D-MEM
PhD-UNIMI

Doctorate program
Milan
EXPERIMENTAL
MEDICINE

UNIVERSITÀ DEGLI STUDI DI MILANO

PhD Course in Experimental Medicine

CYCLE XXXIV

PhD thesis

EPIGENETIC CHARACTERIZATION OF TUMOR INFILTRATING CD4+ REGULATORY T CELLS

Candidate: **Dr. Ramona Bason R12321**

Tutor: **Prof. Massimiliano Pagani**

Supervisor: **Grazisa Rossetti, PhD**

Director: **Prof. Nicoletta Landsberger**

Academic Year 2020-2021

SUMMARY

ABSTRACT

In recent years, the role of CD4+ regulatory T cells (Treg cells) in inhibiting the anti cancer activity of effector T cells has become increasingly evident and they are therefore currently considered promising targets for cancer immunotherapy. Despite Treg cell depletion has been reported to increase anti-tumor specific immune responses and to reduce tumor burden, some relevant issues still remain to be addressed, for a safer, more effective clinical application of these therapies.

Previous findings in our lab identified unique transcriptional profiles of human Treg cells infiltrating colorectal cancer (CRC) and non-small cell lung cancer (NSCLC) supporting the existence of an underlying regulatory hubs that specifically shape tumor Treg identity and represent potential target for their functional modulation specificially in the tumor microenvironment.

Treg cell transcriptome represents just a layer of the machinery that drive the acquisition of their tumor cell state. Ultimately, a comprehensive understanding of the regulatory networks that govern tumor Treg gene expression programs can provide insights on how a more selective inhibition can be achieved by interfering with the hubs that translate the cues coming from the tumor microenvironment.

To define the epigenetic blueprints specific for tumor infiltrating Treg cells we have integrated distinct histone marks (H3K4me3, H3K4me1, H3K36me3, H3K27ac, H3K27me3) with ATAC-seq derived chromatin accessibility data we generated in Treg cells isolated from peripheral blood, normal and tumor tissues. To infer genome chromatin states we have employed ChromHMM a machine learning based approach that predicts the molecular structure of promoters and enhancers based on data from chromatin accessibility assays and a set of histone modification. We focused in particular on active enhancer regions, characterized by the co-presence of H3K27Ac and H3K4me1. Transcription factor footprinting analysis on the identified regulatory regions showed TF groups with distinct binding activity profile across Treg populations that are clearly lost or gained specifically in tumor Treg cells. Based on our findings we are implementing CRISPR based epigenetic modulation of selected enhancers that represent a novel strategy to reverse Treg mediated immunosuppression in the tumor microenvironment.

The new knowledge coming from this project cues will lead to a better understanding of tiTreg plasticity and to the identification of novel potential targets that can guide the rational development of tiTreg reprograming and innovative therapeutics for cancer with increased efficacy and reduced adverse effects.

DISCLOSURE FOR RESEARCH INTEGRITY

I hereby state that the research presented in this thesis was conducted following the European Code of Conduct for Research Integrity.

ABBREVIATIONS

Ab	Antibody
ALI	Air-Liquid Interface culture
APC	Antigen Presenting Cell
ATAC-seq	Assay for Transposase-Accessible Chromatin using sequencing
cAMP	cyclic Adenosine Monophosphate
BAM	Binary file
CD	Cluster of Differentiation
ChIP-seq	Chromatin Immunoprecipitation followed by sequencing
CFSE	Carboxyfluorescein Diacetate Succinimidyl Ester
CMS	Consensus Molecular Subtype
CRC	Colorectal Cancer
CTLA-4	Cytotoxic T-Lymphocyte Antigen 4
DOC	Na-Deoxycholate
MDSC	Myeloid-derived Suppressor Cells
EGFR	Epidermal Growth Factor Receptor
FACS	Fluorescence-Activated Cell Sorting
FBS	Fetal Bovine Serum
FOXP3	Forkhead box p3
GITR	Glucocorticoid-induced TNF Receptor
H3K4me1	Monomethylation of Lysine 4 of Histone H3
H3K4me3	Trimethylation of Lysine 4 of Histone H3
H3K27me3	Trimethylation of Lysine 27 of Histone H3
H3K36me3	Trimethylation of Lysine 36 of Histone H3
H3K27ac	acetylation of lysine 27 of histone H3
HAT	Histone Acetyltransferase,
HD	Healthy Donor
HDAC	Histone Deacetylases
HDM	Histone Demethilase
HKMT	Histone Lysine Methyltransferase

IDO	Indolamine 2,3-Dioxygenase
Ig	Immunoglobulin
IL	Interleukin
IFN	Interferon
IPEX	Immunodisregulation, Polyendocrinopathy, Enteropathy, X-linked Syndrome
K	Tumoral tissue
KRAS	v-Ki-ras2 Kirsten rat Sarcoma Viral Oncogene
LAG-3	Lymphocyte-Activation Gene 3
LPMC	Lamina Propria Mononuclear Cell
mAb	monoclonal Antibody
MHC	Major Histocompatibility Complex
MSI-H	High Microsatellite Instability
N	Normal tissue
NAT-Treg	Normal Tissue-infiltrating Treg
NFR	Nucleosome Free Regions
NK	Natural Killer
NSCLC	Non-Small Cell Lung Cancer
nTreg	natural T regulatory cell
OCR	Open Chromatin Regions
PB	Peripheral Blood
PBMC	Peripheral Blood Mononuclear Cells
PBS	Phosphate-Buffered Saline
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PD1	Programmed Cell Death 1
PD-L1	Ligand of PD1
PIC	Protease Inhibitor Cocktail
PMSF	Phenilmetilsulfonil Fluoride
PPAR-γ	Peroxisome Proliferator-Activated Receptor γ
pTreg	Peripheral T regulatory cell
RT-PCR	Real Time PCR

STAT	Signal Transducer and Activator of Transcription
TAM	Tumor Associated Macrophage
TCR	T cell Receptor
Th	helper T Lymphocyte
TIL	Tumor-infiltrating T Lymphocyte
TME	Tumor Microenvironment
TNF	Tumor Necrosis Factor
TNFR2	TNF Receptor 2
TGF-β	Transforming Growth Factor β
Tr1	Type 1 regulatory T cell
Treg	T regulatory cell
TSS	Transcription Start Site
VEGF	Vascular Endothelial Growth Factor

1 INTRODUCTION

1.1 Cancer and immune system

1.1.1 Cancer

Improvement of cancer treatments represent a pressing medical need as cancer is still the second leading cause of death worldwide(1). Tumor development and progression is the result of the accumulation of genetic and epigenetic alterations that lead to changes in cellular and tissue functions(2). Specific patterns of alterations are associated with environmental factors, such as tobacco smoke, alcohol intake, mutagenic chemicals, ultraviolet light, viruses and bacteria(3,4). These changes mainly occur in genes that positively (proto-oncogenes) or negatively (tumor suppressor genes) regulate cell division and are involved in DNA repair mechanisms (DNA repair genes). The dysfunction of these genes is responsible for uncontrolled cell proliferation(5) that leads to the formation of tissue masses called tumors. One of the main characteristics of tumors is the phenotypic and functional heterogeneity which is caused by genetic alterations and environmental stimuli. Both intra-tumor and inter-tumor heterogeneity make cancer difficult to diagnose and treat efficiently(6,7).

1.1.1.1 Non-Small Cell Lung Cancer

Lung cancer is among the most dangerous cancers in terms of cancer-related deaths for both men and women. Its death rate exceeds that of the three most common cancers (colon, breast, and pancreatic) combined(1). Over half of patients diagnosed with lung cancer die within one year of diagnosis and the 5-year survival is less than 18%(8). Tobacco smoking is the leading risk factor for lung cancer(9).

There are two main subtypes of lung cancer: small-cell lung carcinoma and non-small-cell lung carcinoma (NSCLC); this last subtype accounts for 85% of lung cancers and is the first cause of cancer-related mortality. NSCLC is further classified into three types: squamous-cell carcinoma, adenocarcinoma, and large-cell carcinoma, with adenocarcinoma being the most common type (around 40% of all lung cancers in men and women)(9).

Most commonly mutated oncogenes, such as Epidermal growth factor receptor (EGFR) and v-Ki-ras2 Kirsten rat sarcoma viral oncogene (KRAS), can be considered as genetic drivers of NSCLC, thus representing therapeutic targets(10,11). Moreover, it has been observed a high somatic mutation burden in NSCLC, that could be responsible for the generation of tumor-specific neoantigens recognized by the effector cells of the immune system. In this regard, immune checkpoint inhibitors against PD-1 have gained promising results in patients with NSCLC. Unfortunately, not all the patients respond to therapy, often developing resistance to the therapy itself, due to the downregulation of the major histocompatibility complex (MHC), to defects in interferon-γ (IFN-γ) signaling and an alteration of indolamine 2,3-dioxygenase (IDO) expression, an enzyme fundamental for T cell metabolism (12,13).

The immune infiltrate of NSCLCs has been recently extensively characterized through flow cytometry. Several distinct immune cell types have been identified: T cells dominate the immune cell composition and, in particular, CD4+ T cells are the most abundant, but also CD8+ and CD4−CD8− double-negative T cells have been found. The presence of tumor-infiltrating macrophages, different subsets of dendritic cells and high density of neutrophils have been also described in the immune contexture of NSCLC(14). importantly, data from several studies involving approximately thousands of patients have been analyzed regarding the prognostic relevance of the different cell types found in the tumor microenvironment of NSCLC. According to the results, the presence of regulatory T cells (Treg cells) and M2 macrophages in the immune infiltrate has been correlated with mostly negative prognosis, whereas the presence of tertiary lymphoid structures and CD8+ T cells has been associated with positive prognosis(15).

1.1.1.2 Colorectal Cancer

Colorectal cancer (CRC) is the third most common type of cancer in terms of incidence and the second most dangerous cancer in terms of cancer-related deaths. The 5-year overall survival rate for CRC patients it is 63%(1). Its incidence is age-related, with a sharp peak in the seventh decade of life; obesity, sedentary lifestyle, consumption of red meat, alcohol and tobacco are the leading risk factors for CRC(4).

CRC is a heterogeneous disease and has been stratified into different classification systems based on genetic and molecular profiles. A new molecular classification of CRC has recently been proposed, based on the gene expression profiles of tumor and stromal cells. This consensus molecular subtype (CMS) classification divides CRCs into four main groups: CMS1, which includes hypermutated tumors with high microsatellite instability (MSI-H) and high rates of immune infiltrates; CMS2 consists of all the canonical CRC subtypes characterized by upregulation of the Myc and Wnt signal; CMS3 includes tumors with KRAS mutations and disruption of metabolic pathways; finally, CMS4 is the mesenchymal and more aggressive subtype, defined by a strong activation of the epithelial-mesenchymal transition pathways, angiogenesis and stemness; this CMS includes tumors with activated TGF-β pathways and also the presence of immune infiltrates(16).

Recently the composition of immune and stromal components in the different molecular subtypes of CRC has been analyzed. In this study, CMS2 and CMS3 were immunologically cold, as they were free of lymphoid and myeloid infiltration and had poor expression of MHC genes. In contrast, CMS1, associated with a good prognosis, overexpressed specific genes for cytotoxic lymphocytes and CMS4, associated with a poor prognosis, overexpressed markers of lymphocytes and monocytes, and displayed characteristic signs of angiogenesis, inflammation and immunosuppression as well as the maximum density of fibroblasts(17).

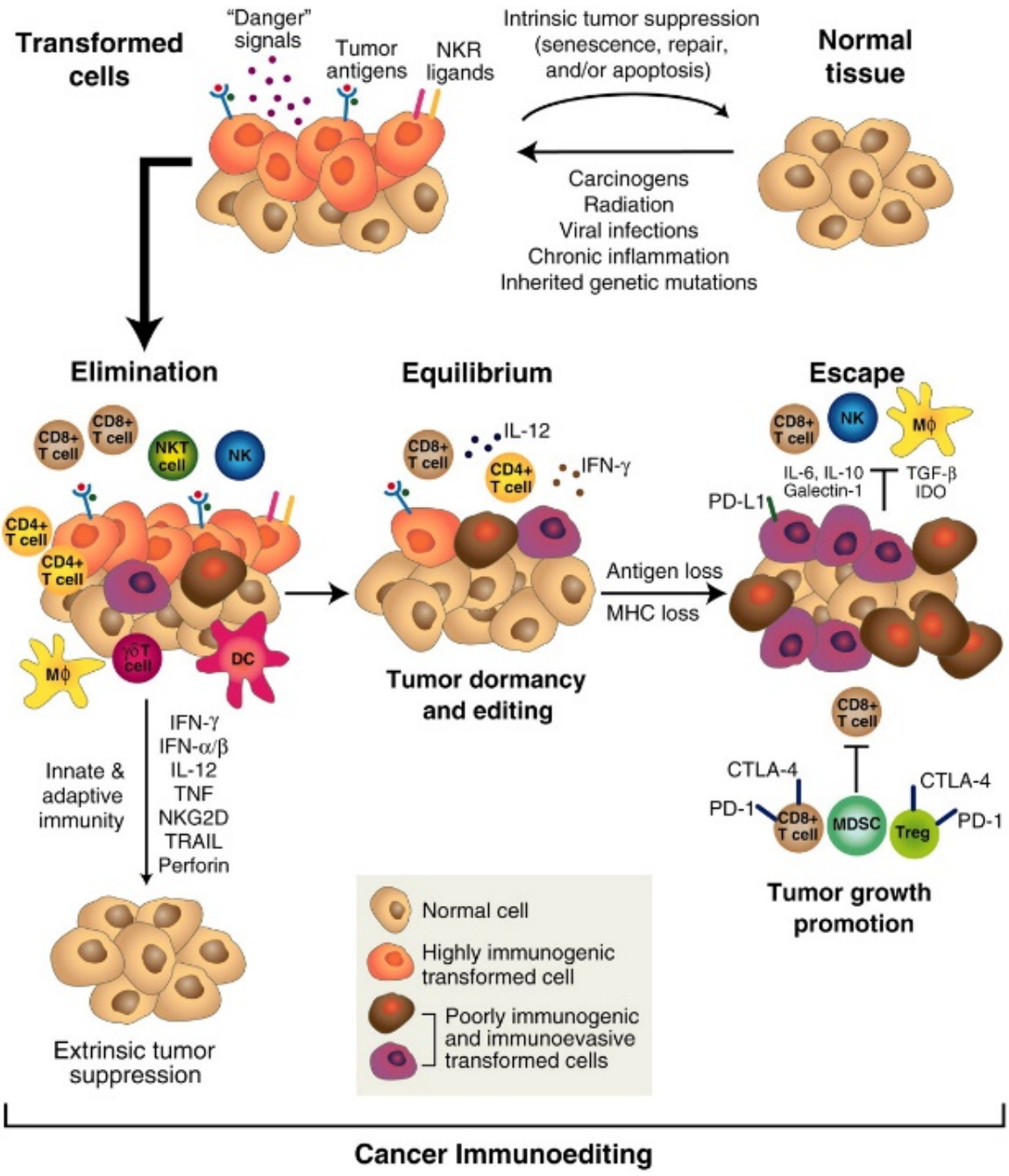
Cataloging each tumor according to its molecular subtype is important to facilitate the development of targeted treatments for patients. For example, since CMS2 and CMS3 CRC patients are poorly immunogenic, further studies would be needed to evaluate new therapeutic strategies that could enhance the antitumor immune response specifically in these molecular subtypes

1.1.2 Cancer Immunoediting

The idea that the immune system could recognize and eliminate nascent malignant cells was first conceived by Paul Ehrlich, in 1909(18). Ehrlich's intuition found experimental confirmation in the 1960s, with the existence of tumor antigens capable of mediate the rejection of transplanted tumors among syngenic mice(19). Ehrlich's original idea was then framed within a formal hypothesis (20,21) and subsequently led to the concept of immunosurveillance against tumors(22).

However, immunosurveillance represents only one dimension of a complex interaction between the immune system and cancer. In fact, the immune system is able to act in all phases of carcinogenesis, exercising not only a protective role but also modifying the immunogenic profile of the tumor, facilitating its tumor growth and spread through a process called "immunoediting"(23–25). This dynamic process consists of three stages: elimination, balance and evasion (**Figure 1**).

In the first stage, elimination, the transformed cells are destroyed by a competent immune system. Sporadic cancer cells that manage to survive immune destruction can then enter a phase of equilibrium in which the immune system controls the development of the tumor but does not eliminate it completely. The evasion phase represents the third and final phase of the process, in which immunologically selected tumors begin to progressively grow, become clinically evident and establish an immunosuppressive tumor microenvironment(26,27).



Schreiber R.D., Science (2011)

Figure 1 Cancer immunoediting. Cancer immunoediting is characterized by three sequential phases: elimination, equilibrium and escape.

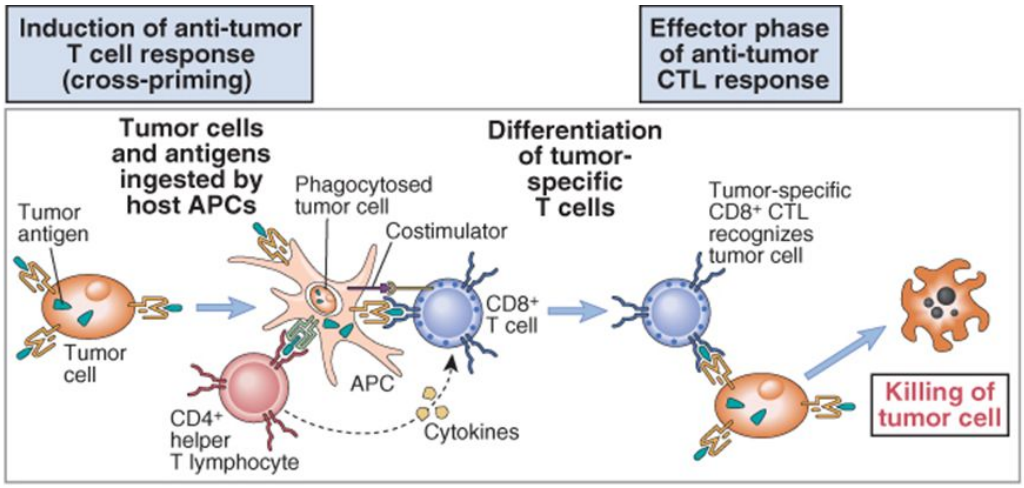
1.1.2.1 Elimination

The elimination phase represents the classic concept of immune surveillance, in which the innate and adaptive immune system recognizes and eliminates cancer cells to prevent the development of the tumor mass. The tumor microenvironment itself is capable of recruiting immune system cells through various mechanisms.

The tumor releases inflammatory molecules that recruit cells of innate immunity, such as natural killer (NK), NKT and T cells, which are stimulated to produce IFNγ. IFNγ acts directly against tumor cells, determining an antiproliferative and pro-apoptotic effect, and indirectly, by stimulating the production of angiostatic cytokines and the recruitment of macrophages M1 and dendritic cells (DCs)(27).

Furthermore, tumor cells can express tumor-specific (selectively expressed by tumor cells) or tumor-associated surface antigens (normal cell protein expressed abnormally or unregulated in tumors) when acquire genetic and epigenetic alterations (19). Antigen presenting cell (APC) can recognize and process these tumor antigens and present them to CD8+ T cells and CD4+ T helper cells (Th), the main effectors of adaptive antitumor immunity (**Figure 2**). APCs express costimulatory B7 molecules capable of providing the signals necessary for the differentiation of CD8+ T lymphocytes into cytotoxic T lymphocytes, capable of recognizing and killing tumor cells in the absence of co-stimulation. More recent studies investigated the role of CD4+ Th cells in antitumor immunity. CD4+ T cells can provide cytokines necessary for the correct development of cytotoxic T lymphocytes; moreover, after activation of the tumor-specific antigen, they release tumor necrosis factor (TNF) and IFN-γ, capable of increasing the expression of MHC class I molecules and the activation of macrophages(28). Furthermore, the tumor itself can release cytokines (INF-α and INF-β)(29) which activate APCs (especially DCs) and promote effector T cells(26).

If the destruction of the tumor cells is complete, the elimination phase represents the end point of the tumor immunoediting process.



Abbas A. et al., Cellular and Molecular Immunology (2014)

Figure 2 T responses against tumor. T-responses against tumor can be induced by cross-priming, in which tumor cells or antigens are processed and presented by APCs to T cells. In some cases, costimulatory molecules expressed by APCs and CD4 helper T cells provide the signals necessary for the differentiation and development of CD8 T cells, which kill cancer cells.

1.1.2.2 Equilibrium

Some variants of cancer cells can survive the elimination phase and enter in the next equilibrium phase, in which the adaptive immune system controls the growth of cancer cells without completely eliminating them. Therefore, the tumor appears clinically invisible and can maintain this latent state for long periods, up to the whole life of the host(30); as such, it may represent a second stable endpoint of tumor immunoediting.

Several analyzes have revealed that only adaptive immunity (specifically CD4 and CD8 T cells) is responsible for maintaining tumor cells in this phase of equilibrium(26,31). However, the adaptive immune system, by exerting selective pressure on transformed cells, may favor the selection of less immunogenic and more surviving tumor clones(26).

1.1.2.3 Escape

During the equilibrium phase of immunoediting, the immunoselected tumor clones develop the ability to elude the antitumor immune response; nowadays this ability is recognized as one of the "hallmarks of cancer"(32). In this phase of evasion, the tumor continues to grow without limitations, until it becomes visible from a clinical

point of view. Immune evasion can occur through various mechanisms which can be divided into mechanisms intrinsic to the tumor cell or extrinsic mechanisms mediated by immune cells with immunosuppressive activity(28).

Intrinsic mechanisms

Tumor cells, in combination with the immunoselection process, select variants of poorly immunogenic tumor cells with the aid of different modalities: (a) through the selection of cells that do not express tumor antigens; (b) through the negative modulation of MHC molecules, which present antigens to effector T cells; (c) through the loss of the components of the antigen processing machinery within the tumor cell necessary to generate the epitope and load it onto the MHC class I molecule (26).

Alternatively, evasion may result from the creation of an immunosuppressive state within the tumor microenvironment(33). Cancer cells can promote the development of this state by producing immunosuppressive cytokines such as vascular endothelial growth factor (VEGF), the growth transformant -β (TGF-β), IL-10 or IDO(26).

Immuno-evasion mechanisms may include the ability to activate mechanisms that inhibit the immune response. Cancer cells are able to increase the expression of the Fas ligand, which by binding to the Fas receptor induces the apoptosis of T lymphocytes. Furthermore, tumor cells are able to activate inhibitory immune checkpoints such as Cytotoxic T-Lymphocyte Antigen 4 (CTLA-4) and PD1, inhibitory receptors present on T lymphocytes. Indeed, PD1 ligand (PD-L1) is expressed by various human tumors and inhibits the activation of lymphocytes. A possible explanation of the role for CTLA-4 in this context is that tumor antigens are presented by APCs with low levels of expression of the costimulatory molecules B7, but sufficient to bind and activate the inhibitory CTLA-4 receptor, because the receptor has a higher affinity(28).

Extrinsic mechanisms

Several cell populations capable of suppressing antitumor immunity have been described within the tumor microenvironment: macrophages, myeloid-derived suppressor cells (MDSCs) and CD4+ T regulatory (Treg) cells.

Tumor-associated macrophages (TAMs) have the M2 phenotype and promote tumor growth and invasiveness by altering the microenvironment, promoting angiogenesis (releasing VEGF and TGF-β), and suppressing T8 cell responses.

MDSCs include different cell types, including DC precursors, monocytes and neutrophils. They suppress innate immune responses, releasing IL-10 which inhibits M1 macrophages and T cell immune responses, and the release of nitric oxide and reactive oxygen species(34).

In our laboratory we are mainly focused on the role of CD4+ regulatory T lymphocytes in dampening anti-tumor immunity through the inhibition of effector T responses(35).

1.2 T regulatory cells

T regulatory cells (Treg cells) were initially identified as a subpopulation of T cells with strong immunosuppressive activity(36). They were called “regulatory” by Sakaguchi and colleagues, who defined Treg cells as a population representing about 10% of CD4+ T cells, expressing high levels of CD25 (α chain of the IL-2 receptor) and able to protect from autoimmune diseases in thymectomised mouse models(37,38). Furthermore, they constitutively express CTLA-4, a negative regulator of activated T cells essential for response suppression in vivo(39,40), and GITR (glucocorticoid-induced TNF receptor), a major inhibitor of Treg cell function(41). CD25, CTLA -4 and GITR are common surface receptors used for Treg labeling and selection. Other markers constitutively activated or highly expressed by Treg cells include LAG-3 (Lymphocyte-Activation Gene 3)(42), CD127(43) , PD1, Helios, galectin-1, OX-40, Ly6, 4-1BB , TNFR2 (TNF Receptor 2), TGF-βR1 (TGF-β Receptor 1) (44) and, more recently, neurophylline 1(45,46). Treg also shows a decrease in CD40L expression, an increase in CTLA-4 expression mediated by TCR activation and production of cytokines only if activated by allogeneic APCs; the cytokines produced are mainly IL-10, TGF-β and low levels of IFN-γ(47).

Treg cells require the action of TGF-β and IL-2 for their development and survival. TGF-β stimulates the expression of the transcription factor FOXP3 (Forkhead box p3), essential for the differentiation of Treg cells(48). IL-2, in addition, activates the transcription factor STAT5 (Signal Transducer and Activator of Transcription 5), which stimulates the expression of the *FOXP3* gene and other genes relevant for Treg functions(49). Since Treg cells are unable to produce IL-2, they depend on an exogenous source produced by CD4 + cells and dendritic cells. IL-2, produced by activated (non-regulatory) T cells, contributes to the maintenance, expansion and activation of Treg cells which, itself, limits the expansion of T cells(50).

To date, the FOXP3 transcription factor is considered the most specific marker of Treg cells, although it can also be transiently expressed by activated T cells(51). FOXP3 plays an important role in both the development and functionality of Treg cells(52–54). In humans, the presence of loss-of-function mutations in the *FOXP3* gene leads to the development of an autoimmune lymphoproliferative disease known as IPEX (Immunodisregulation, Polyendocrinopathy, Enteropathy, X-linked

Syndrome), characterized by activation of CD4 + cells and excessive cytokine production inflammatory(55).

1.2.1 Origin and development of Treg cells

Treg cells can be classified based on their site of origin. In most cases, Treg cells develop from T precursors present in the thymus and are called natural Treg cells (nTreg); differentiation in nTreg cells is supported by high affinity bound between TCR and self-peptides presented by MHC class II molecules and depends on the availability of IL-2(56,57). Otherwise, some Treg cells can differentiate from naive CD4+ T cells in peripheral lymphoid organs, following stimulation with non-self-antigens in the absence of strong activation of innate immunity; in this case they are defined as peripheral Treg cells (pTreg cells). Differentiation in pTreg cells is supported by high concentrations of TGF-β and retinoic acid, which promote FOXP3 expression in peripheral naïve T lymphocytes CD4+(56,58). In the gut, short-chain fatty acids produced by commensal bacteria support the development of pTreg cells(59). It has also been suggested that pTreg differentiation is important in the maternal-fetal tolerance(60). Overall, these evidences suggest that nTreg cells are involved in maintaining self-tolerance, while pTreg cells inhibit immune responses to harmless non-self antigens, such as commensal bacteria, environmental antigens, and food antigens; in fact, a defect in the generation of pTreg cells leads to allergy and inflammation in the intestine and lungs(61).

1.2.2 Treg cells in immunological context

1.2.2.1 Treg Functions

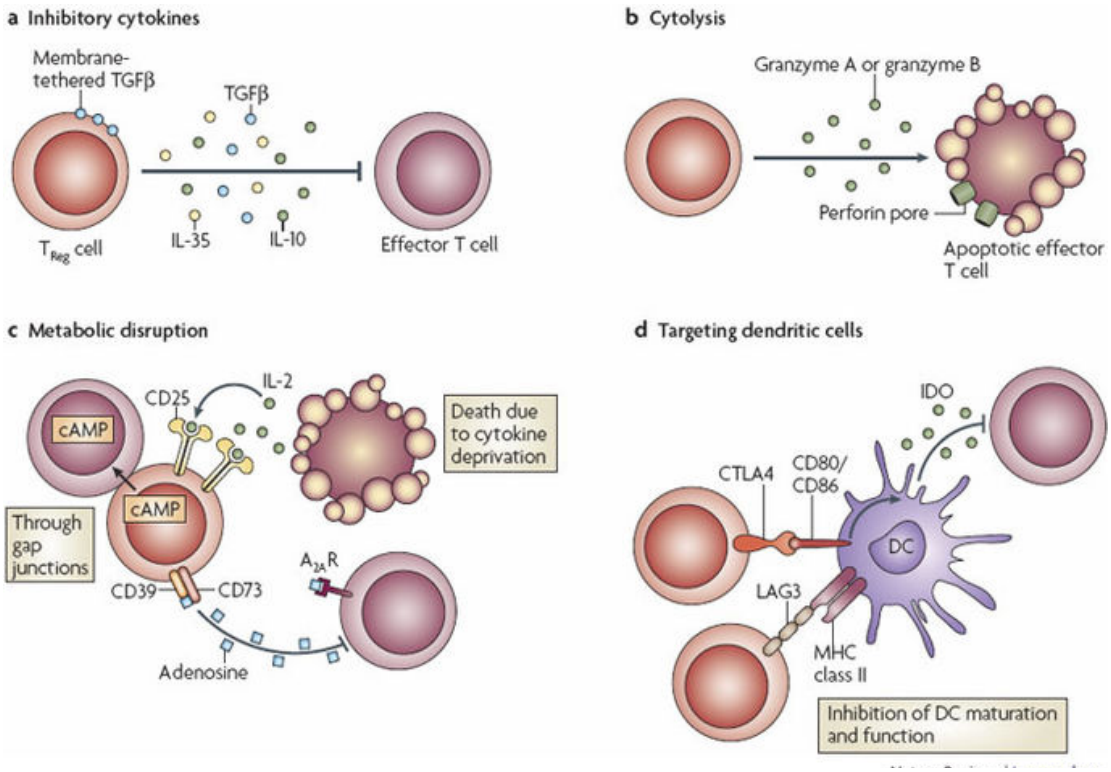
Treg cells play an important role in suppressing immune responses and maintaining immune homeostasis. Treg cells are key mediators for the maintenance of immune tolerance, both towards autologous and allogeneic antigens. Indeed they are involved in the peripheral tolerance mechanism, inhibiting autoreactive T lymphocytes, and play a central role in the prevention of autoimmune diseases, such as type I diabetes(62). They are able to limit chronic inflammatory diseases, such as asthma and intestinal inflammation diseases and are very important in maintaining homeostasis of the intestinal microbiome(63). Inevitably, the

suppressor mechanism itself can lead to negative consequences, such as inhibition of antitumor immune responses, contributing to tumor progression(64).

Treg cells may also perform functions beyond their suppressor role: they are essential in the non-lymphatic tissue protection and, based on the tissue localization, they can adopt specialized functions and exhibit unique phenotypes and transcriptional profiles. Currently, one of the best characterized examples is the Treg population infiltrating visceral adipose tissue, capable of capturing lipids and modulating metabolism in adipose tissue thanks to the high expression of the transcription factor PPAR-γ (Peroxisome Proliferator-Activated Receptor γ) and CD36 scavenger receptor(65,66). It has recently been shown that Treg cells have the ability to preserve the functional and structural integrity of the lung during a viral infection (67)and are involved in muscle regeneration following tissue damage(68). Furthermore, Treg cells have been described as neuroprotective, modulating oxidative stress and inflammation in microglial cells, improving neurodegeneration in Parkinson's disease(69).

1.2.2.2 Treg suppressive mechanisms

Treg cells can exert their suppression function through several mechanisms: a) by inhibitory cytokines; b) by cytolysis; c) by metabolic disruption and d) by modulation of dendritic cell (69) (**Figure 3**).



Vignali D.A.A. et al., Nature reviews. Immunology (2008)

Figure 3 Mechanisms of Treg-mediated cell suppression. Schematic representation of the four main suppressive mechanisms used by Treg cells: a) production of inhibitory cytokines, such as TGF-β; b) cytolysis by release of perforins and granzymes; c) metabolic disruption, such as IL-2 depletion for effector T cells and the release of adenosine; d) modulation of the activity of dendritic cells, through the CTLA-4-B7 interaction that allows the production of IDO and LAG3-MHC class II.

Suppression by inhibitory cytokine production

Treg cells exert suppressive activity on several immune cell types by secreting high levels of inhibitory cytokines, such as IL-10(70), IL-35 (71)and TGFβ(72).

IL-10 has been described as the major immunomodulatory cytokine. It inhibits the production of inflammatory cytokines, such as IL-12, and acts on effector T cells and DCs by binding to IL-10 receptor. On effector T lymphocytes it exerts a suppressive function inhibiting their proliferation; on DCs further promotes the production of IL-10. It also promotes phagocytic activity by elimination of cellular debris at inflammatory sites. The inhibitory effect of IL-10 plays a key role in peripheral tolerance to self-antigens, allergens and also tumor antigens(73,74).

TGF-β is a multifunctional cytokine, able to regulate several immune cells: (a) it inhibits the IL-12 production, (b) suppresses effector T cells differentiation and proliferation and (c) promotes T cell differentiation into Treg, inducing the expression

of FOXP3, or into Th17, in the presence of IL-6. It also inhibits macrophages, dendritic cells and NK cells activity. TGF-β is also expressed on the cell surface of Treg cells, allowing the interaction between Treg cells and target cell(73,74).

IL-35 is an immunomodulatory cytokine, mainly expressed by Treg cells. It suppresses Th cells proliferation and promotes naïve T cells differentiation into Treg cells with strong suppressive activity(71,75).

Suppression by cytolysis

Similar to NK cells and CD8+ T lymphocytes, Treg cells exhibit cytotoxic activity and are able to suppress effector T cells by inducing apoptosis thanks to the release of perforin, granzyme A or granzyme B(76). Perforin molecules form a transmembrane pore across the lipid bilayer through which granzymes can enter target cells. Furthermore, granzymes can also gain access to target cells via endocytosis. Within the cell, granzymes can induce apoptosis through caspase-dependent or caspase-independent mechanisms(77).

Suppression through metabolic disruption

Treg cells express high levels of the IL-2 surface receptor (CD25) and through local IL-2 depletion necessary for effector T cells survival induce their apoptosis (78). The selective expression of CD39 and CD73 ectoenzymes on the cell surface of Treg cells has been shown to induce the production of the inhibitory molecule adenosine, which binds and activates the adenosine receptor 2A, suppressing effector T cells(24,79,80). Treg cells have also been shown to suppress the effector responses by transferring cAMP (cyclic adenosine monophosphate) inside the T cells through gap junctions(81).

Suppression of dendritic cell modulation

Treg cells can modulate the maturation and function of dendritic cells, which are required for the activation of effector T cells. CTLA-4, expressed constitutively on the surface of Treg cells, has been shown to bind CD80 and CD86, expressed on DCs. This interaction induces IDO production by DCs which catalyzes the conversion of the tryptophan into kynurenines. IDO production correlates with the inhibition of immune responses both by depletion of tryptophan from the microenvironment, which is necessary for T cell proliferation and expansion, and by the release of metabolites that can cause apoptosis of T lymphocytes(82).

Furthermore, LAG-3, also known as CD223, is a CD4 homolog able to bind with high affinity to the class II MHC molecules expressed by immature DCs. This interaction suppresses the maturation of DCs and their immunostimulant capacity(83)

1.2.3 Tumor infiltrating T reg cells

Treg cells are significantly enriched within the tumor mass, making up 20-30% of total CD4+ cells(84). This enrichment is particularly important as Treg cells suppress in a dose-dependent manner. It was initially proposed that the presence of tumor-infiltrating T lymphocytes (TILs) within the tumor microenvironment correlated with a good prognosis(85); it is by now established that it is rather the balance between effector T cells and Treg cells that determines the tumor pathology outcome(86). In fact, the decrease in the CD8+/Treg ratio, and the increase of tumor-infiltrating Treg cells (tiTreg cells), is associated with a poor prognosis in several cancers, including breast cancer(87) and NSCLC(88). However, with respect to CRC, the role of Treg cells is controversial: in some studies the infiltration of Treg into the tumor microenvironment correlates with a better prognosis(89), while other studies have shown the opposite(90,91). These conflicting results could arise from an improper interpretation of the heterogeneity of the Treg cell population with different roles. In fact, Saito and colleagues detected during CRC development the presence of two Treg populations with different suppressive capacity(92).

1.2.3.1 Mechanisms of Treg accumulation in tumor microenvironment

Different mechanisms promote the accumulation of Treg cells within the tumor microenvironment: (a) increase in recruitment, (b) expansion and (c) de-novo differentiation. Expansion and differentiation can occur locally, within the tumor microenvironment, or distantly, in the lymph nodes that drain the tumor(84).

Several evidences suggest that Treg cells show a greater infiltration capacity within the tumor microenvironment than the T effector cells(84,93). This preferential recruitment depends on mechanisms driven by chemokines, cytokines with chemotactic activity. Indeed, Treg cells express different chemokine receptors, such as CCR4 and CCR8. Tumor cells and TAMs have been shown to be able to recruit Treg into the tumor microenvironment by releasing CCL22 and CCL17 that are

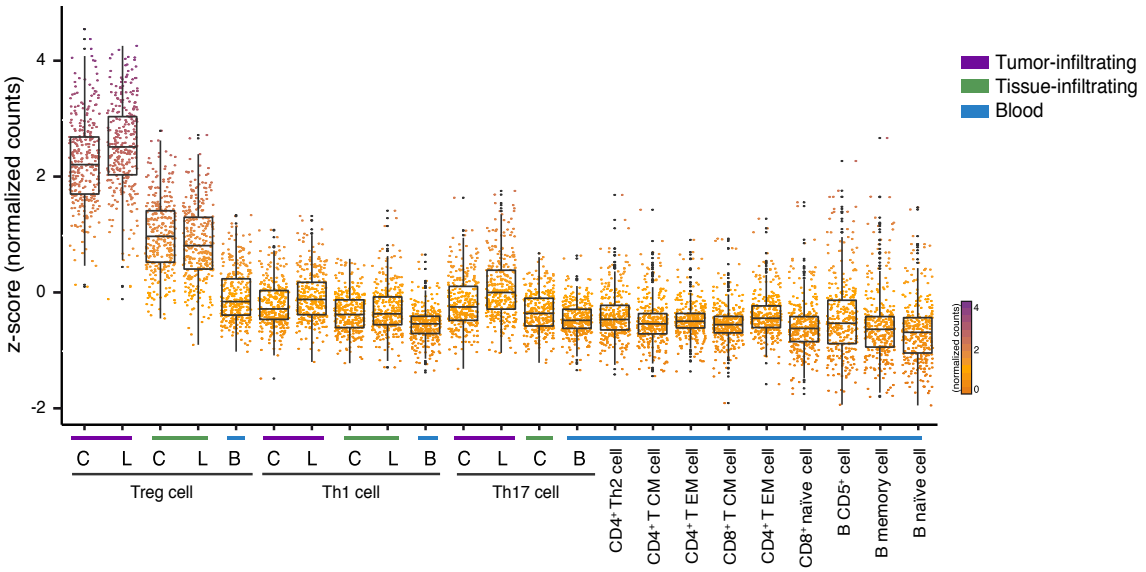
CCR4 ligands. Indeed, CCL22 blockade has been shown to play a role in decreasing Treg infiltration in ovarian cancers(94).

A second mechanism is represented by cytokine-mediated Treg expansion. In this context proliferation of Treg cells in the tumor microenvironment is promoted by the release of IL-2 from effector T cells (93). Other cytokines have been involved such as TGF-β, released by DCs, that plays a role in the expansion of Treg cells in draining lymph nodes in mouse models of cancer(95).

Finally differentiation of CD4+CD25- T cells into Treg(96)has been reported. Indeed, tumor cells themselves may contribute to the induction of Treg cells through the release of high levels of TGF-β(97) and IDO(98), suggesting that tumors exploit different complementary mechanisms to effectively elude the immune response.

1.2.3.2 Molecular profile of tumor infiltrating Treg cells

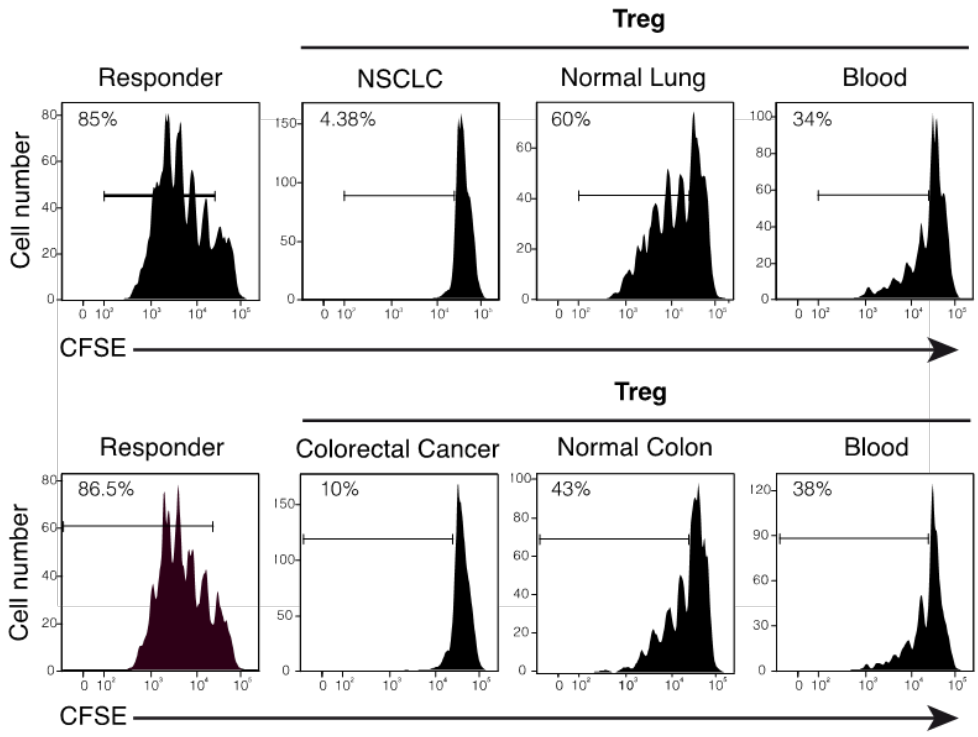
A recent study in our laboratory focused on the transcriptional characterization of human Treg cells infiltrating CRC and NSCLC(99). In this study, the gene expression profiles of tiTreg cells were compared with Treg cells infiltrating healthy tissues and peripheral blood isolated from the same patients. This comparison revealed a striking similarity between tiTreg from different tumors and a molecular signature of more than 300 genes that specifically characterize them (**Figure 4**). The more selectively expressed genes (LAYN, MAGEH1 and CCR8) appear to be negatively correlated with the patient's prognosis, confirming the pro-tumoral role of Treg cells. While LAYN and MAGEH1 function in Treg cells is still uncharacterized, CCR8 is a receptor that plays a role in the recruitment and migration of Treg cells in tissues. Among the genes specifically expressed by tiTreg cells, several immune checkpoints (GITR, OX40, TIGIT, LAG3 and TIM-3) were found with some of the respective ligands (OX40LG, Galectin-9 and CD70) suggesting a higher activation state compared to the counterpart of healthy tissue and peripheral blood.



De Simone M. et al., Immunity (2016)

Figure 4 tiTreg gene signature. This plot shows the expression values (z-score) of genes that are preferentially expressed (Wilcoxon Mann Whitney test p < 2.2 × 10–16) in Treg cells infiltrating CRC (C) and NSCLC (N) compared to peripheral blood (B) and other immune cell subsets isolated from peripheral blood.

Functional studies were also conducted on these cells demonstrating a strong suppressive potential compared to Treg cells isolated from the normal adjacent tissue and from the peripheral blood of the same patient (**Figure 5**). Many of these genes, including CCR8, have also been confirmed by other studies performed on different kind of cancer, such as breast cancer(100).



De Simone M. et al., Immunity (2016)

Figure 5 Treg *in vitro* suppressive assay. These flow cytometry histograms show the *in vitro* suppressive assay for Treg cells isolated from NSCLC (top) and colorectal cancer (bottom) patients. CD4+ naive T cells (Responder) were labeled with carboxyfluorescein diacetate succinimidyl ester (CFSE), *in vitro* activated with anti-CD3 antibody plus DCs and cocultured for 4 days with an equal number of Treg cells isolated from tumor tissue (second column), its adjacent healthy counterpart (third column) and peripheral blood (fourth column) of the same patient. Suppression activity correlates with the progressive dilution of fluorescent dye. Percentage of proliferating cells is indicated.

1.2.3.3 Treg as novel therapeutic tools for cancer treatment

As Treg cells have been found with high frequency in different tumor types and are potent suppressors of effector T cells specific for tumor antigens, they are emerging as potential targets for immunotherapy.

Several strategies have been proposed and attempted toward Treg cell depletion and the inhibition of their suppressive functions to reduce tumor burden. Candidate molecules for Treg-specific immunotherapy include CTLA-4, PD-1, GITR, CCR4, OX-40, LAG3, and CD25(101).

CTLA-4 is constitutively expressed by Treg and plays a vital role in their suppression activity. Recent studies have suggested that the efficacy of the anti-CTLA-4 antibody

is determined not only by the enhancement of the effector activity of T cells, but also by the depletion of Treg cells in the tumor microenvironment by antibody-mediated cellular cytotoxicity(102) or, moreover by the destabilization of Treg cell function in glycolysis-low tumors by promoting CD28 signalling(103).

Alternatively, impairment of Treg cell recruitment by blocking chemokine receptors has been proposed. As a way of example, a monoclonal antibody against the CCR4 chemokine receptor(104) has been developed to inhibit the infiltration and accumulation of Treg cells within the tumor microenvironment; however, the redundancy and promiscuity of chemokine systems could impair therapeutic efficacy and specificity.

Overall immunotherapy based on immune checkpoint inhibitors and tumor Treg reduction in the tumor microenvironment has generated promising clinical results, but there are some critical issues that still need to be addressed. Indeed, most of potential targets are, to some extent, shared between Treg cells and effector T cells. Therefore, Treg cell targeting could lead to depletion or functional impairment of effector T cells. Furthermore, Treg cells are physiologically involved in maintaining self-tolerance and immune homeostasis, therefore novel therapeutic targets should be selected to avoid systemic exhaustion of Treg cells to avoid severe autoimmunity(105).

Our current knowledge underlies the importance of characterizing the molecular features of tumor infiltrating Treg cells to reveal both specific gene expression patterns and the regulatory networks that shape Treg identity in the tumor microenvironment to identify specific and effective approaches for their functional modulation.

1.3 Epigenetic regulation of gene expression

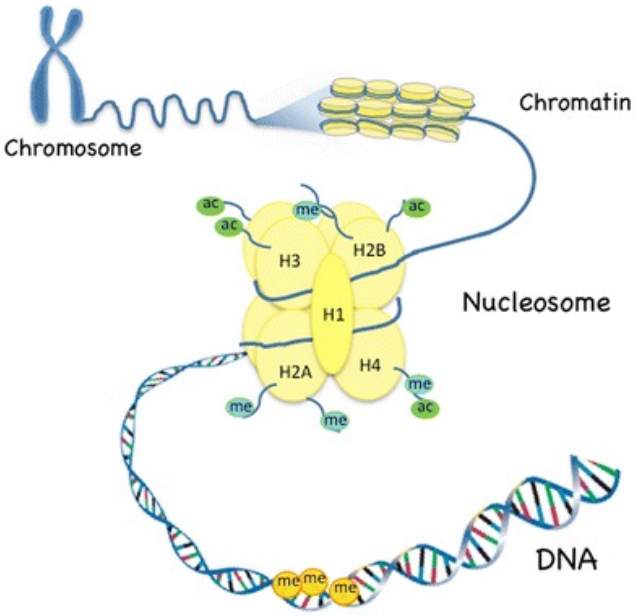
Considering all development stages, in our organism there are at least 200 different cell types, each with its own specialized structure and functions. However, with rare exceptions, cells of the same organism have the same genome, or rather nuclear DNA is made up of the same nucleotide sequence(106). The difference between distinct cell types is determined by the transcriptome, defined by both protein coding and non-coding RNAs. The gene expression profile is coordinated by a series of epigenetic mechanisms which finely regulate gene expression both from the qualitative point of view (which genes to transcribe) and from the quantitative point of view (how much to activate or how much to repress the transcription); these epigenetic mechanisms can be faithfully transmitted to daughter cells, after each cell division, in a process called "epigenetic memory" of the cell identity. The term "epigenetics" has taken on multiple meanings over the years, describing very different phenomena (107). In 1942 Conrad H. Waddington was the first to coin the term "epigenetics", to indicate the mechanisms that lead to changes in the phenotype without altering the genotype(108,109). Currently known and described epigenetic mechanisms include reversible chemical modifications of DNA and associated proteins, which regulate the activation of genes through chromatin remodeling(110). These mechanisms include chromatin accessibility, nucleosomes remodeling, DNA methylation, histone modifications and the regulatory action of particular non-coding RNAs (miRNA, siRNA, piRNA, lncRNA)(111). All these mechanisms are essential for physiological cell fate transition and maintenance of tissue specific gene expression. In the next paragraphs we will provide an overview on specific epigenetic processes (histone modifications and cis-regulatory elements) relevant to the biological questions we are addressing in this project.

1.3.1 Chromatin structure and organization

In humans, genomic DNA is associated with histone and non-histone proteins, forming a polymer called chromatin organized into 23 pairs of chromosomes. Considering all the chromosomes, the DNA molecule of each single cell is about 2 meters long; the genome is subject to different degrees of compaction (**Figure 6**) since it is confined within the cell nucleus, whose diameter measures only 6 μm(112).

The first level of chromatin organization begins with the wrapping of 145-147 base pairs (bp) of DNA around a histone octamer, forming the central nucleosomal particle. The histone octamer is composed of four different histone proteins: H2A, H2B, H3 and H4; specifically, it consists of the tetramer $(H3)_2(H4)_2$ and two H2A-H2B heterodimers(113). The histone octamers are connected to each other by a stretch of DNA (DNA linker) of variable length (10-70 bp) and are associated with histone H1, which allows a more compact and orderly structure of chromatin(114). The central body histones have a N-terminal tails which undergo a various of covalent post-translation modifications, called histone modifications, which impact on the chromatin structure and consequently on the state of transcriptional activation or inactivation. Chromatin in general can appear in more or less condensed forms as a consequence of the different degree of compaction of the DNA around protein complexes called histones.

The condensed state of chromatin limits accessibility to genomic sequences, creating intrinsic barriers for DNA transcription and replication. For this reason, chromatin is an extremely dynamic but controlled structure, both in metaphase (mitosis), where it assumes the condensed form, and in interphase(115).



Larsson L., Current Oral Health Reports (2017)

Figure 6 Overview of the chromatin structure and compaction. The genome is exposed to different degrees of compaction thanks to the help of histone proteins (H2A, H2B, H3, H4 and H1) that wrap and fold the DNA. Further increasing levels of compactness form to the highly condensed mitotic chromosome.

1.3.2. DNA methylation

The DNA methylation consists in the addition of a methyl group (-CH3) on a nucleobase of the DNA, without changing the sequence. The best-studied DNA methylation is the methylation of cytosine to form 5-methylcytosine; this reaction is catalyzed by DNA methyltransferase (DNMT), an enzyme that uses S-adenosyl methionine (SAM) as a donor of the methyl group. DNA methylation exist mostly in the CpG dinucleotide and 70-80% of CpGs are methylated in human genome(116).

DNA methylation can be mitotically inherited and plays an important role in repression of gene transcription, especially when distributed around the transcription start sites (TSS)(117). DNA methylation is crucial for embryonic development(118) and in tumoral cell the methylation pattern is altered compared to the normal tissue; typically, there is hypermethylation of tumor suppressor genes and hypomethylation of oncogenes(119).

1.3.3 Histone modifications

Histone proteins are exposed to a series of post-translational modifications that regulate the degree of compaction and the accessibility of chromatin. These modifications consist in the reversible covalent addition of chemical groups or small proteins on the amino acid residues of the tails or of the globular nucleus of histone proteins. Histone modifications include a) methylation of arginine (R) residues; b) methylation, acetylation, ubiquitination, ADP-ribosylation and sumoylation of lysine (K) residues; phosphorylation of serine (S) and threonine (T)(120). With the advancement of technologies for the study of proteomics, the number of identified histone modifications keep growing and new histone modifications, such as propionylation and butyration, have recently been described(121).

The enzymes actively involved in the insertion or removal of the histone modification are known as "writer" and "eraser" respectively; while proteins capable of binding and interpreting histone modification are called "readers" (**Figure 7**)(122). Based on their nature and location, histone modifications can be associated with transcriptional repression or activation (**Figures 8-9**).

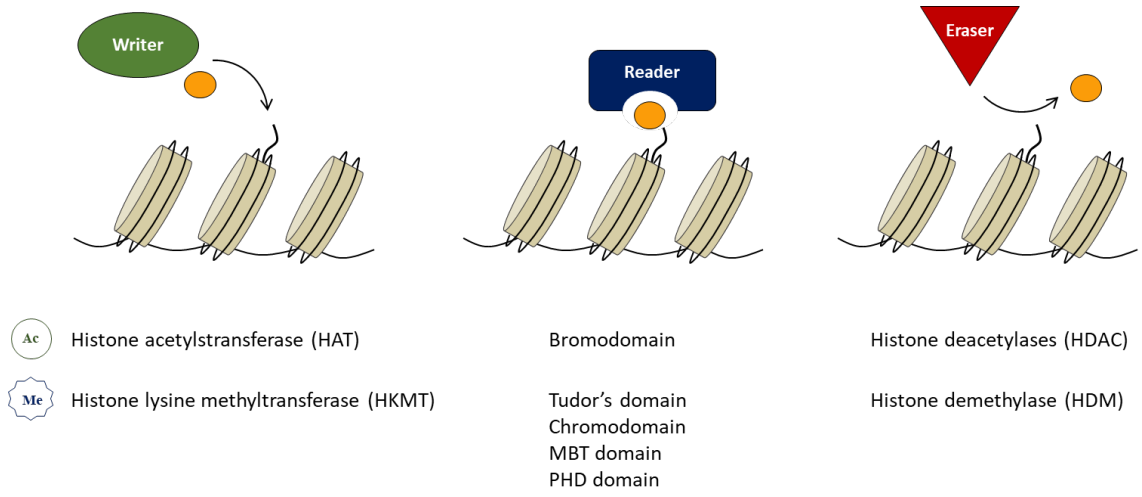


Figure 7 Proteins involved in histone modifications. Overview of proteins actively involved in the addition (writer), removal (*eraser*) and interpretation (*reader*) of post-translational modifications.

1.3.3.1 Histone acetylation

The histone acetylation consists in the addition of an acetyl group (-COCH3) (ac) on a lysine residue at the N-terminal of a histone protein. This reaction is catalyzed by

histone acetyltransferase (HAT), an enzyme that uses acetyl-coenzyme A as a donor of the acetyl group. HATs are classified into two categories: nuclear HATs, responsible for the acetylation of histones assembled in chromatin, and cytoplasmic HATs, responsible for the acetylation of newly translated, not yet assembled histones(123).

Acetylation of lysine residues removes a positive charge from the histone, reducing electrostatic interactions with negatively charged DNA. This mechanism allows a relaxation of the chromatin which facilitates the access of the transcription factors and of the entire transcription machinery. This is why histone acetylation is always associated with transcriptionally active regions and its lack is associated with transcriptionally repressed ones(124).

On the other hand, histone deacetylase (Histone Deacetylase, HDAC) catalyzes the reverse reaction, removing the acetyl group from the histone protein, inhibiting the transcription process(125). Acetylation of histones generates binding sites for proteins containing domains known as bromodomains(123). The bromodomain-containing protein can recruit other proteins, including nucleosome remodeling complexes or other enzymes, which act on DNA(126).

One of the best known and described modifications is the acetylation of lysine 27 on histone H3 (H3K27ac). This modification is associated with transcriptionally active regions and is mainly present on the promoter of the transcribed genes; in co-presence with the monomethylation of lysine 4 of histone H3 (H3K4me1), it describes particular regulatory regions called enhancers. Enhancers are key regulatory elements that act in cis by increasing the transcription frequency of the gene they control. Enhancers do not need to be closed to their promoter and can perform their function several thousand bases from the gene they control(127).

1.3.3.2 Histone methylation

Histone methylation consists of the transfer of a methyl group (-CH3) (me) onto a lysine or arginine residue at the N end of a histone protein, usually H3 and H4. This reaction is catalyzed by histone methyltransferase, an enzyme that uses S-adenosyl-methionine (SAM) as a donor of the methyl group. Arginine can be mono or dimethylated, while lysine can be mono, di or trimethylated(124).

The enzyme responsible for the methylation of lysine is histone lysine methyltransferase (histone lysine methyltransferase, HKMT) which generally carries out its enzymatic activity through the catalytic domain SET (SUV39, Enhancer of zeste, Trithorax)(128,129). There are several HKMTs, specific for a given lysine of a given histone. For example, the EZH2 protein (subunit of the polycomb 2 repressive complex, PRC2) is a specific methyltransferase for histone H3 (H3K27) lysine 27(130); G9a and Suv39h are specific for H3K9(131,132), Set1 is specific for H3K4(133) and Setd2 for H3K36(134).

On the other hand, histone demethylase (histone demethylase, HDM) catalyzes the reverse reaction, removing the methyl group from the histone protein. HDM enzymes are classified into two families based on enzyme activity. The first family, LSD (Lysine-Specific Demethylase), consists of the FAD-dependent amino oxidase (Flavin Adenine Dinucleotide), which can act only on mono and dimethylated lysines. The second family consists of oxygenizes containing Jumonji C (JmjC) domains, which can also act on trimethylated lysines(135). Similarly to HKMTs, HDMs show relative specificity for the methylated lysine residue.

Methylated lysines are recognized by several proteins; these include proteins containing chromodomaines, Tudor domains, PHD (Plant Homeodomain) domains, MBT (Malignant Brain Tumor) domains, PWWP domains, and WD40 domains(136). The same methylation can be recognized by several reader proteins.

Unlike acetylation, histone methylation does not modify the charge of proteins and induces structural changes in chromatin through electrostatic mechanisms. Methylation is associated with both transcriptional activation and repression, based on the lysine involved and the degree of methylation and the type of proteins they are able to recruit. In general, the tri-methylation of H3K4 and H3K36 mark active regions(137,138), while the methylation of H3K9 and H3K27 mark repressive regions (**Figure 8**)(130,132,139).

A very relevant type of modification is the H3K4me1; it is mainly distributed towards the 3' of actively transcribed genes; alone it is related to weakly active enhancers(140,141).

H3K4me3 is mainly distributed around TSS and the promoter of actively transcribed genes; it promotes transcription through interactions with RNA polymerase II(142).

H3K4me3 is associated with the H3K27ac on the active promoter. Several studies suggest that H3K4me3 recruits HAT by promoting histone H3 acetylation(109,143).

H3K36me3 is distributed along the body of active transcribed genes, with an increase towards the 3' of the gene. This peculiar distribution derives from the fact that for histone lysine methylation Setd2 is associated with RNA polymerase II, predominant in this area. H3K36me is able to recruit the FACT complex (Facilitates Chromatin Transcription)(144), chaperone involved in chromatin remodeling, and PRC2 (Polycomb Repressive Complex 2)(145), a protein complex responsible for the H4K4me3, that restore a repressed chromatin configuration and prevent spurious transcription after stretching. H3K36me3 is related to the regulation of histone acetylation, through the recruitment of HDAC to active transcription sites(143).

H3K27me3 distributes from the promoter, where it has a major peak, continuing along the body of transcriptionally repressed genes. H3K27me3 therefore prevents acetylation of lysine at the level of the promoter and of specific regulatory sequences called enhancers

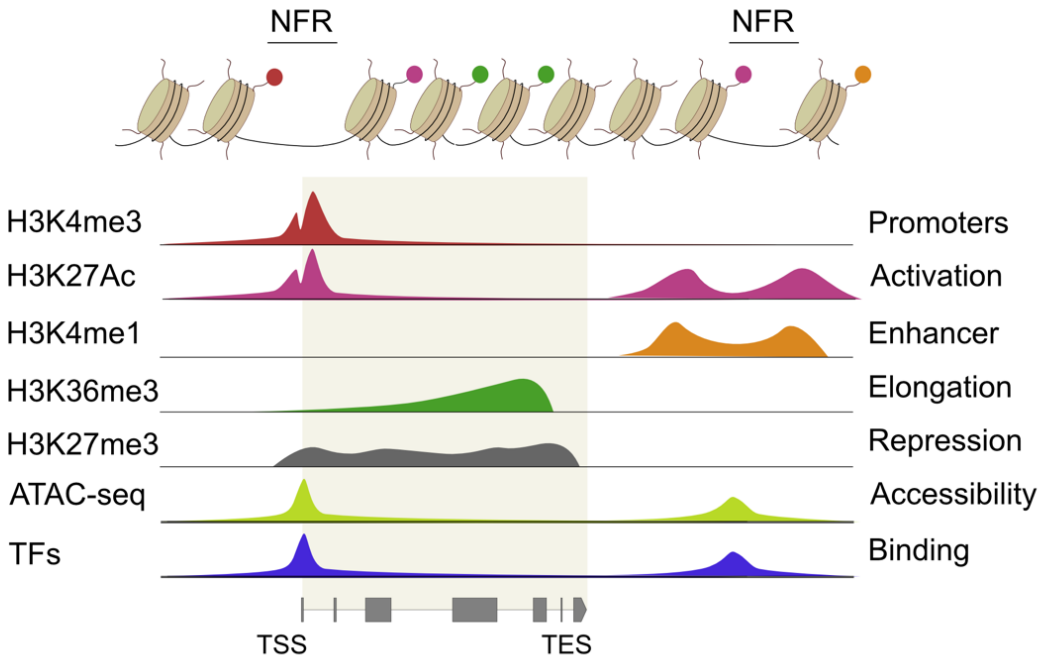
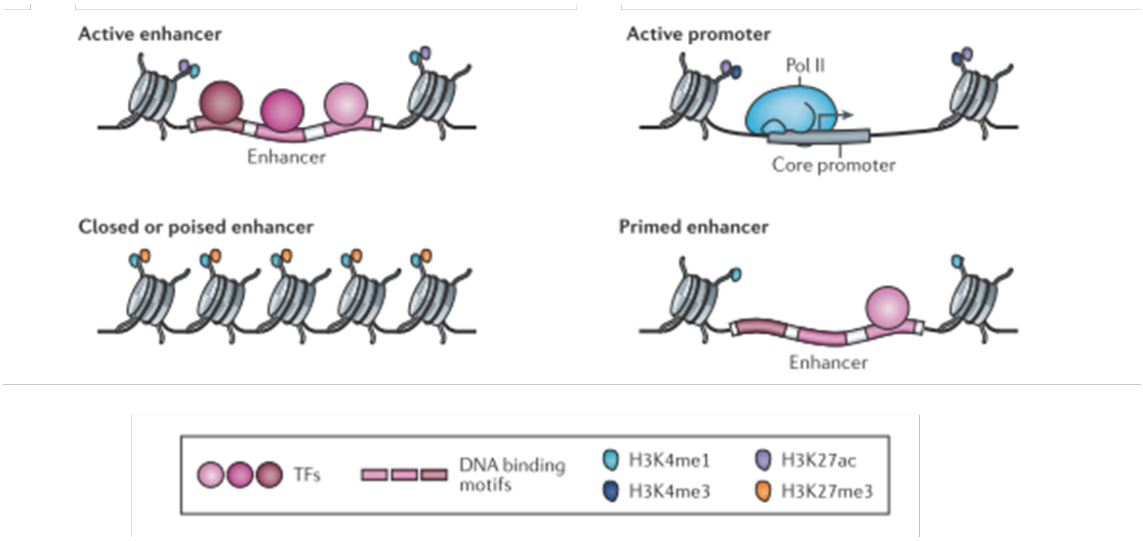


Figure 8 Overview of histone modification distribution. Schematic overview of histone modification distribution on representative gene: each line reports the name and the role of histone modification, followed by the accessible regions detected by ATAC-seq, and finally the profile of representative transcription factor (TF). NFR: nucleosome-free region; TSS: transcription start site; TES: Transcription end site.

In embryonic stem cells, developmental genes are in a so-called poised state: silenced but ready to be transcribed. This state is defined by the coexistence of H3K4me3 (activating change) and H3K27me3 (repressive change) at the TSS level. This type of chromatin state, while on the one hand guarantees gene silencing, on the other hand allows rapid transcriptional activation, following specific evolutionary signals, when H3K27 is demethylated; genes that must be kept inactive will instead keep H3K27me3 at the expense of H3K4me3(146).

Within enhancers regions, H3K4me1 in the presence of H3K27ac or H3K27me3 makes it possible to discriminate between active or poised enhancers respectively (**Figure 9**)(140,141). Furthermore, H3K27Ac, which is related to an active transcription, can be found in association with H3K4me3.



Adattata da D. Shlyueva, G. Stampfel and A. Stark, Nature (2014)

Figure 9 Histone modifications and regulatory elements. The copresence of H3K27ac and H3K4me1 is associated with active enhancer, H3K27ac e H3K4me3 are associated with promotor od transcribed genes, H3K4me1 e H3K27me3 with poised enhancer and H3K4me1 alone is associate with weak enhancers.

1.3.4 Epigenetic landscape of Treg cells

Treg cells begin to establish a specific epigenetic landscape in the thymus, from the earliest stages, before the expression of FOXP3(147,148); during the acquisition of non lymphoid tissue Treg phenotype, Treg cells show specific chromatin accessibility profile based on the tissue in which they reside(149). As discussed above, gene expression is regulated by epigenetic mechanisms, mainly including DNA methylation, histone methylation and acetylation.

Treg cells show a specific DNA methylation pattern which distinguishes them from naïve T cells. In particular, Treg cell lineage is associated with DNA hypomethylation within the FOXP3 locus, specifically the conserved non coding sequence 2 (CNS2), that is necessary for the stable expression of FOXP3 itself (150,151). Moreover, Treg cells are characterized by a specific DNA hypomethylation also at Treg cell-associated gene loci, such as FOXP3, CTLA-4, TNFRSF18 (GITR) and IKZF2 (Helios), which contributes to the stable maintenance of Treg phenotype and function(152).

Histone methylation seems not to be specific for Treg cell lineage (152). However, Treg cells show the largest unique H3K4me3 and H3K27me3 islands compared to naïve T cells (153). In activated Treg cells, Foxp3-bound sites show an enrichment for H3K27me3(154), and conversely Treg cells that lose their Foxp3 signature show a decrease of H3K4me3 within the downregulated Treg cell signature genes, such as FOXP3 and CTLA-4, and an increase of H3K4me3 within the th2-associated genes, such as IL-4 and Il-5(155).

About histone acetylation, H3K27ac modification in foxp3 locus seems to be important for Treg differentiation and function. Indeed, only Treg cells shows H3K27ac deposition at the FOXP3 promoter(156).

In conclusion, Treg cell development is a complex mechanism defined by the integration of multiple cues involved in determining FOXP3 induction and Treg-specific epigenetic landscape.

2 AIM OF THE PROJECT

Tumor-infiltrating lymphocytes (TILs) constitute a major component of immune infiltrates in several tumors and are crucially involved in cancer progression. However, the molecular and epigenetic changes that underlie the phenotypic diversity of TILs remain largely unknown. In particular, tumor-infiltrating Treg cells (tiTreg cells) have been correlated with poor prognosis in many cancer types, indicating that selective targeting of tiTreg cells could provide therapeutic benefits. In this respect, the tumor microenvironment is shown to imprint distinct transcriptional features upon tiTreg cells of different tumors compared to activated Treg cells found in corresponding normal adjacent tissues or in the peripheral blood. Characterization of the epigenome can therefore provide insights into the activation of specific transcriptional programs and the regulation of gene expression.

Based on these premises, with this project we seeked to unravel the chromatin landscape of human Treg cells infiltrating CRC and NSCLC by employing ChIPmentation and ATAC-seq techniques in order to identify the regulatory networks that define the Treg cell state shaped by the tumor microenvironment.

The project aims to develop an integrated multi-omics approach in order to pursue the following specific goals:

i) draw a detailed picture of tiTreg cells epigenetic landscape by ChIPmentation and ATAC-seq;

ii) accurately map active and repressed regulatory elements (enhancerome) through de novo chromatin state discovery combined with chromatin accessibility datasets in Treg across diseased and normal tissues;

iii) reconstruct crucial nodes of the regulatory network defining the cell state of tiTreg cells by footprinting analysis;

iv) assess the effect of key regulatory regions modulation on tiTreg cell identity toward the identification of novel potential therapeutic targets.

3 MATERIALS AND METHODS

3.1 Treg isolation from human primary tissues

Treg cells were isolated from tumors (NSCLC or CRC), non-neoplastic counterparts and peripheral blood of the same patients, and from peripheral blood from healthy donors.

Human lung and colorectal biopsies were obtained respectively from Humanitas Research Hospital (Rozzano) and San Paolo Hospital (Milan), following ethical approval from their Institutional Review Boards. Informed consent was obtained from all patients prior to acquisition of the samples. Samples were confirmed to be tumor or normal based on pathologist assessment and were obtained prior to treatment.

3.1.1 Dissociation of human primary tissues

Fresh tissues were processed by mechanical and enzymatic dissociation in order to obtain a single-cell suspension. The tissue dissociation was necessary for solid tissues such as NSLC tissue, CRC tissue and their non-neoplastic counterparts. All steps were performed under sterile conditions**.**

3.1.1.1 Colon tissue dissociation

Colon specimens were cut into pieces of 5-7 mm with disposable scalpel and incubated in 50 ml of 0,16 mM DTT (Sigma-Aldrich) in completed HBSS (ThermoScientific) (Gentamicina 1:200; Penicillin-Streptomycin (Sigma-Aldrich)) 10 minutes at room temperature on magnetic stirrer. Then they were washed twice in 25 ml of completed HBSS for 5 minutes, and incubated twice in 30 ml of 1mM EDTA solution (Sigma-Aldrich) in completed HBSS for 50 min at room temperature on a magnetic stirrer; finally washed twice with completed HBSS. The tissue pieces were cut more finely with disposable scalpel and incubated in 50 ml of type D collagenase solution 0.5 mg/mL (Roche Diagnostic) in RPMI (Sigma-Aldrich) overnight at 37°C on a magnetic stirrer placed in thermostatically controlled incubator.

The following day, the cell suspension, containing tumor infiltrating lymphocytes, was filtered through 100 µm cell strainer, centrifuged at 500 x g for 10 minutes. The pellet was resuspended in 4 ml of PBS and the lymphocytes were then isolated by Percol density-gradient centrifugation (Farmacia).

3.1.1.2 Lung tissue dissociation

Lung tissues were dissociated by using Tumor Dissociation Kit (human) and the gentleMACS™ Dissociator (Miltenyi Biotech). The experimental procedure was carried out following the standard protocols provided in the manufacturer's instructions.

Tumor Dissociation Kit human

Lung specimens were cut into small pieces of 2-4 mm with disposable scalpel, transferred into the gentleMACS C Tube containing the enzyme mix (200 µl of Enzyme H, 100 µl of Enzyme R and 25 µl of Enzyme A) in 4.7 ml of RPMI without serum and incubated onto gentleMACS Dissociator, using h_tumor_01 program. Then C Tube was placed for 30 minutes at 37°C under continuous rotation using the MACSmix Tube Rotator. Then, C tube was incubated onto gentleMACS Dissociator using Program h_tumor_02. Next, C Tube was incubated for 30 minutes at 37°C under continuous rotation using the MACSmix Tube Rotator and, for a second time, C tube was attached onto gentleMACS Dissociator using h_tumor_02 program.

After dissociation, the cell suspension was filtered through 100 µm cell strainer and centrifugated at 500 x g for 10 minutes. The pellet was resuspended in 20 ml of PBS and the lymphocytes were then isolated by Ficoll density-gradient centrifugation (Ficoll- Paque™ Plus, GE Healthcare 17-1440-02)

3.1.2 Isolation of mononuclear cells

Lamina propria mononuclear cells (LPMC) and peripheral blood mononuclear cells (PBMC) were isolated from tumor tissue (NSCLC and CRC), non-neoplastic counterparts and peripheral blood (PB) of the same patients, and from PB from healthy donors by density gradient centrifugation

3.1.2.1 Percoll density gradient centrifugation

Colon LPMC were isolated by discontinuous density gradient centrifugation on Percoll. After collagenase D incubation, the cell suspension was transferred into 50 ml tube through 100 µm cell strainer and centrifuged 10 minutes at 500 g. The supernatant was discarded, and the pellet was resuspended with 2-4 ml (depending on the pellet size) of 40% Percoll solution; the Percoll gradient was prepared using four steps, from the top layer down: 30%, 40%, 60%, and 100% Percoll solution, prepared in 15 ml tube as follow below:

2 ml Percoll 100% 45 ml pure Percoll + 4,1 ml 10X PBS + 0,9 ml HBSS

2 ml Percoll 60% 6 parts of Percoll 100% + 4 parts of HBSS

2 ml Percoll 40% 4 parts of Percoll 100% + 6 parts of 1X PBS

2 ml Percoll 30% 3 parts of Percoll 100% + 7 parts of HBSS

The Percoll-density gradient was centrifugated 30 minutes at 650 x g at room temperature without brake. The T cell fraction was recovered from the interface between the 60% and 40% Percoll layers (**Figure 10**), diluted with HBSS + Pen/Strep (final volume 30 ml) and centrifuged 10 minutes at 650 x g at RT. The cell pellet is now ready for further processing for FACS.

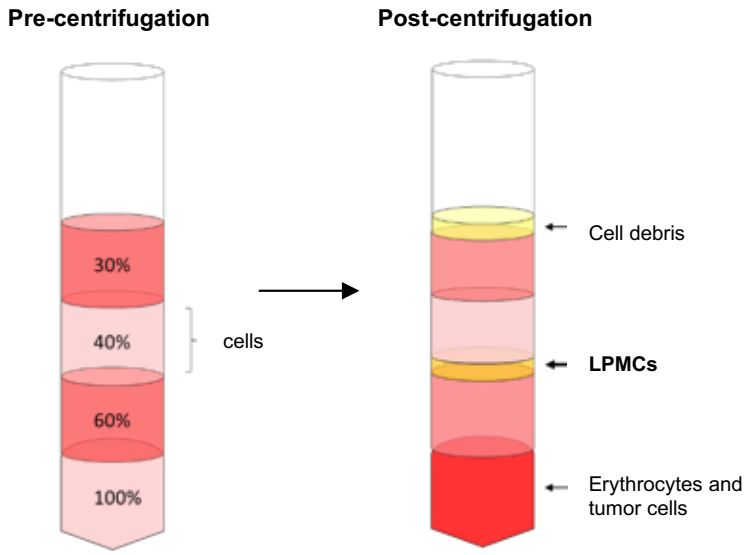


Figure 10 Schematic illustration of leucocyte isolation by discontinuous Percoll-gradient centrifugation. Left tube shows tube before centrifugation. Right tube shows tube after centrifugation: leukocytes are placed at the interface between 40% and 60% Percoll layers, tumor cells are in the bottom of the tube.

3.1.2.2 Ficoll density gradient centrifugation

PBMC and lung LPMC were isolated by Ficoll density gradient centrifugation. Blood samples were transferred to a 50 ml tube and diluted with PBS. Ficoll-Plaque were carefully layered on diluted blood samples or lung single cell suspensions, with ratio 1:3. Then each tube was centrifuged at 805g for 20 min at room temperature without brake. After centrifugation, the upper layer containing plasma and platelets was carefully removed, leaving the mononuclear layer undisturbed at the interface (**Figure 11**). The mononuclear cell layer was transferred to a new 50 ml tube, washed with PBS and centrifugated at 500 x g for 10 min at RT. The isolated mononuclear cell pellet was washed with PBS and centrifuged at 500 x g for 10 min at RT.

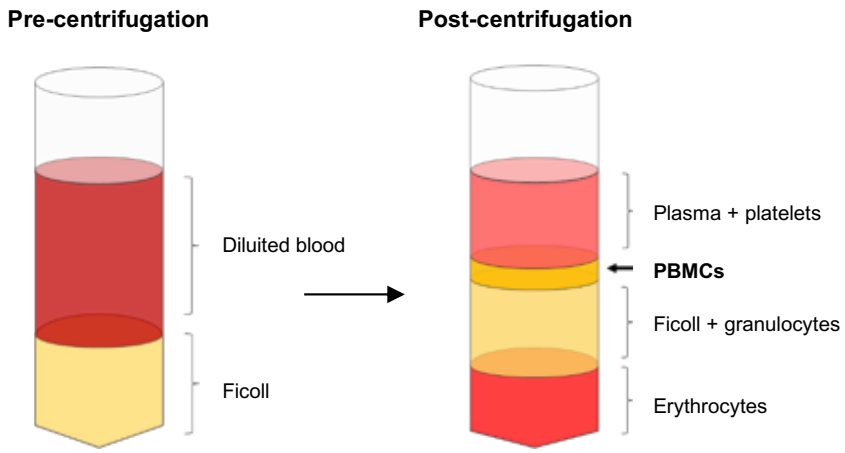


Figure 11 Schematic illustration of the leukocyte isolation of by Ficoll-gradient centrifugation. Left tube shows the gradient before centrifugation. Right tube shows the separation after centrifugation: leukocytes are located at the interface between plasma and Ficoll layers.

3.1.3 Isolation of human Treg cells by FACS

Treg cells were purified by Fluorescence Activated Cell Sorting (FACS) using the following fluorochrome conjugated antibodies: anti-CD4 APC/Cy7 (Biolegend clone OKT4), anti- CD127 PE (Miltenyi, clone MB15-18C9) and anti-CD25 FITC (eBioscience, clone 4E3), using FACSAria II (BD biosciences). Treg cell were sorted as CD25+ and CD127 dim (**Figure 12**, **a**).

To validate the expression of specific Treg markers, the following fluorochrome conjugated antibodies were used after sorting: anti-OX40 (Biolegend, clone Ber-ACT35), anti-CTLA-4 PE (eBioscience, clone eBio20A (20A, A3)) and anti-FOXP3 APC (eBioscience, PCH101 clone) (**Figure 12**, **b**).

In addition to the Treg cells, other four different populations of immune cells have been isolated starting from the tissue specimens by using flow cytometry, according to their specific sorting markers:

- CD4+ T conventional cells were isolated as CD3+ CD4+ CD45RO+ cells;
- CD4+ Tr1 as CD3+ CD4+ CD27+ CD195+ cells;
- CD8+ memory T cells as CD3+ CD8+ CD45RO+ TIM3- CD39- cells;
- CD8+ exhausted T cells as CD3+ CD8+ CD45RO+ TIM3+ CD39+ cells.

Extracellular staining

After purification of the mononuclear cells, the cell pellet was resuspended with the antibody mix (1 µl of each body for 10^6 cells) and incubated at 4°C for 20 minutes protected from light. The pellet was washed with MACS buffer, centrifuged at 500 x g for 5 minutes and then resuspended with 200 µl of MACS buffer.

Intracellular staining

After cell sorting, a small aliquot of Treg cells was centrifuged at 450 x g for 5 minutes. The pellet was resuspended with 200 µl of Fixation / Permeabiliization Concentrate (eBioscence) diluted 1: 4 in Fixation / Permeabiliization Diluent and incubated for 30 minutes at 4°C. The sample was then centrifuged at 450 x g for 5 minutes and the pellet was resuspended with 100 µl of Permeabilization Buffer 10X (eBioscience) diluted 1:10 in H_2O. Anti-FOXP3 antibody was added by incubating for 30 minutes at room temperature, protected from light. A wash was carried out

with 200 µl of PBS (450 x g for 5 minutes) and finally the cell pellet was resuspended in 200 µl of MACS buffer.

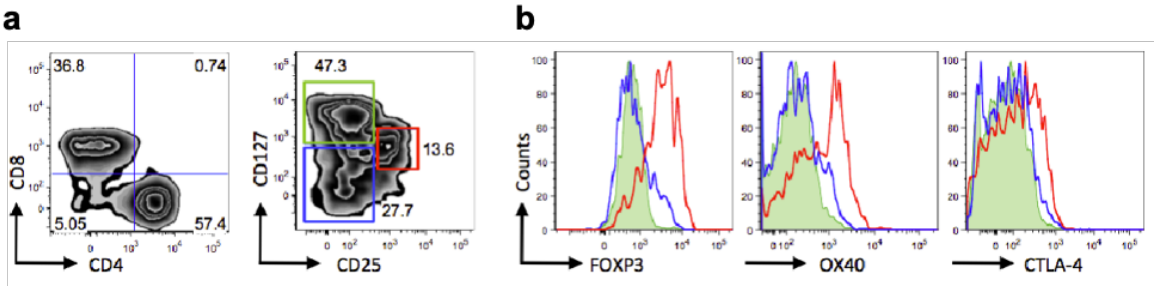


Figure 12 Isolation of human Treg cells by FACS. (**a**) Representation of the sorting strategy of Treg cells CD4+CD25+CD127dim (red box). (**b**) Validation of specific Treg markers performed on Treg isolated as CD4+CD25+CD127dim (red plot).

3.2. Treg *in vitro* expansion

Ex vivo Treg cells were *in vitro* expanded in presence of $1x10^6$ Feeders cells/ml and 1,6 $x10^5$ Rosi cells/ml, with IL2 (200U/ml) (Miltenyi biotec) and anti-CD3 antibody (30ng/ml) (clone OKT3, BioLegend, 317326) in complete RPMI (RPMI 1640, EuroClone ECM2001L), 10% FBS, 1% Pen/Strep, 10 mM Na/Pyruvate (Euroclone ECM0542D), non- essential amino acids 1X, (EuroClone ECB3054D)). After three days, OKT3 was removed and replaced with fresh medium + IL2 (200U/ml).

3.3 Epigenetic techniques

3.3.1 bulk ATAC-seq

To measure genome-wide patterns of chromatin accessibility, nucleosome positioning and perform footprinting analysis, we performed ATAC-seq(157) on *ex vivo* Treg cells, Tr1, CD4+, memory CD8+ and exhausted CD8+ T cells. ATAC-seq dataset consisted of 39 sequenced samples, derived from 9 patients with different clinical conditions. Primary human NSCLC and CRC tissues were obtained from 5 and 3 patients, respectively, together with normal adjacent tissues and peripheral blood for all of them. Cells isolated from one healthy donor were also processed.

Sample preparation

Immediately after sorting, T cells were centrifuged at 500 x g for 15 minutes at 4°C and then the pellet was resuspended in 100 µl of cold PBS. 5000 cells were aliquoted in 100 µl of cold PBS and centrifuged at 500 x g for 30 minutes at 4°C. The supernatant was discarded and the pellet was resuspended in 25 µl of cold ATAC lysis buffer (10mM Tris-HCl pH 7.4 (Invitrogen), 10 mM MgCl2 (Invitrogen), 3 mM MgCl2 (Invitrogen), 0.1% NP-40 (Thermo Scientific) in nuclease-free water (NF H_2O) (Ambion)) and cells were centrifugated at 500 x g for 30 minutes at 4°C.

Tagmentation

The nuclei pellet was resuspended in 5 µl of ATAC tagmentation mix, made by 2.5 µl Tagmentated DNA buffer 2X (Illumina), 1 µl Tn5 enzyme (Illumina) and 1.5 µl NF H_2O. Sample was incubated at 37°C for 1 hour.

Proteinase K treatment

5 µl ATAC clean up buffer (900 mM NaCl, 30 mM EDTA), 2 µl proteinase K 800 U/ml (New England BioLabs), 2 µl SDS 5% (Ambion), 20 µl NF H_2O were added to each sample, which was incubated at 40°C for 30 minutes.

The tagmentated DNA was purified using MinElute PCR purification kit (QIAGEN) following the standard protocols provided in the manufacturer's instructions, and DNA was eluted with 10 µl Buffer EB (10 mM Tris-HCl pH 8.5).

Library pre-amplification and preparation

In order to amplifying DNA, standard PCR on tagmentated DNA was carried out on Veriti Thermal Cycler (Applied Biosystems) and Illumina Nextera primers (i7 index and i5 index). 50 µl reaction for a PCR was performed: 10 µl of tagmentated DNA, 25 µl of NEBNext High-Fidelity 2x PCR Master Mix (New England BioLabs), 2.5 µl Index 1 (i7) Adapter (Illumina), 2.5 µl of Index 2 (i5) Adapters (Illumina), nuclease-free water up to 50 µl. PCR involved 1 cycle of 72°C for 5 minutes, 98°C for 30 seconds, followed by 5 cycles of 98°C for 10 seconds, 63°C for 30 seconds, 72°C for 1 minute.

Pre-amplification qPCR

To reduce GC and size bias in PCR and normalizing amplification cycles, the appropriate number of PCR cycles is determined using Quantitative Real-Time PCR (qPCR) allowing us to stop amplification prior to saturation. qPCR was performed using StepOnePlus System (Applied Biosystems). For amplification, 5 µL of amplified DNA were incubated with 0.09 µL of SYBR Green 100X (Applied Biosystem), 5 µL of NEBNext High-Fidelity 2x PCR Master Mix (New England BioLabs), 0.25 µL of Index 1 (i7) Adapter (Illumina), 0.25 µl of Index 2 (i5) Adapters (Illumina). The reaction mixture was brought up to a final volume of 15 µL with nuclease-free water. Thermocycling conditions were: 1 cycle of 98°C for 30 second; 25 cycles of 98°C for 10 seconds, 63°C for 30 seconds and 72°C for 1 minute. To rule out contamination from buffers and tubes, a negative control with water instead of the tegmented DNA template was used. We used 1/3 of the maximum fluorescent intensity to determining N cycles of library amplification. The remaining 45 µl PCR reaction were amplified according to the cycle number determined by qPCR, using the same thermocycle conditions without the initial activation step at 72°C for 5 minutes.

Post amplification clean up and double size selection

To remove any contaminants from DNA samples and reduce high molecular fragments, the tagmented DNA was purified using SPRI beads (Beckman Coulter). 27.5 µl of SPRI beads (0.55X) were added to the sample, mixed 15 times and incubated 5 minutes at room temperature. Placed on magnet, the supernatant was

transferred into new tube. 72.5 µl of SPRI beads (2X) were added to the sample, mixed 15 times and incubated 5 minutes at room temperature. Beads on magnet were washed two times with 100 µl of fresh 80% EtOH for 30 seconds. The beads were let dry for 2 minutes and then 22 µl of elution buffer was added, mixed 15 times and incubated for 5 minutes at room temperature and the supernatant was recovered.

The ATAC library quality was resolved using electrophoresis Bioanalyzer, and the DNA concentration was measured with a Qubit fluorometer (Life Technologies).

Sequencing

Properly amplified genomic DNA libraries were sequenced on the HiSeq4000 Illumina platform, using the paired-end sequencing mode with 125 bp reads, with an average sequencing depth of 240 million reads for each DNA library.

3.3.2 ChIPmentation

Crosslinking

After cell sorting, T reg cells were centrifuged at 500 x g for 5 minutes. The pellet was resuspended in 150 µl of PBS with 10% FBS (Gibco). To cross link the proteins to the DNA, fresh formaldehyde (Thermo Scientific) was added to a 1% final concentration to cell suspension, which was rocked for 10 minutes at room temperature. To quench the formaldehyde and terminates the cross-linking reaction, 0.125M glycine (Sigma-Aldrich) was added to cell suspension, and rocked for 5 minutes at room temperature. Cells were collected by centrifugation at 500 x g for 5 minutes at 4°C; pellet was washed with 150 µl of cold PBS plus protease inhibitors (Phenilmetilsulfonil Fluoride, PMSF and Protease Inhibitor Cocktails, PIC). The crosslinked cell pellet can be snap-frozen in liquid nitrogen and stored at -80°C.

Sonication

Cell pellet was resuspended in 130 µl of sonication buffer (10 mM Tris pH 8.0, 0.25% SDS, 2 mM EDTA, PMSF, PIC) and was sonicated in MicroTUBE (Covaris) using M220 Focused-ultrasonicator (Covaris) until most fragments are in the range af 200-700 bp (settings: duty factor 10%, peak incidence power 75 Watt, cycles per bust

200, 10-30 minutes). The fragment size was analyzed on 2100 Bioanalyzer (Agilent) using High Sensitivity chip, and DNA concentration was quantified by Quantus fluorometer (Promega). The sonicated chromatin can be snap frozen in liquid nitrogen and stored at -80°C for up to 3 months.

After sonication, lysate was diluted with 1.5 volumes of equilibration buffer (10 mM Tris-HCl pH 8.0, 233 mM NaCl, 1.66% TritonX-100, 0.166% DOC, 1 mM EDTA, PMSF, PIC). To eliminate debris, the sample was centrifugated at 14000 x g for 10 minutes at 4°C and the supernatant was transferred in a new tube.

Immunoprecipitation

Each immunoprecipitation was performed in 0.2 ml PCR tube using 150 µl of chromatin with RIPA-LS plus protease inhibitors (PIC and PMSF). Volume of sonicated chromatin corresponding to 10000 T reg cells was incubated with 0.1/0.2 µg of histone mark antibodies (H3K27Ac abcam 4729; H3K4me3 Millipore 07-473; H3K4Me1 DIAGENODE C15410194; H3K36me3 DIAGENODE C15410192; H3K27me3 07449 Millipore, rabbit IgG Sino Biological), overnight at 4 °C on a tube rotator. Immunocomplexes were recovered the following day with 10 µl Protein G-Dynabeads (Invitrogen), pre-blocked using RIPA-LS with 0,1% BSA, for 2 hours at 4°C. The beads were then washed twice with 150 µl of cold RIPA-LS (10 mM TrisHCl pH 8.0, 140 mM), twice with cold RIPA-high salt (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 500 mM NaCl, 1% Triton X-100, 0.1% SDS, 0.1% Na-Deoxycholate (DOC)), twice with cold RIPA-LiCl (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0, 250 mM LiCl, 0.5% NP-40, 0.5% DOC), once with cold 10 mM Tris-HCl pH 8.0. Finally, the beads were resuspended in 150 µl of 10 mM Tris-HCl pH 8.0 and the solution was transferred to a new tube. Beads bound to the magnet and the supernatant was discarded.

Tagmentation

The beads were resuspended in the tagmentation reaction (5 µl of 5X tagmentation buffer (50 mM Tris-HCl pH 8.0, 0.25 mM MgCl2, 50% v/v dimethylformamide), 19.5 µl of nuclease free water and 0.5 µl of Tn5 enzyme, (Illumina)) and incubated 10 minutes at 37°C. After chromatin tagmentation, 150 µl of RIPA-LS were added to tagmentation reaction and incubated on ice for 5 minutes to inactivate Tn5. The

beads were washed twice with 150 µl of cold RIPA-LS, and then twice with TE1X (10mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0).

Reverse cross-linking

After washing, the beads were resuspended in 48 µl of ChIP elution buffer (10 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0, 300 mM NaCl, 0.4% SDS) supplemented with 2 µl of Proteinase K 800U/ml, and incubated at 55°C for 1 hour and at 65°C overnight, for reverse crosslinking. The following day, the beads were placed on magnet and the supernatant was transferred to a new 1.5 ml DNA lo-bind tube. The beads were resuspended again with 19 µl of ChIP elution buffer and 1 µl of Proteinase K, and incubated 1 hour at 55°C. The beads were placed on magnet and the supernatant was combined with the first eluate.

The immunoprecipitated DNA fragments were purified by MinElute kit (Qiagen) and eluted in 22 µl EB Buffer.

Pre-amplification qPCR

To reduce GC and size bias in PCR and normalizing amplification cycles, the appropriate number of PCR cycles is determined using Quantitative Real-Time PCR (qPCR) to avoid saturation. qPCR was performed using StepOnePlus System (Applied Biosystems). For amplification, 2 µL of ChIPmentation DNA were incubated with 0.1 µL of SYBR Green 100X (Applied Biosystem), 5 µL of NEBNext High-Fidelity 2X PCR Master Mix (New England BioLabs), 0.3 µL of 25 µM Custom Barcodes Adapter 1 (Index i5) and 0.3 µL of 25 µM Custom Barcodes Adapter 2 (Index i7) (**Tables 1-2**). The reaction mixture was brought up to a final volume of 10 µl with nuclease-free water. Thermocycling conditions were: 1 cycle of 72°C for 5 minutes and 98°C for 30 seconds; 25 cycles of 98°C for 10 seconds, 63°C for 30 seconds and 72°C for 30 seconds. We used 1/3 of the maximum fluorescent intensity to determining N cycles of library amplification**.**

Library amplification and preparation

In order to amplify DNA and prepare a ChIPmentation library in one step, standard PCR on tagmentated DNA was carried on Veriti Thermal Cycler (Applied Biosystems). 50 µl reaction for a PCR was performed: 20 µl of tegmentated DNA,

25 µl of NEBNext High-Fidelity 2X PCR Master Mix (New England BioLabs), 1.5 µl 25 µM of Custom Barcodes Adapter 1 (Index i5) and 1.5 µL of 25 µM Custom Barcodes Adapter 2 (Index i7) (**Tables 1-2**), nuclease-free water to make up to 50 µl. PCR involved 1 cycle of 72°C for 5 minutes, 98°C for 30 seconds, followed by N cycles (determined by qPCR) of 98°C for 10 seconds, 63°C for 30, and 72°C for 30 minutes.

Table 1 Custom Barcodes Adapters 1 (index i5) sequences

v2_Ad1.1_TAGATCGC	AATGATACGGCGACCACCGAGATCTACACTAGATCGCTCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.2_CTCTCTAT	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.3_TATCCTCT	AATGATACGGCGACCACCGAGATCTACACTATCCTCTTCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.4_AGAGTAGA	AATGATACGGCGACCACCGAGATCTACACAGAGTAGATCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.5_GTAAGGAG	AATGATACGGCGACCACCGAGATCTACACGTAAGGAGTCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.6_ACTGCATA	AATGATACGGCGACCACCGAGATCTACACACTGCATATCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.7_AAGGAGTA	AATGATACGGCGACCACCGAGATCTACACAAGGAGTATCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.8_CTAAGCCT	AATGATACGGCGACCACCGAGATCTACACCTAAGCCTTCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.9_TGGAAATC	AATGATACGGCGACCACCGAGATCTACACTGGAAATCTCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.10_AACATGAT	AATGATACGGCGACCACCGAGATCTACACAACATGATTCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.11_TGATGAAA	AATGATACGGCGACCACCGAGATCTACACTGATGAAATCGTCGGCAGCGTCAGATGTGTAT
v2_Ad1.12_GTCGGACT	AATGATACGGCGACCACCGAGATCTACACGTCGGACTTCGTCGGCAGCGTCAGATGTGTAT

Table 2 Custom Barcodes Adapters 2 (index i7) sequences

v2_Ad2.1_TAAGGCGA	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.2_CGTACTAG	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.3_AGGCAGAA	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.4_TCCTGAGC	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.5_GGACTCCT	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.6_TAGGCATG	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.7_CTCTCTAC	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.8_CAGAGAGG	CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.9_GCTACGCT	CAAGCAGAAGACGGCATACGAGATAGCGTAGCGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.10_CGAGGCTG	CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.11_AAGAGGCA	CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCGTGGGCTCGGAGATGTG
v2_Ad2.12_GTAGAGGA	CAAGCAGAAGACGGCATACGAGATTCCTCTACGTCTCGTGGGCTCGGAGATGTG

To discard impurities from DNA samples and reduce high molecular fragments, the DNA was purified using SPRI beads (Beckman Coulter). 90 µl of SPRI beads (1.8X) were added to the sample, mixed 15 times and incubated 10 minutes at room temperature. Placed on magnet, the supernatant was discarded and the beads were washed two times with 100 µl of fresh 80% EtOH for 30 seconds. The beads were let dry for 2 minutes and then 22 µl of elution buffer was added, mixed 15 times and

incubated for 10 minutes at room temperature. Placed on magnet, the supernatant was recovered in new tube.

The ChIPmentation library quality was resolved using the High sensitivity DNA kit on Bioanalyzer 2100 instrument (Agilent) and the DNA concentration was measured with a Qubit fluorometer (Life Technologies) performed according to the manufacturer's instructions.

Sequencing

Properly amplified ChIPmentation libraries were sequenced on the HiSeq2500 Illumina platform, using the single-end sequencing mode with 50-75 bp reads, with an average sequencing depth of 40 million reads for each DNA library.

3.3.3. ChIP-seq

ChIP protocol was performed following the same steps of ChIPmentation protocol (see below) without the tagmentation step and library amplification.

ChIP-seq libraries were constructed with TruSeq ChIP Library Preparation Kit (Illumina), according to the manufacturer's instructions and sequenced on Illumina HiSeq2500 platform. Library preparation for ChIPmentation was performed using Nextera primers (Illumina) and enriched libraries were purified using 1.8 V of SPRI beads and sequenced with Illumina HiSeq2500.

3.4 Epigenomic data analysis

3.4.1 ChIP-seq data processing and quality control

The quality control of the reads was performed using FastQC v0.11.7(158) and MultiQCv1.5 (http://www.bioinformatics.babraham.ac.uk/projects/). The reads were aligned to the human hg38 reference genome (GENCODE Release 25 basic gene annotation) using Bowtie v1.2.2(159), sorted using SAMtoolsv1.8 (160)and directly converted into binary files (BAM). PCR duplicate reads were marked and removed using SAMtoolsv1.8. The peaks were called with MACS2 v2.1.(161) using matched input DNA as a control. Peaks overlapping ENCODE blacklisted regions hg38 were removed. Peaks found in un-placed and un-localized scaffolds were removed. For the visualization of ChIP-seq tracks, Bedgraph tracks were generated using MACS2 bdgcmp function, converted into bigwig using UCSC bedClip and bedGraphToBigWig functions. The pyGenomicTrack(162) tool was used for the visualization of the tracks.

3.4.2 Density and heatmap plot for each histone modification

Filtered and sorted BAM files were used to generate normalized coverage tracks using the bamCoverage function from the deepTools(163) suite. The average signal profile and the heatmap plot along the gene body were calculated using the computeMatrix function in scale-regions mode with default parameters and the GENCODE Release 25 basic gene annotation.

3.4.3 Correlation analysis of histone marks

To obtain the correlation heatmap of all the histone modifications, a consensus peakset was generated using DiffBind v2.10.0(164) and merging together only peaks detected in at least two tracks. Then, a count matrix of 180250 peaks x 48 samples was created by counting the number of reads per peak for each sample using the dba.count function with default parameters. The correlation heatmap and the principal component analysis (PCA) were produced using dba.plotHeatmap (distMethod="pearson") and dba.plotPCA respectively, with default parameters.

3.4.4 ATAC-seq data processing and quality control

The sequence quality along all sequenced reads was estimated using FastQC v0.11.9(158) and MultiQCv1.9. To remove masked adapter sequences and to cut sequences from low-quality score regions, TrimGalorev0.6.5 was used. In particular, base calls with Q score < 20 were trimmed from the 3' end of the reads and reads shorter than 36 bp were filtered out. The reads were aligned to the human hg38 reference genome (UCSC hg38 assembly) using BWA-mem (Burrows-Wheeler Alignment), sorted using SAMtoolsv1.10(160) and directly converted into binary files (BAM). Multi-mapped reads, PCR duplicate and low-quality reads were marked and removed using Picardv2.23.1 and SAMtoolsv1.10. Sequences of mitochondrial origin as well as reads mapped to ENCODE hg38 blacklisted regions were discarded using BEDtoolsv2.29.2 and SAMtoolsv1.10. Peaks were called with MACS2v2.2.7.1(161), using *-BAMPE -nomodel* options in order to allow the use of the actual insert size of reads pairs to build the fragment pileup. The Q-value cutoff to call significant regions was set to 0.05 and the *-broad* option was set to false to call sharp peaks. The annotation of peaks relative to genomic features was performed using HOMERv4.11 annotatePeaks and the annotation GTF file from UCSC hg38 assembly. A consensus peak-set across all samples was created using BEDtoolsv2.29.2 *merge*. Peaks found in un-placed and un-localized scaffolds were removed. Starting from the original consensus peak-set obtained with Bedtools, we retained all the peaks present in at least two replicate samples out of the fully paired ATAC-seq samples of the Treg dataset. We then filtered all the peaks mapped to Y chromosome and to Pseudo-Autosomal Regions (PARs), resulting in 71,160 tested Treg consensus peaks.

For the visualization of ATAC-seq tracks of each sample, the GenomeCoverageBed function from bedtoolsv2.29.2. was used to generate bedGraph tracks, scaling the coverage by a constant factor calculated as 1 million divided by the number of mapped reads. The tracks were converted into bigwig using ENCODE bedGraphToBigWigv357. IGV was used for exploring and browsing the tracks.

3.4.5 Annotation of open chromatin regions

Differentially accessible regions were annotated to their putative target genes using public H3K27ac HiChIP data on enhancer-promoter-associated chromatin contacts. The H3K27ac HiChIP loops were specific for human peripheral blood Tregs isolated from healthy donors(165). The remaining regions that were not annotated with H3K27ac HiChIP, were annotated to their nearest and/or overlapping coding or non-coding gene.

3.4.6 Principal Component Analysis and Cluster Analysis

PCA plot and hierarchical cluster were generated using read counts against all the peaks of the Treg consensus peak-set. Raw read counts were normalized using the variance-stabilizing transformation built into the DESeq2 package (https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8). Limma package (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4402510/) was use to correct for batch effect related to patient-to-patient variability.

3.4.7 Differential accessibility analysis of ATAC-seq data

Differentially accessible regions were identified using the DESeq2 package(166) (version 1.22.2). DESeq-derived size factors computed for each sample were used to normalize read counts per consensus peak-set across replicates. Log2FC of each chromatin accessibility region was estimated for each cell type comparison, while controlling for patient-to-patient variability in the design formula. We considered as differentially activated all the regions with a padj ≤ 0.01.

3.4.8 ChIPmentation data processing and quality control

The quality control of the reads was performed using FastQC v.0.11.9 and MultiQCv1.9 (http://www.bioinformatics.babraham.ac.uk/projects/). The reads were aligned to the human hg38 reference genome (UCSC hg38 assembly) using BWA-mem v.0.7.17, sorted using SAMtoolsv1.10 and directly converted into binary files (BAM). Multi-mapped reads, PCR duplicates and low-quality reads were marked and removed using Picardv2.23.1 and SAMtools. The peaks were called with MACS2v2.2.7.1. using matched IgG sample as a control. The P-value to call

significant regions was set at 0.01. For H3K27ac and H3K4me3 samples, the -*broad* option was set as *false* to call sharp peaks, whereas for H3K27me3, H3K4me1 and H3K36me3 samples broad peaks were called. Peaks overlapping ENCODE hg38 blacklisted regions were removed. Peaks found in un-placed and un-localized scaffolds were removed. For the visualization of ChIPmentation-seq tracks of each sample, bedGraph tracks were generated using GenomeCoverageBed function from bedtoolsv2.29.2., scaling the coverage by a constant factor calculated as 1 million divided by the number of mapped reads. The tracks were converted into bigwig using ENCODE bedGraphToBigWigv357. IGV (Integrative Genomic Viewer) was used for visualizing and browsing the tracks.

3.4.9 De novo chromatin state characterization

De novo chromatin state characterization was performed using a multivariate Hidden Markov Model, a machine learning approach (ChromHMM v1.12(167)) considering five histone modifications (H3K4me3, H3K27ac, H3K4me1, H3K36me3 and H3K27me3) across our ChIPmentation library. The IgG samples were downsampled to a maximum depth of 10 million reads. The read counts for all the considered samples were computed in non-overlapping 200-bp bins across the entire genome. The binarization step was performed comparing ChIPmentation-seq read counts to corresponding IgG sample read counts as control to adjust the binarization threshold locally and thus reduce the technical noise. Chromatin state annotation tracks were generated separately for each cell population and biological replicate. Several models were trained in parallel considering 8, 10, 12 and 15 number of states. Of these, the 15-state model captured the key interactions between histone marks and identified an active enhancer state, thus it was used for downstream analyses.

3.4.10 Identification of active enhancers

To identify the highly active enhancer elements, all the “Active Enhancer” and “Genic Enhancer” regions from ChromHMM states were selected. These two states are defined by the co-presence of high levels of H3K27ac and H3K4me1 signal in the case of active enhancers and the co-occupancy of H3K27ac, H3K36me3 and H3K4me1 for genic enhancers. The combined active and genic enhancer states

were filtered for regions with a length of more than 200 bp. A consensus peakset was built for peripheral blood and tumor-infiltrating Treg samples using DiffBind v2.10.0(164) by merging together enhancers detected in at least two biological replicates. This generated a consensus peakset of 13.050 regions across seven samples.

3.4.11 Overlap of ChromHMM states and ATAC-seq data

Bedtoolsv2.29.2 was used to overlap ATAC-seq accessibility peaks with ChromHMM-defined enhancer regions. The ATAC-seq peaks were considered as overlapping if at least 50% of the interval base-pairs intersected with an enhancer region. The ATAC-seq peak start- and end-positions of were then used to define the enhancer regions.

3.4.12 RNA-seq data analysis

For tiTregs and PB-Tregs, we used RNA expression data derived from bulk RNA-seq experiment, previously obtained in our laboratory and published(99). Raw counts of each library were normalized using DESeq2-derived size factors and the mean expression level across replicate samples was obtained. Ensamble gene ID were matched to the TF gene symbols using org.Hs.eg.db Bioconductor package(168).

3.4.13 Footprint analysis on ATAC-seq data

Transcription factor footprinting analysis was performed with TOBIAS, a python-based suite of command-line tools. To ensure maximum reproducibility, a Nextflow-based pipeline wrapping all the main functionalities of the suite was used (https://github.molgen.mpg.de/loosolab/TOBIAS-nextflow/tree/master/TOBIAS_MAPOKS). The ATAC-seq data was corrected for intrinsic Tn5 bias in two steps using the ATACorrect module: first, by defining each Tn5 cutsite as the 5' end of a read shifted by +5 bp at the plus strand and -4 bp at the minus strand to center the transposition event. Then, by calculating expected Tn5 cutsites per region, using a dinucleotide weight matrix (DWM) to estimate the background probability of Tn5 insertion based on its sequence preference, and subtracting these expected cuts from the observed signals to yield corrected tracks.

A continuous combined footprint score was computed using ScoreBigwig module, with a function taking into account both read depletion due to TF binding and accessibility signal. Once base-wise footprint scores were computed, the BINDetect module was used to scan the sequence information of accessible regions previously called by MACS2 for known TFs binding motifs. Their position weight matrices (PWMs) were downloaded from the JASPAR CORE database (2021). Once a motif was confidently matched (pval < 1e−4) to a DNA sequence within open chromatin regions, the base-wise score was converted by TOBIAS into a motif-wise score by taking the highest score among the ones of the bases composing the motif. A global threshold to the score was set to distinguish between bound and unbound sites, with a significance p-value set to as 0.05. For the differential footprinting analysis between multiple conditions, the individual site scores were summarized with the differential binding module by TOBIAS in order to identify differentially bound TFs. TF networks were obtained with the TOBIAS CreateNetwork module and drawn with Cytoscapev3.8.2 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC403769/).

3.5 Epigenetic editing in cell line

3.5.1 Plasmids

lenti_dCas9-KRAB-MeCP2 was a gift from Andrea Califano (Addgene plasmid #122205; http://n2t.net/addgene:122205 ; RRID:Addgene_122205) (**Figure 13**) and was modified to use the dCpf1 protein instead of dCas9 as a DNA-binding domain, as well as to express three copies of the HA tag and a Nucleoplasmin-derived Nuclear Localization Signal instead of the SV40-derived sequence to generate the lentiviral dCpf-KRAB-MeCP2 plasmid.

The lentiviral dCpf-KRAB-MeCP2 construct was generated by PCR-amplification of dCpf1 from MMW1578: CAG-human dLbCpf1(D832A)-NLS-3xHA, a gift from Keith Joung (Addgene plasmid #104563; http://n2t.net/addgene:104563; RRID: Addgene_104563) with primers generating MluI and BamHI restriction sites and Q5 high-fidelity DNA polymerase (New England BioLabs); dCpf1 fragment was cloned into lenti_dCas9-KRAB-MeCP2 backbone lacking the dCas9 repressor, HA tag and 2 NLS domains after digestion using MluI and BamHI restriction enzymes; this intermediate dCpf1-KRAB-MecP2 plasmid has finally undergone BamHI digestion and synthetic dsDNA insertion, containing 3xHA + 2XNLS, to obtain the final dCp1-KRAB-MeCP2 plasmid (**Figure 45** , Results).

CROPseq-Guide-Puro was a gift from Christoph Bock (Addgene plasmid #86708; http://n2t.net/addgene:86708 ; RRID:Addgene_86708) and was modified in-house to swap the selection marker from Puro to mCherry, resulting in CROPseq-Guide-mCherry (**Figure 14**) used for guide expression lentiviral vector for Cas9 experiments. In addition, the CROPseq-Guide-mCherry plasmid was modified to remove the Cas9 sgRNA scaffold sequences from the 3' of the cloning site and express instead the Cas12a/Cpf1 crRNA scaffold sequences at the 5' of the cloning site, resulting in Cpf1CROPseq-Guide-mCherry (**Figure 15**) used for guide expression lentiviral vector for Cpf1 experiments.

CRISPRoff-v2.1 (DNMT3A-DNMT3L-dCas9-BFP-KRAB) was a gift from Luke Gilbert (Addgene plasmid #167981; http://n2t.net/addgene:167981; RRID:Addgene_167981)(169) (**Figure 16**).

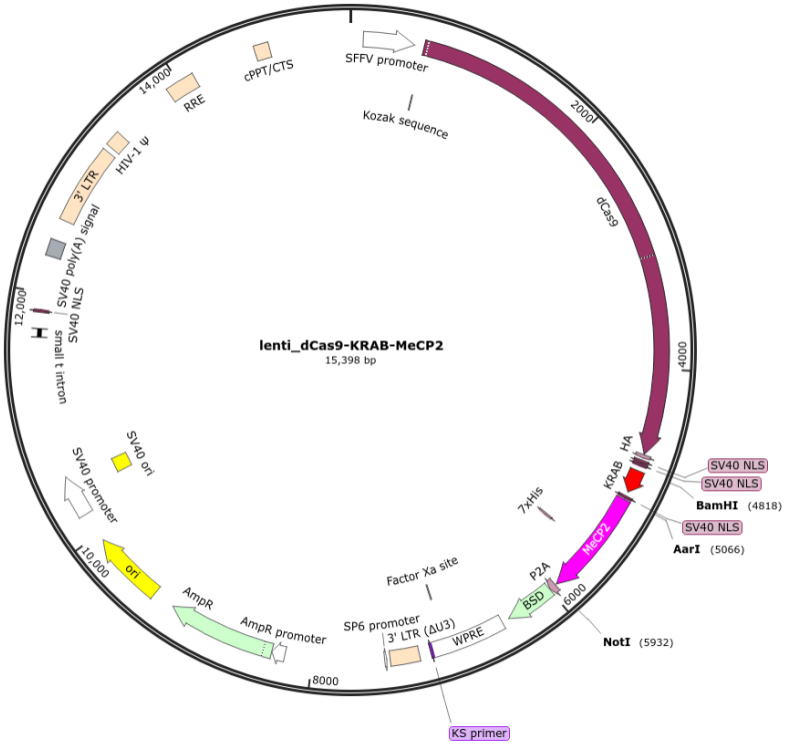


Figure 13 dCas9-Krab-MeCP2 plasmid. Lentiviral vector lenti_dCas9-Krab-MeCP2 (Addgene plasmid # 122205)

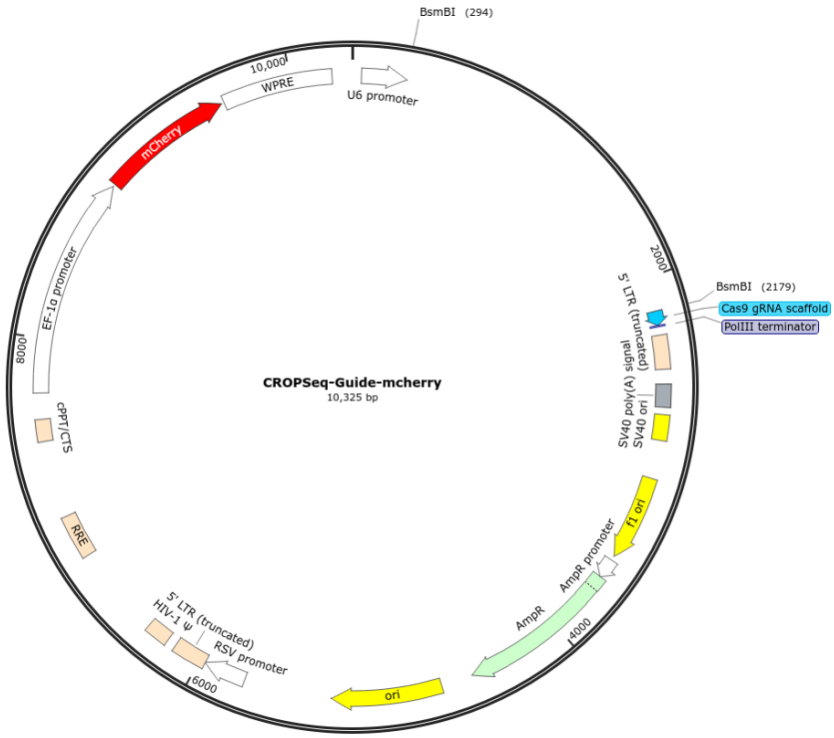


Figure 14 CROPseq-Cas9Guide plasmid. CROPseq-Guide-mCherry guide expression lentiviral vector for Cas9 experiments (modified Addgene plasmid # 86708)

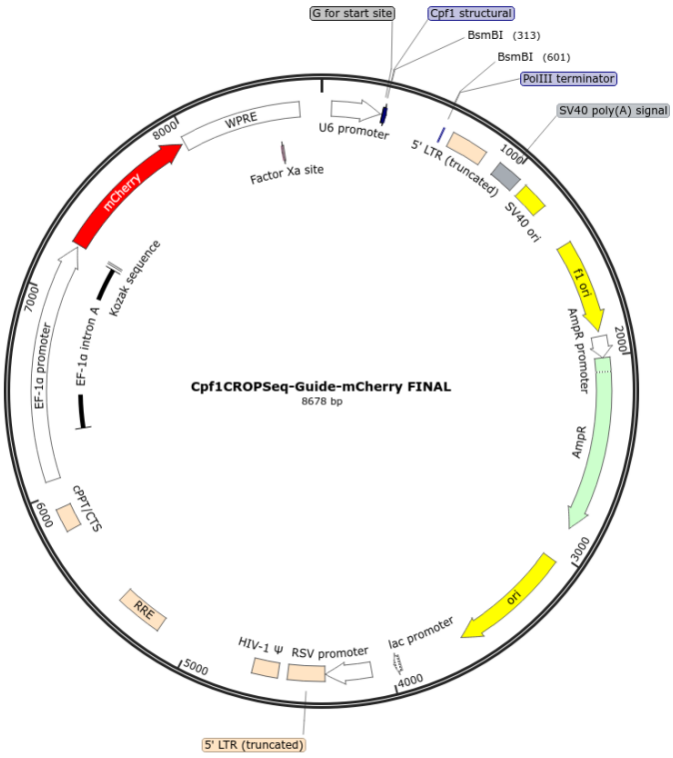


Figure 15 CROPseq-Cpf1Guide plasmid. CROPseq-Guide-mCherry guide expression lentiviral vector for Cpf1 experiments (modified Addgene plasmid # 86708)

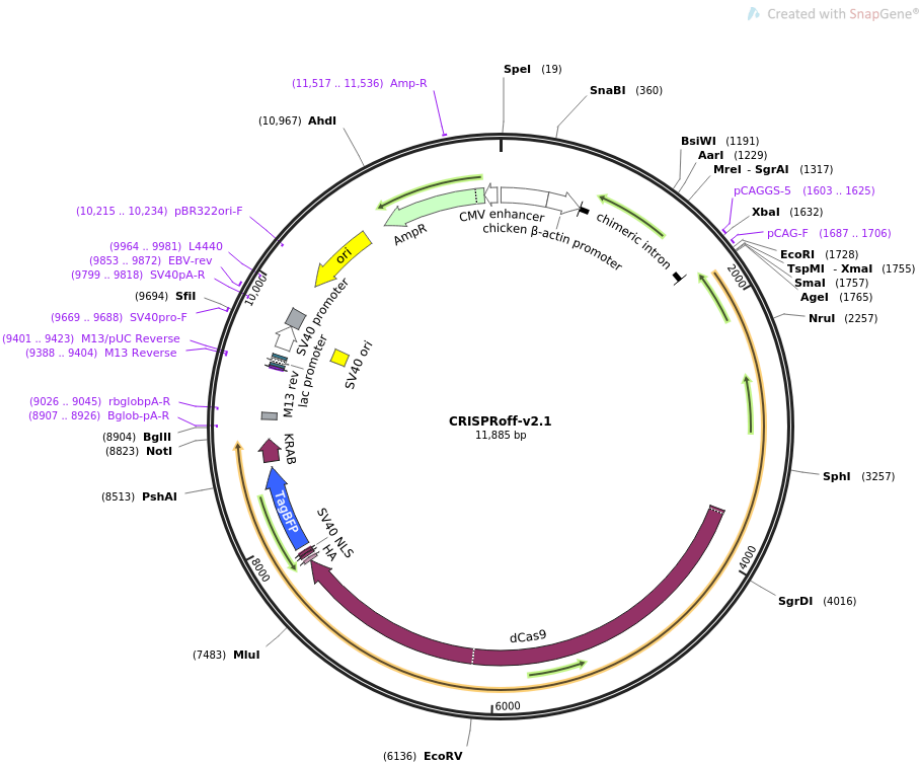


Figure 16 CRISPRoff-v2.1 plasmid. CRISPRoff-v2.1 plasmid (DNMT3A-DNMT3L-dCas9-BFP-KRAB) (Addgene plasmid #167981).

3.5.2 Guide design and cloning

Promoter-directed guide sequences for Cas9 systems were obtained from Yeo et al., 2018. Promoter-directed guide sequences for Cpf1 systems were designed to match the seed region of the Cas9 guides as closely as possible, selecting the highest scored seed-matched candidates from a list of potential target sequences extracted via the online tool GPP sgRNA Designer (Broad Institute, https://portals.broadinstitute.org/gpp/public/), set to search for CRISPRi-optimized guides for Cpf1.

Oligos were designed to match the sticky ends produced by the BsmBI sites present on the CROPseq-Guide-mCherry and Cpf1CROPseq-Guide-mCherry plasmids. All oligos were ordered from Eurofins with standard synthesis and purification conditions; their sequences are reported below in **tables 3** and **4**.

Table 3 Cas9 guide sequences

Cas9 guide	gRNA sequence	Forward oligo	Reverse oligo
Cas9_METg1	TGAGCAGATGCGGAGCCGAG	CACCGTGAGCAGATGCGGAGCCGAG	AAACCTCGGCTCCGCATCTGCTCAC
Cas9_METg2	ACTGGTTCCTGGGCACCGAA	CACCGACTGGTTCCTGGGCACCGAA	AAACTTCGGTGCCCAGGAACCAGTC
Cas9_CHEKg1	GGTGGAGGAATGGTACCAGG	CACCGGGTGGAGGAATGGTACCAGG	AAACCCTGGTACCATTCCTCCACCC
Cas9_CHEKg2	GGGTCTAGATTAGTGAGGGA	CACCGGGGTCTAGATTAGTGAGGGA	AAACTCCCTCACTAATCTAGACCCC

Table 4 Cpf1 guide sequences

Cpf1 guide	gRNA sequence	Forward oligo	Reverse oligo
Cpf1_METg1	TGAGCAGATGCGGAGCCGAGTGG	TAGATTGAGCAGATGCGGAGCCGAGTGG	AAAACCACTCGGCTCCGCATCTGCTCAA
Cpf1_METg2	cCTTTCGGTGCCCAGGAACCAGT	TAGATCCTTTCGGTGCCCAGGAACCAGT	AAAAACTGGTTCCTGGGCACCGAAAGGA
Cpf1_CHEKg1	GGAGGGTGGAGGAATGGTACCAG	TAGATGGAGGGTGGAGGAATGGTACCAG	AAAACTGGTACCATTCCTCCACCCTCCA
Cpf1_CHEKg2	TAAAGAAGGAGTTCGGGGTCCTAG	TAGATTAAAGAAGGAGTTCGGGGTCTAG	AAAACTAGACCCCGAACTCCTTCTTTAA

Oligo annealing and phosphorylation

Oligos were annealed and phosphorylated in a mix containing 1 µl of each of the forward and reverse oligo 100 µM stock solutions, 1 µl of 10X T4 Ligase Buffer (New England Biolabs), 0.5 µl (5U) of T4 Polynucleotide Kinase (PNK) (New England Biolabs) and 6.5 µl of nuclease-free water. The mix was incubated for 30 minutes at 37°C (phosphorylation), then heated up to 95°C and cooled down to room temperature at a controlled rate of 5°C/minute. The process was conducted on an Applied Biosystems Veriti thermocycler with previously characterized ramp rate.

Backbone digestion

5 µg of the guide-expression backbone were digested in a 50 µl reaction volume with 50 U of BsmBI at 55 °C for 3 hours. The digested backbone was loaded on 0.8% agarose gel, cut (size of ~8500 base pairs) and purified with a Wizard SV Gel and PCR Clean-Up System (Promega) according to manufacturer's instructions.

Ligation

The annealed, phosphorylated oligos and the cleaved, purified backbone were used in a ligation reaction: 1 µl annealed oligos diluted 1:200, 50 ng of the digested backbone, 1 µl of T4, 2 µl of T4 Ligase Buffer, nuclease-free water up to 20 µl. A control '0:1' (as opposed to 3:1, meant as the ratio of insert to backbone) reaction in which water was added instead of the oligos was always carried out to check for self-ligation of the backbone.

T4 ligation reaction was incubated at room temperature for 20 minutes; then, 5 µl of the ligation volume were transformed in TOP10 competent cells (Invitrogen) by heath shock, according to manufacturer's instructions. Specifically, 100 µl of the transformation volume were dispensed on an Ampicillin-selective LB-Agar plate. and cultured overnight at 37°C in a humidity-controlled incubator.

Positive colony selection and culture

The following day, we selected picked positive colonies performing a standard PCR. PCR was carried out using 2X Master Mix Standard GL (Genespin), forward primer 5'-GGCCTATTTCCCATGATTCC-3'the reverse primer 5'-GAGTTTGGACAAACCACAAC-3' and the colony DNA template dipped in nuclease-free water. PCR was performed on a GeneAmp PCR System 9700 (Applied Biosystems) thermocycler, set as follows: 5 minutes at 95°C, then 25 cycles of 95°C for 30 seconds, 55°C for 30 seconds, 72°C for 2 minutes and 30 seconds, followed in turn by 5 minutes at 72°C and a 4°C hold for storage. In addition, the ligation reactions and undigested backbone were also amplified as respectively positive and negative controls. The positive colonies were identified as specific fragment on a 1.5% agarose gel.

The positive colonies were cultured in 2ml microcentrifuge tube containing 500 µl of LB broth + Ampicillin over the day and then transferred in 3,5 ml of LB broth + Ampicillin in 50 ml Falcon tubes for overnight culturing at 37°C in a shaking incubator set at 225 rpm. After the culturing, 4 ml of the culture volume were precipitated and the plasmid DNA was extracted via PureLink Quick Plasmid Miniprep Kit (Invitrogen) according to manufacturer's instructions and resuspending pellets in the minimum possible volume. After quantitation of the purified plasmid with a Nanodrop 1000 (ThermoFisher Scientific) spectrophotometer, an aliquot of the resulting DNA was sent for Sanger sequencing via the GATC Light Run program (Eurofins Genimics) to confirm the sequence of the insert and flanking regulatory regions. The resulting sequences were aligned to the backbone via the Clustal X desktop application to verify adherence to the projected sequence.

Plasmids that conformed to the expectations were amplified starting from the residual overnight culture volume by seeding in a hundred-fold higher volume of LB broth + Ampicillin, cultured overnight in identical conditions and purified via PureLink HiPure Plasmid Midiprep Kit (Invitrogen), according to manufacturer's instructions. All centrifugation steps for the latter procedure were executed on an Avanti J-20 XP ultracentrifuge (Beckman Coulter) using JA-14 and JA-25.50 rotors and Nalgene polypropylene copolymer tubes and bottles.

These plasmids were used for epigenetic editing in HEK293T cell line.

3.5.3 HEK293T cell line

HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Euroclone, Lonza) with 10% south american Fetal Bovine Serum (FBS) (Euroclone), L-glutamine 2 mM (Euroclone), Penicillin/Streptomicin 1% (Euroclone), Non-Essential Amino Acids 1% (Euroclone) and sodium pyruvate 1mM (Euroclone) and maintained at 37°C with 5% CO_2. Cell detachment was performed after one wash with Phosphate Buffered Saline (PBS) 1X (Lonza) via short (2-3 minute) digestion with warm trypsin-EDTA (Euroclone) followed by neutralization with at least 4 volumes of warm complete medium. Cells were routinely tested for mycoplasma contamination.

3.5.3.1 HEK transfection with Lipofectamine

HEK293T cells were used for epigenetic editing experiments executed both as single transfection with dCpf1KRAB-MeCP2 and co-transfections of one epigenetic repressor and one guide-bearing plasmids.

For transient expression experiments, 400000 HEK cells were seeded in 6-well plates (Nunclon). The following day, were transfected with Lipofectamine 2000 (Life Technologies, cat. #11668019) as per manufacturer's instruction using 2.5 µg of transcriptional repressor and an equimolar amount of each guide. The ratio of µl Lipofectamine/µg DNA ration was maintained 2:3. Cells were incubated for either five hours; then the lipoplex-containing medium was removed and replenished with 3ml complete DMEM. Depending on the experiment, cells were harvested 24, 48 or 72 hours after medium change.

HEK cells were selected added 10 µg/ml of blasticidin (Invivogen) to the culture medium from 24 hours after transfection.

3.5.3.2 Subcellular fractionation

The subcellular fractionation was based on the Gagnon protocol(170) and then optimized in our laboratory for our specific system.

HEK293T cells were detachment, counted and resuspended in PBS at the concentration of 10^6/ml in microcentrifuge tube. 10% of the resuspension volume was placed in new microcentrifuge tube for the total protein extract; both were

centrifuged at 500 x g for 5 minutes at 4°C. After the centrifugation, the supernatant was removed and the pellet for the total protein extract was momentarily stored at -20°C, while the second tube was used for the fractionation. In the first lysis step, the cells were resuspended in 100 µl per million cells of Hypotonic Lysis Buffer (HLB) (10 mM Tris pH 7.5, 10 mM NaCl, 3 mM MgCl, 0.3% NP40, 10% Glycerol, plus PIC), and centrifuged at 800 g for 8 minutes at 4°C. After centrifugation, the supernatant was recovered in a new microcentrifuge tube for the extraction of the cytoplasmic extract. 5M NaCl was added to the cytoplasmic extract to a final concentration of 150 mM to salt out proteins. The fraction was then stored at -20°C. The nuclear pellet was washed with 500 µl HLB + PIC and centrifuged at 200 g for 2 minutes at 4°C. This step was repeated once more with an identical volume of HLB + PIC and three more times with the same volume of PBS. After the last wash, the total protein extract was recovered and the nuclear and total protein pellets were resuspended in 50 µl RIPA Buffer (50 mM Tris HCl pH8, 1% NP40, 0.5% Na-Deoxycholate, 0.1% SDS, 150 mM NaCl, plus PIC) and Turbo DNAse and incubated for 20 minutes on ice, then incubate at 37°C for 15 minutes and lastly centrifuged at 1200 g for at least 15 minutes at 4°C. The supernatants containing fractional protein extracts can be recovered at this point and stored at -20°C until used.

3.6 Gene expression analysis

For epigenetic editing experiments, HEK293T cells were lysated for the extraction of total RNA at 48 hours after transfection.

3.6.1 RNA extraction

Total RNA was isolated from HEK293T cells using Trizol Reagent (Invitrogen) according to the manufacturer's instructions. RNA was quantified by NanoDrop 1000 (Biophotometer, Eppendorf).

3.6.2 Reverse Transcription

5 µg of total RNA were retrotranscribed using SuperScript III Reverse Transcriptase, Oligo(dT) Primers and RNAse H (Invitrogen) according to manufacturer's instructions. This procedure results in 21 µl of cDNA suspension.

3.6.3 Quantitative Real-Time PCR

Quantitative Real-Time PCR was performed using 96well-plate on a LightCycler 96 (Roche) thermocycler in the SYBR Green chemistry (LightCycler 480 SYBR Green I Master - Roche). For each well, 10 ng of cDNA was incubated with 1X Master Mix, 10 uM of forward and reverse primers up to a final volume of 20 µl nuclease-free water. Thermocycling conditions were: 3 minutes at 95°C; then 40 cycles of 30 seconds at 95°C, 30 seconds at 55°C, and 30 seconds at 72°C.

Reaction was carried out using:

- MET1 (forward primer) 5'- AGCAATGGGGAGTGTAAAGAGG -3' and MET1 (reverse primer) 5'- CCCAGTCTTGTACTCAGCAAC -3',
- CHEK1 (forward primer) 5'- ATATGAAGCGTGCCGTAGACT -3' and CHEK1 (reverse primer) 5'- TGCCTATGTCTGGCTCTATTCTG -3'
- ACTB (forward primer) 5'- CATGTACGTTGCTATCCAGGC -3' and ACTB (reverse primer) 5'- CTCCTTAATGTCACGCACGAT -3', taking from the Yeo and colleagues' work(171).

Each target was quantitated in triplicate for each condition and normalized to ACTB expression from the same plate. A negative control with water instead of the cDNA template were included in each plate. Analysis was conducted on LibreOffice Calc after exporting of the experiment statistical data in .csv format from the LightCycler96 software; graphs were drawn in GraphPad Prism.

We used a relative quantification method $2^{-(\Delta Ct)}$ to calculate the gene expression values.

Where:

$\Delta Ct = Ct_{guide} - Ct_{ACT}$

Analysis of relative expression for n=1 experiments was conducted using the following formula: $2^{-(\Delta\Delta Ct)}$

where:

$\Delta\Delta Ct = \Delta Ct_{guide} - \Delta ct_{no\ guide}$

$\Delta Ct_{guide} = Ct_{guide} - Ct_{ACT}$

$\Delta Ct_{guide} = Ct_{no\ guide} - Ct_{ACT}$

3.7 Staining for FACS analysis

Co-transfection efficiency, i.e. the fraction of cells positive for both the guide-bearing plasmid and the transcriptional repressor in double plasmid delivery experiments via Lipofectamine 2000, was extracted via flow cytometry on a FACS Canto II (Becton Dickinson) cytometer. Gating was performed to exclude debris and doublets and include only Pacific Blue+ mCherry+ cells.

Samples were transferred in a “U-bottom” 96 well-plate and centrifuged 5 minutes at 450 x g at room temperature. Cells were then washed with MACS buffer (Miltenyi biotec), centrifuging for 5 minutes at 500 x g at room temperature. Cells were stained with Fixable Viability Stain FVS780 (1:1000) (BD Biosciences), incubating for 10’ at room temperature protected from light. Cells were then washed once with MACS buffer and once with FBS, spinning at 450 x g for 5 minutes at room temperature. Cells were incubated for 20 minutes at 37°C protected from light and then washed with MACS buffer and spun at 450 x g for 5 minutes at room temperature. Cells were firstly fixed used solution constituted by Fixation/ Permeabilization concentrate (Invitrogen) diluted 1:4 in Fixation/ Permeabilization Diluent (Invitrogen) and incubated for 30 minutes at 4°C in the dark. After washing with MACS buffer at 450 x g for 5 minutes at room temperature, cells were permeabilized using Permeabilization Buffer, diluted 1:10 (Invitrogen) and we proceeded with intracellular staining, adding 0.5 µl of intracellular antibodies for $1x10^6$ cells. In particular, we used Pacific Blue (PB) anti-HA antibody (BioLegend) for dCas9 or dCpf1 identification. Cells were incubated for 30 minutes at room temperature, protected from light, and then washed once with MACS buffer, spinning at 453 x g for 5 minutes at room temperature. Finally, cells were transferred in FACS tubes and samples were analysed using FACS CANTO II (BD biosciences).

3.8 Western Blotting

The localization of the exogenous proteins was visualized via western blot following the fractionation protocol described above. Protein extracts were prepared in a final volume of 25 µl with 4X Bolt LDS sample buffer (Invitrogen) and 10X Bolt Reducing Agent (Invitrogen), and denaturated at 70°C for 10 minutes; then samples were loaded on NuPAGE Novex 4–12% Bis-Tris polyacrylamide gels (Life Technologies) and running using Bolt MOPS SDS Running Buffer 1X (Invitrogen). PageRuler Plus Prestained Protein Ladder (ThermoFisher Scientific) was used as a marker.

The polyacrylamide gel was wet transferred onto a nitrocellulose/ methanol treated-PVDF membrane using Bolt Transfer Buffer 1X (3h 150 mA constant current), or it was dry transferred using iBlot 2 NC Regular Stacks (Invitrogen) in an iBlot 2 machine (Invitrogen) set on the preset program P0 (20 V for 1 minute, 23 V for 4 minutes, 25 V for 2 minutes), according to manufacturer's instructions.

Once the transfer was completed, the membrane was removed from the sandwich assembly or the iBlot2 stack and quickly dipped in Ponceau S, then washed once by dipping in TBST (50 mM Tris, 150 mM NaCl and 0.1% Tween-20) to enhance coloration contrast. The membrane was cutting, in relation to marker bands of the ladder, below the 115 kDa band for the high molecular weights (HA-tagged constructs) and across the 50 kDa band for the mid and low molecular weights (laminin A/C and -βtubulin, respectively).

After cutting, the bands were moved to a suitable vessel placed on a rocking platform for 3 consecutive washes (10 minutes) in TBST, and nonspecific antibody binding was blocked with 5% milk in TBST at room temperature for one hour. The membranes were finally incubated overnight at 4°C with primary antibody in 5% milk in TBS-T: rabbit anti-HA (Sigma-Aldrich #H6908) and rabbit anti-Lamina A (Sigma-Aldrich #L1293) were used at 1 1:1000 dilution while mouse anti-βtubulin (Sigma-Aldrich #T8328) was used at a 1:10000 dilution.

The following day, membrane labeled with primary antibodies was washed 3 times with TBST and incubated with anti-rabbit HRP-conjugated antibody (Alexa fluor 680 (Invitrogen A-21076); Alexa fluor 488 (Invitrogen A-32731)) diluted 1:5000 in 5%milk/TBST or anti-mouse HRP-conjugated antibody (Alexa fluor 680 (Invitrogen A-21057); Alexa fluor 488 (Invitrogen A-32723)) diluted 1:10000 in 5%milk/TBST.

The secondary antibody incubation was performed one hour at room temperature protected from light on the rocking platform. The secondary antibody suspensions were discarded and the membranes were washed three times with TBST, maintained in the dark.

Images were acquired with automated settings for the channel of choice on either iBright 2 (Invitrogen) or Azure 280 (Biosystems) instruments and exported as TIF files for elaboration.

Analysis was performed with the ImageJ – FIJI software; contrast was maximized, and then the gel analysis function was used to extract profiles from the individual bands present in the single membrane slices. Areas under the profile peaks were extracted and compared after normalization for cytoplasmic/nuclear contamination expressed as a ratio between nuclear and cytoplasmic marker peak areas in each fraction. Protein nuclear localization was expressed as a percentage of the sum total area under the peaks of HA positivity in both fractions (where present) normalized for the contamination values.

4 RESULTS

4.1 Set-up and optimization of technological approaches to probe tiTreg cells epigenetic status

One of the major issues concerning the definition of the epigenetic profile of tumor infiltrating Treg cells is the limited number of cells that can be isolated from tumor sites. In order to perform epigenetic analysis of human tiTreg we therefore employed ChIPmentation and ATAC-seq protocol that take advantage to the Tn5 transposase to amplify the tagmented chromatin and allow to perform histone modification and chromatin accessibility analysis starting from few cells. The ChIPmentation protocol was initially shown to work with 10000 cells for H3K27me3 and H4K4me3, but from 10 millions of cells for H3K27ac, H3K36me3 and H3K4me1(172); while published ATAC-seq protocols were scaled down to 50000 cells(173). The number of Treg cells isolated from tumor samples ranges from 50000 to 100000, the protocols therefore required further optimization in line with our experimental settings.

4.1.1 ChIPmentation set-up and optimization

To set-up and optimize ChIPmentation on tumor infiltrating lymphocytes and in particular on tiTreg, we first performed ChIPmentation analysis for a set of four histone modifications that identify the active promoters and the enhancer regions (H3K4me3, H3K27ac, H3K4me1 and H3K27me3) and compared the same datasets generated with a standard ChIP-seq protocol. We started from the same chromatin of *in vitro* expanded CD4+ Treg cells isolated from peripheral blood of a healthy donor. For each immunoprecipitation we used 2 million cells for ChIP-seq, and 1000 or 10000 cells for ChIPmentation. Consistent with a good enrichment signal, the genome-wide distribution of each histone modifications is comparable between ChIP-seq on 2 million cells and ChIPmentation on 10000 cells: moreover, the histone modification distribution reflected their expected localization in relation to the gene body as well as start (TSS) and end (TES) sites: the localization of the active and sharp histone modifications is around the TSS while for the broad ones is spread along the gene body (**Figure 17**).

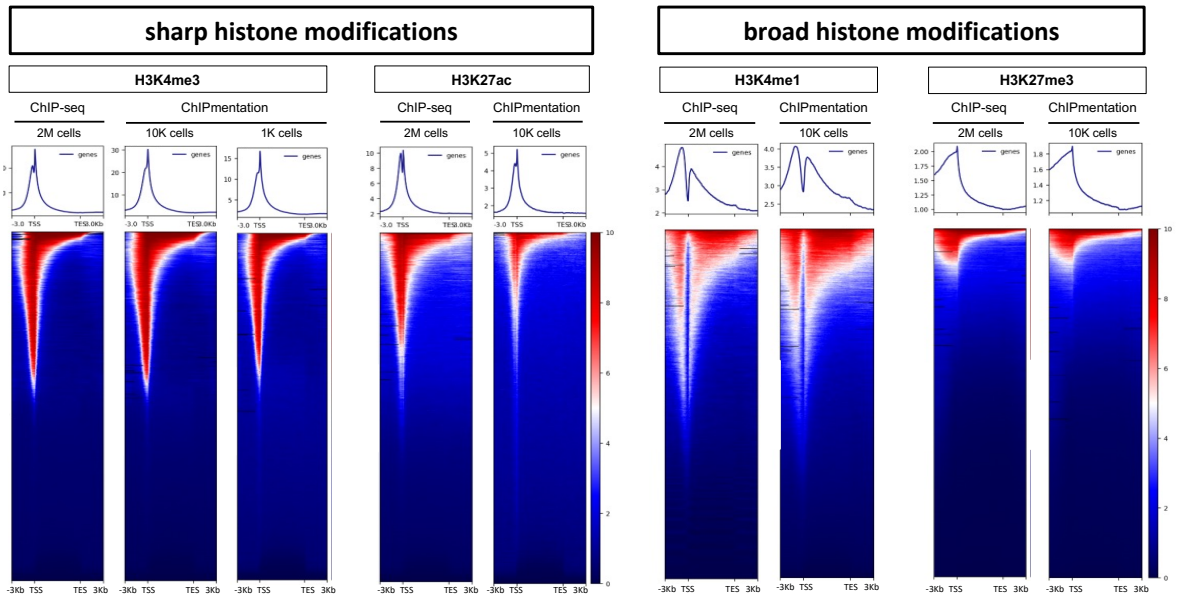


Figure 17 Genome-wide histone modification distribution. Representative density plots (top) of average intensity and corresponding heatmaps (bottom) displaying the relative distribution of H3K4me3, H3K27ac, H3K4me1, H3K36me3, and H3K27me3 signals at regions surrounding +/- 3kb of the gene body for all the genes present in GENCODEv25. Immunoprecipitation was performed on 2 million (2M) Treg cells for ChIP-seq and 1000 (1K) or 10000 (10K) Treg cells for ChIPmentation.

Correlation analysis confirmed the clustering of the same histone marks for ChIP-seq and ChIPmentation and showed the clear separation between the branches relating to the repressive marker H3K27me3 and the block of histone marks defining active regulatory regions (H3K4me3, H3K27ac, and H3K4me1) (**Figure 18**).

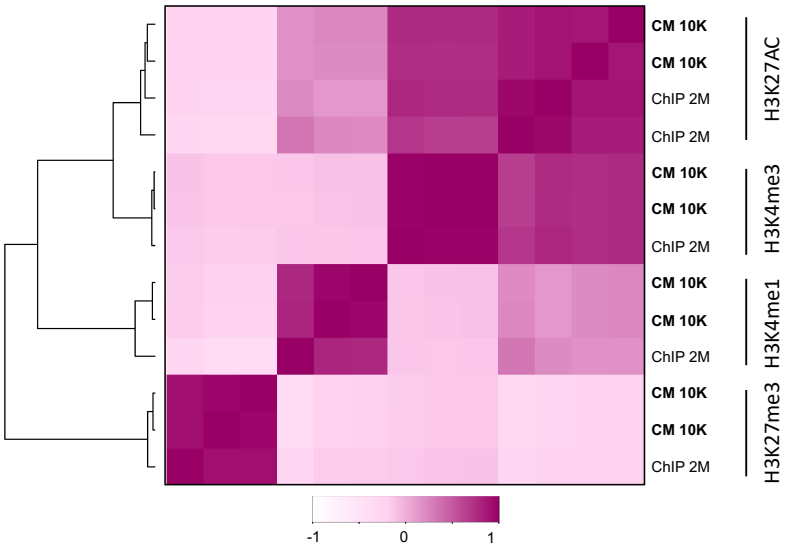


Figure 18 Histone modification correlation analysis. Pearson correlation heatmap of ChIP-seq (ChIP) and ChIPmentation (CM) data for four histone modifications (H3K27ac, H3K4me3, H3K4me1 and H3K27me3) across 10000 (10K) or 2 million (2M) in vitro expanded peripheral blood Treg cells.

We report as an example an active transcribed gene in Treg cells the ChIP and ChIPmentation traces for TGFβ receptor, an important gene for immunosuppressive function of Treg cells. The distribution of each histone modification and the information it provides about a specific genomic locus was comparable between ChIP-seq and ChIPmentation traces. Moreover, the histone modification distribution reflected their expected localization (**Figure 19**).

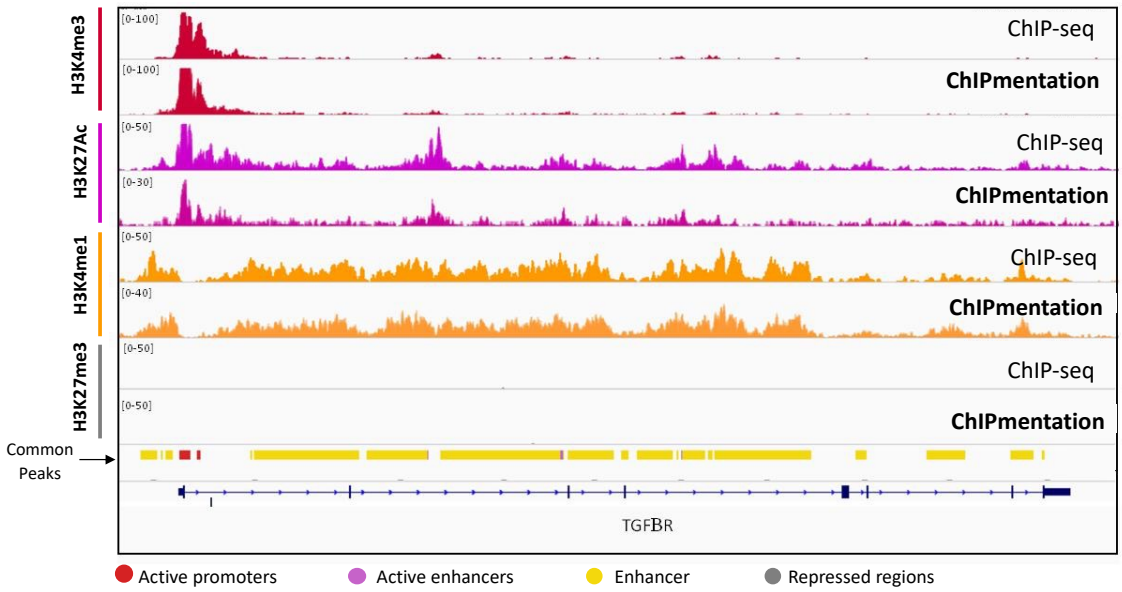


Figure 19 ChIP-seq and ChIPmentation tracks for TGFβR. ChIP-seq and ChIPmentation tracks showing H3K4me3, H3K27ac, H3K4me1 and H3K27me3 distribution for TGFβ receptor.

4.1.2 ATAC-seq set-up and optimization

ATAC-seq protocol was set-up initially using *ex vivo* CD4+ Treg cells isolated from peripheral blood of healthy donors. We first performed the ATAC-seq protocol using the same condition of the published protocol, starting from 50000 Treg cells(173). The distribution of ATAC- library on 50000 Treg cells reflected the distribution in relation to the free-nucleosome chromatin peak (200 bp) and mono, di- and tri-nucleosome chromatin peaks. (**Figure 20**, **a**).

We then scaled down the number of cells (20000, 10000, 5000, 2500, 1250 and 625 Treg cells from peripheral blood of healthy donor) while maintaining the same experimental conditions of published protocol (**Figure 20**, **b-f**). However, the

tagmentation efficiency was lower and the ATAC library of worse quality than in the experiment performed on 50000 cells.

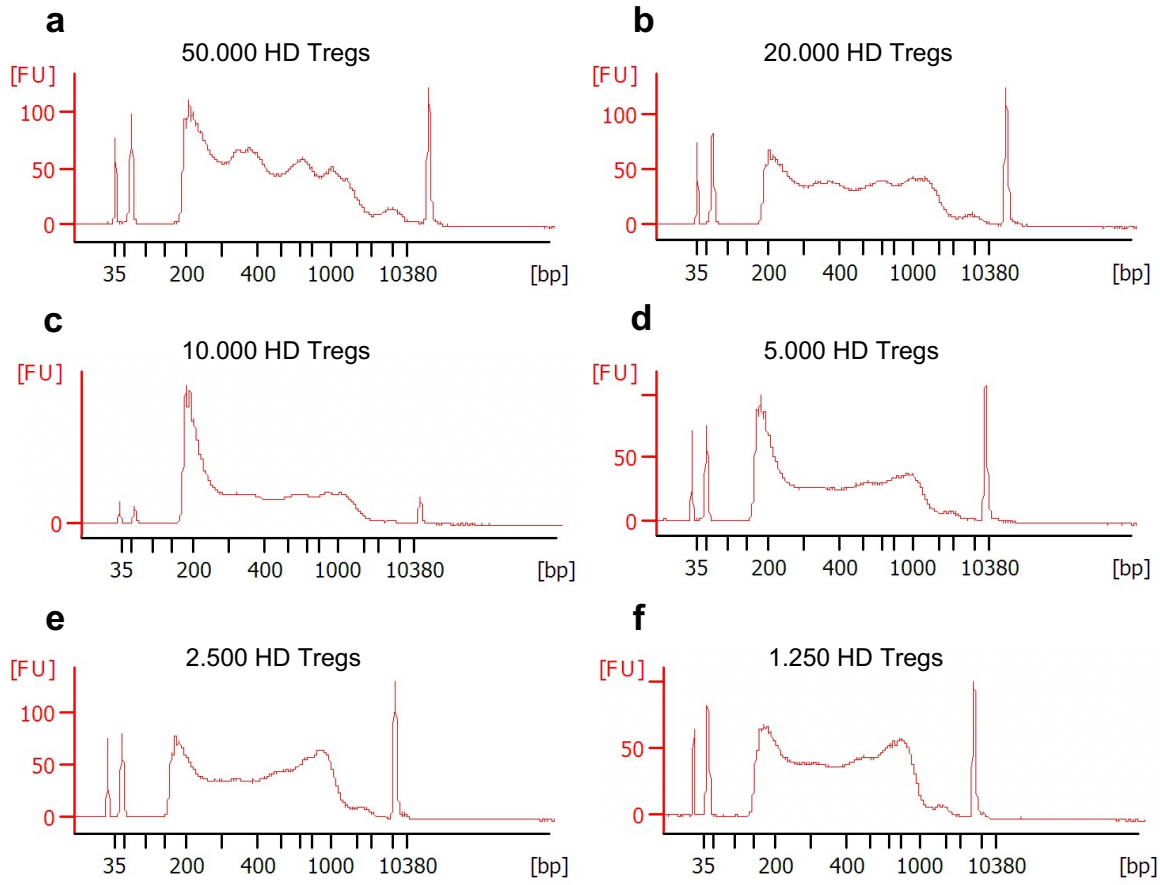


Figure 20 ATAC profiles using the Buenrostro's protocol. The electrophorese profiles, obtained by Bioanlayzer, of ATAC library performed using published protocol conditions and different amount of *ex vivo* Treg cells isolated from peripheral blood of healthy donor (HD).

We thus modified the experimental conditions and understood that the reaction volume was the factor that most influenced the tagmentation efficiency. We have progressively reduced the tagmentation reaction volumes from 25 µl of the standard protocol to 5 µl, maintaining a similar concentration of the Tn5 enzyme. By reducing the volume, we thus obtained a good tagmentation profile, similar to the expected one, even starting from 5000 Treg cells (**Figure 21**, **a**)

Following our modified protocol, we were finally able to obtain a good quality ATAC library also on Treg cells isolated from peripheral blood, tumor tissue and its normal adjacent tissue of CRC patient (**Figure 21**, **b-d**).

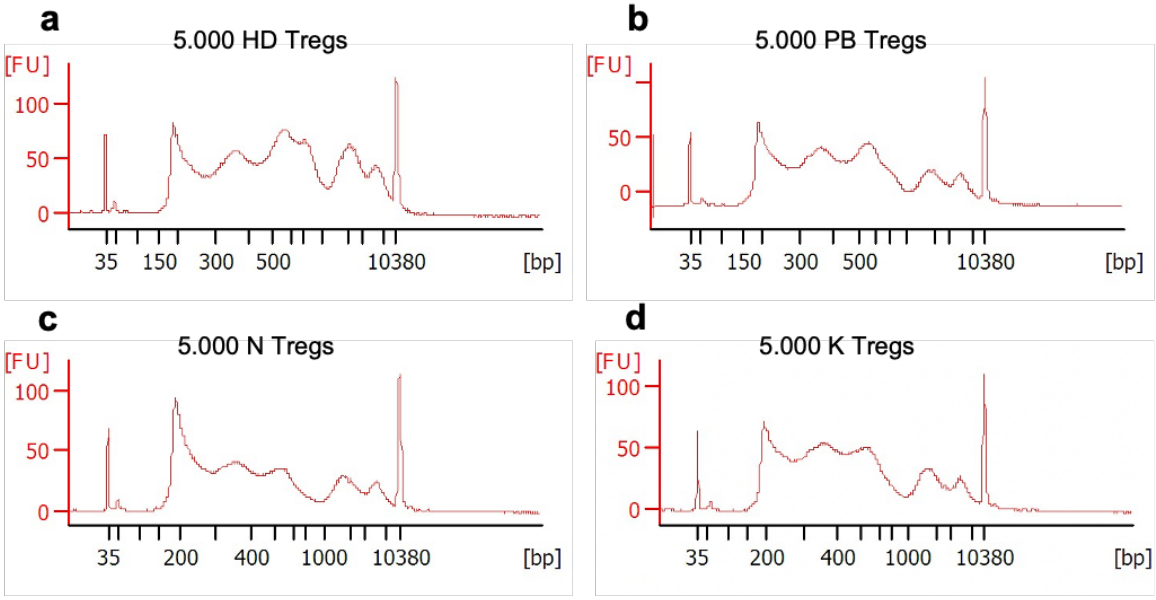


Figure 21 ATAC profiles using the modified protocol. The electrophorese profiles, obtained using Bioanlayzer, of ATAC library performed on 5000 *ex vivo* Treg cells isolated from peripheral blood (PB), normal colon tissue (N) and colorectal cancer tissue (K) of a CRC patient, after optimization of the experimental conditions.

We then sequence the libraires at 240 million reads and preliminary data analysis confirmed a good quality of ATAC-seq libraries (**Table 5**) on 5000 Treg cells. Indeed, more than 70% of the reads were aligned to the reference genome once and the mitochondrial reads were low, which is one of the parameters used to estimate the quality of the data as the reads mapping to the mitochondrial genome, which is more accessible due to the lack of chromatin packaging, have to be discarded.

Table 5 ATAC-seq quality control information

Sample	Initial reads	% Aligned reads	%multi-mapping reads	% not aligned reads	% mitochondrial reads
Treg from peripheral blood of healthy donor	236.758.772	72.91%	23.19%	3.90%	5.843
Treg from peripheral blood of CRC patient	185.316.452	70.16%	25.72%	4.12%	2.543
Treg from CRC normal tissue	199.400.126	70.39%	22.77%	6.83%	1.967
Treg from CRC tumor tissue	218.768.564	74.26%	21.88%	3.86%	3.668

The quality control revealed that reads length was lower than 100 bp, with a 45bp peak; these reads correspond to accessible chromatin regions, not protected by nucleosomes. The peak at 150bp derives from reads associated to DNA wrapped around one nucleosome, while the higher length peaks derive from DNA associated with progressively more compact chromatin regions (two nucleosome, three nucleosomes, etc.). Furthermore, the reads distribution profile of each sample was comparable to the reference one (**Figure 22**)

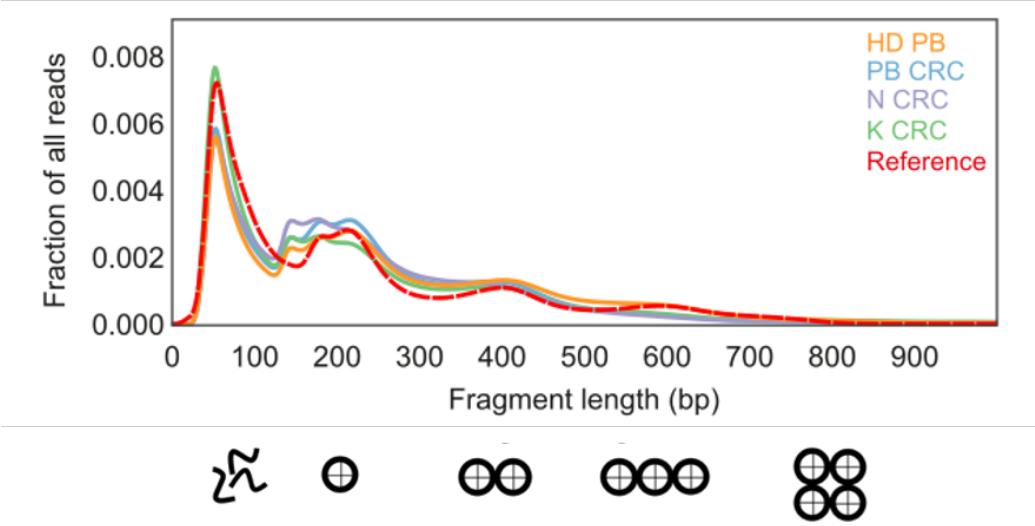


Figure 22 ATAC-seq reads size distribution. Reads size distribution of sequenced ATAC-seq performed on 5000 *ex vivo* Treg cells isolated from peripheral blood of healthy donor (HD) and from peripheral blood (PB), normal colon tissue (N) and colorectal cancer tissue (K) of the same patient.

As a further quality control of our datasets, after peak calling, we checked for ATAC-seq tracks on specific regions, we showed enrichment peaks, corresponding to accessible chromatin regions, only on constitutively expressed genes by Treg cells, such as CD3, and not on repressed genes, such as MYOD which is instead involved in muscle differentiation (**Figure 23**). Moreover, ATAC-seq tracks confirmed the presence of open chromatin region on genes specifically expressed by tiTreg cells such as CCR8 and MAGH1 (**Figure 24**)

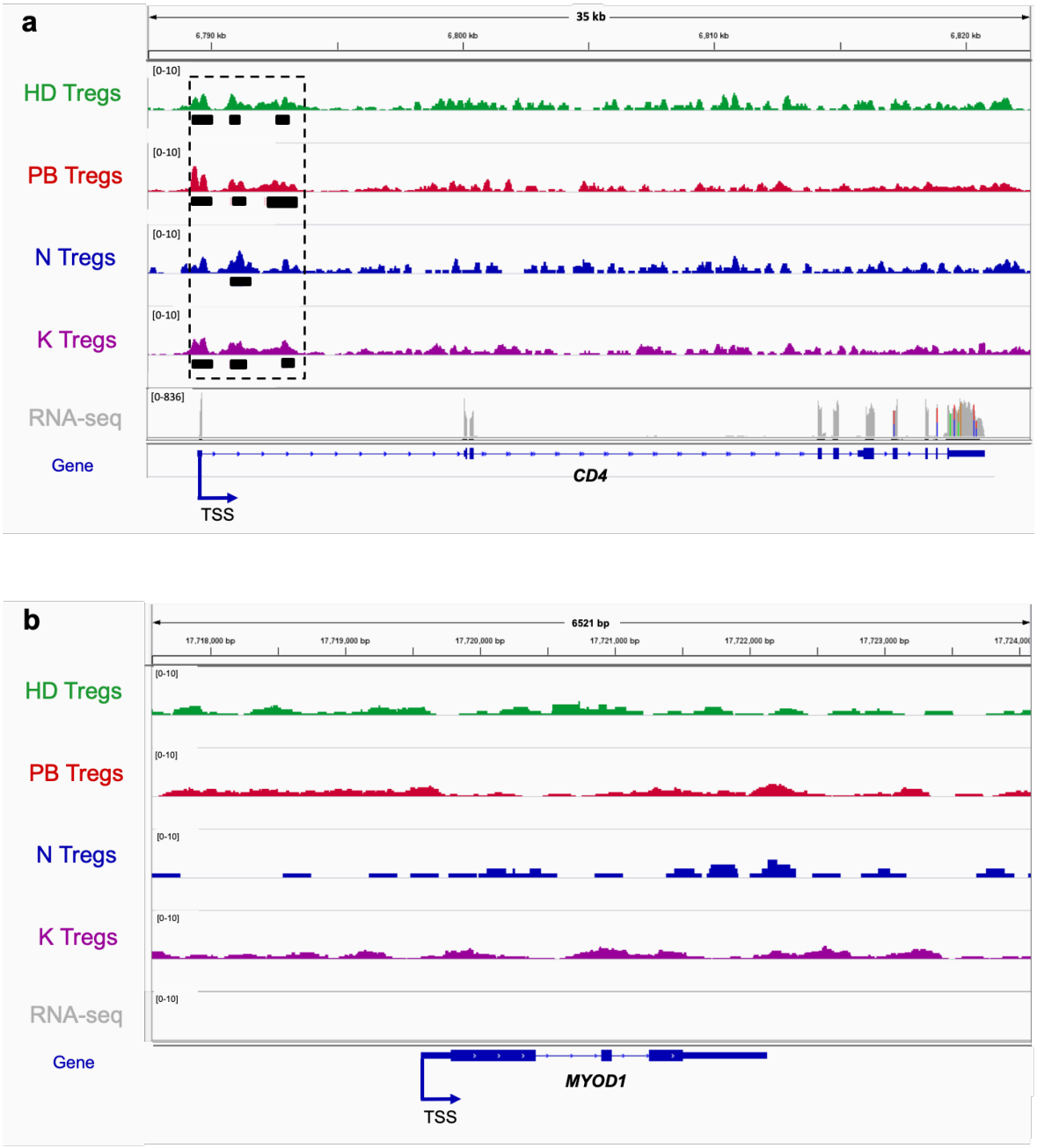


Figure 23 ATAC-seq tracks at *CD4* and *MYOD* gene loci. ATAC-seq tracks, at *CD4* and *MYOD* gene loci, obtain from *ex vivo* Treg cells isolated from peripheral blood of healthy donor (HD), peripheral blood (PB), normal tissue (N) and tumor tissue (K) of CRC patient.

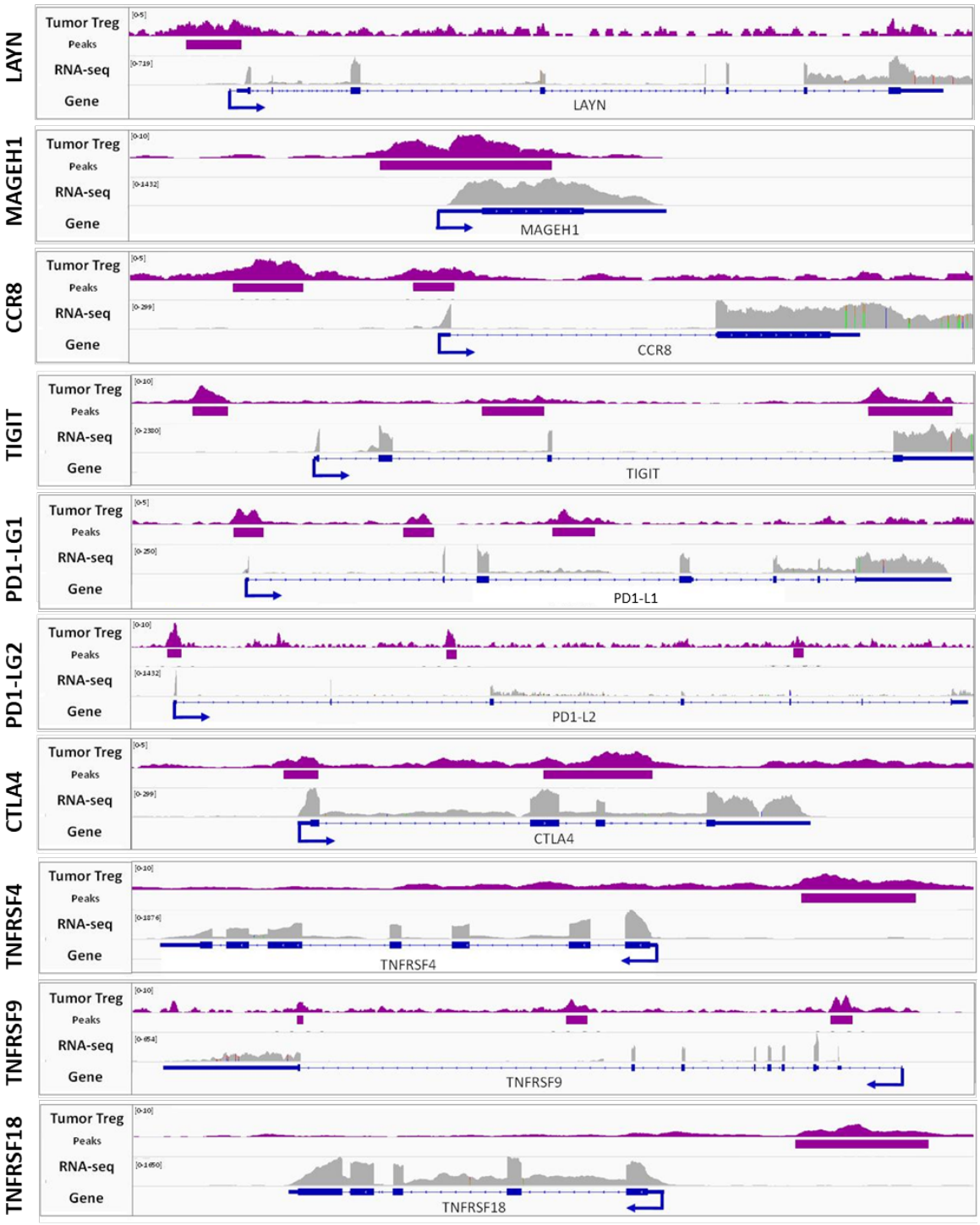


Figure 24 ATAC-seq tracks at tiTeg signature genes. ATAC-seq tracks, at different genes differentially transcribed by tiTreg, obtain from ex vivo Treg cell infiltrating CRC

4.2 Epigenetic characterization of tiTreg

In order to characterize the epigenetic landscape of human tiTreg we collected and generated 14 ATCA-seq libraries for peripheral blood, normal adjacent tissue and tumor-infiltrating Treg cells isolated from both CRC and NSCLC patients. In addition, 30 ChIPmentation libraries were generated for Treg populations obtained from NSCLC patients. In parallel with the Treg cell populations, we isolated and analyzed additional T lymphocyte populations including CD4+ conventional T cells, CD4+ Tr1 cells, CD8+ memory T cells, and CD8+ exhausted T cells. The latter populations facilitated the discovery of active enhancers (see section 4.2.2) and allow for future inter-lymphocyte comparisons in identifying potential molecular targets highly specific for tiTreg cells (see Discussion). The complete dataset comprised of a total of 53 ATAC-seq and 118 ChIPmentation libraries (**Table 6**).

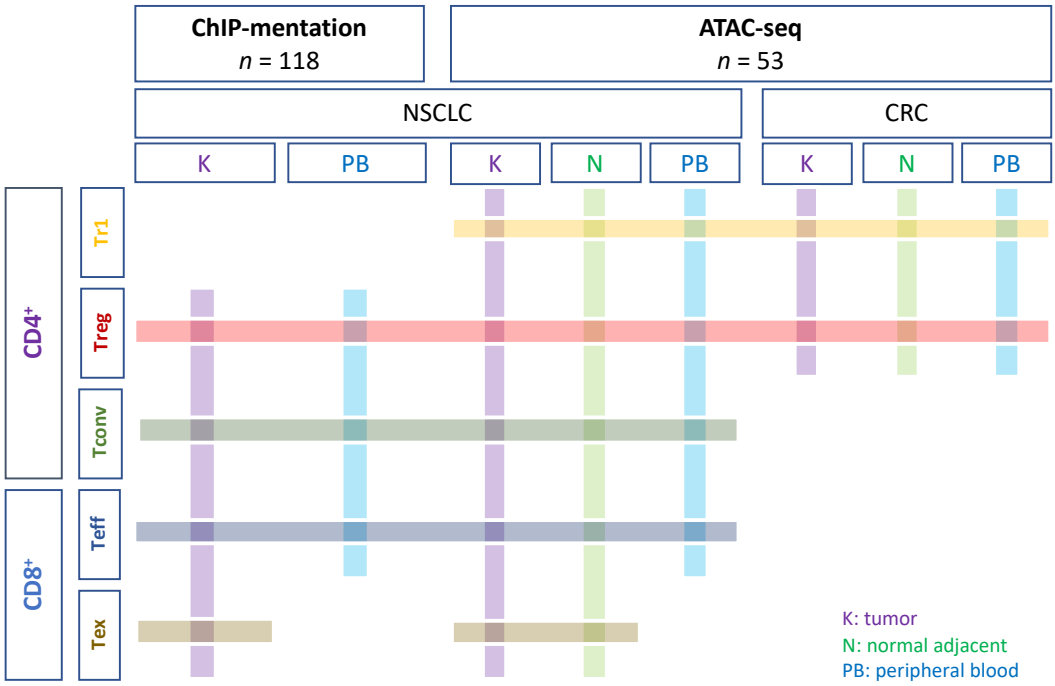


Table 6 Summary of ATAC-seq and ChIPmentation libraries. Summary of ATAC-seq and ChIPmentation libraries used in the study for different $CD4^+$ and $CD8^+$ T cell populations isolated from CRC and NSCLC patients.

4.2.1 Chromatin accessibility landscape of the CD4+ T regulatory subset

The complete ATAC-seq dataset consisted of 53 sequenced samples, derived from11 human subjects with different clinical conditions. Primary tumors were obtained from 7 NSCLC and 3 CRC patients, respectively, together with matching normal adjacent tissues and peripheral blood samples. Additionally, peripheral blood from a healthy donor subject was also collected. Five different populations of immune cells have been isolated starting from the tissue specimens (**Table 6**), as described in material and method.

To capture the chromatin accessibility landscape of tiTreg cells, we analyzed ATAC-seq data performed on Treg cells isolated from 3 NSCLC patients, 3 CRC patients and a healthy donor. (first column of **Table 7**).

		CD4+ Treg	CD4+ T Conventional	CD8+ T Exhausted	CD8+ T Memory	CD4+ Tr1
NSCLC	K	**3** (#181,#180, #186)	**4** (#180,#169,#187, #185)	**3** (#180,#169,#185)	**3** (#180,#169, #187)	**2** (#181,#184)
	N	**3** (#181,#180,#186)	**3** (#180,#169, #187)	**2** (#180,#185)	**3** (#180,#169, #187)	**2** (#181,#184)
	PB	**2** (#181,#186)	**3** (#180,#169, #187)	-	**3** (#180,#169, #187)	**2** (#181,#184)
CRC	K	**3** (#188, #183,#182)				**2** (#183,#182)
	N	**3** (#188, #183,#182)				**2** (#183,#182)
	PB	**2** (#188,#183)				**2** (#182,#183)
HD	PB	**1** (#189)				

Table 7 Summary of ATAC-seq libraries for different clinical conditions and T cell populations. ATAC-seq libraries, with subject ID in parenthesis, for different T lymphocytes isolated from tumor tissue (K), its normal adjetiant counterpart (N) and peripheral blood (PB) of NSCLC and CRC patients and also from PB of healthy donor (HD), for a total of 53 samples.

4.2.1.1 ATAC-seq quality control

The alignment of the reads to the reference genome (hg38) and the peak calling on the mapped fragments confirmed a good quality of ATAC-seq libraries (**Figure 25**). Indeed almost all samples had more than 50 million mapped reads, the minimum cut-off suggested for open chromatin detection and differential analysis, whereas

some of them had nearly 200 million reads, which is recommended for TF computational footprinting(174). Concerning the peak calling, almost all the samples had a number of peaks ranging from 25,000 to 50,000 with some samples reaching 75,000 called peaks (**Figure 25**, bottom panel).

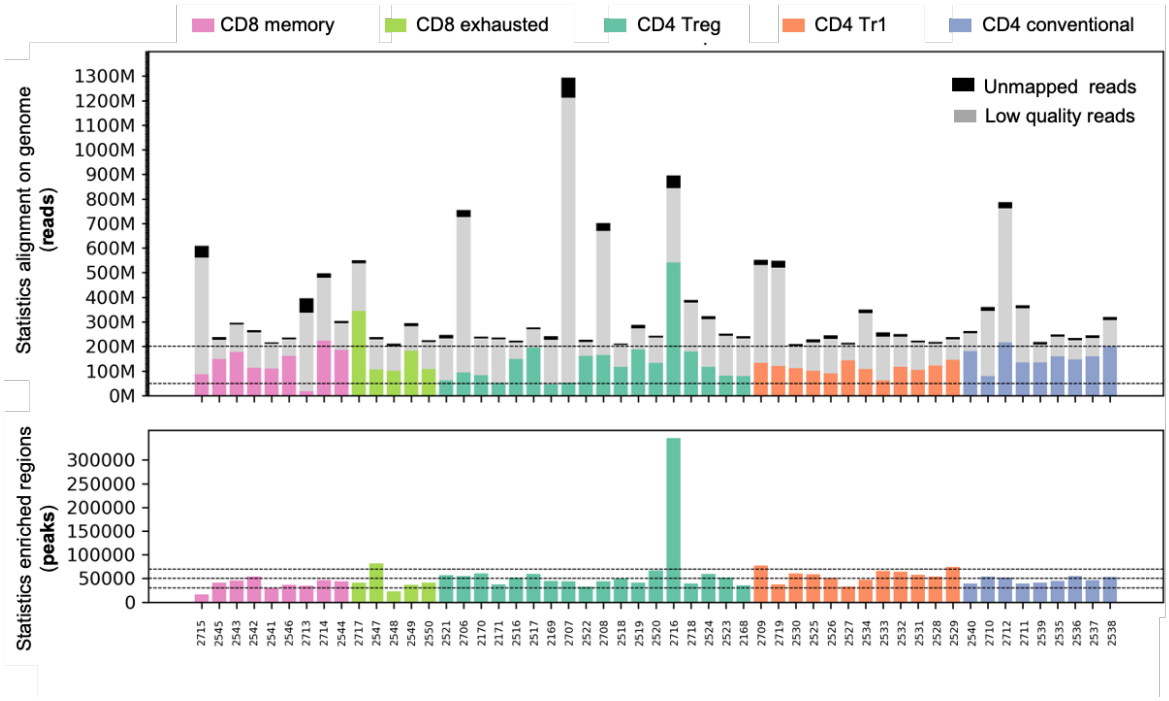


Figure 25 Summarized report of ATAC-seq data quality control. For each sample the number of reads (top) and the number of called peaks (bottom) are represented in bar plots. The black portion of each bar represents the fraction of unmapped reads, the grey one the fraction of low-quality reads (i.e., not uniquely mapped, duplicate reads and reads originated from mitochondrial DNA or repetitive blacklisted regions) that is filtered out in the post- alignment processing. Lastly, the colored bars represent the final pool of reads that passed all the quality controls. Pink color labels CD8 memory T cells, light green color CD8+ exhausted T cells, green color CD4+ Treg cells, orange color CD4+ Tr1 cells and the blue one CD4+ conventional T cells. The dotted lines indicate the important thresholds for each of the metrics, as suggested by ENCODE guidelines for good ATAC-seq experiments.

The fragment length distribution was compared across all the samples to evaluate the presence of a fragment length distribution pattern typical of ATAC-seq data. As expected(174), the majority of samples showed decreasing and periodic peaks in the first 100 bp, at ~ 200 bp and 400 bp, corresponding to nucleosome-free regions (NFR), and mono- and di-nucleosomes, respectively (**Figure 26**). Collectively, these results indicate an overall good quality of our sequencing reads, aligned reads and called peaks, permitting further downstream analyses of the ATAC- seq data.

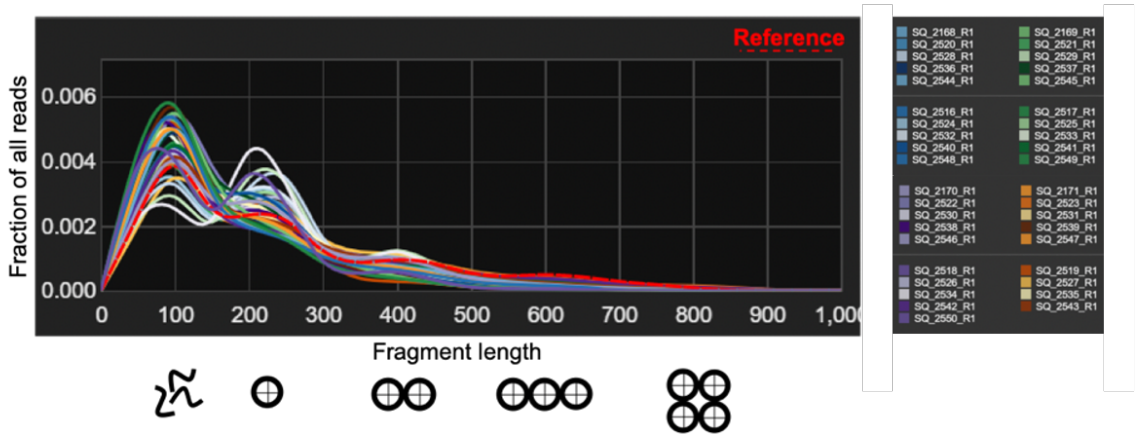


Figure 26 Plot reporting the ATAC-seq fragment length distribution. Decreasing and periodic peaks in the first 100 bp, at ~ 200 bp and 400 bp correspond to nucleosome-free regions, and mono- and di-nucleosomes, respectively.

4.2.1.2 Sample-to-sample distances based on chromatin accessibility

To determine whether CD4 Treg cells isolated from different tissues have different chromatin accessibility profiles, we focused our analysis on the fully paired samples in our ATAC-seq dataset, consisting of 12 peripheral blood (PB), normal tissue-infiltrating (NAT) and tiTreg samples for 4 patients. A consensus peakset of 71,160 open chromatin regions (OCR) was generated by merging overlapping ATAC-seq peaks present in at least two replicates. We first explored the sample-to-sample relationships, by performing PCA on CD4+ Treg populations alone, including tiTreg samples from two NSCLC and two CRC patients, their matched normal-tissue infiltrating Treg counterparts, and peripheral blood Treg cell samples. Notably, the first principal component (PC1) captured the majority of the variance (89%) clearly distinguishing tumor-infiltrating from peripheral blood Treg cells. Normal Tregs were grouped in between and were separated from tiTreg by PC2, which only accounted for a small percentage of the total variance (**Figure 27**, **a**). These results were also confirmed by a hierarchical clustering analysis of the sample-to-sample distances (**Figure 27**, **b**). From the dendrogram we could clearly observe that samples from the same tissue of origin (tumor, normal or peripheral blood) were clustered together, independently of the type of tumor. In addition, tiTreg were clustered separately from their counterparts in normal tissue and peripheral blood. This is in agreement with the distinct transcriptional profile of tiTreg(99,100) and suggests that

their well-defined gene expression signature is likely to be driven by a specific epigenetic reprogramming.

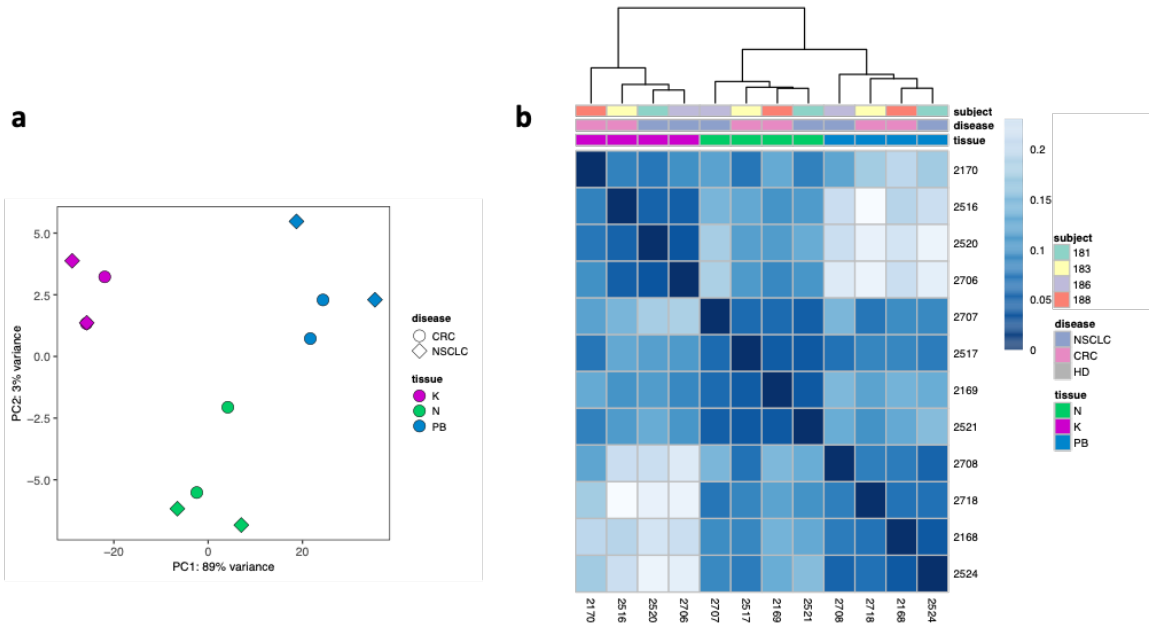


Figure 27 Sample-to-sample distances based on chromatin accessibility. (**a**) PCA performed on rlog-normalized counts for all the Treg samples using the first 100 highly variable intervals after batch effect correction with Limma. Color labelling represents the tissue of origin, whereas different shapes indicate the cancer types; (**b**) Hierarchical clustering analysis with Euclidean distances performed on Treg samples calculating the maximum distances between samples. PB: peripheral blood, N: normal adjacent tissue, K: tumor

4.2.1.3 Comparison of tumor Treg infiltrating different cancer types

We first asked whether Treg cells infiltrating diverse cancer types have a similar epigenomic profile. To this end, we initially performed differential accessibility analysis of tiTreg versus PB Treg separately for CRC and NSCLC. Both analyses revealed around ~10K differential accessible regions (DARs) at *P* adjusted < 0.05, with similar number of significantly gained and lost regions in CRC and NSCLC tiTreg cells (**Figure 28**). In comparing the log2 fold-change of DARs (**Figure 29**, **a-b**), we found a good concordance in the gained and lost regions identified for CRC or NSCLC tiTreg cells. Of the gained regions in CRC, 64.5% are also gained in NSCLC, whilst 33.1% are not significantly gained but show a trend of increased accessibility. Importantly, only 2.4% of gained regions in CRC display reduced accessibility in NSCLC. The same applied to lost regions in CRC tiTreg cells. Further confirming the common epigenomic profile of tiTreg cells, a direct comparison

between CRC and NSCLC tiTreg cells revealed only a small number of DARs (**Figure 29**, **c**). Pathway analysis on these differential regions did not return any significant terms (data not shown).

Overall, these findings suggest that tiTreg cells display a shared chromatin accessibility profile independently of the tumor type they infiltrate. In light of this, tiTreg cells isolated from CRC and NSCLC were considered together in downstream analyses.

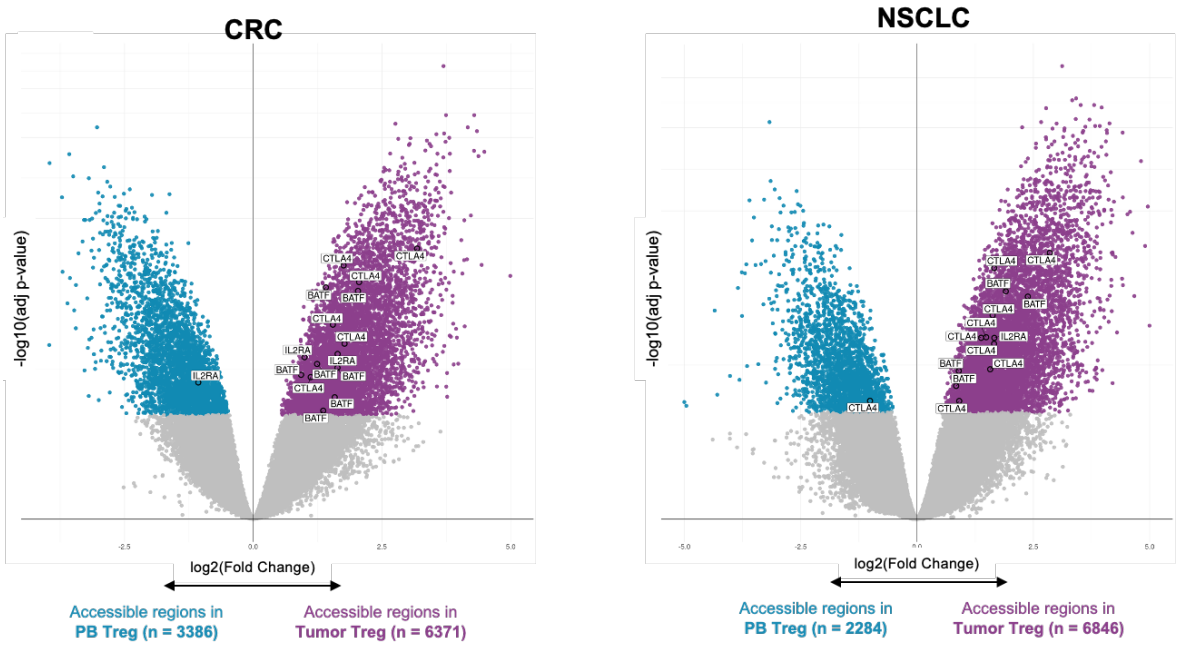


Figure 28 Differential accessible regions for tiTReg vs PBTreg. Volcano plot showing differentially accessible regions (DARs) for TI-Treg cells vs PB-Treg cells separately in CRC (left) and NSCLC (right) patients.

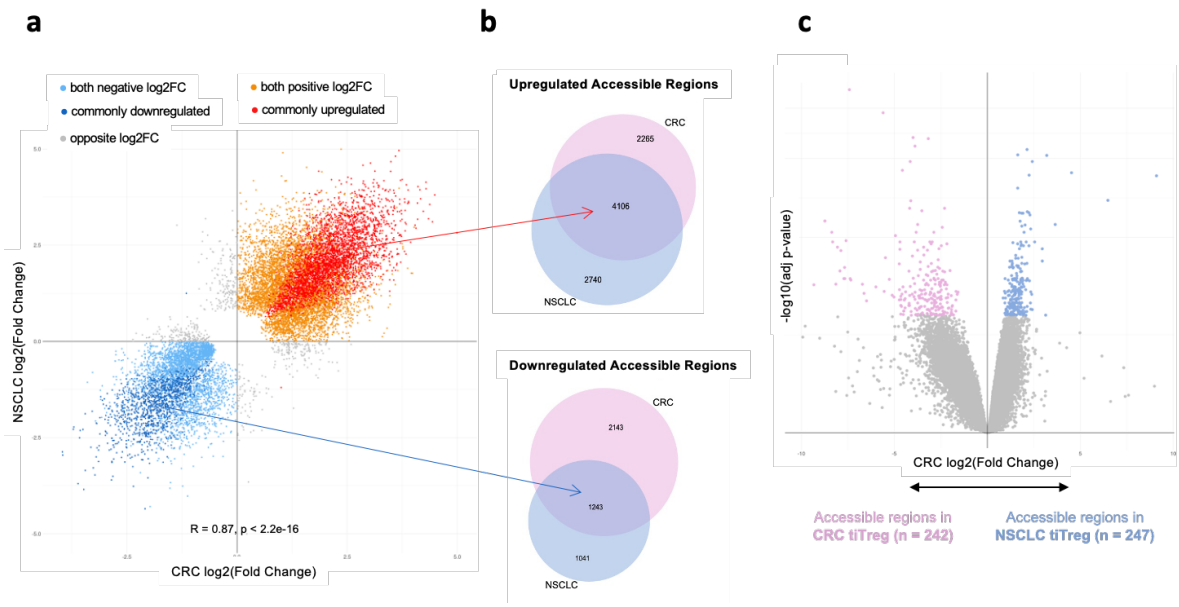


Figure 29 Comparison of DARs for CRC and NSCLC tiTreg cells. (**a**) Scatterplot displaying the correlation of log2(Fold Change) of DARs between tiTreg and PB-Treg in CRC (x-axis) and NSCLC (y-axis). Each dot represents a DAR coloured according to its accessibility concordance (gain or loss) between CRC and NSCLC tiTreg cells. Regions significantly gained or lost in both CRC and NSCLC tiTreg cells are coloured in red and blue, respectively. Non-concordant regions gained in CRC but lost in NSCLC (or vice versa) are coloured in grey. Regions gained in one cancer type that display increased accessibility also in the second cancer type, though not with statistical significance, are coloured in orange. Similarly, regions lost in one cancer type that show reduced accessibility also in the second cancer type are coloured in light blue. (**b**) Venn diagrams displaying the overlap of DARs in tiTreg cells vs PB Treg between CRC and NSCLC separately for gained (top) and lost (bottom) regions. (**c**) Volcano plot showing DARs for NSCLC tiTreg cells vs CRC tiTreg cells.

4.2.1.4 Identification of a tiTreg-specific chromatin accessibility profile

The next step was to identify genomic regions with significant differences in chromatin accessibility between i) tumor-associated and normal-tissue Treg cells and ii) tumor-associated and peripheral blood Treg cells. Comparison of ATAC-seq data between tiTreg and NAT Treg revealed differences in open chromatin regions, with 3863 and 1484 regions (*P* adjusted < 0.05) showing increased accessibility in tiTreg and NAT Treg, respectively (**Figure 30**, **a**). As expected, the number of gained and lost regions in tumor- versus tissue-infiltrating Tregs (DARs = 5347) is lower compared to the much higher number of DARs in tiTreg versus PB Tregs (DARs = 18,950). To identify target genes, we annotated the open chromatin regions to their putative interacting promoter regions using H3K27ac HiChIP data on human PB Treg(165) as well as based on the nearest and/or overlapping genes.

The DARs likely represent regulatory elements accessible to TFs that are involved in the regulation of gene expression of specific targets. The majority are localized either inside intronic regions, across the gene body or in intergenic locations, whereas a smaller fraction was assigned to exons, promoters/TSS and TTS (**Figure 30**, **b**). The observed proportions are consistent with the fact that exonic regions constitute a very small fraction of the genome(175). Considering the overall distribution of our regions, mainly distal from TSS, it is likely that a great fraction of the DARs map to distal regulatory elements such as enhancers, silencers and distal promoter elements. However, due to the complexity of the epigenetic regulation additional information is required, such as histone modifications, to confirm the real identity of the regulatory elements.

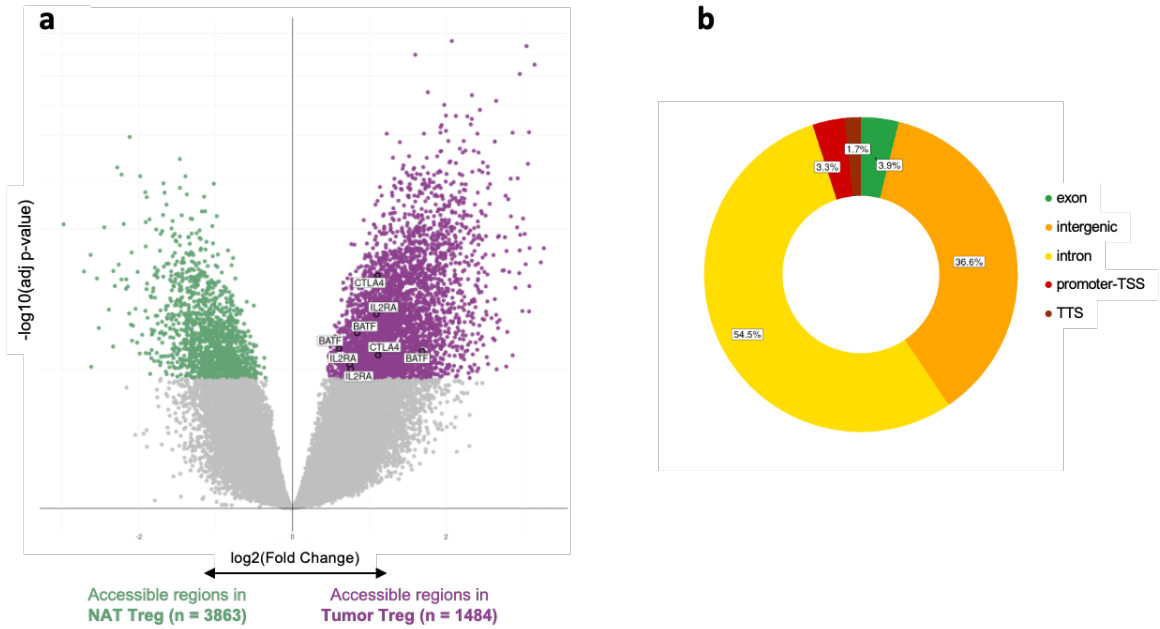


Figure 30 tiTreg-specific chromatin accessibility profile. (**a**) Volcano plot displaying gained DARs in tiTreg cells (purple dots) (CRC and NSCLC combined) and gained DARs in normal-tissue Treg cells (green dots). (**b**) Donut chart showing relative proportion of different genomic features overlapping with DARs (both up- and down- regulated) in tiTreg samples

4.2.2 Identification of tiTreg enhancer regions based on histone modification profiling

Our complete ChIPmentation dataset consisted of 118 sequenced samples, derived from primary tumors of 4 NSCLC patients and peripheral blood of four healthy donor subjects. Four different populations of immune cells (CD4 Treg, CD4 Conventional, CD8 Memory and CD8 Exhausted) were isolated starting from the tissue specimens (**Table 6**), as described in material and methods. Data for a complete set of the most commonly used histone marks was generated including histone marks defining active regulatory regions (H3K4me3, H3K27ac, and H3K4me1), the elongation marker H3K36me3, and the repressive marker H3K27me3.

To identify active enhancers in tiTreg cells, we analyzed five histone modifications (H3K4me3, H3K27ac, H3K4me1, H3K36me3, and H3K27me3) derived from ChIPmentation data performed on Treg cells isolated from three NSCLC patients and four healthy donors (**Table 8**).

	H3K4me3	H3K4me1	H3K36me3	H3K27me3	H3K27ac
tiTreg	3	3	1	2	3
PBTreg	4	3	3	4	4

Table 8 Treg ChIPmentation libraires. ChIPmentation libraries of Treg populations for independent biological replicates across five histone modifications.

4.2.2.1 ChIPmentation quality control

The alignment of the reads to the reference genome and the peak calling on the mapped fragments confirmed the good quality of the ChIPmentation libraries (**Figure 31**). Indeed, almost all samples had between 10 and 20 million of mapped reads. Concerning the peak calling, almost all the samples had a number of peaks ranging from 20,000 to 40,000 (**Figure 31**, bottom panel). In particular, the overall number of peaks for broad histone modifications, such as H3K36me3 and H3K27me3, was greater than 30,000, and for the sharp ones was greater than 20,000, indicating the good quality of samples for downstream analysis.

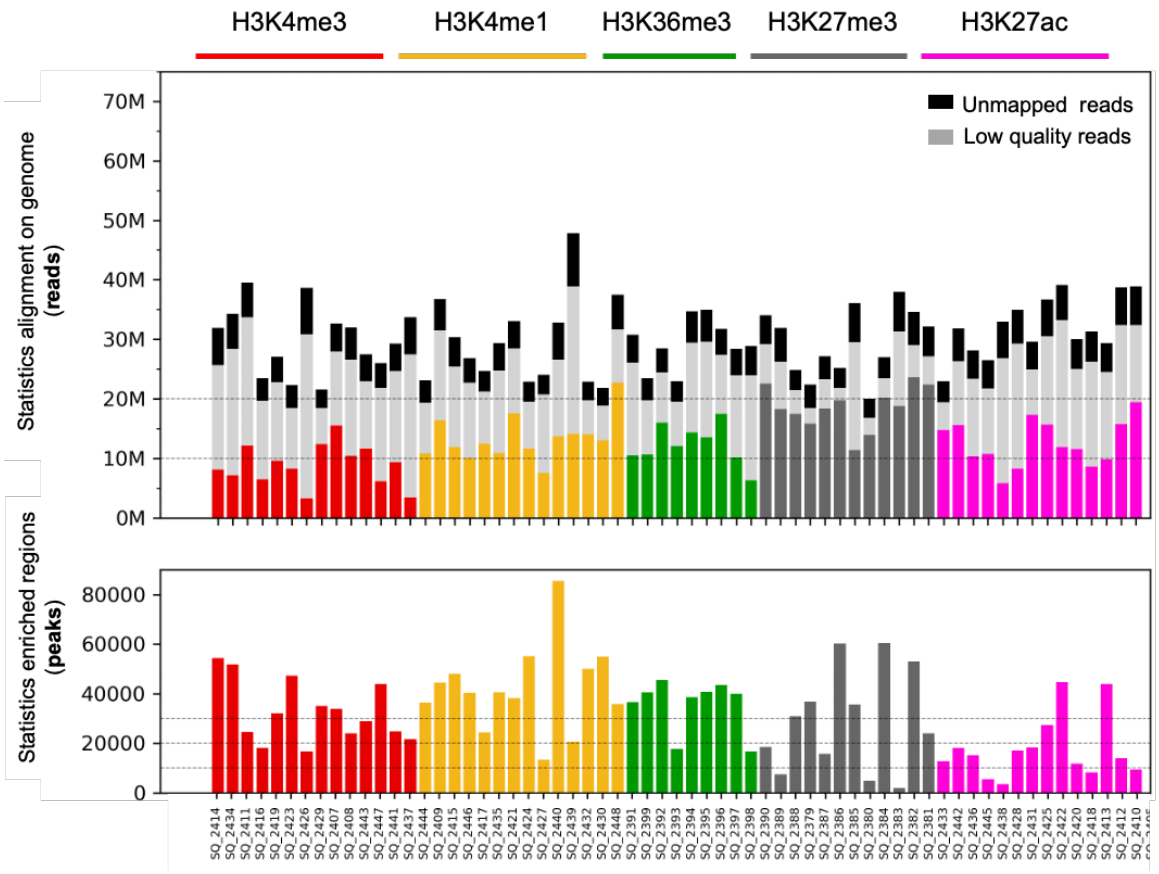


Figure 31 Summarized report of ChIPmentation data quality control. For each sample the number of reads (top) and the number of called peaks (bottom) are represented in bar plots. The black portion of each bar represents the fraction of unmapped reads, the grey one the fraction of low-quality reads (i.e., not uniquely mapped, duplicate reads and reads originated from mitochondrial DNA or repetitive blacklisted regions) that is filtered out in the post- alignment processing. Lastly, the colored bars represent the final pool of reads that passed all the quality controls. Red color labels H3K4me3, yellow color H3K4me1, green color H3K36me3, gray color H3K27me and pink color H3K27ac.

4.2.2.2 De novo chromatin state discovery using chromHMM

To provide a systematic characterization of human tiTreg at the epigenomic level, a multi-factorial integrative analysis of genome-wide ChIPmentation data for a core set of five histone modifications (H3K4me3, H3K27ac, H3K4me1, H3K36me3, and H3K27me3) was performed. The global distribution of histone marks was consistent with their expected localization in relation to the transcription start (TSS), and end (TES) sites as well as the gene body (as described in 4.1.1).

To capture the epigenomic layer of NSCLC in a systematic manner rather than based on a single epigenomic feature, we implemented machine learning approaches to perform *de novo* chromatin state characterization on the complete

ChIPmentation data for Treg, including four ChIPmentation data for peripheral blood Treg and three NSCLC infiltrating Treg (**Table 8**). To strengthen our *de novo* discovery strategy, we built the chromatin state model including additional ChIPmentation-seq data for $CD4^+$ and $CD8^+$ T cells isolated from tumor tissues of NSCLC patients and PB of healthy donors (**Table 6**).

Using ChromHMM(167), we explored the combinatorial patterns of the five histone marks in a 15-state model and predicted specific genomic features with high resolution and robustness across samples. The heatmap in **Figure 32** shows the frequency in which different histone modifications are co-present in the same genomic region (histone marks emission probability). Based on the co-presence or absence of different histone marks, each state was assigned an annotation term in accordance with the Roadmap Epigenomics Consortium nomenclature(176). We defined three promoter states (Active, Flanking, and Flanking Downstream TSS) displaying strong enrichment for H3K4me3, and either H3K27ac or H3Kme1. The state with an enrichment of H3K4me1 and H3K27ac was annotated as “Active Enhancer”, whilst the additional presence of H3K36me3 over the gene body defined “Genic Enhancers”. As “Weak Enhancers”, we defined the state characterized by the presence of H3K4me1 alone. The transcription and repression states were characterized by the presence of H3K36me3 and H3K27me3, respectively. The state with enrichment for both the active mark H3K4me3 and the repressed mark H3K27me3 was defined as “Bivalent TSS”. “Quiescence” states marked regions without any significant enrichment of histone marks. To confirm the robustness of our results, we verified that the proportion of each chromatin state was comparable across all T lymphocyte samples (**Figure 33**).

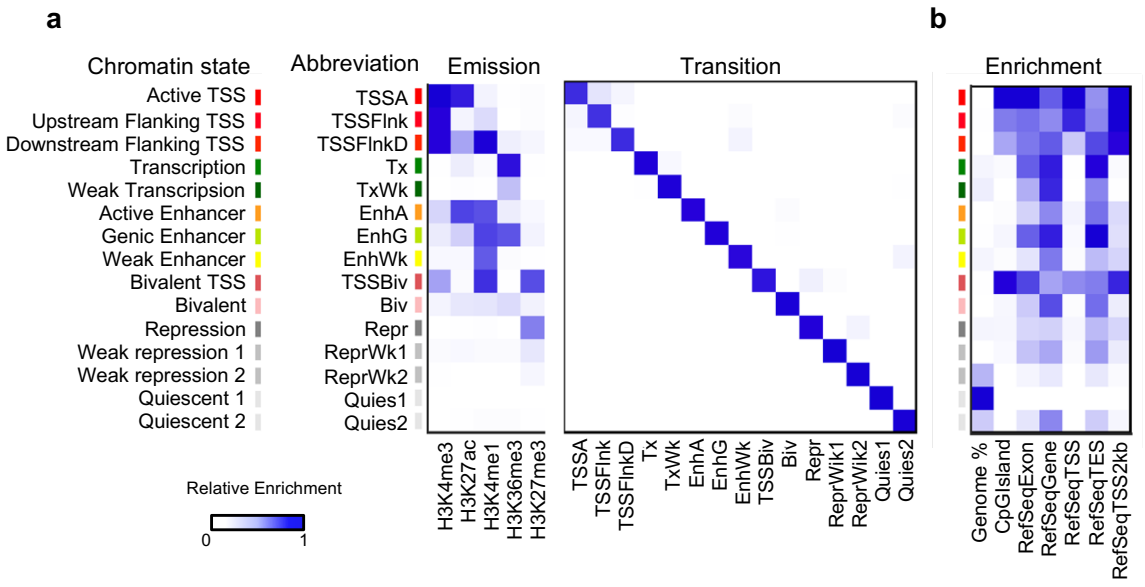


Figure 32 De novo chromatin state characterization of tumor and PB T lymphocytes. (**a**) Combinatorial pattern of histone marks in a 15-state model using ChromHMM. The heatmaps show the frequency of the histone modifications found in each state (Emission plot, left) and the probability that a state is found in the proximity of another state (Transition plot, right). A darker blue color corresponds to a greater fold enrichment. TSSA, Active TSS; TSSFlnk, Flanking TSS; TSSFlnkD, Flanking Downstream TSS; Tx, Elongation; TxWk, Weak elongation; EnhA, Active enhancer; EnhG, Genic enhancer; EnhWk, Weak enhancer; TSSBiv, Bivalent TSS; Repr, Repressed; ReprWk, Weak repressed; Quies, Quiescent. (**b**) Enrichment of each chromatin state for specific types of functional elements, such TSS, TES, exons, gene bodies. Representative heatmap for a tiTreg replicate.

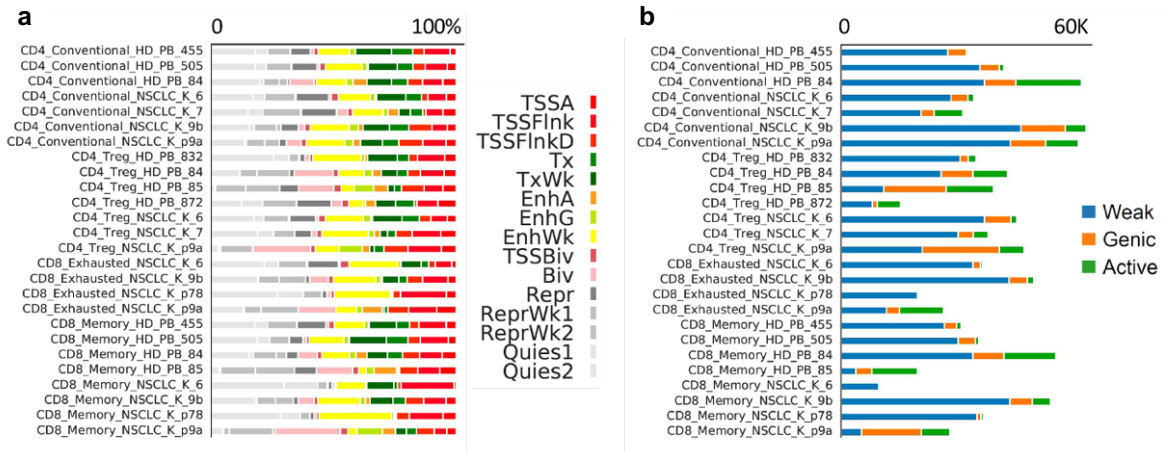


Figure 33 Summary of ChromHMM-defined states in T lymphocytes. (**a**) Distribution of the 15 ChromHMM states and (**b**) number of active, genic, and weak enhancers for each replicate and T cell population.

The histone modification pattern of CD25, a marker of Treg cells, is an example of an open and active chromatin profile that supports gene expression (**Figure 34**, **a**). Active transcription in both tumor and PB Treg is indicated by a ChromHMM profile that associates with active chromatin states around the TSS, both active TSS and active enhancer states. This is supported by the presence of active histone marks (H3K4me3/H3K4me1/H3K27ac) at the TSS flanking region (**Figure 34**, **a**; boxed area).

On the contrary, CD20 (**Figure 34**, **b**), a marker of B cells, is actively silenced in Treg cells evident by the loss of active histone marks (H3K4me4, H3K4me1, H3K27ac and H3K36me3) and the accumulation of the H3K27me3 repressive mark. **Figure 35** shows instead an example of differential epigenomic regulation at the *CCR8* locus, a gene that is highly expressed in tumor Treg compared to PB Treg(99,100).

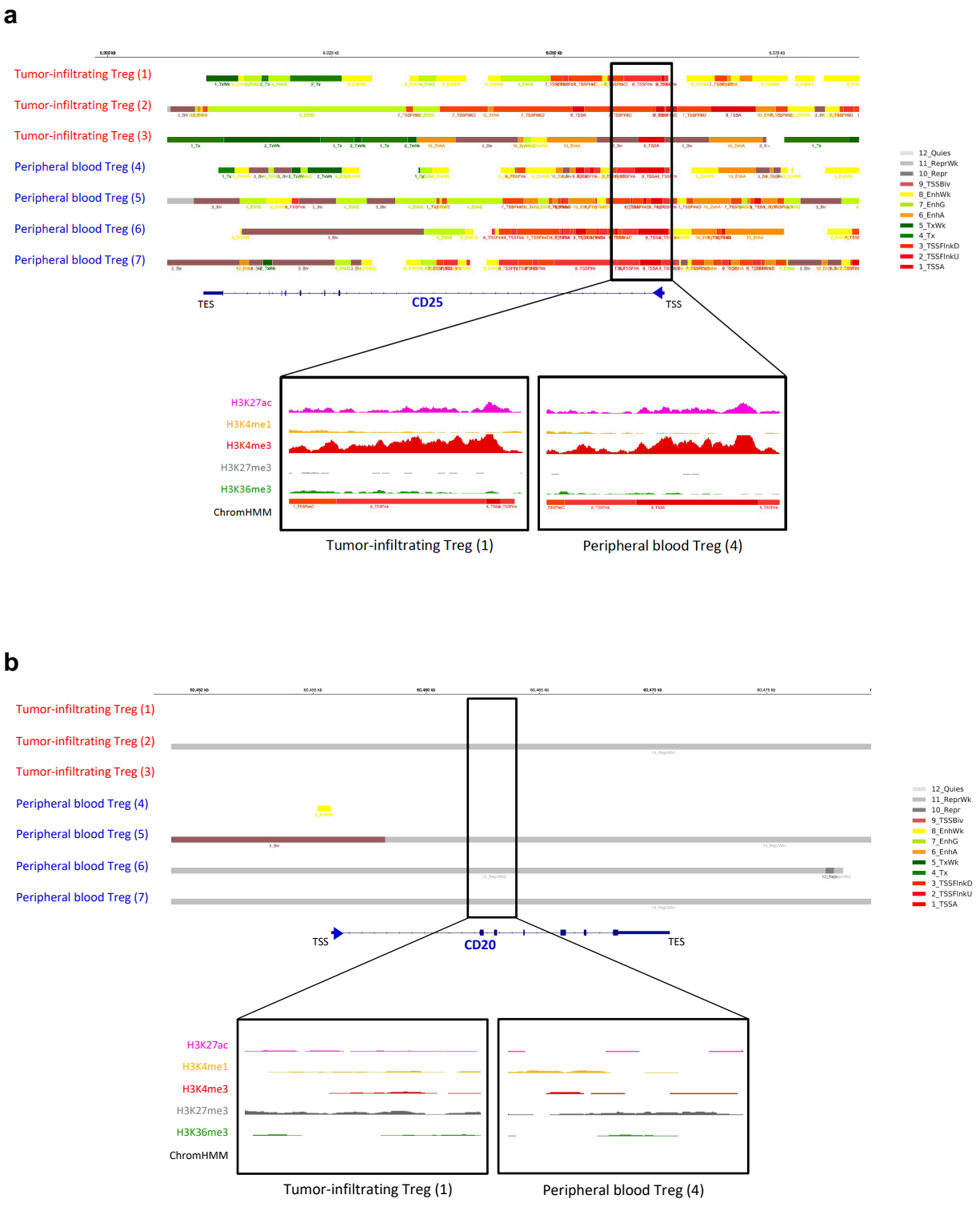


Figure 34 Epigenomic overview of the (a) CD25 and (b) CD20 gene loci. ChromHMM chromatin states for four PB (blue) and three tumor (red) Treg samples from NSCLC patients. The ChromHMM tracks denote regions identified as promoters (red), active enhancers (orange), weak enhancers (yellow), genic enhancers (lime green), elongation (green), repressed (gray) or quiescent (white) states. The expanded regions show the H3K27ac, H3K4me1, H3K4me3, H3K27me3, and H3K36me3 profiles for two samples representative of tiTreg and PB Treg, respectively.

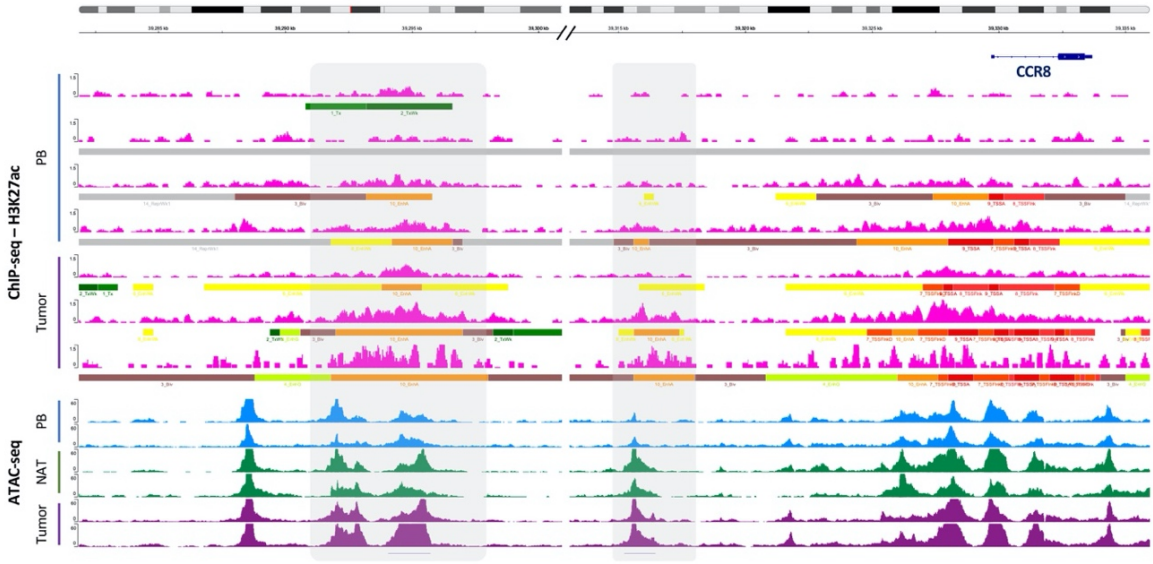


Figure 35 Epigenomic overview of the CCR8 gene locus. (Top panel) Representative tracks of H3K27ac profiles (fuchsia) and ChromHMM chromatin states for four PB (blue) and three tumor (magenta) Treg samples from NSCLC patients. The ChromHMM tracks denote regions identified as promoters (red), active enhancers (orange), weak enhancers (yellow), genic enhancers (lime green), elongation (green), repressed (gray) or quiescent (white) states. (Bottom panel) The ATAC-seq tracks show the profiles of accessible chromatin in PB (blue), normal tissue (green), and tumor (magenta) Treg cells from NSCLC patients. The shaded regions indicate active enhancers upstream of the CCR8 gene.

4.2.2.3 *Identification of active enhancers in Treg cells*

We next sought to gain further insights into the Treg enhancerome (**Figure 36**). Using the ChromHMM-defined "Active" and "Genic" enhancer states, enriched in H3K27ac, we identified a total of 13,050 enhancers observed in at least two peripheral blood and/or tumor Treg. Of these, 720 enhancers were specifically activated in tiTreg, whilst 1309 were present only in PB Treg (**Figure 37**, **a**). To identify common regulatory elements in tumor Tregs, we looked at the concordance of enhancers across tiTreg replicates. Among all enhancers identified in tiTreg, 17,04% were common to all tiTreg, 53,14% were shared by at least two replicates and 46,8% were unique to one tiTreg (Figure 37 B). Unsupervised clustering of the H3K27ac signals for the 13,050 enhancers revealed that the enhancerome profile of tiTreg is distinct from that of PB Treg (**Figure 37**, **c**).

Overall, we exploited machine learning approaches to perform a systematic reconstruction of chromatin states in Treg cell populations. In doing so, we created a reference resource of active regulatory elements (including promoters and enhancers), as well as elongated, repressed and quiescent regions, and further focused on the characterization of the Treg active enhancers to provide leads on the transcriptional regulation of tiTregs.

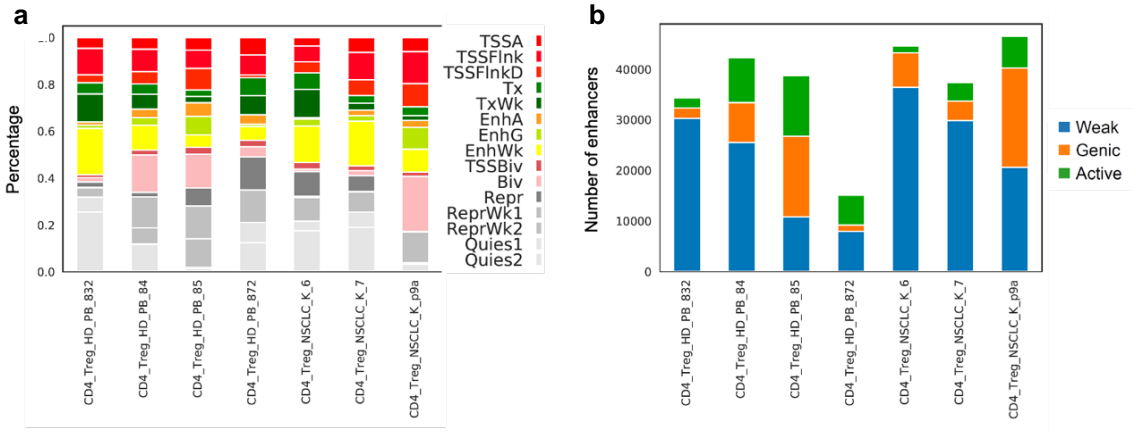


Figure 36 Summary of ChromHMM-defined states in Treg cells. (A) Distribution of the 15 ChromHMM states and (B) number of active, genic, and weak enhancers for each Treg replicate.

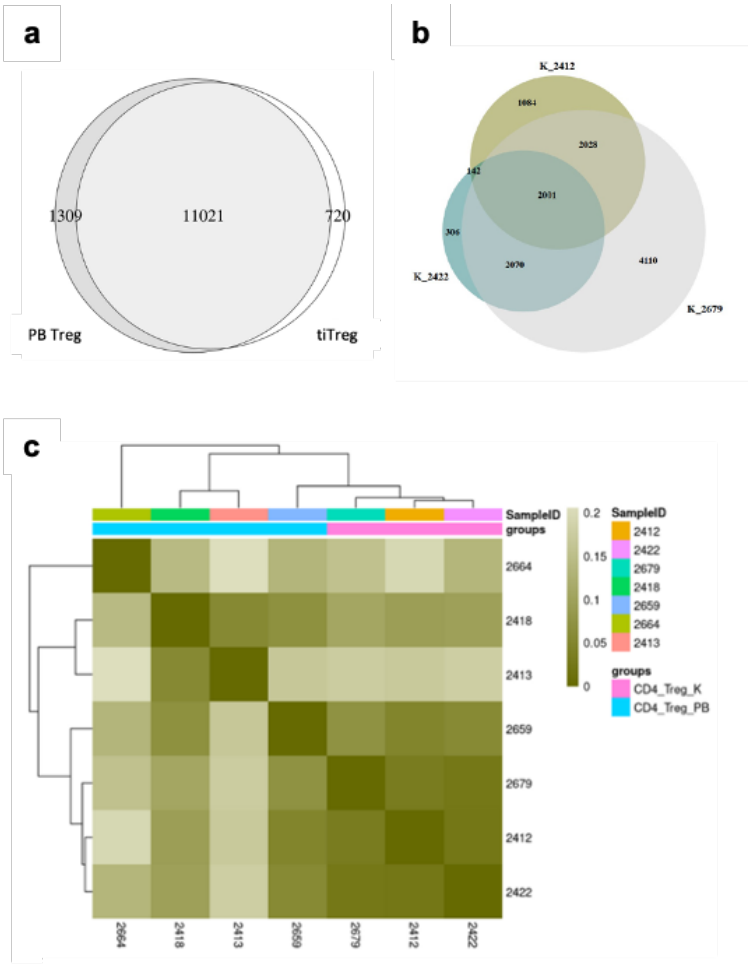


Figure 37 ChromHMM-defined enhancers in human Treg cells. Venn diagrams showing (**a**) the number of shared and specific enhancers between PB and tiTreg, and (**b**) enhancer concordance between different tiTreg replicates. (**c**) Unsupervised clustering analysis and Pearson correlation heatmap of H3K27ac ChIPmentation data for the 13,050 ChromHMM-defined active and genic enhancers distinguish tiTreg from PB Treg.

4.2.3 Definition of tiTreg enhancerome by mapping chromatin states to Treg accessibility landscape

We then compared the ChromHMM states with the chromatin accessibility data generated for PB, NAT and tiTreg cells (see section 4.2.1). **Figure 38 b** shows that the genomic regions associated with the “Active TSS” and “Active Enhancer” states are the most enriched in ATAC-seq signals, whilst the more inactive regions (e.g. “Repressed”, “Quiescence”) are depleted for open chromatin regions. Integration of the two different epigenetic layers showed that ~70% of active enhancers in Treg cells encompass chromatin regions that are accessible. Overall, these results suggest that the ChromHMM-defined chromatin states of Treg cells constitute a

robust atlas of genome-wide regulatory elements that enables the refinement and more accurate interpretation of ATAC-seq-defined open chromatin regions.

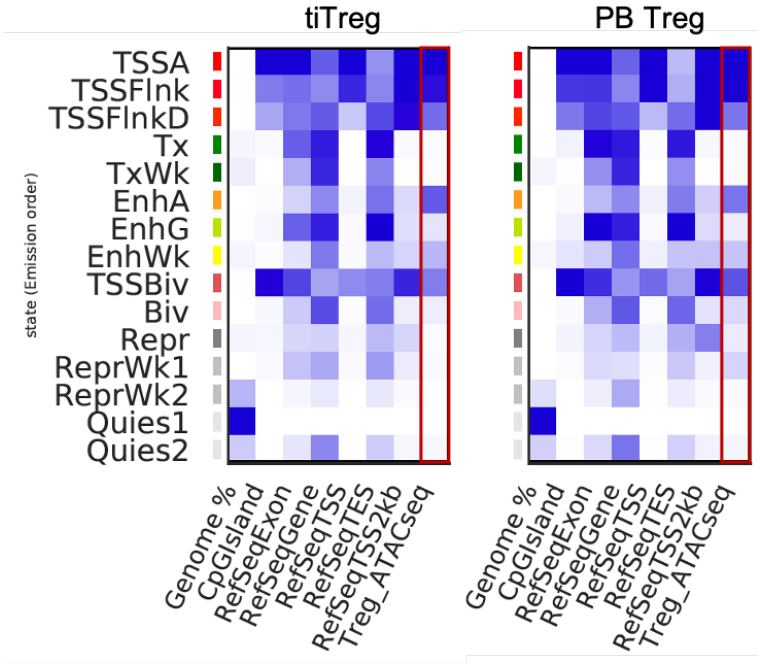


Figure 38 ChromHMM defined states concur with chromatic accessibility data. Heatmaps for a representative tiTreg (left) and PB Treg (right) replicate showing the enrichment of active enhancer and TSS states ("TSSA", "TSSFlnk", "TSSFlnkD", "EnhA", "TSSBiv") in ATAC-seq signals for Treg cells ('Treg_ATACseq' column in red). A darker blue color corresponds to a greater fold enrichment.

We next asked to what extent the chromatin regions which have been found accessible in Tregs by ATAC-seq are also annotated to the active enhancer chromatin states (**Figure 39**). Of the total 71,160 OCRs which represent the accessible peak universe of tissue- and circulating- Tregs, almost 22% (n =15,887) are marked as enhancer regions. Importantly, if we consider the subset of OCRs which display gained accessibility in tiTregs compared to NAT-Tregs, almost 32% of them are also annotated as active enhancers. Focusing on the OCRs that are linked to tiTreg signature genes, the active enhancer-annotated are about 35%.

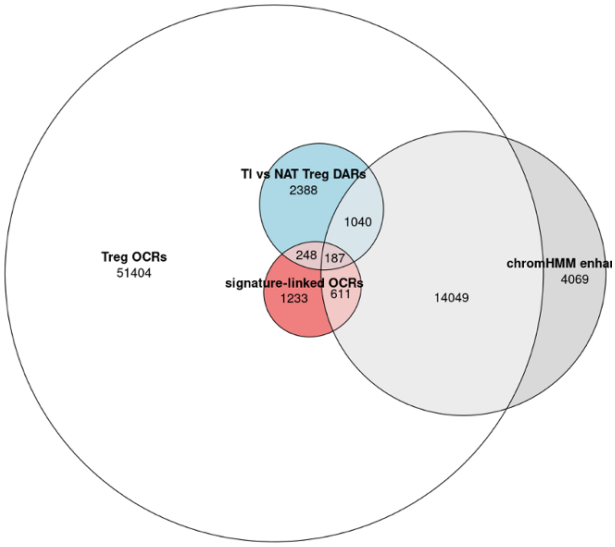


Figure 39 Open chromatin regions in Treg cells. Venn diagram displaying the overlap between all Treg OCRs, tiTreg signature genes-linked OCRs, gained differentially accessible regions (DARs) in tiTregs vs NAT-Tregs and regions annotated to active-enhancer chromatin states according to ChromHMM analysis.

Based on the integration of OCRs and ChromHMM-defined enhancers, we defined a tiTreg active enhancerome as the set of enhancer regions with gained accessibility in tiTregs versus NAT-Tregs (n = 1227). In light of our interest in addressing the epigenetic modulation of our previously identified tiTreg-specific gene signature(99), we also selected another set of regions representing all the signature gene-linked OCRs annotated as active enhancers (n = 798). We focused our investigations on both these sets of enhancers in parallel with the aim to dissect their transcription factor-mediated regulation.

4.2.3 Deciphering tiTreg transcription factor regulatory networks

The identification of cis-regulatory enhancer regions with enforced activity in tiTregs compared to normal cells, prompted us to investigate the trans-active regulatory layer that controls gene expression through DNA binding activity to these elements, in an effort to dissect the gene regulatory networks of transcription factors (TFs) and enhancers affecting the tiTreg phenotype. In this regard, the ATAC-seq assay can be further exploited thanks to its ability to putatively capture TF binding dynamics, since proteins bound to Tn5-digested chromatin regions leave behind footprints as a result of protecting the DNA molecule from being digested. A footprint can be

quantified as a drop in read density within otherwise accessible regions, which can be near or within a motif sequence for a given TF.

We thus performed a TF footprint analysis on our ATAC-seq data of Tregs in order to assess the identity of TFs that govern the tiTreg-specific gene regulatory network and are potentially relevant for establishing and maintaining their phenotype. In particular, we searched for TFs going through changes in their binding activity from circulating-/NAT- Tregs to tumor Tregs.

4.2.4.1 Global changes in transcription factor binding across Treg phenotypes

We started from a global approach in order to explore TF activities in all the accessible regulatory elements identified in tissue- and circulating-Tregs. We investigated binding dynamics for all the human TFs present in the Jaspar CORE database. In total, TOBIAS framework was used to calculate footprint scores for a list of 633 TFs across PB-, NAT- and tiTregs. We then clustered TFs into co-active groups throughout the different conditions and we found 7 distinct clusters of specific binding patterns (**Figure 40**). Importantly, two of these patterns peak in tiTregs: cluster 1 (n = 255) is the largest cluster that displays an increasing gradient of binding activity from PB-Tregs to tiTregs, with NAT-Tregs in the middle, whereas cluster 4 (n = 62) shows a strong activity restricted to tiTreg only. Therefore, for the factors within these clusters it is suggested a preferential-to-specific function in the establishment of tiTreg phenotype. Among the TFs showing a gradient of activity, we noticed factors such as BATF and other members of the AP-1 family that are known to contribute to Treg biology(149,177–179); FOXA1, which programs the immunosuppressive properties of FOXP3- regulatory T cells and has been implicated in the suppressing capacity of tiTregs deriving from lung cancer(180,181); FOXO3, a known activator of FOXP3; and RORα, whose expression is up-regulated in human skin Treg cells compared to circulating Tregs(182). Interestingly, we also found some TFs that are part of the tiTreg-specific transcriptional signature, such as HOXA1 and VDR.

In the cluster with highly-active TFs in tiTregs, we found, among others, FOXP3, IRF1, IRF4, NFKB1, NFKB2 and SOX4. Notably, also FOXP3 and SOX4 have been characterized as signature genes of the tiTreg phenotype.

On the opposite spectrum, five binding clusters are active in either PB- or NAT-Tregs. Amongst them, cluster 3 (*n* = 82) displays the highest activity in PB and NAT Tregs and appears strongly repressed in tiTregs, suggesting a potential downregulation of the TFs in this cluster that specifically occurs at the tiTreg state. Cluster 2 (n = 120), instead, shows high binding activity only in PB-Tregs, suggesting a prominent role for the corresponding TFs in blood-circulating Tregs that is diminished in tissue-infiltrating Tregs. Finally, TFs in cluster 6 (n = 38), which peaks in NAT-Tregs, display comparable activity in tiTregs but lack activity in PB-Tregs, indicating that they are likely involved in regulating the acquisition of the tissue-Treg phenotype.

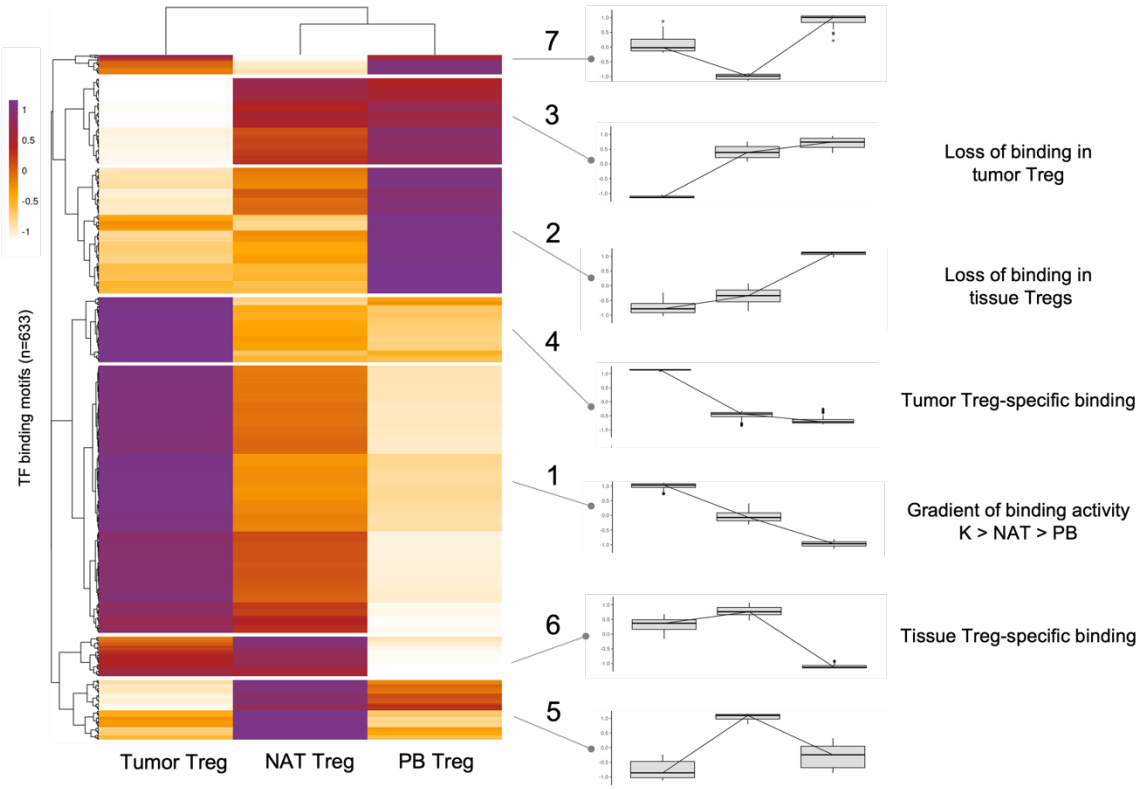


Figure 40 Heatmap of unsupervised clustering of TF binding activities across Treg cell populations (left panel). Each row represents one TF present in the Jaspar database, each column a Treg experimental condition. Clustering is based on mean footprint scores across Treg conditions, calculated for each TF as the mean of the footprint scores of all its binding sites identified along the consensus peakset of accessible chromatin regions called by MACS2 in TI-, NAT- and PB- Tregs. TOBIAS scores are *Z*-score transformed across rows. White color indicates low activity, purple color indicates high activity. Clusters of different TF binding activity were identified based on the dendrogram on the left of the heatmap. In order to visualize cluster trends, each cluster is associated with a mean trend line (left to right) and condition-specific boxplots (right panel).

We were also able to gain insights on the TF binding dynamics found with footprint scores looking at the aggregated bias-corrected footprint profiles of TFs along their consensus motifs. Here we report as a prominent example the footprint profile of BATF obtained from the base-wise aggregation of footprint scores for all the BATF motif occurrences across Treg conditions (**Figure 41**). A clear footprint is visible over the BATF motif within an increasingly accessible chromatin region from PB-Treg to tiTreg conditions.

Taken together, these data show that we were able to unveil massively parallel TF binding dynamics across Treg phenotypes and we identified TFs with tumor-specific binding activity or silencing.

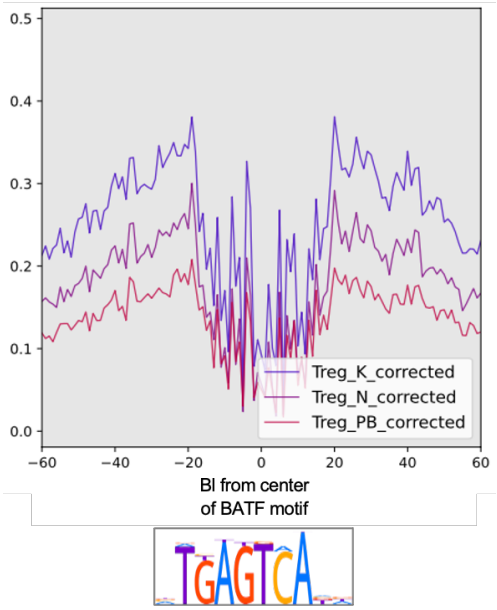


Figure 41 BATF footprinting. Footprint profile across Treg conditions (top) and motif consensus (bottom) for BATF, a tiTreg-specific TF. The aggregated footprinting plot is centered around the binding motif. Bias-corrected ATAC-seq footprint profiles for each Treg condition are labeled with different colors.

4.2.4.2 Assessing TF regulatory activities in tiTreg active enhancers

After setting out to classify TF binding dynamics within the tiTreg accessible global landscape and having found that there are several TFs with a preferential-to-specific binding activity in tiTregs, we focused our attention on the 1227 active enhancers in gained OCRs (section 4.2.3), with the aim of discovering which are the core TFs regulators of the tiTreg cell state that bind to their specific enhancer regions.

In particular, we wanted to investigate differentially bound TFs in our regions of interest between tumor-infiltrating and NAT-Tregs. We therefore utilized the TOBIAS pair-wise differential binding scores to assess the most significant changes in TF activities in both directions (**Figure 42**, **a**). In addition, we also integrated bulk mRNA expression data derived from our previous experiments, matching the TFs analyzed by TOBIAS with their RNA expression in tiTreg and PB-Tregs, respectively. In this way, we focused just on expressed TFs for the Treg cells. We observed, first of all, that TFs from the same families often display similar binding kinetics, like for instance the AP-1 and *forkhead* box (FOX) transcription factor families and the Ets family, which display higher binding activity in tiTregs and NAT-Tregs, respectively. This is not surprising since TF families often possess highly similar binding motifs. Among TFs displaying higher activity in tiTreg enhancers and being expressed in this condition, we identified some interesting factors that also have tiTreg-preferential binding properties in the global analysis of accessible regions, such as FOXP3, FOXO3, NFKB2 and BATF. We then asked what are the TFs that bind more frequently to tiTreg active enhancers and whether this data is concordant with their mRNA expression and binding score (**Figure 42**, **b**). Interestingly, we found that TFs with higher expression levels are more likely to have also a greater number of binding sites compared to the lowly-expressed TFs, that conversely tend to bind less frequently to the enhancers. BATF and the AP-1 family stand out as the top most frequently bound TFs and they have also a relatively higher binding score. Moreover, among the TFs with the highest occurrence we also identified IRF1, another TF with the tiTreg specific binding pattern in the global analysis.

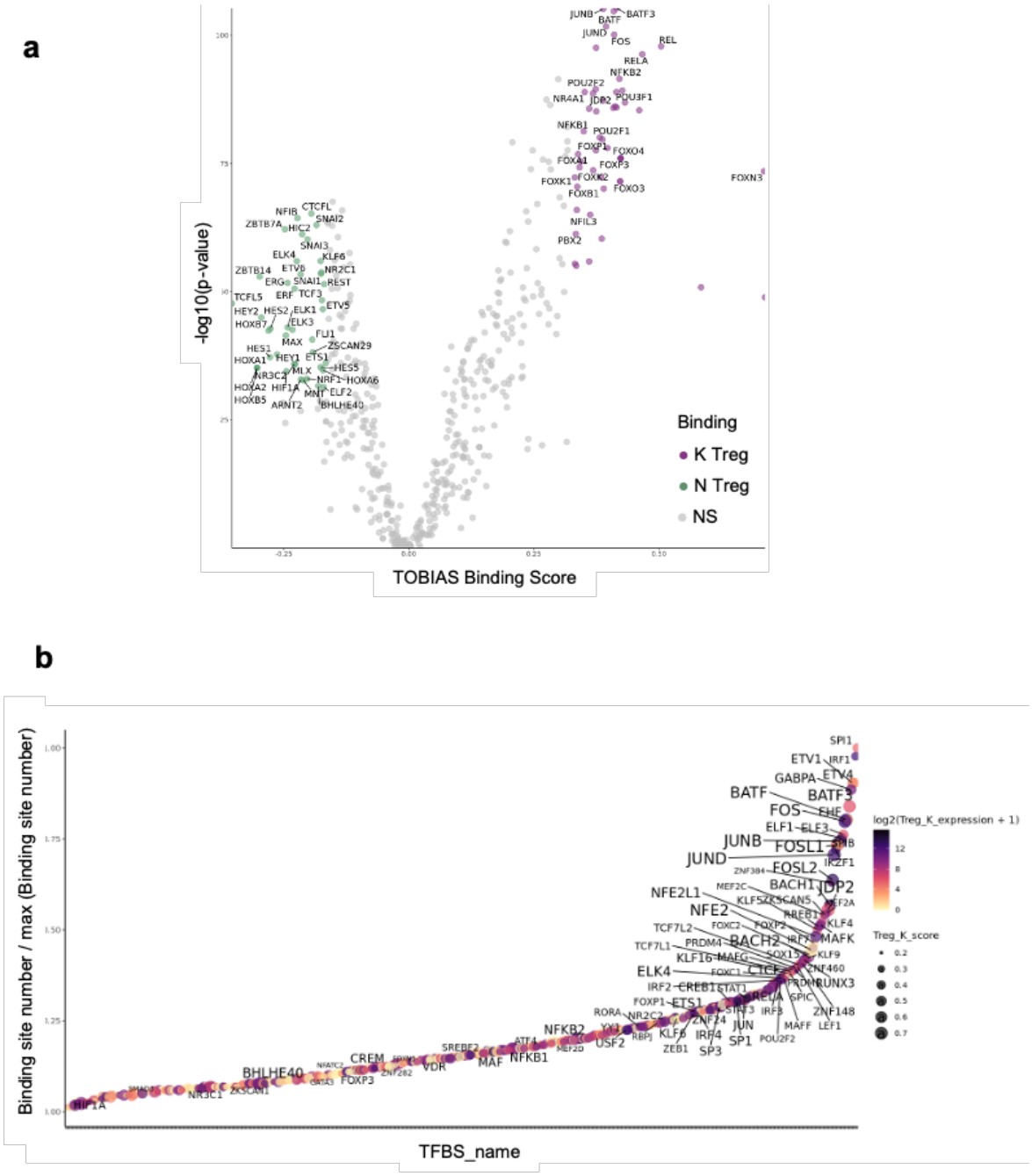


Figure 42 The binding dynamics of TF regulating tiTregs-specific active enhancers. (**a**) Pairwise comparison of TF activity between tumor-infiltrating (K) and NAT (N) Tregs. The volcano plots show the differential binding activity against the −log10(*p* value) (both provided by TOBIAS) of all investigated TF motifs; each dot represents a TF. TFs with top 10% positive and negative differential binding scores are colored based on the condition in which they display higher binding activity. For tiTreg, only colored TFs with top 50% mean of normalized expression in each condition are labeled, for NAT-Treg conditions, all colored dots are labelled and they are non-zero expressed TFs in either tumor or PB Tregs (expression data are not available for NAT-Tregs). tiTreg-specific TFs are labeled in purple, NAT-Treg-specific TFs in green. (**b**) Dot plot with TFs ranked in ascending order based on the number of binding sites in tiTreg condition on the x-axis and binding site number scaled on the maximum across TFs on the y-axis. Font and dot size are proportional to the tiTreg mean footprint score of the TF across all the bound enhancer regions. Dots are colored based on mRNA expression level in tiTregs, with yellow and purple indicating low and high expression, respectively.

We next wanted to address the same questions on TF regulation concerning the previously selected 798 active enhancers that are likely to modulate the expression of the tiTreg signature genes and we asked what are the differences and similarities in TF binding dynamics when focusing on the gene signature-linked enhancers alone compared to the genome-wide set of tiTreg active enhancers. When we looked at the differential binding activities between tiTregs and NAT-Tregs in these regions of interest (**Figure 43**, **a**), the AP-1 family members previously found were reconfirmed as the factors with the greatest change in their activity. Nevertheless, we also spotted some new actors: FOSL2, that is another AP-1 member that was not identified in the previous analysis, as well as BACH1 and BACH2 together with the small Maf proteins (MAFF, MAFK and MAFG), that are known interactors(183), RELB and IRF9 among the others. On the other hand, FOXO3 and FOXP3, as well as other FOX TF family members previously identified for the 1227 gained enhancers, do not display significant differences in the signature-annotated enhancers.

We again looked at the TFBS frequency in our regions of interest for tiTreg cells and this time we noticed that the top frequently bound TFs are not represented by the AP-1 family factors, conversely to what occurs when considering the total amount of tiTreg active enhancers. Instead, although BATF and its affine factors display a relatively high frequency of binding, we found at the first place SPI1, followed by IRF1, ETV1, ETV4 and GABPA. These results suggest that the signature genes could be regulated by enhancers that have some distinct affinities for TF binding compared to all the tiTreg enhancers, possibly relying on slightly different TF regulatory network, and indicate that a deeper investigation of these factors is required.

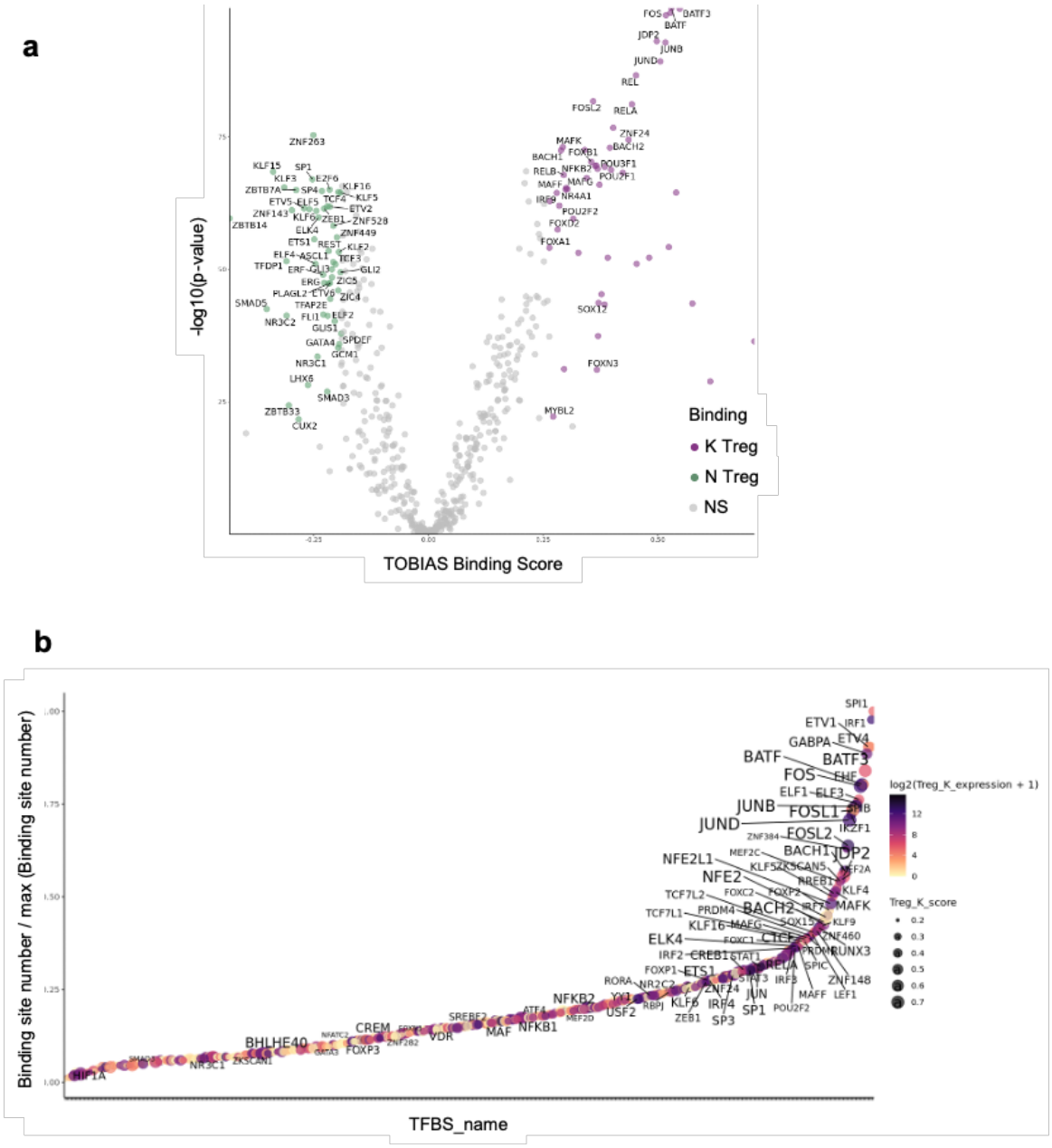


Figure 43 The binding dynamics of TF regulating tiTreg signature-linked active enhancers. (**a**) Pairwise comparison of TF activity between tumor-infiltrating (K) and NAT (N) Tregs. The volcano plots show the differential binding activity against the −log10(*p* value) (both provided by TOBIAS) of all investigated TF motifs; each dot represents a TF. TFs with top 10% positive and negative differential binding scores are colored based on the condition in which they display higher binding activity. For tiTreg, only colored TFs with top 50% mean of normalized expression in each condition are labeled, for NAT-Treg conditions, all colored dots are labelled and they are non-zero expressed TFs in either K or PB Tregs (expression data are not available for NAT-Tregs). tiTreg-specific TFs are labeled in purple, NAT-Treg-specific TFs in green. (**b**) Dot plot with TFs ranked in ascending order based on the number of binding sites in tiTreg condition on the x-axis and binding site number scaled on the maximum across TFs on the y-axis. Font and dot size are proportional to the tiTreg mean footprint score of the TFs across all the enhancer regions. Dots are colored based on mRNA expression levels in tiTreg condition, with yellow and purple indicating low and high expression, respectively.

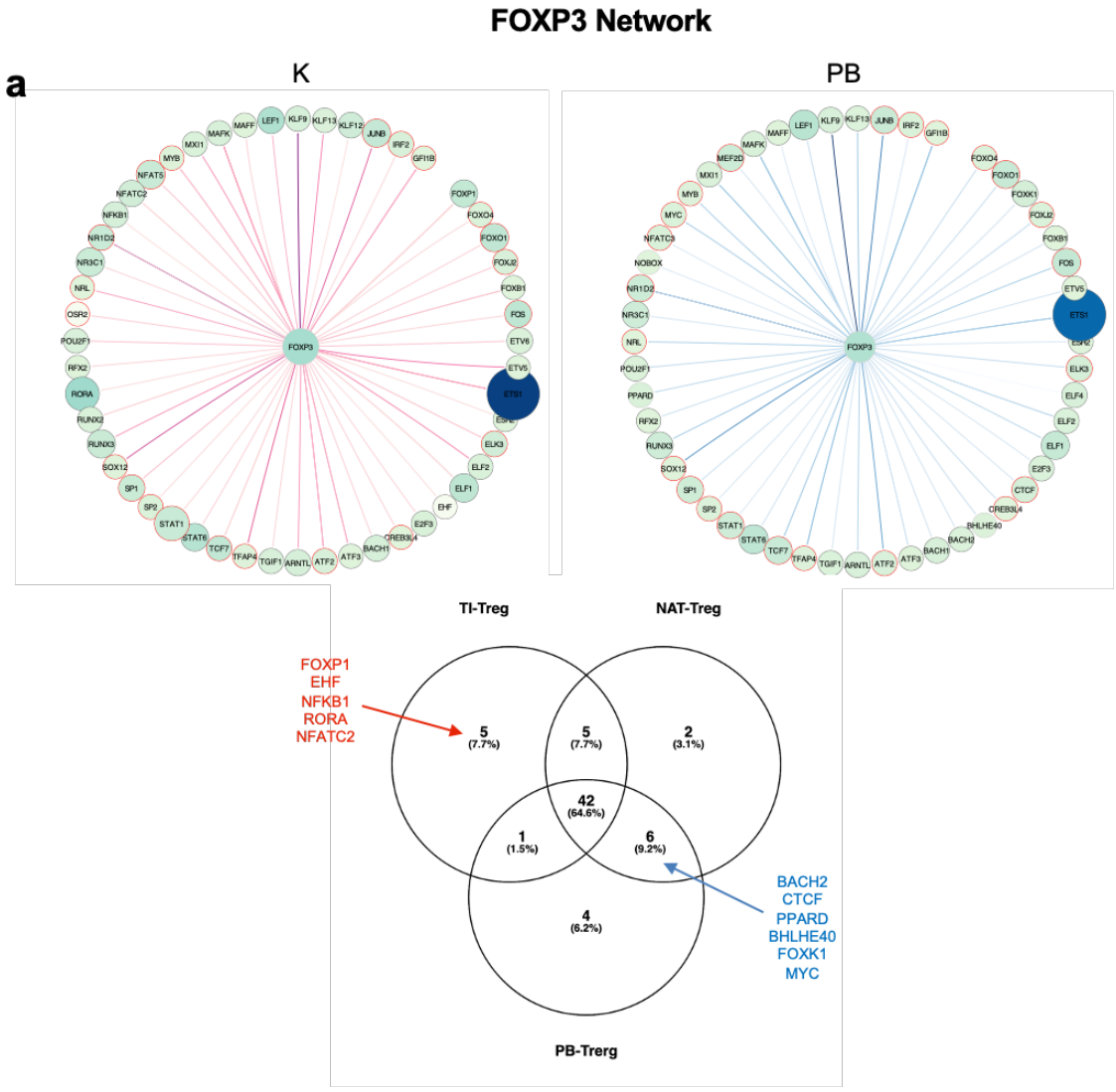
4.2.4.3 Assessing TF-TF regulatory networks that drive tiTreg cell state

The transcription factor FOXP3 is essential for Treg differentiation and function (52,54,55) whilst BATF has a crucial role in regulating differentiation and maintenance in nonlymphoid tissues (149,177,179). In our analyses, we have found that these factors also display a preferential binding activity in tiTreg compared to Tregs infiltrating the normal adjacent tissues. To interrogate the FOXP3- and BATF-mediated transcriptional circuity in tiTregs compared to PB- and NAT-Tregs, we constructed TF networks based on footprint data generated in TOBIAS. Using genome-wide ATAC-seq signals, we initially identified TF-encoding genes containing either FOXP3 or BATF motifs/footprint in OCRs (**Figure 44**, **a** and **b**). We then incorporated transcriptional data (RNA-seq) of Treg cells(99) to gain insight in the regulatory connections with target TFs.

By comparing the TF networks for FOXP3 (**Figure 44**, **a**), first of all we observed a similar number of nodes and edges between different Treg phenotypes, meaning a similar number of FOXP3-regulated TFs. However, some factors were predicted to be regulated by FOXP3 in tiTregs only (i.e., FOXP1, EHF, NFKB1, RORA, NFATC2), or only in PB- and NAT- Tregs (e.g., BACH2, CTCF, PPARD). Interestingly, the most highly expressed FOXP3 target was predicted to be ETS1, that in turn has been reported to have a role in the regulation of *Foxp3* expression(184).

Comparison of the TF networks for BATF (**Figure 44, b**), revealed instead almost twice the number of edges in tiTregs compared to PB-Tregs and an increased number of nodes (i.e., target TFs) in the latter condition (PB-Tregs: 71 nodes and 123 edges, NAT-Tregs: 90 nodes and 196 edges, tiTregs: 100 nodes and 234 edges). This means that BATF not only regulates a greater number of TFs in tiTregs than in the other conditions, but it also binds to a higher number of regulatory elements of the target TF genes or with more than one binding site at the same regulatory region. Moreover, condition-specific BATF targets were identified only in tiTregs (e.g., FOXA1, STAT5B, ETV6). On the other hand, three target TFs were identified only in PB- and NAT- Tregs but not in tiTregs (i.e., ZNF274, TP53 and CREB5). Interestingly, BATF was predicted to have itself as a target TF only in tiTregs and not in PB-Tregs, meaning that a BATF gene regulatory region is bound

by BATF factor in this condition. This result suggests the existence of a reinforcing mechanism for BATF-drived gene regulation through a positive feedback loop in tiTregs.



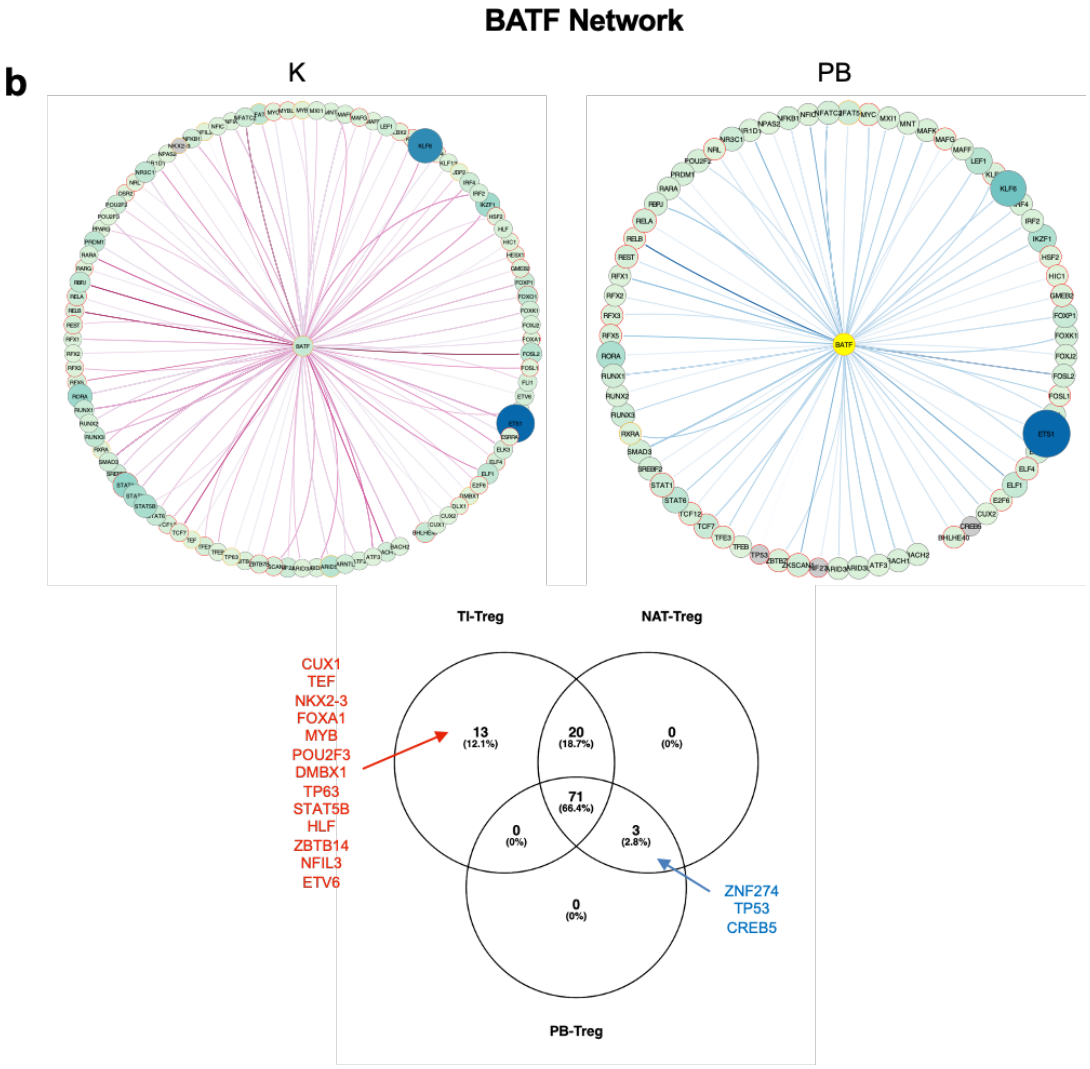


Figure 44. FOXP3- (a) and BATF- (b) centered networks of TFs were generated in TOBIAS. TF-TF networks (top) for tiTregs (left) and PB-Tregs (right): Border color of the target node indicates the presence of FOXP3/BATF (source node) binding footprint(s) in the promoter (red), distal elements (orange) or intronic elements (grey) of the indicated TF. Node color and size for each gene represents the mRNA-expression level of the indicated TF (light green and dark blue for low and high expression, respectively). Arrow width indicates the number of FOXP3/BATF binding footprints for the target TF. Arrow color represents the binding strength and is related to the footprint score of the FOXP3/BATF motif (purple shades for tiTreg, blue shades for PB-Tregs). Venn diagrams (bottom) display the overlap between FOXP3/BATF target TFs identified in the three different Treg conditions. Red-colored TFs are tiTreg-specific targets, blue-colored TFs are uniquely absent in tiTregs.

4.3 Functional assessment of tiTreg regulatory network modulation

Understanding the mechanisms that shape Treg cell identity and enhance their immunosuppressive capacity at tumor sites might provide new avenues for their reprogramming. To this aim we will employ single cell CRISPR pooled screening to assess the effects of tiTreg enhancer functional modulation. We will take advantage of a catalytically inactivate version of Cas9 (dCas9), fused to epigenetic repressor for enhancer modulation in CRISPR-pooled screening coupled to a single cell.

4.3.1 dCpf1-KRAB-MeCP2 protein for transcriptional repression of endogenous genes

The CRISPR/Cas9 system, has become a powerful tool for genome engineering, while the use of catalytically inactive versions of Cas9 and orthologous proteins, such as Cpf1, provides unique opportunities to functionally interrogate the epigenome. To this regard, George Church's lab have recently described a strategy that allows to permanently silence gene expression by means of targeted epigenetic editing. This strategy exploits engineered transcriptional repressors (ETRs), chimeric proteins containing a programmable RNA-guided DNA binding domain from the class II CRISPR/Cas system, and the combination of epigenetic repressor domains KRAB and MeCP2 that increases the silencing effect compared to single domain alone(171). We have modified this CRISPR-dCas9 version, replacing dCas9 with dCpf1(185) (**Figure 45**) that because of its intrinsic RNase activity, can process multiple individual gRNAs for the same or different enhancer regions, enhancing silencing efficiency.

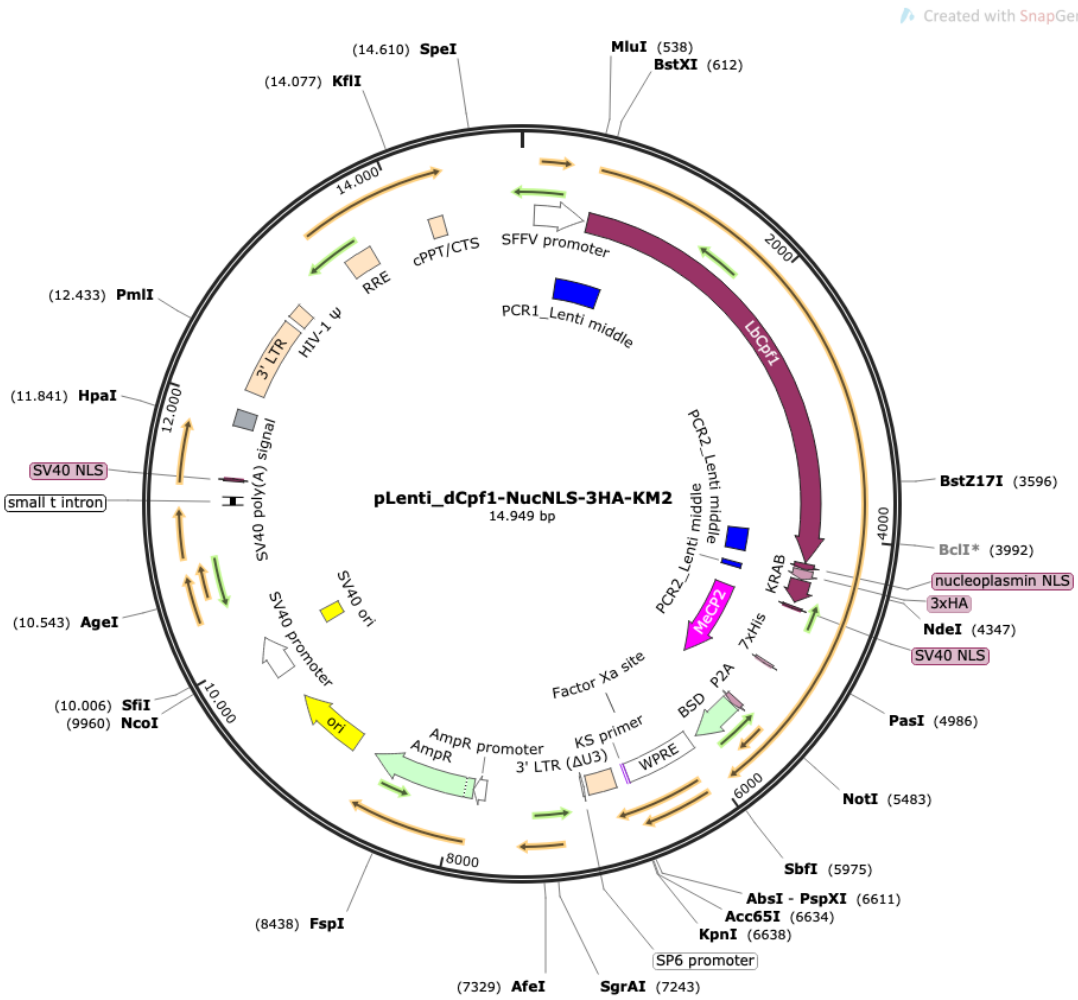


Figure 45 dCpg1-KRAB-MeCP2 plasmid map.

4.3.1.1 dCpf1-KRAB-MeCP2 protein expression and nuclear localization

The first step was to assess the dCpf1-KRAB-MeCP2 (dCpf1KM in short) chimeric protein expression at different time points after transfection into HEK cells. FACS and western blot analysis showed efficient expression of dCpf1KM protein in HEK cells at 24, 48 and 74 hours after transfection (**Figure 46**). Since the highest protein expression was observed 48 hours post-transfection, we performed the following experiments (nuclear localization and repression activity assay) at this timepoint.

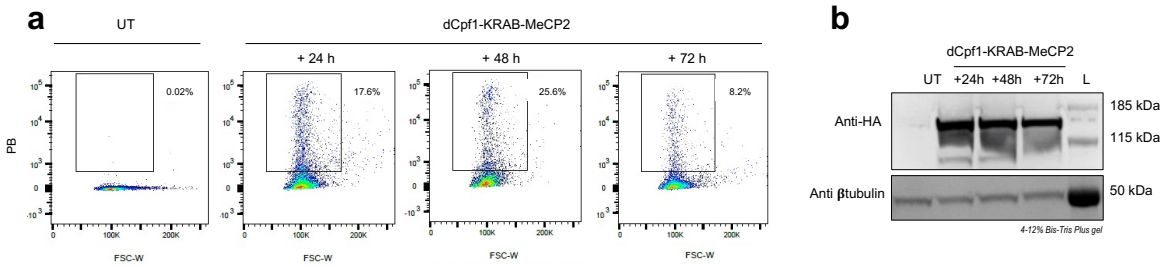


Figure 46 dCpf1-KRAB-MeCP2 protein expression. (**a**) FACS and (**b**) western blot analyses of dCpf1-KRAB-MeCP2 protein expression in HEK293T cells 24, 48 and 72 hours after transfection.

Next, to assess whether the repressor protein was correctly localized in the nucleus we performed a subcellular fractionation followed by western blot analysis. For comparison we also performed the experiment with the same construct with dCas9 fused to KRAB and MeCP2 (dCas9KM in short). We found that only 40% dCpf1KM was enriched in the nuclear fraction compared to about 90% dCas9KM (**Figure 47**).

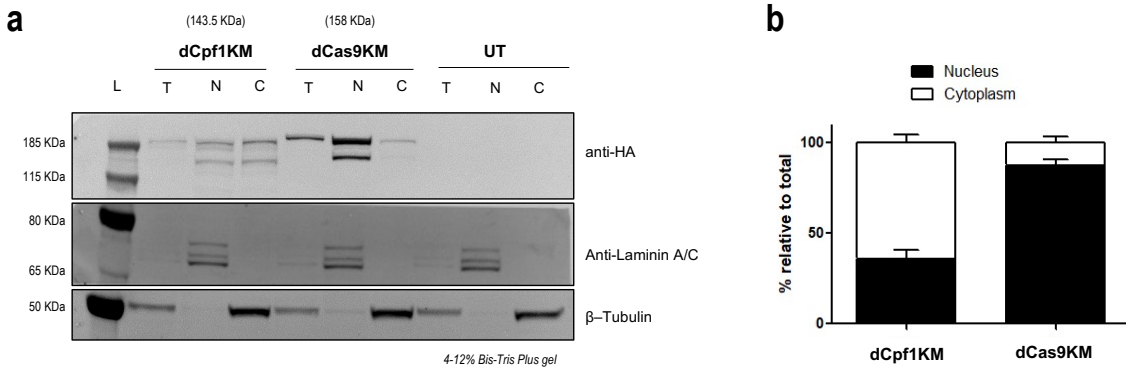


Figure 47 dCpf1KM and dCas9KM subcellular fractionation. (**a**) representative western blot of subcellular fractionation showing the localization of engineered repressors dCas9-KRAB-MeCP2 (dCas9KM) and dCpf1-KRAB-MeCP2 (dCpf1KM); UT: Untransfected, L: ladder; T: 10% of Total protein extract; N: Nuclear protein extract; C: Cytoplasmic protein extract. (**b**) Quantification of nuclear (black) and cytoplasmic (white) fraction of dCPf1KM and dCas9KM. The percentage of relative localization to total was obtained from n=2 experiments.

4.3.1.2 dCpf1-KRAB-MeCP2 promoter-directed transcriptional repression of endogenous genes

Given the relatively inefficient accumulation of dCpf1KM in the nucleus we wonder whether its repressor activity was impaired at least in HEK cells. To assess the promoter-directed transcriptional repression, HEK cells were co-transfected with the lentiviral constructs dCpf1KM and Cpf1CROPseq-Guide-mCherry (for guide

delivering). For this validation experiment, we designed a set of guideRNAs directed versus the promoters of checkpoint kinase CHEK and the tyrosine kinase MET genes that are expressed in HEK cells at the steady state. FACS analysis showed a good co-transfection efficiency compatible with the detection of repressive effects (**Figure 48**, **a**). Propitiously, the non-efficient nuclear localization of dCpf1KMn version had not impact on the repressor activity. Indeed, Real Time PCR showed that dCpf1KM was able to induce repression of endogenous gene in HEK293T cell line 48 hours after co-transfection. Nonetheless dCas9KM was more efficient in comparison to dCpf1KM we therefore concluded that at this stage a strategy based on the employment of dCas9 fusion proteins would be more effective (**Figure 48**, **b**).

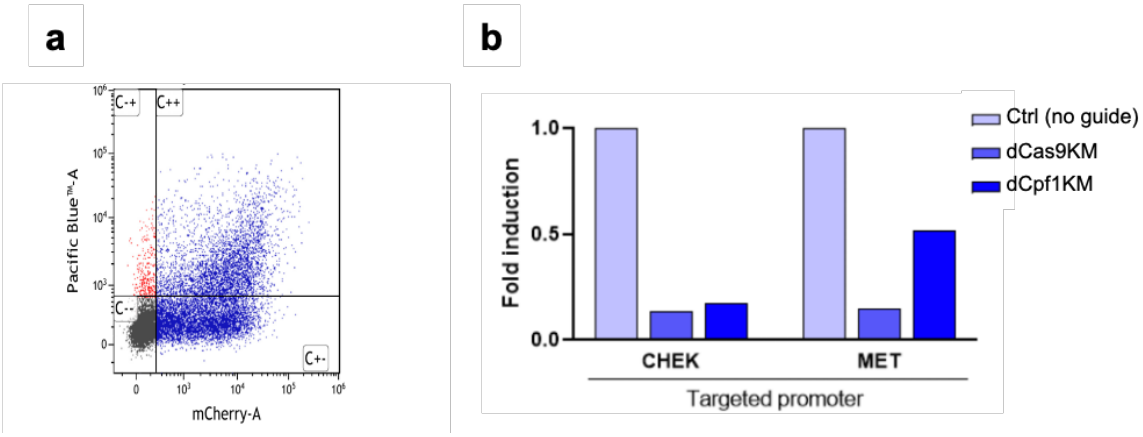


Figure 48 dCpf1-KRAB-MeCP2 transcriptional repression of endogenous genes. (**a**) FACS showing the co-transfection efficiency of dCpf1KM (PB-conjugated anti-HA antibody) and the reporter mCherry present in the CROPseq-Guide-mCherry plasmids used for guide expression (**b**) CHEK and MET expression after dCpf1-KRAB-MeCP2 silencing in HEK293T cells 48 hours after co-transfection. Fold induction values obtained from n=1 experiment.

4.3.2 CRISPRoff v2.1 for the transcriptional repression

Since epigenetic editing is a growing faster field novel tools are continuously under development. Recently an improved version of a catalytically inactive Cas9 combined to KRAB domain with the DNA methyltransferase Dnmt3 results in a greater repression compared to dCas9-KRAB or dCAs9-Dnmt3 alone(169).

We tested this construct (in short CRISPR-off v2) on HEK cells to test repression on the same targets MET and CHEK. With CRISPRoff v2 we obtained a strong and

reproducible repression of endogenous gene in HEK cell line when targeting their promoters (**Figure 49**).

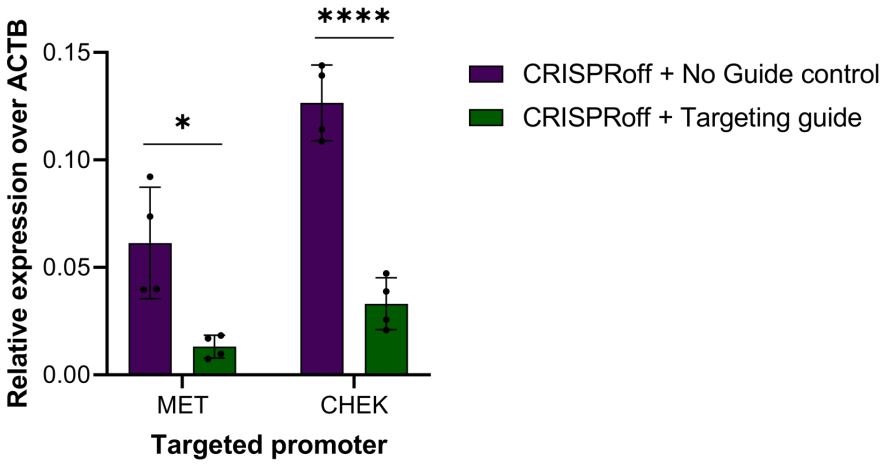


Figure 49 CRISPRoff transcriptional repression of endogenous genes CHEK and MET expression after CRISPRoff silencing in HEK293T cell 48 hours after co-transfection. Relative expression values obtained from n=4 experiments.

We are still investigating whether replacing dCas9 with dCpf1 from CRISPRoff will impair its nuclear localization and repressor ability. Moreover, we are addressing which are the best conditions (namely guideRNA combinations) to effectively repress enhancer regions that we identified through the integration of ATAC-seq and ChIPmentation datasets.

5 DISCUSSION

The immunosuppressive activity of Treg cells in tumors is a major obstacle to effective antitumour immunity. One of the major goals we have been pursuing in our lab is the characterization of the molecular blueprints of tumor infiltrating Treg (tiTreg) cells toward the identification of novel potential therapeutic targets for their specific modulation. Indeed, the present work stems from our previous findings showing that tiTregs adapt their transcriptome in response to the cues of the tumor microenvironment. This is in line with the observation that Treg cells display context-dependent altered stability, plasticity and tissue-specific heterogeneity, which affects their suppressive functions. In addition, Treg cells that are resident in healthy tissues are engaged also in tissue repair and other non-immune functions, that are mediated by the establishment of gene expression profiles different from Treg cells in lymphoid organs.

Treg-specific epigenetic changes that allow Treg cells to acquire these adaptive properties are largely uncharacterized. We thus decided to investigate the features of the epigenetic landscape that instruct the tiTreg phenotype in the two most frequent cancers in both genders: non-small-cell lung cancer (NSCLC) and colorectal cancer (CRC).

One of the strengths of our work is that epigenetic characterization by ChIPmentation and ATAC-seq is carried out on lymphocytes isolated from patients' tumor and peripheral blood without any further manipulation and expansion. However, sample preparation starting from human tissues is a difficult task and often involves a very limited amount of starting material. This implied a technological set-up to optimize the proper experimental conditions that took several months and finally ensured reproducible and good-quality results that could be analyzed and compared among patients and tissue of origin. Indeed, this technological set-up was successfully employed in deciphering transcription factor occupancy at enhancer regions in our recent studies (186).

Tumor-resident Treg cells share many of their epigenomic features with their counterparts in normal nonlymphoid tissues (187), thus differences between tiTreg and circulating Treg cells cannot be solely attributed to the tumor microenvironment.

By analyzing ATAC-seq datasets we identified genomic regions with a significant difference in chromatin accessibility between the tumor-associated and normal-tissue Treg cells. These regions likely represent regulatory elements accessible to transcription factors (TFs) underlying transcriptional programs related to the tumor microenvironment rather than to transcriptional changes that occur in response to tissue infiltration. Interestingly open chromatin regions are largely shared between Treg cells infiltrating NSCLC and CRC suggesting that drivers of Treg epigenetic plasticity are shared across tumor types. These findings support our previous observations on the conservation of tiTreg gene signature in different tumor types (99,100). They also provide a rationale for the discovery of tiTreg-specific epigenetic targets for cancer immunotherapy that can be effective across different cancer types.

By the implementation of ChromHMM (167) we managed to extract biologically informative chromatin states from the ChIPmentation datasets for distinct histone markers (H3K4me3, H3K4me1, H3K36me3, H3K27ac, H3K27me3). Gene expression changes are fundamentally controlled by modulating access of the transcriptional machinery to promoters and enhancers through chromatin structure modulation. Indeed, highly specific epigenetic modifications and enhancer activity are fundamental for the specification of Treg identity and the acquisition of their suppressive function (188). Enhancers in particular consolidate developmental and environmental cues to coordinate gene expression in a tissue-specific manner, thus it is reasonable to hypothesize that tissue microenvironment can have profound effects on the enhancer repertoire. Indeed, enhancer dysfunction in immune cells, mediated by the tumor microenvironment, is increasingly recognized as a major driver of tumor initiation and progression. Therefore, understanding the epigenetic reprogramming of tiTreg cells is a key step for their functional characterization and modulation. In line, therapeutic approaches that target epigenomic modifications show promising clinical efficacy in the management of solid tumors (189). To uncover the regulatory elements that act as key determinants in the definition of tiTreg cell state vs tissue infiltrating Treg cells, we focused on the combinatorial patterns of histone modifications that precisely identify active enhancers from other genomic elements within chromatin accessible regions.

TF footprinting analysis on ATAC-seq data shed light on the regulatory layer upstream enhancer activation mediated by the transcriptional network at the crossroad between environmental cues and enforcement of tumor Treg phenotype. Interestingly clustering based on genome-wide TF footprint scores identifies TF groups with diverse binding activity profiles across different Treg populations which are either lost or gained in tiTreg cells with respect to tissue infiltrating and peripheral blood Treg cells or to peripheral blood Treg cells alone.

TFs with increasing binding activity patterns from peripheral blood to tumors or higher binding activity in peripheral blood can add valuable information on the regulatory networks that specify Treg cell states. Interestingly, clustered TFs into co-active groups showed a large cluster displaying an increasing gradient of binding activity from peripheral blood Treg (PB-Treg) cells to tiTreg cells, with normal tissue infiltrating Treg (NAT-treg) cells in the middle. A second smaller cluster displayed a binding pattern restricted to tiTreg cells only suggesting that the factors within these clusters have a preferential-to-specific function in the establishment of tiTreg phenotype.

We further anlayzed the differences and similarities in TF binding dynamics of the putative regulatory regions for the tiTreg signature genes that we previously identified by bulk RNA transcriptomics (99). Our findings suggest that despite most of the signature genes are enriched but not unique for tiTreg, the underlying regulatory network for their putative enhancers is instead slightly different and likely reflects the specific influence of the tumor microenvironment. TF networks reconstructed from footprint data father support this hypothesis. As a way of example, we presented the TF networks for FOXP3, the master transcription factor of Treg cells, whose expression is required for the maintenance of Treg cell identity across tissues (52,54). Consistently, networks analysis showed a similar number of FOXP3-regulated TFs in tiTreg cells compared to PB- and NAT-Treg cells. Not all FOXP3 predicted targets, though, are shared across the three conditions. FOXP1 for instance, a key determinant for prolonged stability of the FOXP3 locus is a predicted FOXP3 target in tiTreg only (190). Factors that are predicted to be regulated in PB- and NAT- Treg cells are also of interest. BACH2, for instance, is rather intriguing as it promotes the quiescence and long-term maintenance of resident Treg cells by competing with AP-1 factors for DNA binding. This finding

supports the involvement of BACH2 in a FOXP3 driven mechanism that modulate Treg functional state following lineage commitment (191).

Reconstruction of TF network for BATF, that belongs to the gene signature, shows that BATF potentially regulates a high number of TF in tiTregs compared to PB- and NAT- Treg cells with targets (e.g. FOXA1 and STAT5B) unique for tiTreg cells.

Altogether, we were able to define the direct TF regulation of two key factors for Treg phenotype, FOXP3 and BATF (149,192), gaining some insights in the TF-TF networks that drive Treg functions from our experimental accessibility data. Importantly, our results suggest that, although these driver factors regulate Treg cell functional specification and homeostasis, with FOXP3 in particular being a lineage specification factor, they might acquire additional roles uniquely related to tiTreg plasticity in the tumor microenvironment. A deeper investigation of their networks and, especially, of the factors in turn regulated by the target TFs specifically acquired in tiTreg cells, is required to shed light on this mechanism potentially relevant for tiTreg phenotype.

Our findings suggest that the definition of tiTreg enhancerome and of enhancer transcriptional regulators is key to provide the rationale for the identification of novel potential therapeutic targets that lie along these regulatory axes reducing the risks deriving from systemic targeting of Tregs observed with immunomodulatory approaches based on monoclonal antibodies. Additional ATAC-seq and ChIPmentation datasets we generated on $CD4^+$ T conventional cells and $CD8^+$ T cells will help to further refine our efforts to regulatory hubs that are unique to Treg cells.

To this aim, we are implementing tools for the modulation of potential actionable targets derived from Treg enhancerome comprehensive analysis. A first approach is based on the epigenetic modulation of selected enhancers that represent a novel original strategy to reverse Treg mediated immunosuppression. At this stage we have developed the tools based on CRISPR technology to modify the epigenetic landscape around candidate enhancer sequences that take advantage of a catalytically inactive version of Cas9 (dCas9) fused to potent transcriptional repression KRAB and DNA methyltransferase Dnmt3 (CRISPRoff) (193) and we already validated it in human cell lines. The implications of this approach are even

broader if we consider recent findings suggesting that highly interconnected enhancer communities can delineate lineage choices and cell identity as coordinators of specific transcriptional programs (194). Capture HiC experiments targeted on Treg enhancerome can be envisaged to provide insights on the organization of enhancer communities in tiTreg cells and refine the choice of enhancer regions (or hubs) that might represent key vulnerabilities of Treg cells.

On the other hand the identification of the regulatory networks upstream of tiTreg enhancers reveals other potential targets (i.e. TFs) to severe the communication routes that translate the cues coming from the tumor microenvironment into the acquisition of tiTreg identity. To this end multiple approaches can be exploited for a selective inhibition that include the employment of shRNAs, siRNAs or small molecules.

To assess mechanistically the impact of selected actionable target modulation, it is important to establish models that closely represent the tumor microenvironment including stromal and immune cells and can therefore be exploited to address therapeutic strategies targeting this compartment. Air-liquid interface culture (ALI) derived organoids maintain primary tumor cells for several weeks with the associated tumor microenvironment (TME) and the most relevant immune populations of the TME are maintained for days (195). Indeed, the preservation of primary tumor epithelium with native autologous endogenous tumor infiltrating lymphocytes (TIL) provide the opportunity to model immune cell responses testing TIL activation expansion and tumor cell killing. In perspective ALIs therefore represent ideal model to assess how modulation of potential actionable targets identified with the multilayered analyses presented in the present work could impact TME-intrinsic immune cell responses. Remarkably, anti-PD-1/PD-L1 response rates in ALIs models have been shown to correlate well with patient responses, confirming that ALIs represent a good methodological approach for in vitro human modeling of immune-modulatory approaches. The nature of the target, which cannot be predicted in advance, will guide the choice of the modulation strategy (e.g. antibodies, small molecules, shRNAs). As a readout we envisage to assess TIL activation and exhaustion by high dimensional flowcytometry analysis as well as cell killing capacity upon target modulation. Cytokines and chemokines released in the culture supernatant can also be measured to assess the functional status of Treg

cells. Moreover, the impact of target modulation at the transcriptional level can be assessed by single-cell transcriptome analysis.

Alternatively, 3D in vitro tumoroid cultures can be exploited to assess the interplay between tumor cells and CRISPR/dCas9 perturbed Treg cells. Contrary to ALIs, tumoroids are deprived of the stromal and immune-cell component, but defined co-culturing conditions can be used to study the interplay between tumoroids and selected immune cell types to mimic a specific tissue immune-context. Co-cultures are modular systems and their complexity can be adapted to address specific biological questions, therefore they are well suited to model the interactions between immune and tumor cells under controlled experimental conditions.

Overall, several lines of research showed how at the steady state, Treg cells establish and maintain a strong regulatory phenotype, that is resistant to reprogramming. Howewer, Treg cells remian responsive to environmental cues and tailor their suppressive capacity in a context-dependent fashion. The comprehensive characterization of tiTreg epigenome and specifically of regulatory regions suggest that the inherent plasticity of Treg cells can be largely explained by the combination of the chromatin states and by the TF binding dynamics that can be achieved in response to external stimuli. Our findings point toward the existence of vulnerable nodes that sustain Treg cell activity and imply that appropriate modulatory strategies that target these nodes can ultimately alter the functional Treg pool specifically in the tumor milieu.

Converging evidences suggest that TME factors including hypoxia, tissue stiffness, nutrient availability, and cell–cell interactions among malignant and non-malignant cells such as infiltrating immune cells, stromal cells, and endothelial cells are shaping the epigenetic landscape and heavily affecting the identity and function of T cells in the tumor ecosystem. Indeed, definition of the epigenetic regulatory layer that drives T cell gene expression, will contribute to clarify the largely uncharacterized functional mechanisms underlying the effects of immune modulatory therapies such as PD1 and CTLA4 targeting antibodies. Indeed, these new treatments have revolutionized cancer therapies in very recent years, but their relatively low rate of success claims for a better understanding of their impact on both effector and regulatory cells. We expect that our studies on tumor infiltrating Treg “enhancerome” will provide insights on key regulatory nodes underlying the

acquisition of their hyperactivated phenotype and will guide the rational definition of improved clinical protocols for the administration of these powerful and promising immune-modulatory drugs.

ACKNOWLEDGEMENTS

I would like to thank my tutor, the Professor Massimiliano Pagani, for giving me the opportunity to work on challenging projects, in a laboratory that allows any researchers to give free rein to their imagination, and my supervisors Marco De Simone, who believed in me before myself, and Grazisa Rossetti, source of all knowledge; I am extremely grateful for their invaluable advice, continuous support, and patience during my PhD study.

I would like to thank all my colleagues, in particular Federica Gervasoni, Michaela Fakiola and Carolina Dossena for their huge effort for the bioinformatic analysis, and Giulia della Chiara, Claudia D'Oria and Jacopo Arrigoni for their kind help for this project.

I would also like to thank Maria Cristina Crosti and Monica Moro, at INGM FACS facility, and Maria Grazia Totaro, at IFOM FACS facility, for their technical support.

My gratitude extends to the Università degli Studi di Milano for the funding opportunity to undertake my PhD study at the INGM, during the first year, and IFOM Institute, during the last two years.

Ramona

REFERENCES

1. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. Int J cancer. 2021 Apr;
2. Hideyuki Takeshima and Toshikazu Ushijima. Accumulation of genetic and epigenetic alterations in normalcells and cancer risk. npj Precis Oncol. 2019;
3. de Martel C, Georges D, Bray F, Ferlay J, Clifford GM. Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. Lancet Glob Heal. 2020 Feb;8(2):e180–90.
4. Lewandowska AM, Rudzki M, Rudzki S, Lewandowski T, Laskowska B. Environmental risk factors for cancer - review paper. Ann Agric Environ Med. 2019 Mar;26(1):1–7.
5. Hassanpour SH, Dehghani M. Review of cancer from perspective of molecular. J Cancer Res Pract. 2017 Dec 1;4(4):127–9.
6. Liu J, Dang H, Wang XW. The significance of intertumor and intratumor heterogeneity in liver cancer. Exp Mol Med. 2018;50(1):e416–e416.
7. Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. Nature. 2013;501(7467):328–37.
8. Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ CK (eds). SEER Cancer Statistics Review. 1975-2018, Natl Cancer Institute Bethesda, MD, [Internet].
9. Zappa C, Mousa SA. Non-small cell lung cancer: current treatment and future advances. Transl lung cancer Res. 2016 Jun;5(3):288–300.
10. Pao W, Miller VA. Epidermal growth factor receptor mutations, small-molecule kinase inhibitors, and non-small-cell lung cancer: current knowledge and future directions. J Clin Oncol Off J Am Soc Clin Oncol. 2005 Apr;23(11):2556–68.
11. Turke AB, Zejnullahu K, Wu Y-L, Song Y, Dias-Santagata D, Lifshits E, et al. Preexistence and clonal selection of MET amplification in EGFR mutant

NSCLC. Cancer Cell. 2010 Jan;17(1):77–88.

12. Holmgaard RB, Zamarin D, Munn DH, Wolchok JD, Allison JP. Indoleamine 2,3-dioxygenase is a critical resistance mechanism in antitumor T cell immunotherapy targeting CTLA-4. J Exp Med. 2013 Jul;210(7):1389–402.

13. Gettinger S, Choi J, Hastings K, Truini A, Datar I, Sowell R, et al. Impaired HLA Class I Antigen Processing and Presentation as a Mechanism of Acquired Resistance to Immune Checkpoint Inhibitors in Lung Cancer. Cancer Discov. 2017 Dec;7(12):1420–35.

14. Stankovic B, Bjørhovde HAK, Skarshaug R, Aamodt H, Frafjord A, Müller E, et al. Immune Cell Composition in Human Non-small Cell Lung Cancer. Front Immunol. 2018;9:3101.

15. Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. Nat Rev Clin Oncol. 2017 Dec;14(12):717–34.

16. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. Nat Med. 2015 Nov;21(11):1350–6.

17. Becht E, de Reyniès A, Giraldo NA, Pilati C, Buttard B, Lacroix L, et al. Immune and Stromal Classification of Colorectal Cancer Is Associated with Molecular Subtypes and Relevant for Precision Immunotherapy. Clin cancer Res an Off J Am Assoc Cancer Res. 2016 Aug;22(16):4057–66.

18. Ehrlich P. Ueber den jetzigen stand der Karzinomforschung. Ned Tijdschr Geneeskd. 1909;5:273–290.

19. Old LJ, Boyse EA. Immunology of Experimental Tumors. Annu Rev Med [Internet]. 1964 Feb 1;15(1):167–86. Available from: https://doi.org/10.1146/annurev.me.15.020164.001123

20. Burnet M. Cancer: A Biological Approach. Br Med J. 1957;1(5022):779–847.

21. Thomas L. Discussion of cellular and humoral aspects of the hypersensitive states. HS Lawrence. 1959;(529–532).

22. Burnet FM. The concept of immunological surveillance. Prog Exp Tumor Res.

1970;13:1–27.

23. Shankaran V, Ikeda H, Bruce AT, White JM, Swanson PE, Old LJ, et al. IFNγ and lymphocytes prevent primary tumour development and shape tumour immunogenicity. Nature. 2001 Apr 26;410:1107.

24. Deaglio S, Dwyer KM, Gao W, Friedman D, Usheva A, Erat A, et al. Adenosine generation catalyzed by CD39 and CD73 expressed on regulatory T cells mediates immune suppression. J Exp Med. 2007 Jun 11;204(6):1257–65.

25. Dunn GP, Koebel CM, Schreiber RD. Interferons, immunity and cancer immunoediting. Nat Rev Immunol. 2006 Nov 1;6:836.

26. D Schreiber R, J Old L, J Smyth M. Cancer Immunoediting: Integrating Immunity's Roles in Cancer Supression and Promotion. Vol. 331, Science (New York, N.Y.). 2011. 1565–1570 p.

27. Dunn GP, Old LJ, Schreiber RD. The Immunobiology of Cancer Immunosurveillance and Immunoediting. Immunity. 2004 Jan 3;21(2):137–48.

28. Abbas AK, Lichtman AH, Pillai S. Cellular and molecular immunology. 8th editio. Philadelphia :Saunders, editor. 2014.

29. Matzinger P. Tolerance, Danger, and the Extended Family. Annu Rev Immunol. 1994 Apr 1;12(1):991–1045.

30. Aguirre-Ghiso JA. Models, mechanisms and clinical evidence for cancer dormancy. Nat Rev Cancer. 2007 Nov;7(11):834–46.

31. Farrar JD, Katz KH, Windsor J, Thrush G, Scheuermann RH, Uhr JW, et al. Cancer Dormancy. VII. A Regulatory Role for CD8+ T Cells and IFN-γ in Establishing and Maintaining the Tumor-Dormant State. J Immunol. 1999 Mar 1;162(5):2842 LP – 2849.

32. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell. 2011 Jan 3;144(5):646–74.

33. Radoja S, Rao TD, Hillman D, Frey AB. Mice Bearing Late-Stage Tumors Have Normal Functional Systemic T Cell Responses In Vitro and In Vivo. J Immunol. 2000 Mar 1;164(5):2619 LP – 2628.

34. Vesely MD, Kershaw MH, Schreiber RD, Smyth MJ. Natural Innate and

Adaptive Immunity to Cancer. Annu Rev Immunol. 2011 Mar 22;29(1):235–71.

35. Terabe M, Berzofsky JA. Immunoregulatory T cells in tumor immunity. Curr Opin Immunol. 2004;16(2):157–62.

36. Gershon RK, Cohen P, Hencin R, Liebhaber SA. Suppressor T Cells. J Immunol. 1972 Mar 1;108(3):586 LP – 590.

37. Sakaguchi S, Takahashi T, Nishizuka Y. Study on cellular events in post-thymectomy autoimmune oophoritis in mice. II. Requirement of Lyt-1 cells in normal female mice for the prevention of oophoritis. J Exp Med. 1982 Dec 1;156(6):1577–86.

38. Sakaguchi S, Sakaguchi N, Asano M, Itoh M, Toda M. Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. J Immunol. 1995 Aug 1;155(3):1151 LP – 1164.

39. Read S, Malmström V, Powrie F. Cytotoxic T Lymphocyte–Associated Antigen 4 Plays an Essential Role in the Function of CD25+CD4+ Regulatory Cells That Control Intestinal Inflammation. J Exp Med. 2000 Jul 17;192(2):295 LP – 302.

40. Takahashi T, Tagami T, Yamazaki S, Uede T, Shimizu J, Sakaguchi N, et al. Immunologic Self-Tolerance Maintained by CD25+CD4+ Regulatory T Cells Constitutively Expressing Cytotoxic T Lymphocyte–Associated Antigen 4. J Exp Med. 2000 Jul 17;192(2):303 LP – 310.

41. McHugh RS, Whitters MJ, Piccirillo CA, Young DA, Shevach EM, Collins M, et al. CD4+CD25+ Immunoregulatory T Cells: Gene Expression Analysis Reveals a Functional Role for the Glucocorticoid-Induced TNF Receptor. Immunity. 2002 Dec 29;16(2):311–23.

42. Huang C-T, Workman CJ, Flies D, Pan X, Marson AL, Zhou G, et al. Role of LAG-3 in Regulatory T Cells. Immunity. 2004 Dec 29;21(4):503–13.

43. Seddiki N, Santner-Nanan B, Martinson J, Zaunders J, Sasson S, Landay A, et al. Expression of interleukin (IL)-2 and IL-7 receptors discriminates between

human regulatory and activated T cells. J Exp Med. 2006 Jul 10;203(7):1693–700.

44. Piccirillo CA, Thornton AM. Cornerstone of peripheral tolerance: naturally occurring CD4+CD25+ regulatory T cells. Trends Immunol. 2004;25(7):374–80.

45. Bruder D, Probst-Kepper M, Westendorf AM, Geffers R, Beissert S, Loser K, et al. Frontline: Neuropilin-1: a surface marker of regulatory T cells. Eur J Immunol. 2004;34(3):623–30.

46. Plitas G, Rudensky AY. Regulatory T Cells: Differentiation and Function. Cancer Immunol Res. 2016 Sep 2;4(9):721–5.

47. Vignali DAA, Collison LW, Workman CJ. How regulatory T cells work. Nat Rev Immunol. 2008 Jul;8(7):523–32.

48. Wan YY, Flavell RA. TGF-β and Regulatory T Cell in Immunity and Autoimmunity. J Clin Immunol. 2008 Nov 16;28(6):647–59.

49. Zorn E, Nelson EA, Mohseni M, Porcheray F, Kim H, Litsa D, et al. IL-2 regulates FOXP3 expression in human CD4(+)CD25(+) regulatory T cells through a STAT-dependent mechanism and induces the expansion of these cells in vivo. Blood. 2006 Sep 1;108(5):1571–9.

50. Sakaguchi S, Yamaguchi T, Nomura T, Ono M. Regulatory T Cells and Immune Tolerance. Cell. 2008 Dec 28;133(5):775–87.

51. Morgan ME, van Bilsen JHM, Bakker AM, Heemskerk B, Schilham MW, Hartgers FC, et al. Expression of FOXP3 mRNA is not confined to CD4+CD25+ T regulatory cells in humans. Hum Immunol. 2005;66(1):13–20.

52. Hori S, Nomura T, Sakaguchi S. Control of Regulatory T Cell Development by the Transcription Factor Foxp3; Science. 2003 Feb 14;299(5609):1057–1061.

53. Khattri R, Cox T, Yasayko S-A, Ramsdell F. An essential role for Scurfin in CD4+CD25+ T regulatory cells. Nat Immunol. 2003 Mar 3;4:337.

54. Fontenot JD, Gavin MA, Rudensky AY. Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. Nat Immunol. 2003 Mar 3;4:330.

55. Bennett CL, Christie J, Ramsdell F, Brunkow ME, Ferguson PJ, Whitesell L,

et al. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. Nat Genet. 2001 Jan 1;27:20.

56. Josefowicz SZ, Lu L-F, Rudensky AY. Regulatory T Cells: Mechanisms of Differentiation and Function. Annu Rev Immunol. 2012 Mar 26;30(1):531–64.

57. Hsieh C-S, Zheng Y, Liang Y, Fontenot JD, Rudensky AY. An intersection between the self-reactive regulatory and nonregulatory T cell receptor repertoires. Nat Immunol. 2006 Mar 12;7:401.

58. Chen W, Jin W, Hardegen N, Lei K, Li L, Marinos N, et al. Conversion of Peripheral CD4(+)CD25(−) Naive T Cells to CD4(+)CD25(+) Regulatory T Cells by TGF-β Induction of Transcription Factor Foxp3. J Exp Med. 2003 Dec 15;198(12):1875–86.

59. Arpaia N, Campbell C, Fan X, Dikiy S, van der Veeken J, deRoos P, et al. Metabolites produced by commensal bacteria promote peripheral regulatory T cell generation. Nature. 2013 Dec 19;504(7480):451–5.

60. Samstein RM, Josefowicz SZ, Arvey A, Treuting PM, Rudensky AY. Extrathymic Generation of Regulatory T Cells in Placental Mammals Mitigates Maternal-Fetal Conflict. Cell. 2012 Dec 28;150(1):29–38.

61. Josefowicz SZ, Niec RE, Kim HY, Treuting P, Chinen T, Zheng Y, et al. Extrathymically generated regulatory T cells control mucosal Th2 inflammation. Nature. 2012 Feb 8;482(7385):395–9.

62. Sakaguchi S, Sakaguchi N, Shimizu J, Yamazaki S, Sakihama T, Itoh M, et al. Immunologic tolerance maintained by CD25+ CD4+ regulatory T cells: their common role in controlling autoimmunity, tumor immunity, and transplantation tolerance. Immunol Rev. 2001;182(1):18–32.

63. Coombes JL, Robinson NJ, Maloy KJ, Uhlig HH, Powrie F. Regulatory T cells and intestinal homeostasis. Immunol Rev. 2005;204(1):184–94.

64. Kretschmer K, Apostolou I, Jaeckel E, Khazaie K, Von Boehmer H. Making regulatory T cells with defined antigen specificity: role in autoimmunity and cancer. Immunol Rev. 2006;212(1):163–9.

65. Cipolletta D, Feuerer M, Li A, Kamei N, Lee J, Shoelson SE, et al. PPARγ is

a major driver of the accumulation and phenotype of adipose-tissue T(reg) cells. Nature. 2012 Jun 28;486(7404):549–53.

66. Feuerer M, Herrero L, Cipolletta D, Naaz A, Wong J, Nayer A, et al. Lean, but not obese, fat is enriched for a unique population of regulatory T cells that affect metabolic parameters. Nat Med. 2009 Jul 26;15:930.

67. Arpaia N, Green JA, Moltedo B, Arvey A, Hemmers S, Yuan S, et al. A Distinct Function of Regulatory T Cells in Tissue Protection. Cell. 2015 Aug 27;162(5):1078–89.

68. Burzyn D, Kuswanto W, Kolodin D, Shadrach JL, Cerletti M, Jang Y, et al. A Special Population of Regulatory T Cells Potentiates Muscle Repair. Cell [Internet]. 2013 Dec 28;155(6):1282–95. Available from: http://dx.doi.org/10.1016/j.cell.2013.10.054

69. Corthay A. How do Regulatory T Cells Work? Scand J Immunol. 2009 Oct 25;70(4):326–36.

70. Asseman C, Mauze S, Leach MW, Coffman RL, Powrie F. An Essential Role for Interleukin 10 in the Function of Regulatory T Cells That Inhibit Intestinal Inflammation. J Exp Med. 1999 Oct 4;190(7):995–1004.

71. Collison LW, Workman CJ, Kuo TT, Boyd K, Wang Y, Vignali KM, et al. The inhibitory cytokine IL-35 contributes to regulatory T-cell function. Nature. 2007 Nov 22;450:566.

72. Nakamura K, Kitani A, Strober W. Cell Contact–Dependent Immunosuppression by Cd4(+)Cd25(+)Regulatory T Cells Is Mediated by Cell Surface–Bound Transforming Growth Factor β. J Exp Med. 2001 Sep 3;194(5):629–44.

73. Taylor A, Verhagen J, Blaser K, Akdis M, Akdis CA. Mechanisms of immune suppression by interleukin-10 and transforming growth factor-β: the role of T regulatory cells. Immunology. 2006;117(4):433–42.

74. Arce-Sillas A, Álvarez-Luquín DD, Tamaya-Domínguez B, Gomez-Fuentes S, Trejo-García A, Melo-Salas M, et al. Regulatory T Cells: Molecular Actions on Effector Cells in Immune Regulation. J Immunol Res. 2016 May 19;2016:1720827.

75. Olson BM, Sullivan JA, Burlingham WJ. Interleukin 35: A Key Mediator of Suppression and the Propagation of Infectious Tolerance. Front Immunol. 2013 Oct 18;4:315.

76. Grossman WJ, Verbsky JW, Tollefsen BL, Kemper C, Atkinson JP, Ley TJ. Differential expression of granzymes A and B in human cytotoxic lymphocyte subsets and T regulatory cells. Blood. 2004 Nov 1;104(9):2840 LP – 2848.

77. Gondek DC, Lu L-F, Quezada SA, Sakaguchi S, Noelle RJ. Cutting Edge: Contact-Mediated Suppression by CD4+CD25+ Regulatory Cells Involves a Granzyme B-Dependent, Perforin-Independent Mechanism. J Immunol. 2005 Feb 15;174(4):1783 LP – 1786.

78. Pandiyan P, Zheng L, Ishihara S, Reed J, Lenardo MJ. CD4+CD25+Foxp3+ regulatory T cells induce cytokine deprivation–mediated apoptosis of effector CD4+ T cells. Nat Immunol. 2007 Nov 4;8:1353.

79. Borsellino G, Kleinewietfeld M, Di Mitri D, Sternjak A, Diamantini A, Giometto R, et al. Expression of ectonucleotidase CD39 by Foxp3+ Treg cells: hydrolysis of extracellular ATP and immune suppression. Blood. 2007 Aug 15;110(4):1225 LP – 1232.

80. Kobie JJ, Shah PR, Yang L, Rebhahn JA, Fowell DJ, Mosmann TR. T Regulatory and Primed Uncommitted CD4 T Cells Express CD73, Which Suppresses Effector CD4 T Cells by Converting 5′-Adenosine Monophosphate to Adenosine. J Immunol. 2006 Nov 15;177(10):6780 LP – 6786.

81. Bopp T, Becker C, Klein M, Klein-Heßling S, Palmetshofer A, Serfling E, et al. Cyclic adenosine monophosphate is a key component of regulatory T cell–mediated suppression. J Exp Med. 2007 Jun 11;204(6):1303 LP – 1310.

82. Fallarino F, Grohmann U, Hwang KW, Orabona C, Vacca C, Bianchi R, et al. Modulation of tryptophan catabolism by regulatory T cells. Nat Immunol. 2003 Oct 26;4:1206.

83. Liang B, Workman C, Lee J, Chew C, Dale BM, Colonna L, et al. Regulatory T Cells Inhibit Dendritic Cells by Lymphocyte Activation Gene-3 Engagement of MHC Class II. J Immunol. 2008 May 1;180(9):5916 LP – 5926.

84. Oleinika K, Nibbs RJ, Graham GJ, Fraser AR. Suppression, subversion and escape: the role of regulatory T cells in cancer progression. Clin Exp Immunol. 2013 Jan 3;171(1):36–45.

85. Clemente CG, Mihm MC, Bufalino R, Zurrida S, Collini P, Cascinelli N. Prognostic value of tumor infiltrating lymphocytes in the vertical growth phase of primary cutaneous melanoma. Cancer. 1996;77(7):1303–10.

86. Sato E, Olson SH, Ahn J, Bundy B, Nishikawa H, Qian F, et al. Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. Proc Natl Acad Sci United States Am. 2005 Dec 20;102(51):18538–43.

87. Bates GJ, Fox SB, Han C, Leek RD, Garcia JF, Harris AL, et al. Quantification of Regulatory T Cells Enables the Identification of High-Risk Breast Cancer Patients and Those at Risk of Late Relapse. J Clin Oncol. 2006 Dec 1;24(34):5373–80.

88. Petersen RP, Campa MJ, Sperlazza J, Conlon D, Joshi M-B, Harpole DH, et al. Tumor infiltrating Foxp3+ regulatory T-cells are associated with recurrence in pathologic stage I NSCLC patients. Cancer. 2006;107(12):2866–72.

89. Salama P, Phillips M, Grieu F, Morris M, Zeps N, Joseph D, et al. Tumor-Infiltrating FOXP3+ T Regulatory Cells Show Strong Prognostic Significance in Colorectal Cancer. J Clin Oncol. 2009 Jan 10;27(2):186–92.

90. Sinicrope FA, Rego RL, Ansell SM, Knutson KL, Foster NR, Sargent DJ. A low intraepithelial effector (CD3+):regulatory (FoxP3+) T-cell ratio predicts adverse outcome of human colon carcinoma. Gastroenterology. 2009 Oct 3;137(4):1270–9.

91. Suzuki H, Chikazawa N, Tasaka T, Wada J, Yamasaki A, Kitaura Y, et al. Intratumoral CD8+ T/FOXP3+ cell ratio is a predictive marker for survival in patients with colorectal cancer. Cancer Immunol Immunother. 2010;59(5):653–61.

92. Saito T, Nishikawa H, Wada H, Nagano Y, Sugiyama D, Atarashi K, et al. Two FOXP3+CD4+ T cell subpopulations distinctly control the prognosis of colorectal cancers. Nat Med. 2016 Apr 25;22:679.

93. Quezada SA, Peggs KS, Simpson TR, Allison JP. Shifting the equilibrium in cancer immunoediting: from tumor tolerance to eradication. Immunol Rev. 2011 May;241(1):104–18.

94. Curiel TJ, Coukos G, Zou L, Alvarez X, Cheng P, Mottram P, et al. Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. Nat Med. 2004 Aug 22;10:942.

95. Ghiringhelli F, Puig PE, Roux S, Parcellier A, Schmitt E, Solary E, et al. Tumor cells convert immature myeloid dendritic cells into TGF-β–secreting cells inducing CD4(+)CD25(+) regulatory T cell proliferation. J Exp Med. 2005 Oct 3;202(7):919–29.

96. Zhou G, Levitsky HI. Natural Regulatory T Cells and De Novo-Induced Regulatory T Cells Contribute Independently to Tumor-Specific Tolerance. J Immunol. 2007 Feb 15;178(4):2155 LP – 2162.

97. Liu VC, Wong LY, Jang T, Shah AH, Park I, Yang X, et al. Tumor Evasion of the Immune System by Converting CD4+CD25− T Cells into CD4+CD25+; T Regulatory Cells: Role of Tumor-Derived TGF-β. J Immunol. 2007 Mar 1;178(5):2883 LP – 2892.

98. Curti A, Pandolfi S, Valzasina B, Aluigi M, Isidori A, Ferri E, et al. Modulation of tryptophan catabolism by human leukemic cells results in the conversion of CD25− into CD25+ T regulatory cells. Blood. 2007 Apr 1;109(7):2871 LP – 2877.

99. De Simone M, Arrigoni A, Rossetti G, Gruarin P, Ranzani V, Politano C, et al. Transcriptional Landscape of Human Tissue Lymphocytes Unveils Uniqueness of Tumor-Infiltrating T Regulatory Cells. Immunity. 2016;45(5):1135–47.

100. Plitas G, Konopacki C, Wu K, Bos PD, Morrow M, Putintseva EV, et al. Regulatory T Cells Exhibit Distinct Features in Human Breast Cancer. Immunity. 2016;45(5):1122–34.

101. Tanaka A, Sakaguchi S. Regulatory T cells in cancer immunotherapy. Cell Res. 2017 Jan 20;27(1):109–18.

102. Selby MJ, Engelhardt JJ, Quigley M, Henning KA, Chen T, Srinivasan M, et

al. Anti-CTLA-4 Antibodies of IgG2a Isotype Enhance Antitumor Activity through Reduction of Intratumoral Regulatory T Cells. Cancer Immunol Res. 2013 Jul 1;1(1):32 LP – 42.

103. Zappasodi R, Serganova I, Cohen IJ, Maeda M, Shindo M, Senbabaoglu Y, et al. CTLA-4 blockade drives loss of T(reg) stability in glycolysis-low tumours. Nature. 2021 Mar;591(7851):652–8.

104. Kurose K, Ohue Y, Wada H, Iida S, Ishida T, Kojima T, et al. Phase Ia Study of FoxP3+ CD4 Treg Depletion by Infusion of a Humanized Anti-CCR4 Antibody, KW-0761, in Cancer Patients. Clin Cancer Res. 2015 Oct 1;21(19):4327 LP – 4336.

105. Nishikawa H, Sakaguchi S. Regulatory T cells in tumor immunity. Int J Cancer. 2010;127(4):759–67.

106. Consortium IHGS. Finishing the euchromatic sequence of the human genome. Nature. 2004 Oct 21;431:931.

107. Deans C, Maggert KA. What Do You Mean, “Epigenetic”? Genetics. 2015 Apr 1;199(4):887 LP – 896.

108. Waddington CH. The Epigenotype. Int J Epidemiol. 2012 Feb 1;41(1):10–3.

109. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nat Rev Genet. 2016 Jun 27;17:487.

110. Larsson L. Current Concepts of Epigenetics and Its Role in Periodontitis. Curr Oral Heal Reports. 2017 Nov 6;4(4):286–93.

111. Goldberg AD, Allis CD, Bernstein E. Epigenetics: a landscape takes shape. Cell. 2007 Feb;128(4):635–8.

112. Bloom K, Joglekar A. Towards building a chromosome segregation machine. Nature. 2010 Jan 28;463(7280):446–56.

113. Kornberg RD, Thonmas JO. Chromatin Structure: Oligomers of the Histones. Science (80-). 1974 May 24;184(4139):865 LP – 868.

114. Hergeth SP, Schneider R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. EMBO Rep. 2015 Nov 16;16(11):1439–53.

115. Li B, Carey M, Workman JL. The Role of Chromatin during Transcription. Cell. 2007 Jan 14;128(4):707–19.

116. Lyko F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. Nat Rev Genet. 2018 Feb;19(2):81–92.

117. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012 May;13(7):484–92.

118. Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol. 2019 Oct;20(10):590–607.

119. Gonzalo S. Epigenetic alterations in aging. J Appl Physiol. 2010 Aug;109(2):586–97.

120. Kouzarides T. Chromatin Modifications and Their Function. Cell. 2007 Dec 17;128(4):693–705.

121. Kebede AF, Schneider R, Daujat S. Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest. FEBS J. 2015 May 1 [cited 2017 Dec 18];282(9):1658–74.

122. Torres IO, Fujimori DG. Functional coupling between writers, erasers and readers of histone and DNA methylation. Curr Opin Struct Biol. 2015 Dec 9;35:68–75.

123. Marmorstein R, Zhou M-M. Writers and Readers of Histone Acetylation: Structure, Mechanism, and Inhibition. Cold Spring Harb Perspect Biol. 2014 Jul;6(7):a018762.

124. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011 Mar 15;21(3):381–95.

125. Seto E, Yoshida M. Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. Cold Spring Harb Perspect Biol. 2014 Apr;6(4):a018713.

126. Erdel F. How Communication Between Nucleosomes Enables Spreading and Epigenetic Memory of Histone Modifications. BioEssays. 2017;39(12):1700053-n/a.

127. Plank JL, Dean A. Enhancer Function: Mechanistic and Genome-Wide

Insights Come Together. Mol Cell. 2014 Jan 23;55(1):5–14.

128. Trievel RC, Beach BM, Dirk LMA, Houtz RL, Hurley JH. Structure and Catalytic Mechanism of a SET Domain Protein Methyltransferase. Cell. 2002;111(1):91–103.

129. Sadakierska-Chudy A, Filip M. A Comprehensive View of the Epigenetic Landscape. Part II: Histone Post-translational Modification, Nucleosome Level, and Chromatin Regulation by ncRNAs. Neurotox Res. 2015 Feb;27(2):172–97.

130. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, et al. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. Science (80-). 2002 Nov 1;298(5595):1039 LP – 1043.

131. Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, et al. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev . 2002 Jul 15;16(14):1779–91.

132. Peters AHFM, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schöfer C, et al. Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability. Cell. 2001 Jan 14;107(3):323–37.

133. Soares LM, Radman-Livaja M, Lin SG, Rando OJ, Buratowski S. Feedback Control of Set1 Protein Levels Is Important for Proper H3K4 Methylation Patterns. Cell Rep. 2014 Jan 14;6(6):961–72.

134. Edmunds JW, Mahadevan LC, Clayton AL. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. EMBO J. 2008 Jan 23;27(2):406–20.

135. Klose RJ, Zhang Y. Regulation of histone methylation by demethylimination and demethylation. Nat Rev Mol Cell Biol. 2007 Mar 7;8:307.

136. Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. Cell Res. 2011 Apr 22;21(4):564–78.

137. Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. Spatial Distribution of Di- and Tri-methyl Lysine 36 of Histone H3 at Active Genes. J Biol Chem. 2005 May 6;280(18):17732–6.

138. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, et al. Genomic Maps and Comparative Analysis of Histone Modifications in Human and Mouse. Cell. 2005 Jan 14;120(2):169–81.

139. Rice JC, Briggs SD, Ueberheide B, Barber CM, Shabanowitz J, Hunt DF, et al. Histone Methyltransferases Direct Different Degrees of Methylation to Define Distinct Chromatin Domains. Mol Cell. 2003 Jan 14;12(6):1591–8.

140. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci. 2010 Dec 14 [cited 2021 Jun 3];107(50):21931–6.

141. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. 2011;

142. Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol. 2003 Dec 7;6:73.

143. Zhang T, Cooper S, Brockdorff N. The interplay of histone modifications – writers that read. EMBO Rep. 2015 Nov 1;16(11):1467 LP – 1481.

144. Carvalho S, Raposo AC, Martins FB, Grosso AR, Sridhara SC, Rino J, et al. Histone methyltransferase SETD2 coordinates FACT recruitment with nucleosome dynamics during transcription. Nucleic Acids Res. 2013 Mar 15;41(5):2881–93.

145. Cai L, Rothbart SB, Lu R, Xu B, Chen W-Y, Tripathy A, et al. An H3K36 methylation engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. Mol Cell. 2013 Feb 7;49(3):571–82.

146. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. Cell. 2006 Jan 14;125(2):315–26.

147. Kitagawa Y, Ohkura N, Kidani Y, Vandenbon A, Hirota K, Kawakami R, et al. Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment. Nat Immunol. 2017;18(2).

148. Ohkura N, Sakaguchi S. Transcriptional and epigenetic basis of Treg cell development and function: its genetic anomalies or variations in autoimmune diseases. Cell Res. 2020;30(6):465–74.

149. Delacher M, Imbusch CD, Hotz-Wagenblatt A, Mallm J-P, Bauer K, Simon M, et al. Precursors for Nonlymphoid-Tissue Treg Cells Reside in Secondary Lymphoid Organs and Are Programmed by the Transcription Factor BATF. Immunity. 2020 Feb;52(2):295-312.e11.

150. Floess S, Freyer J, Siewert C, Baron U, Olek S, Polansky J, et al. Epigenetic control of the foxp3 locus in regulatory T cells. PLoS Biol. 2007 Feb;5(2):e38.

151. Zheng Y, Josefowicz S, Chaudhry A, Peng XP, Forbush K, Rudensky AY. Role of conserved non-coding DNA elements in the Foxp3 gene in regulatory T-cell fate. Nature. 2010;463(7282):808–12.

152. Ohkura N, Hamaguchi M, Morikawa H, Sugimura K, Tanaka A, Ito Y, et al. T cell receptor stimulation-induced epigenetic changes and Foxp3 expression are independent and complementary events required for Treg cell development. Immunity. 2012 Nov;37(5):785–99.

153. Wei G, Wei L, Zhu J, Zang C, Hu-Li J, Yao Z, et al. Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4+ T cells. Immunity. 2009 Jan;30(1):155–67.

154. Arvey A, van der Veeken J, Samstein RM, Feng Y, Stamatoyannopoulos JA, Rudensky AY. Inflammation-induced repression of chromatin bound by the transcription factor Foxp3 in regulatory T cells. Nat Immunol. 2014;15(6):580–7.

155. He H, Ni B, Tian Y, Tian Z, Chen Y, Liu Z, et al. Histone methylation mediates plasticity of human FOXP3(+) regulatory T cells by modulating signature gene expressions. Immunology. 2014 Mar;141(3):362–76.

156. Feng Y, van der Veeken J, Shugay M, Putintseva E V, Osmanbeyoglu HU, Dikiy S, et al. A mechanism for expansion of regulatory T-cell repertoire and its role in self-tolerance. Nature. 2015 Dec;528(7580):132–6.

157. Buenrostro J, Wu B, Chang H, Greenleaf W. ATAC-seq: A Method for

Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol. 2015 Jan 5;109:21.29.1-21.29.9.

158. FastQC. 2015. https://qubeshub.org/resources/fastqc

159. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25.

160. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009 Jul;25(14):1754–60.

161. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based Analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.

162. Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nat Commun. 2018;9(1):189.

163. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016 Jul;44(W1):W160-5.

164. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature. 2012 Jan;481(7381):389–93.

165. Mumbach MR, Satpathy AT, Boyle EA, Dai C, Gowen BG, Cho SW, et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. Nat Genet. 2017;49(11):1602–12.

166. Love MI, Anders S, Huber W. Differential analysis of count data - the DESeq2 package. Vol. 15, Genome Biology. 2014. 550 p.

167. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods. 2012;9(3):215–6.

168. Carlson M. org.Hs.eg.db: Genome wide annotation for Human. R package version 3.8.2. 2019;

169. Nuñez JK, Chen J, Pommier GC, Cogan JZ, Replogle JM, Adriaens C, et al. Genome-wide programmable transcriptional memory by CRISPR-based

epigenome editing. Cell. 2021 Apr 29;184(9):2503-2519.e17.

170. Gagnon KT, Li L, Janowski BA, Corey DR. Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading. Nat Protoc. 2014;9(9):2045–60.

171. Yeo NC, Chavez A, Lance-Byrne A, Chan Y, Menn D, Milanova D, et al. An enhanced CRISPR repressor for targeted mammalian gene regulation. Nat Methods. 2018;15(8):611–6.

172. Schmidl C, Rendeiro AF, Sheffield NC, Bock C. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. Nat Methods. 2015 Oct 17;12(10):963–5.

173. Buenrostro J, Wu B, Chang H, Greenleaf W. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide.

174. Yan F, Powell DR, Curtis DJ, Wong NC. From reads to insight: a hitchhiker's guide to ATAC-seq data analysis. Genome Biol. 2020 Feb;21(1):22.

175. Shabalina SA, Spiridonov NA. The mammalian transcriptome and the function of non-coding DNA sequences. Genome Biol. 2004;5(4):105.

176. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317–30.

177. Wang C, Thangamani S, Kim M, Gu B-H, Lee JH, Taparowsky EJ, et al. BATF is required for normal expression of gut-homing receptors by T helper cells in response to retinoic acid. J Exp Med. 2013 Mar;210(3):475–89.

178. Bao K, Carr T, Wu J, Barclay W, Jin J, Ciofani M, et al. BATF Modulates the Th2 Locus Control Region and Regulates CD4+ T Cell Fate during Antihelminth Immunity. J Immunol. 2016 Dec;197(11):4371–81.

179. Vasanthakumar A, Moro K, Xin A, Liao Y, Gloury R, Kawamoto S, et al. The transcriptional regulators IRF4, BATF and IL-33 orchestrate development and maintenance of adipose tissue-resident regulatory T cells. Nat Immunol. 2015 Mar;16(3):276–85.

180. Liu Y, Carlsson R, Comabella M, Wang J, Kosicki M, Carrion B, et al. FoxA1

directs the lineage and immunosuppressive properties of a novel regulatory T cell population in EAE and MS. Nat Med. 2014;20(3):272–82.

181. Liang J, Tian C, Zeng Y, Yang Q, Liu Y, Liu Y, et al. FOXA1+ regulatory T cells: A novel T cell subset that suppresses antitumor immunity in lung cancer. Biochem Biophys Res Commun. 2019;514(1):308–15.

182. Malhotra N, Leyva-Castillo JM, Jadhav U, Barreiro O, Kam C, O'Neill NK, et al. RORα-expressing T regulatory cells restrain allergic skin inflammation. Sci Immunol. 2018 Mar;3(21).

183. Richer MJ, Lang ML, Butler NS. T Cell Fates Zipped Up: How the Bach2 Basic Leucine Zipper Transcriptional Repressor Directs T Cell Differentiation and Function. J Immunol. 2016 Aug 15;197(4):1009 LP – 1015.

184. Mouly E, Chemin K, Nguyen HV, Chopin M, Mesnard L, Leite-de-Moraes M, et al. The Ets-1 transcription factor controls the development and function of natural regulatory T cells. J Exp Med. 2010 Sep;207(10):2113–25.

185. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell. 2015 Oct;163(3):759–71.

186. Della Chiara G, Gervasoni F, Fakiola M, Godano C, D'Oria C, Azzolin L, et al. Epigenomic landscape of human colorectal cancer unveils an aberrant core of pan-cancer enhancers orchestrated by YAP/TAZ. Nat Commun. 2021;12(1).

187. Delacher M, Simon M, Sanderink L, Hotz-Wagenblatt A, Wuttke M, Schambeck K, et al. Single-cell chromatin accessibility landscape identifies tissue repair program in human regulatory T cells. Immunity. 2021 Apr;54(4):702-720.e17.

188. Kawakami R, Kitagawa Y, Chen KY, Arai M, Ohara D, Nakamura Y, et al. Distinct Foxp3 enhancer elements coordinate development, maintenance, and function of regulatory T cells. Immunity. 2021 May;54(5):947-961.e8.

189. Ahuja N, Easwaran H, Baylin SB. Harnessing the potential of epigenetic therapy to target solid tumors. J Clin Invest. 2014 Jan;124(1):56–63.

190. Ghosh S, Roy-Chowdhuri S, Kang K, Im S-H, Rudra D. The transcription

factor Foxp1 preserves integrity of an active Foxp3 locus in extrathymic Treg cells. Nat Commun. 2018;9(1):4473.

191. Grant FM, Yang J, Nasrallah R, Clarke J, Sadiyah F, Whiteside SK, et al. BACH2 drives quiescence and maintenance of resting Treg cells to promote homeostasis and cancer immunosuppression. J Exp Med. 2020 Sep;217(9).

192. Koizumi S-I, Ishikawa H. Transcriptional Regulation of Differentiation and Functions of Effector T Regulatory Cells. Cells. 2019 Aug;8(8).

193. Nuñez JK, Chen J, Pommier GC, Cogan JZ, Replogle JM, Adriaens C, et al. Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. Cell. 2021 Apr;184(9):2503-2519.e17.

194. Madsen JGS, Madsen MS, Rauch A, Traynor S, Van Hauwaert EL, Haakonsson AK, et al. Highly interconnected enhancer communities control lineage-determining genes in human mesenchymal stem cells. Nat Genet. 2020 Nov;52(11):1227–38.

195. Neal JT, Li X, Zhu J, Giangarra V, Grzeskowiak CL, Ju J, et al. Organoid Modeling of the Tumor Immune Microenvironment. Cell. 2018 Dec;175(7):1972-1988.e16.

LIST OF FIGURES AND TABLES

DISSEMINATION OF RESULTS

The ultimate long-term goal of this proposal is the identification of novel targets for the modulation of anti-tumor immune responses. The topic of these novel therapies to treat cancer, although of great interest to general public, is often hardly intelligible and unfortunately represents a fertile ground for fake news. Moreover, even medical doctors who are at the front-line in the communication with patients are often not fully aware of how research in the field is progressing. In this perspective I plan to share the results obtained during my PhD to:

1) engage the scientific and medical community through

- regularly scheduled meeting reports, invited seminars in other institutions, attendance to scientific international and national meetings
- practical courses on the technological and bioinformatics platforms used in the laboratory
- result publications on international peer-reviewed journals under open-access policy

2) engage the civil society by promoting knowledge sharing and the public awareness through

- social media (e.g Facebook, Twitter) and Institute website
- participation of public events organized by AIRC for students and general public to make specialized information more accessible

Summary

The combination of genetic mutations and epigenetic modifications that are peculiar to all tumors generate anti-tumor T cell responses in cancer patients that are usually disabled upon tumor progression by suppressive mechanisms triggered by the interplay between malignant cells and the tumor microenvironment. Among the cells of the immune system that infiltrate tumor sites, regulatory T cells (Treg), which are physiologically engaged in the maintenance of immunological self-tolerance, are potent suppressors of anti-tumor immunity. Depletion of tumor infiltrating Treg cells can boost specific immune responses against tumors. Therefore, to identify novel and more specific therapeutic strategies aimed at Treg cell elimination from the tumor microenvironment we performed a detailed molecular characterization of these cells that unveiled the regulatory networks that shape their identity and ultimately affect their function.

Riassunto

Le cellule tumorali, accumulando mutazioni genetiche ed epigenetiche, vengono riconosciute come cellule anomale, e dunque eliminate, dal nostro sistema immunitario, in particolare dalle cellule T effettrici. Tuttavia, nelle fasi più avanzate, il tumore è in grado di eludere il sistema immunitario reclutando cellule T regolatorie (Treg) in grado di sopprimere la risposta immunitaria contro il tumore. La riduzione o la soppressione delle cellule Treg all'interno del tumore potrebbe quindi migliorare le risposte immunitarie contro il tumore. Per questo motivo abbiamo deciso di caratterizzare da un punto di vista molecolare queste cellule in modo tale da comprendere in dettaglio quali sono I meccanismi regolatori che definiscono la loro identità e in influenzano la loro funzione.