

UNIVERSITY OF MODENA AND REGGIO EMILIA

PhD Programme of Clinical and Experimental Medicine

Curriculum: Translational Medicine

XXXIV Doctorate Cycle

Director of the PhD Programme: Prof. Giuseppe Biagini

Metastatic melanoma and renal cell carcinoma: focus on the role of CD8+ T cells in responses to immune checkpoint inhibitors

Supervisor:

Prof. Andrea Cossarizza

PhD Candidate:

Domenico Lo Tartaro

Co-Supervisor:

Prof. Giorgio De Santis

Academic year 2020/2021

TABLE OF CONTENTS

ABSTRACT (ENGLISH VERSION)	4
ABSTRACT (ITALIAN VERSION)	6
1. INTRODUCTION	9
1.1 CD8+ T CELLS IN CANCER IMMUNITY	9
1.2 CD8+ T CELL EXHAUSTION	14
1.3 IMMUNE CHECKPOINT INHIBITOR	17
1.4 THE ROLE OF TUMOR IMMUNE MICROENVIRONMENT (TIME) IN CANCER IMMUNOTHERAPY	21
1.5 THE ROLE OF MUCOSAL ASSOCIATED INVARIANT T CELLS (MAIT) IN CANCER IMMUNOTHERAPY	23
1.6 CD8+ T STEM CELL MEMORY IN CANCER IMMUNOTHERAPY	27
1.7 SINGLE-CELL APPROACHES TO IMMUNE PROFILE	29
1.8 COMPUTATIONAL ANALYSIS OF HIGH-DIMENSIONAL FLUORESCENT DATA	34
A) <i>PhenoGraph</i>	36
B) <i>FlowSOM</i>	37
C) <i>FAUST</i>	39
1.9 COMPUTATIONAL ANALYSIS OF HIGH-DIMENSIONAL SCRNA-SEQ DATA	42
2. AIM OF THE STUDY 1 (MELANOMA)	46
3. AIM OF THE STUDY 2 (MRCC)	47
4. METHODS 1 (MELANOMA)	48
4.1 PATIENTS' SELECTION	48
4.2 SAMPLES COLLECTION	49
4.3 POLYCHROMATIC FLOW CYTOMETRY	49
4.4 UNSUPERVISED FLOW CYTOMETRY DATA ANALYSIS	51
4.5 CELL SORTING AND SINGLE-CELL RNA-SEQUENCING (SCRNA-SEQ) LIBRARY PREPARATION	51
4.6 SCRNA-SEQ ANALYSIS	52
4.7 SURFACE PROTEIN IMPUTATION	53
4.8 IN SILICO ANALYSIS	53
4.9 STATISTICAL ANALYSIS	54
4.10 DATA AVAILABILITY	54
4.11 CODE AVAILABILITY	54
5. METHODS 2 (MRCC)	55
5.1 PATIENTS' SELECTION	55
5.2 SAMPLES COLLECTION	56
5.3 POLYCHROMATIC FLOW CYTOMETRY	56
5.4 HIGH-DIMENSIONAL FLOW CYTOMETRY DATA ANALYSIS USING FULL ANNOTATION USING SHAPE-CONSTRAINED TREES (FAUST)	57
5.5 STATISTICAL ANALYSIS	58
6. RESULTS 1 (MELANOMA)	60
6.1 A HIGHER PROPORTION OF ACTIVATED EFFECTOR MEMORY CD8+T CELLS IN RESPONDERS	60
6.2 MAIT CELLS ARE MORE ABUNDANT IN RESPONDERS AS REVEALED BY SCRNA-SEQ	64
6.3 ACTIVATED MAIT CELLS WITH HOMING PROPERTIES ARE MORE ABUNDANT IN RESPONDERS	71
6.4 LEVEL OF MAIT CELLS BEFORE THERAPY IDENTIFIES RESPONDER PATIENTS	72
7. RESULTS 2 (MRCC)	81
7.1 MRCC PATIENTS DISPLAYED A DECREASE OF TSCM AND INCREASE OF EM CELL PERCENTAGES AFTER THERAPY	81
7.2 ICI THERAPY INDUCE STRONG CD8+ T CELL ACTIVATION IN MEMORY SUBSET BUT NOT IN MORE UNDIFFERENTIATED POPULATION	94
7.3 CYTOKINE PRODUCTION IS INCREASED IN MORE UNDIFFERENTIATED CELLS AFTER THERAPY	94

8.	DISCUSSION 1 (MELANOMA)	98
9.	DISCUSSION 2 (MRCC)	101
10.	GENERAL DISCUSSION AND CONCLUSIONS.....	104
11.	REFERENCES	107
12.	SUPPLEMENTARY TABLES	119
13.	ACKNOWLEDGEMENTS.....	123

ABSTRACT (ENGLISH VERSION)

CD8+ T lymphocytes play a central role in immunity to cancer through their capacity to kill malignant cells. However, prolonged exposure to cognate antigens in tumor microenvironment contributes to induce severe CD8+ T cell exhaustion (Tex). During the last decade cancer treatment has been revolutionized by immune check point inhibitors (ICIs) that block the activity of inhibitor receptors, such as PD-1 present on the surface of Tex cells, reinvigorating them. Despite observations of durable responses to ICI therapy, not all patients respond to the treatment. Therefore, we focused our attention on identifying alterations that occur in circulating CD8+ T lymphocytes of metastatic melanoma (MM) and metastatic renal-cell carcinoma (mRCC) patients during treatment with anti-PD1, in order to understand how and why some patients respond or not to ICI therapy.

The first cohort comprised 28 patients with MM, 17 of whom were defined responders (R) whereas 11 non-responders (NR). Cryopreserved peripheral blood mononuclear cells (PBMCs), obtained prior to initiating therapy (T0) after the first (T1) and second therapy cycles (T2), were studied by 30 parameter high-dimensional flow cytometry in combination with single cell RNA-sequencing (scRNAseq). The second cohort comprised 19 patients with mRCC, 5 of which were defined responders (R) whereas 14 non-responders (NR). Cryopreserved PBMCs, obtained prior to initiating therapy (T0) after the first (T1), second (T2) and third therapy cycle (T3), were studied by 28 flow cytometry and analyzed using Full Annotation Using Shape-constrained Trees (FAUST), a novel non-parametric method for unsupervised discovery of cell population.

In R MM patients we found an increase of proliferating effector memory (EM) cells expressing high level of Ki67, ICOS, CD95, HLA-DR, CD71, CD98, CXCR6, granzysin and CD38, both before and after first and second cycle of ICI therapy. scRNA-seq revealed the presence of activated mucosal associated invariant T (MAIT) cells expressing CD69 and CXCR4, and their increase in R compared to NR MM patients, before and after ICI therapy. Along with increased percentage of peripheral MAIT cells we also observed greater ability to produce interferon gamma (IFN- γ) and granzyme-B (GRZM-B) in R patients, before starting therapy but not after. *In silico* analysis of public datasets revealed the presence of MAIT cells within primary and metastatic lesions and their increased level within lesions regressing after ICI. Finally, to associate our finding with clinical outcomes, we correlated the median level (1.7% of CD8+ T cells) of circulating MAIT cells with the response to therapy. Using this value as a threshold, MM patients who exhibited MAIT frequency above the threshold showed a better response to therapy.

Flow cytometry reveals that CD8+ T cells from mRCC patients could be classified in 61 cell clusters. Due to the relatively low number of patients responding to therapy, we found similar percentages of clusters in R and NR at all timepoints. However, considering all patients treated with anti-PD1 (n=19), therapy induced a redistribution of different subpopulations of CD8+ T cells. In particular, T stem cell memory (T_{SCM}) decreased after the third cycle of therapy. On the contrary, the EM compartments increased after therapy. In terms of functionality, we observed that, if compared with EM, after therapy T_{SCM} retained more polyfunctionality, producing simultaneously tumor necrosis factor (TNF) and IFN- γ .

In conclusion, we provide evidence that MM and mRCC patients differently respond to anti-PD1 therapy, the MM are characterized by more activated MAIT in R patients, while the mRCC are characterized by a pool of circulating likely-exhausted T_{SCM} in both R and NR.

ABSTRACT (ITALIAN VERSION)

I linfociti T CD8+ giocano un ruolo centrale nell'immunità al cancro mediante la loro capacità di uccidere le cellule maligne. Tuttavia, la prolungata esposizione agli antigeni nel microambiente tumorale contribuisce a indurre un grave esaurimento delle cellule T CD8+ (Tex). Durante gli ultimi due decenni il trattamento del cancro è stato rivoluzionato dagli inibitori dei checkpoint immunitari (ICIs), che bloccano l'attività dei recettori inibitori, come il PD-1, presenti sulla superficie delle cellule Tex, rinvigorendole. Nonostante l'osservazione di risposte durature alla terapia ICI, non tutti i pazienti rispondono al trattamento. Abbiamo focalizzato la nostra attenzione sull'identificazione delle alterazioni che si verificano nei linfociti T CD8+ circolanti di pazienti con melanoma metastatico (MM) e carcinoma renale metastatico (mRCC) durante il trattamento con anti-PD1, al fine di capire come e perché alcuni pazienti rispondono o no alla terapia.

La prima coorte comprendeva 28 pazienti MM, 17 dei quali definiti responsivi (R) e 11 non responsivi (NR) alla terapia. Le cellule mononucleate da sangue periferico (PBMCs) crioconservate, ottenute prima dell'inizio della terapia (T0) dopo il primo (T1), secondo (T2) e terzo ciclo di terapia (T3), sono state studiate mediante citometria a flusso a 30 parametri in combinazione con il sequenziamento dell'RNA a singola cellula (scRNA-seq). La seconda coorte comprendeva invece 19 pazienti mRCC, 5 dei quali definiti responsivi (R) e 14 non responsivi (NR). Le cellule mononucleate da sangue periferico (PBMCs) crioconservate, ottenute prima dell'inizio della terapia (T0) dopo il primo (T1), secondo (T2) e terzo ciclo di terapia (T3), sono state studiate mediante citometria a flusso a 30 parametri e analizzate utilizzando Full Annotation Using Shape-constrained Trees (FAUST), un nuovo metodo non parametrico per la scoperta non supervisionata di popolazioni cellulari.

Nei pazienti R con MM abbiamo riscontrato un aumento delle cellule effettrici di memoria (EM) proliferanti esprimenti alti livelli di Ki67, ICOS, CD95, HLA-DR, CD71, CD98, CXCR6, granulicina e CD38, sia prima che dopo il primo e il secondo ciclo di terapia. L'scRNA-seq ha rivelato la presenza delle cellule T invarianti associate alla mucosa (MAIT) attivate esprimenti il CD69 e CXCR4 e il loro aumento nei pazienti MM R rispetto ai NR, sia prima e dopo la terapia con ICI. Insieme all'aumento in percentuale abbiamo anche osservato una maggiore capacità delle MAIT nei pazienti R di produrre interferone-gamma (IFN- γ) e granzima-B (GRZM-B), solo al T0. L'analisi *in silico* di dataset pubblici ha rivelato la presenza delle MAIT all'interno delle lesioni primarie e metastatiche, e dopo terapia il loro livello era aumentato nelle lesioni che regredivano. Infine, abbiamo correlato il livello mediano delle cellule MAIT circolanti (1,7% delle cellule T CD8+) con la risposta alla terapia.

Utilizzando questo valore come soglia, i pazienti MM che presentavano una percentuale di MAIT al di sopra della soglia hanno mostrato una risposta migliore alla terapia con ICI.

L'analisi delle cellule T CD8⁺ dei pazienti con mRCC ha rivelato la presenza di 61 clusters. A causa del numero basso di pazienti che rispondono alla terapia, abbiamo trovato percentuali simili tra i clusters dei pazienti R e NR sia prima che dopo terapia. Tuttavia, analizzando insieme tutti i pazienti trattati con anti-PD1 (n=19), si osserva come la terapia abbia indotto una ridistribuzione di diverse sottopopolazioni di cellule T CD8⁺. In particolare, le cellule T staminali di memoria (T_{SCM}) sono diminuite dopo il terzo ciclo di terapia. Al contrario, le sottopopolazioni di cellule EM sono aumentate in percentuale dopo la terapia. In termini di funzionalità, abbiamo osservato che, se comparate con le cellule EM, le T_{SCM} dopo la terapia mantengono una più alta polifunzionalità, producendo simultaneamente il fattore di necrosi tumorale (TNF) e l'IFN- γ .

In conclusione, abbiamo fornito l'evidenza che i pazienti MM e mRCC rispondono in modo diverso alla terapia con anti-PD1, gli MM sono caratterizzati da più alti livelli di cellule MAIT attivate nei pazienti R, mentre gli mRCC sono caratterizzati da un pool probabilmente esaurito di T_{SCM} circolante sia nei responder che nei non responder.

Box 1: ABBREVIATION INDEX

AIDS : acquired immunodeficiency syndrome; **AKT** : Protein kinase B; **ANOVA** : Analysis of Variance; **APC** : antigen presenting cells; **ATAC-seq** : Transposase-Accessible Chromatin using sequencing; **BCC** : basal cell carcinoma; **BCL** : base call; **BTC** : biliary tract cancer; **CAR** : chimeric antigen receptor; **CCL** : chemokine (C-C motif) ligand; **CCR** : chemokine (C-C motif) receptor; **CD(number)**: Cluster of Differentiation; **CITE-seq** : Cellular Indexing of Transcriptomes and Epitopes by Sequencing; **CM** : central memory; **CTLA-4** : Cytotoxic T-Lymphocyte Antigen 4; **CODEX** : Co-detection by Indexing; **CRC** : colorectal cancer; **CRISPR** : Clustered Regularly Interspaced Short Palindromic Repeats; **CSV** : comma-separated value; **CTL** : cytotoxic T lymphocytes; **CREM** : cAMP responsive element modulator; **CXCL** : chemokine (C-X-C motif) ligand; **CXCR** : chemokine (C-X-C motif) receptor; **DC** : dendritic cell; **DMSO** : Dimethyl sulfoxide; **DNA** : deoxyribonucleic acid; **cdNA** : complementary deoxyribonucleic acid; **DP** : double positive; **DUSP** : Dual-specificity phosphatases; **EDTA** : ethylenediaminetetraacetic acid; **EFHD** : EF-Hand Domain Family Member D; **EM** : effector memory; **ENTPD** : Ectonucleoside triphosphate diphosphohydrolase; **EOMES** : Eomesodermin; **FACS** : fluorescence-activated single cell sorting; **FAUST** : Full Annotation Using Shape-constrained Trees; **FBS** : fetal bovine serum; **FCS** : Flow Cytometry Standard; **FDA** : Food and Drug Administration; **FGFBP2** : fibroblast growth factor- binding protein 2; **FOXO1** : Forkhead box O1; **FOXP3** : forkhead box P3; **GEO** : Gene Expression Omnibus; **GITR** : glucocorticoid-induced TNFR-related protein; **GLMM** : generalized linear mixed model; **GMB** : glioblastoma multiforme; **GNLY** : granulysin; **GRZM(A-B-H-K)** : granzyme(A-B-H-K); **HCC** : hepatocellular carcinoma; **HEPES** : 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; **HIV** : human immunodeficiency virus; **HLA** : human leukocyte antigen; **ICI** : immune check point inhibitor; **ICOS** : Inducible T Cell Costimulator; **IDO** : Indoleamine-2,3-dioxygenase; **IFN** : Interferon; **IMC** : Imaging Mass Cytometry; **IMDC** : international metastatic RCC database consortium; **IRs** : Inhibitory receptors; **IRF4** : Interferon Regulatory Factor 4; **ITIM** : Immunoreceptor Tyrosine-based Inhibition Motif; **ITSM** : immunoreceptor tyrosine-based switch motif; **KNN** : k-nearest neighbors; **KLR(B-G)** : killer cell lectin-like receptor; **LAG-3** : lymphocyte activation gene 3; **LCMV** : lymphocytic choriomeningitis virus; **LDH** : lactate dehydrogenase; **LYZ** : lysozyme; **MAIT** : mucosal associated invariant T; **MAPK** : mitogen-activated protein kinase; **MB-DNN** : multiple branch deep neural network; **MDR1** : Multi-Drug Resistance Gene 1; **MDSC** : myeloid-derived suppressor cells; **MFI** : median fluorescence intensity; **MHC** : major histocompatibility class; **MM** : metastatic melanoma; **mRCC** : metastatic renal cell carcinoma; **MPEC** : memory precursor; **MR1** : major histocompatibility complex class I-related gene protein; **MST** : minimum spanning tree; **NFAT** : nuclear factor of activated T-cells; **NSCLC** : non-small cell lung carcinoma; **ORR** : objective response rate; **PBMC** : peripheral blood mononuclear cell; **PBS** : phosphate-buffered saline; **PC** : principal components; **PCA** : principal component analysis; **PCR** : polymerase chain reaction; **PD-1** : programmed cell death protein 1; **PLZF** : promyelocytic leukemia zinc finger protein; **PMA** : phorbol myristate acetate; **PRF** : perforin; **QC** : quality control; **ROR-t** : RAR-related orphan receptor t; **ROS** : reactive oxygen species; **SSN** : shared nearest neighbor; **SD** : stable disease; **SEM** : standard error of measurement; **SHP2** : src homology region 2 domain-containing phosphatase-2; **SIRT1** : sirtuin 1; **SLO** : secondary lymphoid organs; **SNP** : single nucleotide polymorphisms; **SOM** : Self-Organizing Maps; **SPICE** : Simplified Presentation of Incredibly Complex Evaluation; **TAA** : tumor-associated antigens; **TAM** : tumor-associated macrophages; **TCR** : T cell receptor; **TE** : terminal effector; **Teff** : effector CD8+ T cells; **Tex** : T cell exhausted; **TIGIT** : T cell immunoglobulin and ITIM domain ; **TILs** : tumor-infiltrating lymphocytes; **TIM-3** : T cell immunoglobulin and mucin domain-containing 3; **TIME**: tumor immune microenvironment; **TLS** : tertiary lymphoid structures; **TM** : transitional memory; **TME** : tumor microenvironments; **T_{RM}** : tissue-resident memory T cell; **t-SNE** : t-distributed stochastic neighbor embedding; **Tmem** : long-term memory CD8+ T cells; **Tpe** : T cell pre-exhausted; **TSA** : tumor-specific antigens; **UMAP** : Uniform Manifold Approximation and Projection; **UMI** : unique molecular identifier; **VISTA** : short for V-domain Ig suppressor of T cell activation.

1. INTRODUCTION

1.1 CD8+ T cells in cancer immunity

During the last two decades, the immune system's role in controlling tumor growth was in part deciphered. The description of the correlation between the immune presence and clinical benefit in patients with cancer strongly suggests that the immune system plays a crucial role in killing malignant cells. In particular, the more the neoantigen load, the better the response to therapy. However, the intensity and efficacy of the immune response can vary upon neoantigens' clonality. It seems that neoantigens derived from clonal mutations, which appear early during tumorigenesis, may elicit more effective antitumor responses than neoantigens derived from sub-clonal mutations, which are acquired later in tumorigenesis [1]. This means that intratumoral heterogeneity may impact the response to new therapies such as immune checkpoint inhibitor (ICI). To reinforce this notion, clinical epidemiology supports the existence of an antitumoral immune response in some types of disease. Firstly, evidence for immunosurveillance can be found in patients with acquired immunodeficiencies, like that caused by the human immunodeficiency virus (HIV), the cause of AIDS, who has an increased frequency of virus-associated malignancies, including Kaposi's sarcoma, lymphomas, urogenital cancers, and cervical cancers due to different strains of papillomavirus [2]. Secondly, higher cancer prevalence has been observed in transplanted recipients treated with immunosuppressive drugs. Immunosuppression to prevent transplant rejection is associated with a 3- to 100-fold increased risk of developing certain types of cancer, mainly lymphomas [3]. Since Virchow noted the presence of lymphocytes infiltration in solid tumors in 1863 [4], additional support for cancer immunosurveillance and immunoediting is evident in uncountable reports that describe the presence of immune cells infiltrating the tumors and that correlate their frequency with patient prognosis. The presence of T cells inside tumors was observed, in the late 1990s, in patients with melanoma and then described for several other malignancies, including ovarian, colorectal (CRC), or lung cancer [5-7]. Afterward, great attention has been paid to investigating the role of tumor infiltrating lymphocytes (TILs), particularly CD8+ T cells. This effort resulted in identifying CD8+ TILs frequency as a bona fide indicator of improved prognosis and increased overall survival for several types of tumors.

The generation of efficient adaptive immune response arises in secondary lymphoid organs (SLO) where the professional antigen-presenting cells (APC), i.e., mature dendritic cells (DCs), present antigens to CD4 and CD8+ T cells [8]. In particular, DCs in tumor sites after tumor antigens

encounter and after receiving adequate stimulus, such as inflammatory cytokines, begin the maturation process by starting the lysosomal proteolysis of antigen (Ag) and migrating to secondary lymphoid organs [9]. Tumor antigens are antigenic substances produced by tumor cells, able to trigger an immune response in the host. There are two major classes of tumor antigen which are targeted by T cell immunotherapies 1) tumor-specific antigens (TSA), found on cancer cells only, not on healthy cells 2) tumor-associated antigens (TAA), which have elevated levels on tumor cells, but are also expressed at lower levels on healthy cells [10]. In order to ensure safety and effectiveness, immunotherapy strategies should therefore elicit efficient T cell responses against antigenic peptides that are present on tumor cells but not on healthy cells to avoid autoimmune side effects. In that regard, antigens encoded by mutated genes are among the safest. In SLO, the DCs express the processed Ag on the cell surface through the major histocompatibility class I or II (MHC I-II); at this point, the complex Ag-MHC may be bound by T cell receptor (TCR) expressed on the surface of T cells [9]. These steps allow lymphocytes proliferation and differentiation into effector T (Teff) cells and their migration into the tumor, leading to the destruction of malignant cells. As well as reported in the literature, cytotoxic T cells expressing cell-surface CD8⁺ are the most potent effectors in the anticancer immune response and represent the backbone for a successful novel cancer treatment, such as immune checkpoint inhibitor (ICI) therapy [11]. In particular, CD8⁺ T cells, also called cytotoxic T lymphocytes (CTL), are able to kill infected or cancer cells using three primary mechanisms:

1. Secretion of cytokines, primarily tumor necrosis factor-(TNF) and interferon-(IFN)- γ , which have antitumor effects;
2. Releasing of granules containing cytotoxic proteins such as perforin and granzymes;
3. Inducing of apoptosis via Fas/FasL interactions.

Specifically, activated CD8⁺ T cells express a high level of FasL on cell-surface, which may be engaged by Fas (also known as CD95) expressed on malignant cells [12].

Naive CD8⁺ T cell is considered immature and, unlike activated or memory T cells, has not encountered its cognate antigen within the periphery. Their function is to patrol tissues, recirculating continuously from the blood to lymph nodes or other secondary lymphoid tissues looking for novel pathogens and cancer cells. After cognate Ag-encounter, naïve CD8⁺ T cells proliferate by clonal expansion and differentiate into effector cells, which express receptors that enable them to migrate and enrich the tumor region [13]. We classified peripheral subsets of CD8⁺

T cells based on the distinct expression of cell-surface markers such as CCR7 (CD197), CD62L, CD45RA, CD45RO, CD28, CD27, CD95, CD161, CXCR5, and CD38. In a pivotal study analyzing the surface expression of CD45RA and CD27, the presence of three discrete subsets identified as naïve (CD45RA⁺ CD27⁺), memory (CD45RA⁻ CD27⁺), and effector (CD45RA⁺ CD27⁻) was observed among the peripheral CD8⁺ T cells [14]. Interestingly, compared to naïve, memory cells were able to secrete not only interleukin (IL)-2 but also interferon IFN- γ but not to exert effector function. On the other hand, effector cells had high cytolytic activity and the ability to produce IFN- γ without in vitro pre-stimulation. Moreover, two subsets of memory cells have been identified so far, named central memory (CM, CCR7⁺CD62L⁺CD45RA⁻) with lymph node homing propriety and limited effector functions, and effector memory (EM, CCR7⁻CD62L⁻CD45RA⁻) with high effector function and one subset of terminal effector (TE, CCR7⁻CD62L⁻CD45RA⁺) CD8⁺ T cells [15]. The development of techniques for measuring a higher number of markers at the single-cell level allowed detailed phenotypic analysis. Indeed, other peripheral memory subsets have been identified, such as transitional memory (TM), which constitutes a subset that appears to be more differentiated than CM cells but not as fully differentiated as EM cells in terms of phenotype and effector function. Another recently characterized peripheral CD8⁺ T cell subset is represented by a mucosal-associated invariant T cell (MAIT), a class of innate-like T cells expressing a semi-invariant TCR and CD161, showing effector phenotype and unique antigen specificity conferring them a potential antitumor role. Furthermore, over the past decade, the idea that memory T cells might not be confined solely to the CD45RA⁻ T cell compartment led to the discovery of T stem cell memory (TSCM), which represent a rare subset endowed with stem cell-like ability to self-renewal and a total capacity to reconstitute the entire composition of T cell subsets [16]. Interestingly, TSCM also showed superior antitumor responses in a humanized mouse model compared with both CM and EM, thus opening their use in adoptive T cell therapy to treat cancer and infectious diseases [16].

Although a high number of published human immunological studies have been performed on CD8⁺ TILs, the ability to monitor anti-tumor CD8⁺ T cell responses in the blood has tremendous therapeutic potential. A recent study showed a positive correlation with TCR repertoire and durable response to anti-PD1 alone or in combination with anti-CTLA-4. In particular, they found that responding patients have more large clones (those occupying > 0.5% of repertoire) post-treatment than non-responding patients or controls, and clinical response is associated with the large CD8⁺ T cell clone count three weeks after treatment [17]. On the other hand, the recent finding showed that the frequency of circulating PD-1⁺TIGIT⁺ CD8⁺ T-cells after one month of anti- PD-1 therapy

was associated with clinical response and increased overall survival [18]. Other striking results of single-cell technologies have been obtained in blood samples from cancer patients treated with ICI. In those with melanoma, circulating Ki67+, CD8+ T cell response was correlated with tumor burden [19]. Similar results were found in non-small cell lung carcinoma (NSCLC) patients treated with anti-PD-1. After therapy, increased Ki-67+ PD-1+ CD8+ T cells were observed in responding patients displaying an effector-like phenotype (HLA-DR+, CD38+, Bcl-2low), costimulatory molecules (CD28+, CD27+, ICOS+), high levels of PD-1, and co-expression of CTLA-4. Tracking TCR clones and transcriptional phenotypes in basal cell carcinoma (BCC) revealed clonal expansion of CD8+ CD39+ T cells, which co-expressed markers of chronic T cell activation and exhaustion. However, in this case, the expansion of T cell clones did not derive from pre-existing TILs, but from novel clonotypes that had not previously been observed in the same tumor [20]. TCR is a heterodimer composed of two chains, α , and β , resulting from genetic recombination of the V(D)J genes and is responsible for the specificity of each T cell against cognate antigens. Generally, the diversity of TCR $\alpha\beta$ repertoire is associated with efficient protection against several pathogens [21], and more recently, the clonality of both peripheral blood and tumor TCR $\alpha\beta$ repertoire has also been associated with the improved clinical outcome under anti-PD-1 or anti-CTLA-4 immunotherapy [22-24]. This evidence suggests that the response to anti-PD-1 depends on the intrinsic capacity of tumors to recruit novel T cell clones, which replace pre-existing exhausted T cells that have insufficient capacity to reinvigorate in response to therapy [25].

Our current knowledge of the CD8+ T cell states present in human tumors and the specific role of different CD8+ T cell populations is far from being completely unraveled. Indeed, for many years it has been thought that TILs were a homogenous group of effector-cytotoxic cells that produce a large amount of IFN- γ and the protease granzyme (GRZM)-B. Conversely, a recent finding has highlighted the existence of more than one TIL subsets (also referred to as Tc) in tumor microenvironments (TME), each one with a distinct effector or cytotoxic function. These CD8+ Tc subsets comprise IFN- γ producing Tc1, interleukin (IL)-4 producing Tc2, IL-9 producing Tc9, IL-17 producing Tc17, and IL-22 producing Tc22. In particular, Tc1, Tc2, and Tc22 were cytotoxic subsets identified by distinct expression of transcription factor and effector cytokines. Specifically, Tc1 expresses TBET, EOMES, and STAT4 and a high level of granzyme, while Tc2 and Tc22 express GATA3 and STAT6 accompanied with a high level of granzyme production [26,27]. On the other hand, Tc9 and Tc17 express a low level of GRZM-B and have a poor cytolytic function [28,29]. To date, only Tc1, Tc2, Tc17, and Tc22 were detected in the TME. In particular, the IFN- γ -producing Tc1 was the most described TIL subset,

probably due to their exceptional cytotoxic potential combined with the large amount of IFN- γ released. Indeed, the presence of Tc1 cells in cancer tends to correlate with a more favorable prognosis [30]. Despite the literature reporting predominantly Tc1 cells, other Tc has been detected in cervical and ovarian cancer, such as Tc2 or Tc22, respectively [26,31]. Tc2 increase in TME remains to be further investigated, even though it was reported that their percentage might negatively correlate with tumor progression [32]. On the other side, the highest level of Tc22 CD8⁺ T cells in TME patients with ovarian cancer strongly and positively correlates with recurrence-free survival. A more recently described Tc subsets in TME was Tc17 [33]; an interesting study reported that Tc17 was enriched in lung cancer patients responding to anti-PD-1 therapy. These cells displayed tissue-resident memory (TRM) characteristics in terms of transcriptome and protein expression of CD103, CD69, CD49a, PD-1, pSTAT3, and Aiolos. TRM was defined as memory lymphocyte lineage that occupies tissues without recirculating, providing the first response against infections or tumors. A recent study also revealed the presence of CD8⁺ CXCR5⁺ TIM-3⁻ TILs with stem-like qualities, able to exert anti-tumor immunity through their rapid production of effector molecules such as IFN- γ and TNF [34]. To date, it is unclear whether all the TIL subsets detected within the TME are polarized in SLO and recruited to the TME, or this happens directly within TME. Despite that, evidence supporting the latter hypothesis exists, suggesting that the tertiary lymphoid structure within TME directly may be involved in the plasticity of recruited CD8⁺ T cells [35].

The development of high-dimensional profiling techniques, such as single-cell RNA sequencing (scRNA-seq), has allowed the field to go substantially beyond this relatively coarse profiling of CD8⁺ T cells based on the expression of just a few protein markers and has been used to profile T cell deeply infiltrates in human tumors. In recent studies, CD8⁺ TILs were dissected on the basis of their transcriptome in three major subpopulations. In particular, 'memory' cells (also called naïve-like) marked by the expression of the genes CC chemokine receptor 7 (CCR7), lymphoid enhancer-binding factor 1 (LEF1), interleukin-7 receptor (IL7R), TCF7 (encoding TCF-1), SELL (encoding L-selectin/CD62L); 'cytotoxic' cells expressing, among other genes, perforin 1 (PRF1), granzyme A (GZMA), GZMB and natural killer cell granule 7 (NKG7), CX3C chemokine receptor 1 (CX3CR1), killer cell lectin-like receptor subfamily G member 1 (KLRG1) and fibroblast growth factor-binding protein 2 (FGFBP2) and granulysin (GNLY) [11,36,37]. In addition to the memory/naïve-like and cytotoxic CD8⁺ T cell states, the third pool of heterogeneous T cells displaying features of 'dysfunction' or 'exhaustion' was observed. These cells are characterized by the increased expression of genes such as PDCD1 (encoding PD-1), HAVCR2 (encoding TIM-3), LAG3, CTLA4, TIGIT, and LAYN [11,36,37].

Interestingly, the frequencies of memory/naïve-like T cell in tumor tissue predicts a better response and survival in melanoma patients treated with immune checkpoint therapy [11]. Moreover, data obtained from melanoma samples and peripheral blood of patients treated with anti-CTLA-4 and anti-PD-1 revealed that could be observed treatment-specific effects. Indeed, while anti-PD-1 mainly induced the expansion of specific tumor-infiltrating exhausted-like CD8+ T cell subsets, anti-CTLA-4 led to the expansion of an ICOS+ Th1-like CD4+ effector subset other than engaging specific subsets of exhausted-like CD8+ T cells [38]. It was also reported that the population of CD8+, CD45RO+, PD-1+, TBET+, EOMES+ T cells increased after treatment only in TILs if compared to the peripheral blood [38] and that the gene expression signature of EOMES+, CD69+, CD45RO+ T cells were associated with more significant tumor shrinkage in both anti-CTLA-4 and anti-PD-1 therapies [39]. Likewise, in a cohort of patients with NSCLC treated with anti-PD-1, the presence of PD-1++ T cells within the tumor was strongly predictive for both response and survival. PD-1++ T cells indeed produce C-X-C Motif Chemokine Ligand 13 (CXCL13), which mediates immune cell recruitment to tertiary lymphoid structures (TLSs) [40]. Similarly, in a cohort of patients with GMB treated with anti-PD-1, an enhanced expression of chemokine transcripts, higher immune cell infiltration, and augmented TCR clonal diversity among tumor-infiltrating TILs was reported [41]. In summary, a large variability can be observed among different patients' cohorts concerning the abundance of different T cell functional states. After therapy with ICI, an increase in CD8+ T cells with an effector-like phenotype expressing inhibitory/costimulatory molecules and proliferation markers has been described in several cancer settings. However, only in a few cases, this cellular immune response was correlated with a measurable clinical response.

1.2 CD8+ T cell exhaustion

The term 'exhausted' is referred to a distinct CD8+ T cell development stage that occurs in animals and humans during chronic infection and cancer. The concept of memory T cell exhaustion was developed in the 1960s, but only in 1993, this biological effect was linked in lymphocytic choriomeningitis virus (LCMV) infected mice with persistent Ag-stimulation [42].

Commonly, during acute infection, naïve CD8+ T cells develop first into memory precursor (MPEC) expressing KLRG-1^{lo} and CD127^{hi} and then into effector subset (Teff, also known as cytotoxic) expressing KLRG-1^{hi} and TBET+ [43]. Teff cells represent the critical element in controlling the disease thanks to their high cytotoxic capacity to produce IFN- γ , granulysin (GNLY), perforin (PRF), and multiple granzyme family molecules GRZM-A, GRZM-K, GRZM-B, GRZM-H, and GRZM-M. In

particular, during activation and response, Teff transiently expresses inhibitory receptors, such as PD-1, and switch from mitochondrial oxidative phosphorylation driving quiescent state to glycolysis to meet their increased energetic demand [43]. After clearance of the primary infection, most Teff dies by apoptosis, but 5–10% survive and differentiate into long-term memory CD8⁺ T cells (Tmem). Instead, during chronic Ag-stimulation, the cells' development stage is slightly different, driven by high antigen load and long-time antigen exposure. Indeed, MPEC cells fails to differentiate into Teff, giving rise to a different subset known as pre-exhausted (Tpe) CD8⁺ T cells expressing PD-1^{int}, TCF-1⁺ (encoded by TCF7), CD44^{hi}, and EOMES^{lo} [43]. In particular, Tpe retains at least in part both proliferative capacities expressing Ki67 and effector function producing IL-2, TNF, IFN- γ , and cytotoxic molecules [43]. Subsequently, a few weeks after the chronic infection (or cancer), Tpe progressively loses memory potential becoming terminally exhausted (Tex), expressing PD-1^{hi}, CD44^{int}, and EOMES^{hi}, but not KLRG-1, CD127, and TCF-1. Tex lost the IL-2, TNF, and IFN- γ production completely while still retaining a low level of the ability to degranulate measured by their CD107a expression [43]. There is also evidence that in vivo Tex shows unique IRs expression and epigenetics landscape if compared to Tpe. In fact, unlike the Tpe, Tex is characterized by co-expression of PD-1 and other IRs such as TIGIT, LAG-3, TIM-3, CTLA-4, ICOS, and VISTA. In particular, IRs actively compete with costimulatory receptors like CD28, CD27, OX40, CD226, and 4-1BB, preventing T cell stimulation [43]. Interestingly, a recent study revealed the presence of stem-like cytotoxic CD8⁺ T cells infiltrating human primary non-small cell lung cancer. These cells identified as CCR7⁻, CD45RA⁻, CXCR5⁺, and TIM-3⁻ are characterized by co-expression of other IRs such as TIGIT, PD-1, and low level of the transcription factor, such as EOMES (85). Functionally showed enhanced effector functions rapidly producing IFN- γ , IL-2, TNF cytokines, and the degranulation marker CD107a upon PMA/ionomycin stimulation ex vivo. Moreover, these cells expressed a high level of TCF-1 and are endowed with enhanced stemness, generating after in vitro stimulation all subsets of T cells [34]. Interestingly the CXCR5⁺, TIM-3⁻, EOMES^{lo} showed huge similarities with previously described Tpe, suggesting that CD8⁺ TILs are functionally heterogeneous, and that this heterogeneity can be linked to the response to ICI therapy.

On the other side, epigenetic studies confirm that Tex is a unique lineage of cells different from Teff, Tmem, and Tpe due to their different chromatin landscape. Specifically, Tex contains a thousand differentially accessible chromatin regions that actively control the expression of IRs and the accessibility to the region of promoters and enhancers. These regions can be bound by transcription factors (TF) such as TBET, EOMES, FOXO1, NFAT, Tox, ikzf2, Blimp-1, IRF-4, and Id3 [44].

Although some of the previous TFs are also expressed in Teff and Tmem at a low level, such as TBET and EOMES, recent studies suggest a re-evaluation of their function in Tex accompanied by the presence of other transcription factors that may precisely modulate their role. Indeed, data suggest that many other TFs or transcriptional pathways have been implicated in CD8+ T exhaustion during chronic Ag-stimulation, such as FOXO1 and NFAT [45,46]. In particular, the prolonged stimulation of TCR is accompanied by decreases in AKT and mTOR activation and enhanced intracellular calcium influx. The intracellular increase of calcium causes the nuclear localization of NFATs by altering the balance with AP-1 favoring high NFAT and low AP-1. This mechanism triggers the unconventional role of partnerless-NFAT to bind the PD-1, Lag-3, and Tim-3 promoters inducing their active transcription [45,46]. On the other hand, less activated AKT results in reduced phosphorylation of FOXO1, also accompanied by its deacetylation mediated by SIRT-1. Both post-translational modifications of FOXO1 result in a more stabilized protein that can migrate into the nucleus blocking glycolytic and cytotoxic capacity [47]. Notably, functional alterations in metabolism are established early in chronic infection before significant exhaustion manifestations occur. Indeed, early Tex cells may not meet bioenergetic demand under persisting antigen signaling due to PD-1 mediated intracellular signaling that directly constrains glucose uptake and its use [48]. These dysregulations skewed early Tex cells to an altered metabolic phenotype prone to mitochondrial depolarization and high ROS production. Metabolic alterations are not irreversible in early Tex; indeed defects could be partially reversed by improving glucose uptake, removing PD-1 signals, or enforcing the metabolic regulator PGC1 α [43]. Recent data also reinforced the evidence that partnerless-NFAT, under chronically TCR stimulation, drives the downstream development of exhausted T cells inducing the expression of nuclear factor thymocyte selection-associated high mobility group box (TOX) [49-52]. TOX expression was negatively correlated with the development of KLRG1+ terminal Teff cells and instead was associated with high expression of inhibitory receptors and key Tex cell transcription factors. As reported before, the establishment and maintenance of Tex cells depend on proliferative hierarchy mediated by TCF-1, TBET, and EOMES. Interestingly, the selective deletion of TOX in Tex cells of mice was associated with reduced expression of genes that codify for inhibitory receptors such as PDCD1, ENTPD1, HAVCR2, CD244, and TIGIT, while retained high expression of transcription factors TCF-1 [49-51,53]. Along with TOX family proteins, other factors cooperate in imposing CD8+ T cell exhaustion, among them, NRA4 transcription factors displayed a critical role in modulating the expression of IRs and chromatin accessibility [54].

Nowadays, even though the impressive efficacy of ICI drugs, a still large fraction of patients does not respond to the treatments. Despite T-cell exhaustion being associated with reduced response to standard antitumor care, current knowledge, first in tumor models and then in humans, has shown that blocking IRs, such as CTLA-4 and PD-1, mediate tumor regression. In particular, tumor-infiltrating lymphocytes, showing exhausted phenotype, are reinvigorated by anti-PD-1 or CTLA-4 treatment achieving a high response rate in melanoma, non-small cell lung carcinoma, and numerous other types of cancers [36,53]. This effect was associated at least in part with the heterogeneity of dysfunctional CD8⁺ T cells, composed of several subpopulations, including Tpe and Tex. Indeed, recent data suggest that Tpe but not Tex contains cells capable of responding to PD-1 or CTLA-4 blockade, consistent with their less functionally exhausted characteristics than Tex counterpart [56]. To date, despite the effort to dissect the molecular initiator of exhaustion and interrogate the pathways that restrain these cells from developing into T_{eff} or T_{mem}, not all underlying mechanisms have been understood, and more data are needed to unravel the complex relationship between T-cell exhaustion and chronic Ag-stimulation.

1.3 Immune checkpoint inhibitor

Before the era of immune checkpoint inhibitors, solid tumor immunotherapy was in an awful situation. The first documented use of immunotherapy dates back to 1891 when William Coley, a surgeon at Memorial Sloan Kettering, whom pioneered the first nonsurgical treatment of cancer, treats sarcoma with the toxins of bacterial skin infection to stimulate the body's natural defenses against cancer. Years after, in 1959, the immunologist Lloyd J. Old discovers that certain bacterial products could boost an animal's immunity to cancer. For the successive 27 years, immunotherapy was based on the use of immune cytokines, such as interleukin-2 and alpha-interferon, or cancer vaccine. All these treatments showed poorly effective patients' responses and/or high toxicity.

A radically new vision of cancer treatments that revolutionized the oncology field is to be attributed to the birth of immune checkpoint inhibitor therapy. In particular, Tasuku Honjo in 1992 and James Allison in 1996, independently, showed that blocking a programmed cell death 1 (PD-1) or Cytotoxic T-Lymphocyte Antigen 4 (CTLA-4) molecule on immune cells, could cure cancer in mice. CTLA-4 is homologous of the costimulatory receptor CD28, both expressed by CD4⁺ and CD8⁺ T cells, which mediate antagonistic functions in T cell activation, and share two ligands, namely CD80 (also known B7-1) and CD86 (also known B7-2), expressed on APCs. CD28 interacts with the CD80 dimer with relatively high affinity and the CD86 monomer with lower affinity, to mediate T cell

activation in conjunction with TCR signals. Conversely, CTLA-4 interacts with both ligands with higher affinity and avidity than CD28, to inhibit T cell activation. Along with CTLA-4, the PD-1/PD-L1 system constitutes another immune checkpoint pathway mainly operating by controlling immune homeostasis. PD-1 mediates an inhibitory signal in T cells after binding to its ligands, PD-L1 and PD-L2, which are expressed on APCs and cancer cells [57]. Physiologically these two inhibitory receptors (IRs) were expressed transiently on activated T cells and when bound by their cognate ligand, PD-L1/ PD-L2 for PD-1 or CD80/86 for CTLA-4 respectively, suppress T-cell receptor (TCR) signal restrain immune function at the end of acute infections to limit immunopathology or autoimmunity. Unfortunately, these mechanisms can also be used by malignant cells to escape killing by cytotoxic T cells. Indeed, was widely demonstrated that TILs were chronically stimulated by persistent tumor antigens within the tumor microenvironment (TME). This process induces a state of T cell dysfunction, known as T cell exhaustion, marked by the upregulation on the cell surface of inhibitory receptors, such as PD-1 and CTLA-4, and other inhibitors receptors [42]. On the other side, malignant cells express on their surface in response to inflammatory signals a high level of IRs ligands, such as PD-L1/ PD-L2 or CD80/86. When CD8⁺ T cells go to meet and kill cancer cells, through the binding of T cell receptor (TCR) with human leukocyte antigen class I (HLA I), the interaction of PD-1 or CTLA-4 with their cognate partner prevents CD8⁺ T cells from destroying the tumor cells. This phenomenon, collectively known as tumor immunosuppression, inhibits the activation of T lymphocytes and enhances the immune tolerance of tumor cells, thereby achieving tumor immune escape (**Figure 1**). The cascade signaling pathways, initiated upon ligation of CTLA-4 or PD-1 engagement were extensively studied. Upon binding CD80 or CD86, CTLA-4 recruits the phosphatases SHP2 and PP2A via the YVKM motif in its cytoplasmic domain. SHP2 recruitment results in attenuation of TCR signaling by dephosphorylating the CD3 ζ chain. PP2A recruitment results in downstream dephosphorylation of AKT, further dampening the T cell activation pathway. PD-1 ligation by PD-L1 or PD-L2 also recruits SHP2 to the ITSM domain, resulting in membrane-proximal decreases in TCR signaling through the dephosphorylation of CD3, ZAP70, and protein kinase C θ [58].

From this new immunological knowledge were born the immune checkpoint inhibitors drugs, a monoclonal antibody that blocks IRs, such as PD-1, CTLA-4, or PD-L1, avoiding 'off' signal triggered by tumor cells. Indeed in 2011, following clinical trials led by Memorial Sloan Kettering physician-scientist Jedd Wolchok, the first ICI treatment for advanced melanoma was approved by the food and drugs administration (FDA) [59]. This drug, also known as 'ipilimumab', acts by binding the CTLA-

4 preventing its interaction with the CD80/86. Three years later, in 2014, the FDA approves two new immune therapies for melanoma, pembrolizumab, and nivolumab, which target the cell-surface inhibitory receptor PD-1. Just one year later, the FDA approves the latter two drugs also for the treatment of non-small cell lung cancer, and renal cell carcinoma. In 2016 the first monoclonal antibody act to block the PD-L1 (Atezolizumab) was approved by FDA for the treatment of non-small cell lung cancer.

The best examples of stable responses are those observed in patients with advanced melanoma. In these patients, it was reported that 3-year overall survival was 34 and 52% for ipilimumab and nivolumab, respectively [60,61]. The 3-year overall survival was 60, 55 and 41% for nivolumab plus ipilimumab, nivolumab alone, or ipilimumab alone [62]. In advanced RCC, the 2-year overall survival of patients treated with nivolumab plus ipilimumab was 28%, and complete responses were 11% [63-65]. In other cancers, responses to immune checkpoint monotherapies were not as impressive as in melanoma. This means that despite considerable advancements in the clinical care of some tumors, epidemiologic data and ongoing clinical trials suggest that most of the patients receiving ICI do not derive benefit or durable responses, and the mechanisms at the basis of this lack of responsiveness are multiple and still not completely known. Indeed, following the success of the first wave of ICI therapy, new targets were discovered such as lymphocyte activation gene 3 (LAG-3), T cell immunoglobulin and mucin domain-containing 3 (TIM-3) protein, Indoleamine-2,3-dioxygenase (IDO), short for V-domain Ig suppressor of T cell activation (VISTA), OX40 (also known as CD134/TNFRSF4), glucocorticoid-induced TNFR-related protein (GITR; also known as CD357/TNFRSF18), inducible T-cell costimulatory (ICOS), 4-1BB (CD137) and T cell immunoglobulin and ITIM domain (TIGIT) [66]. All the above-mentioned targets are currently being investigated to assess their activity as a single agent or in combination with others ICI drugs, radiation, and chemotherapy. The goal of combination immunotherapies is to produce a durable antitumor response in patients who do not respond to monotherapy. Despite that, malignant cells may use multiple mechanisms of immune evasion, which may vary during the therapy in response to the selective pressure of treatment. Indeed, although infiltrating lymphocytes are generally positively prognostic in many tumor types, the barely infiltrating CD8+ T cells are clearly insufficient to eliminate the tumor. Tertiary lymph node-like structures have been described in tumors at the tumor margin and appear to be related to improved immunotherapy and survival in melanoma [35]. These new pieces of knowledge should be used in the future to build tumor-specific therapy combining different ICIs to target different mechanisms of immune evasion. Therefore, the

development of a system to predict response that considers not only PD-L1 status and lymphocyte profile but also a broader picture of the immune milieu, such as Treg, tumor-associated macrophages, and B cells, will be invaluable for directing therapeutic combinations.

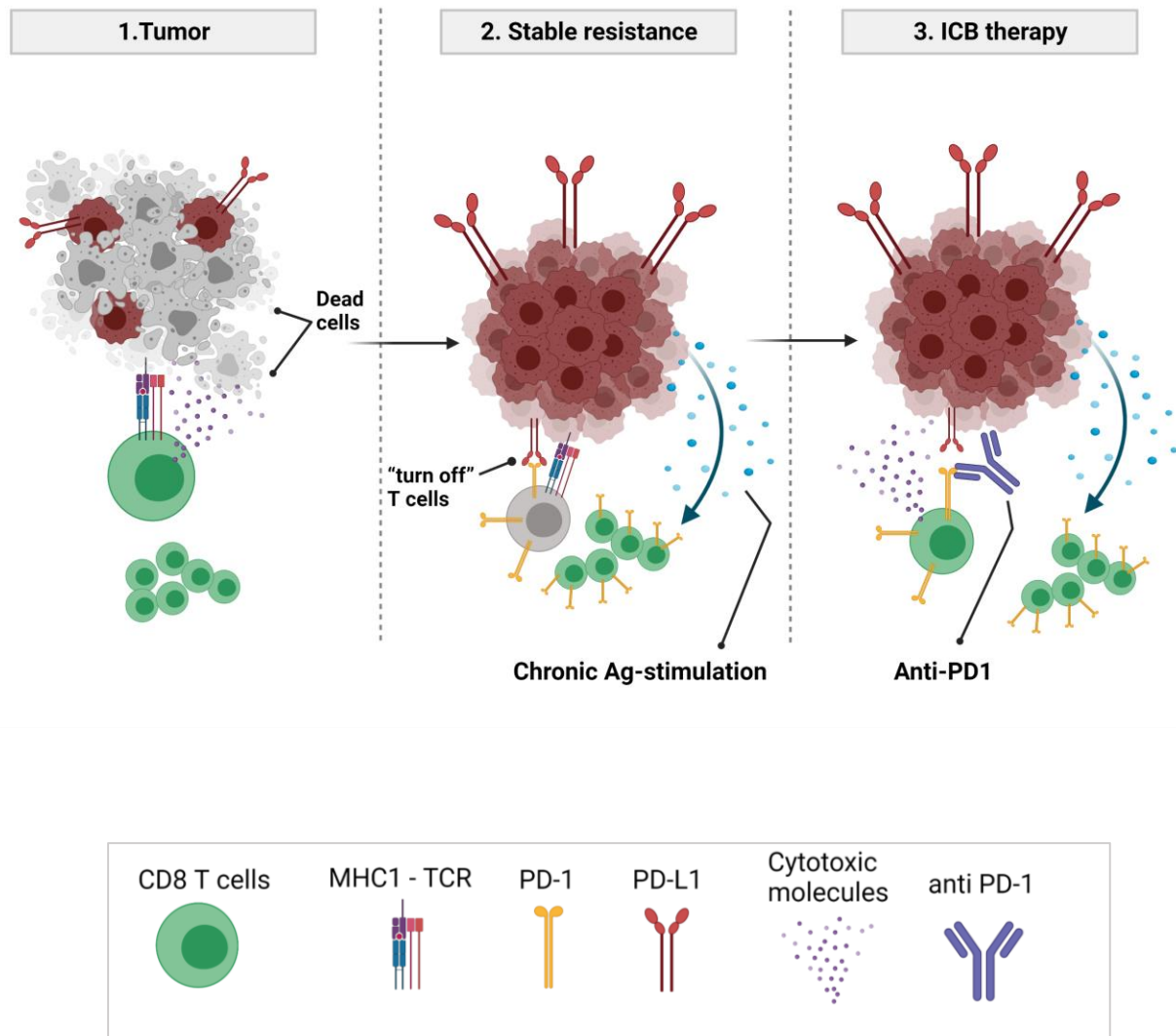


Figure 1. Mechanism of action of immune checkpoint inhibitor, such as anti-PD-1 or anti-CTLA-4.

1) CD8+ T cells recognize and kill tumor cells; 2) chronic Ag-stimulation induce the expression of immune checkpoint molecules, such as PD-1, on the surface of CD8+ T cells, subsequently the interaction among PD-1 and PD-L1 (expressed by tumor cells) turn off CD8+ T cells; 3) anti-PD-1 inhibit PD-1/PD-L1 interaction reinvigorating CD8+ T cells.

1.4 The role of tumor immune microenvironment (TIME) in cancer immunotherapy

Tumors contain different cell populations in constant evolution. This diversity is commonly referred to as tumor heterogeneity and is considered the main driver of resistance to therapy and metastasis [67]. The complete comprehension of this heterogeneity would be critical to optimize existing therapeutic interventions and find new strategies to break down relapses and mortality. The recent development of technologies based on sequencing individual cells has been crucial to address tumor heterogeneity and elucidate how cells are organized into multicellular systems. Single-cell profiles revealed that human tumors include subpopulations of genetically diverse malignant cells. A profusion of different cell types from the surrounding tissues and the immune system, each with a precise role in the pathology, is present within the tumor microenvironment. The immune components of the tumor microenvironment have been recently annotated and finely characterized in different kinds of malignancies, including melanoma, clear cell RCC (ccRCC), NSCLC, breast cancer (BC), hepatocellular carcinoma (HCC), glioblastoma multiforme (GMB), microsatellite instability-stable CRC [67,68]. In most of these tumors, immune cells were mainly T cells, whereas myeloid cells, B cells, and NK cells were found at lower frequencies [53,67-70].

In ccRCC, in-depth immunophenotyping analysis identified T cells and tumor-associated macrophages (TAMs) as the main components of the TME [68]. Concerning T cells, eight CD4⁺ clusters, eleven CD8⁺ clusters, one CD4⁺/CD8⁺ double-positive cluster, and one CD4[−]/CD8[−] double negative cluster were identified [68]. Notably, whereas PD-1 had a broad expression in both CD8⁺ and CD4⁺ T cell population, other inhibitory molecules, like TIM-3, CTLA-4, and 4-1BB (CD137), were expressed only by a few PD-1⁺ subsets, indicating that a pre-exhaustion status is also present in ccRCC [68]. Interestingly, in CD8⁺ and CD4⁺ T cells, PD-1 is co-expressed with CD38, which mediates immunosuppression by activating nitric oxide synthetase, which catalyzes nitric oxide production from arginine. Although CD38 has traditionally been linked to T cell activation, these data suggest that its expression is not restricted to activated cells but can be extended to exhausted T cells, at least in ccRCC and BC [67,70]. In BC, has been observed a higher frequency of PD-1^{hi}CTLA-4⁺CD38⁺ T cells in tumor biopsies if compared to juxta-tumoral tissues, thus confirming that PD-1 and CD38 are both expressed in exhausted cells [70]. Indeed, CD8⁺ T cells expressing a high level of PD-1 also expressed the co-inhibitory molecules TIM-3 and CTLA-4, and the activation markers HLA-DR and CD38, which were not expressed by CD8⁺ T cells expressing intermediate levels of PD-1 [70]. Despite the presence of pre-exhausted TILs within the ccRCC lesions, these cells did not correlate with

increased progression-free survival, whereas patients showing poor progression-free survival had high frequencies of TAMs (in the paper reported as M-11 or M-1) [68].

In melanoma, the scRNA-seq analysis revealed that tumor-infiltrating immune cells were heterogeneous, comprising a high number of T cells, B cells, macrophages, and natural killer (NK) cells. Regarding T lymphocytes, CD8+, CD4+, and regulatory T lymphocytes (Treg) have been identified based on their respective transcriptome signature (such as FOXP3 for Tregs). Moreover, CD8+ and CD4+ T cells were dissected based on the specific gene signature in naïve, cytotoxic, and exhausted T cells. The exhausted signature was elevated within CD8+ T cells along with the overall T cell activation level of these cells and could be linked to the immunotherapeutic response. Indeed, patients who expressed a reduced level of CTLA-4 showed an initial response and subsequently the development of resistance to CTLA-4 blockade with ipilimumab. Furthermore, a considerable correlation was observed among clonal expansion and exhaustion state of CD8+ T cells, such that low-exhaustion T cells were not present in clonally expanded T cells but only in clonally non-expanded T cells [36]. In a recent study, the heterogeneity of tumor-infiltrating immune in melanoma patients treated with anti-PD-1 was defined more in-depth. Interestingly, among CD45+ immune cells were observed eleven different clusters comprising B cells, plasma cells, TAMs, Tregs, dendritic cells, memory/naïve-like, cytotoxic and exhausted CD4+ or CD8+ T cells. When we consider regression or progression of each lesion, B cells, and memory/naïve-like were more frequent in responder lesions, while TAMs, dendritic cells, and exhausted CD8+ T cells were more frequent in non-responder lesions [11]. The highest presence of B cells in responder lesions remark their potential role with improved survival and response to ICI therapy [71]. On the contrary, it was shown that TAMs in melanoma lesions were polarized toward a non-inflammatory phenotype and promoted tumor growth [72], suggesting their involvement in reduced response to ICI therapy. Notably, the refined analysis of CD8+ T cells reveals that responder lesions were enriched by memory/naïve-like expressing TCF7, FOXO1, and GZMB genes, while non-responder lesions were enriched by exhausted CD8+ T cells expressing a high level of TIM-3, PRDM1, TOX, and ENTPD1 (encoding for CD39). Moreover, the chromatin states of exhausted and memory/naïve-like CD8+ T cells reveal that the latter were enriched of TCF1 (encoded by TCF7) binding region in open chromatin peaks as observed for pre-exhausted CD8+ T cells [73]. Together these data confirm that the response to immunotherapy in melanoma is associated with higher pre-exhausted CD8+ T cells phenotype who positively correlates with improved cytotoxic function [11].

In non-small cell lung carcinoma (NSCLC) tumor samples, different subsets of CD8⁺ T cells, conventional CD4⁺ T cells, and Treg cells have been found [74]. Each subset is characterized by a specific gene expression signature, reflecting a specific functional status. Two main clusters were found at high frequency: (i) exhausted CD8⁺ T cells, characterized by increased expression of effector molecules, including GZMA, GNLY, perforin, GZMB, NKG7, and inhibitory receptors, like lymphocyte activating LAG-3, TIGIT, PD-1, and CTLA-4; (ii) suppressive CD4⁺ Treg cells, characterized by increased expression of costimulatory molecules, including CD28 and ICOS, and inhibitory receptors like CTLA-4 and TIGIT [59]. Moreover, two CD8⁺ T cell subsets exhibited a functional state that precedes exhaustion and is called the pre-exhaustion state. These subsets do not express CTLA-4 and express mild levels of TIGIT, PD-1, and the transcription factor TOX, which is a critical driver of exhaustion [49-51]. Whether or not pre-exhausted subsets could be more effectively reinvigorated by ICI than fully exhausted subsets are still not known. Furthermore, the expression level of PD-1 or other inhibitory receptors does not necessarily correlate with exhaustion-dependent dysfunction. It was indeed reported that in lung tumors, proliferating tissue-resident memory T cells express PD-1, T-Cell Immunoglobulin and Mucin Domain-Containing Protein-3 (TIM-3) and are negative for CD127 (the α chain of the IL-7 receptor). These cells can upregulate TCR activation-induced genes, exhibiting a transcriptional signature indicative of effector, survival, and tissue-residency properties, and producing cytokines like IL-2, IFN- γ , and TNF [52]. In early-stage triple-negative BC, among CD8⁺ T cells infiltrating the TME, TRM cells display high levels of genes encoding for cytotoxic molecules, including GZMB and PRF1, high levels of genes encoding for the inhibitory checkpoint, as well as high levels of genes associated with proliferation [75]. This evidence means that the expression of inhibitory receptors is not a unique feature of exhausted T cells as several highly functional effector cells also express those receptors. All these data support the idea that exhausted T cells are heterogeneous, and that T cell exhaustion and T cell dysfunctionality are a gradual, rather than a discrete state [43].

1.5 The role of mucosal associated invariant T cells (MAIT) in cancer immunotherapy

Several populations of T cells, sometimes referred to as 'unconventional' T cells, recognize non-peptide antigens presented by specialized MHC class I-like molecules. Mucosal-associated invariant T cells, described for the first time in 1999, belong to 'unconventional' T cells and are restricted to the non-polymorphic MHC class I-like protein MR1 [76,77]. Functional studies revealed that MR1 may bound vitamin B metabolite, such as 5-(2-oxopropylideneamino)-6-D-ribityl-amino

uracil (5-OP-RU), synthesized by a range of microorganisms [78]. Vitamin B metabolism pathways are highly conserved across species, and therefore antigen presentation in the context of MR1 molecule allows for the recognition of a group of pathogens upon their shared metabolic signature. Furthermore, MR1 binds and presents pathogen antigens without processing; in particular antigen-MR1 interaction results in a transient relocation of MR1 molecules from the endoplasmic reticulum to the surface of cells [78]. Many hematopoietic and non-hematopoietic cells express a low surface level of MR1, despite that nothing was known about how MR1 may regulate MAIT activity [79]. Approximately 85-90% of circulating MAIT cells are constituting by CD8+CD4⁻ in healthy adults. The remaining 10% of total MAIT are double negative CD8⁻CD4⁻, while CD8+CD4⁺ and CD8⁺CD4⁺ represent the remaining 5% of total circulating MAIT cells. Several analyses have shown that MAIT cells express certain transcription factors, such as promyelocytic leukemia zinc finger protein (PLZF) (encoded by ZBTB16), eomesodermin (encoded by EOMES), RAR-related orphan receptor t (ROR-t) (encoded by RORC), and TBET (encoded by TBX21), making them distinct from other conventional CD4⁺ and CD8⁺ T cells [80]. The expression of surface markers almost overlaps with conventional T cells such as the following activation and exhaustion markers: CD69, HLA-DR, CD38, CD25, CD44, CD39, PD-1, or several chemokine receptors such as CCR5, CCR6, CCR9, and CXCR6. Despite that, several surface proteins can be used to uniquely identify MAIT cells, such as CD161, Va7.2 TCR, IL-7R, IL-18R, and MDR1 [80]. Currently, the use of Va7.2 TCR has been replaced by using MR1 tetramers loaded with the bacterial ligand 5-OP-RU. Transcriptionally were characterized by specific expression of SLC4A10, KLRB1, IL18RAP, CD40LG, IL23R, CCR6, DUSP1, ZBTB16, and RORC [53].

In humans, MAIT cells were first described in the intestinal mucosa and liver [77]; however, they can be found in many other organs such as skin, colon, and lung. It was shown that MAIT cells are selected by interacting with MR1-expressing thymocytes but not by interacting with thymic epithelial cells nor thymus-resident immune cells. MAIT cells that egress from the thymus display a naïve phenotype but, in the periphery, rapidly expand acquiring memory characteristics enriching mucosal tissues. Interestingly, the selection is maintained in germ-free bred mice, suggesting the possible existence of endogenous MR1 ligand which is to date still unknown.

Functionally, MAIT cells may be activated using two different mechanisms:

- 1) by TCR-dependent activation using MR1 on an MR1-expressing cell such as APC, endothelial cells, and tumor cells [55];

- 2) by TCR-independent activation using virally infected cells, stressed cells, or other immune cells which release IL-12, IL-18, and IL-15 cytokines [81].

These two activation modes may act in concert or independently, depending upon the stimulus. After activation, MAIT cells release cytokines molecules such as TNF, IL-8, IL-17a, and IFN- γ , as well as cytotoxic effector granules like GRZM-A, GRZM-K, GRZM-B, GRZM-H, GRZM-M, and perforin. When MR1 is engaged, MAIT cells predominantly release TNF, IL-17a, GRZM-B, GRZM-H, and perforin, while under IL-12/IL-18 activation, MAIT cells mostly release IFN- γ , GRZM-A, GRZM-K, GRZM-B, GRZM-H, GRZM-M, and perforin [82]. Cytotoxic characteristics and their high percentage in peripheral blood suggest that MAIT cells may also be involved in killing malignant tumor cells. Indeed, during the last years, an increased interest in studying MAIT cells in many types of cancers such as colorectal cancer, hepatocellular carcinomas, multiple myeloma, and melanoma was observed. Several studies of colorectal cancer patients have found increased frequencies of MAIT cells within malignant tissues compared to healthy adjacent tissue [83,84], accompanied by their decrease in blood and liver metastasis [85]. This observation could indicate differences in MAIT cell infiltration of primary tumor versus metastatic sites, or that perhaps MAIT cells in the metastases are more exhausted and tend as a final stage of exhaustion to be deleted. This is consistent, at least in part with a recent study showing that TCR-induced activation of MAIT cells prompts the expression of both CD39 and PD-1, two markers of exhausted T cells [86].

A recent paper studying esophageal cancer patients also reported decreased MAIT cell frequencies in the blood, with increased frequencies within esophageal tumors accompanied by reduced expression of IFN- γ and TNF [87]. Interestingly, this study evidences the inverse correlation between the frequency of infiltrating MAIT cells and patients' death risk. Other studies reported conflicting evidence regarding the role of MAIT cells in cervical cancer, as some report decreased MAIT cells in the blood and others report increased MAIT cells in the blood of cervical cancer patients compared to healthy individuals [88,89]. Furthermore, it was also observed a positive correlation between CD8+CD4-CD38+ circulating MAIT and myeloid-derived suppressor cells (MDSCs), implying that interaction with other cells may also modulate the role of MAIT cells in malignancies. In other cancers type, such as hepatocellular carcinomas, was observed a reduction of infiltrating MAIT compared to adjacent normal tissue [53,90]. In addition, MAIT cells that infiltrate hepatocellular carcinomas showed functionally impairment upregulating inhibitory molecules like PD-1, CTLA-4, TIM-3 and secreting significantly less IFN- γ , GRZM-M, and perforin, highlighting their

reprogramming to shift away from antitumor immunity toward the direction of tumor promotion [90]. In a recent study, in an *in vivo* melanoma mice model, was observed that MAIT cells promote the initiation and expansion of lung metastases via MR1. Furthermore, they also showed that tumor initiation, growth, and metastases were suppressed in MR1^{-/-} mice, suggesting that MAIT cells display tumor-promoting function by inhibiting T and/or NK cells and that blocking MR1 may represent a new therapeutic strategy for cancer immunotherapy [91].

In another study, it was observed that MR1-restricted T cell clones belonging to the family of MAIT cells (defined as "MR1" T cells). Interestingly, this rare population is able to killing several cancer cell lines expressing low levels of MR1 while remaining inert towards non-cancerous cells [92]. As such, all the findings showing the role of MAIT cells should be interpreted with caution when assessing the applicability to humans, due principally to contradictory evidence reported in the literature. As broadly reported in the literature, MAIT cells express many IRs, such as PD-1, suggesting a potential importance of these cells in responses to immune checkpoint inhibitors. In squamous cell carcinoma of the skin and basal cell carcinoma, following PD-1 blockade tumor-residing MAIT cells upregulated expression of genes that codify for activation, cytotoxicity, and chemokines molecules such as HLA-DRB1, IFNG, GNLY, GZMH, CCL4 and CCL5. Despite being activated, their presence is associated with poor overall survival in squamous cell carcinomas patients [93]. Interestingly, another study has found a significant association between immunotherapeutic responses and commensal microbiome composition in humans, showing that responders to anti-PD-1 therapy have favorable microbiome compositions and that antibiotic consumption is associated with poor response to ICI therapy [94]. Recently, it was observed in stool samples from metastatic melanoma patients that responders to ICI therapy the enrichment of *Bifidobacterium longum*, *Collinsella aerofaciens*, and *Enterococcus faecium*. Additionally, transferring fecal material from responding patients to germ-free mice improved the immune response and correlated with the tumor regression [95]. In light of data demonstrating vitamin-B synthesis by *Bifidobacterium*, these results suggest a potential positive effect of commensal flora to potentiate antitumor T cell responses and therefore also of MAIT cell that responds explicitly to vitamin-B metabolites through MR1 receptor. Despite the increasing numbers of data reported in the literature, there are still many routes to take with MAIT cell research to characterize better their role and pathways involved in tumor response in patients naïve or treated with ICI therapy.

1.6 CD8⁺ T stem cell memory in cancer immunotherapy

Recently discovered T memory stem (TSCM) cells represent a rare subset of memory lymphocytes endowed with the stem cell-like ability to self-renew and the multipotent capacity to reconstitute the entire spectrum of memory and effector T cell subsets [16]. These cells were found within the naive T cells compartment (CD45RO⁺CCR7⁺CD45RA⁺CD62L⁺CD27⁺CD28⁺ and IL-7Rα⁺), but unlike them, TSCM expressed large amounts of CD95, IL-2Rβ (also known as CD122), CXCR3, Ki67, and long telomeres [16,96]. To date, several studies in mice and non-human primates have shown their low differentiated state among memory subsets, placing TSCM at the apex of the hierarchical system of memory T lymphocytes [16]. In particular, TSCM after the exposure to IL-15 retained higher amounts of CD45RA, CD62L, and CCR7 than CM cells. The 60% of those cells preserved at the end of stimulation (12-15 days) their phenotypic identity (CCR7⁺CD62L⁺CD45RA⁺CD45RO⁺CD95⁺) if compared to those of CM that reached only 30%. Moreover, after α-CD3/CD2/CD28 stimulation, only the TSCM can differentiate into CM and EM, reconstructing the entire heterogeneity of the memory T cell subset [16]. In terms of the transcriptome, TSCM expressed high LEF1 and FOXP1, both regulating the quiescence of naive T cells, accompanied by low levels of EOMES, TBX21, PRF1, KLRG1, and GZMA which identify more differentiated memory T cells [97-99]. Interestingly, genetically engineered TSCM, expressing the mesothelin-specific chimeric antigen receptor (CAR), injected in mice model triggered tumor regression if compared with not treated ones [16,100]. A recent study also showed that the response and the long-term CAR T-cell persistence in leukemic patients are strongly related to the percentage of TSCM present pre-infusion within the pool of genetically engineered CD8⁺ T cells [101]. An elegant study using mice models (C57BL/6) showed that the antitumor response against solid malignancies induced by ICI therapy relies not on the reversal of T cell exhaustion programs but on the proliferation of a CD8⁺ stem-like TIL subset. In particular, they described a subset of tumor-reactive TILs bearing hallmarks of exhausted cells and long-term memory cells, including expression of the checkpoint protein PD-1 and the transcription factor TCF1 (encoded by TCF7). These stem-like tumor-reactive TCF1⁺ PD-1⁺ CD8⁺ T cells were also detected within the blood and primary tumor site of melanoma patients, which correlated positively with TCF1⁺PD-1⁺ effector cells present in the tumor and stroma [56]. In agreement with these data, patients that respond to Ipilimumab and/or Nivolumab showed a higher fraction of infiltrating TCF7⁺PD-1⁺ cells among CD8⁺ TILs than non-responders [36]. Moreover, a meta-analysis of the TCGA database indicates that a TCF7/PD-1 signature correlates with improved patient survival to immune checkpoint inhibitor [102]. The presence of infiltrating stem-like CD8⁺ T cells was also

observed in NSCLC. These cells expressed CXCR5⁺ TIM-3[−]PD-1^{lo}TIGIT^{lo} while retaining gene networks responsible for stem-like plasticity and cytotoxicity such as TCF7 [34]. Interestingly, using the glycolytic capacity as negative prognostic markers in NSCLC tumor (the more is the uptake, the worst is the prognosis), the frequencies of CXCR5⁺ CD8⁺ T cells are more prevalent in low glycolytic tumors [34,103]. In a recent study, infiltrating stem-cell like TCF1⁺ CD8⁺ T cells resides preferentially in tumors niches enriched by dendritic cells expressing a high level of major histocompatibility complex II (MHC-II) [104]. Loss of tumor niche negatively impacts the percentage of CD8⁺ TILs infiltrating the tumor, suggesting that stem-cell like TCF1⁺ CD8⁺ T cells are critical for the effector differentiation process. Indeed, patients with progressive disease lack these immune niches, suggesting that niche breakdown may be a critical immune escape mechanism [104]. Together, these data suggest that the progressive differentiation of T lymphocytes leads to a gradual loss of function and therapeutic potential. These imply that poorly differentiated immune cells, such as TSCM may have more application potential in clinical treatment.

TSCM accounts for 2-3% of total circulating CD8⁺ T cells. Despite their rarity, significant heterogeneity within these cell populations is recognized. Recent data displayed the presence of two subsets of human circulating stem-like CD8⁺ memory T cells referred to as functional or dysfunctional lineages. The former lacking the inhibitory receptors PD-1 and TIGIT showed greater activation and cytotoxicity, whereas the second expressing PD-1 and TIGIT were committed to a dysfunctional state showing reduced functional proprieties [105]. Functional TSCM produced cytokines (IFN- γ , IL-2, and TNF) at higher frequencies and higher levels on a per-cell basis than did dysfunctional TSCM in response to stimulation with anti-CD3 plus CD28. Moreover, dysfunctional TSCM displayed also epigenetic alteration remarked by high accessible genes related to T cell dysfunction and terminal differentiation such as TOX, TOX2, TIGIT, PDCD1, NFATC2, EOMES, and MAF [105]. Another group also reported that TSCM expressing surface CD122^{hi} and CXCR3^{lo} displayed greater proliferation, multipotency, and enhanced polyfunctionality with higher frequencies of triple-positive TNF, IL-2, IFN- γ cytokine-producing cells upon exposure to antigen stimulation [106]. The relative scarcity of circulating TSCM cells limits their use in tumor therapy, despite that several protocols have been developed to expand this cell type in vitro for adoptive T cell therapy. Nowadays, one of the most common cytokines used to expand therapeutic T cell products for patients is IL-2. However, high IL-2 levels in the cell culture medium diminish the overall production of early memory T cells by reducing long-term memory T cells and increasing effectors T cells [107]. Instead, the use IL-15 promote the proliferation of TSCM cells by down-regulating the

expression of PD-1 and promoting the ability of T cells to produce IFN- γ , IL-2, TNF, and granzyme B. Indeed, CAR-T cells may expand only under the homeostatic stimulus, mediated by IL-15 and not by IL-7, preserving a less-differentiated TSCM phenotype. Moreover, recent data showed that the combined use of IL-7 and IL-15 can preserve the TSCM phenotype and enhance the effectiveness of CAR-T cells. Current human data displayed that also the IL-21 can promote in vitro generation of CD8⁺ TSCM cell population from naive precursors cells; the proportion of TSCM cells obtained with IL-21 stimulation was ~5-fold higher than that obtained with IL-2 stimulation and 2-fold higher than that obtained with IL-15 stimulation [108]. The signaling mediated by IL-21R on the surface of CD8⁺ T cells involved the phosphorylation of STAT3. The latter migrating to the nucleus induces the expression of TBX21 and SOCS1 but reduces EOMES and GATA3 expression. Indeed, adoptive immunotherapy of IL-21-induced CD8⁺ TSCM cells in a humanized melanoma mouse model produced a superior antitumor response and tumor regression [108]. Recently, the treatment of primary tumors with liquid nitrogen and radiofrequency (also known as cryo-thermal therapy) induced potent antitumor immunity-enhancing the cytolytic function of CD8⁺ T cells and induced differentiation of CD8⁺ T cells into TSCM [109,110]. The potential antitumor effect of TSCM shed light on long-term for in vivo future application in clinical settings. Several data demonstrated that CAR T cells expressing engineered CXCR1 or CXCR2 receptors markedly enhance migration and persistence of T cells in the tumor, which induce in a pre-clinical model a complete regression accompanied by long-term immunological memory against solid tumors [111]. Indeed, in vivo tracking of humans T cells, reveals safety and long-term survival of TSCM cells supporting their use for tumor-targeted adoptive immunotherapy [112]. Their characteristics can be fundamental to overcoming the current therapeutic failure of using CAR T cells as a curative approach for patients with solid and non-solid malignancies. Indeed, recent data showed that CAR T cell expansion during a clinical response is coupled to progressive restriction of gene regulatory programs that control T cell fate potential coupled with extensive exhaustion epigenetic programs that skewed DNA methylation to memory genes such as LEF1 and TCF7 [107].

1.7 Single-Cell Approaches to Immune Profile

During the last decade, the immunology field has been revolutionized by the introduction of new high-throughput technology such as high-parameter flow cytometry, mass cytometry, and single-cell RNA sequencing (scRNA-seq).

In 2001, using flow cytometry, immunologists were able to detect 11 different markers on the surface of millions of peripheral blood mononuclear cells (PBMC) [113], and just three years

later, this number was scaled up to 17 [114]. In 2009 the first mass cytometer CyTOF (cytometry by time of flight) was introduced. It allows simultaneous analysis of more than 40 molecules, including cytokines and transcription factors with minimal compensation [115]. Mass cytometry overcomes the fluorescent labeling limit by utilizing lanthanide isotopes attached to antibodies. This method could theoretically allow the use of 40 to 60 different labels and has been demonstrated for 40 labels. Mass cytometry is fundamentally different from flow cytometry: cells suspension is carried into nebulizer (plasma) through the fluidic system, ionized and associated isotopes are quantified via time-of-flight mass spectrometry. Although this method permits the use of a large number of labels, it currently has a lower throughput capacity (500 cells/s) than flow cytometry (up to 30.000 cells/s) [115]. Mass spectrometry can discriminate isotopes of different atomic weights with high accuracy, enabling more features to be evaluated simultaneously, so the quantity of reporter ions in a particular mass channel represents the marker expression with little signal overlap between parameters [115]. However, these methods reveal the different percentages of cell populations in different pathophysiological onsets, barely identifying different clones [27]. Indeed, flow and mass cytometry perfectly capture the phenotype of cells but fails to snap their biological complexity. The spectrum of phenotypic diversity of immune cells within the TME and in the blood can better be appreciated by analyses at the single-cell level to explore cellular heterogeneity in terms of gene expression and chromatin accessibility. Recently, the implementation of high-parameter cytometry to imaging platforms led to the development of two promising technologies, known as Imaging Mass Cytometry (IMC) and Co-detection by Indexing (CODEX) [116-118]. The former is used to process tissues, employs antibodies tagged with lanthanides, and combines a high-resolution laser ablation system with CyTOF [118]. The latter (CODEX) employs oligonucleotide-conjugated antibodies and the cyclic addition and removal of complementary fluorescently labeled DNA probes used to simultaneously visualize up to 60 markers in situ [119]. Both enable the study of cell-cell interactions and intercellular networks, thus providing information regarding the spatial distribution of cells within a tissue. These two techniques are instrumental in investigating the TME not only in terms of cells composition but also in terms of their spatial localization. Indeed, a recent study on patients who have biliary tract cancer (BTC) and were treated with anti-PD-1 showed that infiltrating PD-1+CD8+ T cells were localized adjacent to tumor cells, and the number of infiltrating CD8+ T cells positively correlate with the clinical response to PD-1 blockade therapy during the early-stage of the disease [120].

Along with flow and mass cytometry, scRNA-seq technology has transformed our view of biology, shed light on the complex cell type landscape, and become invaluable to further understand cellular mechanisms. In particular, scRNA-seq allows the analysis of hundreds of messenger RNAs (mRNAs) in a single experiment, enabling the reconstruction of a high-resolution map of cells according to their molecular signature [121]. The first example of single-cell digital gene expression profiling was published in 2009, and since then, tremendous efforts have been made to improve experimental protocols and bioinformatics pipelines to analyze data [122]. A canonical scRNA-seq protocol involves several steps, including single-cell isolation, cell lysis to obtain RNA, reverse transcription into the first-strand cDNA, second-strand synthesis, cDNA amplification, and sequencing, making it extremely complex if compared to standard flow cytometry protocol [123].

Although different techniques can isolate single cells, the use of microdroplet-based microfluidics is now widely diffused among the majority of commercial scRNA-seq platforms and allows the isolation of individual cells into aqueous droplets in an oil phase. In every droplet, cells are lysed in a hypotonic buffer, and mRNAs are captured by poly-dT primers. When reverse transcription occurs, cDNA molecules originating from cells of the same donor are identified using short DNA barcode tags. Then, second strands are generated, and the small amount of synthesized double-strand cDNA is amplified by conventional polymerase chain reaction (PCR) or in vitro transcription, depending on the technology [123]. Some protocols improved read coverage across transcripts analyzing the full-length mRNA, which significantly enhances detailed analyses of alternative transcript isoforms and identification of single nucleotide polymorphisms (SNPs) with high sensitivity and accuracy [124]. Based on different brands and principles used to isolate single cells, library preparation could be different in some steps, but currently, most protocols use the Illumina platform for sequencing and the Nextera kit for library preparation. Once the libraries are prepared, sequencing is performed [123]. A critical parameter to be considered is the sequencing depth, which affects the sensitivity of scRNA-seq to detecting the complete cell heterogeneity. The depth of single-cell sequencing (i.e., the number of transcripts detected in each cell) diminishes as the number of libraries included in a sequencing run increases, owing to a finite sequencing capacity per run. The number of required cells increases proportionally to the sample complexity. Considering T cells, which include different clones and functionally different subsets, the relationships between RNA profiles and cell subtypes could be challenging to observe. For this reason, starting from a large number of cells may increase the statistical power and the chance to describe different patterns. Moreover, power calculation and group size estimation could be

investigated to predict the degree of heterogeneity that the scRNA-seq experiment will reveal. Unbiased cell-type classification within a mixed population of distinct cell types can be achieved with as few as 10,000 to 50,000 reads per cell [125]. Indeed, if the read depth is kept low and the cell number is higher, the possibility to detect a rare population represented at a frequency of <1% is logically increased. Thus, on the one hand, if the study aims to identify rare cell populations or to look for cells in a mixed population, lower read depth is cheaper and more practical. On the other hand, detailed information on gene expression in every single cell could not be provided. Moreover, the efficiency of the method itself that including capturing of poly-A tails, reverse transcription to cDNA, and amplification, cannot be calculated and can range between 10% and 40% [126], thus meaning that if a gene is expressed at a low level, probably it will not be detected. To overcome this problem, read depth could be increased even if, beyond a certain point, the fraction of PCR duplicates increases. However, the confidence on whether a gene is truly expressed or not depends on how many mRNA molecules are detectable, which depends on many factors, including mRNA stability.

During sequencing, cell libraries are pooled together (multiplexed), and upon completion of sequencing, base call (BCL) files can be used to demultiplex (using sample index present in each DNA adapter) and obtain per-sample FASTQ files to perform secondary analysis. FASTQ files should be aligned to reference genome (hg19 or hg38 for human samples) by using the aligner such as STAR, bwa, BBMap, bowie2 [127-130], and barcode combined with unique molecular identifiers (UMIs) are used to produce counts of captured mRNA molecules for cell (count data). Filtering empty beads or noise from counting data can be performed by analyzing the distribution of the number of UMI counts per cell barcode (knee calling; also known as "knee plot"). The "knee plot" is standard quality control (QC) for the scRNA-seq sequencing step. Cells are ordered on the x-axis according to the number of distinct UMIs observed to make the plot. The y-axis displays the number of distinct UMIs for each barcode. High-quality barcodes are located on the left-hand side of the plot, and thresholding is performed by identifying the "knee" on the curve. On the right-hand side, past the inflection point, are barcodes that have relatively low numbers of reads and are considered to have had a failure in capture and to be too noisy for further analysis. All count data obtained by these processes can be used to perform further QCs, analyze, and interpret single-cell RNA-seq data [123].

Despite its numerous pros and great potential, scRNA-seq suffers from the caveat that mRNA and protein expression do not always directly correlate (coefficient of correlation = 0.4). For this reason, recent technological advances have been made to detect mRNAs and proteins

simultaneously. For example, Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-seq), RNA Expression and Protein Sequencing (REAP-seq), Antibody sequencing (Ab-seq) enable the measurement of proteins and mRNAs in individual cells by using antibodies labeled with DNA barcodes instead of fluorochromes, thus avoiding the limitations dictated by the possible spectral overlap of fluorescent signals [131-133]. Quantifying proteins together with mRNAs allows overcoming the lack of correlation that sometimes exists, between mRNA and protein levels, thus providing a more readout of cellular phenotype at the single-cell level. Indeed, proteins, not mRNAs, are the real targets of drugs, and mRNA abundance cannot necessarily resemble protein abundance [131]. Moreover, in specific settings, the measurement of protein abundance is more sensitive for markers with low levels of mRNA transcripts [131]. Thus, CITE-seq, REAP-seq, and Abseq give an unbiased view of the mRNA and protein profile at the single-cell level, which is necessary to precisely identify the cellular function and provide essential insights into the pathophysiology of multiple disorders. However, sample preparation requires more than two days and cells need to be fixed or lysed, therefore excluding the possibility to perform further analysis [131-133]. Another possibility to investigate both mRNA and proteins is the combination of scRNA-seq and high-parameter flow cytometry. The combinatorial use of scRNA-seq and high-parameter flow cytometry in the same sample would likely have a considerable impact in the field of immunotherapy, as it could be associated with the benefit of combining phenotype and transcriptome data simultaneously. However, the bottleneck of the combinatorial use of scRNA-seq and high-parameter flow cytometry is represented by the inability to link the knowledge gained by the two methods. Interestingly, to date, new computational methods based on deep neural networks may be used to transfer learning from existing single-cell multi-omic resources to new scRNA-seq or flow cytometry datasets and impute surface protein abundances or chromatin accession [134]. To accurately impute surface protein abundance from scRNA-seq data, a deep neural network such as cTP-net employs two steps: (1) denoising the scRNA-seq count matrix and (2) imputation based on the denoised data through a transcriptome-protein mapping. To learn the mapping from a cell's transcriptome to the relative abundance of a given set of surface proteins, cTP-net employs a multiple branch deep neural network (MB-DNN) [134]. MB-DNN contains two internal layers designed to learn and encode features that are shared across proteins 1) first has dimension 1000 nodes, 2) has dimension 128 nodes. The remaining layers are protein-specific, with 64 nodes for each protein that feeds into a one-node output layer giving the imputed value. All layers except the last layer are fully connected

(FC) using rectified linear unit (ReLU) activation function. A multi-layer neural network distorts the input space to make the data classes linearly separable, improving the learning speed [134].

Recently other omics approach was introduced in the field of single-cell analysis, such as the Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq). This method is employed to assess genome-wide chromatin accessibility. ATAC-seq identifies accessible DNA regions by probing accessible chromatin with hyperactive mutant Tn5 transposase that inserts sequencing adapters into open regions of the genome [135]. It has been used to map gene regulation in cell-to-cell variability and rare cell phenotypes, including healthy and malignant immune cells. The use of multi-omics approaches is fundamental to studying multiple characteristics simultaneously. For example, a recent study on cancer cell lines combined pooled Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) screens with scRNA-seq and surface protein measurements to investigate the molecular networks that regulate PD-L1 expression and pathway alteration in response to IFN- γ stimulation [136].

The main advantage of scRNA-seq is to be totally unbiased as it analyses the expression of thousands of genes, while high-parameter flow and mass-cytometry look at 10–40 markers that are pre-selected based on a priori knowledge. Despite that, flow-cytometry remain the most versatile and high-throughput approach for single-cell analysis and immune profile for several reasons, such as relatively low cost with a high dynamic detection range and the simplicity of sample preparation.

1.8 Computational analysis of high-dimensional fluorescent data

Thanks to these methods mentioned above, it is possible to simultaneously analyze from 20 up to 5000 features in a hundred thousand or millions of single cells. To analyze this large amount of data, new computational methods are required. Machine learning (ML) has met this need. ML constitutes a family of computational models that aim to associate a set of input features to a set of discrete output features. Briefly, ML can be divided into supervised or unsupervised learning. Supervised learning refers to fitting a model to data, previously labeled with some 'ground truth' propriety, derived from experimental evidence or assigned by a human. Unsupervised learning refers to methods able to recognize patterns or information without providing the system with information to drive the analysis. To date, state of the art in the field of immunology moved towards the use of unsupervised learning such as FlowSOM, Phenograph, or Full Annotation Using Shaped-constrained Trees (FAUST) clustering algorithms for the identification of a single cell population.

Fluorescence-based flow cytometry can measure up to 28 fluorescence simultaneously [34,137-138], thus representing a reemerging tool for single-cell analysis. This allows for combining fast analysis and high throughput at a relatively low cost with a high dynamic detection range. Flow cytometry suffers from autofluorescence and spreading error (SE) generated by errors in the measurement of photons mainly at red and far-red wavelengths, limiting barcoding and the detection of dim markers. Consequently, developing 28-color fluorescent antibody panels for flow cytometry is laborious and time-consuming. Despite that, this technology is especially interesting for high content immune-monitoring that, owing to a significant increase in immunotherapeutic approaches to the treatment of cancer, has become paramount to defining the correlation among the clinical response and the changes occurring at the level of the immune system. The computational approaches allow for the interrogation of large datasets generated in these types of studies and for the unbiased analysis of the data, possibly leading to the detection of rare cell types [139]. Fluorochromes with overlapping emission spectra cannot be avoided when working with a high-dimensional flow cytometry panel. This effect leads to unwanted fluorescence 'spilling' into a secondary detector [140]. This technical artifact can be reduced by applying the mathematical correction of this spillover, also known as 'compensation'. Compensation is correctly applied when the median fluorescence intensity of the compensated positive population is equal to that of the negative population. Both spillover and SE can be reduced by the accurate design of panels that include the selection of fluorochromes with less overlapping emission spectra or assign the detection of an Ag that is not too bright for dyes emitting at red and far-red wavelengths (e.g., PE-Cy7 and APC-Cy7) [140].

Incorrect compensation is a common source of artifacts in flow cytometry analysis, especially when using computational processes, as it might lead to the appearance of 'new' subpopulations. Instead, uncontrolled SE cause that cells in the spread interval to occupy different regions of the t-SNE or UMAP plot (these methods will be discussed below) even though they are not biologically different. Together with the quality assurance of instrument performance [141], even the differences in background fluorescence (batch effect) can induce technical artifacts in computational analysis. To improve the output of the automated gating strategy, data have to be transformed applying correct algorithms such as biexponential, arcsinh, and HyperLog transformation. Although there is still no consensus on which transformation should be used, the optimal transformation depends on the original data, and a step of optimization is necessary to reduce intra-metacluster variability [142].

Before starting the unsupervised analysis, a pre-processing step must be performed using FlowJo software (or R by using FCS package) in which data are correctly compensated and checked the irregularities during acquisition interrogating all fluorescence parameters versus Time. Then were excluded doublet, dead cells, and fluorochrome aggregates. The latter show as random events (sometimes in the diagonal of a bimodal plot among the fifth and sixth decade) are generally brighter than the fluorescence that is usually observed for known populations. Generally, the pre-processing step is equal across the different unsupervised clustering algorithms used, such as PhenoGraph, FlowSOM, and FAUST [139]. Once these steps are correctly performed, the cells of interest, such as CD4+ or CD8+ T cells, can be selected for the unsupervised step. Below will be explained in detail the theory and the analytical steps used by different clustering algorithms.

A) PhenoGraph

PhenoGraph provides a computationally efficient graph-based method for identifying subpopulations in high-dimensional single-cell data. To build the graph, it finds the k-nearest neighbors (KNN) for each cell (using Euclidean distance), resulting in N sets of k-neighborhoods [143]. The algorithm defines the underlying cellular populations by a graph (or network) in which each node represents an individual cell, and each edge represents the number of shared neighbors. PhenoGraph then partitions the graph into distinct populations based on the Louvain community detection method, where cells with similar phenotypes are clustered together [143]. In this way, each community represents a unique population of cells with similar phenotypic features, and the connections between communities represent the correlation between the various populations. Tuning the iper-parameter 'k' (representing the max number of nearest neighbors of each cell) is helpful in selecting the number of clusters that better fit the biological subpopulations within the data. Lower k-values are best for detecting rare subpopulations, whereas higher k-values mainly detect the most prevalent subpopulations [143]. In the Phenograph pipeline, the data need to be transformed in FlowJo v.10 (or later), optimizing the biexponential transformation by increasing or decreasing the 'width basis' per fluorochrome (channel) to minimize the width of the negative population (to minimize the batch effect) [139]. Then to reduce the processing time (in our experience 4-5 hours for 1.000.000 cells), the downsample step should be

performed for each file to reduce the number of cells of the population of interest (i.e., CD8+ or CD4 T cells) to 10.000-20.000 cells, and then export the downsampled data as a comma-separated value (CSV) file before uploading them into Rstudio [139]. For each k-value used in the analysis, as output PhenoGraph generates a heatmap that will be used to compare the phenotype of the different clusters and to select the optimal k-value. The value plotted to reveal the clusters' identities in the heatmap is the integrated MFI (iMFI). This value for each marker is calculated by multiplying the percentage of positive cells by the non-scaled MFI of the positive cells and scoring it on a scale ranging from 0 (minimal value) to 1 (maximum value) [139].

In general, if there are multiple clusters with the same phenotype, a higher k-value is necessary; yet rare populations might be missed by using a k-value that is too high. Once selected the k-value that best fit the data biologically, define groups among samples (e.g., healthy and diseased), if applicable, and use the frequency of the clusters per sample to generate a dot plot or boxplot graph and perform statistical analysis by using the generalized linear mixed model (GLMM) or Mann-Whitney *U*-test.

B) FlowSOM

FlowSOM uses consensus clustering to organize cells and analyzes cytometry data with Self-Organizing Maps (SOM), a type of artificial neural network [144]. FlowSOM clustering algorithm includes four computational steps: (1) Scaling within each marker; (2) Building up a SOM with nodes representing the overall composition of neighboring cells and assigning cells to the closest node; (3) Visualizing the SOM by building a minimal spanning tree to connect the nodes into a graph; (4) Calculating the meta-clustering of the nodes and automatically suggesting the best number of clusters for each dataset (Consensus Clustering) [144]. In FlowSOM pipeline, the data need to be arcsinh transformed ($\text{channel MFI} = \text{arcsinh}(\text{value}/\text{cofactor})$). This arcsinh transformation makes the distributions more symmetrical and comparable in their range of expression, which is fundamental for clustering in FlowSOM [144]. Another feature of arcsinh transformation is that it can be tweaked by user-controlled settings and behave differently for regions around zero. This user-controlled setting is represented by cofactor (or scale argument) for arcsinh scales. For data values on either side of zero to the magnitude of the cofactor, data are displayed in a linear-like fashion, tracking with

the raw data values. For values beyond the cofactor, data are displayed in a log-like fashion. Generally, all channel cofactors should be tested before starting the analysis [145]. Indeed, once exported the data for the cell population of interest from all samples (as FCS, not as CSV), all files should be uploaded into the Cytobank web application available for this pre-analytical step. It is also possible to do data transformation of original exported FCS files directly in R. In this case, transformation cofactors would have to be assessed manually for each channel of the dataset, and their appearance would need to be compared to the original appearance in FlowJo to control for correct assignment of the cofactors [139]. Of note, cofactors should be those that best fit the data distribution among all files. Then upload the files in the R environment to concatenate and arcsinh transform the data using the pre-calculated cofactor. Check data quality by histogram overlays of all markers among all samples and all groups to exclude samples that could lead to false-positive artifacts (high batch effect). At this point, the data are ready to be analyzed by FlowSOM [144]. The algorithm automatically scales the transformed MFI of each channel to values between 0 and 1 by applying low (e.g., 1%) and high (e.g., 99%) percentiles as a limit. This additional transformation of the arcsinh-transformed data can sometimes give a better visual representation of relative differences in marker expression as the ranges of marker intensities can vary substantially. Of note, all computations (differential testing, hierarchical clustering) are still performed on arcsinh-transformed and not on scaled expressions [144].

Generally, FlowSOM, as a result returns a heatmap showing the meta-clusters, where the number of nodes 'k' (meta-clusters) is set to the expected number of cell types $\pm 15\%$. However, for visualization purposes, it is advantageous to include more nodes than the expected number of clusters and manually merge meta-clusters. The merging of 30 or more clusters can be arduous. To simplify this task and reduce excessive clustering, can be used the ConsensusClusterPlus package ('elbow criterion'). The latter criterion analyzes several values for 'k' that are tried out, and for each meta-clusters, the variance in the clusters is calculated. The goal of the elbow criterion is to detect the point where the variance stops decreasing sharply and only decreases slowly from that point on [144]. If the number of clusters is very low, the variance will be high. The variance will decrease enormously if the number of clusters is increased. If the number of clusters is correct, the variance will be relatively low.

C) FAUST

FAUST represents a recently introduced method for single-cell cytometry analysis, which performs unbiased cell population discovery and annotation. The fundamental assumption underpinning FAUST is that biologically relevant cell clusters into homogeneous modal groups along the markers measured in a panel [146]. Generally, clustering methods such as PhenoGraph and FlowSOM often require investigators to pre-specify the number of clusters expected in a sample. However, when these methods are applied to the analysis of a new panel to describe the immune landscape in cancer or other diseases, such information is often unavailable. The proposed solution is to partition a dataset into a considerable number of clusters in order to capture its main structure, but this over partitioning can reflect a method's assumptions rather than biological distribution. Another challenge for many methods is represented by the output, as the assigned phenotype of clusters is arbitrary based on the observed scaled MFI on the heatmap. Instead, FAUST output is interpretable in that its final product is an annotated matrix of counts for an experiment, with rows corresponding to samples and columns to annotated selected clusters [146]. The algorithm consists of two main phases. Both phases are performed on a sample-by-sample basis. The first phase consists of deriving a standardized set of thresholds for informative markers in the cytometry experiment (markers boundary). The second phase of FAUST consists of discovering and down-selecting phenotypes in terms of the informative markers (selected based on depth score, see below). This step is done by growing numerous trees whose leaves correspond to candidate phenotypes and then selecting a subset of leaves across trees that jointly partition the sample [146].

Unlike the previous methods, with FAUST we import directly into R environment the Flowjo gatingset (containing compensated data and gating strategy nodes) and not the CSV or FCS files. Then we need to set the population of interest as 'starting node', such as CD8+ T cells, and specify which markers to use in the analysis. The critical parameter to set is the 'experimental unit', which determines whether any groups of samples should be concatenated prior to discovering and annotating phenotypes [146]. By default, the experimental unit is set to 'individual sample'. However, in longitudinal experiments where samples from a single subject are exposed to multiple stimulations at multiple timepoints, concatenating samples from a single subject across stimulation conditions

may be desirable (modifying the 'imputation hierarchy' parameter). Another fundamental parameter is the annotation boundaries (markers boundary). The boundary is the regions where marker expression on cells can be clearly divided into positive or negative. FAUST can automatically estimate the boundary or use a supervised estimated list of prior (biological) knowledge of marker distribution. During the imputation step of annotation boundaries, FAUST standardizes the number of annotation boundaries for each marker by majority vote. The most frequently occurring number of annotation boundaries across experimental units is chosen as the 'standard number' [146]. When the annotation boundaries do not match the standard boundary in an experimental unit, FAUST imputes or delete the non-conforming boundaries. This step can be bypassed via the use of a supervised estimated list of boundaries that incorporate prior biological information. The boundary selection needs to be carefully controlled to avoid over or under-representation of each subpopulation.

Once parameters have been set, FAUST grows an annotation forest for each experimental unit. FAUST defines the annotation forest as the exhaustive collection of all "reasonable" gating strategies in an experimental unit by using previously selected markers. Once grown, FAUST computes a 'depth score' for each marker that quantifies how consistently a marker separates into subpopulations in the annotation forest [146]. The depth score is used to select a subset of markers in the panel that are consistently high-scoring across experimental units, then estimates a standard number of sample-specific annotation thresholds for each selected marker. FAUST ultimately uses this annotation system to define its clusters and match them across experimental units [146]. Note that the high-scoring markers across experimental units are chosen manually by observing the scoreLine plot selecting both 'depth score threshold' and 'quantile threshold'. After marker selection, FAUST grows multiple discovery forests for each experimental unit. Unlike the trees in the annotation forest (where the depth score is calculated), these trees are not depth constrained. Consequently, each tree in the forest is a proposed clustering of the experimental unit, and each leaf within a tree corresponds to a unique phenotype. The per-sample computational time is in line with FlowSOM, reaching 20 to 30 minutes to compute the phenotype of 10.000.000 cells. Unlike PhenoGraph and FlowSOM, the annotated count matrix contains the complete phenotype of obtained clusters. Despite the cluster annotation's clarity, the output

generally contains hundreds of clusters that need to be carefully managed to exclude tiny clusters or clusters present only in less than fifty percent of patients.

Given that cytometry data are multi-dimensional, meaning each individual cell is quantified by multiple parameters (i.e., the number of fluorescences used) in order to visualize these parameters in a low-dimensional manner simultaneously (i.e., 2 or 3 dimensions), several classical methods have been applied that reduce high dimensional data into low dimensional space. They could be divided into linear and non-linear transformations. The first one is used to separate the data according to the linear combination of the original variables, while the second one is applied when the high dimensional data contains non-linear relations among them. An example of a linear method commonly used in biology includes principal component analysis (PCA), employed in immunology for the first time in 2008 to allow subjects' classification using T cell subpopulations [147].

Nowadays, for the analysis of complex datasets, are preferred non-linear transformations such as t-distributed stochastic neighbor embedding (t-SNE) or Uniform Manifold Approximation and Projection (UMAP) [148,149]. Both t-SNE and UMAP have been designed to preserve the "local structure predominantly", that is, to group the neighboring points favoring a very informative view of the heterogeneity in the data. The critical idea of t-SNE is to use random walks on the k-nearest neighbors' graph of the data to find a smaller set of landmarks, which are points that can serve as representatives of the surrounding points. This idea is fundamental to reduce the size of the data set by an order of magnitude, improving the computational speed [148]. Instead, UMAP is based mainly on manifold theory and topological data analysis. At a high level, UMAP uses local manifold approximations and patches together their local fuzzy simplicial set representations to construct a topological representation of the high dimensional data. UMAP then optimizes the layout of the data representation in the low dimensional space to minimize the cross-entropy between the two topological representations [149]. Indeed, the quality of data visualization, i.e., distances between data points within clusters (local structure), is comparable between t-SNE and UMAP provided that hyperparameters are adequately tuned. Conversely, the distance between clusters (global structure) is preserved correctly only by UMAP. This means that in the UMAP plot, the Euclidean distances among the clusters correlate with their similarities [150]. We can color (embedding) the cells by clusters type obtained using the appropriate algorithm for all the previously described

methods. The embedding step is also beneficial to select the clusters that need to be manually merged.

1.9 Computational analysis of high-dimensional scRNA-seq data

Computational analysis of scRNA-seq data includes four computational steps: 1) Quality control (QC), 2) normalization, 3) scaling and regress-out the source of technical variation, 4) integration (optional), 5) clustering 6) dimensional reduction. Before analyzing the single-cell expression data, all count matrix files obtained after FASTQ alignment need to pass through a QC step. During this crucial phase, we should exclude cells with a low (empty droplet) or high (doublet) number of genes or UMIs [123].

Damaged cells should also be excluded using a threshold on the percentages of mitochondrial genes and rRNA genes, as this could represent a potential factor for biased clustering. It is noteworthy that a unique threshold for datasets containing heterogeneous cell populations does not exist (different cell states could be misinterpreted as low-quality outlier cells). Generally, considering and reconsidering QC decisions multiple times is the best solution, starting with more permissive thresholds and changing them to investigate the effect of more stringent ones. Usually, the percentage of mitochondrial genes for human immune cells may not exceed 10% of the total genes count [123].

Normalization is a necessary step in single-cell data analysis, as it allows to transform non-normally distributed data into normally like distributed data which can be used for downstream analysis. The choice of method depends on available information and the type of data. Some methods may express certain advantages over others. For data obtained using droplet-based 3' enrichment methods, log-normalized approach (LogNormalize), which is part of one of the most popular tools (Seurat), and Scran approach are widely used to normalize the data. LogNormalize method [151], by default, divides gene counts for each cell(y) by the total genes expression(x) and then multiplies this by the scale factor (as default 10.000) adding '+1' to account for zero counts, and natural log-transforms the result " $\log\{(\frac{y}{x} * 10.000) + 1\}$ ". Instead, Scran method uses a deconvolution approach that partitions cells into pools, normalizes across cells in each pool, then uses the pool-based size factor to define individual cell factors [152]. By performing a uniform scaling per cell, these methods assume that the underlying RNA content is constant for all cells in

the dataset and that a single scaling factor can be applied for all genes. ScRNA-seq data exhibits significant cell-to-cell variation due to technical factors, including the number of molecules detected in each cell, which can confound biological heterogeneity with technical effects. To address this SCTransform normalization approach has been introduced. It uses Pearson's residuals from "regularized negative binomial regression", where cellular sequencing depth is utilized as a covariate in a generalized linear model, successfully removing the influence of technical characteristics from downstream analyses while preserving biological heterogeneity [153].

Normalization accounts only for the effect of count sampling, while the other sources of variation (such as biological or technical covariates) may still remain in analyzed data. As shown in the Seurat webpages, first, we should shift the mean expression of each gene across cells to zero and then scale the expression of each gene, so that the variance across cells is one (this allows us to equalize gene expression so that highly expressed genes do not dominate). To improve the performance of clustering or trajectory inference algorithms, during this step, we can also regress out the source of biological variation, such as cell-cycle effects on the gene expression, or the percentage of the mitochondrial gene, or other technical variations (number of UMI or count depth).

In many experiments, cells are divided into distinct groups that derive from different experimental conditions, different sequencing lanes, or different times of preparation. This could introduce a source of variation which is distinct from biological differences. Thus, to avoid this issue, integration methods present in Seurat (including, such as Canonical Correlation Analysis, Mutual Nearest Neighbors, or Harmony) can be used [154-156]. These methods identify shared subpopulations across datasets but fail when only a subset of cell types is shared across datasets, or significant technical variation masks shared biological signals. To overcome this limitation, a recent integration method has been introduced. It identifies cross-dataset pairs of cells in a matched biological state ('anchors'), enabling the integration of single-cell measurements across scRNA-seq technologies and different modalities [157]. Each anchor pair is scored using the shared nearest neighbor (SNN) approach, which measures the consistency of anchors across the neighborhood structure of each dataset. High-scoring correspondences, therefore, represent cases where many similar cells in one dataset are predicted to correspond to the same group of similar cells in a second dataset, reflecting increased robustness in the association between the anchor cells. Anchors and their scores are then used to compute "correction" vectors that can be used to integrate the transcriptome of two or more different datasets [157].

At this point, we need to identify the subpopulation through the clustering process. In the first step are selected the most variable genes (showing high cell-to-cell variations) and used to perform the linear dimensional reduction of the data by the Principal Component Analysis (PCA) [158] or by diffusion maps [159]. The PCA finds the linear combination of a set of variables having the maximum variance and produces several Principal Components (PC) that summarize our data structure. To reduce the following time-consuming step, such as clustering and non-linear dimensional reduction, select the number of PC that best fit diversity in the data. This might be done using "Elbow plot," a ranking of principal components based on the percentage of variance explained by each one, or "jackstraw" a visualization tool for comparing the distribution of p-values for each PC with a uniform distribution. The next step is clustering, which allows to organize cells into clusters and decipher the identity of clustered cells. The most important clustering algorithms are graph-based clustering approaches, such as SNN-Cliq [160] and PhenoGraph [143]. A similar approach is recently implemented in Seurat and Scanpy based on the Louvian algorithm [161] that has become the standard approach to cluster single-cell datasets. In-depth to identify each subpopulation first is computed k-nearest neighbors for a given dataset (using the most variable PC computed before). Identify clusters of cells by a shared nearest neighbor (SNN) modularity optimization-based clustering algorithm. A key parameter in this step is the 'resolution' that sets the 'granularity' of the downstream clustering, with increased values leading to a greater number of clusters (typically, a good result is obtained between 0.4-1.2).

Also, for scRNA-seq, tSNE and UMAP are the most used tools to perform the non-linear dimensional reduction. As for flow and mass cytometry data (see above), the goal of these algorithms is to learn the underlying manifold of the data in order to place similar cells together in low-dimensional space. Cells within the graph-based clusters determined above should co-localize on these dimensions' reduction plots. Furthermore, considering that cells are typically unsynchronized in single-cell assays, it is possible to reconstruct dynamic cellular trajectories that identify differentiation or cell cycle state. Many algorithms can be used, such as Monocle [162], Scorpis [163], or Wanderlust [164], that incorporate unsupervised data-driven approaches and allow understanding of which genes are involved during these processes. Briefly, these algorithms introduced the strategy of ordering single cells in pseudotemporal analysis (also known pseudotime), placing them along with a trajectory corresponding to a biological process [162]. These approaches allow the identification of genes that exhibit differential expression over the course of the biological process and cluster them based on their expression dynamics. Once identified and

annotated each cluster, differential analysis of cell population abundance can be performed. Moreover, if data contain different conditions or groups, a comparative analysis may be performed to identify differences in genes expression within each cluster across the condition.

2. AIM OF THE STUDY 1 (melanoma)

CD8+ T cells can drive adaptive immune responses against several types of human malignancies, in particular those with higher mutational burden and neoantigen load [165]. These cells are activated by tumor antigens, undergo expansion, and localize and kill infected or cancer cells. However, prolonged exposure to cognate antigens often contracts the effector capacity of T cells and attenuates their therapeutic potential. This process, collectively known as T-cell exhaustion, is characterized by limited proliferation, cytokine production and effector capacity, metabolic rearrangement, increased inhibitory receptors expression, and genome-wide accumulation of epigenetic modifications at effector and memory-related gene loci [42]. Among inhibitory receptors, programmed death-1 (PD-1) has been extensively studied and is now targeted by therapies with monoclonal antibodies that are capable to reinvigorate T cells in several cancer settings. However, immune checkpoint inhibitors (ICI) mediate tumor regression only in a subset of patients, and the mechanisms at the basis of therapeutic resistance are poorly known [166]. A number of studies have initially focused on the mutational load of the tumor as well as on the quality of the cells infiltrating the tumor microenvironment and revealed that increased mutational burden and the presence of CD8+ T cells with stem-like qualities [157,167], among others, can predict the response to ICI [11,24,168-170]. However, tumoral tissue may not always be accessible, thereby making the quest of circulating biomarkers an absolute need. In this regard, recent studies have shown that responding patients have larger clones (those occupying > 0.5% of repertoire) post-treatment than non-responding patients or controls, and this correlates with effector memory T-cell percentage [17], suggesting that peripheral T-cell expansion could predict tumor infiltration and clinical response [171].

Over the last decade, a pressing need to deeply interrogate immune cells either in the tumor microenvironment and/or in blood has led investigators to integrate data obtained from traditional approaches with those obtained with new, more advanced, single-cell technologies, capable to define characteristics of immune cells at an unprecedented degree of resolution [172].

For these many reasons, our project is focused on unraveling the key alterations that occur in peripheral CD8+ T lymphocytes of metastatic melanoma patients during longitudinal treatment with anti-PD1, identifying a possible predictive marker of successful ICI therapy.

3. AIM OF THE STUDY 2 (mRCC)

Renal-cell carcinoma (RCC) accounts for about 5% and 3% of all malignancies in men and women, respectively, and at diagnosis, almost 15% of these tumors show metastasis. Metastatic renal-cell carcinoma (mRCC) has a poor prognosis, although a progressive improvement in median overall survival (OS) has been observed after the introduction of targeted therapy. Since 2015, the management of mRCC was revolutionized by the introduction of immune checkpoints inhibitors (ICI), i.e., anti-cytotoxic T-lymphocyte-associated protein 4 (CTLA-4, or CD152), anti-programmed cell death-1 (PD-1, or CD279), and anti-programmed death-ligand 1 (PD-L1, or CD274) monoclonal antibodies. During the past few years, a significant increase in life expectancy, together with lasting responses, allowed the approval of different combinations of drugs for treating mRCC patients, in both first and second-line settings, i.e., 1) nivolumab (anti-PD-1), administered either alone or in combination with ipilimumab (anti-CTLA-4), or with the multikinase inhibitor cabozantinib; 2) or pembrolizumab (anti-PD-1), given in combination with axitinib or lenvatinib; and 3) avelumab (anti-PD-L1), used in combination with axitinib [62,63,173,174]. ICI has dramatically modified several patients' prognoses with mRCC, but unfortunately, only a subset of patients benefits from these treatments. Basic and clinical research focuses on the immunological effects of checkpoint inhibitors to identify predictors or early indicators of response. Monitoring peripheral immune cells is very attractive, given the ease of blood sample collection and the possibility of following different parameters over time. Therefore, we used high-parameter flow cytometry to address the impact of PD-1 inhibition on circulating CD8⁺ T cell dynamics in patients with mRCC.

4. METHODS 1 (Melanoma)

4.1 Patients' selection

Patients were selected at the 'Centro Oncologico Modenese' of the Azienda Ospedaliero Universitaria di Modena and Reggio Emilia. Evaluation of entry criteria and subsequent enrollment in the study occurred after the clinical visits. All patients at the baseline were treatment naïve. For the study we enrolled 28 patients with metastatic melanoma (25 out of 28 were BRAF wild type) treated with standard-of-care nivolumab or pembrolizumab. As reported by American Cancer Society, the immune checkpoint inhibitors are typically the first drugs tried 'standard-of-care', especially in people whose cancer cells do not have BRAF gene changes. According to the RECIST 1.1 criteria, responders (n = 17) were defined as patients with complete response (CR), partial response (PR), stable disease (SD), or mixed response (MR) of >6 months with no progression, and non-responders (n = 11) as patients with progressive disease (PD). Specifically, among responders, 41.2% had CR, 35.3% had a PR, 17.6% had SD, and 5.9% (which corresponds to one patient) had a MR. The clinicopathologic characteristics of patients are reported in **Table 1**. The mean age of the entire cohort was 71 ± 12 years and plasma lactate dehydrogenase (LDH) level was 418.7 ± 134.7 .

Variable	Non-responder (NR) (n=11)	Responder (R) (n=17)	Total (n=28)
Demographic characteristics			
Age, mean yr $\pm$ (sd)	71.0 $\pm$ 12.3	70.0 $\pm$ 12.4	71.0 $\pm$ 12.1
Male Sex, %	45.4	76.4	64.3
Clinical characteristics			
M stage (%)			
M1a	27.3	23.5	25.0
M1b	27.3	35.3	32.1
M1c	18.2	29.4	25.0
M1d	27.3	11.8	17.8
LDH level (U/L)	423.9 $\pm$ 150.0	401.2 $\pm$ 124.1	418.7 $\pm$ 134.7
Previous therapy (%)	0	0	0
Tumor burden (cm, range)	10.5 (2.1–43.6)	10.5 (1.5–37.8)	10.4 (1.5–43.6)

Table 1. Demographic and clinical characteristics of metastatic melanoma patients.

4.2 Samples collection

All patients provided written informed consent for sample and clinical data collection according to the Azienda Ospedaliero Universitaria di Modena and Reggio Emilia ethical approval protocol #Prot AOU 0005400/18. For each patient, at the different time-point, was collected up to 30mL of peripheral blood into ethylenediaminetetraacetic acid (EDTA) tubes. Briefly the samples were collected before starting ICI therapy (hereafter indicated as T0), 14 days after the first or the second cycle of ICI therapy (hereafter indicated as T1 and T2, respectively). PBMCs were isolated by Ficoll density centrifugation, cryopreserved in FBS supplemented with 10% DMSO and stored in liquid nitrogen until use [175].

4.3 Polychromatic flow cytometry

- a) **T cells landscape.** Multicolor flowcytometry panel was designed to fine characterize the T cells landscape using conjugated antibodies CD3, CD4, CD8, CD45RA, CD197, CD28, CD27, CD127, CD95, CD98, CD71, CD25, HLA-DR, CD38, CD39, CXCR6, CCR4, Ki67, TBET, granulysin, PD-1, BTLA, CD244, and ICOS. According to the therapy the panel was optimized to identify the expression of PD-1 on T cells isolated from patients treated with anti-PD-1 (either nivolumab or pembrolizumab) using an anti IgG4 to recognize the anti-PD-1 bound to PD-1 [19]. All antibodies were previously titrated to define the optimal concentration, the relative information was reported in the **supplementary Table 1**. Briefly the high-dimensional flow cytometry was performed as reported below. Cryopreserved samples were thawed in R10 medium, i.e., RPMI supplemented with 10% fetal bovine serum (FBS), 100 U/mL penicillin, 100 µg/mL streptavidin, 2mM L-glutamine, 20mM HEPES (ThermoFisher, Eugene, OR) and 20 µg/mL DNase I from bovine pancreas (Sigma-Aldrich, St. Louis, MO). After washing with phosphate buffer saline (PBS), cells were stained immediately with the Zombie Aqua Fixable Viability kit (BioLegend, San Diego, CA) for 15 min at room temperature. Then, cells were washed and stained with the combination of monoclonal antibodies (mAbs) listed before and in supplementary the **supplementary Table 1**. Chemokine receptors were stained for 20 min at 37 °C, for 20 min at room temperature. Intracellular detection of Ki67, granulysin and TBET was performed following fixation of cells with the FoxP3 transcription factor staining

buffer set (eBioscience/ThermoFisher) according to manufacturer's instructions and by incubating with specific mAbs for 30 min at 4 °C. Samples were acquired on a FACS Symphony A5 flow cytometer (BD Biosciences) equipped with five lasers (UV, 350 nm; violet, 405 nm; blue, 488; yellow/green, 561 nm; red, 640 nm) and capable to detect 30 parameters. Flow cytometry data were compensated in FlowJo by using single stained controls (BD Compbeads incubated with fluorochrome-conjugated antibodies) [115].

- b) **Deep phenotyping of MAIT cells.** MAIT cells were identified from PBMC using optimized 18-parameter/16-color flow cytometry panel, including CD3, CD8, TCR V α 7.2, CD161, CD45RO, CD197, CD28, CD27, CD127, CD95, CD25, HLA-DR, CD38, CXCR4, Ki67, granulysin and CD69. Cryopreserved samples were thawed as explained in the paragraph a) and stained immediately with PromoFluor-840, viability probe (PromoCell - PromoKine) for 20 min at room temperature. Then, cells were washed and stained with the combination of monoclonal antibodies (mAbs) listed in **supplementary Table 2**. Chemokine receptors were stained for 20 min at 37 °C, whereas all the other markers were stained for 20 min at room temperature. Intracellular detection of Ki67 and granulysin was performed following fixation of cells with the FoxP3/ transcription factor staining buffer set (eBioscience, ThermoFisher) according to manufacturer's instructions and by incubating with specific mAbs for 30 min at 4 °C. Samples were acquired on a Cytoflex LX flow cytometer (Beckman Coulter, Hialeah, FL) equipped with six lasers (UV, 355 nm; violet, 405 nm; blue, 488; yellow/green, 561 nm; red, 638 nm; IR, 808 nm) and capable to detect 21 parameters. Flow cytometry data were compensated in FlowJo by using single stained controls, as above.
- c) **In vitro cytokine and cytotoxic activity by MAIT cells.** Simultaneously with the step b), PBMC were rested for 4 h at 37 °C and then in vitro stimulated with anti-CD3/CD28 (1 μ g/mL) (Miltenyi, Bergisch Gladbach, Germany) and suboptimal concentration of IL-12 (2 ng/mL) (Miltenyi) and IL-18 (50 ng/mL) (R&D System, Minneapolis, MN) and a combination of those [176]. A 11 parameter/10-color flow cytometer panel was optimized to identify MAIT cells producing Granzyme (GRZM)-A, GRZM-B, TNF and IFN- γ that were detected after 16 h of incubation (see **supplementary Table 3**). For the quantification of intracellular cytokines, cells were fixed with BD Cytofix/Cytoperm Fixation/ Permeabilization Solution kit (BD Biosciences) according to the manufacturer's instructions. Samples were acquired on an Attune NxT acoustic flow cytometer (ThermoFisher) equipped with four lasers (violet, 405

nm; blue, 488; yellow/green, 561 nm; red, 640 nm) and capable to detect 14 parameters. Flow cytometry data were compensated in FlowJo by using single stained controls as above.

4.4 Unsupervised flow cytometry data analysis

Flow Cytometry Standard (FCS) 3.0 files were analyzed using FlowJo software version 9.6. Aggregates and dead cells were removed from the analyses and identify CD3+CD8+T cells were gated. 10,000 CD8+T cells per sample were exported and biexponentially transformed in FlowJo version 10. Further analyses were performed by a custom-made script that makes use of Bioconductor libraries and R statistical packages. Data were analyzed using the Phenograph algorithm coded in the Cytokit package (version 1.6.5) in R (version 3.3.3). Parameter K was set at 60. Phenograph clusters were visualized using tSNE. Clusters representing <0.5% were not analyzed in subsequent analysis. New FCS files (one for each cluster), originated from Phenograph analyses, were imported and analyzed in FlowJo to determine the frequency of positive cells for each marker and the corresponding median fluorescence intensity (MFI). These values were multiplied to derive the integrated MFI (iMFI, rescaled to values from 0 to 100). ggplot2 R package was used to generate the heatmap, showing the iMFI of each marker per cluster.

4.5 Cell Sorting and single-cell RNA-sequencing (scRNA-seq) library preparation

Cryopreserved samples were thawed and washed two time in R10 medium supplemented with 20 µg/mL DNase I from bovine pancreas (Sigma-Aldrich). Then PBMC were washed with phosphate buffer saline (PBS) to remove the medium containing DNase I and stained with the Red Live Dead Fixable Viability kit (Thermo- Fisher) for 15 min at room temperature. PBMC were washed with PBS and stained with mAb anti-CD3-PE and anti-CD8-FITC. Viable CD3+CD8+T cells were sorted by using eS3 sorter (Bio-Rad Laboratories, Hercules, CA) equipped with two lasers (blue, 488; yellow/green, 561 nm; all tuned at 100 mW). Cell sorting was performed with 96-99% purity. Sorted CD3+CD8+T cells were immediately loaded on ddSEQ single-cell isolator (Bio-Rad Laboratories) to isolate and barcode single cells. scRNA-seq libraries were prepared by using the Illumina Bio-Rad SureCell WTA 3' Library Prep Kit (Illumina, San Diego, CA, manufactured for Bio-Rad) following manufacturer's instructions [123]. Briefly, after barcoding, RNA was reverse transcribed and cleaned up. Then, second strand cDNA was synthesized and tagmented. Tagmented DNA was amplified, and final

indexed libraries were quantified by using the high sensitivity DNA kit (Agilent, Santa Clara, CA) on a bioanalyzer (Agilent). Sequenced libraries were loaded on an Illumina NextSeq 550.

4.6 scRNA-seq analysis

A total of 74,404 single cells were obtained from 20 patients (9 NR and 11 R at T0, T1, T2 therapy cycles) after SureCell RNA Single-Cell (v 1.1.0) pre-processing and UMI quantification. Downstream analysis was performed in R using Seurat v3.0 [157]. A Seurat object containing all 74,404 was created, then the dataset was split in three subsets T0, T1, T2. For each subset, unwanted cells were filtered out. Specifically, cells that express >10% of mitochondrial genes, cells having <130 or >1500 detected genes, cells having >2500 UMI count and cells in which ribosomal protein-coding genes represented >65% of gene content, were removed. Next, feature expression-measured were normalized by dividing them by the total expression in each cell and multiplying by a factor of 10,000 (Log-Normalize). To promote the identification of common cell types and enable comparative analyses, the three datasets were integrated using IntegrateData function of Seurat package, yielding an expression matrix of 56,142 cells by 17,745 genes. Principal components were selected using the jackstraw and Elbow methods. The dimensional reduction was performed using Uniform Manifold Approximation and Projection (UMAP) on the previously selected principal components. Unsupervised clustering was performed by finding the K-nearest neighbors (KNN) and then, to group the cells, a modularity optimization-based algorithm was applied. A cluster of 231 cells featuring genes related to the myeloid lineage (expressing LYZ) and 4210 Natural Killer cells (expressing TYROBP) was excluded from the analysis for a final matrix of 51,701 cells. The best resolution was selected using clustree package. Differentially expressed genes were identified using the FindAllMarkers function, and the top 15 genes for each cluster were visualized in the heatmap. Differential expression analysis was performed between each cluster and all other cells using a Wilcoxon rank-sum test. Genes were selected to be significant as $\log_{2}FC > 0.3$ and adjusted p-value < 0.05 . Cells from each single cluster were selected and re-clustered to identify the presence of subpopulation. Comparative analyses across conditions inside of each cluster was performed using FindMarkers, genes were considered as significant with $\log_{2}FC > 0.3$ and adjusted p-value < 0.05 . Furthermore, 4000 cells were randomly selected from all 51,701 cells to perform trajectory and pseudotime analysis using Monocle v2.

4.7 Surface protein imputation

A deep neural network (cTP-net) was used to impute the surface protein expression of our CD8+ T cells dataset. The neural network was trained on PBMC processed using multiomics approach (CITE-seq and REAP-seq) and use the transfer learning process from scRNA-seq transcriptome to protein prediction. cTP-net impute the following list of surface proteins: CD3, CD4, CD45RA, CD45RO, CD16, CD14, CD11c, CD19, CD8, CD34, CD56, CD57, CD2, CD11a, CD123, CD127-IL7Ra, CD161, CD27, CD278-ICOS, CD28, CD38, CD69, CD79b and HLA-DR. The prediction of surface proteins on our dataset was performed using integrated and normalized data.

4.8 In silico analysis

The scRNA-seq data were retrieved from the Gene Expression Omnibus (GSE) 148190. The analysis was restricted to K383, K409 and K411 samples containing blood (B), lymph nodes metastasis (LN) or Tumor (T) lymphocytes. The datasets used were GSM4455931, GSM4455932, GSM4455933, GSM4455935, GSM4455937 and GSM4455938. Data from each dataset were cleaned selecting the cells expressing <10% of mitochondrial genes, read counts of at least 200 genes and <3000 genes. Then all datasets were integrated and normalized yielding a total of 26,757 (11,614 of B, 12,915 of LN and 2170 of T). We performed clustering and dimensional reduction using UMAP (see methods) finding ten clusters at the resolution of 0.3. Signature of each cluster was obtained by using “FindConservedMarkers” function coded in the Seurat R package. MAIT signature was confirmed by using GeneOverlap [177].

CD8+T cells derived from 48 human melanoma samples (accession number GSE120575) were re-analyzed using R based Seurat v3.0. Raw data were normalized (Log-Normalize) and top 6000 variable genes were selected and used for dimensionality reduction (UMAP) and clustering analysis. Instead, all genes were used for testing differential expression. Unsupervised clustering was performed by finding the K-nearest neighbors (KNN) and then, to group the cells, a modularity optimization-based algorithm was applied. Among the six clusters identified at the resolution of 0.3, we detected MAIT cells based on the expression of *SLC4A10*, *CD69*, *CXCR4* and *KLRB1*. Then the frequencies of MAIT cells, between responder and non-responder, were computed and the Mann-Whitney *U*-test was used to perform the statistical analysis.

4.9 Statistical analysis

Statistical analyses were performed using GraphPad Prism version 6 (GraphPad Software Inc., La Jolla, USA), unless specified otherwise. Significance of differences for the frequency of single Phenograph clusters was determined using two-way ANOVA with Bonferroni post-hoc test. To compare distributions of manually gated subsets significance was determined by paired Wilcoxon t test, unless otherwise specified in the figure legends. Simplified Presentation of Incredibly Complex Evaluation (SPICE) software (version 6, kindly provided by Dr. Mario Roederer, Vaccine Research Center, NIAID, NIH, Bethesda, MD, USA) was used to analyze flow cytometry data on T-cell polyfunctionality [178]. Comparison of the curves of response to therapy was performed by Log-Rank (Mantel-Cox) test and p-value was considered significant <0.05. Finally, dividing patients by gender did not result in any statistical difference.

4.10 Data availability

Original .fcs files concerning cytofluorimetric analysis (T cells landscape, deep phenotyping of MAIT cells and in vitro assay) are deposited at the flowrepository.org in the following folders: FR-FCM-Z36A; for MAIT cells: FR-FCMZ36B. The scRNA-seq data have been deposited in GEO under the accession code GSE166181.

4.11 Code availability

<https://github.com/DomenicoSkyWalker89/CD8-T-lymphocytes-MAIT>

5. METHODS 2 (mRCC)

5.1 Patients' selection

The study was conducted on 19 patients with metastatic renal carcinoma treated with nivolumab at the Division of Oncology – Azienda Ospedaliero Universitaria di Modena and Reggio Emilia. According to RECIST, 5 patients were classified as responders (R) and 14 as non-responders (NR). The clinicopathologic characteristics of patients are reported in **Table 2**. According to the international metastatic RCC database consortium (IMDC) 60% of R patients had a good prognosis and 64.3% of NR patients had an intermediate prognosis. The mean age of the total cohort was 69 years (range 55-81). Eighty-four percent of patients showed lung metastasis. All patients had previously received other therapies: 52.7% received one previous therapy, 31.5% received two previous therapy and 15.8% received three previous therapies. In particular every patient was treated before starting ICI therapy with different tyrosine kinase inhibitor: 1) Sunitinib, 2) Cabozantinib, 3) Pazopanib and 4) Axitinib (the latter only in third line). All patients were completely washed-out for 1 month before starting ICI therapy.

Variable	RCC (n=19)
Demographic characteristics	
Age, mean yr (range)	69 (55-81)
Male Sex, N (%)	15 (79.0)
Clinical characteristics	
Years from diagnosis, mean (range)	10.8 (3-29)
Fuhrman grade of renal-cell carcinoma	
1 or 2, N (%)	7 (36.8)
3 or 4, N (%)	7 (36.8)
Unknown, N (%)	5 (26.3)
Tumour stage	
T1, N (%)	6 (31.5)
T2, N (%)	5 (26.3)
T3, N (%)	5 (26.3)
Tx, N (%)	3 (15.8)
Node stage	
N0, N (%)	8 (42.1)
Nx, N (%)	11 (57.9)
Median number of metastatic sites (range)	2 (1-5)
Location of metastasis	
Lung, N (%)	16 (84.0)
Liver, N (%)	2 (15.9)
Lymph nodes, N (%)	10 (52.6)
Pancreas, N (%)	4 (21.0)
Blood parameters	

Basal white blood cells, N/uL	6,483
Basal neutrophils, mean N/uL	4,439
Basal lymphocytes, mean N/uL	1,544
Basal CRP, mean mg/L	6.13
Previous therapy	
One previous therapy, N (%)	10 (52.7%)
Two previous therapies, N (%)	6 (31.5%)
Three previous therapies, N (%)	3 (15.8%)

Table 2. Demographic and clinical characteristics of mRCC patients.

5.2 Samples collection

All patients provided written informed consent for sample and clinical data collection according to the Azienda Ospedaliero Universitaria di Modena and Reggio Emilia ethical approval protocol #Prot AOU 195/2016. For each patient, at the different time-point, was collected up to 30mL of peripheral blood into ethylenediaminetetraacetic acid (EDTA) tubes. Briefly the samples were collected before starting ICI therapy (hereafter indicated as T0), 14 days after the first, the second or the third cycle of ICI therapy (hereafter indicated as T1, T2 and T3, respectively). PBMCs were isolated by Ficoll density centrifugation, cryopreserved in FBS supplemented with 10% DMSO and stored in liquid nitrogen until use [175].

5.3 Polychromatic flow cytometry

- a) **T cells landscape.** Multicolor flowcytometry panel was designed to fine characterize the T cells landscape using conjugated antibodies CD3, CD4, CD8, CD45RA, CD197, CD28, CD27, CD127, CD95, CD98, CD71, CD25, HLA-DR, CD38, CD39, CXCR6, CCR4, Ki67, TBET, granulysin, PD-1, BTLA, CD244, and ICOS. According to the therapy the panel was optimized to identify the expression of PD-1 on T cells isolated from patients treated with anti-PD-1 (either nivolumab or pembrolizumab) using an anti IgG4 to recognize the anti-PD-1 bound to PD-1 [19]. All antibodies were previously titrated to define the optimal concentration, the relative information was reported in **supplementary Table 1**. Briefly the high-dimensional flow cytometry was performed as reported below. Cryopreserved samples were thawed in R10 medium, i.e., RPMI supplemented with 10% fetal bovine serum (FBS), 100 U/mL penicillin, 100 µg/mL streptavidin, 2mM L-glutamine, 20mM HEPES (ThermoFisher, Eugene, OR) and 20 µg/mL DNase I from bovine pancreas (Sigma-Aldrich, St. Louis, MO). After washing with

phosphate buffer saline (PBS), cells were stained immediately with the Zombie Aqua Fixable Viability kit (BioLegend, San Diego, CA) for 15 min at room temperature. Then, cells were washed and stained with the combination of monoclonal antibodies (mAbs) listed in **supplementary Table 1**. Chemokine receptors were stained for 20 min at 37 °C, for 20 min at room temperature. Intracellular detection of Ki67, granulysin and TBET was performed following fixation of cells with the FoxP3 transcription factor staining buffer set (eBioscience/ThermoFisher) according to manufacturer's instructions and by incubating with specific mAbs for 30 min at 4 °C. Samples were acquired on a FACS Symphony A5 flow cytometer (BD Biosciences) equipped with five lasers (UV, 350 nm; violet, 405 nm; blue, 488; yellow/green, 561 nm; red, 640 nm) and capable to detect 30 parameters. Flow cytometry data were compensated in FlowJo by using single stained controls (BD Compbeads incubated with fluorochrome-conjugated antibodies).

- b) **In vitro cytokine production.** Thawed PBMC were rested for 4 hours at 37°C and then in vitro stimulated with anti-CD3/CD28 (1ug/ml) (Miltenyi, Bergisch Gladbach, Germany) and suboptimal concentration of IL-2 (10 ng/mL) (Miltenyi). A 11 parameter/10-color flow cytometer panel was optimized to identify T_{SCM} and different subpopulations of T cells producing TNF, IFN- γ and IL-2 that were detected after 16 hours of incubation (see **supplementary Table 4**). For the quantification of intracellular cytokines, cells were fixed with BD Cytofix/Cytoperm Fixation/Permeabilization Solution kit (BD Biosciences) according to the manufacturer's instructions. Samples were acquired on an Attune NxT acoustic flow cytometer (ThermoFisher) equipped with four lasers (violet, 405 nm; blue, 488; yellow/green, 561 nm; red, 640 nm) and capable to detect 14 parameters. Flow cytometry data were compensated in FlowJo by using single stained controls as above.

5.4 High-dimensional flow cytometry data analysis using Full Annotation Using Shape-constrained Trees (FAUST)

Flow Cytometry Standard (FCS) 3.0 files were imported into FlowJo software version 10, compensated and biexponentially transformed. Then debris, aggregates, doublets and dead cells were excluded. Overlaying fluorescence histograms of all samples were used to identify differences in staining, background fluorescence and transformation. We imported basic FlowJo workspaces into R to create a gating set using flowWorkspace (v4.0.6). 'Live CD8+' T cells gate was selected as

startingNode and manually set channel bounds according to the FlowJo marker's distribution. FAUST makes minimal assumptions about the distribution of markers, according to this it detects the different sub-population in a data-driven fashion. To perform this, FAUST calculate the separation-score 'depth score' for each marker by sample, this value (by marker) indicates how consistently a marker separates populations across all the samples. Next, FAUST selects a subset of high scoring markers in the panel, calculates the sample-specific marker boundary and perform the clustering. In our analysis the following markers showed the highest 'depth score' and were selected for clustering: CD45RA, CD57, TBET, CD28, CD127, CD38, CCR7, GRANULYSIN, CD95, CXCR6 and PD1 (Indicated as PD1_IGG4) (**Figure 2**). The following markers HLA-DR, ICOS and Ki67 were excluded from the analysis due to batch error. Analysis was performed without downsampling. 'FAUST' (v 0.5.1) was run by setting tuning parameters 'samplename' for the experimentalUnit and '1' for nameOccuranceNum 25. The depthScoreThreshold and selectionQuantile analysing the depth score line plots for each channel were selected (**Figure 2**). The final countMatrix was filtered choosing the clusters that were at least 50% of samples with at least 200 cells for each phenotype. The nonlinear dimensional reduction was performed using Uniform Manifold Approximation and Projection (UMAP) algorithm (uwot v.0.1.8) through makeAnnotationEmbedding function. Finally, all data were graphed with ggplot2 (v. 3.3.2) using a custom script.

5.5 Statistical analysis

Statistical analyses were performed using GraphPad Prism version 6 (GraphPad Software Inc., La Jolla, USA), unless specified otherwise. To compare distributions of manually gated subsets significance was determined by Kruskal-Wallis test, unless otherwise specified in the figure legends. Simplified Presentation of Incredibly Complex Evaluation (SPICE) software (version 6, kindly provided by Dr. Mario Roederer, Vaccine Research Center, NIAID, NIH, Bethesda, MD, USA) was used to analyze flow cytometry data on T-cell polyfunctionality [178]. The two-way analysis of variance (ANOVA) with Bonferroni correction was used to test each sub-population for differential abundance during the time-point therapy.

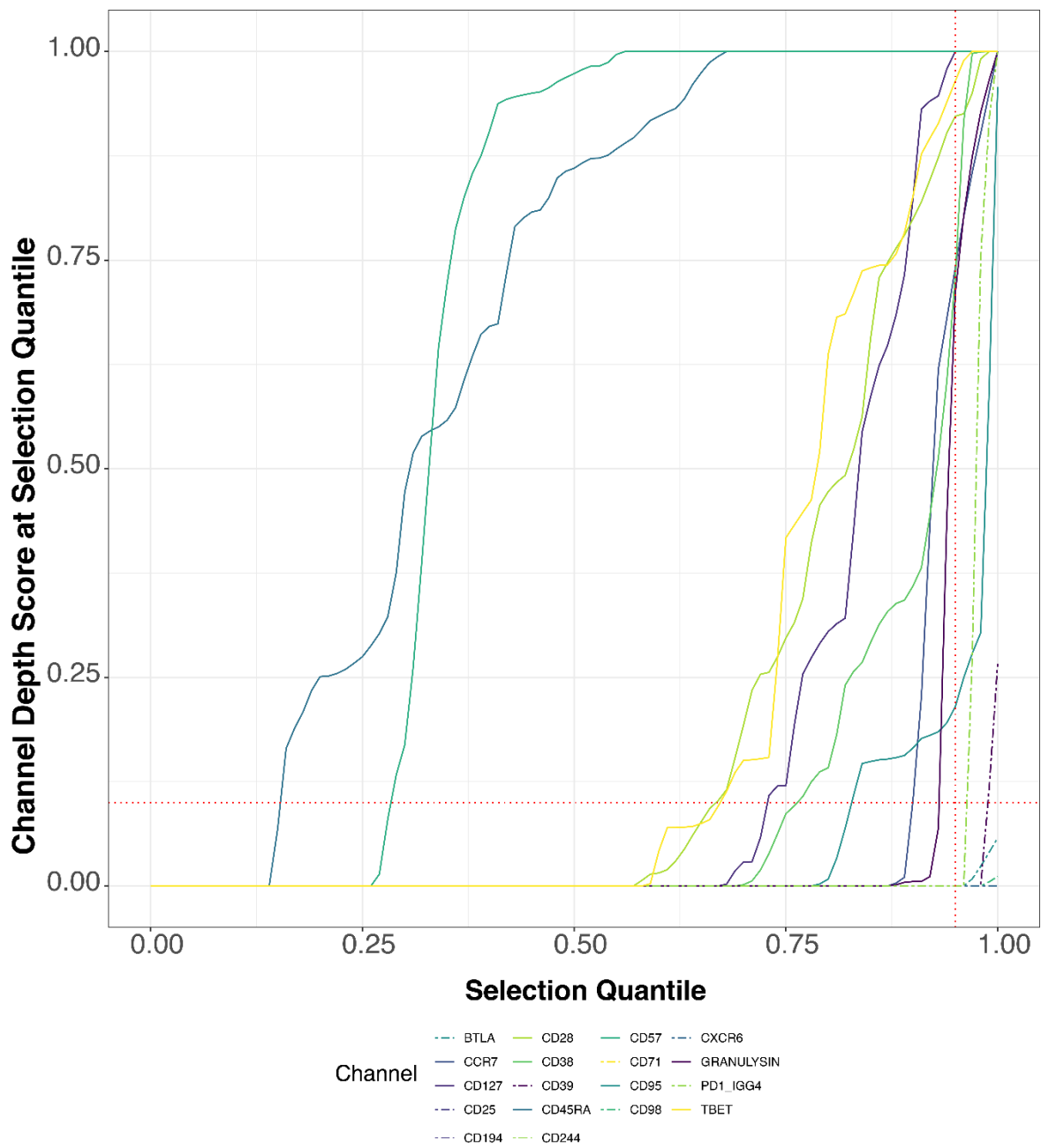


Figure 2. FAUST separation-score 'depth score' for each marker by sample. In our analysis, the following markers showed the highest 'depth score' and were selected for clustering: CD45RA, CD57, TBET, CD28, CD127, CD38, CCR7, GRANULYSIN, CD95.

6. RESULTS 1 (melanoma)

6.1 A higher proportion of activated effector memory CD8+T cells in responders

We initially used high-dimensional flowcytometry to longitudinally define the characteristics of T cells upon PD-1 blockade in melanoma patients (**Figure 3**). Computational analysis of aggregated data from multiple patients and time points identified 28 clusters (individually labeled as C) among CD8+T cells, resolving a broad spectrum of T-cell states, including maturation, activation, and exhaustion. C21, C22, C26 display phenotypic identity proper of subsets of naive T cells, characterized by expression of CD45RA, CCR7, CD27, CD28, negligible expression of CD25 and ICOS, and absence of additional markers [179] (**Figure 4A**). C28 represents recently activated T cells characterized by expression of CD38 and ICOS, but no expression of the late activation marker HLA-DR. T stem-cell memory cells are identified in C10, and their phenotype is similar to that of naive T cells and includes the expression of CD95 [180]. C1, C20, and C27 represent CCR7+CD45RA- central memory T cells characterized by high expression of CD28, CD27, BTLA, CD194, CD25, CD95 and ICOS. C1 displays high levels of CD194, CD28, CD95 while C20 represents a cluster of T central memory (CM) cells that expresses high level of CD39. C6, C17, C11, C2, C24, and C5 represent terminally differentiated T cells, being characterized by the expression of CD45RA, but not of CCR7, CD27, or CD28, and high levels of CD244, CD57 and TBET. These cells also lacked granulysin expression.

C15, C4, C25, C14, C9, C13, C12, C19, and C16 represent effector memory T-cell subsets characterized by the lack of expression of CD45RA, CCR7, and expression of CD25 and CD95 [180]. Among these, C14 expressed PD-1 and CD57, and TBET at intermediate levels, thereby suggesting the identification of replicative senescent cells [17]. C9, C13, C12 are transitional effector memory T cells as they express intermediate levels of CD28 and CD27 [181]. C9 expresses CXCR6, identifying effector memory cells with the capability to migrate to metastasis [182], while C12 is a cytotoxic T-cell subset displaying high level of granulysin. C19 display high levels of CD127, CD39, and CD25, identifying not only metabolically activated, but also tumor-reactive cells [183]. C16 is a cluster of activated and proliferating effector memory T cells characterized by high level of expression of Ki67, ICOS, CD95, HLA-DR, CD71, CD98, CXCR6, granulysin, CD38, intermediate expression of CD127, CD39, CD25, CD28, CD194, CD27, BTLA, TBET and CD244; as shown in **Figure 4B**, this cluster was much more represented in responder (R) patients if compared to non-responders (NR).

Longitudinal analysis did not identify obvious differences in the dynamics of these immune populations between responders and non-responders to anti-PD-1 therapy (**Figure 4B**). Cross sectional analysis identified C16, highly proliferating Ki67+CD71+ effector cells equipped for cytotoxicity (GNLY+), whose relative proportion was higher in responder before starting therapy ($p < 0.001$) (**Figure 4B**). This difference remained stable also after treatment ($p < 0.01$) (**Figure 4B**).

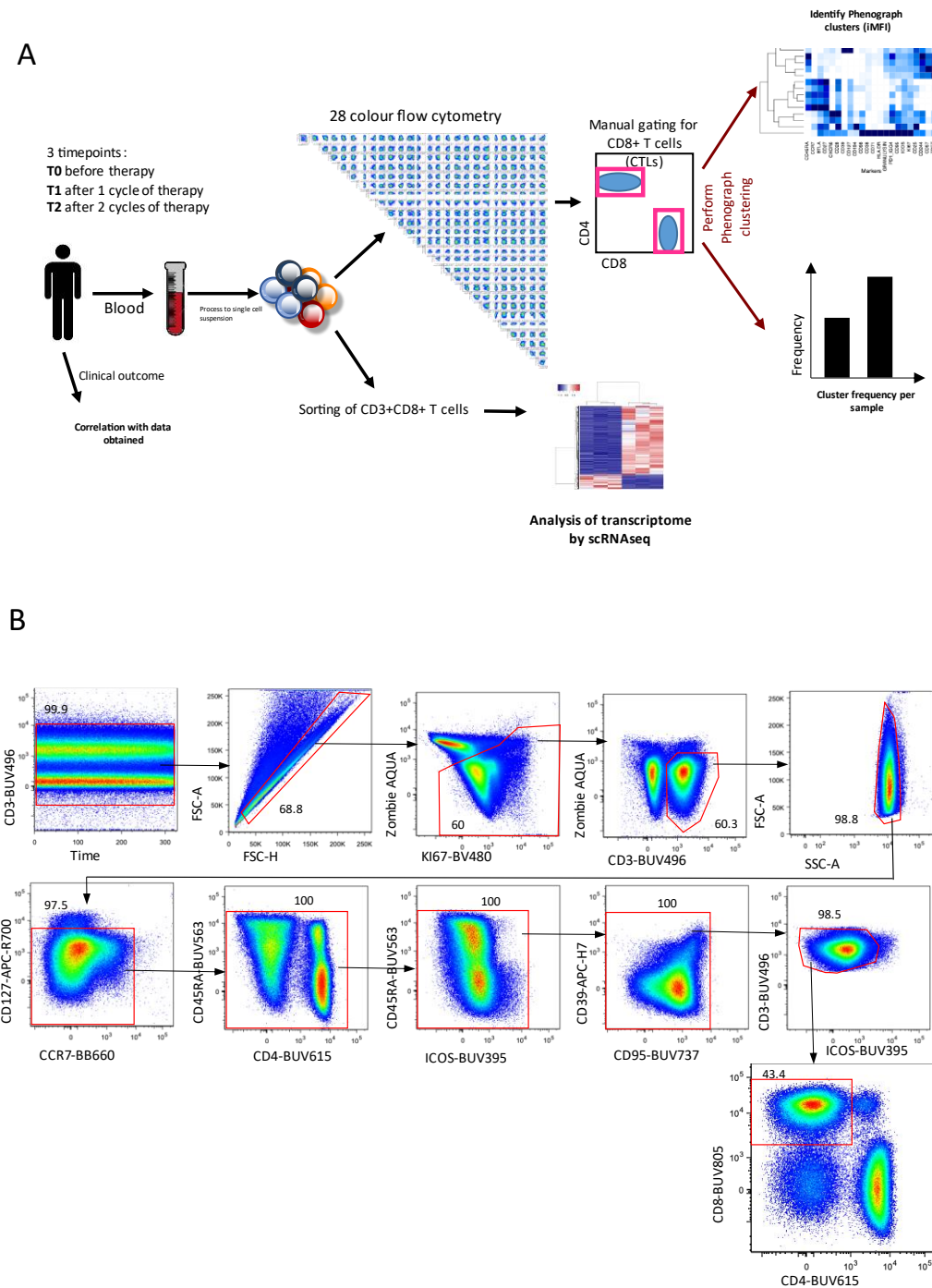


Figure 3.

A) Experimental workflow on PBMC isolated from patients with metastatic melanoma. **B)** Gating strategy used to identify CD8+ T cells.

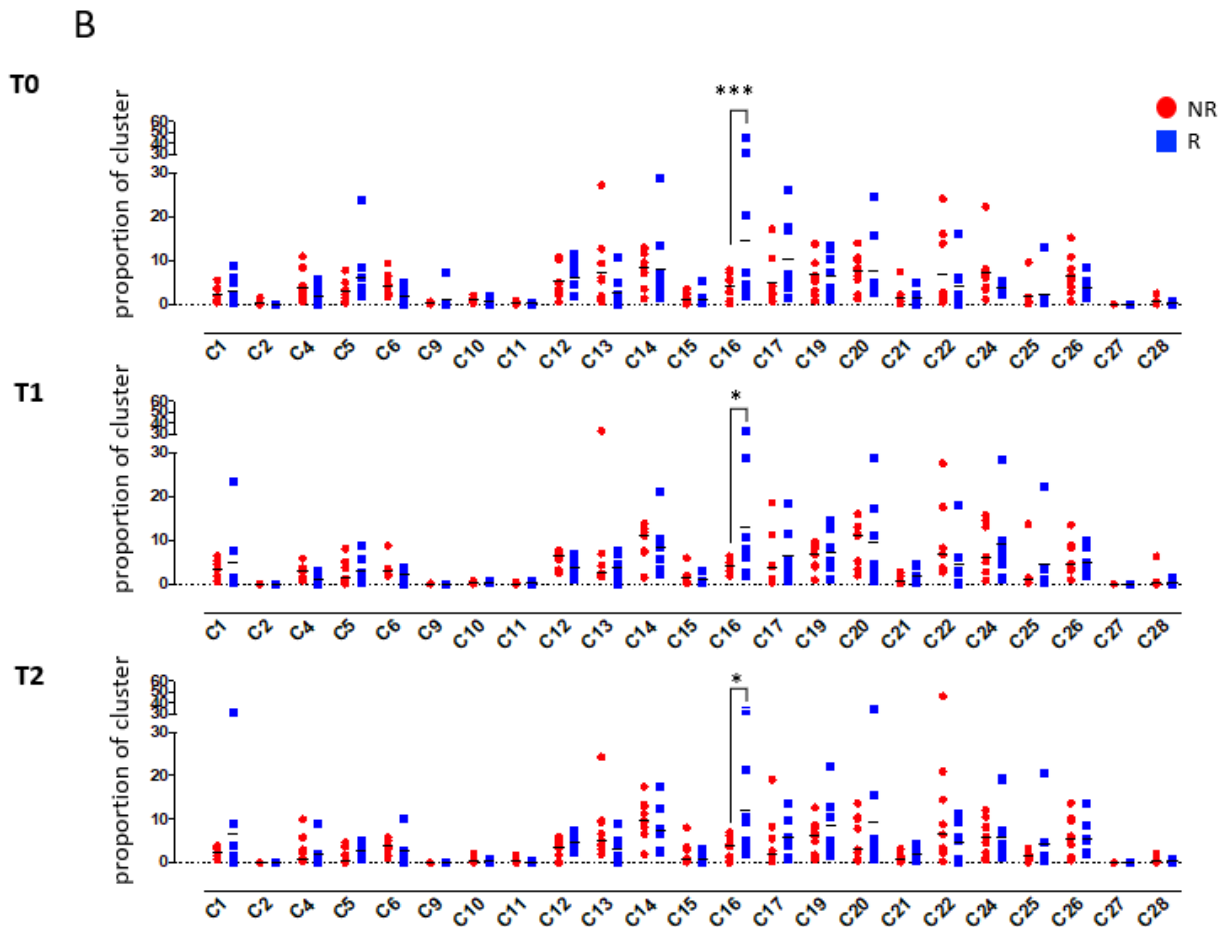
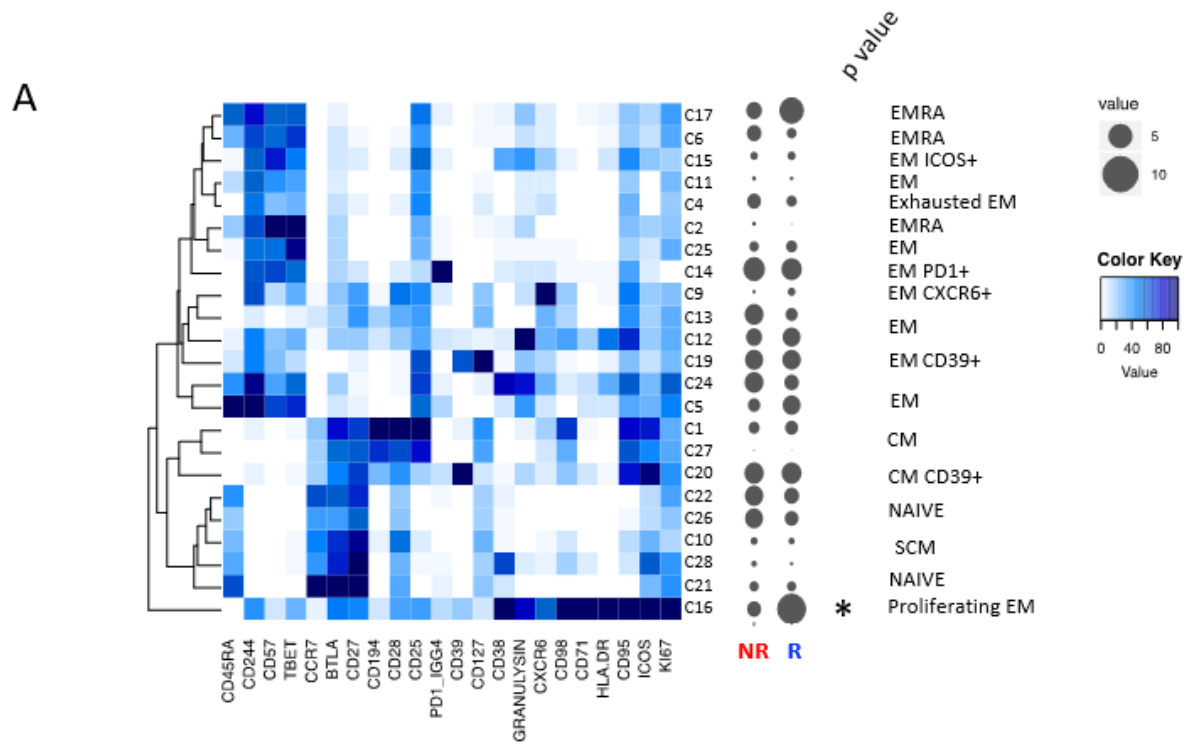


Figure 4. High-dimensional single-cell analysis of CD8⁺ T cells identifies higher proportion of activated effector memory T cells in responders.

A) Heatmap showing the iMFI of specific markers in discrete Phenograph clusters. Ballons indicate the median frequency of each cluster amongst responders and non-responders. Statistical analysis by two-sided Mann–Whitney nonparametric test, Bonferroni’s multiple comparisons test, * $p = 0.046$. **B)** Proportion of cell within each cluster for each individual. Data represent individual values, mean (center bar) $\pm$ SEM (upper and lower bars). Statistical analysis by two-sided Mann–Whitney nonparametric test, Bonferroni’s multiple comparisons test; T0, *** $p < 0.0001$; T1, * $p = 0.046$; T2 $p = 0.037$. T0 = before therapy, T1 = after 1 cycle of therapy, T2 = after two cycles of therapy. Individual measurements of NR = 9, R = 8.

6.2 MAIT cells are more abundant in responders as revealed by scRNA-seq

To further define the dynamics of T cells potentially involved in therapeutic response, we performed scRNA-seq of isolated CD3+CD8+T cells from a total of 20 patients at T0, T1, and T2 after anti-PD-1. After quality control (see Methods), 56,142 cells were deemed suitable for analysis. Contaminating 4210 NK and 231 monocytes cells, expressing *TYROBP*, *FCGR3A*, *KLRB1*, and *LYZ*, respectively, were removed from the analysis. We obtained a total of 51,701 purified CD8+T cells. Using a cTP-net, a deep neural network trained on multi-omics data, we imputed surface protein abundances within the scRNA-seq data to confirm T-cell phenotype [134] (**Figure 5**).

Computational analysis identified eight different cell clusters on the basis of gene expression profiles (**Figure 6A, B**). Naive T cells were identified by expression of *LEF1*, *SELL*, *TCF7* genes while terminally differentiated effector memory cells, with cytotoxic properties were characterized by the expression of *GZMB*, *GNLY*, *NKG7*, *EFHD2*, and *CXCR3* [184]. Two different clusters of effector memory cells were recognizable: one cluster of transitional effector memory (characterized by the expression of *GZMK* and *LYAR*) and one of more mature and activated phenotype with homing properties (expression of *TNFAIP3*, *CXCR4*, *CREM*, *CD69*) [53]. Two clusters of recently activated naive T cells have been characterized: one expressed *GATA3* and *IL7R*, the other *FOS* and *JUN*. Activated and replicating effector memory T-cell clusters were identified by the expression of *HLADRA*, *HLADRB1*, *CD74*, *GZMA*, *PCNA*, *MKI67*, *TOP2A*, *MCM4*, *MCM*. Finally, mucosal-associated invariant T (MAIT) cells with homing properties were identified they expressed high level of *KLRB1*, *SLC4A10*, *MAF*, and *CXCR4* [53]. Pseudotime analysis revealed that the differentiation process started from naive T cells towards terminally differentiated T cells passing through activated naive T cells, transitional effector memory T cells and effector memory T cells (**Figure 7**). In this process, the transcriptionally distinct MAIT cells belong to a different branch of the Pseudotime trajectory compared to the rest of the T cells, albeit mapping close to effector memory T cells, in line with their shared phenotypic identity [162].

No main differences were found between R and NR in the proportion of naive, cytotoxic terminally differentiated and activated naive T cells, both before and after therapy, as revealed by analysis of gene expression profiles by scRNA-seq (**Figure 8**). The proportion of activated effector memory T cells, reminiscent of C16 as defined by flow cytometry, was higher after two cycles of therapy in R compared to NR (**Figure 6C**, left panel). At the same time, activated effector memory T cells from R expressed higher levels of genes indicating activation (*FOS*, *DUSP1*, *FGFBP2*, *HLAC*) and

cytotoxic behavior (GNLY, GZMH), thereby suggesting heightened functional capacity in R (**Figure 6C**, right panels).

The proportion of MAIT cells was higher in R before therapy and after the first cycle of therapy (**Figure 6D**, left panel). This trend was visible also after the second cycle of therapy. Similarly, to EM T cells, also MAIT cells showed overexpression of genes related to cell activation in R compared to NR before (*TNFAIP3*, *NKG7*, *NFKBIA*, *JUND*, *ZNF331*, *RGCC*) or after the first (*ZFP36L2*, *BTG1*, *ARL4C*, *CXCR4*, *ID2*, *FOS*, *ZFP36*) or the second cycle (*DUSP1*, *FOS*, *TNFAIP3*, *GZMK*, *JUND*) of anti-PD-1 therapy (**Figure 6D**, right panels), overall suggesting a dynamic regulation of MAIT cell activation over time.

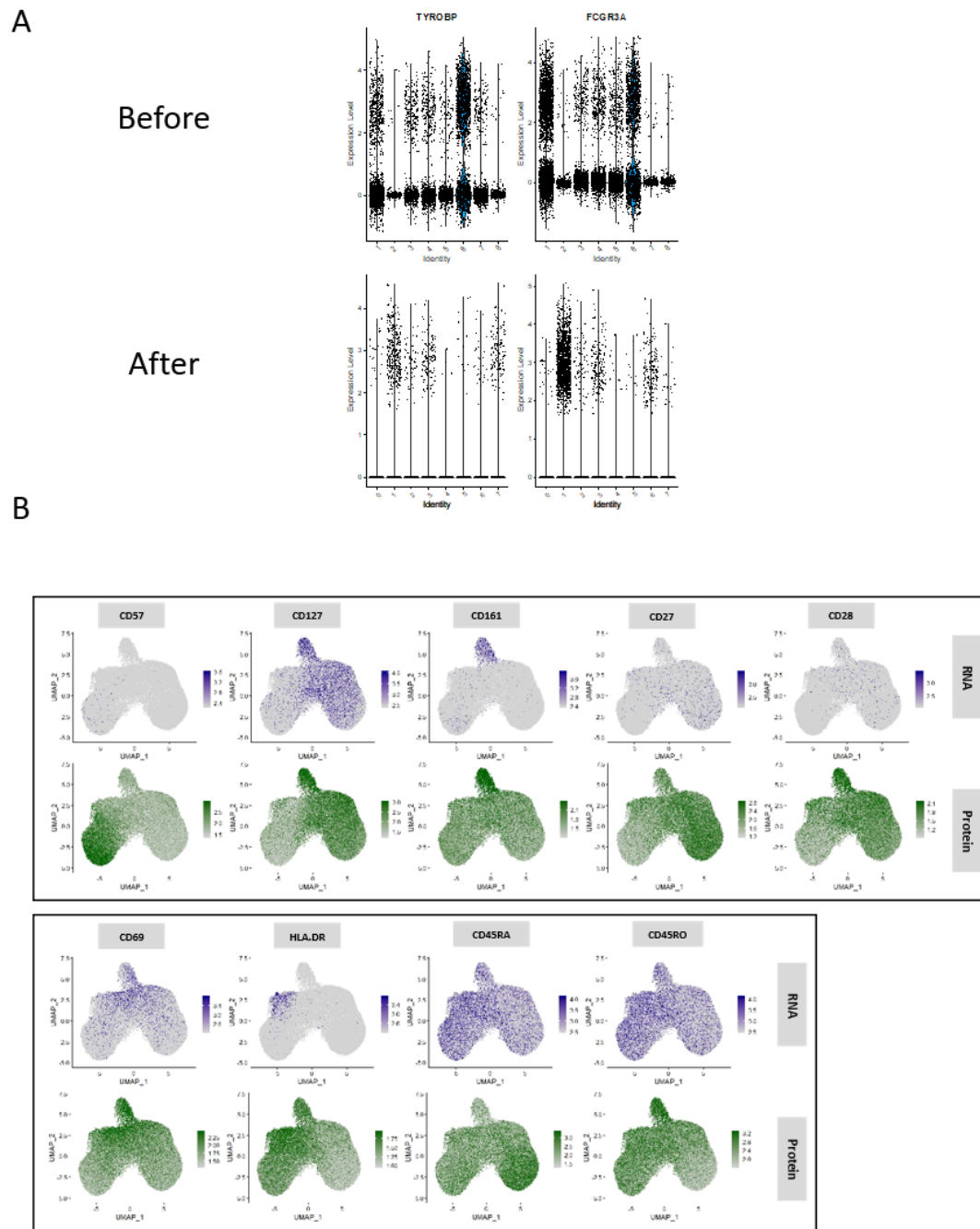


Figure 5. Quality control and imputation of surface protein expression in scRNA-seq data.

A) Gene expression analysis of TYROBP, FCGR3A and KLRB1 before and after the removal of NK cells.

B) Neural network analysis run after the removal of NK cells. This analysis compares gene and protein expression. The most representative genes expressed by CD8+ T cells are shown.

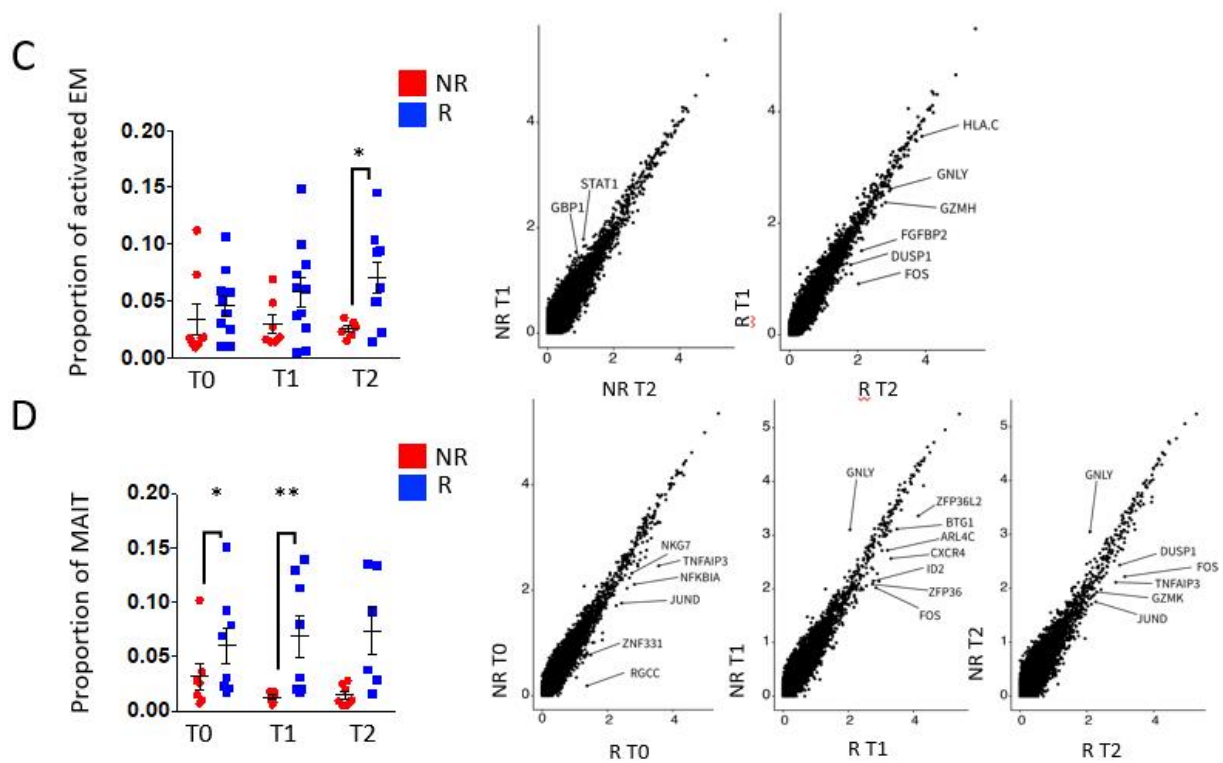
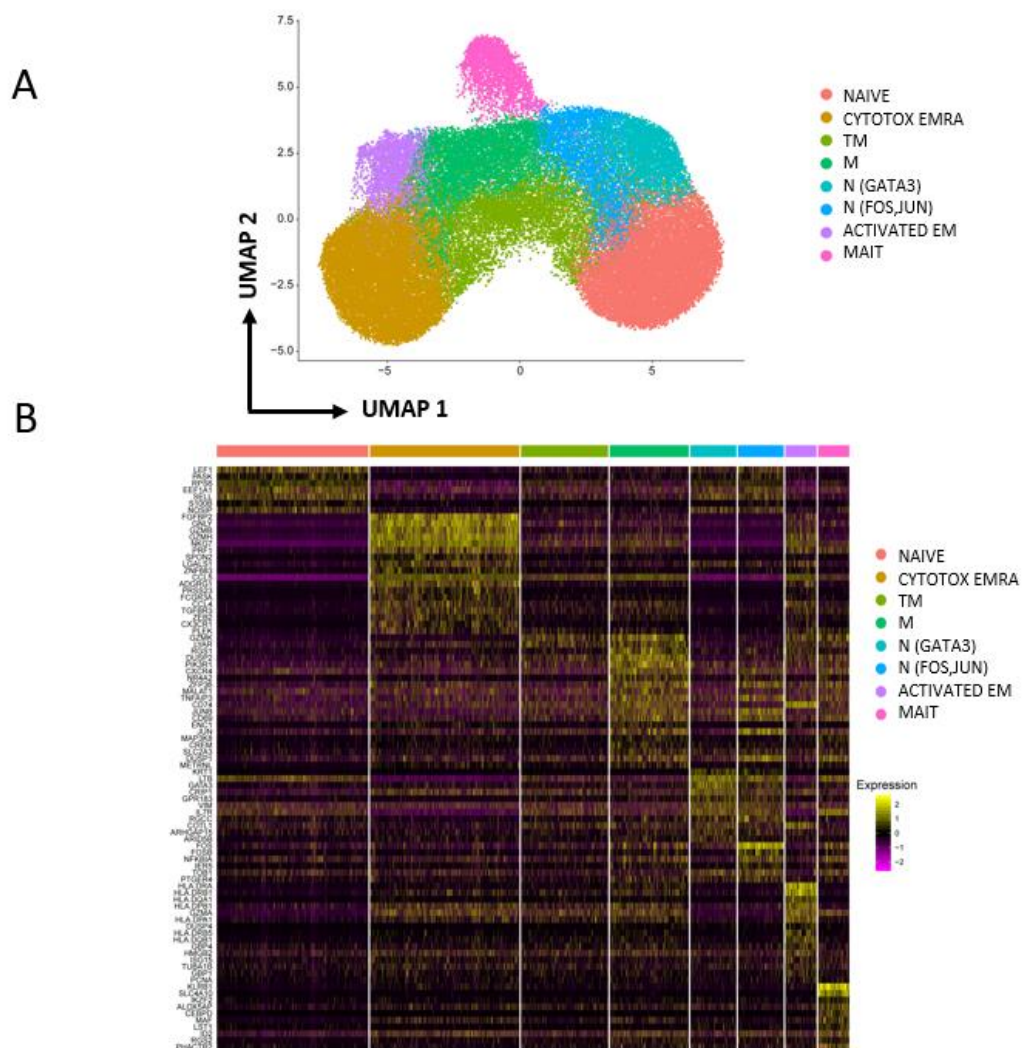


Figure 6. MAIT cells are more abundant in responders as revealed by scRNA-seq.

A) UMAP plot depicting CD8⁺ T-cell heterogeneity. Cells are colored according to the eight clusters defined in an unsupervised manner. **B)** Heatmap displaying scaled-expression values of discriminative gene set per cluster related to CD3⁺CD8⁺ T cells that passed quality control. A list of the most representative genes is shown per each cluster (left). N, naive; EMRA, effector memory expressing CD45RA; TM, transitional memory; M, memory; EM, effector memory; MAIT, mucosal-associated invariant T cells. **C)** Proportion of activated effector memory (EM) CD8⁺ T cells at different time points (left) and differential gene expression in this cluster between responders and non-responders at T2 (right). p-values of the differential expression analyses are reported in source tables. Only genes differentially expressed are reported in the figure. Statistical analysis by Mann–Whitney nonparametric test, Bonferroni’s multiple comparisons test; *p = 0.046. **D)** Proportion of MAIT cells and differential gene expression of this cluster between responders and non-responders at T0, T1, and T2 (right). p-values of the differential expression analyses are reported in source tables. Only genes differentially expressed are reported in the figure. Data represent individual values, mean (center bar) ± SEM (upper and lower bars). Statistical analysis by two-sided Mann–Whitney nonparametric test, Bonferroni’s multiple comparisons test; if not indicated, p-value is not significant. *p = 0.023; **p = 0.0012. Source data are provided as a Source Data file. T0 = before therapy, T1 = after 1 cycle of therapy, T2 = after two cycles of therapy. Individual measurements of NR = 8, R = 11.

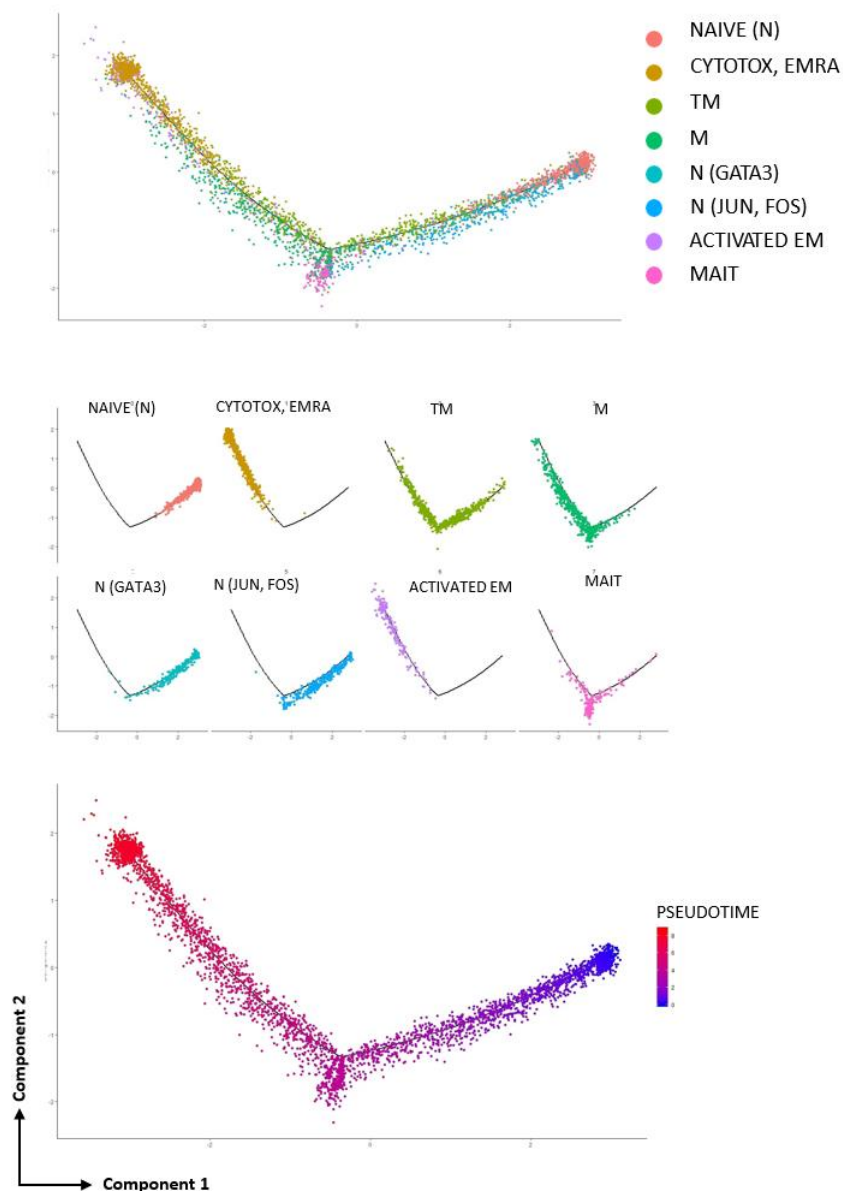


Figure 7. Trajectory analysis of the eight CD3+CD8+ T cells clusters.

CD8+ T cells distribution in a two-dimensional independent space showing the dynamic-programming of different clusters starting from naïve T cells (in blue in the pseudotime plot). Solid black line indicates the main diameter path of the minimum spanning tree (MST) and provides the backbone of Monocle's pseudotime ordering of the cells. Each dot represents an individual cell colored by cluster or by pseudotime.

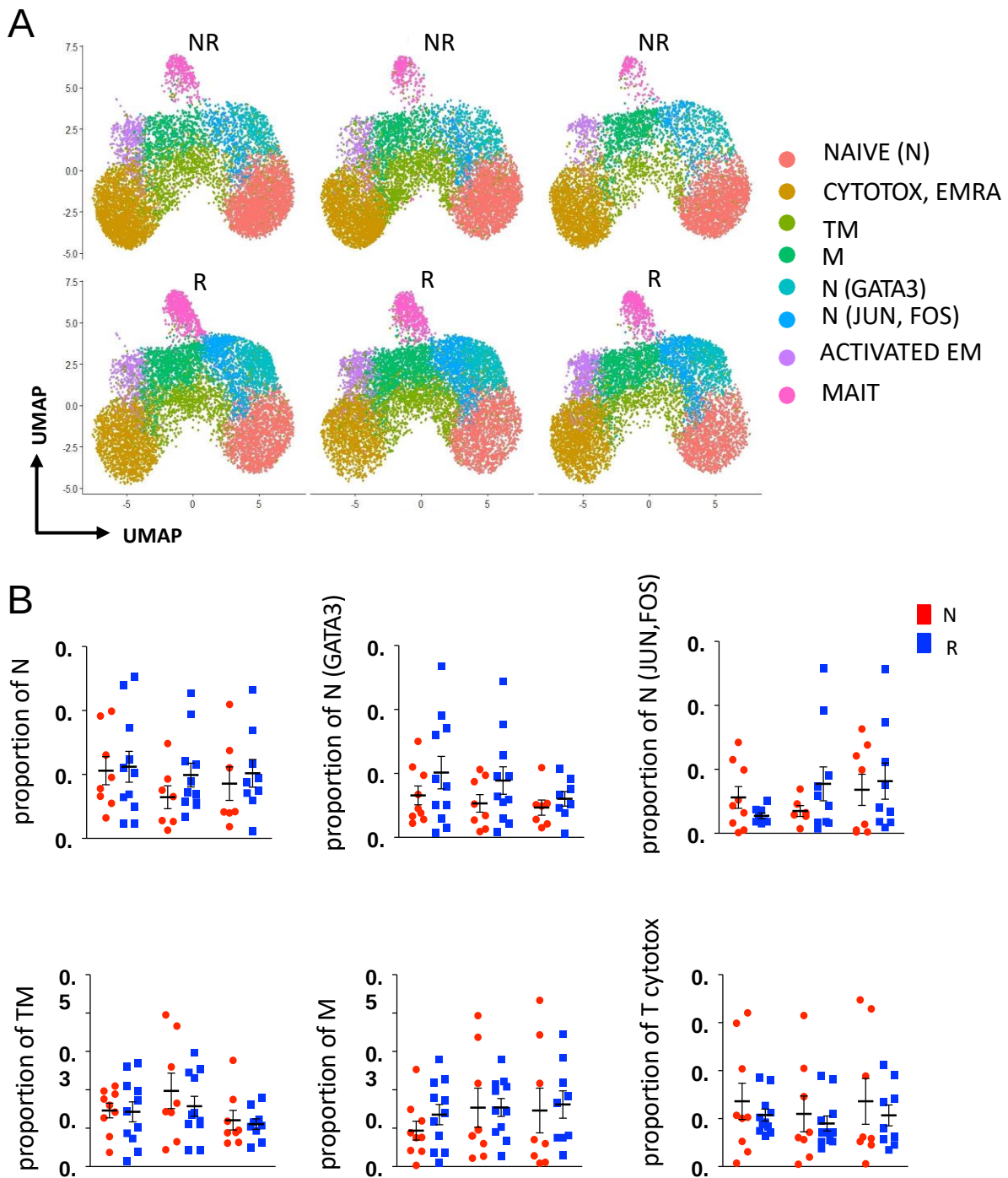


Figure 8. Landscape of CD8⁺ T cells in melanoma patients treated with ICI revealed by scRNA-seq.
A) Distribution of different cell clusters identified by scRNA-seq analysis. **B)** Proportion of different cell clusters identified by scRNA-seq analysis. Data represent individual values, mean (centre bar) $\pm$ SEM (upper and lower bars).

6.3 Activated MAIT cells with homing properties are more abundant in responders

We further subclustered MAIT cells before and after therapy to gain more insights into the cellular dynamics of these cells during the anti-tumor immune response. Our approach identified two different types of MAIT cells with differential expression of genes related to T-cell activation or effector functions *DUSP1*, *ZFP36*, *TNFAIP3*, *ZFP36L2*, *FOS*, *CXCR4*, *NFKBIA*, *CD69*, *TSC22D3*, *BHLHE40* and *JUN*, thereby suggesting the identification of quiescent and activated subsets of cells (**Figure 9A, B**). In line with previous data, R showed a significantly higher proportion of activated MAIT compared to NR not only before therapy, but also after the first and the second cycle (**Figure 9C**).

We next used polychromatic flow cytometry to confirm these findings also at the protein level. In this regard, we analyzed the proportion and phenotype of MAIT cells, identified as CD3+CD8+T cells that expressed TCR α 7.2 and CD161 (**Figure 11A**, left), and found marked expansion of these cells in the circulation of R patients when compared to NR before therapy (**Figure 10A**, right). This difference waned after therapy introduction in line with the data obtained using scRNA-seq.

Moreover, we found that the proportion of MAIT cells expressing the homing receptor CXCR4 increased after two cycles of therapy in R, but not in NR, that had a relevant variability (**Figure 11B**).

To confirm the presence of MAIT cells in the metastasis and primary tumor site we analyzed a public dataset available on Gene Expression Omnibus (GSE148190). This dataset contains single-cell RNA and TCR sequencing of PBMCs and tumor-infiltrating lymphocytes from untreated patients with metastatic melanoma. We used the scRNA-seq data of blood (B), lymph nodes metastasis (LN), and Tumor (T) from patients K383, K409, and K411. A total of 26,757 cells have been analyzed (11,614 of B, 12,915 of LN, and 2170 of T). About 3% of cells in LN and T were identified as MAIT cells expressing CXCR4 gene, suggesting their ability to home the inflamed tissue (**Figure 10B**). We confirmed the signature of the MAIT tumor-infiltrating genes by comparing it with that obtained by us from the blood circulating MAIT (**Figure 12A**). In addition, by analyzing a public dataset of CD8+ T cells obtained by melanoma patients treated with ICI (GSE120575; see “Methods”) [11], we found that MAIT cells increased in the metastatic lesions regressing after ICI compared to those that did not regress compared to baseline, thereby suggesting the potential recruitment of CXCR4-expressing MAIT cells expanding in the circulation (**Figures 12B**).

We next analyzed the effector functional capacity of the MAIT cells following in vitro stimulation with IL-12, IL-18, CD3/ CD28 followed by the detection of the effector molecules GRZM-

B, IFN- γ , and TNF (**Figure 13**). The overall quality of the response of MAIT cells, as assessed by combinatorial cytokine production, was largely similar between R and NR at different time points, where the majority of cells were capable to simultaneously produce GRZM-B, IFN- γ , and TNF. Moreover, R were characterized by higher proportion of cells producing IFN- γ and GRZM-B if compared to NR (**Figure 10C**). Nevertheless, before therapy, the proportion of cells able to produce only GRZM-B was higher in R if compared to NR (**Figure 10D**), thereby corroborating previous evidence that MAIT cells show preferential effector propensity.

6.4 Level of MAIT cells before therapy identifies responder patients

We next evaluated the prognostic significance of the levels of MAIT cells in the circulation as predictive biomarker of the response to anti-PD-1 therapy. Flow cytometric analysis revealed that, within CD8+T cells, the median level of MAIT in the population of patients with metastatic melanoma was 1.7%, thus this value was used as a cutoff to stratify patients. **Figure 14** reports that patients with a frequency of MAIT cells >1.7% had an increased probability to respond than those patient with MAIT cells.

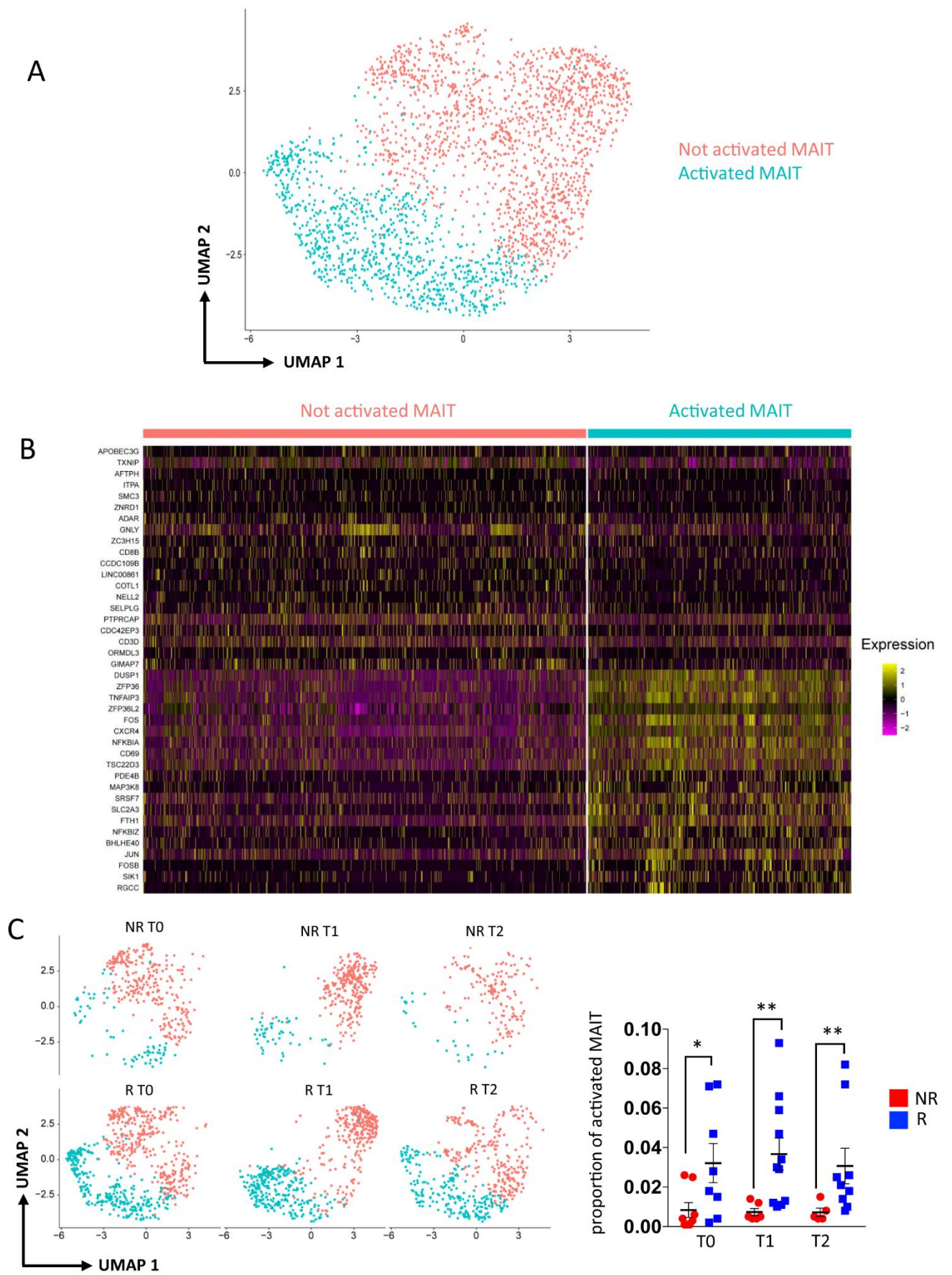


Figure 9. Activated MAIT cells with homing properties are more abundant in responders.

A) UMAP plot of MAIT cells. Not activated MAIT are in salmon and activated ones are in light blue.

B) Heatmap displaying scaled-expression values of discriminative gene set per each cluster of MAIT cells. A list of representative genes is shown on the left.

C) Left part: UMAP plot representing two clusters of MAIT cells between R and NR at T0, T1, T2. Right part: Proportion of activated MAIT cells between R and NR at T0, T1, T2. * $p = 0.04$; T1, ** $p = 0.005$; T2, ** $p = 0.007$. Statistical analysis by two-sided Mann–Whitney nonparametric test, Bonferroni’s multiple comparisons test. T0 = before therapy, T1 = after 1 cycle of therapy, T2 = after two cycles of therapy. Individual measurements of NR = 8, R = 11. Source data are provided as a Source Data file.

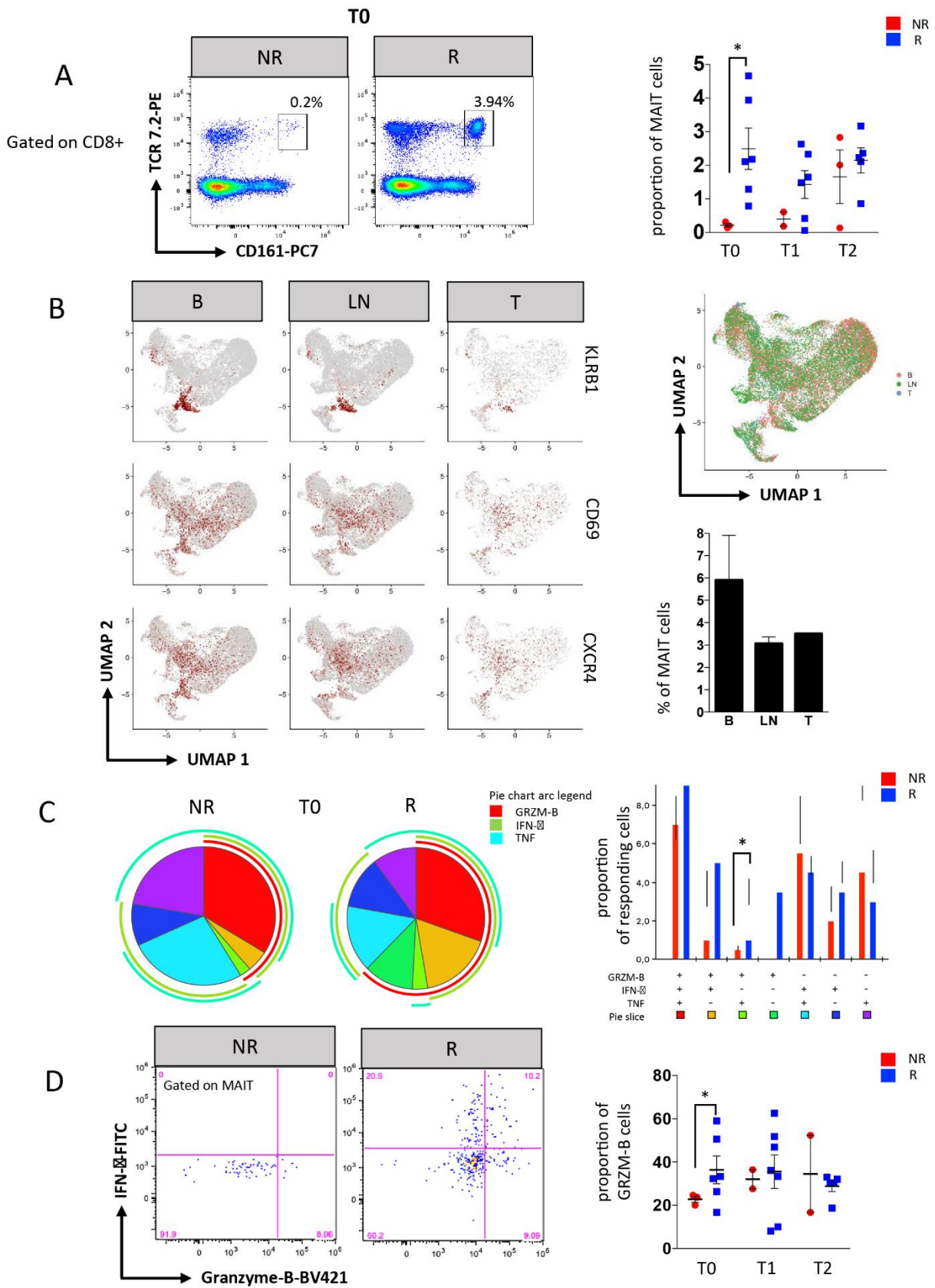
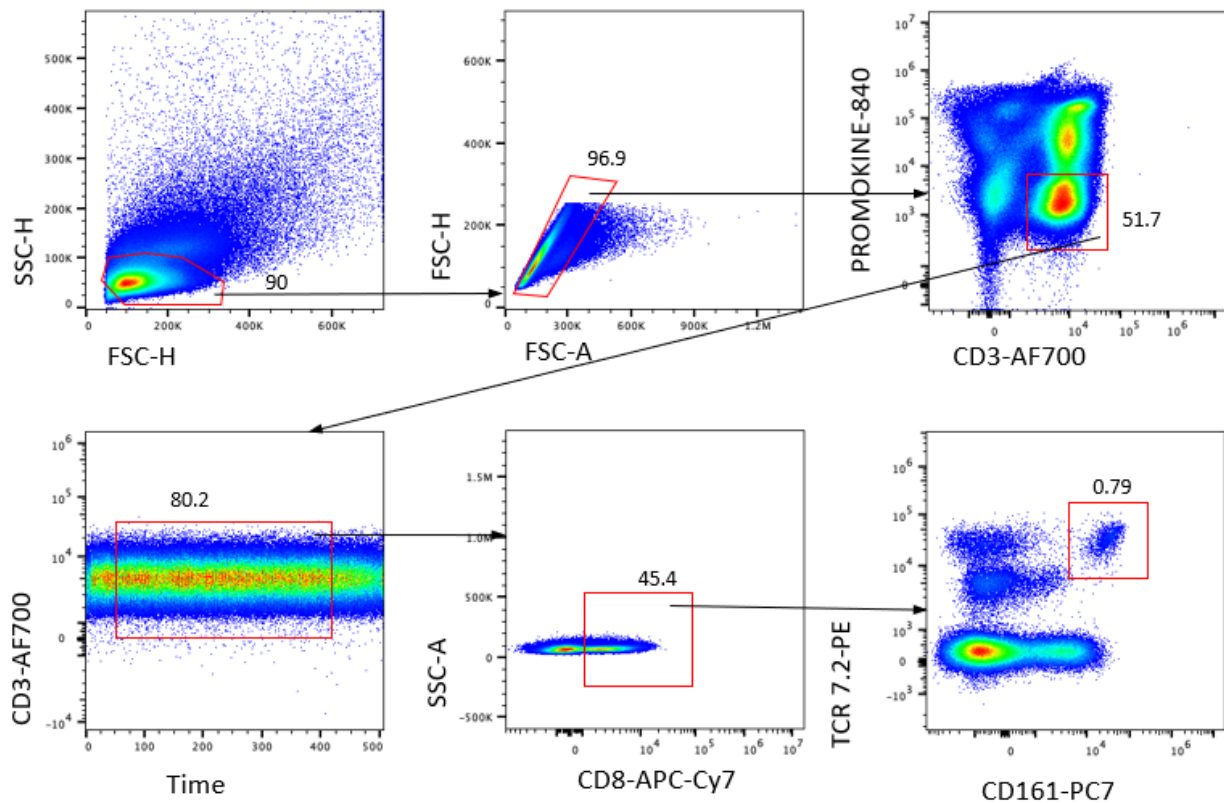


Figure 10. MAIT polyfunctionality evaluated after in vitro stimulation in PBMC of melanoma patients.

A) Left and center panels: representative dot plots of MAIT cells, identified as TCR 7.2+ and CD161+ within CD8+ T cells of one R and one NR at T0. Right part: proportion of MAIT cells in R and NR at T0, T1, T2. Data represent individual values, mean (center bar) $\pm$ SEM (upper and lower bars). Statistical analysis by two-sided Mann–Whitney nonparametric test, Bonferroni’s multiple comparisons test. * $p = 0.016$. NR = 4, R = 8. **B)** Left panel: UMAP representation of PBMC or tumor-infiltrating lymphocytes from patients with metastatic melanoma. Expression of KLRB1, CD69, and CXCR4 MAIT cells in blood (B), lymph nodes metastasis (LN), and Tumor (T) from the K383, K409, and K411 patients (Gene Expression Omnibus, GSE148190). Right panel: UMAP representation with the distribution of cells of B, LN and T, and proportion of MAIT in blood (B), lymphnode (LN) and tumor (T). **C)** Left and central panels: pie charts representing the proportion of MAIT cells producing different combinations of GRZM-B, IFN- γ , and TNF after stimulation at T0. Frequencies were corrected by background subtraction as determined in unstimulated controls; permutation tests (10,000 number of permutations), using SPICE software, show no difference between R and NR. Right panel: frequency of MAIT cells expressing and producing different combinations of GRZM-B, IFN- γ , and TNF after stimulation at T0. Statistical analysis by Wilcoxon rank test; * $p = 0.041$. In the figure, error bars (median value) and upper whiskers (whisker range, SEM) are represented. Individual measurements of NR = 4, R = 6. **D)** Left and central panels: dot plots show the difference between a R and a NR in the proportion of cells that produce IFN- γ and GRZM-B at T0. Right part, proportion of MAIT cells producing GRZM-B at different time points. Data represent individual values, mean (center bar) $\pm$ SEM (upper and lower bars). * $p = 0.028$, Statistical analysis by Mann–Whitney nonparametric test, Bonferroni’s multiple comparisons test. Individual measurements of NR = 4, R = 6. Source data are provided as a Source Data file.

A



B

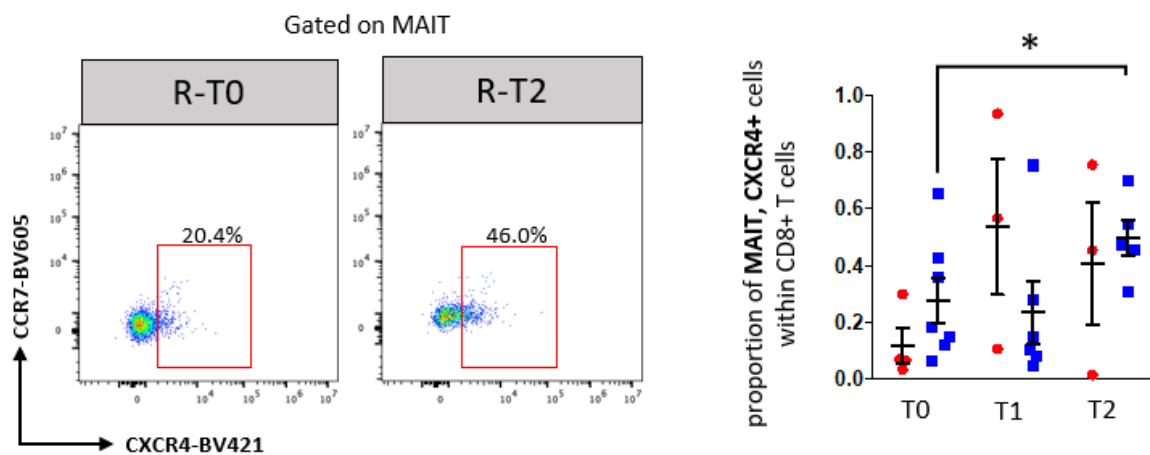
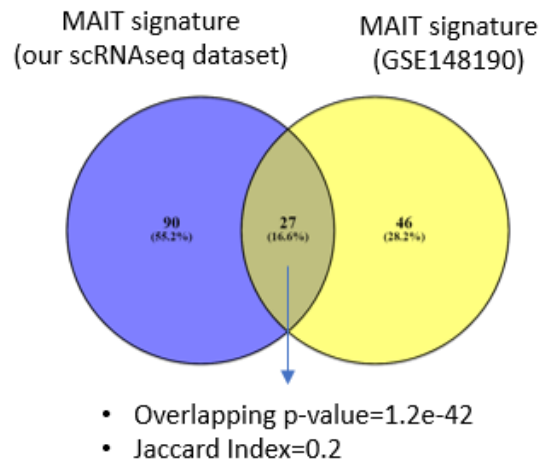


Figure 11.

A) Gating strategy used for the analysis of MAIT cells. **B)** Representative dot plots of MAIT cells expressing CXCR4 of one R analyzed at T0 (left) and T2 (central panel). On the right, proportion of MAIT cells expressing CXCR4 between R and NR at T0, T1, T2. * $p < 0.05$, Statistical analysis by Mann–Whitney nonparametric test and Bonferroni’s multiple comparisons test.

A



B

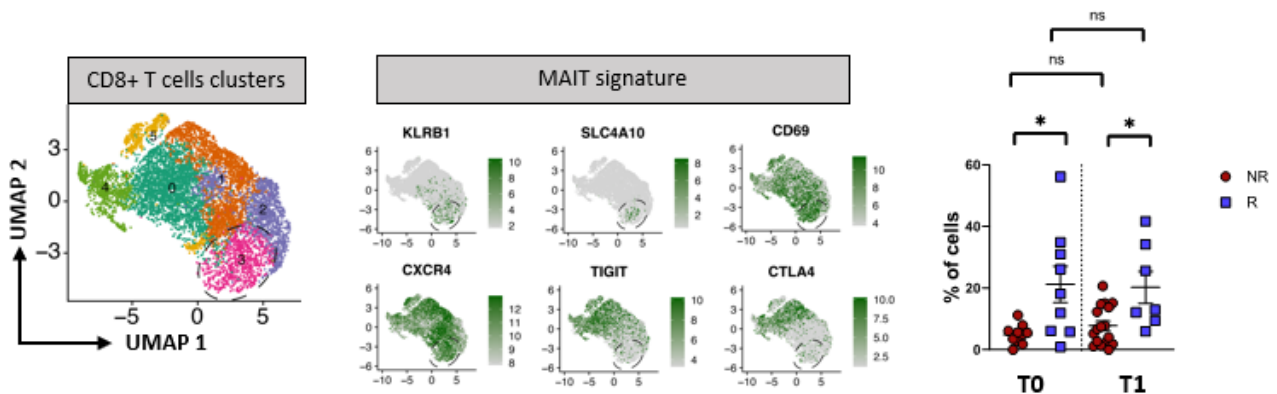


Figure 12. MAIT cells infiltrate melanoma metastatic lesions.

A) MAIT signature in TIL dataset. MAIT signature was confirmed by using GeneOverlap. In blue circle (left) numbers (percentage) of genes identifying MAIT cluster in our scRNAseq dataset; in yellow circle (right) numbers (percentage) of genes identifying MAIT cluster in GSE148190. The intersection between the two circles represents the number of genes (percentage) shared by the MAIT signature found in the two different datasets. **B)** Left: UMAP representation of CD8+ T cell clusters from melanoma lesions of R and NR patients to ICI therapy (GSE120575). Center: score of MAIT signature. Cluster 3 is the one with the highest score (expressing all MAIT genes), thus identifying MAIT cells. Right: Relative proportion (% among CD8) of cluster 3 of MAIT cells in R and NR pre and post ICI. P values were calculated using non-parametric Kruskal-Wallis test and Dunn's multiple comparisons.

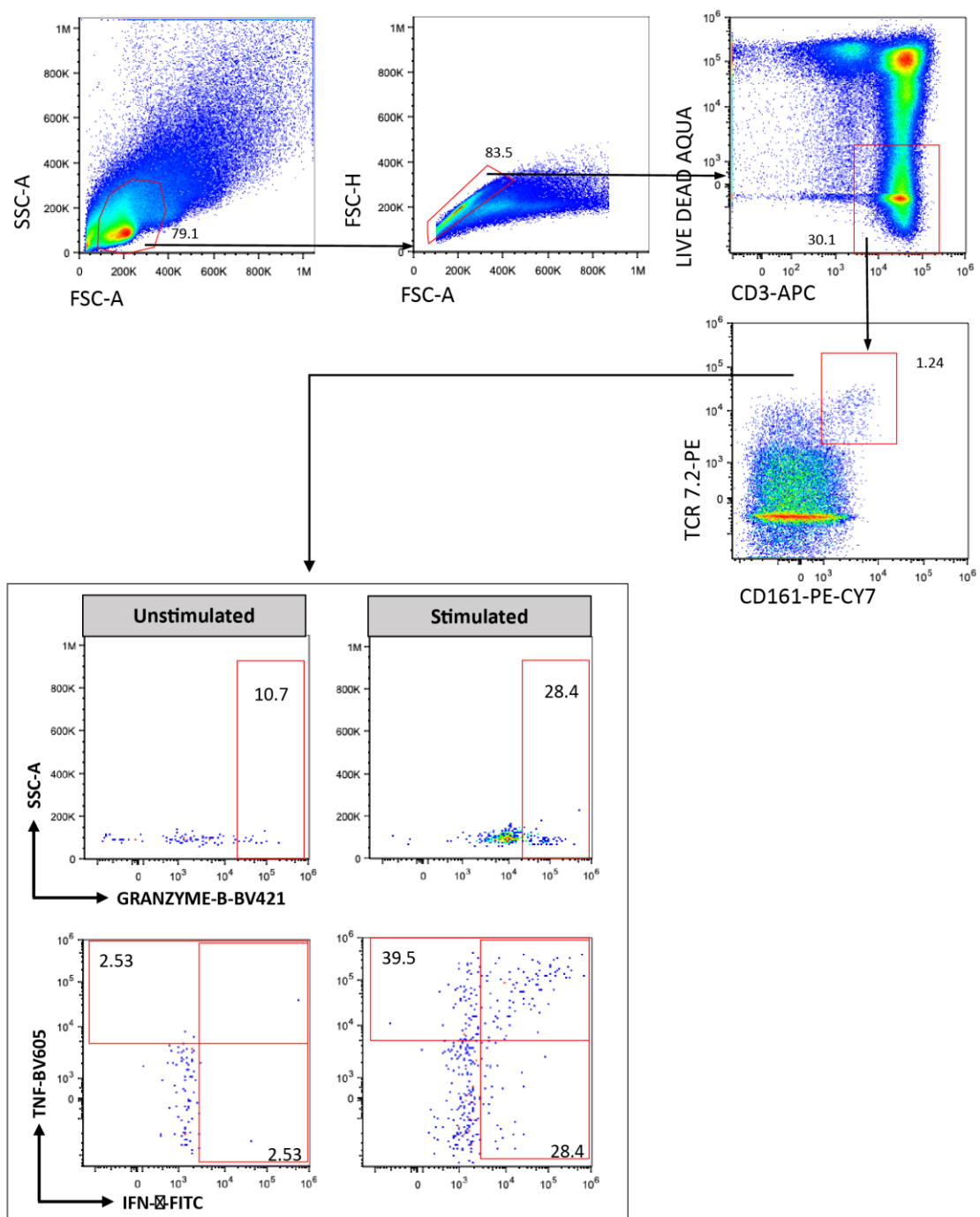


Figure 13.

Gating strategy used for the analysis of cytokine production by MAIT cells.

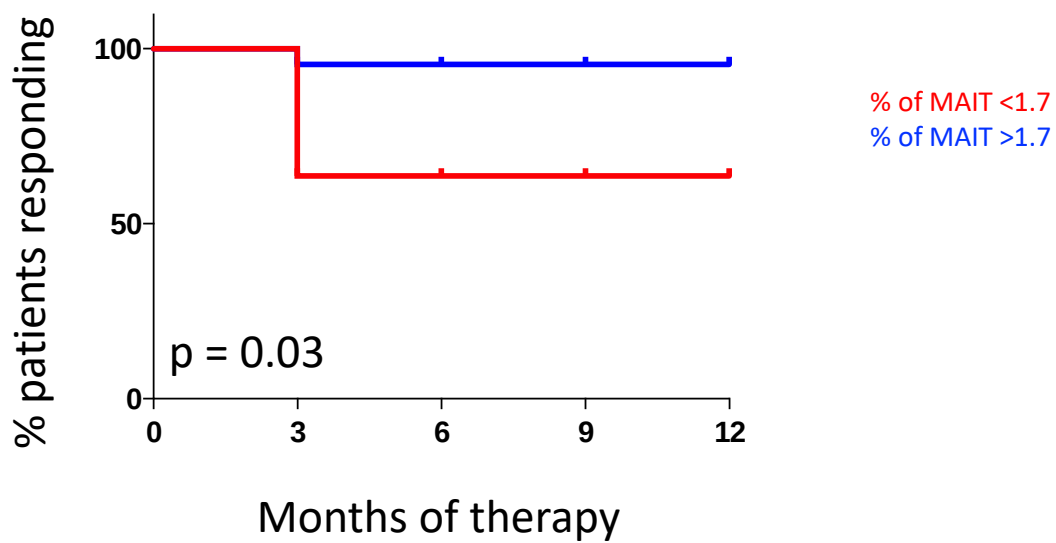


Figure 14. Level of MAIT cells before therapy can predict initial response to therapy.

Analysis of the cohort patients with metastatic melanoma indicates that patients with MAIT cells >1.7% of CD3+CD8+ T lymphocytes showed a better response to therapy compared to those with MAIT < 1.7% ($p = 0.0363$, Log-rank Mantel-Cox test). $N < 1.7\% = 4$; $N > 1.7\% = 8$.

7. RESULTS 2 (mRCC)

7.1 mRCC patients displayed a decrease of TSCM and increase of EM cell percentages after therapy

Changes in CD8⁺ T cell dynamics have been investigated by using 30-parameters flow cytometry and analyzed by Full Annotation Using Shape-constrained Trees (FAUST), a novel non-parametric method for unbiased discovery of cell population in single-cell flow cytometry [33]. Clustering was performed using differentiation and functional markers such as CD45RA, CD57, TBET, CD28, CD127, CD38, CCR7, GRANULYSIN, CD95, PD1 and CXCR6 (see Methods and **Figure 15** for gating strategy). Sixty-one cell clusters were identified that represent the entire differentiation spectrum of CD8⁺ T cells (**Figure 16A** and **Figure 17**). In particular, we identified: 4 clusters of naïve (N) T cell, 3 clusters of TSCM, 2 clusters of central memory T cells (CM), 12 clusters of transitional memory (TM) cells, 13 clusters of effector memory (EM) T cells and 27 clusters of terminally differentiated effector memory (EMRA) T cells (Figure 14A, Figure 15). Overall, the total of N clusters represents a mean value of 5.22% of CD8⁺ T cells (median $\pm$ SEM, 3.6 ± 0.66), TSCM clusters represents 1.76% (0.88 ± 0.26), CM cluster is 1.11% (0.64 ± 0.17), TM clusters are 5.47% (5.03 ± 0.18), EM clusters are 27.7% (24.18 ± 2.06), EMRA are 52.35% (53.55 ± 3.16) (**Figure 16B**), indicating that half of the CD8⁺ T cell population is represented by a terminally differentiated effector memory compartment.

First, statistical analyses were performed dividing patients on the basis of the response to therapy. Responder (R, n = 5) and non-responder (NR, n = 14) patients were compared for all the clusters identified, but due to low number of patients responding to therapy, we found similar percentages of clusters in R and NR (**Figure 18**) at all timepoints. However, considering all patients treated with anti-PD1 (n=19), we show that therapy was able to induce a redistribution of the different subpopulations of CD8⁺ T cells (**Figure 19** and **Figure 20**). Naïve compartment, identified as the cluster of cells that were CD45RA⁺CD57⁺TBET⁺CCR7⁺CD28⁺CD127⁺CD38⁺GRANULYSIN⁺CXCR6⁺CD95⁺PD1⁺IgG4⁻, increased after the third cycle of therapy. It is to note that, even before therapy, in these patients the percentage of naïve T cell was very low. This could be congruent with the relatively advanced age of mRCC patients (range 55-81 years) and with the fact that their immune system could have been challenged by ten years of disease (see **Table 2** for demographic and clinical characteristics of the patients included in the study). One of the underpinnings of T cell immune decline and cancer is that in the elderly there is a reduction in naïve T cells and in their diversity, leading to a dominant

memory T cell pool (81). After therapy, TSCM subset, identified as CD45RA⁺CD57⁻TBET⁻CD28⁺CD127⁺CD38⁻CCR7⁺GRANULYSIN⁻CXCR6⁻CD95⁺PD1⁻IgG4⁻, decreased by 2.5-fold passing from a mean value of 3.26%, measured before therapy, to 1.32%, measured after the third cycle of therapy. TSCM cells are long-living, highly active "super immune cells" with strong antitumor activity. Whereas TSCM cells are considered the most pluripotent, CM cells are capable of give origin to EM subsets and to a cluster of CM T cells (CD45RA⁻CD57⁻TBET⁻CCR7⁺CD28⁺CD127⁺CD38⁻GRANULYSIN⁻CXCR6⁻CD95⁻PD1⁻IgG4⁻) that increased after the first, the second and the third cycle of therapy.

On the contrary, the effector memory compartments increased after therapy. A total of 8 EM clusters were affected by therapy initiation. Discrimination between distinct stages of human CD8⁺ T-cell differentiation has often relied upon the cell surface expression of CD27 and CD28, costimulatory molecules of distinctive function. Interaction of CD28 with CD80 and CD86 on antigen-presenting cells amplifies T-cell proliferation and activation through a series of direct effects. T cell differentiation has been proposed wherein CD27⁺ CD28⁺ CD45RA⁺ naive cells progress through a CD27⁺ CD28⁺ CD45RA⁻ early antigen-experienced phenotype to a CD27⁺ CD28⁻ CD45RA^{-/+} intermediate phenotype and finally to a CD27⁻ CD28⁻ CD45RA^{+/+}-late antigen-experienced phenotype that parallels an increased cytotoxic potential and reduced ability to proliferate [185]. Here, a TM subset, identified as CD45RA⁻CD57⁻TBET⁻CCR7⁻CD28⁺CD127⁺CD38⁻GRANULYSIN⁻CXCR6⁻CD95⁻PD1⁻IgG4⁻, increased from a mean value of 0.21%, measured before therapy, to 1.15%, measured after the third cycle of therapy.

The expression of CD127, the alpha chain of IL-7 receptor, on cell membrane is very important for T cell homeostasis and is a useful marker for measuring long-living memory T cells. Most importantly, its expression allows the distinction between memory and effector T cells early after in vivo priming. EM subsets expressing or not CD127 (CD45RA⁻CD57⁻TBET⁻CCR7⁻CD28⁻CD127^{-/+}CD38⁻GRANULYSIN⁻CXCR6⁻CD95⁻PD1⁻IgG4⁻), increased after the third cycle of therapy. In addition, CD45RA⁻CD57⁻TBET⁻CCR7⁻CD28⁺CD127⁺CD38⁺GRANULYSIN⁻CXCR6⁻CD95⁺PD1⁻IgG4⁻ cells, identified as activated EM, increased after the third cycle of therapy.

Previous studies have demonstrated that both TBET and EOMES are involved in the generation of effector and central memory CD8⁺ T cells. TBET is an essential regulator of effector differentiation and function in multiple different immune lineages. Inflammatory cytokines induce

TBET expression, which in turn regulates distinct differentiation outcomes. For example, TBET can act as a fulcrum between Th1 and Tfh cell differentiation, pathogenic and non-pathogenic Th17 cells, and CD8⁺ effector and memory T cells. TBET induces CXCR3 expression on multiple different lymphocytes promoting an effective immune response by allowing the homing of different cells in the inflamed tissue [187]. EM clusters expressing TBET identified as CD45RA⁺CD57⁺TBET⁺CCR7⁺CD28⁺CD127⁺/CD38⁺GRANULYSIN⁺CXCR6⁺CD95⁺PD1⁺IgG4⁻, increased after therapy, too.

CD45RA⁺CCR7⁻ EMRA cells are found more frequently in the CD8⁺ compartment, remain CCR7⁻CD62L⁻, are generally negative for CD27 and CD28, and display the shortest telomeres among T cells. EMRA cells express markers of senescence, including KLRG-1, CD57, and phosphorylation of histone H2AX, and have low proliferative and functional capacity indicating terminal differentiation [180]. Moreover, CD8⁺ EMRA T cells display a senescence-associated secretory phenotype regulated by p38 MAPK [187]. Seven EMRA clusters were affected by anti-PD1 therapy. A cluster of EMRA expressing or not CD127, identified as CD45RA⁺CD57⁺TBET⁺CCR7⁺CD28⁺CD127⁺CD38⁺GRANULYSIN⁺CXCR6⁺CD95⁺PD1⁺IgG4⁻, increased after the third cycle of therapy as well as a cluster of EMRA cells expressing CD28 (CD45RA⁺CD57⁺TBET⁺CCR7⁺CD28⁺CD127⁺CD38⁺GRANULYSIN⁺CXCR6⁺CD95⁺PD1⁺IgG4⁻) and one expressing TBET (CD45RA⁺CD57⁺TBET⁺CCR7⁺CD28⁺CD127⁺CD38⁺GRANULYSIN⁺CXCR6⁺CD95⁺PD1⁺IgG4⁻). On the contrary, terminally differentiated clusters of EMRA expressing CD57, TBET and CD95 (CD45RA⁺CD57⁺TBET⁺CCR7⁺CD28⁺CD127⁺CD38⁺GRANULYSIN⁺CXCR6⁺CD95⁺PD1⁺IgG4⁻; CD45RA⁺CD57⁺TBET⁺CCR7⁺CD28⁺CD127⁺CD38⁺GRANULYSIN⁺CXCR6⁺CD95⁺PD1⁺IgG4⁻; CD45RA⁺CD57⁺TBET⁺CCR7⁺CD28⁺CD127⁺CD38⁺GRANULYSIN⁺CXCR6⁺CD95⁺PD1⁺IgG4⁻), decreased after therapy.

These results suggest that anti-PD-1 therapy can induce a redistribution of the different subpopulations of T cells, with a decrease of TSCM reservoir and an increase of EM cells able to mount a pro-inflammatory and anti-tumoral immune response.

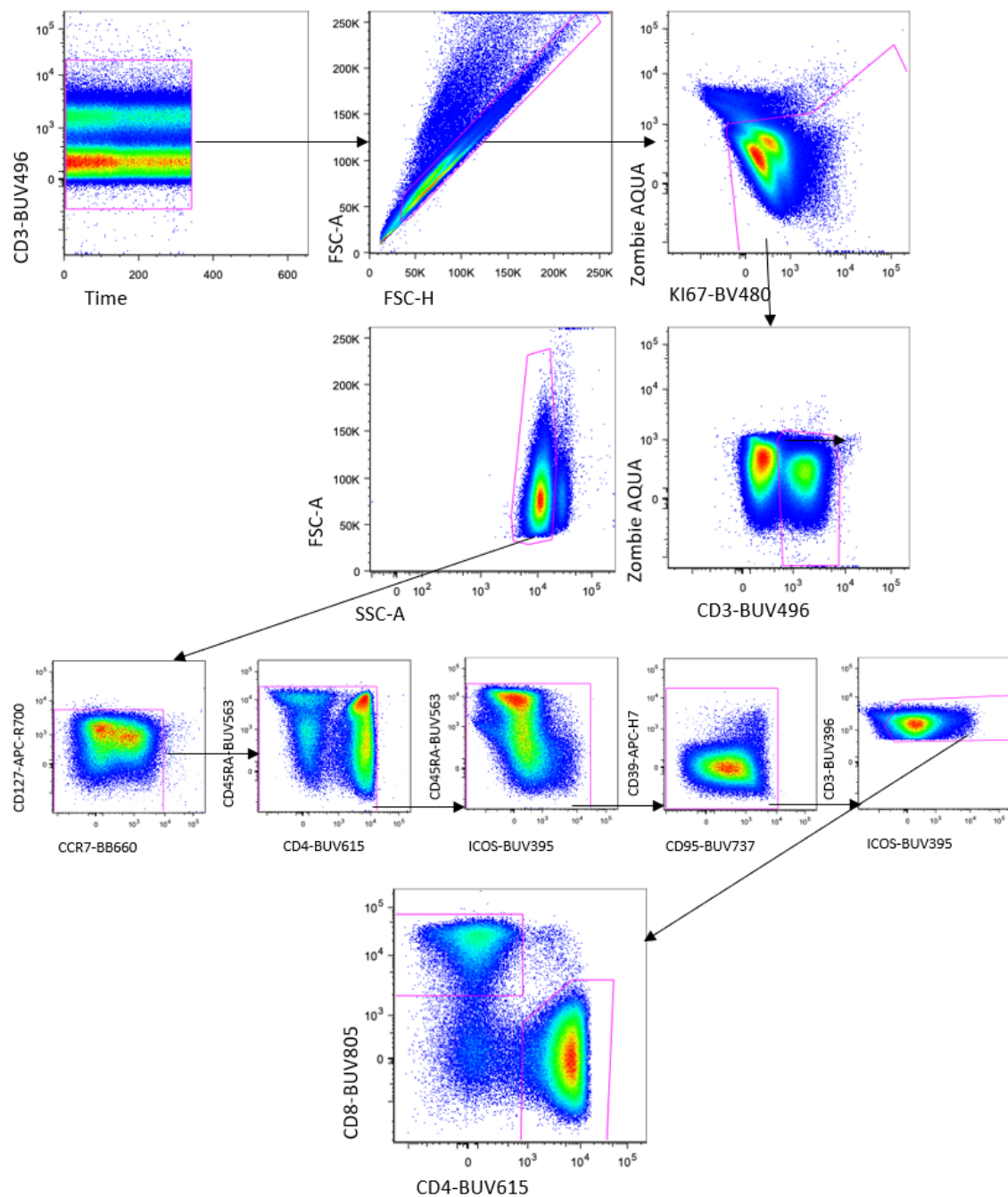


Figure 15.

Gating strategy used to identify CD8⁺ T cells among thawed PBMC of mRCC patients. Briefly, CD3⁺ T cells were identified, turbulences during acquisition were removed as well as doublets and dead cells (Zombie AQUA⁺ cells). In this population lymphocytes were selected according to physical parameters and fluorochrome aggregates were removed. Finally, CD4⁺ and CD8⁺ T cells were identified.

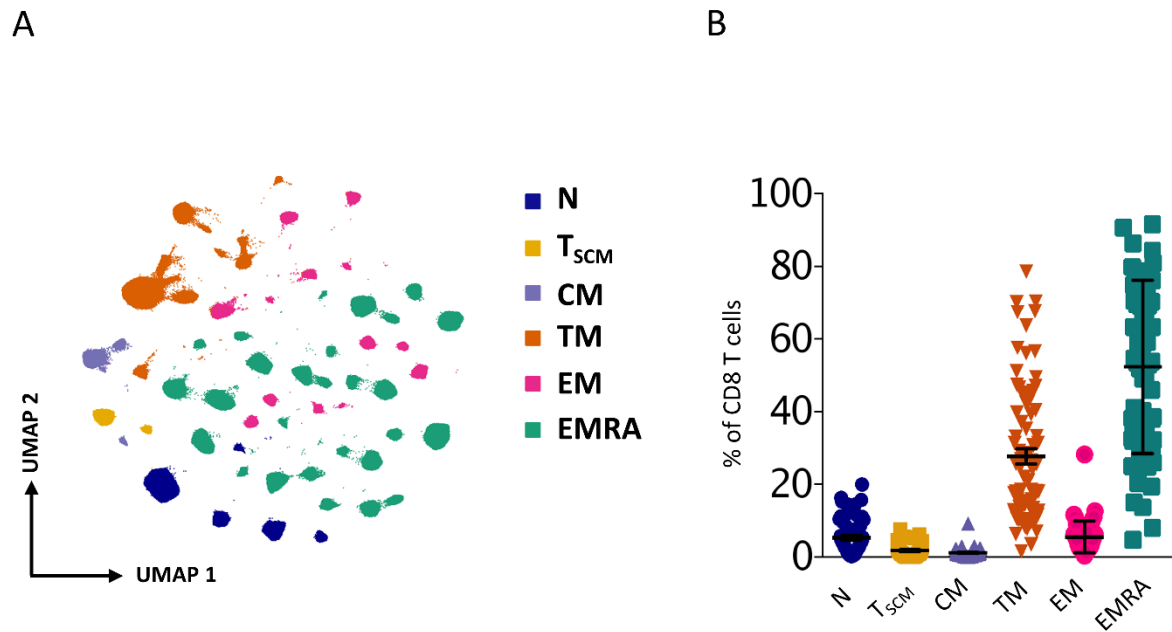


Figure 16.

High-dimensional single cell analysis of CD8+ T cells identifies higher percentage of memory T cells and lower percentage of TSCM after therapy. A) Uniform Manifold Approximation and Projection (UMAP) representation of the CD8+ T cell landscape. N= naïve; TSCM= T cell stem memory; CM= central memory; TM= transitional memory; EM= effector memory; EMRA= terminally differentiated effector memory. RCC patients, n = 19. B) Total percentages of different subpopulations of CD8+ T cells identified by FAUST. Data represent individual values, mean (centre bar) $\pm$ SEM (upper and lower bars). N= naïve; TSCM= T cell stem memory; CM= central memory; TM= transitional memory; EM= effector memory; EMRA= terminally differentiated effector memory.

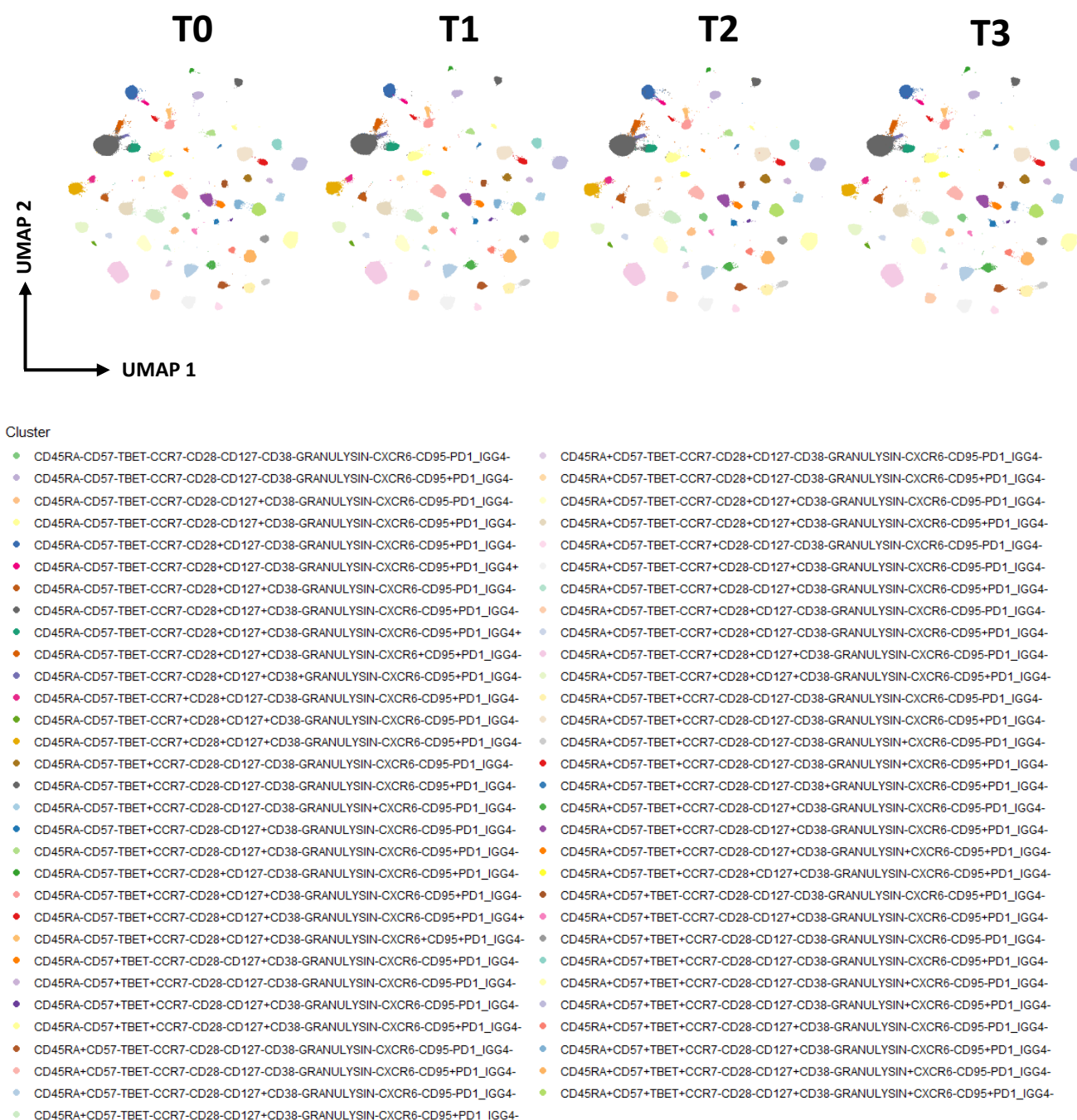
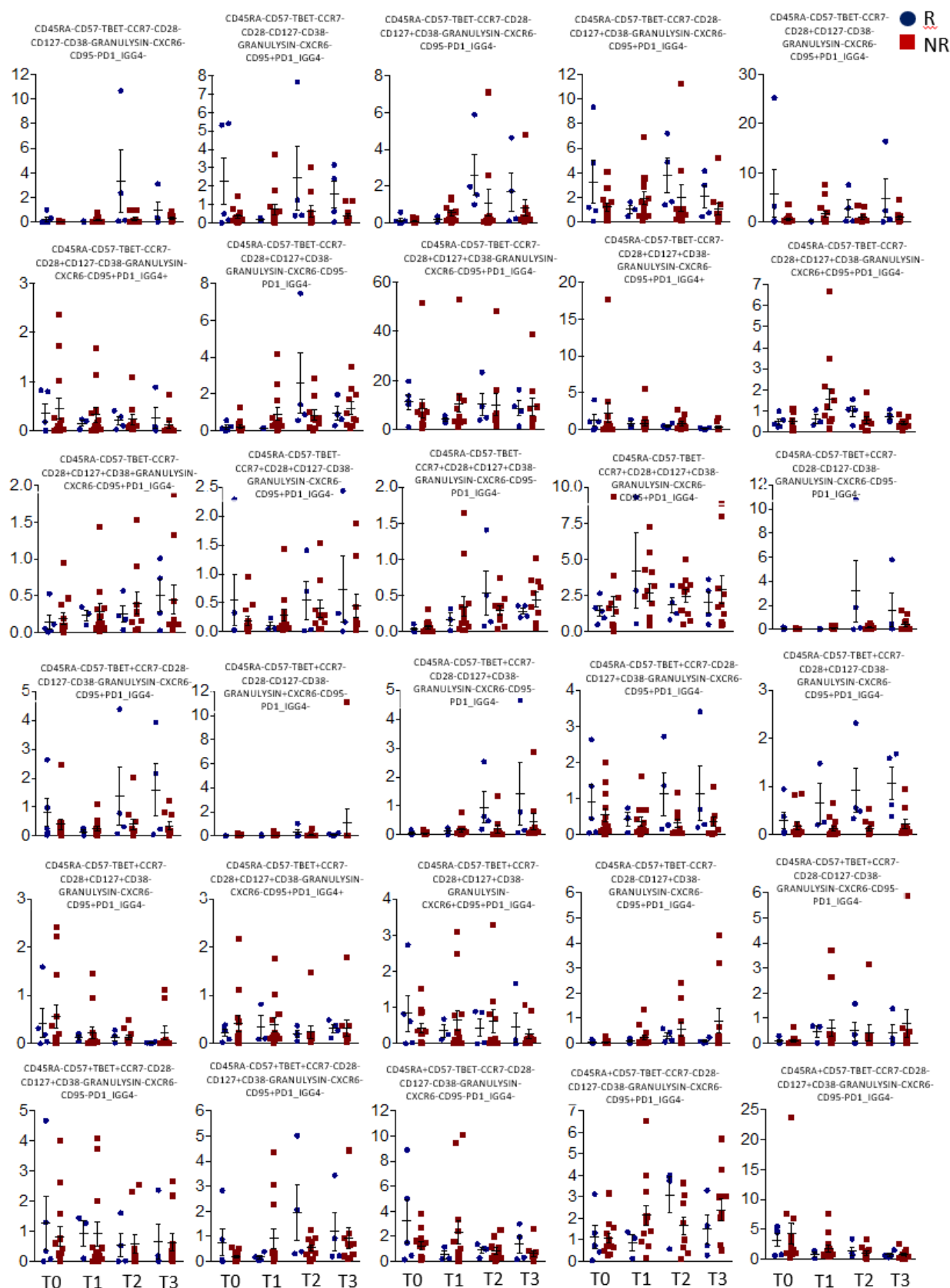
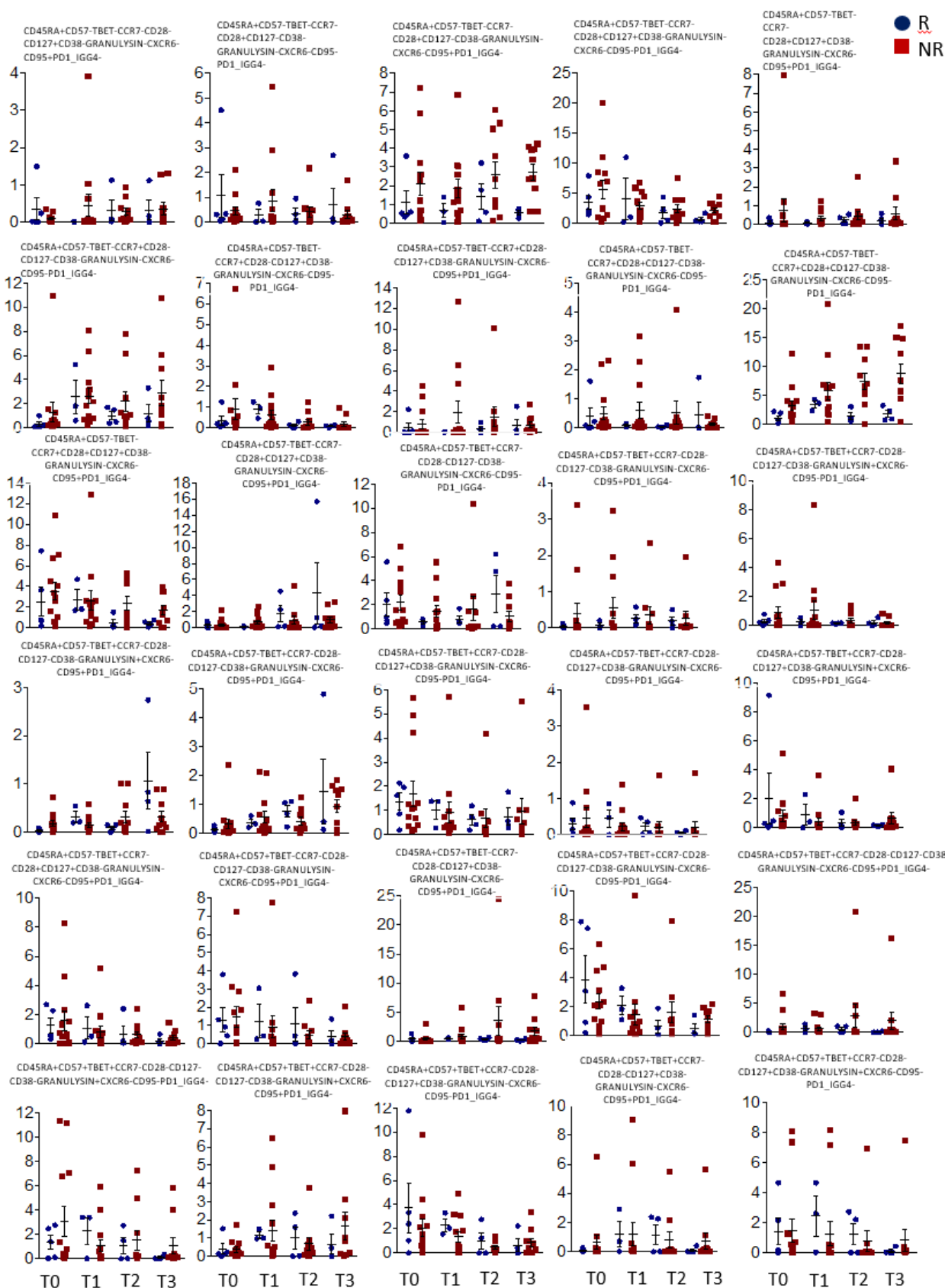


Figure 17. Uniform Manifold Approximation and Projection (UMAP) representations of CD8+ T cell subpopulations embedded with FAUST clusters split by therapy time point. Each color refers to as one of 61 clusters found accompanied by its complete phenotype.

% of CD8+ T cells



% of CD8+ T cells



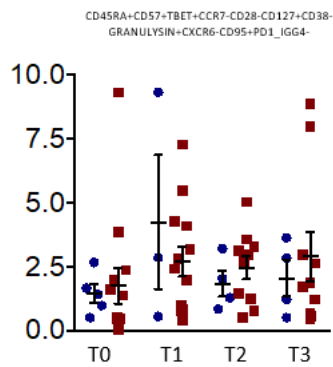


Figure 18. Percentages of different subpopulations (clusters) of CD8+ T cells identified by FAUST in mRCC patients treated with ICI therapy. Data represent individual values, mean (centre bar) $\pm$ SEM (upper and lower bars). T0= before therapy; T1= after 1 cycle of therapy; T2= after 2 cycles of therapy; T3= after three cycles of therapy. R= responder, NR= not responder Two-way anova with Bonferroni correction was used to test each sub-population for differential abundance.. p value: * <0.05 ; ** <0.001 . The name of each population is reported above the graph.

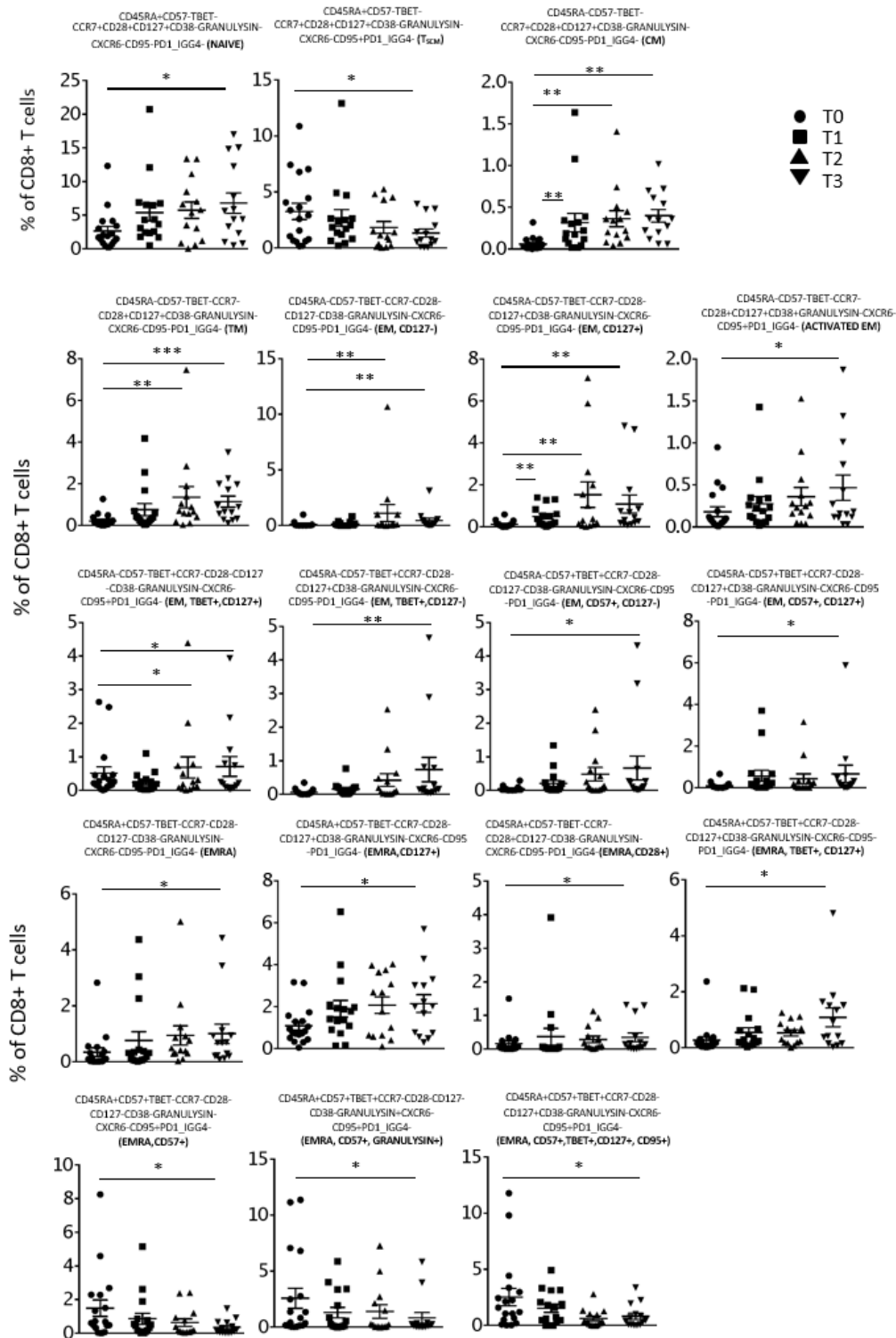
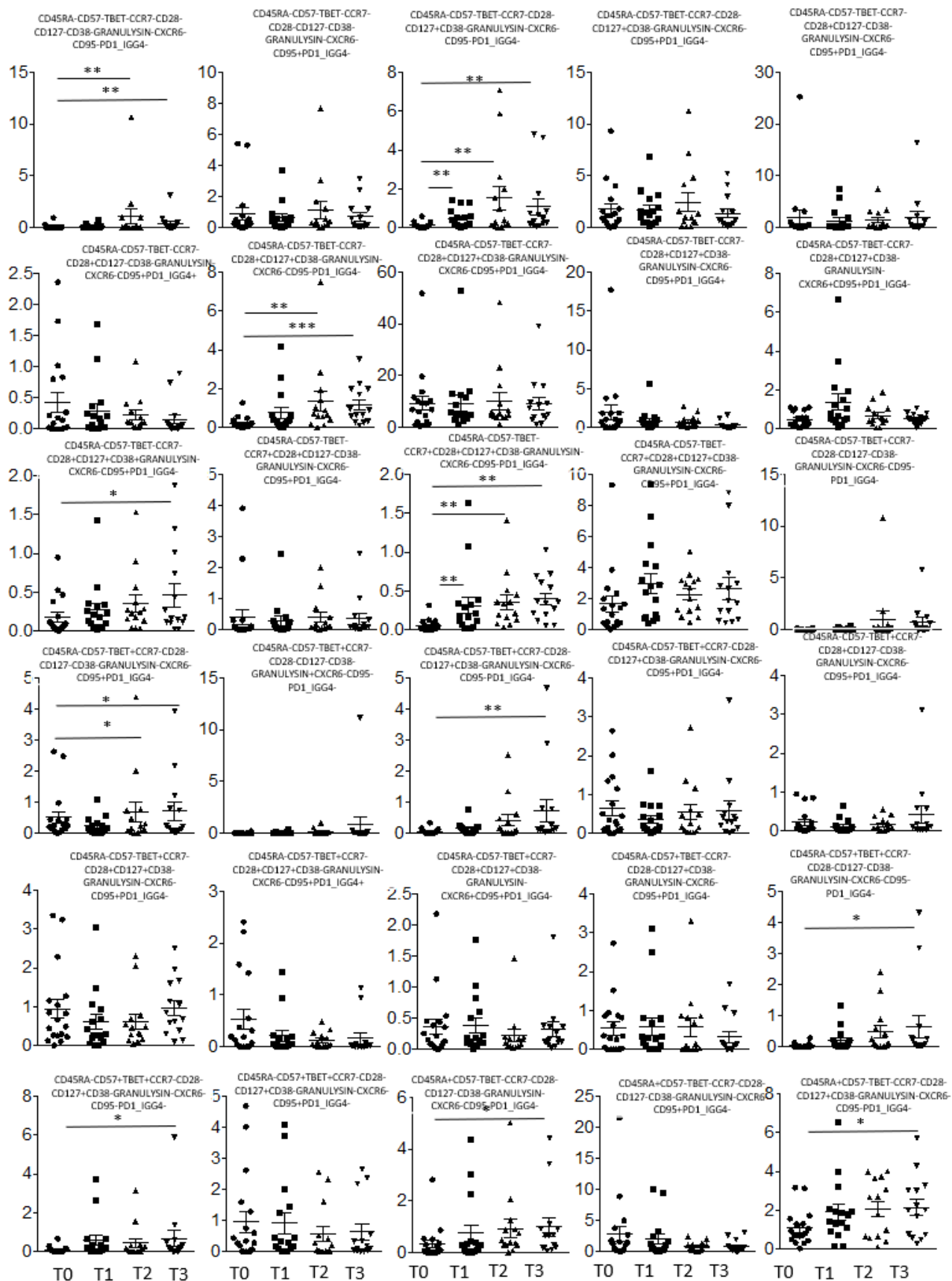
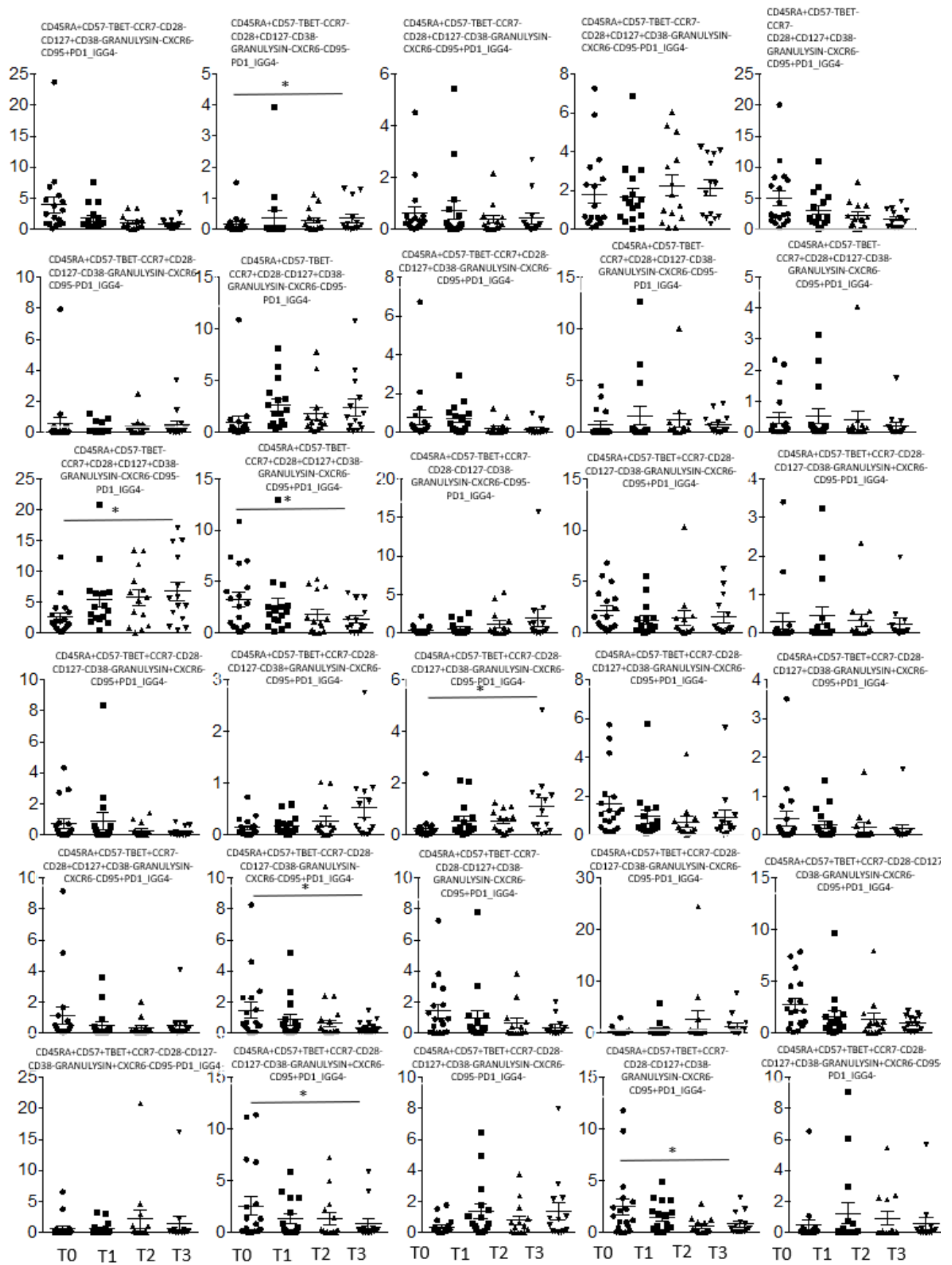


Figure 19. Percentages of different subpopulations of CD8+ T cells identified by FAUST (R and NR were analyzed together). Data represent individual values, mean (central bar) $\pm$ SEM (upper and lower bars). T0= before therapy; T1= after 1 cycle of therapy; T2= after 2 cycles of therapy; T3= after three cycles of therapy. Kruskal-Wallis test with Benjamini-Hochberg correction for multiple comparisons was used to test each sub-population for differential abundance. p value: * <0.05 ; ** <0.001 . The name of each population is reported above the graph.

% of CD8+ T cells



% of CD8+ T cells



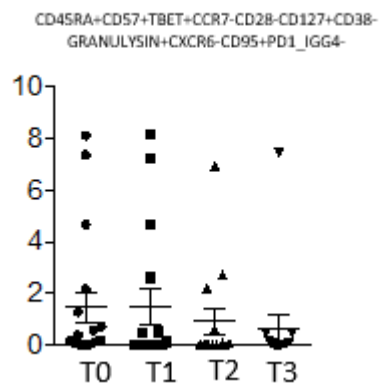


Figure 20. Percentages of different subpopulations of CD8+ T cells identified by FAUST. Data represent individual values, mean (central bar) $\pm$ SEM (upper and lower bars). T0= before therapy; T1= after 1 cycle of therapy; T2= after 2 cycles of therapy; T3= after three cycles of therapy. Kruskal-Wallis test with Benjamini-Hochberg correction for multiple comparisons was used to test each subpopulation for differential abundance. p value: * <0.05 ; ** <0.001 . The name of each population is reported above the graph.

7.2 ICI therapy induce strong CD8+ T cell activation in memory subset but not in more undifferentiated population

We focused our attention on the expression of the activation markers that were not included in the unsupervised clustering, such as BTLA, CD39, CD71, CD98, CD194, CD244, CXCR6, ICOS, PD1, Ki67 and HLA-DR, within N cells (identified as a total of CD45RA+CCR7+), TSCM (identified as CD45RA+CCR7+CD95+), CM (identified as a total of CD45RA-CCR7+), EM (identified as a total of CD45RA-CCR7-) and EMRA (identified as a total of CD45RA+CCR7-). Therapy did not induce changes in terms of expression of different activation markers except for Ki67 and HLA-DR. A population identified as activated and proliferating cells (Ki67+HLA-DR+) increased after therapy within the CM and EM subsets but not in more undifferentiated population such as N and TSCM subsets (**Figure 21**), thus suggesting TSCM pool is unaffected by anti-PD1 therapy and is rather consumed to give origin to EM compartments.

7.3 Cytokine production is increased in more undifferentiated cells after therapy

To ascertain the capability of therapy to reactivate the pro-inflammatory function of CD8+ T cells, PBMC were in vitro stimulated by using anti-CD3 plus anti-CD28. Then, the percentages of CD8+ T cell subpopulations producing one or more cytokines, such as IFN- γ and TNF together with IL-2, were quantified. Overall, similar percentages of CD8+ T cells producing different cytokines were found after therapy, but, looking at the different subpopulations of T cells, we noted that TSCM as well as CM producing IFN- γ and TNF increased after the second cycle of therapy. In terms of polyfunctionality, we found higher percentages of TSCM able to produce simultaneously TNF and IFN- γ . No differences were found in terms of cells producing IL-2. This suggests that the pools of more undifferentiated cells are able to exert pro-inflammatory activity after in vitro stimulation (**Figure 22**). No differences in terms percentages of cells producing IFN- γ and TNF after therapy were found among EM and EMRA (not shown).

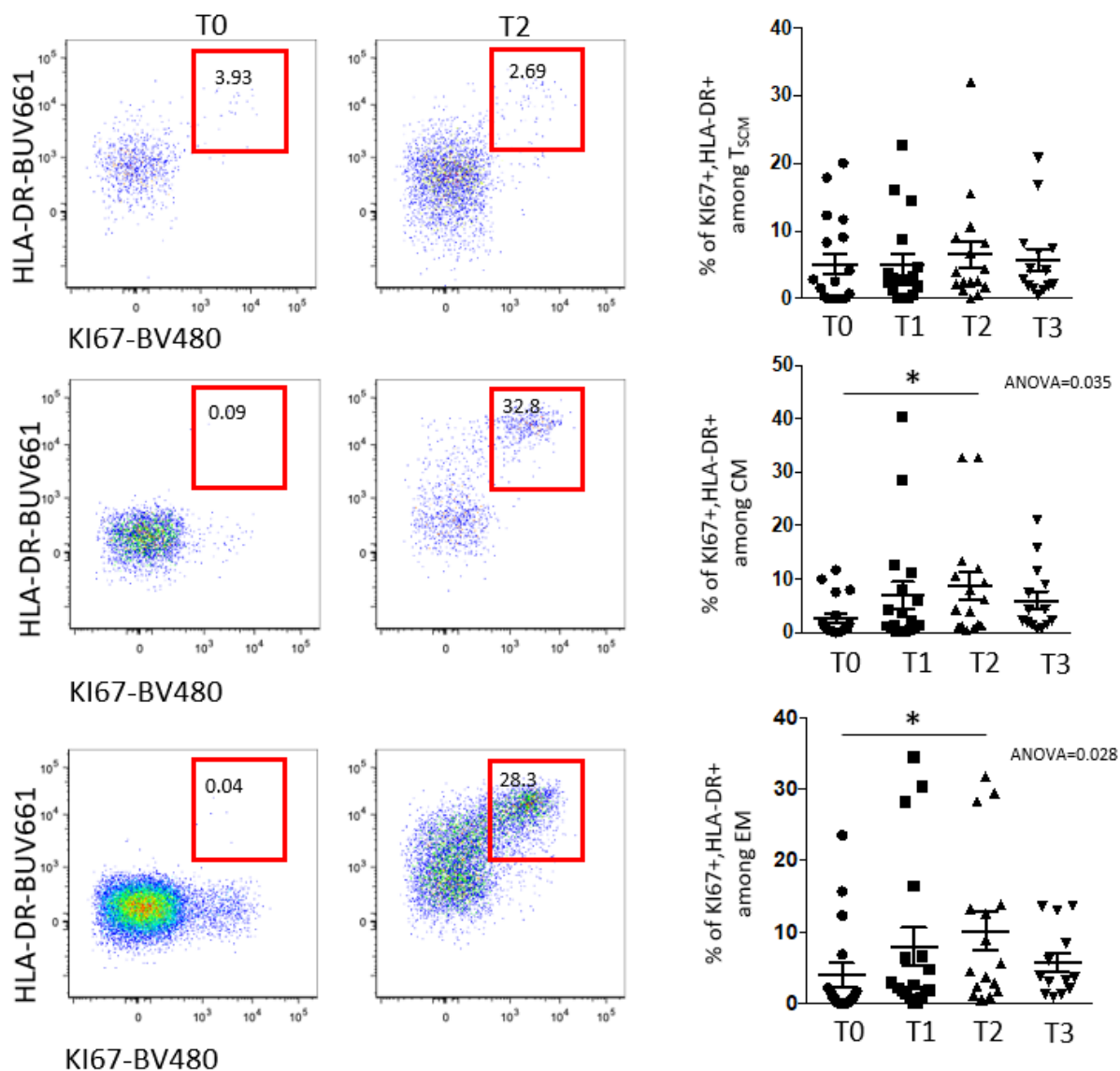
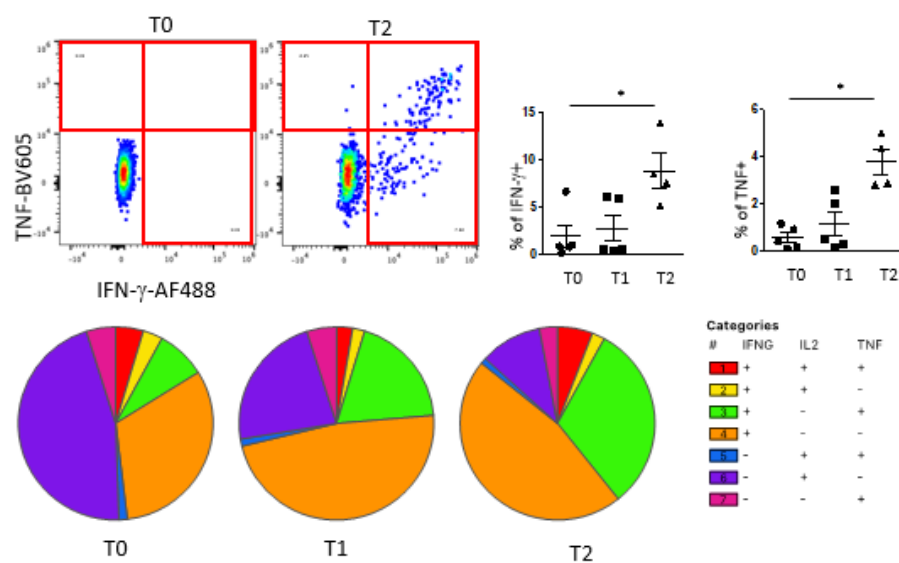


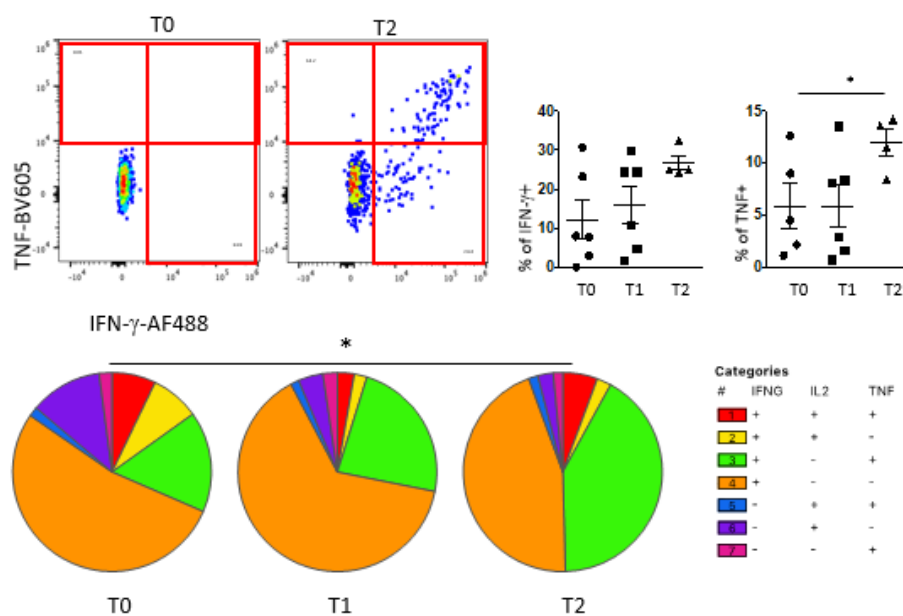
Figure 21.

T cell activation evaluated during ICI therapy. Representative dot plots of KI67+ and HLA-DR+ activated T cells before therapy (T0) and after two cycles of therapy (T2); Individual values of activated cells present in TSCM (upper row), CM (middle row) and EM (lower row). Data represent individual values, mean (center bar) $\pm$ SEM (upper and lower bars). Kruskal-Wallis one-way Anova was applied; exact p-value is reported in the figure. TSCM= T cell stem memory; CM= central memory; EM= effector memory; T0= before therapy; T1= after 1 cycle of therapy; T2= after 2 cycles of therapy; T3= after three cycles of therapy. RCC patients, n = 19.

A



B



C

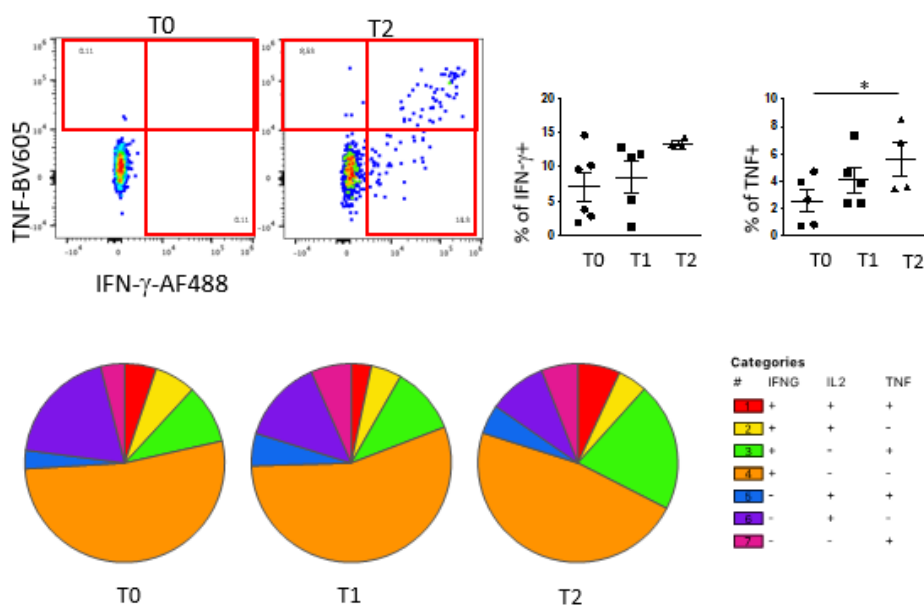


Figure 21.

CD8⁺ T cell subpopulation cytokine production evaluated after in vitro stimulation. **A)** Representative dot plots of TSCM producing TNF and IFN- γ after in vitro stimulation before (T0) and after two cycles of therapy (T2). Percentage of cells producing IFN- γ and TNF is also reported. Data represent individual values, mean (center bar) $\pm$ SEM (upper and lower bars). Kruskal-Wallis one-way Anova was used for statistical analysis (Upper row). Pie charts represent polyfunctional response of TSCM after in vitro stimulation before and after therapy. Permutation test has been used (lower row). RCC patients, n=6. **B)** Representative dot plots of TCM producing TNF and IFN- γ after in vitro stimulation before (T0) and after two cycles of therapy (T2). Percentage of cells producing IFN- γ and TNF is also reported. Data represent individual values, mean (center bar) $\pm$ SEM (upper and lower bars). Kruskal-Wallis one-way Anova was used for statistical analysis (Upper row). Pie charts represent polyfunctional response of TCM after in vitro stimulation before and after therapy. Permutation test has been used (lower row). RCC patients, n=6. **C)** Representative dot plots of EM producing TNF and IFN- γ after in vitro stimulation before (T0) and after two cycles of therapy (T2). Percentage of cells producing IFN- γ and TNF is also reported. Data represent individual values, mean (center bar) $\pm$ SEM (upper and lower bars). Kruskal-Wallis one-way Anova was used for statistical analysis (Upper row). Pie charts represent polyfunctional response of EM after in vitro stimulation before and after therapy. Permutation test has been used (lower row). RCC patients, n=6.

8. DISCUSSION 1 (melanoma)

The main finding of our study is that patients who respond to ICI are characterized by a different composition of T-cell subpopulations compared to those who do not respond, that are detectable before therapy initiation. The most relevant of these differences is at the level of MAIT cells, an innate population of CD3⁺ T cells previously involved in early immunity against infections in peripheral tissue. Although the direct role of MAIT in mediating anti-tumor immune responses in melanoma is still under scrutiny, our data suggest that investigating MAIT cell frequency in the peripheral blood could be considered a possible predictive marker of successful therapy. Following introduction of ICI, R show different proportions of T cells compared to NR, involving the expansion of activated effector memory cells showing features of immune activation, proliferation and effector differentiation, as previously reported by other groups [19].

During the last decade, the immune response mediated by T cells in cancer patients assuming ICI has been deeply investigated by analyzing both tumor-infiltrating lymphocytes and circulating T cells. Patients with melanoma or non-small cell lung cancer are characterized by an exhausted T-cell phenotype along with impaired proliferation and low metabolic activation, and a high oligoclonal repertoire [41,187-188]. Activation of CD8⁺ T cells has been considered a hallmark of response to therapy, and indeed after one cycle of therapy, Ki67 (a marker of cell proliferation) increases among effector memory cells [19,189].

We show here that even if before treatment R and NR were characterized by similar clinical characteristic in terms of tumor burden and LDH level, activated effector memory T cells were more abundant in R, which can reflect a more activated CD8⁺ T-cell compartment. This was particularly evident in MAIT cells. Circulating MAIT cells are a pro-inflammatory and cytotoxic population within effector memory T cells [190] and can represent up to 10% of peripheral CD8⁺ T cells. They recognize microbial proteins presented by non-polymorphic MHC class I related-molecule (MR1) and display homing properties, as they express different homing and cytokine receptors. Furthermore, MAIT cells are deeply involved in patrolling mucosae and orchestrating the immune response in this environment [77,191].

The role of MAIT cells in cancer has been widely analyzed. However, few studies have investigated their role during therapy with ICI. It was found that MAIT cells were decreased in blood and displayed an altered cytokine production in patients with cervical, colorectal, gastric, hepatocellular carcinoma, lung cancer, and multiple myeloma. Moreover, controversial data exist

on the prognostic benefit of MAIT cells in the tumor microenvironment, as in the case of hepatocellular carcinoma [77]. Recent studies also show that MAIT cells promote tumor initiation, growth and metastases via tumor MR1 [91].

To the best of our knowledge, these are the first data that characterize MAIT cells in the peripheral blood of patients treated with anti-PD-1. We found that in R compared to NR, at baseline and after therapy introduction, (i) the proportion of MAIT cells was higher; (ii) MAIT cells displayed enhanced expression of genes related to immune activation and effector functions; (iii) the proportion of MAIT cells expressing CXCR4 was higher in R after two cycles of therapy.

CXCR4-CXCL12 axis plays an important role in the interactions between cancer cells and their microenvironment. This axis modulates the traffic of tumor cells to metastasis, and mediates invasiveness, vasculogenesis, and angiogenesis. However, pre-clinical melanoma models reported that this pathway can be influenced by anti-cancer treatments [192]. Hence, it is possible to hypothesize that, among other activities, the increased expression of CXCR4 on MAIT cells induced by anti-PD-1 therapy could facilitate their migration towards metastases, where they could exert a pro-inflammatory and cytotoxic activity. In line with this hypothesis, we observed that MAIT cells from R expressed CD69, which is not only an activation marker, but also a constitutively expressed marker of tissue residency. In immunotherapy-naïve melanoma patients, the intratumoral presence of CD8+CD103+CD69+ T cells that are able to significantly increase during anti-PD-1 therapy has been associated with improved survival [193]. Accordingly, we have found an increased relative proportion of MAIT cells in melanoma lesions responding to ICI compared to those that did not respond, possibly suggesting the migration of circulating MAIT cells to the tumor site, as it has been observed for conventional CD8+ T cells [19]. Future studies investigating the clonal composition of MAIT cells from the blood and the tumor will be required to confirm this hypothesis.

Recently, a rare population of MR1-restricted T cells belonging to the family of MAIT cells (defined “MR1” T cells) has been described as a rare population capable to respond to a variety of tumor cells of different tissue origin, but not to microbial antigens [194,195]. Thanks to its ability to kill several cancer cell lines expressing low levels of MR1 while remaining inert towards noncancerous cells, this population represents a subset with a great potential for cell therapy approaches in several malignancies [92].

While we acknowledge some limitations associated with this study, such as the relatively low number of patients enrolled in the study and the lack of a possible mechanism responsible of a better prognosis, nevertheless we provide evidence of the association between the frequency and

the effector functions of MAIT cells and the response to ICI in melanoma, thereby suggesting that the circulating levels of MAIT cells in the peripheral blood could serve as a useful, non-invasive biomarker. Thus, further studies will be needed not only to confirm the utility of MAIT as biomarkers, but also to demonstrate their therapeutic potential or to provide actionable information about tumor's biology, which together holds great promise with respect to realizing "personalized" treatment of melanoma. In conclusion, we provide evidence of the association between the frequency and the effector functions of MAIT cells and the response to ICI in melanoma, thereby suggesting that the circulating levels of MAIT cells in the peripheral blood could serve as a useful, non-invasive biomarker. Future studies are also needed to assess whether MAIT cells are directly involved in mediating tumor regression that can be further amplified by targeting PD-1 or alternate immune checkpoints.

9. DISCUSSION 2 (mRCC)

Despite the impressive clinical improvements achieved with immunotherapy in mRCC management, many patients eventually fail to respond to this therapy due to the onset of primary or secondary resistance. Clinical definition of tumor resistance to immunotherapy is clearly established, defining primary resistance as the clinical scenario where cancer does not respond to an immunotherapy strategy, and secondary resistance as the clinical condition in which cancer initially responded to immunotherapy, but after a period of time, it relapsed and progressed [196]. Currently, a crucial unmet need is to recognize primary refractory mRCC early, and thus every attempt to clarify the mechanism of resistance is necessary. The purpose of this study was to characterize the complexity of resistance pathways in a group of these patients. The early identification of primary resistant patients represents a significant challenge in mRCC management to optimize treatment strategies.

Our data suggest that anti-PD1 therapy induces a redistribution of peripheral CD8+ T cell subpopulations in mRCC patients. The most relevant difference involves the decrease of TSCM pool and the increase of the effector memory compartment, spanning from transitional memory cells to terminally differentiated effector memory elements. Indeed, TSCM pool is not reactivated by therapy, although it increases its functionality in terms of the production of multiple pro-inflammatory cytokines. On the contrary, memory subset is reactivated after therapy and increases in terms of percentage during the treatment, while the number of cells able to produce cytokines is not affected by therapy. On the whole, these results outline how in primary refractory patients, an anti-PD1 treatment is active but not effective, resulting in the early progression of disease with almost no clinical benefit.

Clearly, clinical definition of efficacy outcomes may represent a major limitation in biomarker studies. Indeed, the conventional RECIST response criteria (that include tumor shrinkage or progression-free survival), the actual main tool to define formally the best outcome from ICI treatment, likely does not fully represent the benefits and changes of the immune response. Thus, for mRCC patients, increasing efforts in biomarker studies are required because solid biological predictors of clinical response, likely based upon dynamic tracking of peripheral immune cells, are still lacking. Our results, for the first time, show the discrepancy between treatment activity and lacking response in primary refractory mRCC patients.

During the last decade, the immune response mediated by T cells in cancer patients treated with ICI has been deeply investigated by analyzing both tumor-infiltrating lymphocytes and circulating T cells. Recently, it has been shown that CD8⁺ and CD4⁺ tumor-infiltrating lymphocytes expressing PD-1 and TIM-3 were significantly upregulated in the patients with RCC of grade 3 or 4 and that RCC tumors display significant expansion of double-positive (DP) CD4⁺CD8⁺T cells expressing exhaustion markers [197,198]. Moreover, high CD73 expression in the tumor microenvironment portends significantly worse survival outcomes independently of stage and grade (202), while the number of CD8⁺ and FoxP3⁺ T lymphocytes in adjacent normal and intratumoral sites in RCC may serve as a good predictive marker for prolonged free survival [199].

Few data concerning PBMC investigation in RCC patients treated with different therapies have been published. Main findings regard CD3⁺CD8⁺CD137⁺, and CD3⁺CD137⁺PD1⁺ T cell populations, that were associated with clinical responses of RCC patients treated with tyrosine kinase inhibitors and ICI (204, 205). Other data indicate that patients with progressive disease who were treated with anti-PD1 showed an increase in the frequency of CD4⁺ and CD8⁺ T cells expressing TIM-3 [200].

TSCM cells are a rare subset of memory lymphocytes with the stem cell-like ability to self-renewal and the multipotent capacity to reconstitute the entire spectrum of memory and effector T cell subsets [179,200]. In adoptive immunotherapy of cancer, these cells are characterized by an enhanced antitumor activity if compared with CM and EM cells. Here we show for the first time a complex redistribution of circulating T cells after anti-PD1 therapy and, in particular, we indicate that TSCM pool diminishes after therapy. In a humanized mouse model of mesothelioma, the transfer of mesothelin-specific TSCM, CM, or EM revealed that while EM and CM were associated with tumor progression, TSCM triggered tumor regression and eradication [16]. A similar effect has been described in mouse models with established B16 melanoma tumors [201]. Moreover, it has been shown that superior TSCM persistence supports long-lived T cell memory [16,202]. Indeed, two subsets of CD8⁺ TSCM progenitors with distinct fate commitments have been recently discovered [105]. One of these gives rise to a functional pool of cells, while the other is committed to a dysfunctional, exhausted-like lineage [105].

TSCM in mRCC has been less studied so far, but this population has recently attracted attention in this setting, and a first study reported similar percentages between healthy donors and mRCC patients [203]. Here, we found that CD8⁺ TSCM decreased after anti-PD-1 therapy. Other studies showed that intratumoral Tcf1⁺PD-1⁺CD8⁺ T cells have stem-like properties and display

features of exhausted cells, but they can persist and sustain the tumor immune response [56]. Previous data obtained on patients with stage III/IV melanoma and treated with ipilimumab revealed that CD4⁺ TSCM percentage gradually decreased after therapy despite sustained proliferation, suggesting an increase in cell differentiation [204]. This could be consistent with the increase of EM that we found after the third cycle of therapy. Thus, one could hypothesize that the pool of TSCM is consumed to give rise to a population of EM that still maintains the capability to produce pro-inflammatory cytokines.

We are well aware that this study has several limitations. The first is represented by the relatively low number of patients enrolled in the study and by their clinical characteristics, as all of them were treated with anti-PD1 in the second or third line of therapy, and most patients are not responding to therapy. Second, due to ethical and clinical restraints, sampling of tumor biopsies from these patients was not possible. This limits the possibilities for intratumoral immunophenotyping and mechanistic investigations. Third, the number of cells that we could obtain did not allow further functional investigation, and all the metabolic aspects of the cells of interest still await detailed analysis.

10. GENERAL DISCUSSION AND CONCLUSIONS

Cancer is a genetic disease characterized by the uncontrolled proliferation of the body's cells and their spreading into other tissue giving rise to metastasis [205]. Under physiological conditions, the body's cells grow and proliferate to replace the older or damaged cells. Abnormal or damaged cells may proliferate and form tumors when this process breaks down. The better life condition and newest health care treatments have improved the human lifespan dramatically during the last century to more than eighty years to an expected increase of 4.4 years by 2040. Accordingly, the number of new cases and the incidence of most cancers increase with age. In 2020, an estimated 19.3 million new cancer cases will be diagnosed worldwide, of which 377,000 in Italy, and almost 10.0 million will die from the disease [206]. Moreover, by 2040, the number of new cancer cases per year is expected to rise to 28.4 million, 610,000 in Italy, and the number of cancer-related deaths to 16.4 million. The overall economic impact of cancer is enormous; premature death and disability from cancer worldwide impacted \$895 billion in 2008. These high costs from cancer could be mitigated by targeted, cost-effective, and more efficient therapies.

Cancer-related mortality has decreased by 1.8% per year for men and 1.4% for women during the last ten years. This observation may reflect the recent improvement and availability of new cancer treatment, also known as immune checkpoint inhibitor (ICI), acting to reinvigorate the immune system. These therapies showed promising results in melanoma and non-small cell lung carcinoma (NSCLC) patients displaying an objective response rate (ORR) of 40% and 45%, respectively [207,208]. Despite the success of anti-CTLA-4 and anti-PD-1/PD-L1 therapies, only a fraction of patients benefit from ICIs. This evidences now opens a new question of why some patients do not respond to ICI therapy and how we may improve the success of immunotherapy or identify patients who are unresponsive to ICI therapy.

Accordingly, in the first part of this study, an in-depth characterization of the frequency, phenotype, and function of circulating PBMC in metastatic melanoma patients treated with anti-PD-1 in the first line has been performed. As previously reported by other groups, R shows different proportions of T cells after the ICI treatment compared to NR, involving the expansion of activated effector memory cells, proliferation, and more effector characteristics. Furthermore, the more important finding was identifying differences in circulating MAIT cells between R and NR.

We stringently identified MAIT cells using flow cytometry and scRNA-seq and assessed their increase in patients who respond to ICI before therapy initiation. We also found that after two cycles of ICI therapy MAIT cells from R patients displayed enhanced expression of CXCR4, at gene and protein levels, suggesting their increased ability to migrate in inflamed tissue (as tumor site). Moreover, MAIT cells of R patients expressed a higher level of genes related to immune activation and effector functions. Thus, we speculate that activated MAIT can migrate in the tumor, improving ICI response. We performed two different in silico analyses using scRNA-seq data obtained from CD8+ TIL of melanoma patients to confirm our idea. In both datasets, we observed infiltrating MAIT cells, which were higher in R after ICI therapy. We correlated the median level (1.7% of CD8+ T cells) of circulating MAIT cells with the response to therapy. Using this value as a threshold, MM patients who exhibited MAIT frequency above the threshold showed a better response to therapy.

Overall, the most important findings of the first part of our study are related to the ability of ICI therapy to positively modulate the immune system and the ability of MAIT level before therapy to predict the initial response to ICI.

In the second part of this study, we investigated the effect of ICI therapy in a cohort of mRCC patients previously treated with other anticancer drugs. Despite the impressive clinical improvements achieved with immunotherapy in mRCC management, many patients eventually fail to respond to this therapy due to the onset of primary or secondary resistance. We performed an in-depth flow cytometry investigation of phenotypic and functional features of peripheral CD8+ T cells during ICI therapy. The data analysis was performed using FAUST, a novel algorithm, high sensible to identify rare populations across samples.

We did not observe change in the frequencies of the 61 identified clusters between R and NR to ICI. Unfortunately, this may be due to the low number of R patients observed in our cohort. Despite that, analyzing all together the mRCC patients we observed that anti-PD-1 therapy induces a redistribution of peripheral CD8+ T cell subpopulations. We found that mRCC displayed a decrease in TSCM pool and the increase of effector memory compartment, spanning from transitional memory cells to terminally differentiated effector memory elements. TSCM although they its functionality in terms of production of pro-inflammatory cytokines if compared to other subsets, were not reactivated by therapy. On the contrary, the memory subset is reactivated after therapy in terms of HLA-DR and Ki67 expression, while the amount of cells able to produce cytokines is not affected by therapy. TSCM frequency is relatively stable in blood despite the absolute number of

CD8⁺ TSCM decreased with age. Thus, we speculate that the reduction of TSCM, following the ICI therapy, could be related to patients' reduced ability to maintain tumor-specific long-term memory cells. Their reduction could also reflect the differentiation of TSCM cells into EM that still maintains the capability to produce pro-inflammatory cytokines, not only the failure to maintain self-renewal.

Overall, the second part of our study suggests that redistribution of CD8⁺ T cells after ICI therapy in second and third line in mRCC patients exists, and this could help to identify new possible therapeutic strategies able to modulate the TSCM reduction over time. PBMC and TSCM in mRCC have been less studied so far, for this reason further studies are needed to understand better their role to improve future tumor-specific treatment.

In conclusion, we provide evidence that MM and mRCC patients differently respond to anti-PD1 therapy, the MM are characterized by more activated MAIT in R patients, while the mRCC are characterized by a pool of circulating likely-exhausted T_{SCM} in both R and NR.

11. REFERENCES

- 1) McGranahan, Nicholas, et al. "Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade." *Science* 351.6280 (2016): 1463-1469.
- 2) Boshoff, Chris, and Robin Weiss. "AIDS-related malignancies." *Nature Reviews Cancer* 2.5 (2002): 373-382.
- 3) Buell, Joseph F., et al. "Malignancy after transplantation." *Transplantation* 80.2S (2005): S254-S264.
- 4) Virchow, Rudolf. "Krankhaften Geschwulste." *Hirschwald, Berlin* 1 (1863): 443.
- 5) Gadalla, Ramy, et al. "Validation of CyTOF against flow cytometry for immunological studies and monitoring of human cancer clinical trials." *Frontiers in oncology* 9 (2019): 415.
- 6) Spitzer, Matthew H., et al. "Mass cytometry: single cells, many features." *Cell* 165.4 (2016): 780-791.
- 7) Robinson, J. Paul. "Comparative overview of flow and image cytometry." *Current protocols in cytometry* 31.1 (2005): 12-1.
- 8) Mellman, Ira, et al. "Cancer immunotherapy comes of age." *Nature* 480.7378 (2011): 480-489.
- 9) Trombetta, E. Sergio, et al. "Cell biology of antigen processing in vitro and in vivo." *Annu. Rev. Immunol.* 23 (2005): 975-1028.
- 10) Vigneron, Nathalie. "Human tumor antigens and cancer immunotherapy." *BioMed research international* 2015 (2015).
- 11) Sade-Feldman, Moshe, et al. "Defining T cell states associated with response to checkpoint immunotherapy in melanoma." *Cell* 175.4 (2018): 998-1013.
- 12) Schreiber, Robert D., et al. "Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion." *Science* 331.6024 (2011): 1565-1570.
- 13) Von Andrian, Ulrich H., et al. "T-cell function and migration—two sides of the same coin." *New England Journal of Medicine* 343.14 (2000): 1020-1034.
- 14) Hamann, Dörte, et al. "Phenotypic and functional separation of memory and effector human CD8+ T cells." *Journal of Experimental Medicine* 186.9 (1997): 1407-1418.
- 15) Sallusto, Federica, et al. "Two subsets of memory T lymphocytes with distinct homing potentials and effector functions." *Nature* 401.6754 (1999): 708-712.
- 16) Gattinoni, Luca, et al. "A human memory T cell subset with stem cell-like properties." *Nature medicine* 17.10 (2011): 1290-1297.
- 17) Fairfax, Benjamin P., et al. "Peripheral CD8+ T cell characteristics associated with durable responses to immune checkpoint blockade in patients with metastatic melanoma." *Nature medicine* 26.2 (2020): 193-199.
- 18) Simon, Sylvain, et al. "PD-1 and TIGIT co-expression identifies a circulating CD8 T cell population predictive of response to anti-PD-1 therapy in melanoma and Merkel-cell carcinoma patients." (2020): 4476-4476.
- 19) Huang, Alexander C., et al. "T-cell invigoration to tumour burden ratio associated with anti-PD-1 response." *Nature* 545.7652 (2017): 60-65.

- 20) Yost, Kathryn E., et al. "Clonal replacement of tumor-specific T cells following PD-1 blockade." *Nature medicine* 25.8 (2019): 1251-1259.
- 21) Nikolich-Zugich, Janko, et al. "The many important facets of T-cell repertoire diversity." *Nature Reviews Immunology* 4.2 (2004): 123-132.
- 22) De Simone, Marco, et al. "Single cell T cell receptor sequencing: techniques and future challenges." *Frontiers in Immunology* 9 (2018): 1638.
- 23) Hogan, Sabrina A., et al. "Peripheral blood TCR repertoire profiling may facilitate patient stratification for immunotherapy against melanoma." *Cancer immunology research* 7.1 (2019): 77-85.
- 24) Tumeh, Paul C., et al. "PD-1 blockade induces responses by inhibiting adaptive immune resistance." *Nature* 515.7528 (2014): 568-571.
- 25) Harjes, Ulrike. "T cells pass on the baton during checkpoint inhibition." *Nature Reviews Cancer* 19.10 (2019): 546-546.
- 26) Paul, Michael St, et al. "IL6 Induces an IL22 + CD8 + T-cell Subset with Potent Antitumor Function." *Cancer* 8 (2020): 1-13.
- 27) Yang, Yu, et al. "Identification of a distant TBET enhancer responsive to IL-12/Stat4 and IFN γ /Stat1 signals." *Blood, The Journal of the American Society of Hematology* 110.7 (2007): 2494-2500.
- 28) Visekruna, Alexander, et al. "Tc9 cells, a new subset of CD8+ T cells, support Th2-mediated airway inflammation." *European journal of immunology* 43.3 (2013): 606-618.
- 29) Huber, Magdalena, et al. "A Th17-like developmental process leads to CD8+ Tc17 cells with reduced cytotoxic activity." *European journal of immunology* 39.7 (2009): 1716-1725.
- 30) Ayers, Mark, et al. "IFN- γ -related mRNA profile predicts clinical response to PD-1 blockade." *The Journal of clinical investigation* 127.8 (2017): 2930-2940.
- 31) Sheu, Bor-Ching, et al. "Predominant Th2/Tc2 polarity of tumor-infiltrating lymphocytes in human cervical cancer." *The Journal of Immunology* 167.5 (2001): 2972-2978.
- 32) Faghih, Zahra, et al. "IL-17 and IL-4 producing CD8+ T cells in tumor draining lymph nodes of breast cancer patients: positive association with tumor progression." *Iranian Journal of Immunology* 10.4 (2013): 193-204.
- 33) Kryczek, Ilona, et al. "Cutting edge: Th17 and regulatory T cell dynamics and the regulation by IL-2 in the tumor microenvironment." *The Journal of Immunology* 178.11 (2007): 6730-6733.
- 34) Brummelman, Jolanda, et al. "High-dimensional single cell analysis identifies stem-like cytotoxic CD8+ T cells infiltrating human tumors." *Journal of Experimental Medicine* 215.10 (2018): 2520-2535.
- 35) Cabrita, Rita, et al. "Tertiary lymphoid structures improve immunotherapy and survival in melanoma." *Nature* 577.7791 (2020): 561-565.
- 36) Tirosh, Itay, et al. "Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq." *Science* 352.6282 (2016): 189-196.
- 37) Li, Hanjie, et al. "Dysfunctional CD8 T cells form a proliferative, dynamically regulated compartment within human melanoma." *Cell* 176.4 (2019): 775-789.

- 38) Wei, Spencer C., et al. "Distinct cellular mechanisms underlie anti-CTLA-4 and anti-PD-1 checkpoint blockade." *Cell* 170.6 (2017): 1120-1133.
- 39) Gide, Tuba N., et al. "Distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/anti-CTLA-4 combined therapy." *Cancer cell* 35.2 (2019): 238-255.
- 40) Thommen, Daniela S., et al. "A transcriptionally and functionally distinct PD-1+ CD8+ T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade." *Nature medicine* 24.7 (2018): 994-1004.
- 41) Schalper, Kurt A., et al. "Neoadjuvant nivolumab modifies the tumor immune microenvironment in resectable glioblastoma." *Nature medicine* 25.3 (2019): 470-476.
- 42) Wherry, E. John. "T cell exhaustion." *Nature immunology* 12.6 (2011): 492-499.
- 43) McLane, Laura M., et al. "CD8 T cell exhaustion during chronic viral infection and cancer." *Annual review of immunology* 37 (2019): 457-495.
- 44) Sen, Debattama R., et al. "The epigenetic landscape of T cell exhaustion." *Science* 354.6316 (2016): 1165-1169.
- 45) Martinez, Gustavo J., et al. "The transcription factor NFAT promotes exhaustion of activated CD8+ T cells." *Immunity* 42.2 (2015): 265-278.
- 46) Staron, Matthew M., et al. "The transcription factor FOXO1 sustains expression of the inhibitory receptor PD-1 and survival of antiviral CD8+ T cells during chronic infection." *Immunity* 41.5 (2014): 802-814.
- 47) Jeng, Mark Y., et al. "Metabolic reprogramming of human CD8+ memory T cells through loss of SIRT1." *Journal of Experimental Medicine* 215.1 (2018): 51-62.
- 48) Bengsch, Bertram, et al. "Bioenergetic insufficiencies due to metabolic alterations regulated by the inhibitory receptor PD-1 are an early driver of CD8+ T cell exhaustion." *Immunity* 45.2 (2016): 358-373.
- 49) Alfei, Francesca, et al. "TOX reinforces the phenotype and longevity of exhausted T cells in chronic viral infection." *Nature* 571.7764 (2019): 265-269.
- 50) Khan, Omar, et al. "TOX transcriptionally and epigenetically programs CD8+ T cell exhaustion." *Nature* 571.7764 (2019): 211-218.
- 51) Scott, Andrew C., et al. "TOX is a critical regulator of tumour-specific T cell differentiation." *Nature* 571.7764 (2019): 270-274.
- 52) Clarke, James, et al. "Single-cell transcriptomic analysis of tissue-resident memory T cells in human lung cancer." *Journal of Experimental Medicine* 216.9 (2019): 2128-2149.
- 53) Zheng, Chunhong, et al. "Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing." *Cell* 169.7 (2017): 1342-1356.
- 54) Seo, Hyungseok, et al. "TOX and TOX2 transcription factors cooperate with NR4A transcription factors to impose CD8+ T cell exhaustion." *Proceedings of the National Academy of Sciences* 116.25 (2019): 12410-12415.
- 55) Kjer-Nielsen, Lars, et al. "MR1 presents microbial vitamin B metabolites to MAIT cells." *Nature* 491.7426 (2012): 717-723.

- 56) Siddiqui, Imran, et al. "Intratumoral Tcf1+ PD-1+ CD8+ T cells with stem-like properties promote tumor control in response to vaccination and checkpoint blockade immunotherapy." *Immunity* 50.1 (2019): 195-211.
- 57) Wu, Yilun, et al. "PD-L1 distribution and perspective for cancer immunotherapy—blockade, knockdown, or inhibition." *Frontiers in immunology* 10 (2019): 2022.
- 58) Sheppard, Kelly-Ann, et al. "PD-1 inhibits T-cell receptor induced phosphorylation of the ZAP70/CD3 ζ signalosome and downstream signaling to PKC θ ." *FEBS letters* 574.1-3 (2004): 37-41.
- 59) Wolchok, Jedd D., et al. "Ipilimumab monotherapy in patients with pretreated advanced melanoma: a randomised, double-blind, multicentre, phase 2, dose-ranging study." *The lancet oncology* 11.2 (2010): 155-164.
- 60) Mahoney, Kathleen M., et al. "Combination cancer immunotherapy and new immunomodulatory targets." *Nature reviews Drug discovery* 14.8 (2015): 561-584.
- 61) Wolchok, Jedd D., et al. "Overall survival with combined nivolumab and ipilimumab in advanced melanoma." *New England Journal of Medicine* 377.14 (2017): 1345-1356.
- 62) Hodi, Frank Stephen, et al. "Nivolumab plus ipilimumab or nivolumab alone versus ipilimumab alone in advanced melanoma (CheckMate 067): 4-year outcomes of a multicentre, randomised, phase 3 trial." *The Lancet Oncology* 19.11 (2018): 1480-1492.
- 63) Motzer, Robert J., et al. "Avelumab plus axitinib versus sunitinib for advanced renal-cell carcinoma." *New England Journal of Medicine* 380.12 (2019): 1103-1115.
- 64) Motzer, Robert J., et al. "Nivolumab plus ipilimumab versus sunitinib in first-line treatment for advanced renal cell carcinoma: extended follow-up of efficacy and safety results from a randomised, controlled, phase 3 trial." *The Lancet Oncology* 20.10 (2019): 1370-1385.
- 65) Rini, Brian I., et al. "Atezolizumab plus bevacizumab versus sunitinib in patients with previously untreated metastatic renal cell carcinoma (IMmotion151): a multicentre, open-label, phase 3, randomised controlled trial." *The Lancet* 393.10189 (2019): 2404-2415.
- 66) Yost, Kathryn E., et al. "Clonal replacement of tumor-specific T cells following PD-1 blockade." *Nature medicine* 25.8 (2019): 1251-1259.
- 67) Goswami, Sangeeta, et al. "Immune profiling of human tumors identifies CD73 as a combinatorial target in glioblastoma." *Nature medicine* 26.1 (2020): 39-46.
- 68) Chevrier, Stéphane, et al. "An immune atlas of clear cell renal cell carcinoma." *Cell* 169.4 (2017): 736-749.
- 69) Guo, Xinyi, et al. "Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing." *Nature medicine* 24.7 (2018): 978-985.
- 70) Wagner, Johanna, et al. "A single-cell atlas of the tumor and immune ecosystem of human breast cancer." *Cell* 177.5 (2019): 1330-1345.
- 71) Cabrita, Rita, et al. "Tertiary lymphoid structures improve immunotherapy and survival in melanoma." *Nature* 577.7791 (2020): 561-565.
- 72) Bohn, Toszka, et al. "Tumor immunoevasion via acidosis-dependent induction of regulatory tumor-associated macrophages." *Nature immunology* 19.12 (2018): 1319-1329.
- 73) Philip, Mary, et al. "Chromatin states define tumour-specific T cell dysfunction and reprogramming." *Nature* 545.7655 (2017): 452-456.

- 74) Guo, Xinyi, et al. "Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing." *Nature medicine* 24.7 (2018): 978-985.
- 75) Savas, Peter, et al. "Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis." *Nature medicine* 24.7 (2018): 986-993.
- 76) Tilloy, Florence, et al. "An invariant T cell receptor α chain defines a novel TAP-independent major histocompatibility complex class Ib–restricted α/β T cell subpopulation in mammals." *Journal of Experimental Medicine* 189.12 (1999): 1907-1921.
- 77) Treiner, Emmanuel, et al. "Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1." *Nature* 422.6928 (2003): 164-169.
- 78) Patel, Onisha, et al. "Recognition of vitamin B metabolites by mucosal-associated invariant T cells." *Nature communications* 4.1 (2013): 1-9.
- 79) Lamichhane, Rajesh, et al. "Expression and trafficking of MR1." *Immunology* 151.3 (2017): 270-279.
- 80) Cogswell, Dasha T., et al. "MAIT Cells: Partners or Enemies in Cancer Immunotherapy?" *Cancers* 13.7 (2021): 1502.
- 81) Sattler, Arne, et al. "IL-15 dependent induction of IL-18 secretion as a feedback mechanism controlling human MAIT-cell effector functions." *European journal of immunology* 45.8 (2015): 2286-2298.
- 82) Lamichhane, Rajesh, et al. "TCR-or cytokine-activated CD8+ mucosal-associated invariant T cells are rapid polyfunctional effectors that can coordinate immune responses." *Cell reports* 28.12 (2019): 3061-3076.
- 83) Zabijak, Luciane, et al. "Increased tumor infiltration by mucosal-associated invariant T cells correlates with poor survival in colorectal cancer patients." *Cancer immunology, immunotherapy* 64.12 (2015): 1601-1608.
- 84) Sundström, Patrik, et al. "Tumor-infiltrating mucosal-associated invariant T (MAIT) cells retain expression of cytotoxic effector molecules." *Oncotarget* 10.29 (2019): 2810.
- 85) Shaler, Christopher R, et al. "Mucosa-associated invariant T cells infiltrate hepatic metastases in patients with colorectal carcinoma but are rendered dysfunctional within and adjacent to tumor microenvironment." *Cancer Immunology, Immunotherapy* 66.12 (2017): 1563-1575.
- 86) Li, Shamin, et al. "Human tumor-infiltrating MAIT cells display hallmarks of bacterial antigen recognition in colorectal cancer." *Cell Reports Medicine* 1.3 (2020): 100039.
- 87) Melo, Ashanty M., et al. "Mucosal-associated invariant T cells display diminished effector capacity in oesophageal adenocarcinoma." *Frontiers in immunology* 10 (2019): 1580.
- 88) Huang, Wan-Chun, et al. "Less circulating mucosal-associated invariant T cells in patients with cervical cancer." *Taiwanese Journal of Obstetrics and Gynecology* 58.1 (2019): 117-121.
- 89) Lu, Zhimin, et al. "The combined action of monocytic myeloid-derived suppressor cells and mucosal-associated invariant T cells promotes the progression of cervical cancer." *International Journal of Cancer* 148.6 (2021): 1499-1507.

- 90) Duan, Meng, et al. "Activated and exhausted MAIT cells foster disease progression and indicate poor outcome in hepatocellular carcinoma." *Clinical Cancer Research* 25.11 (2019): 3304-3316.
- 91) Yan, Juming, et al. "MAIT cells promote tumor initiation, growth, and metastases via tumor MR1." *Cancer discovery* 10.1 (2020): 124-141.
- 92) Crowther, Michael D., et al. "Genome-wide CRISPR–Cas9 screening reveals ubiquitous T cell cancer targeting via the monomorphic MHC class I-related protein MR1." *Nature immunology* 21.2 (2020): 178-185.
- 93) Yao, Tony, et al. "Leveraging public single-cell and bulk transcriptomic datasets to delineate MAIT cell roles and phenotypic characteristics in human malignancies." *Frontiers in immunology* (2020): 1691.
- 94) Gopalakrishnan, Vancheswaran, et al. "Gut microbiome modulates response to anti–PD-1 immunotherapy in melanoma patients." *Science* 359.6371 (2018): 97-103.
- 95) Matson, Vyara, et al. "The commensal microbiome is associated with anti–PD-1 efficacy in metastatic melanoma patients." *Science* 359.6371 (2018): 104-108.
- 96) Ahmed, Raya, et al. "Human stem cell-like memory T cells are maintained in a state of dynamic flux." *Cell reports* 17.11 (2016): 2811-2818.
- 97) Joshi, Nikhil S., et al. "Inflammation directs memory precursor and short-lived effector CD8+ T cell fates via the graded expression of TBET transcription factor." *Immunity* 27.2 (2007): 281-295.
- 98) Feng, Xiaoming, et al. "Transcription factor Foxp1 exerts essential cell-intrinsic regulation of the quiescence of naive T cells." *Nature immunology* 12.6 (2011): 544-550.
- 99) Pearce, Erika L., et al. "Control of effector CD8+ T cell function by the transcription factor Eomesodermin." *Science* 302.5647 (2003): 1041-1043.
- 100) Morgan, Richard A., et al. "Cancer regression in patients after transfer of genetically engineered lymphocytes." *Science* 314.5796 (2006): 126-129.
- 101) Biasco, Luca, et al. "Clonal expansion of T memory stem cells determines early anti-leukemic responses and long-term CAR T cell persistence in patients." *Nature Cancer* (2021): 1-14.
- 102) He, Kun, et al. "The cryo-thermal therapy eradicated melanoma in mice by eliciting CD4+ T-cell-mediated antitumor memory immune response." *Cell death & disease* 8.3 (2017): e2703-e2703.
- 103) Lopci, Egesta, et al. "Correlation of metabolic information on FDG-PET with tissue expression of immune markers in patients with non-small cell lung cancer (NSCLC) who are candidates for upfront surgery." *European journal of nuclear medicine and molecular imaging* 43.11 (2016): 1954-1961.
- 104) Jansen, Caroline S., et al. "An intra-tumoral niche maintains and differentiates stem-like CD8 T cells." *Nature* 576.7787 (2019): 465-470.
- 105) Galletti, Giovanni, et al. "Two subsets of stem-like CD8+ memory T cell progenitors with distinct fate commitments in humans." *Nature immunology* 21.12 (2020): 1552-1562.

- 106) Zhao, Yanran, et al. "Human CD8 T-stem cell memory subsets phenotypic and functional characterization are defined by expression of CD122 or CXCR3." *European Journal of Immunology* (2021).
- 107) Kaartinen, Tanja, et al. "Low interleukin-2 concentration favors generation of early memory T cells over effector phenotypes during chimeric antigen receptor T-cell expansion." *Cytotherapy* 19.6 (2017): 689-702.
- 108) Chen, Yingshi, et al. "Adoptive transfer of interleukin-21-stimulated human CD8+ T memory stem cells efficiently inhibits tumor growth." *Journal of Immunotherapy* (Hagerstown, Md.: 1997) 41.6 (2018): 274.
- 109) Zhu, Jun, et al. "Cryo-thermal therapy elicits potent anti-tumor immunity by inducing extracellular Hsp70-dependent MDSC differentiation." *Scientific reports* 6.1 (2016): 1-13.
- 110) He, Kun, et al. "The cryo-thermal therapy eradicated melanoma in mice by eliciting CD4+ T-cell-mediated antitumor memory immune response." *Cell death & disease* 8.3 (2017): e2703-e2703.
- 111) Jin, Linchun, et al. "CXCR1-or CXCR2-modified CAR T cells co-opt IL-8 for maximal antitumor efficacy in solid tumors." *Nature communications* 10.1 (2019): 1-13.
- 112) Biasco, Luca, et al. "In vivo tracking of T cells in humans unveils decade-long survival and activity of genetically modified T memory stem cells." *Science translational medicine* 7.273 (2015): 273ra13-273ra13.
- 113) De Rosa, Stephen C., et al. "11-color, 13-parameter flow cytometry: identification of human naive T cells by phenotype, function, and T-cell receptor diversity." *Nature medicine* 7.2 (2001): 245-248.
- 114) Perfetto, Stephen P., et al. "Seventeen-colour flow cytometry: unravelling the immune system." *Nature Reviews Immunology* 4.8 (2004): 648-655.
- 115) Bandura, Dmitry R., et al. "Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry." *Analytical chemistry* 81.16 (2009): 6813-6822.
- 116) Chang, Qing, et al. "Imaging mass cytometry." *Cytometry part A* 91.2 (2017): 160-169.
- 117) Goltsev, Yury, et al. "Deep profiling of mouse splenic architecture with CODEX multiplexed imaging." *Cell* 174.4 (2018): 968-981.
- 118) Giesen, Charlotte, et al. "Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry." *Nature methods* 11.4 (2014): 417-422.
- 119) Black, Sarah, et al. "CODEX multiplexed tissue imaging with DNA-conjugated antibodies." *Nature Protocols* 16.8 (2021): 3802-3835.
- 120) Umemoto, Kumiko, et al. "The potential application of PD-1 blockade therapy for early-stage biliary tract cancer." *International immunology* 32.4 (2020): 273-281.
- 121) Tang, Fuchou, et al. "mRNA-Seq whole-transcriptome analysis of a single cell." *Nature methods* 6.5 (2009): 377-382.
- 122) Tang, Fuchou, et al. "mRNA-Seq whole-transcriptome analysis of a single cell." *Nature methods* 6.5 (2009): 377-382.

- 123) Lo Tartaro, Domenico, et al. "Gene Expression Analysis of T-Cells by Single-Cell RNA-Seq." *Methods in Molecular Biology* (Clifton, NJ) 2285 (2021): 277-296.
- 124) Ramsköld, Daniel, et al. "Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells." *Nature biotechnology* 30.8 (2012): 777-782.
- 125) Pollen, Alex A., et al. "Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex." *Nature biotechnology* 32.10 (2014): 1053-1058.
- 126) Islam, Saiful, et al. "Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq." *Genome research* 21.7 (2011): 1160-1167.
- 127) Dobin, Alexander, et al. "STAR: ultrafast universal RNA-seq aligner." *Bioinformatics* 29.1 (2013): 15-21.
- 128) Li, Heng, et al. "Fast and accurate short read alignment with Burrows–Wheeler transform." *bioinformatics* 25.14 (2009): 1754-1760.
- 129) Bushnell, Brian, et al. BBMap: a fast, accurate, splice-aware aligner. No. LBNL-7065E. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2014.
- 130) Langmead, Ben, and Steven L. Salzberg. "Fast gapped-read alignment with Bowtie 2." *Nature methods* 9.4 (2012): 357-359.
- 131) Peterson, Vanessa M., et al. "Multiplexed quantification of proteins and transcripts in single cells." *Nature biotechnology* 35.10 (2017): 936-939.
- 132) Shahi, Payam, et al. "Abseq: Ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding." *Scientific reports* 7.1 (2017): 1-12.
- 133) Stoeckius, Marlon, et al. "Simultaneous epitope and transcriptome measurement in single cells." *Nature methods* 14.9 (2017): 865-868.
- 134) Zhou, Zilu, et al. "Surface protein imputation from single cell transcriptomes by deep neural networks." *Nature communications* 11.1 (2020): 1-10.
- 135) Buenrostro, Jason D., et al. "ATAC-seq: a method for assaying chromatin accessibility genome-wide." *Current protocols in molecular biology* 109.1 (2015): 21-29.
- 136) Papalex, Efthymia, et al. "Characterizing the molecular regulation of inhibitory immune checkpoints with multimodal single-cell screens." *Nature Genetics* 53.3 (2021): 322-331.
- 137) Mazza, Emilia Maria Cristina, et al. "Background fluorescence and spreading error are major contributors of variability in high-dimensional flow cytometry data visualization by t-distributed stochastic neighboring embedding." *Cytometry Part A* 93.8 (2018): 785-792.
- 138) Mair, Florian, et al. "OMIP-044: 28-color immunophenotyping of the human dendritic cell compartment." *Cytometry Part A* 93.4 (2018): 402-405.
- 139) Brummelman, Jolanda, et al. "Development, application and computational analysis of high-dimensional fluorescent antibody panels for single-cell flow cytometry." *Nature protocols* 14.7 (2019): 1946-1969.
- 140) Roederer, Mario, et al. "Spectral compensation for flow cytometry: visualization artifacts, limitations, and caveats." *Cytometry: The Journal of the International Society for Analytical Cytology* 45.3 (2001): 194-205.

- 141) Perfetto, Stephen P., et al. "Quality assurance for polychromatic flow cytometry using a suite of calibration beads." *Nature protocols* 7.12 (2012): 2067-2079.
- 142) Finak, Greg, et al. "Optimizing transformations for automated, high throughput analysis of flow cytometry data." *BMC bioinformatics* 11.1 (2010): 1-13.
- 143) Levine, Jacob H., et al. "Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis." *Cell* 162.1 (2015): 184-197.
- 144) Van Gassen, Sofie, et al. "FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data." *Cytometry Part A* 87.7 (2015): 636-645.
- 145) Nowicka, Malgorzata, et al. "CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets." *F1000Research* 6 (2017).
- 146) Greene, Evan, et al. "New interpretable machine-learning method for single-cell data reveals correlates of clinical response to cancer immunotherapy." *Patterns* 2.12 (2021): 100372.
- 147) Lugli, Enrico, et al. "Subject classification obtained by cluster analysis and principal component analysis applied to flow cytometric data." *Cytometry Part A: The Journal of the International Society for Analytical Cytology* 71.5 (2007): 334-344.
- 148) Van der Maaten, Laurens, et al. "Visualizing data using t-SNE." *Journal of machine learning research* 9.11 (2008).
- 149) McInnes, Leland, et al. "Umap: Uniform manifold approximation and projection for dimension reduction." *arXiv preprint arXiv:1802.03426* (2018).
- 150) <https://towardsdatascience.com/tsne-vs-umap-global-structure-4d8045acba17>
- 151) Satija, Rahul, et al. "Spatial reconstruction of single-cell gene expression data." *Nature biotechnology* 33.5 (2015): 495-502.
- 152) Lun, Aaron TL, et al. "Pooling across cells to normalize single-cell RNA sequencing data with many zero counts." *Genome biology* 17.1 (2016): 1-14.
- 153) Hafemeister, Christoph, et al. "Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression." *Genome biology* 20.1 (2019): 1-15.
- 154) Butler, Andrew, et al. "Integrating single-cell transcriptomic data across different conditions, technologies, and species." *Nature biotechnology* 36.5 (2018): 411-420.
- 155) Haghverdi, Laleh, et al. "Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors." *Nature biotechnology* 36.5 (2018): 421-427.
- 156) Korsunsky, Ilya, et al. "Fast, sensitive and accurate integration of single-cell data with Harmony." *Nature methods* 16.12 (2019): 1289-1296.
- 157) Stuart, Tim, et al. "Comprehensive integration of single-cell data." *Cell* 177.7 (2019): 1888-1902.
- 158) Pearson, Karl. "LIII. On lines and planes of closest fit to systems of points in space." *The London, Edinburgh, and Dublin philosophical magazine and journal of science* 2.11 (1901): 559-572.

- 159) Coifman, Ronald R., et al. "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps." *Proceedings of the national academy of sciences* 102.21 (2005): 7426-7431.
- 160) Xu, Chen, et al. "Identification of cell types from single-cell transcriptomes using a novel clustering method." *Bioinformatics* 31.12 (2015): 1974-1980.
- 161) Blondel, Vincent D., et al. "Fast unfolding of communities in large networks." *Journal of statistical mechanics: theory and experiment* 2008.10 (2008): P10008.
- 162) Trapnell, Cole, et al. "The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells." *Nature biotechnology* 32.4 (2014): 381-386. APA.
- 163) Cannoodt, Robrecht, et al. "SCORPIUS improves trajectory inference and identifies novel modules in dendritic cell development." *bioRxiv* (2016): 079509.
- 164) Bendall, Sean C., et al. "Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development." *Cell* 157.3 (2014): 714-725.
- 165) Topalian, Suzanne L., et al. "Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy." *Nature Reviews Cancer* 16.5 (2016): 275-287.
- 166) Havel, Jonathan J., et al. "The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy." *Nature Reviews Cancer* 19.3 (2019): 133-150.
- 167) Lugli, Enrico, et al. "Stem, effector, and hybrid states of memory CD8+ T cells." *Trends in immunology* 41.1 (2020): 17-28.
- 168) Miao, Diana, et al. "Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors." *Nature genetics* 50.9 (2018): 1271-1281.
- 169) Miao, Diana, et al. "Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors." *Nature genetics* 50.9 (2018): 1271-1281.
- 170) Davoli, Teresa, et al. "Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy." *Science* 355.6322 (2017).
- 171) Wu, Thomas D., et al. "Peripheral T cell expansion predicts tumour infiltration and clinical response." *Nature* 579.7798 (2020): 274-278.
- 172) Gibellini, Lara, et al. "Single-cell approaches to profile the response to immune checkpoint inhibitors." *Frontiers in immunology* 11 (2020): 490.
- 173) Motzer, Robert J., et al. "Nivolumab versus everolimus in advanced renal-cell carcinoma." *New England Journal of Medicine* 373.19 (2015): 1803-1813.
- 174) Motzer, Robert, et al. "Lenvatinib plus pembrolizumab or everolimus for advanced renal cell carcinoma." *New England Journal of Medicine* 384.14 (2021): 1289-1300.
- 175) Cossarizza, Andrea, et al. "Guidelines for the use of flow cytometry and cell sorting in immunological studies." *European journal of immunology* 49.10 (2019): 1457-1973.
- 176) Tang, Fuchou, et al. "mRNA-Seq whole-transcriptome analysis of a single cell." *Nature methods* 6.5 (2009): 377-382.
- 177) Shen, Li, et al. "GeneOverlap: Test and visualize gene overlaps." *R package version 1.0* (2013): 2013.
- 178) Roederer, Mario, et al. "SPICE: exploration and analysis of post-cytometric complex multivariate datasets." *Cytometry Part A* 79.2 (2011): 167-174.

- 179) Gattinoni, Luca, et al. "T memory stem cells in health and disease." *Nature medicine* 23.1 (2017): 18-27.
- 180) Mahnke, Yolanda D., et al. "The who's who of T-cell differentiation: human memory T-cell subsets." *European journal of immunology* 43.11 (2013): 2797-2809.
- 181) Larbi, Anis, et al. "From "truly naïve" to "exhausted senescent" T cells: when markers predict functionality." *Cytometry Part A* 85.1 (2014): 25-35.
- 182) Deng, Ling, et al. "CXCR6/CXCL16 functions as a regulator in metastasis and progression of cancer." *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* 1806.1 (2010): 42-49.
- 183) Duhén, Thomas, et al. "Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors." *Nature communications* 9.1 (2018): 1-13.
- 184) Kurtulus, Sema, et al. "Checkpoint blockade immunotherapy induces dynamic changes in PD-1– CD8+ tumor-infiltrating T cells." *Immunity* 50.1 (2019): 181-194.
- 185) Powell Jr, Daniel J., et al. "Transition of late-stage effector T cells to CD27+ CD28+ tumor-reactive effector memory T cells in humans after adoptive cell transfer therapy." *Blood* 105.1 (2005): 241-250.
- 186) Kallies, Axel, et al. "Transcription factor TBET orchestrates lineage development and function in the immune system." *Trends in immunology* 38.4 (2017): 287-297.
- 187) Gide, Tuba N., et al. "Distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/anti-CTLA-4 combined therapy." *Cancer cell* 35.2 (2019): 238-255.
- 188) Wieland, Andreas, et al. "T cell receptor sequencing of activated CD8 T cells in the blood identifies tumor-infiltrating clones that expand after PD-1 therapy and radiation in a melanoma patient." *Cancer Immunology, Immunotherapy* 67.11 (2018): 1767-1776.
- 189) Kamphorst, Alice O., et al. "Proliferation of PD-1+ CD8 T cells in peripheral blood after PD-1–targeted therapy in lung cancer patients." *Proceedings of the National Academy of Sciences* 114.19 (2017): 4993-4998.
- 190) Voillet, Valentin, et al. "Human MAIT cells exit peripheral tissues and recirculate via lymph in steady state conditions." *JCI insight* 3.7 (2018).
- 191) Godfrey, Dale I., et al. "The biology and functional importance of MAIT cells." *Nature immunology* 20.9 (2019): 1110-1128.
- 192) Domanska, Urszula M., et al. "A review on CXCR4/CXCL12 axis in oncology: no place to hide." *European journal of cancer* 49.1 (2013): 219-230.
- 193) Edwards, Jarem, et al. "Cd103+ tumor-resident cd8+ t cells are associated with improved survival in immunotherapy-naïve melanoma patients and expand significantly during anti–pd-1 treatment." *Clinical Cancer Research* 24.13 (2018): 3036-3045.
- 194) Koay, Hui-Fern, et al. "A three-stage intrathymic development pathway for the mucosal-associated invariant T cell lineage." *Nature immunology* 17.11 (2016): 1300-1311.
- 195) Reantragoon, Rangsim, et al. "Structural insight into MR1-mediated recognition of the mucosal associated invariant T cell receptor." *Journal of Experimental Medicine* 209.4 (2012): 761-774.

- 196) Kluger, Harriet M., et al. "Defining tumor resistance to PD-1 pathway blockade: recommendations from the first meeting of the SITC immunotherapy resistance Taskforce." *Journal for immunotherapy of cancer* 8.1 (2020).
- 197) Kawashima, Atsunari, et al. "Tumour grade significantly correlates with total dysfunction of tumour tissue-infiltrating lymphocytes in renal cell carcinoma." *Scientific reports* 10.1 (2020): 1-13.
- 198) Menard, Laurence C., et al. "Renal cell carcinoma (RCC) tumors display large expansion of double positive (DP) CD4+ CD8+ T cells with expression of exhaustion markers." *Frontiers in immunology* 9 (2018): 2728.
- 199) Teke, Kerem, et al. "Prognostic value of immunological profile based on CD8+ and FoxP3+ T lymphocytes in the peritumoral and intratumoral subsites for renal cell carcinoma." *International Urology and Nephrology* 52.12 (2020): 2289-2299.
- 200) Juliá, Estefanía Paula, et al. "Peripheral changes in immune cell populations and soluble mediators after anti-PD-1 therapy in non-small cell lung cancer and renal cell carcinoma patients." *Cancer Immunology, Immunotherapy* 68.10 (2019): 1585-1596.
- 201) Klebanoff, Christopher A., et al. "Determinants of successful CD8+ T-cell adoptive immunotherapy for large established tumors in mice." *Clinical Cancer Research* 17.16 (2011): 5343-5352.
- 202) Lugli, Enrico, et al. "Superior T memory stem cell persistence supports long-lived T cell memory." *The Journal of clinical investigation* 123.2 (2013).
- 203) Yan, Cihui, et al. "Memory stem T cells generated by Wnt signaling from blood of human renal clear cell carcinoma patients." *Cancer biology & medicine* 16.1 (2019): 109.
- 204) Felix, Joana, et al. "Ipilimumab reshapes T cell memory subsets in melanoma patients with clinical response." *Oncoimmunology* 5.7 (2016): 1136045.
- 205) <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>.
- 206) Sung, Hyuna, et al. "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries." *CA: a cancer journal for clinicians* 71.3 (2021): 209-249.
- 207) Larkin, James, et al. "Combined nivolumab and ipilimumab or monotherapy in untreated melanoma." *New England journal of medicine* 373.1 (2015): 23-34.
- 208) Borghaei, Hossein, et al. "Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer." *New England Journal of Medicine* 373.17 (2015): 1627-1639.

12. SUPPLEMENTARY TABLES

Specificity	Dye	Clone	Vendor	Catalog Number	Lot number	μl used (in a volume of 100μl)
AQUA	BV510	N/A	Biolegend	423102	B244918	0,06
CCR7	BB660	150503	BD	625454	7136810	1,25
CXCR6	BV421	13B 1E5	BD	566008	7131852	2,50
CD194	PE-CF594	1G1	BD	565391	7276847	0,30
CD25	BB790	M-A251	BD	353736	7150757	2,5
CD244	PC5.5	C1.7	BC	B21171	16	1,25
CD98	BB515	UM7F8	BD	565103	7283973	1,25
CD28	BV786	CD28.2	Biolegend	302950	B249581	0,6
CD39	APC-H7	A1	Biolegend	328226	B246491	1,25
CD127	APC-R700	HIL-7R-M21	BD	565185	7221803	2,5
CD38	BV711	HIT2	Biolegend	303528	B238633	2,5
CD272 (BTLA)	BV650	J168-540	BD	564803	7103949	0,60
CD57	BV605	NK-1	BD	563895	7018755	1,25
CD27	BV570	O323	Biolegend	302825	B244350	1,25
CD71	PE-CY5	M-A712	BD	551143	7058789	0,60
CD8	BUV805	SK1	BD	564912	7311599	1,25
CD95	BUV737	DX2	BD	564710	7032582	0,63
HLA-DR	BUV661	G46-6	BD	565073	7249926	0,03
CD4	BUV615	SK3	BD	624297		0,31
CD45RA	BUV563	HI100	BD	565702	7219651	1,25
CD3	BUV496	UCHT1	BD	564809	7312719	2,50
ICOS CD278	BUV395	DX29	BD	564777	7263764	2,50
PD-1	PE	EH12.1	BD			2,5
IgG4	PE	HP6025	Southern Biotech	9200-09	B3317-YD17B	1,25
TBET	PE-CY7	4B10	eBioscience	25-5825-80	4286798	0,15
GRANULYSIN	APC	DH2	Biolegend	348006	B192033	5,00
KI67	BV480	B56	BD	566109	7254647	5,00

Supplementary Table 1. Reagents for the 28-color panel used for the characterization of T cells.

Specificity	Dye	Clone	Vendor	Catalog Number	Lot Number	μl used (in a volume of 100μl)
PromoFluor-840	Maleimide	N/A	Promokine	PK-PF840-3-05	429P017	0,3
CXCR4	BV421	12G5	BD	562448	4003566	1,25
CCR7	BV605	G043H7	Biolegend	353224	B284011	2,5
CD3	AF700	UCHT1	Biolegend	300424	B279942	0,6
CD8a	APC-CY7	RPA-T8	Biolegend	301016	B274260	1,5
CD161	PC7	B30631	Beckman Coulter	B191B8	200023	2,5
TCR72	PE	3C10	Biolegend	351706	B292236	1,25
CD69	FITC	FN50	Biolegend	310904	B290844	0,6
CD95	BUV395	DX2	BD	740306	9269965	0,6
CD38	BUV496	HIT2	BD	564658	8339559	1,25
HLA-DR	BUV661	G46-6	BD	565073	7249926	0,3
CD127	BV650	A019D5	Biolegend	351325	B238890	0,6
CD25	BV785	BC96	Biolegend	302638	B237289	1,25
CD45RO	PC5.5	UCHL1	Beckman Coulter	B30638	200022	10
Ki67	BV480	B56	BD	566109	7254647	5
GRANULYSIN	AF647	DH2	Biolegend	348006	B192033	5

Supplementary Table 2. mAbs used in flow cytometry panel used for the detection of MAIT cells.

Specificity	Dye	Clone	Vendor	Catalog Number	Lot Number	μl used in a volume of 100μl
Live Dead	AQUA	N/A	ThermoFisher	L34965	2069643	1,25
CD4	AF700	RPA-T4	Biolegend	300526	B274110	0,6
CD8	APC-Cy7	RPA-T8	Biolegend	301016	B274260	0,6
CD161	PE-Cy7	HP-3G10	Biolegend	339918	B204086	0,6
TCR 7.2	PE	3C10	Biolegend	351706	B292236	1,25
CD3	PE-Cy5	UCHT1	Biolegend	300410	B252119	0,6
TNF	BV605	MAb11	Biolegend	502936	B282176	3,75
IFN-γ	FITC	B27	Biolegend	506504	B286029	2,5
GRZB	BV421	QA18A28	Biolegend	396414	B296875	2,5
GRZA	APC	CB9	Biolegend	507220	B277773	5

Supplementary Table 3. mAbs used for investigating the function of MAIT cells.

Specificity	Dye	Clone	Vendor	Catalog Number	Lot Number	μl used in a volume of 100μl
Live Dead	AQUA	N/A	ThermoFisher	L34965	2069643	1,25
CD4	AF700	RPA-T4	Biolegend	300526	B274110	0,6
CD8	APC-Cy7	RPA-T8	Biolegend	301016	B274260	0,6
CD45RA	PE-Cy7	HI100	Biolegend	304126	B248293	1,3
CD95	PE	DX2	Biolegend	305608	B314811	2,50
CD3	PE-Cy5	UCHT1	Biolegend	300410	B252119	0,6
TNF	BV605	MAB11	Biolegend	502936	B282176	3,75
IFN-γ	FITC	B27	Biolegend	506504	B286029	2,5
CCR7	BV421	G043H7	Biolegend	353208	B263161	1,3
IL-2	APC	MQ1-17H12	Biolegend	500310	B294390	3

Supplementary Table 4. mAbs used for investigating the function of CD8+ T cells.

13. Acknowledgements

It is a genuine pleasure to express my deep sense of thanks and gratitude to my mentor, **professor Andrea Cossarizza**, director of the laboratory of General pathology and Immunology, University of Modena and Reggio Emilia. His timely advice, meticulous scrutiny and scientific approach have helped me to a very great extent to accomplish this task.

I owe a deep sense of gratitude to **Dr.ssa Sara De Biasi** and **Dr.ssa Lara Gibellini** for their keen interest on me at every stage of my research. Their prompt inspiration, knowledge and enthusiasm have enabled me to complete my thesis.

It is my privilege to thank all the past and current members of the laboratory: **Marco Mattioli, Anita Neroni, Rosalberta Pullano, Simone Pecorini, Annamaria Paolini** and **Rebecca Borella** for their constant encouragement throughout my research period.

I am extremely thankful to **professor Mattia Forcato**, University of Modena and Reggio Emilia, and **Dr. Greg Finak**, Fred Hutchinson Cancer Research Center, Seattle, USA for providing me necessary technical suggestions during my research pursuit.

I thank profusely all the doctors of the **Azienda Ospedaliero Universitaria di Modena and Reggio Emilia** for their clinical support and for the care and selection of patients.

Finally, it is my privilege to thank all my **family** and **friends Davide, Daniele, Pietro and Emanuele** for their love that they profuse to me during my study period.