified T cell phenotypes in each condition;CM=central memory, EM=effector memory, TEMRA=terminally differentiated effector memory cells expressing CD45RA, TEX=exhausted T cells, TPEX=precursor-exhausted T cells; UMAP representations of (E) PD-1, (F) CD39, (G) DNAM-I and (H) NKG2D expressions in total analyzed T cells. Each data point represents an individual T cell.

8. Bacteria-reactive DN T cells express clonotypes shared between three bacterial stimulations and TILs.

Finally, we performed a comprehensive comparative analysis of the TCR repertoires of bacteria-reactive T cells expanded from the peripheral blood of five CRC patients (antigens F5, B19, B8) against those of their autologous TILs. Using a specialized algorithm, we successfully identified unique TCRs for a total of 1,086 cells, out of which, 840 different clonotypes were detected. However, technical challenges in sequencing resulted in unbalanced cell distributions across the five CRC patients under four different conditions. The proportions of total TCR-identified cells among the four conditions are depicted in Figure 9A.

Figure 9.

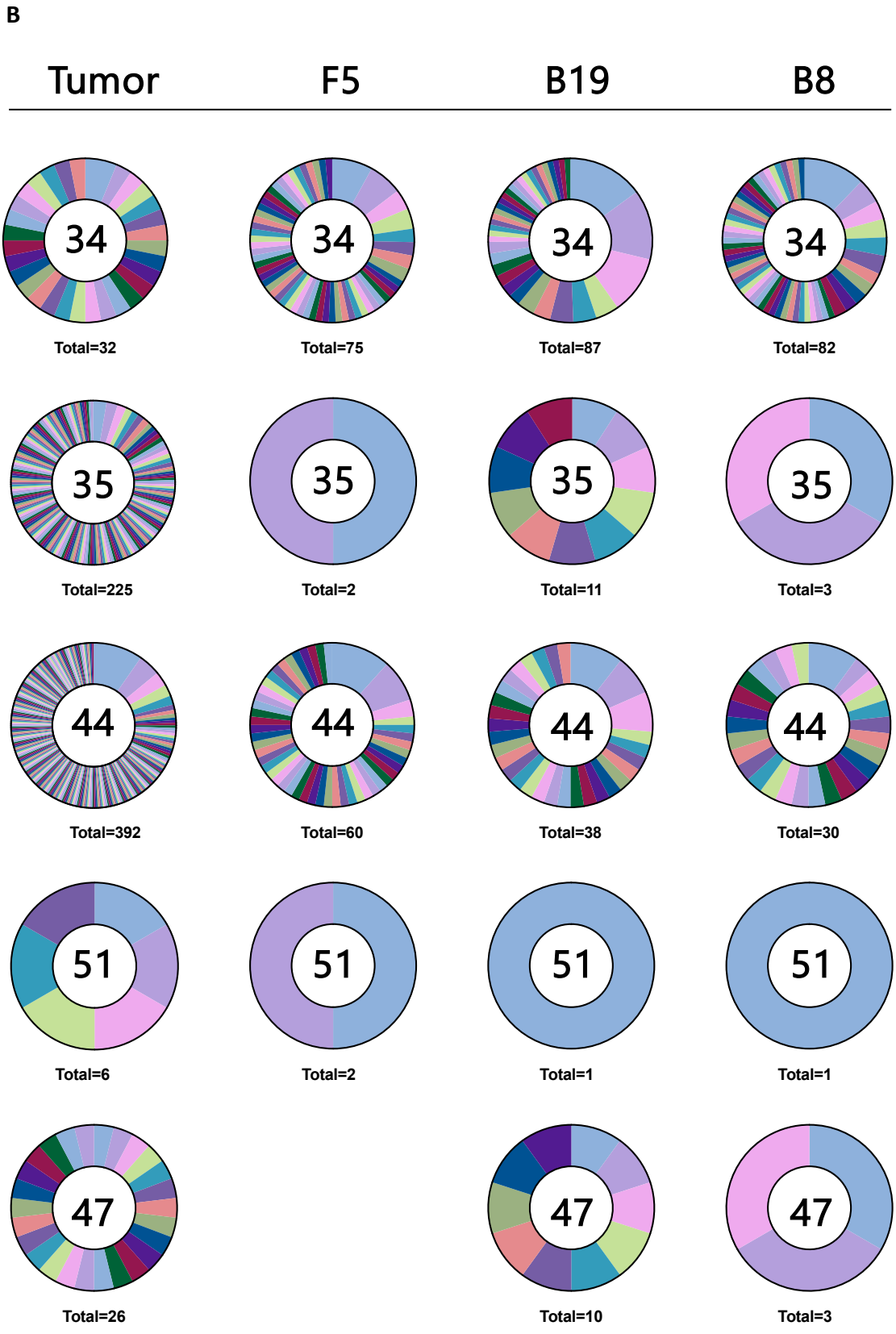


The number of T cell clonotypes identified per CRC patient is illustrated in Figure 9B. Notably, certain conditions yielded low clonotype outputs, hindering further analysis. Conversely, we observed an expansion of specific clonotypes, particularly in CRC patients 34 and 44 (Figure 9B). Interestingly, among TILs of patient 44, we could detect several expanded clonotypes, with a dominant clonotype 1 (Figure 9C), which according to whole transcriptome analysis, appears to align with the CD8+TPEX identity in the human T cell atlas. Clonotype 9, expanded in F5 condition was identified as CD4+, while expanded clonotype 13 induced by B19 was identified as CD8+CM (Figure 9C). Interestingly, B8 induced expansion of clonotype 8, which represents CD8+ MAIT cells, characterized by the expression of the previously described TRAJ33 chain. We have not observed shared clonotypes between the four conditions in this patient.

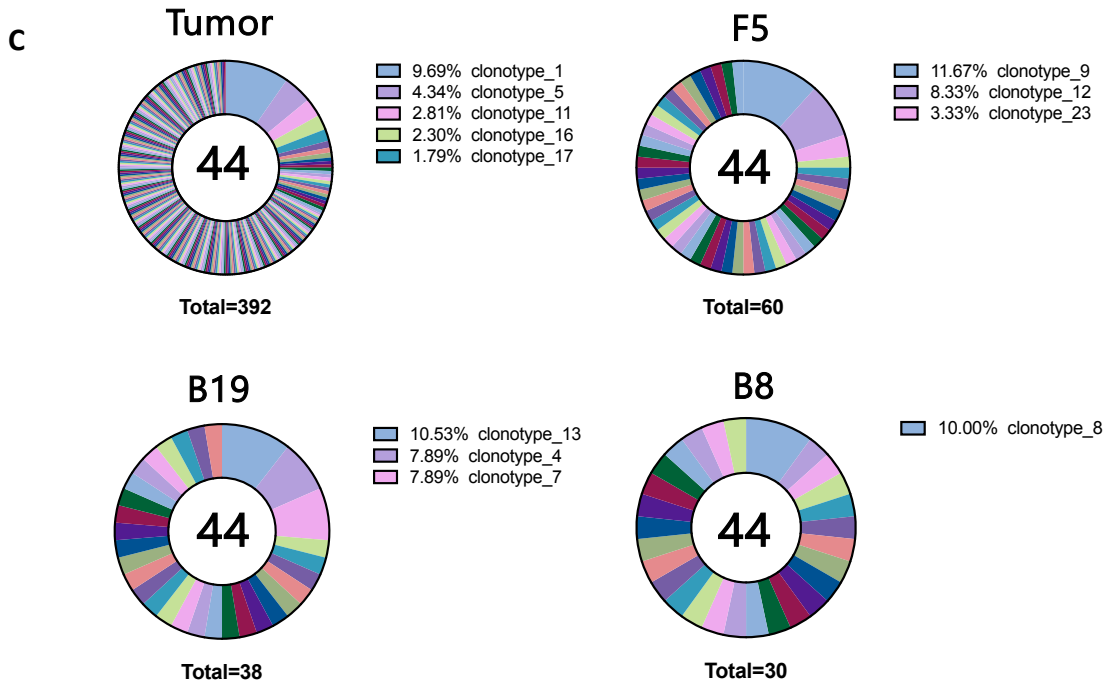
In patient 34 (Figure 9D) dominant clonotypes induced by F5 comprised mostly NKT and CD8+ cells. Clonotype 6, expanded under the B19 condition, was identified as CD8+ MAIT cells, also expressing the TRAJ33 chain, previously detected. Most interestingly, we found an overlap of two dominantly expanded clonotypes, namely clonotypes 2 and 3, between B19 and B8 conditions, both characterized as CD8+ MAIT cells with an invariant TCR featuring the TRAJ33 chain (Figure 9D). This finding implies the cross-reactivity of these unique clonotypes to different bacterial determinants.

Remarkably, when we overlayed all the detected clonotypes from four different conditions, we observed several clonotypes common to the three bacterial stimulations (Figure 9E). Specifically, we identified 2 shared clonotypes between F5 and B8, 14 clonotypes shared between B8 and B19, and 2 clonotypes present in all three bacterial conditions. Most importantly, we identified shared clonotypes between bacteria-expanded cells and TILs. This includes clonotype 18 (shared by F5 and TILs, in patient 34), clonotype 86 (shared by B8 and TILs, in patient 34), and clonotypes 20 and 105 (shared by B19 and TILs, in patient 34 and patient 35, respectively).

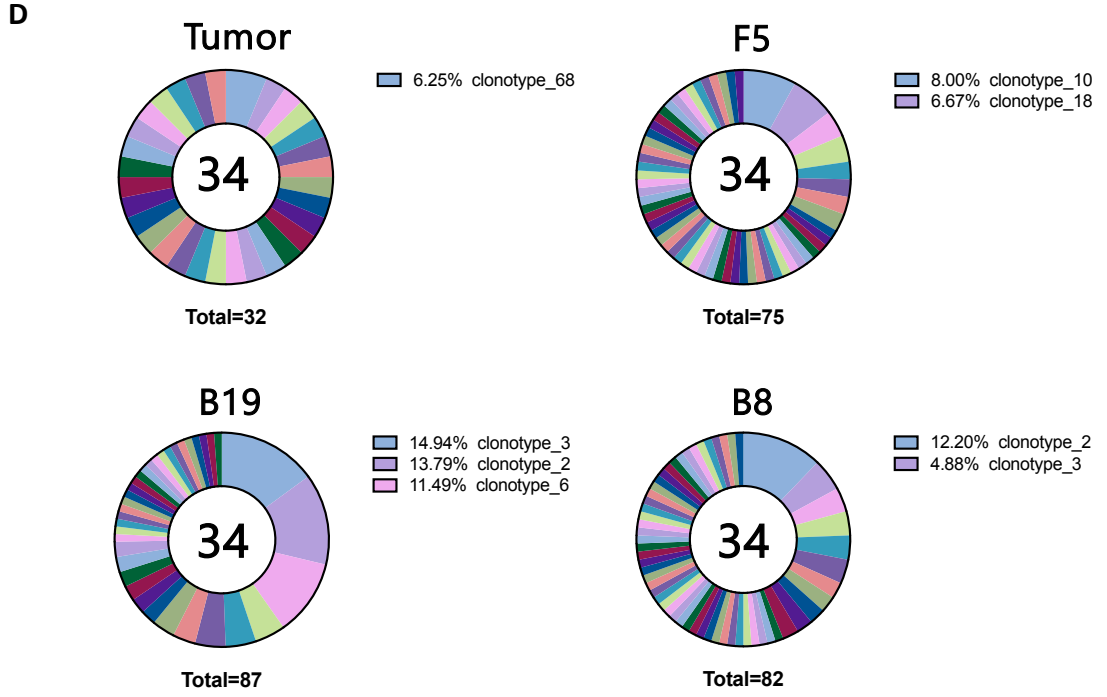
Figure 9.


B
Tumor
F5
B19
B8
34
Total=32
34
Total=75
34
Total=87
34
Total=82
35
Total=225
35
Total=2
35
Total=11
35
Total=3
44
Total=392
44
Total=60
44
Total=38
44
Total=30
51
Total=6
51
Total=2
51
Total=1
51
Total=1
47
Total=26
47
Total=10
47
Total=3

C


Tumor
44
Total=392
9.69% clonotype_1
4.34% clonotype_5
2.81% clonotype_11
2.30% clonotype_16
1.79% clonotype_17
F5
44
Total=60
11.67% clonotype_9
8.33% clonotype_12
3.33% clonotype_23
B19
44
Total=38
10.53% clonotype_13
7.89% clonotype_4
7.89% clonotype_7
B8
44
Total=30
10.00% clonotype_8

D


Tumor
34
Total=32
6.25% clonotype_68
F5
34
Total=75
8.00% clonotype_10
6.67% clonotype_18
B19
34
Total=87
14.94% clonotype_3
13.79% clonotype_2
11.49% clonotype_6
B8
34
Total=82
12.20% clonotype_2
4.88% clonotype_3

E

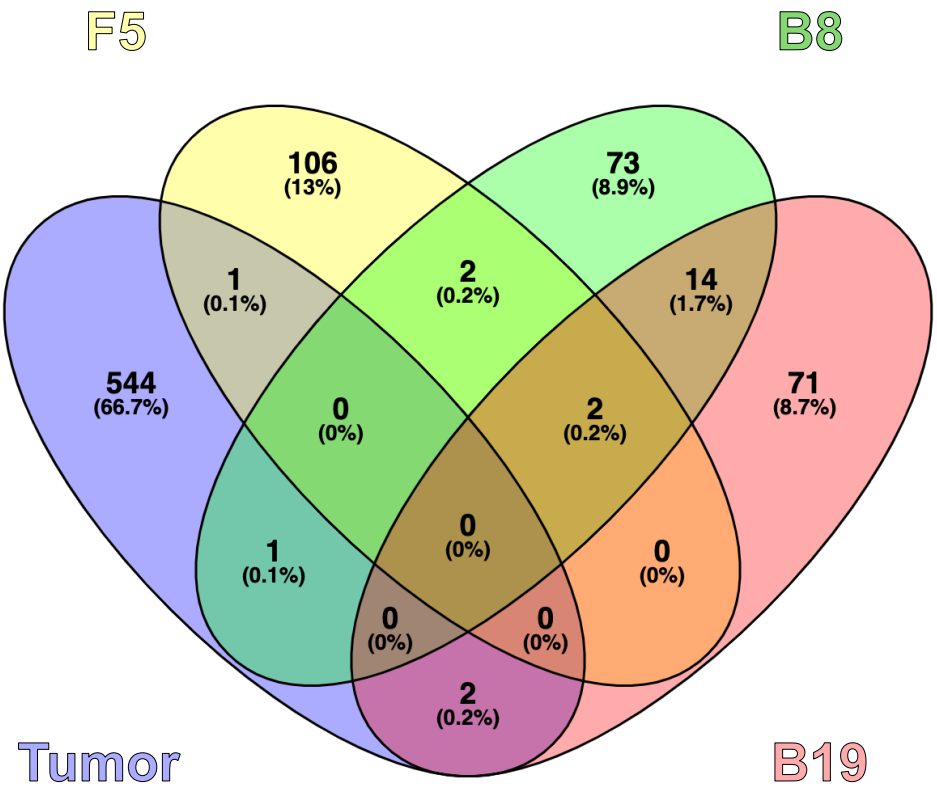


Figure 9: Comparative analysis of bacteria-expanded peripheral blood T cells and tumor-infiltrating T cells – TCR repertoire. Single-cell TCR sequencing was performed on TCRαβ cells isolated from MSS CRC patients (n=5). In total, 1,086 cells were assigned specific clonotype IDs. (A) The proportions of total TCR-identified cells among the four conditions (Tumor, F5, B19, B8) are depicted; (B) The numbers of T cell clonotypes identified in each CRC patient are shown. Segments in the pie charts represent different TCR clonotypes, and the size of the segment corresponds to the abundance of the unique expanded clonotype. CRC patient ID is reported in the center of the pie chart; The four pie charts represent the TCR repertoire of (C) patient CRC 44; (D) patient CRC 34, with annotation of the dominant clonotypes; (E) Venn diagram representing shared clonotypes within all detected clonotypes.

Discussion

The role of T lymphocytes in antitumor immunity, particularly in human CRC, is of paramount importance. Their presence within CRC tumors has been consistently linked to improved clinical outcomes, underscoring their potential as therapeutic targets [34,47]. Research efforts to date have tried to elucidate the antigenic specificities of T cells and understand the factors driving their expansion within certain CRC tumors. Notably, tumors with a high mutational burden (MSI tumors) often exhibit higher densities of TILs and a good response to immune checkpoint inhibitors [28,30,31]. This is most likely due to the abundance of neoantigens emerging in such tumor microenvironments, resulting in enhanced tumor-specific immune recognition and response [146–148]. However, despite these advances, it is critical to acknowledge the limitations in the effectiveness of immune checkpoint blockade therapy in patients with MSS tumors, which account for a large majority, and to further investigate the antigenic specificities of T cells in tumor infiltrates [148–150].

Recent studies have shed light on gut microbial antigens as potential drivers of beneficial immune responses in cancer. Particularly intriguing is the phenomenon of cross-reactivity between tumor and bacterial antigens observed in various malignancies [132–135]. This aspect is particularly pertinent in the context of CRC, which develops within a vast and diverse array of gut microbiota. Moreover, recent findings indicate a significant proportion of so-called bystander T cell activation in MSS tumor infiltrates, with TCR specificity for bacterial bystander antigens being predicted exclusively in MSS tumors[70].

In our study, we have delved into this relatively unexplored domain of gut bacteria-reactive T lymphocytes in CRC. We aimed to unravel the functional profiles, TCR repertoires, and response types of T cells stimulated by CRC-associated bacteria. Employing a comparative approach, we analyzed both bacteria-expanded T cells from peripheral blood and autologous TILs from MSS CRC patients. To our knowledge, this study is the first of its kind, addressing the presence and role of bacteria-reactive T cells within the CRC tumor microenvironment.

Numerous investigations have elucidated the association between specific bacterial species linked to CRC and the prevalence of various T cell phenotypes within CRC tissues, suggesting potentially significant clinical implications [99–102]. Research has identified gut microbiota-specific T cells[111], and their cross-reactivity with multiple bacterial species has been demonstrated [115–118]. Nevertheless, most of these insights have been derived from in vivo studies utilizing murine models. In contrast, our findings are exclusively derived from human ex vivo investigations. Furthermore, the T cell responses that we observed to the bacterial strains used in our study have not been fully elucidated so far.

We successfully identified CRC-associated bacteria-reactive T cells in the peripheral blood of both healthy donors and CRC patients. While some CRC patients exhibited prominent T cell responses, collectively, we have not observed significant differences in the extent of T cell reactivity to bacteria between healthy donors and CRC patients. We may speculate that these particularly responsive T cells of the aforementioned individuals (CRC 32, 37, and 47) had previously encountered these specific bacteria in the gut and generated peripheral memory, providing rapid responses. The relative quantification of these bacteria in primary tumors of these patients (Figure 1C) might support this speculation.

Furthermore, the response magnitude and phenotypes of induced T cells did not differ much between the three bacterial stimulations, suggesting that all three gut bacteria may induce T cell-mediated immune responses (Figure 3B,C). Intriguingly, our study revealed that both conventional and unconventional T cell subsets responded to this bacterial stimulation (Figure 3B,C).

Considering that the bacteria examined in our study are commensals, typically found in the gut, it would be reasonable to anticipate their tolerogenic responses. Nevertheless, what stood out in our observations was the notable absence of such T cell responses, evidenced by the lack of Treg induction (Figure 3B,C) and IL-10 secretion (Figure 5C,D). This could be due to several factors: (a) the potential of molecular mimicry between antigens of commensal and pathogenic bacteria, which may provoke a cytotoxic response against commensal bacteria, (b) the use of single bacterial strains in our experimental setting, whereas diversity and mutualism of the whole gut microbiota are key to preventing excessive inflammation, (c) the lack of the compartmentalization of immune responses within the gut and gut-associated lymphoid tissue

(GALT) in our experimental setting, which typically confines local immune reactions and favors dendritic cell-mediated tolerogenic responses. Nevertheless, a disruption of the normal gut structure, as occurring during CRC, may indeed break the immune tolerance and lead to the triggering of cytotoxic responses.

A significant observation in our study was the pronounced expansion of CD4-CD8- (double-negative, DN) T cells by all three tested CRC-associated bacterial strains. These DN T cells, characterized by the absence of both CD4 and CD8 markers, predominantly comprised TCRαβ effector memory cells, with a smaller fraction being MAIT or TCRγδ cells (Figure 3B-F). These cells, activated by direct interaction with bacteria-loaded antigen-presenting cells, exhibited robust expansion and cytokine production.

In literature, DN T cells are characterized as originating from the thymus and comprising up to 5% of the total T cell population in peripheral blood, but infiltrating various organs [139]. DN T cells can express either TCRαβ or TCRγδ, positioning them as integral components of both the innate and adaptive immune systems. A critical aspect of DN T cells is their ability to infiltrate solid tumors, playing a pivotal role in antitumor responses [139]. This function is facilitated through the expression of innate receptors NKG2D and DNAM-1, and the release of IFNγ and TNFα [142]. Additionally, they can mediate tumor cell killing through the Fas/FasL apoptotic pathway [143,144].

Strikingly, we demonstrated that CRC-associated bacteria trigger the upregulation of innate receptors such as NKG2D and DNAM-1 in proliferating DN T cells. This induction was coupled with the secretion of a spectrum of cytotoxic molecules, highlighting the robust functional capabilities of these cells. Leveraging this property, we investigated the antitumor potential of DN T cells. Our experiments successfully demonstrated that tumor cell exposure to both DN T cell-derived cytotoxic supernatants and DN T cells themselves effectively induced apoptosis in established human CRC cell lines, although in the latter case, the mechanism of DN T cell-mediated direct tumor killing remains to be addressed. Notably, gut bacteria-derived SCFAs have been reported to induce NKG2D ligands surface expression on CRC cells [151], which is predictive of improved patient survival [152,153]. Thus, defined CRC-associated bacteria may indirectly mediate anti-tumor effects on one hand by expanding NKG2D-expressing T cells, and on the other by inducing upregulation of their corresponding

ligands on tumor cells. However, whether metabolites derived from bacteria under investigation may modulate NKG2D ligands expression remains to be investigated.
Importantly, upon analysis of fresh primary tumor specimens, we have identified abundant clusters of DN T cells within TILs. Interestingly, while the frequency of total TILs did not significantly differ between MSS and MSI CRCs, DN T cells appeared to be more abundant in MSS CRCs. This finding may indicate that tumor infiltration by DN T cells is not influenced by the microsatellite status of the tumor, but rather by other determinants. However, a notable observation was the upregulation of T cell markers associated with exhaustion, including PD-1 and CD39, in many of the infiltrating cells, but also in bacteria-induced proliferating DN T cells. This might imply the lack of their effector function within the tumor microenvironment. By targeting these exhaustion pathways, it might be possible to rejuvenate the antitumor activity of DN T cells, enhancing their efficacy against CRC. Several agents targeting CD39, either alone or in combination with other treatments, are progressing into clinical trials, including those for CRCs [154]. Our findings would suggest that a combination of anti-PD-1 and anti-CD39 may represent a novel therapeutic approach for MSS CRCs.

Finally, we employed the single-cell RNA sequencing method to elucidate the unique identities and functional specialization of both bacteria-reactive and tumor-infiltrating T cells, particularly in five patients exhibiting the MSS phenotype. This focus was strategically chosen to dissect the T cell landscape in tumors devoid of neo-antigens, thereby aiming to unravel the potential roles and specificities of "bystander" T cells in such an environment.

An important observation was the lack of the expression of CD4 and CD8 in most of the sequenced T cells, aligning with our data from functional assays obtained by flow cytometry. However, the WT data was correlated to the publicly available human T cell atlas used, which does not include DN T cell characterization. As a result, the algorithm identified DN T cells as CD8+ due to a substantial match in their transcripts, even though they do not express CD8 marker. Therefore, it is essential to consider this context for all the subsequent annotations including CD8+ cells, and acknowledge it as a limitation of this study. In addition, most of the studies involving single-cell sequencing normally exclude DN T cells from their analysis. E.g. cells with a negligible expression of CD8A/B have been annotated as ‘low quality’ and were not the focus of further research in other studies [70].

The whole transcriptome (WT) analysis also revealed that peripheral blood T cells responding to B8 clustered together, overlapping with B19-responsive cells, whereas T cells responding to the F5 were segregated from the former ones (Figure 8A). This pattern suggests shared gene expression features in the T cell responses to these two antigens, potentially due to structural or immunogenic similarities, as both antigens originate from the same bacterial genus. The observed overlap might also indicate cross-reactivity or shared epitopes between B8 and B19 antigens. Certainly, TCR sequencing unraveled multiple clonotypes cross-reactive between these *Bacteroides* species. In contrast, the segregation of T cells responding to the F5 antigen within the plot possibly implies a more different T cell response to this antigen. This could mean that the F5 antigen might be triggering specific signaling pathways, or it induces a specific subset of T cells. Indeed, WT analysis showed a relative expansion of CD4, CD8+TPEX, and NKT subsets among F5-reactive cells, as compared to other subsets. Nevertheless, 2 clonotypes shared between F5-, B8- and B19-induced T cells were also identified.

Due to the exclusive amplification of TCR α and β chains, unfortunately, we could not detect TCRγδ cells in the TCR repertoire of the sequenced cells. Therefore, the information on the TCRγδ clonotypes and possible bacterial antigen specificities and cross-reactivity of the TCRγδ TILs is missing. This also explains the limited numbers of TCR-sequenced cells, given that WT findings revealed a significant presence of TCRγδ cells in our samples.

However, despite having sequenced a relatively low number of T cells, from five patients only, we could identify 4 bacteria-specific T cell clonotypes shared with TILs, thereby conclusively indicating that bacteria-specific T cells are actually included in CRC-infiltrating T cell repertoire.

In summary, our study has shed light on the complex interactions between gut microbiota and T-cell responses in CRC. We found that DN T cells, detected among TILs and particularly abundant in poorly immunogenic MSS CRCs, possess innate-like cytotoxic abilities, and most importantly, can be expanded by CRC-associated bacteria. Our study leaves a few critical questions, such as the molecular mechanism underlying the tumor cell-killing ability of DN T cells, and the restriction element for their antigen recognition, open for future investigation. Nevertheless, our findings suggest that boosting the intratumoral expansion of DN T cells may

represent a novel therapeutic approach for enhancing effective anti-tumor responses. Adoptive transfer of DN T cells combined with immune checkpoint inhibitors, such as anti-PD-1, has been recently shown to result in prominent antitumor effects in experimental tumor models [155].

Cumulatively, our data support and further expand this concept and may pave the way for innovative combination treatments, based on the administration of bacterial antigens and novel immunological inhibitors, such as anti-CD39, potentially boosting the intratumoral expansion of DN T cells in poorly immunogenic CRCs.

Bibliography:

1. 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 (2021).
2. Favoriti, P. *et al.* Worldwide burden of colorectal cancer: a review. *Updates in Surgery* vol. 68 7–11 Preprint at https://doi.org/10.1007/s13304-016-0359-y (2016).
3. Keum, N. N. & Giovannucci, E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nature Reviews Gastroenterology and Hepatology* vol. 16 713–732 Preprint at https://doi.org/10.1038/s41575-019-0189-8 (2019).
4. Patel, S. G., Karlitz, J. J., Yen, T., Lieu, C. H. & Boland, C. R. The rising tide of early-onset colorectal cancer: a comprehensive review of epidemiology, clinical features, biology, risk factors, prevention, and early detection. *The Lancet Gastroenterology and Hepatology* vol. 7 262–274 Preprint at https://doi.org/10.1016/S2468-1253(21)00426-X (2022).
5. O'Sullivan, D. E. *et al.* Risk Factors for Early-Onset Colorectal Cancer: A Systematic Review and Meta-analysis. *Clinical Gastroenterology and Hepatology* vol. 20 1229-1240.e5 Preprint at https://doi.org/10.1016/j.cgh.2021.01.037 (2022).
6. Fleming, M., Ravula, S., Tatishchev, S. F. & Wang, H. L. Colorectal carcinoma: Pathologic aspects. *Journal of Gastrointestinal Oncology* vol. 3 153–173 Preprint at https://doi.org/10.3978/j.issn.2078-6891.2012.030 (2012).
7. Fearon, E. F. & Vogelstein, B. *A Genetic Model for Colorectal Tumorigenesis*. *Cell* vol. 61 (1990).
8. Hossain, M. S. *et al.* Colorectal Cancer: A Review of Carcinogenesis, Global Epidemiology, Current Challenges, Risk Factors, Preventive and Treatment Strategies. *Cancers* vol. 14 Preprint at https://doi.org/10.3390/cancers14071732 (2022).
9. Hu, Z. *et al.* Quantitative evidence for early metastatic seeding in colorectal cancer. *Nat Genet* **51**, 1113–1122 (2019).
10. Cheung, K. J. & Ewald, A. J. A collective route to metastasis: Seeding by tumor cell clusters. *Science* vol. 352 167–169 Preprint at https://doi.org/10.1126/science.aaf6546 (2016).
11. Ciardiello, F. *et al.* Clinical management of metastatic colorectal cancer in the era of precision medicine. *CA Cancer J Clin* **72**, 372–401 (2022).
12. Zhou, H. *et al.* Colorectal liver metastasis: molecular mechanism and interventional therapy. *Signal Transduction and Targeted Therapy* vol. 7 Preprint at https://doi.org/10.1038/s41392-022-00922-2 (2022).
13. Fearnhead, N. S., Wilding, J. L. & Bodmer, W. F. *Genetics of Colorectal Cancer: Hereditary Aspects and Overview of Colorectal Tumorigenesis*. *British Medical Bulletin* vol. 64 https://academic.oup.com/bmb/article/64/1/27/293605 (2002).
14. Dekker, E., Tanis, P. J., Vleugels, J. L. A., Kasi, P. M. & Wallace, M. B. *Colorectal Cancer*. *www.thelancet.com* www.thelancet.com.

15. Cunningham, D. *et al.* Colorectal cancer. *The Lancet* vol. 375 1030–1047 Preprint at https://doi.org/10.1016/S0140-6736(10)60353-4 (2010).
16. Markowitz, S. D. & Bertagnolli, M. M. Molecular Basis of Colorectal Cancer. *New England Journal of Medicine* **361**, 2449–2460 (2009).
17. Schmitt, M. & Greten, F. R. The inflammatory pathogenesis of colorectal cancer. *Nature Reviews Immunology* vol. 21 653–667 Preprint at https://doi.org/10.1038/s41577-021-00534-x (2021).
18. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat Med* **21**, 1350–1356 (2015).
19. Dienstmann, R. *et al.* Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nature Reviews Cancer* vol. 17 79–92 Preprint at https://doi.org/10.1038/nrc.2016.126 (2017).
20. Fadaka, A. O., Pretorius, A. & Klein, A. Biomarkers for Stratification in Colorectal Cancer: MicroRNAs. *Cancer Control* vol. 26 Preprint at https://doi.org/10.1177/1073274819862784 (2019).
21. Nagtegaal, I. D., Quirke, P. & Schmoll, H. J. Has the new TNM classification for colorectal cancer improved care? *Nature Reviews Clinical Oncology* vol. 9 119–123 Preprint at https://doi.org/10.1038/nrclinonc.2011.157 (2012).
22. Maguire, A. & Sheahan, K. Controversies in the pathological assessment of colorectal cancer. *World Journal of Gastroenterology* 9850–9861 Preprint at https://doi.org/10.3748/wjg.v20.i29.9850 (2014).
23. Lugli, A. *et al.* Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. *Modern Pathology* **30**, 1299–1311 (2017).
24. Galon, J. *et al.* Cancer classification using the Immunoscore: A worldwide task force. *Journal of Translational Medicine* vol. 10 Preprint at https://doi.org/10.1186/1479-5876-10-205 (2012).
25. Kuipers, E. J. *et al.* Colorectal cancer. *Nat Rev Dis Primers* **1**, (2015).
26. Weitz, J. *et al.* Colorectal cancer. in *Lancet* vol. 365 153–165 (Elsevier B.V., 2005).
27. Gatalica, Z., Vranic, S., Xiu, J., Swensen, J. & Reddy, S. High microsatellite instability (MSI-H) colorectal carcinoma: a brief review of predictive biomarkers in the era of personalized medicine. *Familial Cancer* vol. 15 405–412 Preprint at https://doi.org/10.1007/s10689-016-9884-6 (2016).
28. Ooki, A., Shinozaki, E. & Yamaguchi, K. Immunotherapy in Colorectal Cancer: Current and Future Strategies. *J Anus Rectum Colon* **5**, 11–24 (2021).
29. Johdi, N. A. & Sukor, N. F. Colorectal Cancer Immunotherapy: Options and Strategies. *Frontiers in Immunology* vol. 11 Preprint at https://doi.org/10.3389/fimmu.2020.01624 (2020).
30. Lenz, H.-J. *et al.* First-Line Nivolumab Plus Low-Dose Ipilimumab for Microsatellite Instability-High/ Mismatch Repair-Deficient Metastatic Colorectal Cancer: The Phase II CheckMate 142 Study. *J Clin Oncol* **40**, 161–170 (2021).
31. André, T. *et al.* Pembrolizumab in Microsatellite-Instability–High Advanced Colorectal Cancer. *New England Journal of Medicine* **383**, 2207–2218 (2020).

32. Cercek, A. *et al.* PD-1 Blockade in Mismatch Repair–Deficient, Locally Advanced Rectal Cancer. *New England Journal of Medicine* **386**, 2363–2376 (2022).
33. Galon, J., Fridman, W. H. & Pages, F. The adaptive immunologic microenvironment in colorectal cancer: A novel perspective. *Cancer Research* vol. 67 1883–1886 Preprint at https://doi.org/10.1158/0008-5472.CAN-06-4806 (2007).
34. Fridman, W. H., Pagès, F., Sauts-Fridman, C. & Galon, J. The immune contexture in human tumours: Impact on clinical outcome. *Nature Reviews Cancer* vol. 12 298–306 Preprint at https://doi.org/10.1038/nrc3245 (2012).
35. Al-Mterin, M. A. & Elkord, E. Myeloid-derived suppressor cells in colorectal cancer: prognostic biomarkers and therapeutic targets. *Exploration of Targeted Anti-tumor Therapy* vol. 3 497–510 Preprint at https://doi.org/10.37349/etat.2022.00097 (2022).
36. Yin, K., Xia, X., Rui, K., Wang, T. & Wang, S. Myeloid-Derived Suppressor Cells: A New and Pivotal Player in Colorectal Cancer Progression. *Frontiers in Oncology* vol. 10 Preprint at https://doi.org/10.3389/fonc.2020.610104 (2020).
37. Koelzer, V. H. *et al.* Phenotyping of tumor-associated macrophages in colorectal cancer: Impact on single cell invasion (tumor budding) and clinicopathological outcome. *Oncoimmunology* **5**, (2016).
38. Li, S. *et al.* Tumor-associated macrophages remodeling EMT and predicting survival in colorectal carcinoma. *Oncoimmunology* **7**, (2018).
39. Herrera, M. *et al.* Cancer-associated fibroblast and M2 macrophage markers together predict outcome in colorectal cancer patients. *Cancer Sci* **104**, 437–444 (2013).
40. Galdiero, M. R. *et al.* Occurrence and significance of tumor-Associated neutrophils in patients with colorectal cancer. *Int J Cancer* **139**, 446–456 (2016).
41. Hirt, C. *et al.* Colorectal carcinoma infiltration by myeloperoxidase-expressing neutrophil granulocytes is associated with favorable prognosis. *Oncoimmunology* **2**, (2013).
42. Sconocchia, G. *et al.* Tumor infiltration by FcIγRIII (CD16)+ myeloid cells is associated with improved survival in patients with colorectal carcinoma. *Int J Cancer* **128**, 2663–2672 (2011).
43. Chengzeng, Y. I. N. *et al.* Prognostic significance of CD8+ tumor-infiltrating lymphocytes and CD66b+ tumor-associated neutrophils in the invasive margins of stages I-III colorectal cancer. *Oncol Lett* **24**, (2022).
44. Rottmann BG *et al.* Clinicopathological significance of neutrophil-rich colorectal carcinoma. *J Clin Pathol.* (2023) doi:10.1136/jclinpath-2021-207702.
45. Hu, X. *et al.* A Risk Signature With Inflammatory and T Immune Cells Infiltration in Colorectal Cancer Predicting Distant Metastases and Efficiency of Chemotherapy. *Front Oncol* **9**, (2019).
46. Gulubova, M. V. *et al.* Role of dendritic cells in progression and clinical outcome of colon cancer. *Int J Colorectal Dis* **27**, 159–169 (2012).
47. Galon, J. *et al.* Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome. *Science (1979)* **313**, 1960–1964 (2006).

48. Pagès, F. *et al. Effector Memory T Cells, Early Metastasis, and Survival in Colorectal Cancer and the Departments of General and Digestive Surgery (A. N Engl J Med* vol. 25 www.genome.tugraz.at (2005).
49. Bindea, G. *et al.* Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* **39**, 782–795 (2013).
50. Tosolini, M. *et al.* Clinical impact of different classes of infiltrating T cytotoxic and helper cells (Th1, Th2, Treg, Th17) in patients with colorectal cancer. *Cancer Res* **71**, 1263–1271 (2011).
51. Camus, M. *et al.* Coordination of intratumoral immune reaction and human colorectal cancer recurrence. *Cancer Res* **69**, 2685–2693 (2009).
52. Ogino, S. *et al.* Lymphocytic reaction to colorectal cancer is associated with longer survival, independent of lymph node count, microsatellite instability, and CpG island methylator phenotype. *Clinical Cancer Research* **15**, 6412–6420 (2009).
53. Nosho, K. *et al.* Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer, and prognosis: Cohort study and literature review. *Journal of Pathology* **222**, 350–366 (2010).
54. Masuda, K. *et al.* Multiplexed single-cell analysis reveals prognostic and nonprognostic T cell types in human colorectal cancer. *JCI Insight* **7**, (2022).
55. Di, J. *et al.* Phenotype molding of T cells in colorectal cancer by single-cell analysis. *Int J Cancer* **146**, 2281–2295 (2020).
56. Frey, D. M. *et al.* High frequency of tumor-infiltrating FOXP3+ regulatory T cells predicts improved survival in mismatch repair-proficient colorectal cancer patients. *Int J Cancer* **126**, 2635–2643 (2010).
57. Salama, P. *et al.* Tumor-infiltrating FOXP3+ T regulatory cells show strong prognostic significance in colorectal cancer. *Journal of Clinical Oncology* **27**, 186–192 (2009).
58. Michel, S. *et al.* High density of FOXP3-positive T cells infiltrating colorectal cancers with microsatellite instability. *Br J Cancer* **99**, 1867–1873 (2008).
59. Liu, J. *et al.* IL-17 is associated with poor prognosis and promotes angiogenesis via stimulating VEGF production of cancer cells in colorectal carcinoma. *Biochem Biophys Res Commun* **407**, 348–354 (2011).
60. Amicarella, F. *et al.* Dual role of tumour-infiltrating T helper 17 cells in human colorectal cancer. *Gut* **66**, 692–704 (2017).
61. Pagès, F. *et al.* In situ cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. *Journal of Clinical Oncology* **27**, 5944–5951 (2009).
62. Galon, J. & Bruni, D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nature Reviews Drug Discovery* vol. 18 197–218 Preprint at https://doi.org/10.1038/s41573-018-0007-y (2019).
63. Mlecnik, B. *et al.* Multicenter International Society for Immunotherapy of Cancer Study of the Consensus Immunoscore for the Prediction of Survival and Response to Chemotherapy in Stage III Colon Cancer. (2020) doi:10.1200/JCO.19.
64. Bremers, A. J. A. *et al.* T cell responses in colorectal cancer patients: Evidence for class II HLA-restricted recognition of shared tumor-associated antigens. *Int J Cancer* **88**, 956–961 (2000).

65. Bonertz, A. *et al.* Antigen-specific Tregs control T cell responses against a limited repertoire of tumor antigens in patients with colorectal carcinoma. *Journal of Clinical Investigation* **119**, 3311–3321 (2009).
66. Schumacher, T. N. & Schreiber, R. D. *Neoantigens in Cancer Immunotherapy*. www.sciencemag.org.
67. Cohen, C. J. *et al.* Isolation of neoantigen-specific T cells from tumor and peripheral lymphocytes. *Journal of Clinical Investigation* **125**, 3981–3991 (2015).
68. McGranahan, N. *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science (1979)* **351**, 1463–1469 (2016).
69. Simoni, Y. *et al.* Bystander CD8+ T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature* **557**, 575–579 (2018).
70. Borràs, D. M. *et al.* Single cell dynamics of tumor specificity vs bystander activity in CD8+ T cells define the diverse immune landscapes in colorectal cancer. *Cell Discov* **9**, (2023).
71. Belkaid, Y. & Harrison, O. J. Homeostatic Immunity and the Microbiota. *Immunity* vol. 46 562–576 Preprint at https://doi.org/10.1016/j.immuni.2017.04.008 (2017).
72. Camilleri, M. Leaky gut: mechanisms, measurement and clinical implications in humans. *Gut* vol. 68 1516–1526 Preprint at https://doi.org/10.1136/gutjnl-2019-318427 (2019).
73. McGuckin, M. A., Lindén, S. K., Sutton, P. & Florin, T. H. Mucin dynamics and enteric pathogens. *Nat Rev Microbiol* **9**, 265–278 (2011).
74. Abreu, M. T. Toll-like receptor signalling in the intestinal epithelium: How bacterial recognition shapes intestinal function. *Nature Reviews Immunology* vol. 10 131–143 Preprint at https://doi.org/10.1038/nri2707 (2010).
75. Adak, A. & Khan, M. R. An insight into gut microbiota and its functionalities. *Cellular and Molecular Life Sciences* vol. 76 473–493 Preprint at https://doi.org/10.1007/s00018-018-2943-4 (2019).
76. Ansaldo, E., Farley, T. K. & Belkaid, Y. Control of Immunity by the Microbiota. *Annu Rev Immunol* **39**, 449–479 (2021).
77. Song, C. *et al.* Intestinal mucus components and secretion mechanisms: what we do and do not know. *Experimental and Molecular Medicine* vol. 55 681–691 Preprint at https://doi.org/10.1038/s12276-023-00960-y (2023).
78. Portincasa, P. *et al.* Gut Microbiota and Short Chain Fatty Acids: Implications in Glucose Homeostasis. *International Journal of Molecular Sciences* vol. 23 Preprint at https://doi.org/10.3390/ijms23031105 (2022).
79. Feng, Q. *et al.* Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nat Commun* **6**, (2015).
80. Zeller, G. *et al.* Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol Syst Biol* **10**, (2014).
81. Dai, Z. *et al.* Multi-cohort analysis of colorectal cancer metagenome identified altered bacteria across populations and universal bacterial markers. *Microbiome* **6**, 70 (2018).
82. Grivennikov, S. I. *et al.* Adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth. *Nature* **491**, 254–258 (2012).

83. Gallimore, A. M. & Godkin, A. Epithelial Barriers, Microbiota, and Colorectal Cancer. *New England Journal of Medicine* **368**, 282–284 (2013).
84. Castellarin, M. *et al.* Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Res* **22**, 299–306 (2012).
85. Rubinstein, M. R. *et al. Fusobacterium nucleatum* promotes colorectal cancer by inducing Wnt/β-catenin modulator Annexin A1. *EMBO Rep* **20**, (2019).
86. Kostic, A. D. *et al.* Fusobacterium nucleatum Potentiates Intestinal Tumorigenesis and Modulates the Tumor-Immune Microenvironment. *Cell Host Microbe* **14**, 207–215 (2013).
87. Gur, C. *et al.* Binding of the Fap2 Protein of Fusobacterium nucleatum to Human Inhibitory Receptor TIGIT Protects Tumors from Immune Cell Attack Europe PMC Funders Group. *Immunity* **42**, 344–355 (2015).
88. Mima, K. *et al.* Fusobacterium nucleatum in colorectal carcinoma tissue and patient prognosis. *Gut* **65**, 1973–1980 (2016).
89. Hamada, T. *et al.* Fusobacterium nucleatum in colorectal cancer relates to immune response differentially by tumor microsatellite instability status. *Cancer Immunol Res* **6**, 1327–1336 (2018).
90. Gao, Y. *et al.* Fusobacterium nucleatum enhances the efficacy of PD-L1 blockade in colorectal cancer. *Signal Transduct Target Ther* **6**, (2021).
91. Boleij, A. *et al.* The bacteroides fragilis toxin gene is prevalent in the colon mucosa of colorectal cancer patients. *Clinical Infectious Diseases* **60**, 208–215 (2015).
92. Chung, L. *et al.* Bacteroides fragilis Toxin Coordinates a Pro-carcinogenic Inflammatory Cascade via Targeting of Colonic Epithelial Cells. *Cell Host Microbe* **23**, 203-214.e5 (2018).
93. Wu, S. *et al.* A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat Med* **15**, 1016–1022 (2009).
94. Purcell, R. V. *et al.* Colonization with enterotoxigenic Bacteroides fragilis is associated with early-stage colorectal neoplasia. *PLoS One* **12**, (2017).
95. Wilson, M. R. *et al.* The human gut bacterial genotoxin colibactin alkylates DNA. *Science (1979)* **363**, (2019).
96. Hernández-Luna, M. A., López-Briones, S. & Luria-Pérez, R. The Four Horsemen in Colon Cancer. *Journal of Oncology* vol. 2019 Preprint at https://doi.org/10.1155/2019/5636272 (2019).
97. Bennedsen, A. L. B. *et al.* The gut microbiota can orchestrate the signaling pathways in colorectal cancer. *APMIS* vol. 130 121–139 Preprint at https://doi.org/10.1111/apm.13206 (2022).
98. Wong, S. H. & Yu, J. Gut microbiota in colorectal cancer: mechanisms of action and clinical applications. *Nature Reviews Gastroenterology and Hepatology* vol. 16 690–704 Preprint at https://doi.org/10.1038/s41575-019-0209-8 (2019).
99. Luu, K. *et al.* Fecal and Tissue Microbiota Are Associated with Tumor T-Cell Infiltration and Mesenteric Lymph Node Involvement in Colorectal Cancer. *Nutrients* **15**, (2023).

100. Kim, H. S. *et al.* Fusobacterium nucleatum induces a tumor microenvironment with diminished adaptive immunity against colorectal cancers. *Front Cell Infect Microbiol* **13**, (2023).
101. Duggan, W. P. *et al.* Increased Fusobacterium tumoural abundance affects immunogenicity in mucinous colorectal cancer and may be associated with improved clinical outcome. *J Mol Med* **101**, 829–841 (2023).
102. Cremonesi, E. *et al.* Gut microbiota modulate T cell trafficking into human colorectal cancer. *Gut* **67**, 1984–1994 (2018).
103. Montalban-Arques, A. *et al.* Commensal Clostridiales strains mediate effective anti-cancer immune response against solid tumors. *Cell Host Microbe* **29**, 1573-1588.e7 (2021).
104. Lu, Y., Wang, L., Zhang, J., Li, J. & Wan, G. Induction of CD8 T cell cytotoxicity by fecal bacteria from healthy individuals and colorectal cancer patients. *Biochem Biophys Res Commun* **516**, 1007–1012 (2019).
105. Tanoue, T. *et al.* A defined commensal consortium elicits CD8 T cells and anti-cancer immunity. *Nature* **565**, 600–605 (2019).
106. Oh, S. F., Jung, D.-J. & Choi, E. Gut Microbiota-Derived Unconventional T Cell Ligands: Contribution to Host Immune Modulation. *Immunohorizons* **6**, 476–487 (2022).
107. Ulevitch, R. J. *Innate Immune Responses to Microbial Pathogens Molecular Mechanisms of Innate Immunity*. *Immunologic Research* vol. 21 (2000).
108. Kumar, S., Ingle, H., Prasad, D. V. R. & Kumar, H. Recognition of bacterial infection by innate immune sensors. *Critical Reviews in Microbiology* vol. 39 229–246 Preprint at https://doi.org/10.3109/1040841X.2012.706249 (2013).
109. Maksymowych, W. P. & Kane, K. P. *Bacterial Modulation of Antigen Processing and Presentation*. (2000).
110. Bäckhed, F. *et al.* Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host Microbe* **17**, 690–703 (2015).
111. Zegarra-Ruiz, D. F. *et al.* Thymic development of gut-microbiota-specific T cells. *Nature* **594**, 413–417 (2021).
112. Lathrop, S. K. *et al.* Peripheral education of the immune system by colonic commensal microbiota. *Nature* **478**, 250–254 (2011).
113. Sorini, C., Cardoso, R. F., Gagliani, N. & Villablanca, E. J. Commensal bacteria-specific CD4+T cell responses in health and disease. *Frontiers in Immunology* vol. 9 Preprint at https://doi.org/10.3389/fimmu.2018.02667 (2018).
114. Noble, A. *et al.* Altered immunity to microbiota, B cell activation and depleted γδ/resident memory T cells in colorectal cancer. *Cancer Immunology, Immunotherapy* **71**, 2619–2629 (2022).
115. Duchmann, R. *et al. T Cell Specificity and Cross Reactivity towards Enterobacteria, Bacteroides, Bifidobacterium, and Antigens from Resident Intestinal Flora in Humans*. http://gut.bmj.com/.

116. Hegazy, A. N. *et al.* Circulating and Tissue-Resident CD4+ T Cells With Reactivity to Intestinal Microbiota Are Abundant in Healthy Individuals and Function Is Altered During Inflammation. *Gastroenterology* **153**, 1320-1337.e16 (2017).
117. Duhen, T. & Campbell, D. J. IL-1β Promotes the Differentiation of Polyfunctional Human CCR6+CXCR3+ Th1/17 Cells That Are Specific for Pathogenic and Commensal Microbes. *The Journal of Immunology* **193**, 120–129 (2014).
118. Cassotta, A. *et al.* Broadly reactive human CD4+ T cells against Enterobacteriaceae are found in the naïve repertoire and are clonally expanded in the memory repertoire. *Eur J Immunol* **51**, 648–661 (2021).
119. Ruf, B., Greten, T. F. & Korangy, F. Innate lymphoid cells and innate-like T cells in cancer — at the crossroads of innate and adaptive immunity. *Nature Reviews Cancer* vol. 23 351–371 Preprint at https://doi.org/10.1038/s41568-023-00562-w (2023).
120. Pellicci, D. G., Koay, H. F. & Berzins, S. P. Thymic development of unconventional T cells: how NKT cells, MAIT cells and γδ T cells emerge. *Nature Reviews Immunology* vol. 20 756–770 Preprint at https://doi.org/10.1038/s41577-020-0345-y (2020).
121. Constantinides, M. G. & Belkaid, Y. Early-life imprinting of unconventional T cells and tissue homeostasis. *Science* vol. 374 Preprint at https://doi.org/10.1126/science.abf0095 (2021).
122. Gold, M. C. *et al.* Human mucosal associated invariant T cells detect bacterially infected cells. *PLoS Biol* **8**, (2010).
123. Hartmann, N. *et al.* Riboflavin metabolism variation among clinical isolates of streptococcus pneumoniae results in differential activation of mucosal-Associated invariant T cells. *Am J Respir Cell Mol Biol* **58**, 767–776 (2018).
124. Kjer-Nielsen, L. *et al.* MR1 presents microbial vitamin B metabolites to MAIT cells. *Nature* **491**, 717–723 (2012).
125. Tastan, C. *et al.* Tuning of human MAIT cell activation by commensal bacteria species and MR1-dependent T-cell presentation. *Mucosal Immunol* **11**, 1591–1605 (2018).
126. Li, S. *et al.* Human Tumor-Infiltrating MAIT Cells Display Hallmarks of Bacterial Antigen Recognition in Colorectal Cancer. *Cell Rep Med* **1**, (2020).
127. Carding, S. R. & Egan, P. J. γδ T cells: Functional plasticity and heterogeneity. *Nature Reviews Immunology* vol. 2 336–345 Preprint at https://doi.org/10.1038/nri797 (2002).
128. Zhou, Q. H., Wu, F. T., Pang, L. T., Zhang, T. B. & Chen, Z. A Role of γδT cells in liver diseases and its relationship with intestinal microbiota. *World Journal of Gastroenterology* vol. 26 2559–2569 Preprint at https://doi.org/10.3748/WJG.V26.I20.2559 (2020).
129. Simões, A. E., Di Lorenzo, B. & Silva-Santos, B. Molecular determinants of target cell recognition by human γδ T cells. *Frontiers in Immunology* vol. 9 Preprint at https://doi.org/10.3389/fimmu.2018.00929 (2018).
130. Chien, Y. H., Meyer, C. & Bonneville, M. γδ T cells: First line of defense and beyond. *Annual Review of Immunology* vol. 32 121–155 Preprint at https://doi.org/10.1146/annurev-immunol-032713-120216 (2014).

131. Crosby, C. M. & Kronenberg, M. Tissue-specific functions of invariant natural killer T cells. *Nature Reviews Immunology* vol. 18 559–574 Preprint at https://doi.org/10.1038/s41577-018-0034-2 (2018).
132. Bessell, C. A. *et al.* Commensal bacteria stimulate antitumor responses via T cell cross-reactivity. *JCI Insight* **5**, (2020).
133. Fluckiger, A. *et al.* Cross-reactivity between tumor MHC class I–restricted antigens and an enterococcal bacteriophage. *Science (1979)* **369**, 936–942 (2020).
134. Kalaora, S. *et al.* Identification of bacteria-derived HLA-bound peptides in melanoma. *Nature* **592**, 138–143 (2021).
135. Naghavian, R. *et al.* Microbial peptides activate tumour-infiltrating lymphocytes in glioblastoma. *Nature* **617**, 807–817 (2023).
136. Kirkham, L.-A. S., Corscadden, K. J., Wiertsema, S. P., Currie, A. J. & Richmond, P. C. *A Practical Method for Preparation of Pneumococcal and Nontypeable Haemophilus Influenzae Inocula That Preserves Viability and Immunostimulatory Activity*. http://www.biomedcentral.com/1756-0500/6/522 (2013).
137. Messi, M. *et al.* Memory and flexibility of cytokine gene expression as separable properties of human TH1 and TH2 lymphocytes. *Nature Immunology* vol. 4 78–86 Preprint at https://doi.org/10.1038/ni872 (2003).
138. Pizzolato, G. *et al.* Single-cell RNA sequencing unveils the shared and the distinct cytotoxic hallmarks of human TCRVδ1 and TCRVδ2 γδ T lymphocytes. *Proc Natl Acad Sci U S A* **116**, 11906–11915 (2019).
139. Wu, Z. *et al.* CD3+CD4-CD8- (Double-Negative) T Cells in Inflammation, Immune Disorders and Cancer. *Frontiers in Immunology* vol. 13 Preprint at https://doi.org/10.3389/fimmu.2022.816005 (2022).
140. Kupz, A. *et al.* NLRC4 inflammasomes in dendritic cells regulate noncognate effector function by memory CD8 + T cells. *Nat Immunol* **13**, 162–169 (2012).
141. Chen, X., Wang, D. & Zhu, X. Application of double-negative T cells in haematological malignancies: recent progress and future directions. *Biomarker Research* vol. 10 Preprint at https://doi.org/10.1186/s40364-022-00360-w (2022).
142. Yao, J. *et al.* Human double negative T cells target lung cancer via ligand-dependent mechanisms that can be enhanced by IL-15. *J Immunother Cancer* **7**, (2019).
143. Chen, J., Hu, P., Wu, G. & Zhou, H. Antipancreatic cancer effect of DNT cells and the underlying mechanism. *Pancreatology* **19**, 105–113 (2019).
144. Lu, Y. *et al.* Double-negative T Cells Inhibit Proliferation and Invasion of Human Pancreatic Cancer Cells in Co-culture. *Anticancer Res* **39**, 5911–5918 (2019).
145. Vasic, D. *et al. Allogeneic Double-Negative CAR-T Cells Inhibit Tumor Growth without off-Tumor Toxicities*. *Sci. Immunol* vol. 7 https://www.science.org (2022).
146. Pelka, K. *et al.* Spatially organized multicellular immune hubs in human colorectal cancer. *Cell* **184**, 4734-4752.e20 (2021).
147. Mlecnik, B. *et al.* Integrative Analyses of Colorectal Cancer Show Immunoscore Is a Stronger Predictor of Patient Survival Than Microsatellite Instability. *Immunity* **44**, 698–711 (2016).

148. Picard, E., Verschoor, C. P., Ma, G. W. & Pawelec, G. Relationships Between Immune Landscapes, Genetic Subtypes and Responses to Immunotherapy in Colorectal Cancer. *Frontiers in Immunology* vol. 11 Preprint at https://doi.org/10.3389/fimmu.2020.00369 (2020).
149. Nguyen, M. *et al.* An update on the use of immunotherapy in patients with colorectal cancer. *Expert Review of Gastroenterology and Hepatology* vol. 15 291–304 Preprint at https://doi.org/10.1080/17474124.2021.1845141 (2021).
150. Chida, K. *et al.* Transcriptomic Profiling of MSI-H/dMMR Gastrointestinal Tumors to Identify Determinants of Responsiveness to Anti–PD-1 Therapy. *Clinical Cancer Research* **28**, 2110–2117 (2022).
151. Høgh, R. I. *et al.* Metabolism of short-chain fatty acid propionate induces surface expression of NKG2D ligands on cancer cells. *FASEB Journal* **34**, 15531–15546 (2020).
152. McGilvray, R. W. *et al.* NKG2D ligand expression in human colorectal cancer reveals associations with prognosis and evidence for immunoediting. *Clinical Cancer Research* **15**, 6993–7002 (2009).
153. Watson, N. F. S. *et al.* Expression of the stress-related MHC class I chain-related protein MICA is an indicator of good prognosis in colorectal cancer patients. *Int J Cancer* **118**, 1445–1452 (2006).
154. Moesta, A. K., Li, X. Y. & Smyth, M. J. Targeting CD39 in cancer. *Nature Reviews Immunology* vol. 20 739–755 Preprint at https://doi.org/10.1038/s41577-020-0376-4 (2020).
155. Fang, L. *et al.* Targeting late-stage non-small cell lung cancer with a combination of DNT cellular therapy and PD-1 checkpoint blockade. *Journal of Experimental and Clinical Cancer Research* **38**, (2019).

Appendix

Gut

Direct toll-like receptor triggering in colorectal cancer-associated stromal cells elicits immunostimulatory properties leading to enhanced immune cell recruitment

Journal:	*Gut*
Manuscript ID	gutjnl-2023-331759.R1
Article Type:	Letter
Date Submitted by the Author:	n/a
Complete List of Authors:	Djordjevic, Julija; EOC Translational Research Laboratory, Department of Surgery; Università della Svizzera italiana, Faculty of Biomedical Sciences Cisneros Romero, Nubia ; University of Basel Faculty of Science, Department of Biomedicine Cascione, Luciano; Bioinformatics Core Unit, Institute of Oncology Research (IOR), Faculty of Biomedical Sciences, Università della Svizzera italiana Mele, Valentina; University of Basel, Department of Biomedicine Cremonesi, Eleonora; University of Basel Faculty of Medicine, Department of Biomedicine; University Hospital Basel Sorrenti, Elisa; EOC Translational Research Laboratory, Department of Surgery; Università della Svizzera italiana, Faculty of Biomedical Sciences Basso, Camilla; EOC Translational Research Laboratory, Department of Surgery; Università della Svizzera italiana, Faculty of Biomedical Sciences Cianfarani, Agnese; EOC Translational Research Laboratory, Department of Surgery; Repubblica e Cantone Ticino Ente Ospedaliero Cantonale, Department of Surgery Roesel, Raffaello; Repubblica e Cantone Ticino Ente Ospedaliero Cantonale, Department of Surgery; Università della Svizzera italiana, Faculty of Biomedical Sciences Galafassi, Jacopo; Repubblica e Cantone Ticino Ente Ospedaliero Cantonale, Department of Surgery Majno-Hurst, Pietro; Università della Svizzera italiana, Faculty of Biomedical Sciences; Repubblica e Cantone Ticino Ente Ospedaliero Cantonale, Department of Surgery Spagnoli, Giulio; National Research Council, Institute of Translational Pharmacology Christoforidis, Dimitrios; Università della Svizzera italiana, Faculty of Biomedical Sciences; Repubblica e Cantone Ticino Ente Ospedaliero Cantonale, Department of Surgery Iezzi, Giandomenica; EOC Translational Research Laboratory, Department of Surgery; Università della Svizzera italiana, Faculty of Biomedical Sciences
Keywords:	COLORECTAL CANCER, MYOFIBROBLASTS, CHEMOKINES, INTESTINAL

	BACTERIA, IMMUNE RESPONSE

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Letter to Editors

Direct toll-like receptor triggering in colorectal cancer-associated stromal cells elicits immunostimulatory properties leading to enhanced immune cell recruitment

Julija Djordjevic[1,2], Nubia Sarahi Cisneros Romero[3], Luciano Cascione[4], Valentina Mele[3], Eleonora Cremonesi[3], Elisa Sorrenti[1,2], Camilla Basso[1,2], Agnese Cianfarani[1,5], Raffaello Roesel[2,5],Jacopo Galafassi[5], Pietro E. Majno-Hurst[2,5], Giulio C. Spagnoli[6], Dimitrios Christoforidis[2,5,7], Giandomenica Iezzi[1,2].

1 Laboratory of Translational Surgical Research, Ente Ospedaliero Cantonale, Bellinzona, Switzerland

2 Faculty of Biomedical Sciences, Università della Svizzera Italiana, Lugano, Switzerland

3 Department of Biomedicine, University of Basel, Basel, Switzerland

4 Institute of Oncological Research, Bellinzona, Switzerland

5 Department of Surgery, Ente Ospedaliero Cantonale, Lugano, Switzerland

6 Institute of Translational Pharmacology, National Research Council, Rome, Italy

7 Department of Visceral Surgery, CHUV, University of Lausanne, Switzerland

Keywords: Colorectal cancer, tumor-associated stromal cells, gut microbiota, toll-like receptors, chemokines

We have read with interest the study by Corry et al. (1) reporting that upregulation of IFNγ, IFNα, and STAT-1 response pathways, downstream of double-stranded (ds) RNA and/or viral responses, is associated with favorable prognosis in stroma-rich colorectal cancers (CRCs). Furthermore, in vitro stimulation of myeloid cells with poly(I:C), a synthetic dsRNA viral mimetic and toll-like receptor 3 (TLR3) agonist, effectively induces expression of STAT1 and its target genes. Importantly, administration of poly(I:C) enhances immune cell infiltration of liver metastases and reduces metastatic tumor burden in CRC murine model (1).

Thus, TLR3 targeting may represent a novel therapeutic option in patients with stroma-rich CRC. However, TLRs are expressed by a variety of cell types, besides myeloid cells (2). We previously reported that TLR triggering on tumor cells enhances their chemokine production

capacity ultimately favoring intratumoral immune cell recruitment (3). Whether TASCs may also be directly targeted by TLR3- or other TLR-agonists remains to be addressed.

Prompted by this question, we investigated TLR expression profiles of CRC-derived TASCs in publicly available single-cell RNA seq databases (4, Table S1) and upon in vitro expansion (5, Table S2), and assessed their capacity to respond to microbial stimuli.

A substantial fraction of CRC-derived TASCs expressed heterogeneous levels of TLR3, TLR4, and TLR5, while <1% of cells displayed expression of other TLRs (Figure 1A). Comparable expression patterns were also observed when considering only distal, mismatch repair (MMR) proficient CRCs (Figure S1), most likely accounting for stroma-rich consensus molecular subtype 4 (CMS4). Accordingly, in expanded TASCs (Figure S2), TLR3, TLR4, TLR5, and TLR6 gene expression were detected in 8, 10, 5, and 11 out of 12 samples, respectively (Figure 1B). Expression of TLR1 and TLR2 genes was also observed in two and three samples, respectively. No expression of TLR7, TLR8, TLR9, and TLR10 genes was detected (Figure 1B).

TLR functionality was assessed on expanded TASCs, based on IL-6 production at baseline and upon stimulation by TLR agonists. Consistent with TLR expression profiles, stimulation with poly(I:C) strongly upregulated IL-6 expression at both gene and protein levels at all concentrations tested (Figure 1C-F). Similarly, exposure to LPS and flagellin, TLR4, and TLR5 ligands, respectively, resulted in higher expression of IL-6 gene and protein, although to lower extents than poly (I:C). Instead, no significant response was observed upon stimulation with the TLR2/TLR6 agonist FSL-1, the TLR2 agonist PGN, the TLR7/8 agonist imiquimod, and the TLR9 agonist ODNs.

Interestingly, TLR triggering in TASCs resulted in differential modulation of chemokine expression patterns (Figure 2). Expression of myeloid cell-recruiting chemokines, including CCL2, CXCL1, CXCL2, CXCL5, CXCL6, and CXCL8, already detectable at baseline, was boosted at both gene and protein levels. However, while LPS- and flagellin-mediated effects largely varied across different TASC preparations, possibly reflecting heterogeneous expression levels of TLR4 and TLR5, stimulation by poly (I:C) strongly enhanced chemokine production in all TASC samples (Figure 2 A,C). Remarkably, only poly (I:C) induced expression of T cell-recruiting chemokine genes, including CCL5, CXCL9, CXCL10, and CXCL11, whereas other TLR agonists showed poor or no capacity (Figure 2 B, D). Consistently, supernatants from poly(I:C)-stimulated TASCs significantly induced T cell migration in vitro, whereas those from LPS- or flagellin-stimulated TASCs displayed negligible effects (Figure 2E)

Our findings demonstrate that human CRC-associated TASCs express functional TLRs, enabling them to directly sense microbial stimuli and potential therapeutic TLR ligands. Importantly, we provide further evidence in favor of TLR3 targeting as the most powerful approach to enhance chemokine production in various cell types of CRC microenvironment, including TASCs, in addition to myeloid cells (1) and tumor cells (3), thereby effectively promoting recruitment of beneficial immune cells into tumor tissues.

References

1. Corry SM, McCorry AM, Lannagan TR, et al. Activation of innate-adaptive immune machinery by poly(I:C) exposes a therapeutic vulnerability to prevent relapse in stroma-rich colon cancer. Gut 2022:10.1136/gutjnl-2021-326183.
2. Burgueño JF, Abreu MT. Epithelial Toll-like receptors and their role in gut homeostasis and disease. Nat Rev Gastroenterol Hepatol. 2020: 10.1038/s41575-019-0261-4.
3. Cremonesi E, Governa V, Glaus Garzon JF, et al. Gut microbiota modulate T cell trafficking into human colorectal cancer. Gut 2018: 10.1136/gutjnl-2016-313498.
4. Pelka K, Hofree M, Chen JH, et al. Spatially organized multicellular immune hubs in human colorectal cancer. Cell 2021: 10.1016/j.cell.2021.08.003.
5. Mele V, Basso C, Governa V, et al. Identification of TPM2 and CNN1 as Novel Prognostic Markers in Functionally Characterized Human Colon Cancer-Associated Stromal Cells. Cancers 2022: 10.3390/cancers14082024.
6. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. Nat Med 2015: 21:10.1038/nm.3967

Correspondence to Prof. Giandomenica Iezzi Laboratory of Translational Surgical Research, Ente Ospedaliero Cantonale, Bellinzona, Switzerland, Faculty of Biomedical Sciences, Università della Svizzera Italiana, Lugano, Switzerland, giandomenica.iezzi@eoc.ch

Acknowledgments We thank the patients and their families for their consent to use their biological samples for this study and our colleagues at the Department of Surgery, Basel University Hospital for helping in the initial phase of the project. We are also thankful to Dr. Valentina Cecchinato and Prof. Mariagrazia Uguccioni, Institute of Research in Biomedicine

(IRB), Bellinzona, for their help with migration assays, and to Dr. Chiara Arrigoni and Prof. Matteo Moretti for their support with TASC culture.

Contributors Conceptualisation: JD, LC, GCS, DC, GI. Methodology: JD, NSCR, VM, LC, ES, CB, AC, RR, JG. Visualization and original draft: JD, NSCR, GCS, GI. Supervision: DC, GI. Review and editing: JD, GCS, PEMH, DC, GI. Guarantor and overall supervision: GI

Funding This work was supported by Kurt und Senta Herrmann Stiftung, Krebsliga Beider Basel, Stiftung für Krebsbekämpfung, Gebert Rüf Stiftung, and EOC AFRI Senior Research Grant to GI.

Competing interests: None to declare

Patient consent for publication not applicable

Ethics approval Thus study involves the use of human samples and was approved by local ethical authorities (Ethikkommission Nordwest und Zentralschweiz, EKNZ, study protocol n. 2014-388, and Comitato Etico Cantonale Ticino, 2020-00437 I CE 3598). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer-reviewed.

Figure legends

Figure 1. CRC-associated TASCs do express functional TLRs. A. Data relative to cells classified as "Fibroblasts"(n=3032) were retrieved from the largest publicly available scRNA seq database of human primary CRC (GSE178341) (4). This image shows the normalized log-expression of TLR1-10 genes, detected in 610 cells. **B.** TASCs were isolated from human primary CRC samples (n=12), as previously described (5). Following short in vitro expansion, expression of stromal cell markers was confirmed upon phenotypical analysis by flow cytometry (see Supplementary Figure 1). Gene expression of TLR1-10 was assessed on expanded TASC preparations (n=12) by quantitative qRT-PCR, using GAPDH as housekeeping gene. Each dot corresponds to one TASC preparation. Those further used for functional testing are identified by a specific color. **C.** To identify optimal stimulatory conditions, TASCs from one representative preparation were incubated with a panel of TLR agonists, including the TLR2-ligand peptidoglycan (PGN), the TLR3-ligand polyinosinic-polycytidylic acid (poly(I:C), the TLR4-ligand Lipopolysaccharide (LPS), the TLR-ligand flagellin (FLAG), the TLR2/TLR6 ligand and synthetic diacylated lipoprotein FSL-1,the TLR7/8 ligand Imiquimod (IMQ) and the TLR9-ligands CpG oligonucleotide type A (ODN 2216), and CpG oligonucleotide type B (ODN 2006), at the indicated titrated concentrations. After 4 hours cells were collected, and following RNA extraction, IL-6 gene expression was assessed by quantitative qRT-PCR, using GAPDH as housekeeping gene. Cumulative data from 3 independent experiments are shown. **D.** TASCs from five different preparations (each identified by a specific color) were stimulated with poly (I:C), LPS, and FLAG at the indicated concentrations. After 4 hours, IL-6 expression was assessed as detailed above**. E, F.** TASCs were stimulated with TLR agonists as described above. After overnight incubation culture supernatants were collected and IL-6 release was assessed by ELISA assays. **E.** Cumulative data from 3 independent experiments performed with one single TASC preparation. **F.** Data from five different TASC preparations, each identified by a specific color. Statistical significance of observed differences was assessed by Mann-Whitney (C,E) or Friedmann (D,F) tests. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

Figure 2. **TLR triggering induces chemokine production in TASCs** . CRC-derived TASCs were stimulated with poly (I:C), LPS, and flagellin (FLAG) at the indicated concentrations, as described in Figure 1. **A,B.** After 4 hours cells were collected, and following RNA extraction, chemokine gene expression was assessed by quantitative qRT-PCR, using GAPDH as housekeeping gene. Cumulative data (means ± SD) are reported . **C,D.** After an overnight incubation chemokine release in culture supernatants was assessed by Legendplex assay. Data from individual TASC preparations, each identified by a specific color, are reported. Means are indicated by columns. **E.** T cells were isolated from healthy donors and their capacity to migrate towards TASC supernatants was evaluated. Numbers of migrated cells (total T cells, CD8+, or CD4+) towards individual TASC preparations, each identified by a specific color, are reported. Statistical significance of observed differences was assessed by Kruskal-Wallis (A,B) or Friedmann (C-E) tests. *p<0.05, **p<0.01, ***p<0.001.

Letter to Editors

Direct toll-like receptor triggering in colorectal cancer-associated stromal cells elicits immunostimulatory properties leading to enhanced immune cell recruitment

Julija Djordjevic[1,2], Nubia Sarahi Cisneros Romero[3], Luciano Cascione[4], Valentina Mele[3], Eleonora Cremonesi[3], Elisa Sorrenti[1,2], Camilla Basso[1,2], Agnese Cianfarani[1,5], Raffaello Roesel[2,5],Jacopo Galafassi[5], Pietro E. Majno-Hurst[2,5], Giulio C. Spagnoli[6], Dimitrios Christoforidis[2,5,7], Giandomenica Iezzi[1,2].

1 Laboratory of Translational Surgical Research, Ente Ospedaliero Cantonale, Bellinzona, Switzerland

2 Faculty of Biomedical Sciences, Università della Svizzera Italiana, Lugano, Switzerland

3 Department of Biomedicine, University of Basel, Basel, Switzerland

4 Institute of Oncological Research, Bellinzona, Switzerland

5 Department of Surgery, Ente Ospedaliero Cantonale, Lugano, Switzerland

6 Institute of Translational Pharmacology, National Research Council, Rome, Italy

7 Department of Visceral Surgery, CHUV, University of Lausanne, Switzerland

Keywords: Colorectal cancer, tumor-associated stromal cells, gut microbiota, toll-like receptors, chemokines

~~Non-transformed cells play key roles in colorectal cancer (CRC) microenvironment. While immune cell infiltration correlates with improved survival, tumor-associated mesenchymal stromal cells (TASCs) abundance predicts poor prognosis (1,2). Indeed, CRCs characterized by high stromal component and low immune cell infiltration, show the worst prognosis, with relapse rate of about 60% (2).~~

We have read with interest ~~T~~the study by Corry et al. (1~~3~~) report~~inged~~ that upregulation of IFNγ, IFNα, and STAT-1 response pathways, downstream of double-stranded (ds) RNA and/or viral responses, is associated with favorable prognosis in stroma-rich colorectal cancers (CRCs).

Furthermore, in vitro stimulation of myeloid cells with poly(I:C), a synthetic dsRNA viral mimetic and toll-like receptor 3 (TLR3) agonist, effectively induces expression of STAT1 and its target genes. Importantly, administration of poly(I:C) enhances immune cell infiltration of liver metastases and reduces metastatic tumor burden in CRC murine model (13).

Thus, TLR3 targeting may represent a novel therapeutic option in patients with stroma-rich CRC. However, TLRs are expressed by a variety of cell types, besides myeloid cells (24). We previously reported that TLR triggering on tumor cells enhances their chemokine production capacity ultimately favoring intratumoral immune cell recruitment (35). Whether TASCs may also be directly targeted by TLR3- or other TLR-agonists remains to be addressed.

Prompted by this question, we investigated TLR expression profiles of CRC-derived TASCs in publicly available single-cell RNA seq databases (46, Table S1) and upon in vitro expansion (5, Table S27), and assessed their capacity to respond to microbial stimuli.

A substantial fraction of CRC-derived TASCs expressed heterogeneous levels of TLR3, TLR4, and TLR5, while <1% of cells displayed expression of other TLRs (Figure 1A). Comparable expression patterns were also observed when considering only distal, mismatch repair (MMR) proficient CRCs (Figure S1), most likely accounting for stroma-rich consensus molecular subtype 4 (CMS4). Accordingly, in expanded TASCs (Figure S2), TLR3, TLR4, TLR5, and TLR6 gene expression were detected in 8, 10, 5, and 11 out of 12 samples, respectively (Figure 1B). Expression of TLR1 and TLR2 genes was also observed in two and three samples, respectively. No expression of TLR7, TLR8, TLR9, and TLR10 genes was detected (Figure 1B).

TLR functionality was assessed on expanded TASCs, based on IL-6 production at baseline and upon stimulation by TLR agonists. Consistent with TLR expression profiles, stimulation with poly(I:C) strongly upregulated IL-6 expression at both gene and protein levels at all concentrations tested (Figure 1C-F). Similarly, exposure to LPS and flagellin, TLR4, and TLR5 ligands, respectively, resulted in higher expression of IL-6 gene and protein, although to lower extents than poly (I:C). Instead, no significant response was observed upon stimulation with the TLR2/TRLR6 agonist FSL-1, the TLR2 agonist PGN, the TLR7/8 agonist imiquimod, and the TLR9 agonist ODNs.

Interestingly, TRLR triggering in TASCs resulted in differential modulation of chemokine expression patterns (Figure 2). Expression of myeloid cell-recruiting chemokines, including CCL2, CXCL1, CXCL2, CXCL5, CXCL6, and CXCL8, already detectable at baseline, was boosted at both gene and protein levels. However, while LPS- and flagellin-mediated effects largely

varied across different TASC preparations, possibly reflecting heterogeneous expression levels of TLR4 and TLR5, stimulation by poly (I:C) strongly enhanced chemokine production in all TASC samples (Figure 2 A,CB). Remarkably, only poly (I:C) induced expression of T cell-recruiting chemokine genes, including CCL5, CXCL9, CXCL10, and CXCL11, whereas other TLR agonists showed poor or no capacity (Figure 2 BC, D). Consistently, supernatants from poly(I:C)-stimulated TASCs significantly induced T cell migration in vitro, whereas those from LPS- or flagellin-stimulated TASCs displayed negligible effects (Figure 2E) Expression of CXCL12, detected in TASCs at basal conditions, remained unmodified upon TLR agonist stimulation.

Our findings demonstrate that human CRC-associated TASCs express functional TLRs, enabling them to directly sense microbial stimuli and potential therapeutic TLR ligands. Importantly, we provide further evidence in favor of TLR3 targeting as the most powerful approach to enhance chemokine production in various cell types of CRC microenvironment, including TASCs, in addition to myeloid cells (13) and tumor cells (36), thereby effectively promoting recruitment of beneficial immune cells into tumor tissues.

References

1. Fridman WH, Pagès F, Sautes-Fridman C, et al. The immune contexture in human tumours: impact on clinical outcome. Nat Rev Cancer 2012: 10.1038/nrc3245.
2. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. Nat Med 2015: 21:10.1038/nm.3967
3. Corry SM, McCorry AM, Lannagan TR, et al. Activation of innate-adaptive immune machinery by poly(I:C) exposes a therapeutic vulnerability to prevent relapse in stroma-rich colon cancer. Gut 2022:10.1136/gutjnl-2021-326183.
4. Burgueño JF, Abreu MT. Epithelial Toll-like receptors and their role in gut homeostasis and disease. Nat Rev Gastroenterol Hepatol. 2020: 10.1038/s41575-019-0261-4.
5. Cremonesi E, Governa V, Glaus Garzon JF, et al. Gut microbiota modulate T cell trafficking into human colorectal cancer. Gut 2018: 10.1136/gutjnl-2016-313498.
6. Pelka K, Hofree M, Chen JH, et al. Spatially organized multicellular immune hubs in human colorectal cancer. Cell 2021: 10.1016/j.cell.2021.08.003.

5. Mele V, Basso C, Governa V, et al. Identification of TPM2 and CNN1 as Novel Prognostic Markers in Functionally Characterized Human Colon Cancer-Associated Stromal Cells. Cancers 2022: 10.3390/cancers14082024.
6. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. Nat Med 2015: 21:10.1038/nm.3967

Correspondence to Prof. Giandomenica Iezzi Laboratory of Translational Surgical Research, Ente Ospedaliero Cantonale, Bellinzona, Switzerland, Faculty of Biomedical Sciences, Università della Svizzera Italiana, Lugano, Switzerland, giandomenica.iezzi@eoc.ch

Acknowledgments We thank the patients and their families for their consent to use their biological samples for this study and our colleagues at the Department of Surgery, Basel University Hospital for helping in the initial phase of the project. We are also thankful to Dr. Valentina Cecchinato and Prof. Mariagrazia Uguccioni, Institute of Research in Biomedicine (IRB), Bellinzona, for their help with migration assays, and to Dr. Chiara Arrigoni and Prof. Matteo Moretti for their support with TASC culture.

Contributors Conceptualisation: JD, LC, GCS, DC, GI. Methodology: JD, NSCR, VM, LC, ES, CB, AC, RR, JG. Visualization and original draft: JD, NSCR, GCS, GI. Supervision: DC, GI. Review and editing: JD, GCS, PEMH, DC, GI. Guarantor and overall supervision: GI

Funding This work was supported by Kurt und Senta Herrmann Stiftung, Krebsliga Beider Basel, Stiftung für Krebsbekämpfung, Gebert Rüf Stiftung, and EOC AFRI Senior Research Grant to GI.

Competing interests: None to declare

Patient consent for publication not applicable

Ethics approval Thus study involves the use of human samples and was approved by local ethical authorities (Ethikkommission Nordwest und Zentralschweiz, EKNZ, study protocol n. 2014-388, and Comitato Etico Cantonale Ticino, 2020-00437 I CE 3598). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer-reviewed.

Figure legends

Figure 1. CRC-associated TASCs do express functional TLRs. A. Data relative to cells classified as "Fibroblasts"(n=3032) were retrieved from the largest publicly available scRNA seq database of human primary CRC (GSE178341) (46). This image shows the normalized log-expression of TLR1-10 genes, detected in 610 cells. **B.** TASCs were isolated from human primary CRC samples (n=12), as previously described (56). Following short in vitro expansion, expression of stromal cell markers was confirmed upon phenotypical analysis by flow cytometry (see Supplementary Figure 1). Gene expression of TLR1-10 was assessed on expanded TASC preparations (n=12) by quantitative qRT-PCR, using GAPDH as housekeeping gene. Each dot corresponds to one TASC preparation. Those further used for functional testing are identified by a specific color. **C.** To identify optimal stimulatory conditions, TASCs from one representative preparation were incubated with a panel of TLR agonists, including the TLR2-ligand peptidoglycan (PGN), the TLR3-ligand polyinosinic-polycytidylic acid (poly(I:C), the TLR4-ligand Lipopolysaccharide (LPS), the TLR-ligand flagellin (FLAG), the TLR2/TLR6 ligand and synthetic diacylated lipoprotein FSL-1,the TLR7/8 ligand Imiquimod (IMQ) and the TLR9-ligands CpG oligonucleotide type A (ODN 2216), and CpG oligonucleotide type B (ODN 2006), at the indicated titrated concentrations. After 4 hours cells were collected, and following RNA extraction, IL-6 gene expression was assessed by quantitative qRT-PCR, using GAPDH as housekeeping gene. Cumulative data from 3 independent experiments are shown. **D.** TASCs from five different preparations (each identified by a specific color) were stimulated with poly (I:C), LPS, and FLAG at the indicated concentrations. After 4 hours, IL-6 expression was assessed as detailed above**. E, F.** TASCs were stimulated with TLR agonists as described above. After ~~an after~~ overnight incubation culture supernatants were collected and IL-6 release was assessed by ELISA assays. **E.** Cumulative data from 3 independent experiments performed with one single TASC preparation~~s~~. **F.** Data from five different TASC preparations, each identified by a specific color. Statistical significance of observed differences was assessed by Mann-Whitney (C,E) or Friedmann (D,F) tests. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

Figure 2. **TLR triggering induces chemokine production in TASCs** . CRC-derived TASCs were stimulated with poly (I:C), LPS, and flagellin (FLAG) at the indicated concentrations, as described in Figure 1. **A,BC.** After 4 hours cells were collected, and following RNA extraction, chemokine gene expression was assessed by quantitative qRT-PCR, using GAPDH as housekeeping gene. Cumulative data (means ± SD) are reported . **CB,D.** After an overnight incubation chemokine release in culture supernatants was assessed by Legendplex assay. Data from individual TASC preparations, each identified by a specific color, are reported. Means are indicated by columns. **E.** T cells were isolated from healthy donors and their capacity to migrate towards TASC supernatants was evaluated. Numbers of migrated cells (total T cells, CD8+, or CD4+) towards individual TASC preparations, each identified by a specific color, are reported. Statistical significance of observed differences was assessed by Kruskal-Wallis (A,BC) or Friedmann (CB,D-E) tests. *p<0.05, **p<0.01, ***p<0.001.

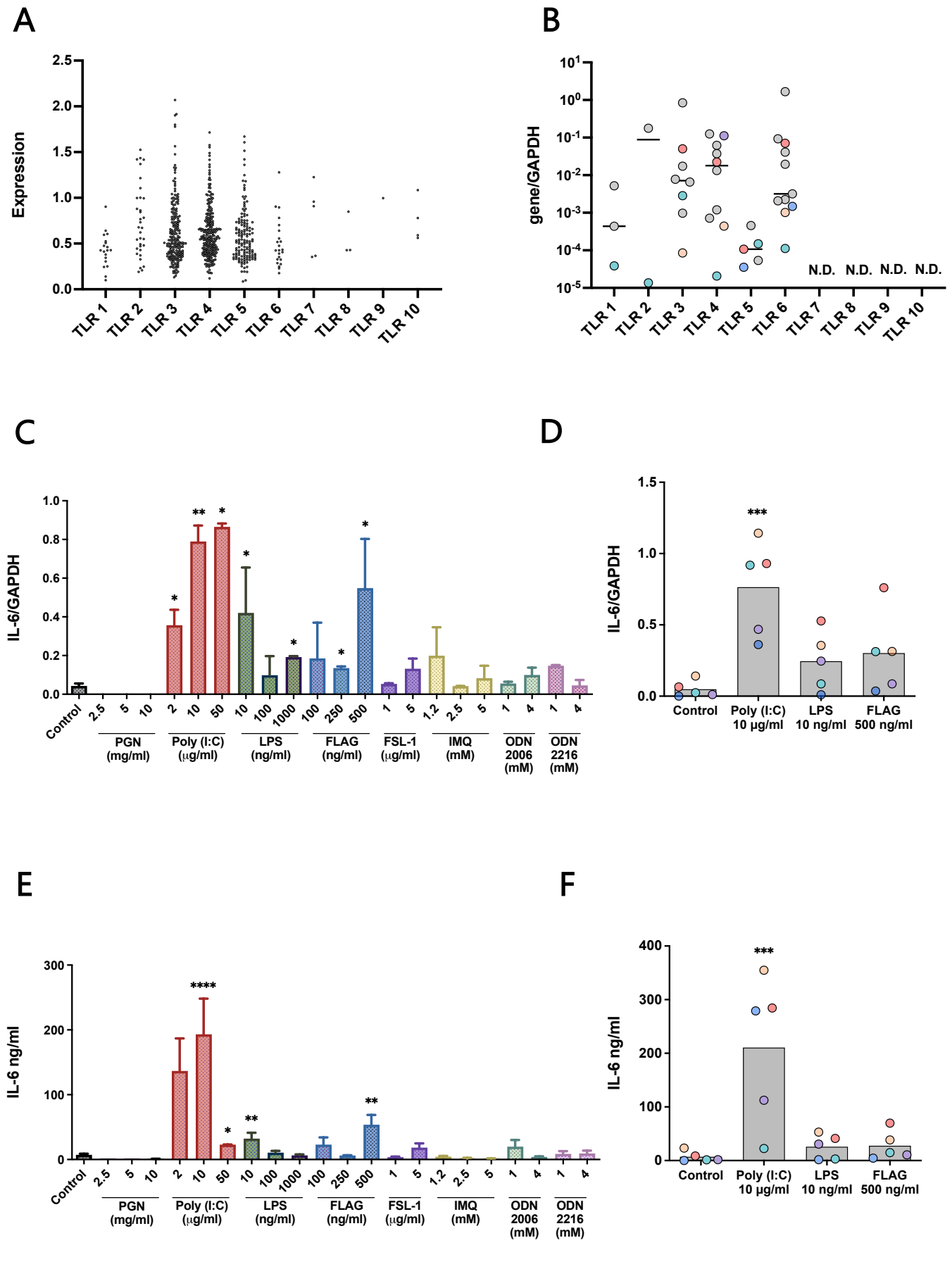

A
Expression
TLR 1
TLR 2
TLR 3
TLR 4
TLR 5
TLR 6
TLR 7
TLR 8
TLR 9
TLR 10
B
gene/GAPDH
N.D.
C
IL-6/GAPDH
Control
PGN (mg/ml)
Poly (I:C) (μg/ml)
LPS (ng/ml)
FLAG (ng/ml)
FSL-1 (μg/ml)
IMQ (mM)
ODN 2006 (mM)
ODN 2216 (mM)
D
IL-6/GAPDH
Control
Poly (I:C) 10 μg/ml
LPS 10 ng/ml
FLAG 500 ng/ml
E
IL-6 ng/ml
F
IL-6 ng/ml

Figure 1

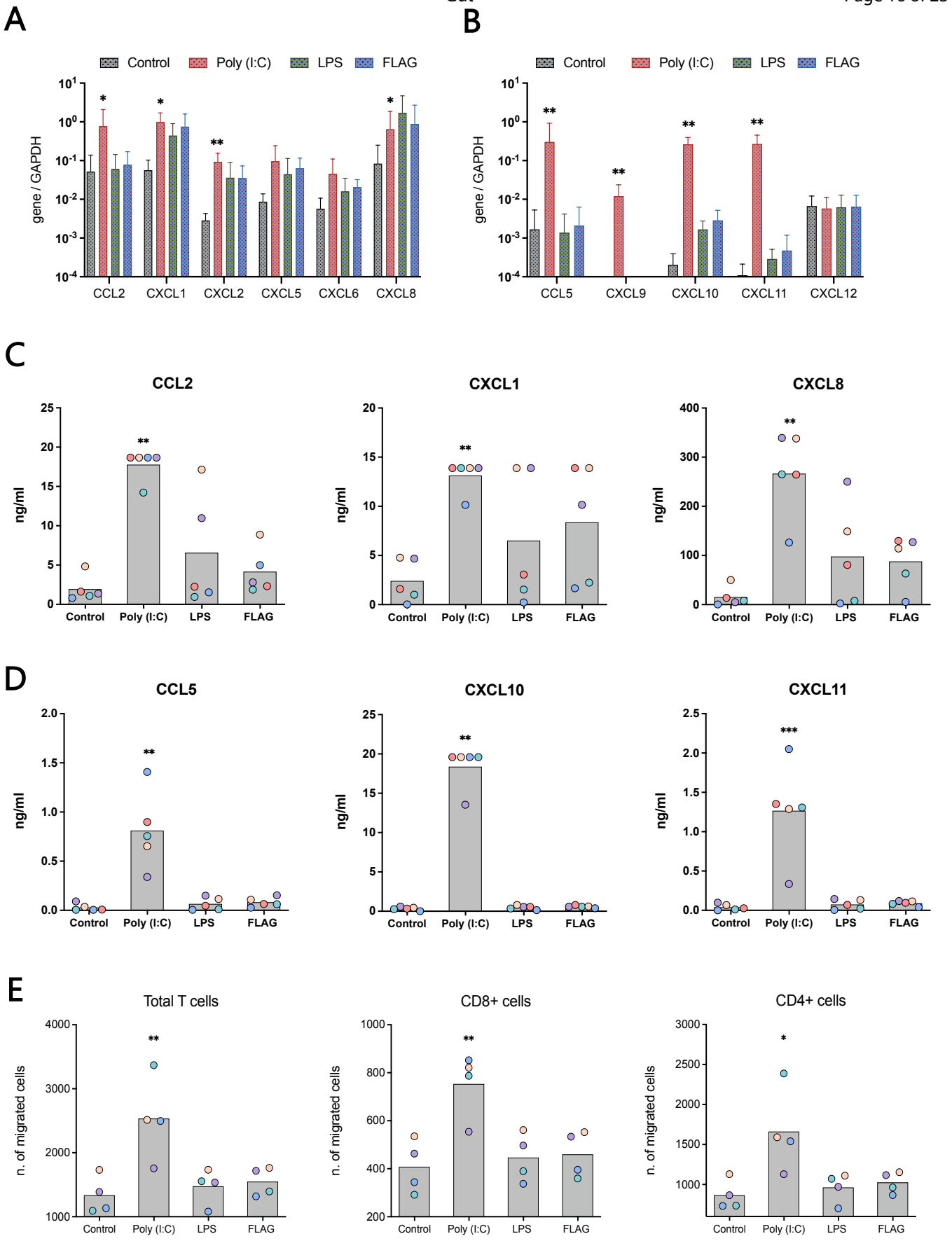

A
Control
Poly (I:C)
LPS
FLAG
gene / GAPDH
CCL2
CXCL1
CXCL2
CXCL5
CXCL6
CXCL8
B
CCL5
CXCL9
CXCL10
CXCL11
CXCL12
C
CCL2
CXCL1
CXCL8
ng/ml
D
CCL5
CXCL10
CXCL11
E
Total T cells
CD8+ cells
CD4+ cells
n. of migrated cells

Figure 2

Supplementary Material

Supplementary Methods

In silico data analysis

Data for in silico analysis were retrieved from GSE178341, the largest public scRNA seq database of human primary CRC currently available (4). Raw count matrix in the HDF5 Feature Barcode Matrix Format, including cluster subtype assignments, were downloaded and processed from GEO (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5387971).

A multi-step process was used, as previously reported (7). Raw single-cell RNA-Seq data were first normalized according to Seurat's normalization techniques and results were log-transformed, thus ensuring an unbiased and comparable gene expression landscape across individual cells. Subsequently, data scaling was applied to harmonize variations in sequencing depths and to capture the true biological diversity present in the dataset. Clinico-pathological characteristics of patients evaluated in this study were downloaded from (4) and are listed in Supplementary Table 1.

TASC isolation and characterization

Clinical specimens were collected from consenting patients undergoing surgical treatment at University Hospital Basel, St. Claraspital Basel, and Ospedale Civico Lugano, all in Switzerland. The use of human samples in this study was approved by the local ethical authorities (Ethikkommission Nordwest und Zentralschweiz, EKNZ and Comitato etico cantonale Ticino). Clinical-pathological characteristics of patients evaluated in this study are listed in Supplementary Table 2.

Tumor-associated stromal cells (TASCs) were isolated from freshly excised CRC samples as previously detailed (5). Briefly, specimens rinsed with PBS (GIBCO) and minced using scissors and scalpels were transferred to digestion medium containing serum-free DMEM (GIBCO), supplemented with Collagenase IV (100x; 20kU/ml; Worthington #CLSS-4), DNAse I (100x; 50 mg/ml; Sigma-Aldrich #D5025), HEPES (10mM; GIBCO #15630-056), Kanamycin (100x; GIBCO #15160-047), Amphotericin B (100x; 250ug/mL; Sigma-Aldrich #A9528), Metronidazol (250x; 200mg/ml, Braun), Cefuroxim (250x; 15mg/ml, Braun) and digested for 1 hour at 37°C on

continuous smooth rotation. Following digestion, suspensions were filtered through a 100µm (Falcon #352360) and then 70µm (Falcon #352350) nylon mesh Cell Strainer.

The small chunks remaining on the top of the 100µm filter were picked up with tweezers, transferred into wells of 12- well plates and cultured in alpha MEM supplemented with 10% Fetal Bovine serum (GIBCO), Kanamycin (100x; GIBCO #15160-047), GlutaMAX-I (10mM) Sodium Pyruvate (100mM), and Fibroblast growth factor-2 (5ng/mL, R&D Systems), at 37°C with 5% CO2.

Outgrowing cells, usually detectable after one week of culture, were further expanded and analyzed by flow cytometry for the expression of stromal markers and the absence of hemopoietic and endothelial markers, upon staining with antibodies specific for the indicated markers (Table S3).

TASCs cell stimulation with TLR agonists

TASCs were plated in 12-well plates (Sigma-Aldrich, St. Louis, MO) (120'000 cells/well in 1 ml) and stimulated with a panel of TLR agonists, including triacylated lipopeptide (Pam3CSK4, Invivogen, San Diego, CA), Lipopolysaccharide (LPS, from *Escherichia coli* O111:B4, Sigma-Aldrich), polyinosinic-polycytidylic acid (poly(I:C), Invivogen, San Diego, CA), synthetic diacylated lipoprotein (FSL-1, Invivogen), purified flagellin from *Salmonella typhimurium* (Invivogen, San Diego, CA), peptidoglycan (PGN-SA from *Staphylococcus aureus,* Invivogen, San Diego, CA), imidazoquinoline amine (Imiquimod R837, Invivogen, San Diego, CA), CpG oligonucleotide type A (ODN 2216, Invivogen, San Diego, CA), and CpG oligonucleotide type B (ODN 2006, Invivogen, San Diego, CA), at the indicated concentrations. After 4 hours, cells were collected for RNA extraction and subsequent analysis of chemokine gene expression. Supernatants from parallel cultures were collected after overnight incubation and used for Legendplex assays.

Quantitative Real-time PCR assays

Total RNA was extracted from untreated or stimulated TASCs using NucleoSpin RNA (MACHEREY-NAGEL). RNA concentration and purity were determined using NanoDrop1 ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE). RNA (1 µg) was reverse transcribed using the Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT, Invitrogen), and cDNA samples were amplified and analyzed by quantitative Real-Time PCR

(RTqPCR), by using the ABI prism™ 7700 sequence detection system, with TaqMan Universal Master Mix and No AmpErase UNG (both from Applied Biosystems). Commercially available primer sequences used in the study are listed in Table S4.

Legendplex protein assay

Chemokine release by TASC was measured using the LEGENDplex™ Human Th Cytokine Panel (13-plex) (BioLegend, San Diego, CA), allowing simultaneous quantification of 13 human chemokines, including CCL2, CCL5, CXCL10, CCL11, CCL17, CCL3, CCL4, CXCL9, CCL20, CXCL5, CXCL1, CXCL11 and CXCL8, according to the manufacturer's instructions. Briefly, beads of two different sizes and differing levels of allophycocyanin (APC) fluorescence, conjugated with antibodies specific for the chemokines of interest, were incubated with 25 µL of supernatants of TASCs (12 x10^4 cells per well) previously stimulated with TLR agonists. Beads were then incubated with phycoerythrin (PE)-labeled streptavidin for 30 min and washed twice prior to sample analysis. Samples were analyzed using a dual laser BD FACS Symphony flow cytometer. Samples were differentiated based on bead size and intensity of APC fluorescence. Protein quantity was determined using PE fluorescence calibrated to a standard curve.

Migration assay

Peripheral blood mononuclear cells (PBMCs) were isolated from healthy donors by density gradient separation, and T cells were subsequently purified by magnetic bead-sorting (Miltenyi Biotec). T cell migration towards TASC-derived supernatnats was assessed using 96-well transwell plates with polycarbonate membrane with a 5 µm pore size. T cells (1.5 x 10^5/well), resuspended in chemotaxis buffer (RPMI supplemented with 25 mM HEPES and 1% BSA) were plated in the upper chamber 80 (µl/well), whereas TASC supernatants were seeded in the lower chamber (235 µl/well). After an incubation period of 1.5 hours, cells migrated to the bottom of the plate were collected, stained with specific antibodies specific for CD3, CD4, and CD8, and counted by flow cytometry.

Statistical analysis

Significance of differential IL-6 and chemokines expression in stimulated versus untreated TASCs at gene and protein levels was tested as indicated. Statistical analysis was performed by using GraphPad Prism 5 software (GraphPad Software). P values < 0.05 were considered significant.

Supplementary References

7. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. Cell. 2021: 10.1016/j.cell.2021.04.048

Supplementary Table 1. Clinico-pathological characteristics of patients in the cohort used for scRNA seq analysis (n=64).

Patients' characteristics		Frequency n (%)
Age, median (range)		62 (35-91)
Sex	Male n(%)	32 (50)
	Female n (%)	32 (50)
Tumor location	Right colon n (%)	44 (69)
	Transverse colon n (%)	7 (11)
	Left colon n (%)	4 (6)
	Sigmoid colon n (%)	9 (14)
Grade	Low (G1 and G2) n (%)	51 (80)
	High (G3) n (%)	13 (20)
T stage	T1 n (%)	2 (3)
	T2 n (%)	10 (16)
	T3 n (%)	32 (50)
	T4 n (%)	20 (31)
N stage	N0 n (%)	36 (56)
	N1 n (%)	20 (31)
	N2 n (%)	8 (13)
M stage	M0 n (%)	61 (95)
	M1 n (%)	3 (5)
AJCC stage	I n (%)	7 (11)
	II n (%)	29 (45)
	III n (%)	25 (39)
	IV n (%)	3 (5)

Supplementary Table 2. Clinico-pathological characteristics of patients whose samples were used for TASC isolation (available for n=11 out of 12).

Patients' characteristics		Frequency n (%)
Age, median (range)		55 (42-83)
Sex	Male n(%)	7 (64)
	Female n (%)	4 (36)
Tumor location	Right colon n (%)	2(18)
	Transverse colon n (%)	1 (9)
	Left colon n (%)	2 (18)
	Sigmoid colon n (%)	6 (55)
Grade	G1 n (%)	1 (9)
	G2 n (%)	9 (82)
	G3 n (%)	1 (9)
T stage	T1 n (%)	0 (0)
	T2 n (%)	1 (9)
	T3 n (%)	9 (82)
	T4 n (%)	1 (9)
N stage	N0 n (%)	2 (18)
	N1 n (%)	5 (46)
	N2 n (%)	4 (36)
M stage	M0 n (%)	11 (100)
	M1 n (%)	0 (0)
AJCC stage	I n (%)	1 (9)
	II n (%)	1(9)
	III n (%)	9 (82)
	IV n (%)	0 (0)

Supplementary Table 3. List of antibodies used in this study.

Name	Clone	Brand
CD90	5E10	BD
CD29	TS2/16	Biolegend
CD73	AD2	BD
CD105	43A3	Biolegend
PDPN	NC-08	Biolegend
FSP-1	EPR2761(2)	Novus Biologicals
aSMA	1A4	Invitrogen
FAP	1E5	Abcam
CD3	OKT3	Biolegend
CD4	RPA-T4	Biolegend
CD8	SK1	Biolegend

Supplementary Table 4. List of primers used in this study.

Gene name	code/seq	Source
TLR1	Hs00413978_m1	AppliedBiosystems
TLR2	Hs00610101_m1	"
TLR3	Hs01551078_m1	"
TLR4	Hs00152939_m1	"
TLR5	Hs01019558_m1	"
TLR6	Hs01039989_s1	"
TLR7	Hs00152971_m1	"
TLR8	Hs00152972_m1	"
TLR9	Hs00152973_m1	"
TLR10	Hs01675179_m1	"
CCL2	Hs00234140_m1	AppliedBiosystems
CCL3	Hs00234142_m1	"
CCL5	Hs00982282_m1	"
CCL7	Hs00171147_m1	"
CCL11	Hs00237013_m1	"
CCL13	Hs00234646_m1	"
CCL26	Hs00171146_m1	"
CXCL1	Hs00236937_m1	"
CXCL2	Hs00236966_m1	"
CXCL5	Hs00171085_m1	"
CXCL6	Hs00237017_m1	"
CXCL8	Hs00174103_m1	"
CXCL9	Hs00171065_m1	"
CXCL10	Hs99999049_m1	"
CXCL11	Hs00171138_m1	"
CXCL12	Hs00171022_m1	"
CXCL16	Hs00222859	"

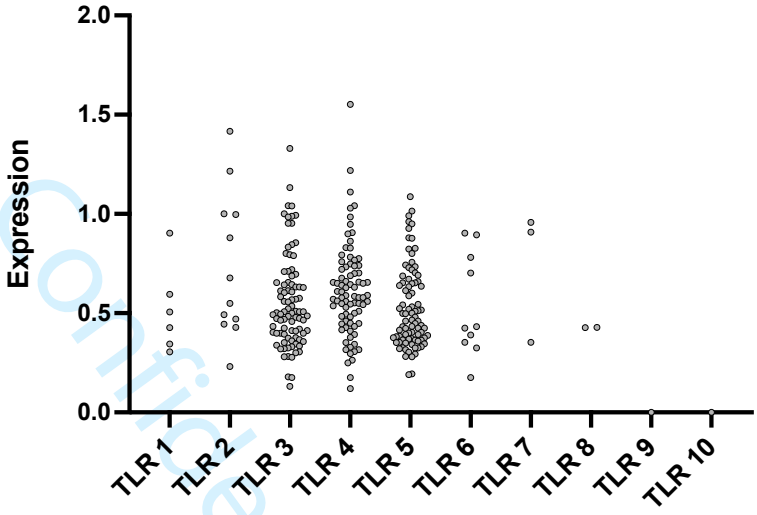


Figure S1. TLR expression on TASCs from distal MMR-proficient CRCs Data relative to cells classified as "Fibroblasts"(n=3032) were retrieved from the largest publicly available scRNA seq database of human primary CRC (GSE178341) (4). Only data relative to distal, MMR-proficient tumors (n=9) were considered. Normalized log-expression of TLR1-10 genes, as detected in 251 cells, is shown.

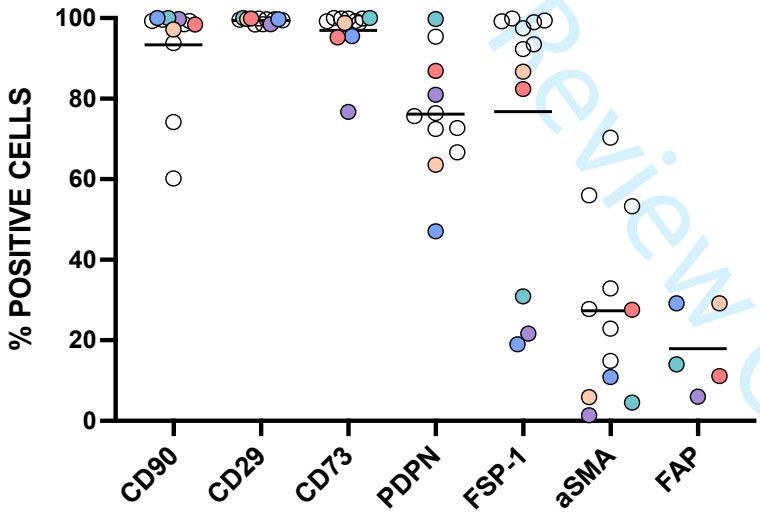


Figure S 21. Phenotypic analysis of expanded TASCs. TASCs were isolated from human primary CRC samples (n=12), as previously described (5). Following short in vitro expansion, phenotypical analysis was performed by flow cytometry. Percentages of cells positive for the indicated markers are reported. Each dot corresponds to one TASC preparation. Those further used for functional testing are identified by a specific color.

Combined tumor-associated microbiome and immune gene expression profiling predict response to neoadjuvant chemo-radiotherapy in locally advanced rectal cancer

Running Title: Combined signature predicts rectal cancer response to nCRT

R. Roesel[1,2,3], F. Strati[4], C. Basso[2,3], S. Epistolio[5], P. Spina[5], J. Djordjevic[2,3], E. Sorrenti[2,3], M.Villa[2,6], A. Cianfarani[1,2], F. Mongelli[1,3], J. Galafassi[1,2], S. G.Popeskou[1,3], F. Facciotti[4], P. E. Majno-Hurst[1,3], A. Franzetti-Pellanda[7], S. De Dosso[8,3], M. Frattini[5], D. Christoforidis[1,3,9]*, and G. Iezzi[2,3].*

*equal contribution

1 Department of Surgery, Ente Ospedaliero Cantonale, Lugano, Switzerland

2 Laboratory for Translational Surgical Research, Ente Ospedaliero Cantonale, Bellinzona Switzerland

3 Faculty of Biomedical Sciences, Università della Svizzera Italiana

4 Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milan, Italy

5 Laboratory of Molecular Pathology, Institute of Pathology, Ente Ospedaliero Cantonale, Locarno, Switzerland

6 Faculty of Biotechnology, University of Milan, Milan, Italy

7 Service of Radiotherapy, Gruppo Ospedaliero Moncucco, Clinica Moncucco, Lugano, Switzerland

8 Department of Medical Oncology, Oncology Institute of Southern Switzerland (IOSI), Ente Ospedaliero Cantonale, Bellinzona, Switzerland.

9 Department of Visceral Surgery, CHUV, University of Lausanne, Switzerland

Corresponding authors:

Prof. Giandomenica Iezzi, Laboratory for Translational Surgical Research, Ente Ospedaliero Cantonale, Via Chiesa 5, Bellinzona, 6500, Switzerland, +41 58 6667113, giandomenica.iezzi@eoc.ch

Prof. Dimitri Christoforidis, Department of Surgery, Ente Ospedaliero Cantonale, Via Tesserete 46, Lugano, 6900, Switzerland, +41 91 8116956, dimitrios.christoforidis@eoc.ch

Statement of translational relevance

The identification of markers predicting response to neoadjuvant chemoradiotherapy (nCRT) in patients with locally advanced rectal cancer (LARC) and the development of therapeutic protocols improving responsiveness remain unmet clinical needs.

In this study, upon combined analysis of intra-tumoral microbiome and immune gene expression profiling of LARC tissues, we identified bacterial species and immune pathways associated with complete response. Furthermore, our integrated analysis unraveled clustering of microbial taxa with each other and with immune cell related genes, leading to the identification of a combined signature of improved predictive capacity, which could correctly classify all non responders.

The possibility to identify unresponsive patients together with the identification of bacterial species associated with responses might pave the way to the development of innovative treatments, based on gut microbiota modulation, to increase effectiveness of nCRT in LARC.

Abstract

Purpose

Locally advanced rectal cancer (LARC) is treated with neoadjuvant chemo-radiotherapy (nCRT) followed by surgery. A minority of patients exhibit complete response (CR) to nCRT, and could be spared from surgery and its functional consequences. Instead, most patients show non-complete response (non-CR) and could benefit of additional treatment to increase CR rates. Reliable predictive markers are still lacking. The aim of this study was to identify novel signatures predicting nCRT responsiveness.

Experimental design

We performed a combined analysis of tumor-associated microbiome and immune gene expression profiling on diagnostic biopsies from 70 patients undergoing nCRT followed by rectal resection, including sixteen (23%) with CR and 54 (77%) with non-CR.

Results

Intratumoral microbiota significantly differed between CR and non-CR groups at genus and species level. Colonization by bacterial species of *Catenibacterium*, *Alloprevotella*, *Coprococcus,* and *Ruminococcus* genera was associated with CR, whereas abundance of *Fusobacterium*, *Pedomicrobium*, *Neissaeria*, and *Streptococcus* species predicted non-CR. Immune gene profiling revealed 41 differentially expressed genes, with significant upregulation of genes related to IFN-gamma and IFN-alpha response in patients with CR.

Integrated microbiome and immune gene profiling analysis unraveled clustering of microbial taxa with each other and with immune cell related genes and led to the identification of a combined signature with improved predictive capacity as compared to individual signatures.

Conclusions

Combined intratumoral microbiome-immune profiling improves prediction of response to nCRT. Correct identification of unresponsive patients and of bacteria promoting responsiveness might lead to novel therapeutic approaches based on gut microbiota pre-conditioning to increase effectiveness of nCRT in LARC.

Abbreviations:

nCRT: neoadjuvant chemoradiotherapy

LARC: Locally Advanced Rectal Cancer

CR : Complete response

CRs: Complete Responders

non-CRs: non-complete responders

TME : Total Mesorectal Excision

DEG: Differentially expressed gene

Introduction

Rectal cancer is common worldwide[1] and is mostly diagnosed at a locally advanced stage (T3-4 and/or N+)[2]. Patients with locally advanced rectal cancer (LARC) are usually treated by a multimodal strategy including neoadjuvant chemo-radiation therapy (nCRT) followed by rectal low anterior resection with total mesorectal excision (TME)[3].

This combined therapeutic approach has significantly decreased rates of local recurrence and improved overall survival [4]. However, rectal surgery, in particular when combined with radiation therapy, leads to significant defecatory, urinary and sexual dysfuntion[5].

In 15–20% of cases, nCRT ultimately results in a pathological complete response (CR), with no viable tumor cells left, as confirmed by histological analysis of the resected tissue[6].

These patients could avoid surgery and be monitored with a "watch and wait" strategy[7–9]. New nCRT protocolos have been developed aiming at increasing response rates[10,11] but accuracy of clinical and radiological assessment of CR after nCRT remains suboptimal and local regrowth occurs in [7–9,12–14] in 20-40% of cases[15–17] . The identification of factors predicting response is needed to effectively tailor nCRT and provide new treatment targets.

A number of histopathologic and genetic biomarkers, including specific gene mutations, non-coding RNA expression and presence of circulating tumor DNA, have been proposed[18–24], but none of them has shown a sufficiently reliable predictive ability.

Recent evidence suggests that effectiveness of nCRT might be associated with activation of the immune system. Indeed, tumor infiltration by CD3+, CD8+ or FOXP3+ lymphocytes has been found to be associated with favorable outcome and improved response to nCRT[25–30]. Radiation and oxaliplatin, currently used in nCRT protocols are indeed known to induce an immunogenic form of cell death, leading to the release of damage-associated molecular patterns (DAMPs) by dying cells, ultimately inducing activation of myeloid cells, including antigen presenting cells (APCs), such as macrophages and dendritic cells[31,32]. By presenting tumor associated antigens to T lymphocytes, APCs may promote the generation of tumor-specific immune responses.

The higher rate of response to nCRT when surgery is delayed[33], is also consistent with the timing of induction of anti-tumor immune responses. However, factors influencing the composition of LARC immune contexture, and the activation of potentially relevant immune cell pathways require better understanding.

Intra-tumoral gut microbiota is known to modulate immune cell infiltration in human colon cancer[34,35]. Due to neoplastic alteration of the mucosal barrier, defined components of the gut microbiota do translocate into tumor tissues. Some bacterial species, such as *Fusobacterium nucleatum,* appear to support tumor progression by promoting tumor cell proliferation[36], enhancing resistance of tumor cells to chemotherapy[37] or favoring the generation of an immunosuppressive environment[38,39]. On the other hand, abundance of other bacterial commensals has been found to be associated with high infiltration by immune cells and favorable outcome[35]or enhanced responsiveness to oxaliplatin[40]. Moreover, administration of specific bacteria has proved to effectively boost anti-tumor responses in CRC models[41,42].

Little is known about LARC-associated microbiota. In a recent study, abundance and persistence of *Fusobacterium nucleatum* in LARC tissues, has been shown to correlate with unresponsiveness to nCRT[43]. In few recent studies exploring fecal microbiota, specific microbial signatures were correlated with response to nCRT[44–46]However, the predictive significance of LARC intra-tumoral microbiota has not been thoroughly investigated, and its relationship with the tumor immune contexture has not been explored yet.

To fill this knowledge gap, in this work, we have characterized LARC-associated intratumoral microbiota and corresponding immune transcriptomic profiles in a cohort of patients treated with nCRT, and have analyzed the capacity of identified microbial and immune related signatures to predict response to nCRT.

Materials and Methods

Patients' population

All patients consecutively admitted at the Regional Hospital of Lugano, of the Ente Ospedaliero Cantonale, between 2012 and 2019 with a diagnosis of non-metastatic LARC (T3 or T4 and/or N+, M0) and, treated, as recommended upon multidisciplinary board discussion, with nCRT (long course: 50,4 Gy over 5 weeks combined with Capecitabine, 1250 mg/m^2) without interruption, followed by surgical resection, were retrospectively considered for inclusion.

Only patients whose diagnostic biopsy was no longer available or not adequate for further molecular characterization, or who did not provide written informed consent were excluded

(Figure1). The study was approved by the local Ethics Committee (Project-ID 2020-02387 / CE 3759).

Clinical and pathological characteristics were collected retrospectively from a prospectively maintained rectal cancer database and from electronic patient records. Local and systemic recurrence rates were determined from data obtained from electronic patient records and from the regional cancer registry.

Pathological response to nCRT was assessed according to Mandard tumor regression grading system[47,48], based on histological evaluation of resected tissues (after TME). Tumor regression grade (TRG) was confirmed by an experienced pathologist (P.S.) on histological re-evaluation. Patients were divided into two groups: complete responders (CRs), i.e. only cases with TRG-1 response, and non-complete responders (non-CRs), with TRG2-5 (see below).

Clinicopathologic characteristics were compared by Fisher's exact test. Survival curves (Supplementary Figure 1) were depicted according to the Kaplan–Meier method and compared by log-rank test.

Sample acquisition and processing

FFPE (Formalin Fixed Paraffin Embedded) tissues from diagnostic biopsies, obtained before any treatment from patients included in the study were retrieved. Three 8 µm-thick and two 10 µm-thick serial sections were obtained for DNA and RNA extraction, respectively. The presence of malignant tissue within the selected area was confirmed by an experienced pathologist (P.S.) upon hematoxylin/eosin staining.

Following sample deparaffinization, genomic DNA (gDNA) and mRNA were extracted from consecutive slides of each specimen, using the respective QIAmp Mini kits (Qiagen, Chatsworth, CA, USA), according to the manufacturers' protocols, and were quantified by NanoDrop Onec (ThermoFisher Scientific, USA). DNA and RNA were then used for microbiome analysis and immune cell gene expression profiling, respectively (see below).

16S rRNA gene sequencing and intratumoral microbiome analysis

Tumor-associated microbiome analysis was conducted based on amplification and sequencing of bacterial 16SrRNA gene. Briefly, two-step PCR libraries were created using the primer pair 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACHVGGGTWTCTA AT-3') to sequence V4 hypervariable region of 16S RNA gene.

Libraries were then sequenced on the Illumina MiSeq platform using a v2 500 cycles kit resulting in 2x250 bp reads. Demultiplexing and trimming of Illumina adaptor residuals for the produced paired-end reads were performed using the Illumina MiSeq Reporter software (v2.6.2). Quality of reads was checked with FastQC software (v0.11.5). Reads were pre-processed using the MICCA pipeline (v.1.7.0) (http://www.micca.org)[49]. Forward and reverse primers trimming and quality filtering were performed using micca trim and micca filter, respectively. Filtered sequences were denoised using the UNOISE algorithm [50] implemented in MICCA operational taxonomic units to determine true biological sequences at the single nucleotide resolution level by generating amplicon sequence variants (ASVs). Bacterial ASVs were taxonomically classified using classify and the Ribosomal Database Project (RDP) Classifier v2.13 [5146]. Multiple sequence alignment (MSA) of 16S sequences was performed using the Nearest Alignment Space Termination (NAST) algorithm [52] , implemented in MICCA msa with the template alignment clustered at 97% similarity of the Greengenes database [53] (release 13_08). Phylogenetic trees were inferred using Micca tree [54]. Sampling heterogeneity was reduced by rarefying samples at the depth of the less abundant sample using micca tablerare. Alpha (within-sample richness) and beta-diversity (between-sample dissimilarity) estimates were computed using the phyloseq R package [55]. Permutational multivariate analysis of variance (PERMANOVA) test was performed by using the adonis function in the R package vegan with 999 permutations. Linear discriminant effect size analysis (LEfSe) was performed to find features (microbial genera) most likely explaining differences between classes [56]. ASVs differential abundance testing was carried out by using the R package DESeq2 [57], on non-rarefied data [58]. Spearman's correlations were tested by using the psych R package[55]. Random Forest analyses[59] of 16S rRNA gene sequencing data were performed by using the randomForest R package, and permutation tests with 1000 permutations were performed to assess model significance [60].

Tissue immune cell profiling

Tissue immune gene profiling was performed by using the nCounter Human PanCancer Immune Profiling Panel (NanoString Technologies), according to the manufacturer protocol. Briefly, tissue-derived mRNA was hybridized at 65°C overnight. Following binding to streptavidin-coated surfaces and washings, cartridges were scanned by using nCounter digital

analyzer (NanoString Technologies). Quality control (QC) metrics were obtained through the NanoStringQC function from nanostringr package [61].

Samples with percent fields of view (FOV) ≤0.75, with Linearity of Positive Controls by Plate (R2) ≤0.95, with average housekeeping gene expression ≤0.60 and with a signal to noise ratio >150 were removed. A total of 61 samples, including 14 CRs and 47 NCRs, passed all QC filters and were valuable for further analysis.

Gene expression data were processed with the NanoStringNorm R package [62]using background subtraction and normalization with housekeeping genes. Batch effects associated with sequencing plate were adjusted with CombatSeq [63].

Differential gene expression analysis was performed by fitting a quasi-likelihood negative binomial generalized model to count data with EdgeR (v3.36) [64].

To test whether a condition was enriched for relevant up/downregulated pathways, the Camera approach was used together with the collection of gene-sets and Gene Ontology (GO) terms from Molecular Signatures Database (MSigDB) [65]. Gene-wise moderated t-statistics and Camera tests were used to assess whether a gene-set is highly ranked relative to condition signature in terms of differential expression (logFC), accounting for inter-gene correlation [66]. To show the enrichment of gene sets among logFC ranked genes, barcode plots were produced using the function implemented in the limma package [67]. As cutoff for statistical significance after multiple testing correction a false discovery rate (FDR) of 1% was used.

IFNγ release assay

Alloprevotella rava (strain DSM 22548) and *Fusobacterium nucleatum* (DSM 15643), expanded under anaerobic conditions, were incubated with peripheral blood mononuclear cells (PBMCs) from healthy donors, purified upon gradient separation, in RPMI 1640 medium supplemented with 5% human AB serum at 10:1 ratio. *IFNG* gene expression was assessed after 6 hours of stimulation by quantitative RT-PCR. Protein release in culture supernatants was measured after 24 hours by ELISA.

Results

Clinicopathological characteristic of the study population

Out of 187 consecutive patients considered, 70 met the inclusion criteria (Figure 1). Of those, 16 (23%) were CRs (TRG 1), and 54 (77%) were non-CRs (TRG 2: 12 (17%); TRG 3: 24 (34%); TRG 4: 18 (26%), TRG5:0 (0%). Clinicopathological characteristics of patients included in the study are listed in Table 1.

No significant differences were observed between CRs and non-CRs regarding median age, sex, distance of the tumor from the anal verge, clinical T or N stage, tumor grade, tumor size, CRM involvement, extramural vascular invasion, or time from nCRT to surgery (Table 1). A trend for more frequent CR in patients with mucinous adenocarcinoma was detectable (3/16 =18.7% vs 1/54=1,9%, p=0.07).

CRs and non-CRs are characterized by a different tumor-associated microbiota

Following amplification and sequencing of 16SRNA gene from gDNA, a total of 13,798,954 sequences were obtained from all samples. Rarefaction curves reached a plateau (Supplementary Figure 2), indicating that sequencing depth was sufficient to capture near-complete sample biodiversity. Sequences were mapped to 4207 ASVs. A total of 153 genera with relative abundance $\geq$0.05% were identified.

No significant differences in the overall diversity and microbial community structure, as measured by alpha- (within-sample richness) and beta-diversity (between-sample diversity) were observed between CRs and nonCRs (Figure 2A-C). However, two genera, i.e. *Catenibacterium* and *Coprococcus*, were significantly enriched in CRs and four genera, i.e. *Pedomicrobium*, *Neisseria,* and two unclassified genera of *Bacillales* and *Microbacteriaceae*, were more abundant in non-CRs (Figure 2D-E).

In agreement with our previous evaluation, an in-depth analysis of the tumor-associated microbiota showed the enrichment in CRs of different amplicon sequence variants (ASVs) belonging to *Catenibacterium* (ASV176) and *Coprococcus* (ASV123) genera. Moreover, *Alloprevotella* (ASV280), *Ruminococcus* (ASV498), and *Enterococococcus* (ASV677) were also enriched (Figure 2F and Supplementary Table 1). On the other hand, 28 different ASVs belonging to, among others, *Porphyromonas* (ASV146)*, Clostridium senso strictu 1* (ASV28), *Fusobacterium* (ASV11, ASV19, ASV154)*,* and *Streptococcus* genera (ASV39, ASV216, ASV1966), were enriched in nonCRs (Figure 2F, and Supplementary Table 1).

Response to nCRT is associated with activation of IFN-γ and IFN-α pathways.

To gain insights into immune contexture features possibly associated with CR, we performed an immune cell gene expression profiling on tissue-derived mRNA by using the PanCancer panel including 770 immune-related genes.

Upon differential analysis of individual gene expression, we could identify 41 differentially expressed genes (DEGs), with 20 genes, including, among others, *TREM2*, *PRAME*, *CXCL11*, *CTSW, FCGR1*, and C*4BPA*, upregulated in CRs, and 21 genes including *CLEC6A, SYCP1*, *CTAGE1, CTCFL, SLAMF6, IL22RA2*, and *EBI3*, overexpressed in non-CRs (Figure 3A and Supplementary Table 2). Gene set enrichment analysis identified numerous upregulated pathways in CRs versus non-CRs (Figure 3B and Supplementary Table 3) many of which were related to IFN-γ and IFN-α responses (Figure 3C).

Combined microbial-immunological signature predicts resistance to nCRT

Previous studies have demonstrated an association between the presence of defined bacterial species within tumor tissues and infiltration by specific immune cell subsets[33,34]. Therefore, we sought to investigate whether tumor-associated bacteria in LARC patients could promote specific immune-related transcriptional responses.

Indeed, immune cell exposure to bacteria associated with CR, such as *Alloprevotella*, resulted in significantly higher IFN-γ expression at both gene and protein level, as compared to *Fusobacterium*, which was abundant in non-CRs (Figure 4A).

Furthermore, combined microbiome and immune cell gene expression profiling analysis, unraveled specific clustering of microbial taxa with each other and with immune cell related genes. In particular, we observed that a cluster of microbial taxa enriched in CRs, including *Catenibacterium*, *Coprococcus, Ruminococcus* and *Alloprevotella,* negatively correlated with the expression of *CHIT1, CLEC6A, CTCFL*, *CYFIP2,* and *PAX5* genes. In addition, *Ruminococcus* (ASV498) abundance positively correlated with expression of *LTK, FCGR1A, CD244* and *C4BPA* (Figure 4B). On the other hand, the cluster of microbial taxa enriched in non-CRs, comprising *Pedomicrobium*, *Hyphomicrobium and* Legionella, positively correlated with the expression of *IL22RA2*, *KIR3DL3, CHIT1, IL4, CTAGE1, SYCP1, CLEC6A* and *CTCFL* (Figure 4B and Supplementary Table 4). *Pedomicrobium* abundance also correlated with *CXCL14* and *SLAMF6* gene expression. Notably, the enrichment of *Nesseria* and *Streptococcus* pathobionts correlated with the expression of *IL22RA2, CLEC6A, CTCFL, CXCL11, C1R* and with the

expression of *STAT1*, *CD209* and *GAGE1* genes, respectively (Figure 4B, and Supplementary Table 4).

Since we were able to associate specific microbial features with differential expression of immune-related genes, we asked whether we could predict responsiveness to nCRT by using their tumor-associated microbiome or immune transcriptional signatures. By using a Random Forest classifier, we observed that the tumor-associated microbiome is a poorer predictor of response to nCRT (OOB= 28.57%, Accuracy=0.715, Kappa=-0.0870, p=0.9) compared to the immune transcriptional profile (OOB= 19.67%, Accuracy=0.822, Kappa=0.346, p=0.002). However, the combination of microbial and immune transcriptional signatures improved the ability of the classifier to predict the response to nCRT (OOB=14.75%, Accuracy=0.901, Kappa=0.672, $p<0.0001$). Most importantly, only the combination of microbial and transcriptional data predicted with no errors non-CR to nCRT (Figure 4B). Of note, abundance of *Catenibacterium*, *Coprococcus,* and *Ruminococcus*, combined with expression levels of genes encoding the adhesion protein CYFIP2, the chemokine receptor CCR6, and the T cell-recruiting chemokine gene CXCL11, were included among top ranked parameters (Figure 4C).

Discussion

In this study we have performed a combined analysis of intra-tumoral microbiome and immune gene expression profiling, and we have identified a novel combined signature predicting response of LARC to nCRT.

We found that tumor tissue colonization by bacterial species of *Catenibacterium*, *Alloprevotella*, *Coprococcus,* and *Ruminococcus* genera is associated with CR to nCRT, whereas abundance of *Fusobacterium*, *Pedomicrobium*, *Neissaeria*, and *Streptococcus* species predicts non-CR. Overexpression of genes related to IFN-γ and IFN-α response was also identified as hallmark of responsiveness. Our analysis unraveled specific clustering of microbial taxa among each other and with the expression of immune cell related genes leading to the identification of highly relevant predictive features. Indeed, while microbiome or immune transcriptional signatures, were per se relatively poor predictors of response, their combination resulted in improved prediction potentially leading to the correct identification of all non-CRs.

A few recent studies have evaluated the impact of components of gut microbiota on nCRT outcome in LARC. Serna et al[43] analyzed rectal cancer tissue samples from 143 patients, specifically for presence and abundance of *Fusobacterium nucleatum* at baseline and after nCRT, using RNA in situ hybridization and quantitative PCR. Although baseline *Fusobacterium nucleatum* abundance was not associated with response, its persistence after nCRT predicted poor outcome. Yi et al [44], performed a whole microbiome analysis on fecal samples of 84 patients. They found that butyrate producing microbes, including *Roseburia, Dorea, and Anaerostipes,* were overexpressed at baseline in responders whereas, consistent with our findings, *Fusobacterium* was associated with poor response. Teng et al also analyzed fecal samples from 116 Chinese LARC patients and found that *Bacteroides vulgatus*-mediated nucleotide biosynthesis was associated with resistance to nCRT [46]. Finally, Sun et al[45], analyzed fecal and blood samples from 39 patients with LARC before, during and after nCRT. At baseline, *Clostridium sensu stricto 1* was found to be significantly enriched in good-responders.

To this date only the study by Takenaka et al.[68], has analyzed the whole tumor-associated microbiome in LARC, in a group of 44 patients. *Hungatella, Flavonifractor,* and *Methanosphaera* bacterial genera were found to be enriched in CRs, whereas *Enhydrobacter, Paraprevotella* and *Finegoldia*, but not *Fusobacterium*, were overrepresented in non-CRs.

However, this cohort included patients of different geographical origin (i.e 18 from Argentina and 26 Brasil) and the composition of the intratumoral microbiota already significantly varied according to the country of origin[68].

Our study represents the first analysis of tumor-associated microbiome in a homogeneous LARC cohort. Our data identifies *Catenibacterium*, *Alloprevotella*, *Coprococcus,* and *Ruminococcus* genera as predictive of CR, whereas enrichment of *Fusobacterium,* together with *Microbacteriaceae*, *Neissaeria*, and *Streptococcus* genera, is found to be associated with non-CR.

In a recent study, the presence of *Alloprevotella* and *Catenibacterium* in gut microbiota of patients with different solid tumors has been reported to be associated with favorable response to treatment with immunological checkpoint inhibitors[69,70]. *Ruminococcus* has also recently been associated with favorable outcome in colon cancer. Furthermore, besides cancer, gut colonization by *Coprococcus* has been shown to be associated with enhanced immune response and favorable outcome in COVID disease[71].

On the other hand, *Neisseria* genus has been found to be enriched in the oral microbiota of patients with oral cavity squamous cell carcinoma[72].

Mechanisms underlying the association of these bacteria with responsiveness and their potential causative role remain to be established. Our data suggests that some of identified bacteria may contribute to shape the tumor immune contexture.

Tumor immune gene expression profiling clearly indicated that CR to nCRT is associated with tumor infiltration by cytotoxic lymphocytes, as suggested by overexpression of genes encoding cathepsin-W, the CXCR3-binding chemokine CXCL11, and the natural killer receptor CD244, and with overexpression of gene sets related to INF-γ and IFN-alpha pathways. This is in line with previous studies reporting an association between tumor infiltration by cytotoxic CD8 and responsiveness to nCRT [25–30] and activation of proinflammatory pathways in responders, including type I IFN [73,74]

It is noted that bacteria associated with CR promoted in vitro expression of INF-γ significantly more effectively than bacteria associated with non-CR, thus supporting the capacity of these bacteria to directly influence the production of proinflammatory cytokines by immune cells.

Also, overexpression of myeloid cell markers, including TREM2, the high affinity immunoglobulin gamma Fc receptor I, and the complement component 4 binding protein

alpha, suggests that tumor-infiltrating myeloid cells, possibly including macrophages, dendritic cells, and neutrophils, also promote responsiveness to nCRT in LARC.

Previous studies in experimental models have shown that specific components of gut microbiota critically influence response to chemotherapeutic drugs, including oxaliplatin, by modulating tumor infiltrating myeloid cell functions [40]. Bacteria abundant in CRs, might also synergize with DAMPs released by nCRT-treated cancer cells, to promote myeloid cell activation and enhance their antigen presentation capacity, ultimately favoring development of lymphocyte-mediated anti-tumor responses. In our cohort, overexpression of HLA-C and HLA-DR molecules and of tumor-associated antigen PRAME was also associated with CR, consistent with the possible induction of tumor-specific immune response.

Thanks to our integrated analysis of intratumoral microbiome and immune profiling, we could also detect correlations between microbial taxa and immune-related genes. In particular, we found direct correlations between abundance of bacterial genera enriched in nonCRs, in particular *Pedomicrobium*, *Hypomicrobium*, *Neisseria*, and *Bacillales*, and expression levels of genes encoding for proteins involved in downmodulation of immune responses, including the IL-22receptor antagonist (encoded by *IL22RA*), the inhibitory cytokine 35 (encoded by *EBI3)* and the inhibitory NK cell receptor KIR3DL3 (encoded by *KIR3DL3*), thus suggesting that these bacteria may induce immunosuppressive effects possibly contributing to nCRT unresponsiveness.

In addition to providing insights into the potential bacteria-immune cell interplay occurring within LARC tissues, integration of microbiome and immune profiling allowed the identification of a combined signature, including 25 variables and displaying improved predictive power as compared to individual microbiome- or immune-related signatures. Abundance of *Catenibacterium*, *Coprococcus,* and *Ruminococcus*, as well as expression of genes encoding for the adhesion protein CYFIP2, the chemokine receptor CCR6, and T cell-recruiting chemokine gene CXCL11, were included among top ranked parameters.

In the study by Sun et al., integrated analysis of microbiome- and immune-related features was also found to predict response to nCRT with higher accuracy in a cohort of 39 patients[45]. However, this work was based on evaluation of fecal microbiota and immune cell parameters in peripheral blood, which may not necessarily mirror the dynamics occurring within the tumor microenvironment. Nevertheless, baseline high levels of *Clostridium senso strictu 1* were also found to be associated with poor response, consistent with our data.

If confirmed, our findings may have a high clinical relevance. The signature identified in our study could correctly classify all non-CRs, with no error. Such a prediction could influence treatment decisions as to the benefit or type of neoadjuvant treatment. Also, this signature could help design pre-neoadjuvant microbiome conditioning by eradicating unfavorable bacterial species, followed by administration of bacteria favoring nCRT effectiveness.

A few limitations of our investigation should be acknowledged. Although it represents the largest study so far exploring the intratumoral microbiome and immune contexture in combination, the number of patients included is relatively small. Moreover, it is a retrospective study. However, the patients' population under investigation is rather homogenous, living in a small geographical area, treated with a standardized nCRT protocol and surgery, which was performed at 10 weeks after completion of nCRT with minimal variation in timing. Intra-tumoral microbiome analysis was necessarily based on sequencing of 16S gene, due to low intra-tissue bacterial loads. While this method allows to identify bacteria that actually infiltrate tumor tissues, it does not provide information about their functional properties. Similarly, large scale immune gene expression profiling provides a comprehensive analysis of functional orientation of tumor microenvironment but does not inform on individual immune cell subsets. Further prospective studies, integrating analysis of fecal and intratumoral microbiome, and large-scale immunophenotyping of tumor infiltrating immune cells are warranted to shed more light on the tumor-microbiota-immune cell interplay within tumor microenvironment. Visualization of potential interactions between microbial species, tumor cells and immune cells by emerging spatial transcriptomics and proteomics[34], may also provide additional information about potentially direct interactions of bacteria predictive of response to nCRT and tumor infiltrating immune cells.

In addition, our analysis has been performed on pre-treatment diagnostic biopsies only, without taking into account changes potentially occurring after treatment, which could also be of predictive value[45]. Nevertheless, the identification of predictive parameters prior to therapy is critical to implement novel interventional strategies, such as gut microbiota conditioning, aimed at enhancing nCRT efficacy.

In conclusion, we have provided evidence that combined analysis of intratumoral microbiome and tumor immune gene expression profiling results in the identification of markers predicting responses to treatment with improved accuracy. The possibility to identify unresponsive patients together with the identification of bacterial species associated with

responses may point towards the development of innovative “pre-neoadjuvant” treatments, based on gut microbiota modulation, to increase effectiveness of nCRT in LARC.

Acknowledgments

We thank the patients and their families for their consent to use their biological samples for this study . We are grateful to Prof. Silvia Monticelli (Institute for Research in Biomedicine, IRB, Bellinzona, Switzerland) and Dr. Andrea Rinaldi (IRB, Bellinzona, Switzerland) for helping with design and performance of Nanostring analysis and to Dr. Ferdinando Bonfiglio (University of Naples Federico II, Naples, Italy) for helping with data analysis. We thank Prof. Franco Cavalli (Bios+, Bellinzona, Switzerland) and Prof. Giulio C. Spagnoli (National Research Council, CNR, Rome, Italy) for critical reading of the manuscript and comments.

Funding

This work was supported by the Ente Ospedaliero Cantonale (AFRI Junior Researcher grant Grant to R.R. and Senior Researcher Grant to G.I.), San Salvatore Foundation (grant to D.C.) and Lega Ticinese contro il cancro (grant to D.C. and G.I.).

Disclosures

The authors do not have any conflict of interest to disclose.

Generative AI was not used in writing the manuscript.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA CANCER J CLIN*. 2018;68:394–424. Available at: https://acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21492
2. AJCC Cancer Staging System Version 9. 2023.
3. Glynne-Jones R, Wyrwicz L, Tiret E, Brown G, Rödel C, Cervantes A, et al. Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology*. 2017;28:iv22–iv40.
4. Rectal Cancer, Version 2.2018, NCCN Clinical Practice Guidelines in Oncology Enhanced Reader.
5. Hernandez MC, Wong P, Melstrom K. Low anterior resection syndrome. *J Surg Oncol* . June 2023;127(8):1271-1276. Available at: https://onlinelibrary.wiley.com/doi/10.1002/jso.27261
6. Glynne-Jones R, Glynne-Jones S. The concept and use of the neoadjuvant rectal score as a composite endpoint in rectal cancer. *Review Lancet Oncol*. 2021;22:314–340. Available at: www.thelancet.com/oncology
7. Fokas E, Appelt A, Glynne-Jones R, Beets G, Perez R, Garcia-Aguilar J, et al. International consensus recommendations on key outcome measures for organ preservation after (chemo)radiotherapy in patients with rectal cancer. *Nat Rev Clin Oncol*. 2021;18(12):805–816. Available at: https://www.surveymonkey.com
8. J Joshua Smith, Julio Garcia-Aguilar, Philip B Paty. Looking Forward, Not Backward, on Watch and Wait for Rectal Cancer-In Reply. *JAMA Oncol* . August 1, 2019;5(8):1231.
9. Dossa F, Moore SE, Baxter NN. Looking Forward, Not Backward, on Watch and Wait for Rectal Cancer. *JAMA Oncol*. August 1, 2019;5(8):1230–1231.
10. Bahadoer RR, Dijkstra EA, van Etten B, M Marijnen CA, Putter H, Meershoek-Klein Kranenbarg E. Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. *Lancet Oncol*. January 2021;22(1):29-42. Available at: www.thelancet.com/oncology
11. Bustamante-Lopez L, Zuhdy M, Tonello M, Dai D, Liu G, Liu H, et al. Clinical feasibility of the therapeutic strategies total neoadjuvant therapy and “watch and wait” in the treatment of rectal cancer patients with recurrence after clinical complete response. *Front Surg*. January 16, 2023;9:1006624.
12. Renehan AG, Malcomson L, Emsley R, Gollins S, Maw A, Myint AS, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. *Lancet Oncol* . February 2016;17(2):174–183.
13. Maas M, Beets-Tan RGH, Lambregts DMJ, Lammering G, Nelemans PJ, Engelen SME, et al. Wait-and-see policy for clinical complete responders after chemoradiation for rectal cancer. *Journal of Clinical Oncology*. December 10, 2011;29(35):4633–4640.
14. Dalton RSJ, Velineni R, Osborne ME, Thomas R, Harries S, Gee AS, et al. A single-centre experience of chemoradiotherapy for rectal cancer: Is there potential for nonoperative management? *Colorectal Disease*. May 2012;14(5):567–571.

15. Garcia-Aguilar J, Patil S, Gollub MJ, Kim JK, Yuval JB, Thompson HM, et al. Organ Preservation in Patients With Rectal Adenocarcinoma Treated With Total Neoadjuvant Therapy. *Journal of Clinical Oncology*. August 10, 2022;40(23):2546–2556.
16. Fernandez LM, São Julião GP, Figueiredo NL, Beets GL, van der Valk MJM, Bahadoer RR, et al. Conditional recurrence-free survival of clinical complete responders managed by watch and wait after neoadjuvant chemoradiotherapy for rectal cancer in the International Watch & Wait Database: a retrospective, international, multicentre registry study. *Lancet Oncol*. 2021;22(1).
17. Smith JJ, Strombom P, Chow OS, Campbell ;, Roxburgh S, Lynn P, et al. Assessment of a Watch-and-Wait Strategy for Rectal Cancer in Patients With a Complete Response After Neoadjuvant Therapy Editor's Note Supplemental content. *JAMA Oncol*. 2019;5(4):185896. Available at: https://jamanetwork.com/
18. Roesel R, Epistolio S, Molinari F, Saletti P, De Dosso S, Valli M, et al. A Pilot, Prospective, Observational Study to Investigate the Value of NGS in Liquid Biopsies to Predict Tumor Response After Neoadjuvant Chemo-Radiotherapy in Patients With Locally Advanced Rectal Cancer: The LiBReCa Study. *Front Oncol*. June 28, 2022;12.
19. Bengala C, Bettelli S, Bertolini F, Sartori G, Fontana A, Malavasi N, et al. Prognostic role of EGFR gene copy number and KRAS mutation in patients with locally advanced rectal cancer treated with preoperative chemoradiotherapy. *Br J Cancer*. 2010;103:1019–1024. Available at: www.bjcancer.com
20. Lopes-Ramos CM, Habr-Gama A, De Souza Quevedo B, Lia N?, Fel M, Cio ?, et al. Overexpression of miR-21-5p as a predictive marker for complete tumor regression to neoadjuvant chemoradiotherapy in rectal cancer patients [Internet]. Available at: www.targetscan.
21. Lee H-H, Chen C-H, Huang Y-H, Chiang C-H, Huang M-Y. Biomarkers of Favorable vs. Unfavorable Responses in Locally Advanced Rectal Cancer Patients Receiving Neoadjuvant Concurrent Chemoradiotherapy. *Cells*. 2022;11(10):1611. Available at: https://doi.org/10.3390/cells11101611
22. Kamran SC, Lennerz JK, Margolis CA, Liu D, Reardon B, Wankowicz SA, et al. Integrative molecular characterization of resistance to neoadjuvant chemoradiation in rectal cancer. *Clinical Cancer Research*. September 15, 2019;25(18):5561–5571.
23. Duldulao MP, Lee W, Nelson RA, Li W, Chen Z, Kim J, et al. Mutations in Specific Codons of the KRAS Oncogene are Associated with Variable Resistance to Neoadjuvant Chemoradiation Therapy in Patients with Rectal Adenocarcinoma. *Society for Surgical Oncology Annual Meeting*. 2012;
24. Chatila WK, Kim JK, Walch H, Marco MR, Chen CT, Wu F, et al. Genomic and transcriptomic determinants of response to neoadjuvant therapy in rectal cancer. *Nat Med*. 2022;28(8).
25. Anitei MG, Zeitoun G, Mlecnik B, Marliot F, Haicheur N, Todosi AM, et al. Prognostic and predictive values of the immunoscore in patients with rectal cancer. *Clinical Cancer Research*. April 1, 2014;20(7):1891–1899.
26. Shinto E, Hase K, Hashiguchi Y, Sekizawa A, Ueno H, Shikina A, et al. CD8+ and FOXP3+ Tumor-Infiltrating T Cells Before and After Chemoradiotherapy for Rectal Cancer. *Ann Surg Oncol*. 2014;21(Suppl 3):S414-21.
27. Mirjolet C, Charon-Barra C, Ladoire S, Arbez-Gindre F, Bertaut A, Ghiringhelli F, et al. Tumor lymphocyte immune response to preoperative radiotherapy in locally advanced rectal cancer: The LYMPHOREC study. *Oncoimmunology*. November 27,

2017;7(3):e1396402. Available at: https://www.tandfonline.com/action/journalInformation?journalCode=koni20
28. Matsutani S, Shibutani M, Kiyoshi |, Hisashi M|, Tatsunari N|, Shigetomi F|, et al. Significance of tumor-infiltrating lymphocytes before and after neoadjuvant therapy for rectal cancer. *Cancer Sci*. April 2018;109(4):966-979. Available at: https://onlinelibrary.wiley.com/doi/10.1111/cas.13542
29. Akiyoshi T, Tanaka N, Kiyotani K, Gotoh O, Yamamoto N, Oba K, et al. Immunogenomic profiles associated with response to neoadjuvant chemoradiotherapy in patients with rectal cancer. *Br J Surg*. September 2019;106(10):1381–1392. Available at: www.bjs.co.uk
30. Sissy C El, Kirilovsky A, Van Den Eynde M, Muşină A-M, Anitei M-G, Romero A, et al. A Diagnostic Biopsy-Adapted Immunoscore Predicts Response to Neoadjuvant Treatment and Selects Patients with Rectal Cancer Eligible for a Watch-and-Wait Strategy A C. *Clin Cancer Res*. October 1, 2020;26(19):5198-5207. Available at: http://clincancerres.aacrjournals.org/
31. Galluzzi L, Buqué A, Kepp O, Zitvogel L, Kroemer G. Immunogenic cell death in cancer and infectious disease. *Nat Rev Immunol*. 2017;17(2):97-111. Available at: www.nature.com/nri
32. Frey B , Rückert M, Deloch L, Rühle PF, Derer A, Fietkau R, et al. Immunomodulation by ionizing radiation-impact for design of radio-immunotherapies and for treatment of inflammatory diseases. *Immunol Rev* . November 2017;280(1):231-248.
33. van Eeghen EE, den Boer F, Bakker SD, Loffeld RJ. Outcome of rectal cancer after radiotherapy with a long or short waiting period before surgery, a descriptive clinical study. *J Gastrointest Oncol*. 2016;7(3):321–5. Available at: http://dx.doi.org/10.21037/jgo.2015.10.08
34. Luis Galeano Niño J, Wu H, LaCourse KD, Kempchinsky AG, Baryiames A, Barber B, et al. Effect of the intratumoral microbiota on spatial and cellular heterogeneity in cancer. *Nature*. 2022;611. Available at: https://doi.org/10.1038/s41586-022-05435-0
35. Cremonesi E, Governa V, Garzon JFG, Mele V, Amicarella F, Muraro MG, et al. Gut microbiota modulate T cell trafficking into human colorectal cancer. *Gut*. November 2018;67(11):1984-1994.
36. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. Fusobacterium nucleatum Promotes Colorectal Carcinogenesis by Modulating E-Cadherin/β-Catenin Signaling via its FadA Adhesin. *Cell Host Microbe*. August 14, 2013;14(2):195–206.
37. Yu TC, Guo F, Yu Y, Sun T, Ma D, Han J, et al. Fusobacterium nucleatum Promotes Chemoresistance to Colorectal Cancer by Modulating Autophagy. *Cell*. July 27, 2017;170(3):548-563.e16.
38. Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, Gamliel M, et al. Binding of the Fap2 protein of fusobacterium nucleatum to human inhibitory receptor TIGIT protects tumors from immune cell attack. *Immunity*. February 17, 2015;42(2):344–355.
39. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, et al. Fusobacterium nucleatum Potentiates Intestinal Tumorigenesis and Modulates the Tumor-Immune Microenvironment. *Cell Host Microbe*. August 14, 2013;14(2):207–215.
40. Iida N, Dzutsev A, Stewart CA, Smith L, Bouladoux N, Weingarten RA, et al. Commensal Bacteria Control Cancer Response to Therapy by Modulating the Tumor Microenvironment [Internet]. Available at: http://science.sciencemag.org/

41. Montalban-Arques A, Katkeviciute E, Busenhart P, Bircher A, Wirbel J, Zeller G, et al. Commensal Clostridiales strains mediate effective anti-cancer immune response against solid tumors. *Cell Host Microbe*. October 13, 2021;29(10):1573-1588.e7.
42. Zhang X, Yu D, Wu D, Gao X, Shao F, Zhao M, et al. Tissue-resident Lachnospiraceae family bacteria protect against colorectal carcinogenesis by promoting tumor immune surveillance. *Cell Host Microbe*. March 8, 2023;31(3):418-432.e8.
43. Serna G, Ruiz-Pace F, Hernando J, Alonso L, Fasani R, Landolfi S, et al. Fusobacterium nucleatum persistence and risk of recurrence after preoperative treatment in locally advanced rectal cancer. *Annals of Oncology*. October 1, 2020;31(10):1366–1375.
44. Yi Y, Shen L, Shi W, Xia F, Zhang H, Wang Y, et al. Gut Microbiome Components Predict Response to Neoadjuvant Chemoradiotherapy in Patients with Locally Advanced Rectal Cancer: A Prospective, Longitudinal Study. *Clin Cancer Res*. March 1, 2021;27(5):1329-1340. Available at: http://clincancerres.aacrjournals.org/
45. Sun Y, Zhang X, Jin C, Yue K, Sheng D, Zhang T, et al. Prospective, longitudinal analysis of the gut microbiome in patients with locally advanced rectal cancer predicts response to neoadjuvant concurrent chemoradiotherapy. *J Transl Med*. December 1, 2023;21(1).
46. Teng H, Wang Y, Sui X, Fan J, Li S, Lei X, et al. Gut microbiota-mediated nucleotide synthesis attenuates the response to neoadjuvant chemoradiotherapy in rectal cancer. *Cancer Cell*. January 9, 2023;41(1):124-138.e6.
47. Mandard A -M, Dalibard F, Mandard J -C, Marnay J, Henry-Amar M, Petiot J -F, et al. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer*. 1994;73(11).
48. Thies S, Langer R. Tumor regression grading of gastrointestinal carcinomas after neoadjuvant treatment. *Front Oncol*. 2013;3 OCT.
49. Albanese D, Fontana P, De Filippo C, Cavalieri D, Donati C. MICCA: a complete and accurate software for taxonomic profiling of metagenomic data. *Sci Rep*. May 2015;19(5)::9743. Available at: www.nature.com/scientificreports
50. Edgar RC. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *Bioxriv*. October 15, 2016; Available at: https://doi.org/10.1101/081257
51. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol*. August 2007;73(16):5261–5267.
52. Desantis TZ, Hugenholtz P, Keller K, Brodie EL, Larsen N, Piceno YM, et al. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Res*. July 1, 2006;34((Web Server issue)):W394-9. Available at: http://greengenes.lbl.
53. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*. July 2006;72(7):5069–5072.
54. Price MN, Dehal PS, Arkin AP. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One*. March 10, 2010;5(3).
55. McMurdie PJ, Holmes S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One*. April 22, 2013;8(4).
56. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation [Internet]. 2011. Available at: http://genomebiology.com/2011/11/6/R60

57. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. Available at: http://www.
58. Mcmurdie PJ, Holmes S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Comput Biol*. 2014;10(4):1003531. Available at: www.ploscompbiol.org
59. Breiman L. Random Forests. *Mach Learn*. 2001;45:5–32.
60. Murphy MA, Evans JS, Storfer A. Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. Vol. 91, Ecology. 2010.
61. Talhouk A, Kommoss S, Mackenzie R, Cheung M, Leung S, Chiu DS, et al. Single-Patient Molecular Testing with NanoString nCounter Data Using a Reference-Based Strategy for Batch Effect Correction. *PLoS One*. April 20, 2016;11(4):e0153844. Available at: https://github.com/
62. Waggott D, Chu K, Yin S, Wouters BG, Liu F-F, Boutros PC. Gene expression NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. *BIOINFORMATICS APPLICATIONS NOTE*. 2012;28(11):1546–1548. Available at: http://cran.r-project.org/web/packages/NanoStringNorm
63. Zhang Y, Parmigiani G, Johnson WE. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genom Bioinform*. 2020;2(3). Available at: https://academic.oup.com/nargab/article/2/3/lqaa078/5909519
64. Mccarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res*. May 2012;40(10):4288–97. Available at: https://academic.oup.com/nar/article/40/10/4288/2411520
65. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* . October 25, 2005;102(43):15545–50. Available at: www.pnas.orgcgidoi10.1073pnas.0506580102
66. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res*. September 1, 2012;40(17):e133. Available at: https://academic.oup.com/nar/article/40/17/e133/2411151
67. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7).
68. Takenaka IKTM, Bartelli TF, Defelicibus A, Sendoya JM, Golubicki M, Robbio J, et al. Exome and Tissue-Associated Microbiota as Predictive Markers of Response to Neoadjuvant Treatment in Locally Advanced Rectal Cancer. *Front Oncol*. March 22, 2022;12.
69. Hamada K, Isobe J, Hattori K, Hosonuma M, Baba Y, Murayama M, et al. Turicibacter and Acidaminococcus predict immune-related adverse events and efficacy of immune checkpoint inhibitor. *Front Immunol*. 2023;14.
70. Roelands J, K Kuppen PJ, Ahmed EI, Mall R, Masoodi T, Singh P, et al. An integrated tumor, immune and microbiome atlas of colon cancer. *Nature Med*. Available at: https://doi.org/10.1038/s41591-023-02324-5
71. Xu X, Zhang W, Guo M, Xiao C, Fu Z, Yu S, et al. Integrated analysis of gut microbiome and host immune responses in COVID-19. *Front Med*. April 2022;16(2):263–275. Available at: https://doi.org/10.1007/s11684-022-0921-6

72. Yan K, Auger S, Diaz A, Naman J, Vemulapalli R, Hasina R, et al. Microbial Changes Associated With Oral Cavity Cancer Progression. *Otolaryngology–Head and Neck Surgery*. June 1, 2023;168(6):1443–1452. Available at: https://doi.org/10.1002/ohn.211
73. Hyuck Jeon S, Kyu Chie E. Characterization of the gene signature correlated with favorable response to chemoradiotherapy in rectal cancer: A hypothesis-generating study. *Cancer Med*. 2023;12:8981–8990. Available at: https://onlinelibrary.wiley.com/doi/10.1002/cam4.5586
74. Rezapour A, Rydbeck D, Byvald F, Tasselius V, Danielsson G, Angenete E, et al. A type I interferon footprint in pre-operative biopsies is an independent biomarker that in combination with CD8 + T cell quantification can improve the prediction of response to neoadjuvant treatment of rectal adenocarcinoma. *Oncoimmunology*. May 10, 2023;12(1):2209473. Available at: https://www.tandfonline.com/action/journalInformation?journalCode=koni20

Table1: Baseline characteristics of study population

	Patients with LARC (n=70)				
	CRs (N=16)	**[TRG1]**	**non-CRs** (N=54)	**[TRG2-5]**	***p***
Age (years):					
Median (IQR)	65	(57-77)	67	(62-74)	0,25
Gender, n. (%):					
Male	12	(75)	35	(64.8)	0,55
Female	4	(25)	19	(35.2)	
Dist. a.v. (cm), n. (%):					
> 5	6	(37.5)	32	(59.3)	0,16
≤ 5	10	(62.5)	22	(40.7)	
MRI Clinical T, n. (%):					
cT2	1	(6.3)	4	(7.4)	0,45
cT3	14	(87.4)	40	(74.1)	
cT4	1	(6.3)	10	(18.5)	
MRI Clinical N, n. (%):					
cN0	4	(25)	10	(18.5)	0,56
cN1	8	(50)	35	(64.8)	
cN2	4	(25)	9	(16.6)	
AJCC Stage Baseline n. (%):					
II	4	(12.6)	10	(18.5)	0,83
III	12	(87.4)	44	(81.5)	
Grading [G] n. (%):					
G2	16	(100)	51	(94.4)	0,79
G3	0	(0)	3	(5.6)	
Histology n. (%):					
Conventional	13	(81.3)	53	(98.1)	0,07
Mucinous	3	(18.7)	1	(1.9)	
CRM Involvement n. (%):					
Positive	6	(37.5)	19	(35.2)	1
Negative	10	(62.5)	35	(64.8)	
EMVI n. (%):					
Positive	1	(6.3)	2	(3.7)	0,55
Negative	15	(93.7)	52	(96.3)	
Tumor Length (cm):					
Median (95% CI)	7	(5,8-8)	5	(4–6)	0,10
Time from nCRT to Surgery (days):					
Median (95% CI)	70	(65-72)	63	(58-70)	0,12
Surgery, n (%):					
Sphincter-Preserving	16	(100)	46	(66.6)	0,37
Non Sphincter-Preserving	0	(0)	8	(33.4)	

Figure 1

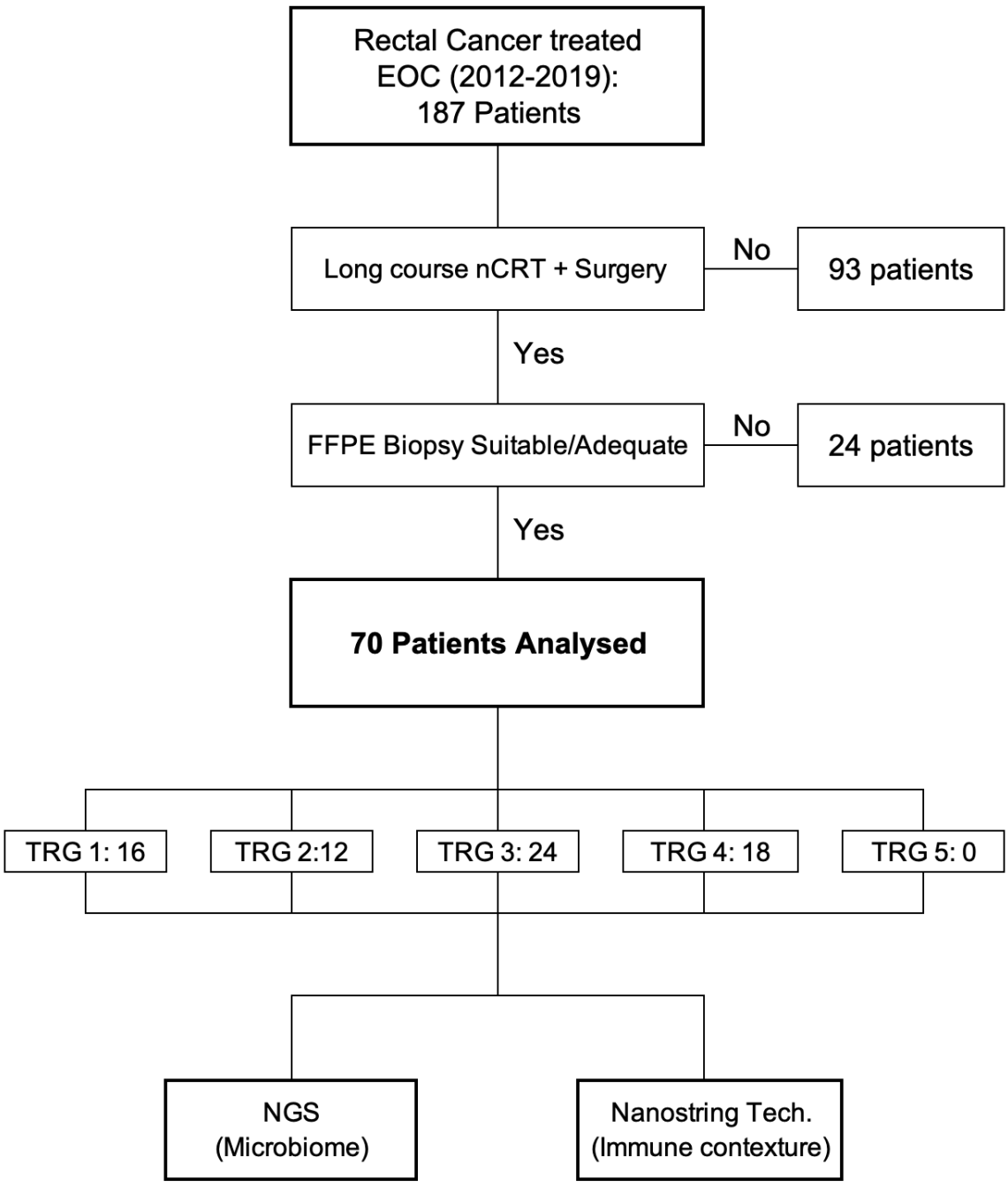


Figure 1: Consort Flow Diagram
EOC: Ente Ospedaliero Cantonale, **nCRT**: Neoadjuvant Chemo-RadioTherapy, **FFPE**: Formalin-Fixed Paraffin-Embedded, **TRG**: Tumor Regression Grade, **NGS**: Next Generation Sequencing.

Figure 2

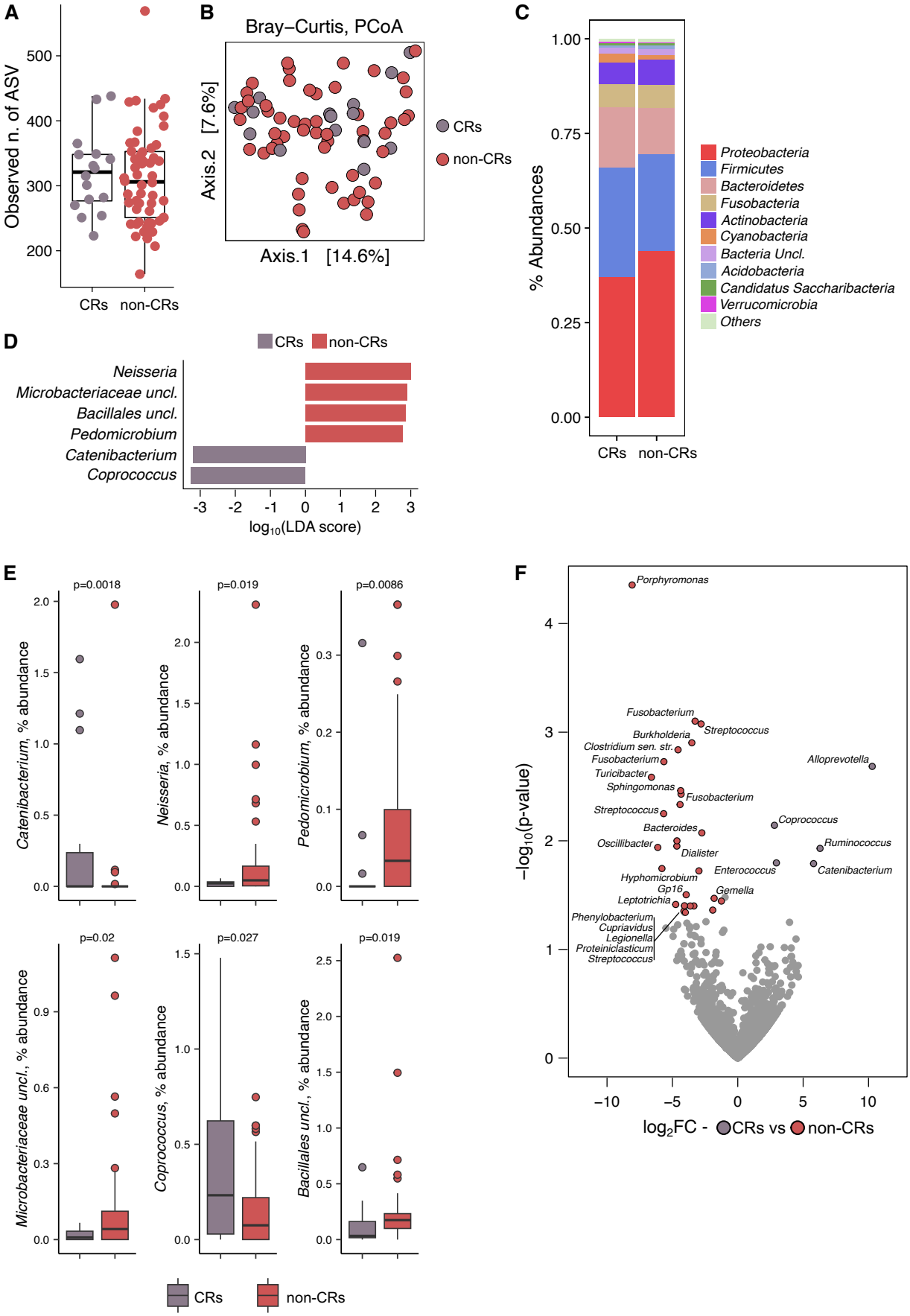

A
Observed n. of ASV
500
400
300
200
CRs
non-CRs
B
Bray–Curtis, PCoA
Axis.2 [7.6%]
Axis.1 [14.6%]
CRs
non-CRs
C
% Abundances
1.00
0.75
0.50
0.25
0.00
CRs
non-CRs
Proteobacteria
Firmicutes
Bacteroidetes
Fusobacteria
Actinobacteria
Cyanobacteria
Bacteria Uncl.
Acidobacteria
Candidatus Saccharibacteria
Verrucomicrobia
Others
D
CRs
non-CRs
Neisseria
Microbacteriaceae uncl.
Bacillales uncl.
Pedomicrobium
Catenibacterium
Coprococcus
-3
-2
-1
0
1
2
3
log10(LDA score)
E
p=0.0018
p=0.019
p=0.0086
p=0.02
p=0.027
p=0.019
Catenibacterium, % abundance
Neisseria, % abundance
Pedomicrobium, % abundance
Microbacteriaceae uncl., % abundance
Coprococcus, % abundance
Bacillales uncl., % abundance
CRs
non-CRs
F
Porphyromonas
Fusobacterium
Streptococcus
Burkholderia
Clostridium sen. str.
Fusobacterium
Alloprevotella
Turicibacter
Sphingomonas
Fusobacterium
Streptococcus
Coprococcus
Bacteroides
Oscillibacter
Ruminococcus
Dialister
Enterococcus
Catenibacterium
Hyphomicrobium
Gp16
Gemella
Leptotrichia
Phenylobacterium
Cupriavidus
Legionella
Proteiniclasticum
Streptococcus
−log10(p-value)
−10
−5
0
5
10
log2FC - CRs vs non-CRs

Figure 2: Tumor-associated microbiome analysis of CRs versus NCRs. A. Observed number of ASVs in complete responders (CR; n=16) and non-complete responders (non-CRs; n=54). **B.** PCoA of microbial beta-diversity as measured by Bray–Curtis dissimilarity index. **C.** Mean relative abundance of the top 10 most abundant Phylum in CRs and non-CRs. All the other less abundant Phylum are reported together and labeled as "others.". **D.** Most discriminant bacterial genera identified by LEfSe analysis. Positive and negative LDA scores indicate taxa enriched in the tumor-associated microbiome of non-CRs and CRs, respectively. Only taxa having a $p<0.05$ (Wilcoxon rank-sum test) and LDA>|2.0| are shown. **E.** Relative abundance of the differentially enriched genera identified by LEfSe analysis. The exact p-value of the pairwise comparisons between CRs and non-CRs are also shown (Wilcoxon rank-sum test). **F.** Volcano plot showing the significantly enriched bacterial amplicon sequence variants (ASVs) ($p<0.05$) by the DEseq2 analysis. The names of the significantly enriched bacterial ASVs classified up to the genus level are reported in Supplementary Table 2.

Figure 3

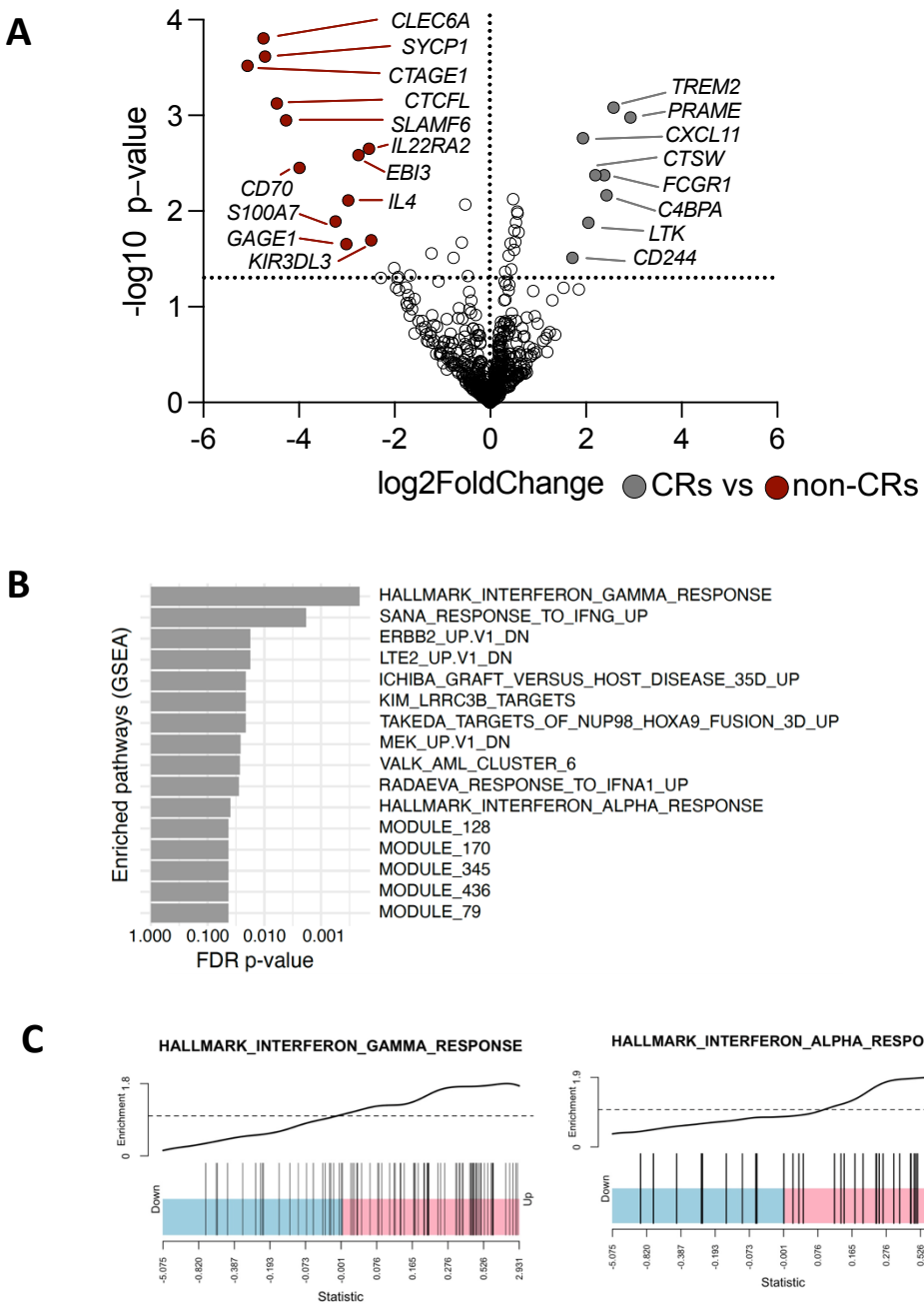


Figure 3: Responsiveness to nCRT is associated with upregulation of immune cell-related genes and activation of IFN-γ and IFN-α responses. A. Volcano plots depicting differentially expressed genes in CRs and in non-CRs. **B.** Significantly enriched transcriptional pathways (FDR-corrected p < 0.05) by GSEA analysis of DEGs. **C.** Enrichment score profiles of IFN-γ and IFN-alpha gene sets.

Figure 4

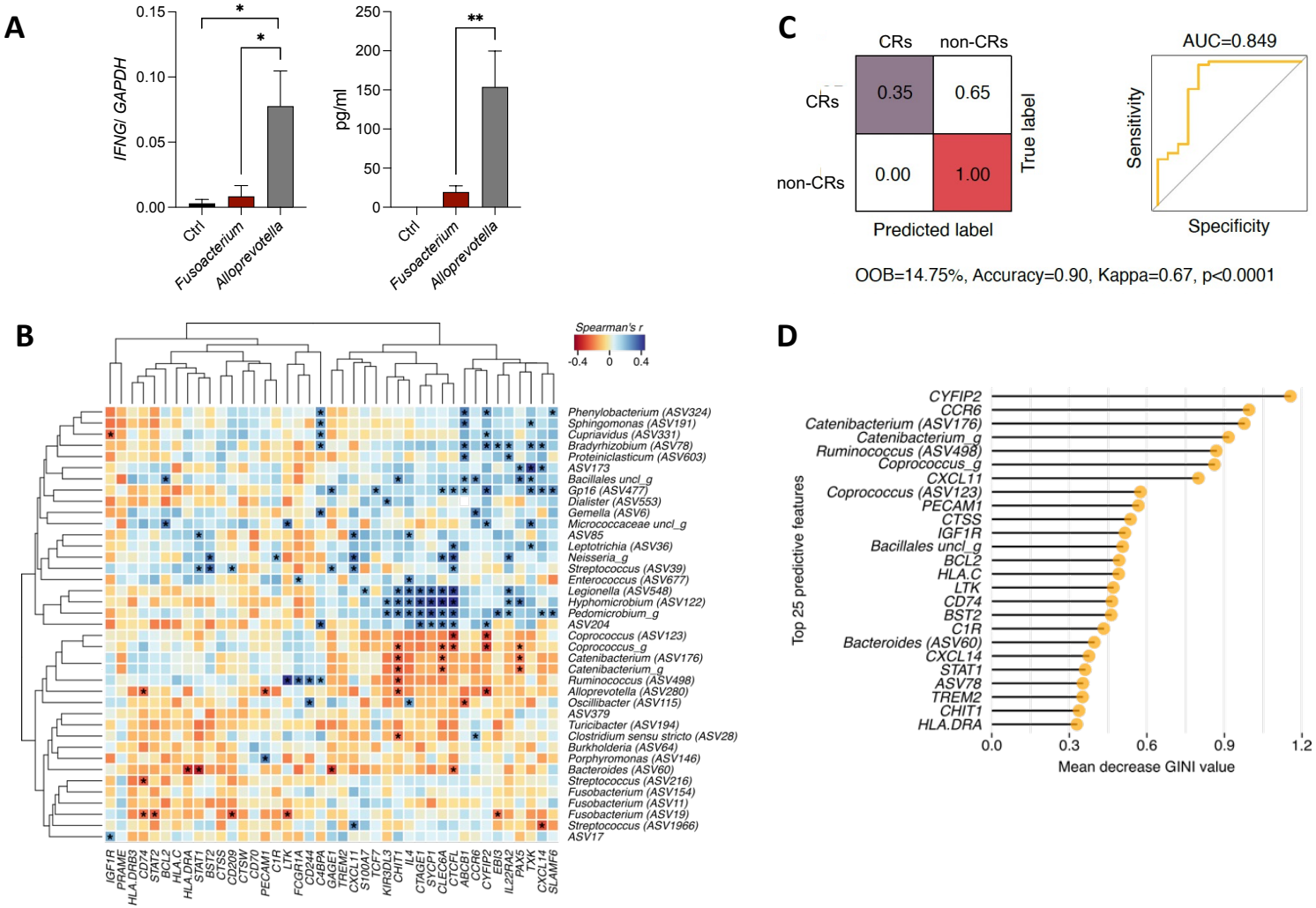
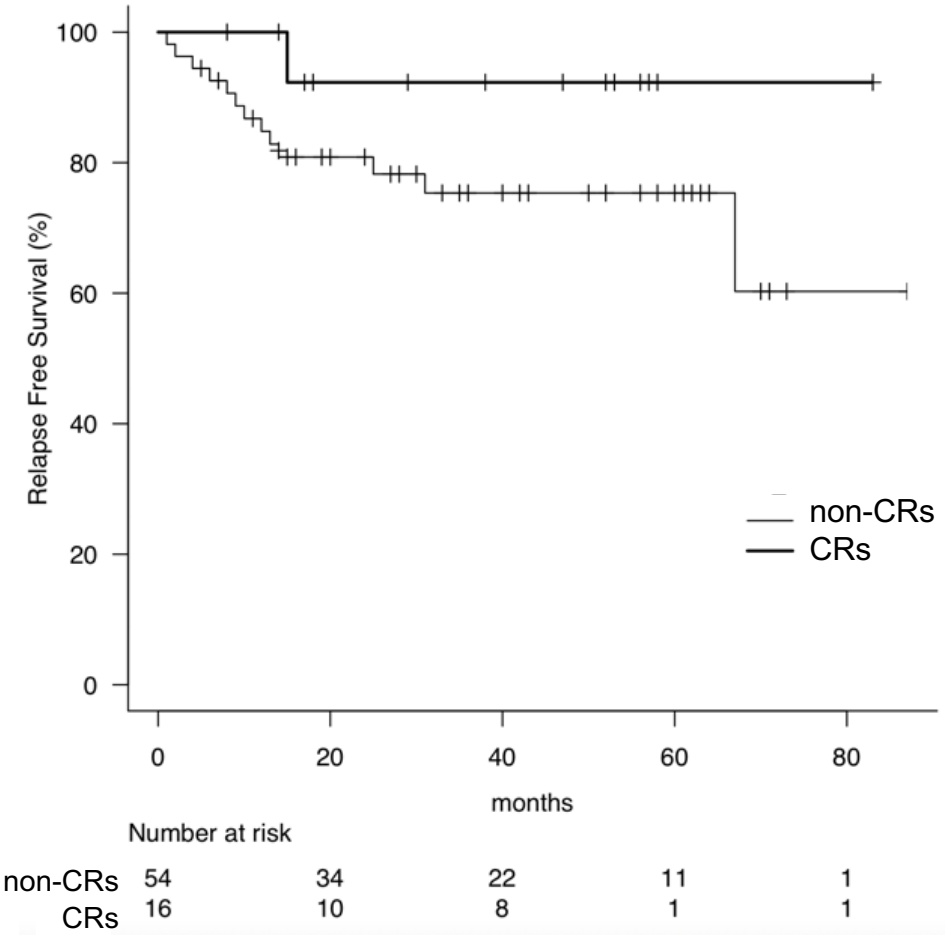
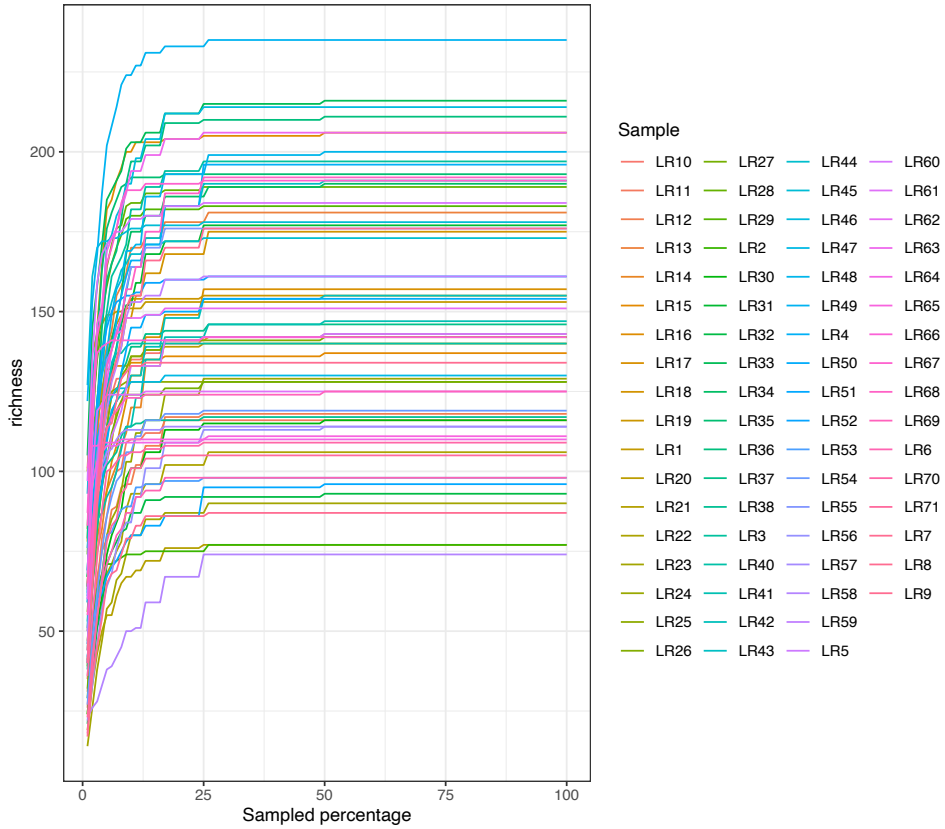


Figure 4: Correlation and predictive power of tumor-associated microbial and immunological signatures to nCRT. A. IFN-γ expression in PBMC unstimulated (Ctrl) or stimulated with the Alloprevotella or Fusobacterium, assessed at gene level after 6h by quantitative PCR (left panel) and at protein level after 24h ELISA (right panel). **B.** Heatmap of Spearman's rho correlations between the relative abundance of the significantly enriched bacterial taxa identified by LEfSE and DEseq2 analysis with the DEGs identified by Nanostring transcriptional profiling. The significant correlations with p-value < 0.05 are indicated with an asterisk. **C.** Statistics, confusion matrix and ROC curve representing the diagnostic accuracy of the Random Forest classification model. **D.** Top 25 features with the highest discriminatory power sorted by mean decrease GINI value from Random Forest analysis



Number at risk

non-CRs	54	34	22	11	1
CRs	16	10	8	1	1


richness
200
150
100
50
0
25
50
75
100
Sampled percentage
Sample
LR10 LR27 LR44 LR60
LR11 LR28 LR45 LR61
LR12 LR29 LR46 LR62
LR13 LR2 LR47 LR63
LR14 LR30 LR48 LR64
LR15 LR31 LR49 LR65
LR16 LR32 LR4 LR66
LR17 LR33 LR50 LR67
LR18 LR34 LR51 LR68
LR19 LR35 LR52 LR69
LR1 LR36 LR53 LR6
LR20 LR37 LR54 LR70
LR21 LR38 LR55 LR71
LR22 LR3 LR56 LR7
LR23 LR40 LR57 LR8
LR24 LR41 LR58 LR9
LR25 LR42 LR59
LR26 LR43 LR5