Resource

Immunity

# Transcriptional Landscape of Human Tissue Lymphocytes Unveils Uniqueness of Tumor-Infiltrating T Regulatory Cells

## Highlights

- Transcriptome analysis performed on tumor-resident $CD4^+$ Th1, Th17, and Treg cells
- Tumor-infiltrating Treg cells are defined by the expression of signature genes
- Treg-specific signature genes correlate with patients' survival in both CRC and NSCLC

## Authors

Marco De Simone, Alberto Arrigoni, Grazisa Rossetti, ..., Hendrik G. Stunnenberg, Sergio Abrignani, Massimiliano Pagani

## Correspondence

abrignani@ingm.org (S.A.), pagani@ingm.org (M.P.)

## In Brief

Tumor-infiltrating regulatory T cells can suppress effector T cells specific for tumor antigens. De Simone et al. (2016) demonstrate that tumor-infiltrating Treg cells display specific gene signatures that were also validated at the single-cell level. These data can contribute to dissect the molecular networks underlying the biology of tumor-infiltrating Treg cells. As part of the IHEC consortium, this study integrates genetic, epigenetic, and transcriptomic profiling in three immune cell types from nearly 200 people to characterize the distinct and cooperative contributions of diverse genomic inputs to transcriptional variation. Explore the Cell Press IHEC webportal at www.cell.com/consortium/IHEC.

## Accession Numbers

E-MTAB-2319
E-MTAB-513
GSE50760
GSE40419
GSE17536
GSE41271

De Simone et al., 2016, Immunity *45*, 1135–1147
November 15, 2016 
http://dx.doi.org/10.1016/j.immuni.2016.10.021

CellPress

Immunity
Resource

# Transcriptional Landscape of Human Tissue Lymphocytes Unveils Uniqueness of Tumor-Infiltrating T Regulatory Cells

Marco De Simone,[1,13] Alberto Arrigoni,[1,13] Grazisa Rossetti,[1,13] Paola Gruarin,[1,13] Valeria Ranzani,[1] Claudia Politano,[1] Raoul J.P. Bonnal,[1] Elena Provasi,[1] Maria Lucia Sarnicola,[1] Ilaria Panzeri,[1] Monica Moro,[1] Mariacristina Crosti,[1] Saveria Mazzara,[1] Valentina Vaira,[1,2,4] Silvano Bosari,[2,4] Alessandro Palleschi,[3] Luigi Santambrogio,[3,4] Giorgio Bovo,[5] Nicola Zucchini,[5] Mauro Totis,[6] Luca Gianotti,[6,7] Giancarlo Cesana,[7] Roberto A. Perego,[7] Nirvana Maroni,[8] Andrea Pisani Ceretti,[8] Enrico Opocher,[8,9] Raffaele De Francesco,[1] Jens Geginat,[1] Hendrik G. Stunnenberg,[10] Sergio Abrignani,[1,11,*] and Massimiliano Pagani[1,12,14,*]

[1]Istituto Nazionale Genetica Molecolare INGM 'Romeo ed Enrica Invernizzi,' Milan 20122, Italy
[2]Division of Pathology, IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan 20122, Italy
[3]Division of Thoracic Surgery, IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan 20122, Italy
[4]Department of Pathophysiology and Organ Transplantation, Università degli Studi di Milano, Milano 20122, Italy
[5]Department of Pathology, San Gerardo Hospital, Monza 20900, Italy
[6]Department of Surgery, San Gerardo Hospital, Monza 20900, Italy
[7]School of Medicine and Surgery, Milano-Bicocca University, Monza 20900 Italy
[8]UO Chirurgia Epatobiliopancreatica e Digestiva Ospedale San Paolo, Milan 20142, Italy
[9]Department of Health Sciences, Università degli Studi di Milano, Milano 20122, Italy
[10]Department of Molecular Biology, Faculty of Science, Radboud University, Nijmegen, The Netherlands
[11]Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milano 20122, Italy
[12]Department of Medical Biotechnology and Translational Medicine, Università degli Studi di Milano, Milano 20129, Italy
[13]Co-first author
[14]Lead contact
*Correspondence: abrignani@ingm.org (S.A.), pagani@ingm.org (M.P.)
http://dx.doi.org/10.1016/j.immuni.2016.10.021

## SUMMARY

**Tumor-infiltrating regulatory T lymphocytes (Treg) can suppress effector T cells specific for tumor antigens. Deeper molecular definitions of tumor-infiltrating-lymphocytes could thus offer therapeutic opportunities. Transcriptomes of T helper 1 (Th1), Th17, and Treg cells infiltrating colorectal or non-small-cell lung cancers were compared to transcriptomes of the same subsets from normal tissues and validated at the single-cell level. We found that tumor-infiltrating Treg cells were highly suppressive, upregulated several immune-checkpoints, and expressed on the cell surfaces specific signature molecules such as interleukin-1 receptor 2 (IL1R2), programmed death (PD)-1 Ligand1, PD-1 Ligand2, and CCR8 chemokine, which were not previously described on Treg cells. Remarkably, high expression in whole-tumor samples of Treg cell signature genes, such as *LAYN*, *MAGEH1*, or *CCR8*, correlated with poor prognosis. Our findings provide insights into the molecular identity and functions of human tumor-infiltrating Treg cells and define potential targets for tumor immunotherapy.**

## INTRODUCTION

The combination of genetic mutations and epigenetic modifications that are peculiar to all tumors generate antigens that T and B lymphocytes can use to specifically recognize tumor cells (Jamal-Hanjani et al., 2013). It is increasingly clear that T lymphocytes recognizing tumor-derived peptides presented by major histocompatibility complex (MHC) molecules play a central role in immunotherapy and in conventional chemo-radiotherapy of cancer (Galluzzi et al., 2015). In fact, anti-tumor T cell responses arise in cancer patients but are disabled upon tumor progression by suppressive mechanisms triggered by the interplay between malignant cells and the tumor microenvironment (Munn and Bronte, 2016). The tumor-dependent immunosuppressive mechanisms depend on the integrated action of infiltrating leukocytes and lymphocytes that upregulate a range of modulatory molecules, collectively called immune checkpoints, whose function is only partially characterized (Pardoll, 2012). Therefore, the search for agonists of co-stimulatory complexes or antagonists of inhibitory molecules to potentiate antigen-specific T cell responses is a primary goal of current anti-tumor research (Sharma and Allison, 2015; Zitvogel et al., 2013). Indeed, clinical trials have unequivocally shown that the blockade of immune checkpoints unleashes the spontaneous anti-tumor immune responses in such a powerful way that it has created a paradigm shift in cancer therapy (Śledzińska et al., 2015; Topalian et al., 2015).

Among the immune checkpoints targeted by blocking strategies, CTLA-4 has been one of the first to be translated into therapeutic applications.

Anti-CTLA-4 monoclonal antibodies (mAb) show remarkable success in metastatic melanoma, and more recently in non-small-cell lung cancer, prostate cancer, renal cell carcinoma, urothelial carcinoma, and ovarian cancer (Carthon et al., 2010;

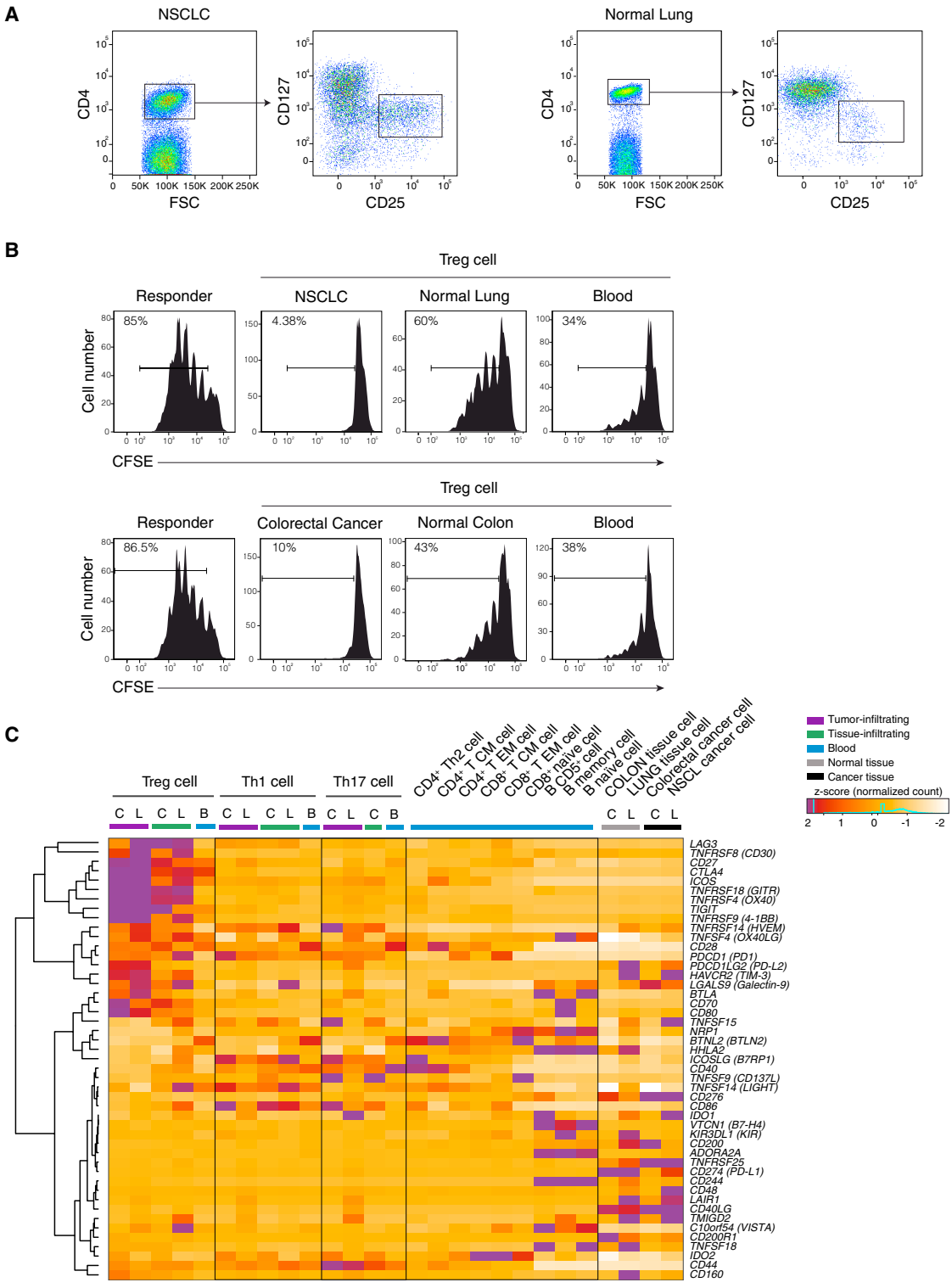



**Figure 1. Purification, Functional Characterization, and Expression of Immune Checkpoints in Tumor Infiltrating Cells**

(A) Representation of the sorting strategy of Treg cells infiltrating tumor or normal tissue.

(B) Representative flow cytometry plots showing suppressive activity of Treg cells isolated from tumor (NSCLC or CRC), normal tissue and blood of the same patient. $4 \times 10^5$ carboxyfluorescein diacetate succinimidyl ester (CFSE)-labeled $CD4^+$ naive T cells from healthy donors were cocultured with an equal number of Treg cells for 4 days with a CD3-specific mAb and $CD1c^+CD11c^+$ dendritic cells. Percentage of proliferating cells is indicated. Data are representative of three independent experiments.

*(legend continued on next page)*

Hodi et al., 2010; van den Eertwegh et al., 2012; Yang et al., 2007). However, the fraction of patients that do not respond remains high, prompting a deeper investigation of the mechanisms underpinning the modulation of immune responses by tumors. Recent experimental evidence shows that anti-CTLA-4 mAb efficacy depends on FcγR-mediated depletion of $CD4^+$ regulatory T cells (Treg cells) within the tumor microenvironment (Peggs et al., 2009; Selby et al., 2013; Simpson et al., 2013; Twyman-Saint Victor et al., 2015).

Treg cells, which are physiologically engaged in the maintenance of immunological self-tolerance and immune homeostasis (Josefowicz et al., 2012; Sakaguchi et al., 2008), are potent suppressors of effector cells and are found at high frequencies in various types of cancers (Fridman et al., 2012; Nishikawa and Sakaguchi, 2010). Treg cells adapt their transcriptional program to the various cytokines to which they are exposed in the inflammatory milieu (Campbell and Koch, 2011). This versatility is controlled by transcription factors generally associated with the differentiation of other effector $CD4^+$ T cell subsets, resulting in various Treg cell populations with unique features and immunomodulatory functions (Duhen et al., 2012; Geginat et al., 2014). Moreover, Treg cells infiltrating non-lymphoid tissues are reported to exhibit unique phenotypes and transcriptional signatures, because they can display functions beyond their well-established suppressive roles, such as metabolic modulation in adipose tissue (Cipolletta et al., 2012) or regulation of tissue repair in skeletal muscle (Burzyn et al., 2013) and in lung tissue (Arpaia et al., 2015).

Treg cell depletion has been reported to increase anti-tumor specific immune responses and to reduce tumor burden (Marabelle et al., 2013; Teng et al., 2010; Walter et al., 2012). Although promising clinical results have been achieved with Treg cell depleting strategies, some relevant issues are to be addressed, for a safer, more effective, and wider clinical application of these therapies. First, severe autoimmunity can occur following systemic Treg cells depletion (Nishikawa and Sakaguchi, 2010), which could be avoided if selective depletion of tumor infiltrating Treg cells were feasible. A second issue concerns the specificity of targeting. Indeed, Treg cells share with effector lymphocytes most of the molecules targeted for therapy, which can possibly deplete also the tumor-specific effector cells. Therefore, the molecular characterization of Treg cells at different tumor sites should help to better define therapeutic targets through a better description of their signature molecules and of the network that regulates Treg cell functions in the tumor microenvironment.

Non-small-cell lung cancer (NSCLC) and colorectal cancer (CRC) are the two most frequent cancers in both genders (Torre et al., 2015). NSCLC has the worst prognosis due to its high mortality rate even in early stages. Although CRC survival rate is highly dependent on the tumor stage at diagnosis, about 50% of patients will progress to metastatic cancer (Gonzalez-Pons and Cruz-Correa, 2015). Both tumors have been targeted with therapies based on monoclonal antibodies to checkpoint inhibitors, but the outcomes have been different. While remarkable clinical success has been obtained in NSCLC, evidence of durable response in CRC is scarce with the exception of mismatch repair-deficient CRC lesions (Jacobs et al., 2015; Kroemer et al., 2015; Le et al., 2015).

Here we provide a comprehensive transcriptome analysis of human $CD4^+$ Treg cells and effector cells (Th1 and Th17) infiltrating NSCLC or CRC and their matched normal tissues. We defined molecular signatures of tumor-infiltrating Treg cells in these two cancer types and confirmed the relevance of these signatures by single-cell analyses. These data could help a better understanding of Treg functional role at tumor sites and pave the way to the identification of therapeutic targets for more specific and safer modulation of Treg cells in cancer therapy.

## RESULTS

### Tumor Infiltrating Treg Cells Upregulate Immune Checkpoints and Are Highly Suppressive

To assess the gene expression landscape of tumor infiltrating $CD4^+$ T cells, we isolated different $CD4^+$ lymphocytes subsets from two different tumors, NSCLC and CRC, from the adjacent normal tissues, and from peripheral blood samples. From all these tissues, we purified by flow cytometry (Figure 1A and S1A and S1B) $CD4^+$ Treg (36 samples from 18 individuals), Th1 (30 samples from 21 individuals), and Th17 (22 samples from 14 individuals) cells (Table 1 and Table S1). To assess Treg cell function, we tested their suppressor activity and showed that Treg cells infiltrating either type of tumor tissues have a remarkably stronger suppressive activity in vitro compared to Treg cells isolated from the adjacent normal tissue and peripheral blood of the same patients (Figure 1B).

The polyadenylated RNA fraction extracted from the sorted $CD4^+$ Treg, Th1, and Th17 cells was then analyzed by paired-end RNA sequencing obtaining about 4 billion mapped "reads" (Table 1). First, we interrogated RNA-sequencing data of $CD4^+$ T cells infiltrating both CRC and NSCLC and their matched normal tissues, to quantitate mRNA expression of known immune checkpoints and their ligands. Second, we analyzed RNA-seq data of CRC and NSCLC, as well as of normal colon and lung samples. We found that several immune checkpoints and their ligands transcripts were strikingly upregulated in tumor infiltrating Treg cells compared to both normal tissue and peripheral blood-derived Treg cells, as well as to T and B lymphocyte subsets purified from peripheral blood mononuclear cells (PBMCs) (Figures 1C and S1C and Table S5). Our findings highlight the specific expression patterns of immune checkpoints and their ligands in tumor infiltrating Treg and effector cells and suggest that their functional relevance should be investigated directly at tumor sites.

### Tumor-Infiltrating Treg Cells Express a Specific Gene Signature

We then asked whether tumor infiltrating Treg cells could be defined by specific gene-expression patterns. First, in order to

(C) *Z*-score normalized RNA-seq expression values of immune checkpoints genes are represented as a heatmap. Cell populations are reported as a color code in the upper part of the graph, while gene names have been assigned to heatmap rows. Hierarchical clustering results are shown as a dendrogram drawn on the left side of the matrix. Colon tissues are indicated as C, lung tissues as L, and peripheral blood as B.
See also Figure S1.

**Table 1. Purification and RNA-Sequencing of Human Primary Lymphocyte Subsets**

| Tissue | Subset | Sorting Phenotype | Number of Samples | Mapped Reads (M) |
|---|---|---|---|---|
| NSCLC | CD4+ Treg | $CD4^+$ $CD127^-$ $CD25^+$ | 8 | 587 |
| | $CD4^+$ Th1 | $CD4^+$ $CXCR3^+$ $CCR6^-$ | 8 | 409 |
| | $CD4^+$ Th17 | $CD4^+$ $CCR6^+$ $CXCR3^-$ | 6 | 206 |
| CRC | $CD4^+$ Treg | $CD4^+$ $CD127^-$ $CD25^+$ | 7 | 488 |
| | $CD4^+$ Th1 | $CD4^+$ $CXCR3^+$ $CCR6^-$ | 5 | 266 |
| | $CD4^+$ Th17 | $CD4^+$ $CCR6^+$ $CXCR3^-$ | 5 | 308 |
| Lung (normal tissue) | $CD4^+$ Treg | $CD4^+$ $CD127^-$ $CD25^+$ | 1 (pool of 6) | 73 |
| | $CD4^+$ Th1 | $CD4^+$ $CXCR3^+$ $CCR6^-$ | 1 (pool of 6) | 76 |
| Colon (normal tissue) | $CD4^+$ Treg | $CD4^+$ $CD127^-$ $CD25^+$ | 7 | 404 |
| | $CD4^+$ Th1 | $CD4^+$ $CXCR3^+$ $CCR6^-$ | 6 | 352 |
| | $CD4^+$ Th17 | $CD4^+$ $CCR6^+$ $CXCR3^-$ | 6 | 284 |
| PB (healthy donor) | $CD4^+$ Treg | $CD4^+$ $CD127^-$ $CD25^+$ | 8 | 259 |
| | $CD4^+$ Th1 | $CD4^+$ $CXCR3^+$ $CCR6^-$ | 5 | 70 |
| | $CD4^+$ Th17 | $CD4^+$ $CCR6^+$ $CXCR3^-$ | 5 | 77 |

For each cell subsets profiled by RNA-sequencing tissue of origin, surface marker combinations used for sorting, number of profiled samples, as well as number of mapped sequencing reads are indicated. M, million; CRC, colorectal cancer; NSCLC, non-small cell lung cancer; PB, peripheral blood.
See also Table S1.

capture the overall similarity between the tumor infiltrating lymphocytes, we performed a principal components analysis (PCA) on the whole transcriptomes. Tumor-infiltrating Treg cells purified from CRC and NSCLC tissues clustered together and were clearly separated from Th1 and Th17 cells purified from CRC and NSCLC tissues (Figures S2A and S2B). PCA showed a distinct grouping of Treg cells purified from different sites; in fact, separation along the first principal component (PC1) clearly divided peripheral blood Treg cells from tissue infiltrating Treg cells (Figure 2A), whereas normal-tissue and tumor-tissue infiltrating Treg cells are mostly divided by the second component (PC2). These findings indicate that tumor-infiltrating Treg cells have specific expression patterns compared not only to other $CD4^+$ T cell subsets but also compared to Treg cells isolated from normal tissues.

In order to identify genes that are preferentially expressed in tumor-infiltrating lymphocytes, we performed self-organizing maps (SOM) analyses that provide a powerful way to define coordinated gene-expression patterns that are visualized in spatial proximity in a 2D mosaic grid heatmap (Wirth et al., 2012). In this way, we analyzed 7,763 genes that were differentially expressed between the different $CD4^+$ T cell subsets purified from PBMCs and tumor tissues (DESeq2 package; FDR < 0.05). Among the different $CD4^+$ T cell subsets (Th1, Th17, and Treg) assessed with SOM, only the tumor-infiltrating Treg cells displayed peculiar gene-expression patterns that were similar between NSCLC and CRC samples (Figures 2B and S2C), thus allowing the identification (FDR < 0.1) of transcripts upregulated in both CRC and NSCLC infiltrating Treg cells (Figure 2C and Table S2). Gene-ontology (GO) analyses of those genes upregulated in tumor infiltrating Treg cells showed significant enrichment for terms related to lymphocytes activation (Figure 2C and Table S3).

To identify signature transcripts of tumor-infiltrating Treg cells, we included in the expression pattern analyses the transcriptome datasets we previously obtained from different T and B lymphocyte subsets purified from PBMCs (Ranzani et al., 2015). In so doing, we obtained a signature of 309 transcripts whose expression is higher in tumor infiltrating Treg cells (Wilcoxon Mann Whitney test $p < 2.2 \times 10–16$) (Figures 2D and S2D and Table S4) compared to the other lymphocyte subsets purified from non-tumoral tissues and from PBMCs of healthy or neoplastic patients.

Altogether, the data show that Treg cells display the most pronounced differences in transcripts expression among $CD4^+$ T cell subsets infiltrating normal and tumor tissues. We defined a subset of signature genes that describe the specific gene-expression profile of tumor infiltrating Treg cells.

## Gene Signature of Tumor-Infiltrating Treg Cells Is Present in Primary and Metastatic Human Tumors

We then look at the single cell level for the differential expression profile of signature genes of tumor infiltrating Treg cells. We isolated $CD4^+$ T cells from 5 CRC and 5 NSCLC tumor samples, as well as from 5 PBMCs of healthy individuals (Table S1), purified Treg cells, and using an automated microfluidic system (C1 Fluidigm) captured single cells (a total of 858 Treg cells: 320 from CRC and 286 from NSCLC; 252 from PBMCs of healthy individuals). We then assessed by high throughput RT-qPCR (Biomark HD, Fluidigm) the expression of 79 genes selected among the highly expressed (> 10 FKPM) tumor Treg cell signature genes (Figures 3A, S3A and S3B).

Notably, we found that the vast majority (75 over 79; 95%) of the tumor-infiltrating Treg cell signatures were co-expressed with bona fide Treg cell markers (i.e., *FOXP3*$^+$ and *IL2RA*) (Figure 3B). The percentage of co-expression between these Treg cell markers and the 79 genes selected among the tumor-infiltrating-Treg-cell signature genes ranged between 81% of *TIGIT* and 0.59% of *CGA* (Figure 3B). The expression of Treg signature genes in the RNA-seq of the whole Treg cell population correlated with the percentage of single cells expressing the different genes (Figure 3C). In order to reduce the "drop-out" effect of the single cell data (i.e., events in which a transcript is detected in one cell but not in another one because the transcript is "missed" during the reverse-transcription step) (Kharchenko et al., 2014), we defined a threshold (median value t = 8.4%) based on the expression distribution for each transcript and

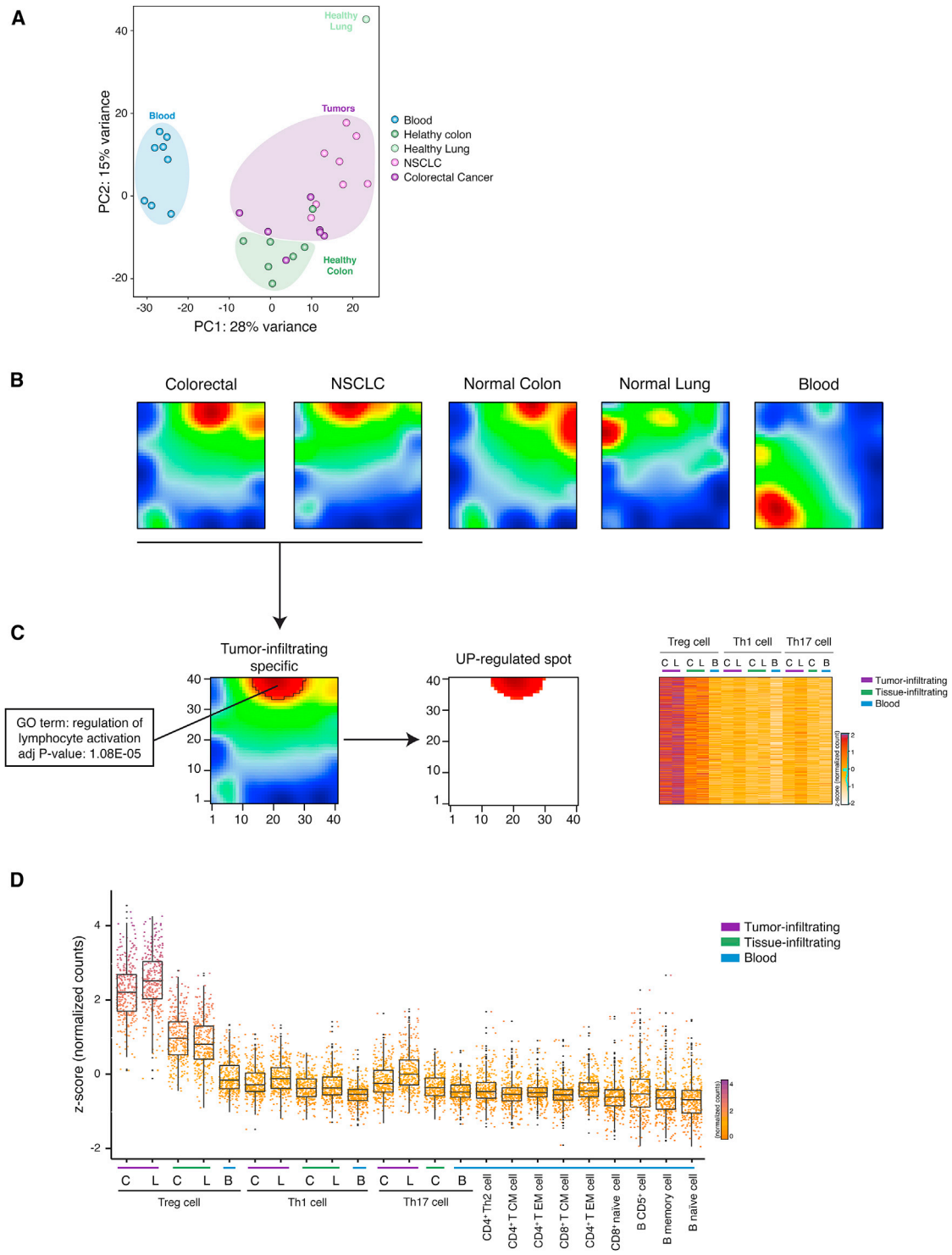



**Figure 2. SOM Analysis Identifies Co-regulated Genes in Tumor Infiltrating Treg Cells**

(A) PCA has been performed on rlog-normalized (DESeq2) counts for all T regulatory cell RNA-seq samples (36 samples from 18 individuals).

(B) Self-organizing maps analysis has been performed on the RNA-seq dataset comprising Treg, Th1, and Th17 cell subsets. Bidimensional SOM profiles are reported for Treg cells.

(C) Group-centered analysis for the identification of upregulated spot (FDR < 0.1) in Treg cells infiltrating both NSCLC and CRC is described as 2D heatmap. Heatmap representing *Z*-score normalized expression values of genes selected from the upregulated spot is shown on the right side of the figure. Top enriched

*(legend continued on next page)*

discarded genes below this threshold (see the Supplemental Experimental Procedures). The forty-five signature transcripts of tumor infiltrating Treg cells detected above this threshold were in most cases significantly overexpressed in Treg cells from both tumors (39 over 45, 87%; Wilcoxon Mann Whitney test $p < 0.05$) or in one tumor type (43 over 45, 96%; Figure 3D). Homogeneity of the purified tissue infiltrating Treg cells can be affected by the carry-over of cells from other lymphocyte subsets. To quantitate this possible contamination, the single cell RT-qPCR analyses of Treg cells was performed including markers specific for other lymphocytes subsets (i.e., Th1, Th2, Th17, Tfh, CD8 T cells, B cells) (Figure S3C). Our data showed that only a very low fraction of the purified single cells displayed markers of lymphocytes subsets different from Treg cells (Figure S3C).

The overlap between the signature genes in the CRC and NSCLC infiltrating Treg cells (Figure 2D) prompted us to assess whether this signature were also enriched in Treg cells infiltrating other tumors. RNA was thus extracted from Treg cells infiltrating breast cancer, gastric cancer, brain metastasis of NSCLC, and liver metastasis of CRC. We found by RT-qPCR that tumor infiltrating Treg signatures genes were mostly upregulated also in these tumors (Figure 3E).

Overall these data show that the tumor-infiltrating Treg cell signature genes are co-expressed at single cell level with *FOXP3* and *IL2RA and* that several primary and metastatic human tumors express the tumor-infiltrating Treg cell signature.

### Gene Signature of Tumor Infiltrating Treg Cells Is Translated in a Protein Signature

We then assessed at the single cell level by flow cytometry the protein expression of ten representative signature genes present in CRC and NSCLC infiltrating Treg cells, adjacent normal tissues, and patients PBMCs. Of the ten proteins, two were proteins (OX40 and TIGIT) whose relevance for Treg cells biology has been demonstrated (Joller et al., 2014; Voo et al., 2013), seven are proteins (BATF, CCR8, CD30, IL-1R2, IL-21R, PDL-1, and PDL-2) whose expression has never been described in tumor-infiltrating Treg cells, and one protein, 4-1BB, is a co-stimulatory receptor expressed on several hematopoietic cells, whose expression on Treg cells has been shown to mark antigen-activated cells (Schoenbrunn et al., 2012). Our findings showed that all these proteins were upregulated (Figure 4A), to different extent, in tumor infiltrating Treg cells compared to the Treg cells resident in normal tissues. Given the increasing interest in the PD1 - PDLs axis as targets for tumor immunotherapy, we assessed the effect of antibodies against PDL-1 and PDL-2 on the suppressive function of tumor-infiltrating Treg cells toward effector $CD4^+$ T cell proliferation in vitro. We found that preincubation of tumor infiltrating Treg cells with monoclonal antibodies against PDL-1 or PDL-2 reduced their suppressive activity as demonstrated by the increased proliferation of effector $CD4^+$ T cells (Figure 4B).

Altogether, our data show there is a molecular signature of tumor infiltrating Treg cells, which can be detected both at the mRNA and at the protein levels.

### Expression of Tumor Treg Signature Genes Is Negatively Correlated with Patient Survival

In an attempt to correlate our findings with clinical outcome, we asked whether the expression of the tumor-Treg signature transcripts correlated with disease prognosis in CRC and NSCLC patients. We therefore interrogated for expression of Treg signature genes transcriptomic datasets obtained from resected tumor tissues of a cohort of 177 CRC patients (GSE17536; Smith et al., 2010) and of a cohort of 263 NSCLC patients (GSE41271; Sato et al., 2013) and correlated high and low gene expression with the 5-year survival data. Among those genes whose expression is highly enriched in tumor-infiltrating Treg cells, we selected *LAYN*, *MAGEH1*, and *CCR8* that are the three genes more selectively expressed (Figure S5A). To normalize for differences in T cell densities within the resected tumor tissues, we used the ratio between expression of the selected signature genes and *CD3G*. We found that high expression of the three signature genes is in all cases correlated with a significantly reduced survival (Figure 5A). We also observed that expressions of the three signature genes increased with tumor staging of CRC patients (Figure 5B).

In conclusion, high expression in the whole-tumor samples of three genes (*LAYN*, *MAGEH1*, and *CCR8*) that are specifically and highly expressed in tumor infiltrating Treg cells correlates with a poor prognosis in both NSCLC and CRC patients.

## DISCUSSION

Diversity of tumor-infiltrating Treg cells should be fully elucidated to understand their functional relevance and prognostic significance in different types of cancer and to possibly improve the therapeutic efficacy of Treg cell modulation through the selective depletion of tumor infiltrating Treg cells. The transcriptome analysis we performed on CRC- and NSCLC-infiltrating T cells showed that tumor-infiltrating Treg cells are different from both circulating and normal tissue-infiltrating Tregs, suggesting that the tumor microenvironment influences specific gene expression in Treg cells. Our findings further support the view that Treg cells from different tissues are instructed by environmental factors to display different gene-expression profiles (Panduro et al., 2016). Indeed the list of signature genes includes a number of molecules that are consistently upregulated in tumor-infiltrating Treg cells isolated from different tumor types, and these signature genes would have not been identified if we had not profiled specifically tumor infiltrating Treg cells.

The number of genes highly expressed in tumor infiltrating cells, as defined by differential expression and SOM analyses, was significantly higher in Treg than in Th17 and Th1 cells, suggesting that Treg cells are more susceptible than other

GO term (DAVID) for genes assigned to upregulated spot is reported with the corresponding significance p value. Colon tissues are indicated as C, lung tissues as L, and peripheral blood as B.

(D) *Z*-score normalized expression values of genes that are preferentially expressed in tumor-infiltrating Treg cells (Wilcoxon Mann Whitney test $p < 2.2 \times 10–16$) over the listed cell subsets are represented as boxed plots. Colon tissues are indicated as C, lung tissues as L, and peripheral blood as B.

See also Figure S2.

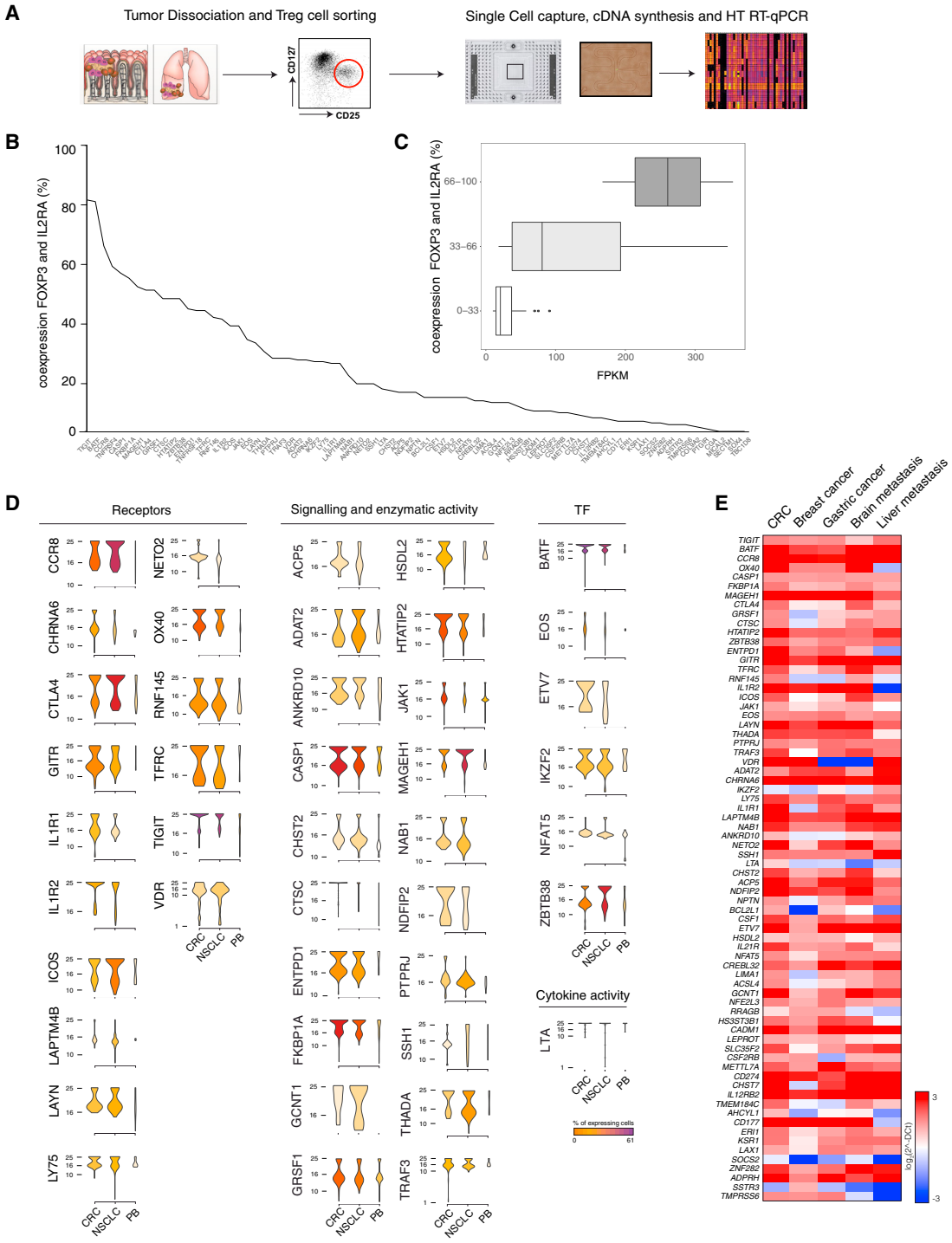



**Figure 3. Single Cell Analysis of Tumor Infiltrating Treg Cells**

(A) Schematic representation of the experimental workflow. Experiments were performed on Treg cells infiltrating CRC, NSCLC, or isolated from peripheral blood of healthy donors (PB); five samples were collected for each tissue.

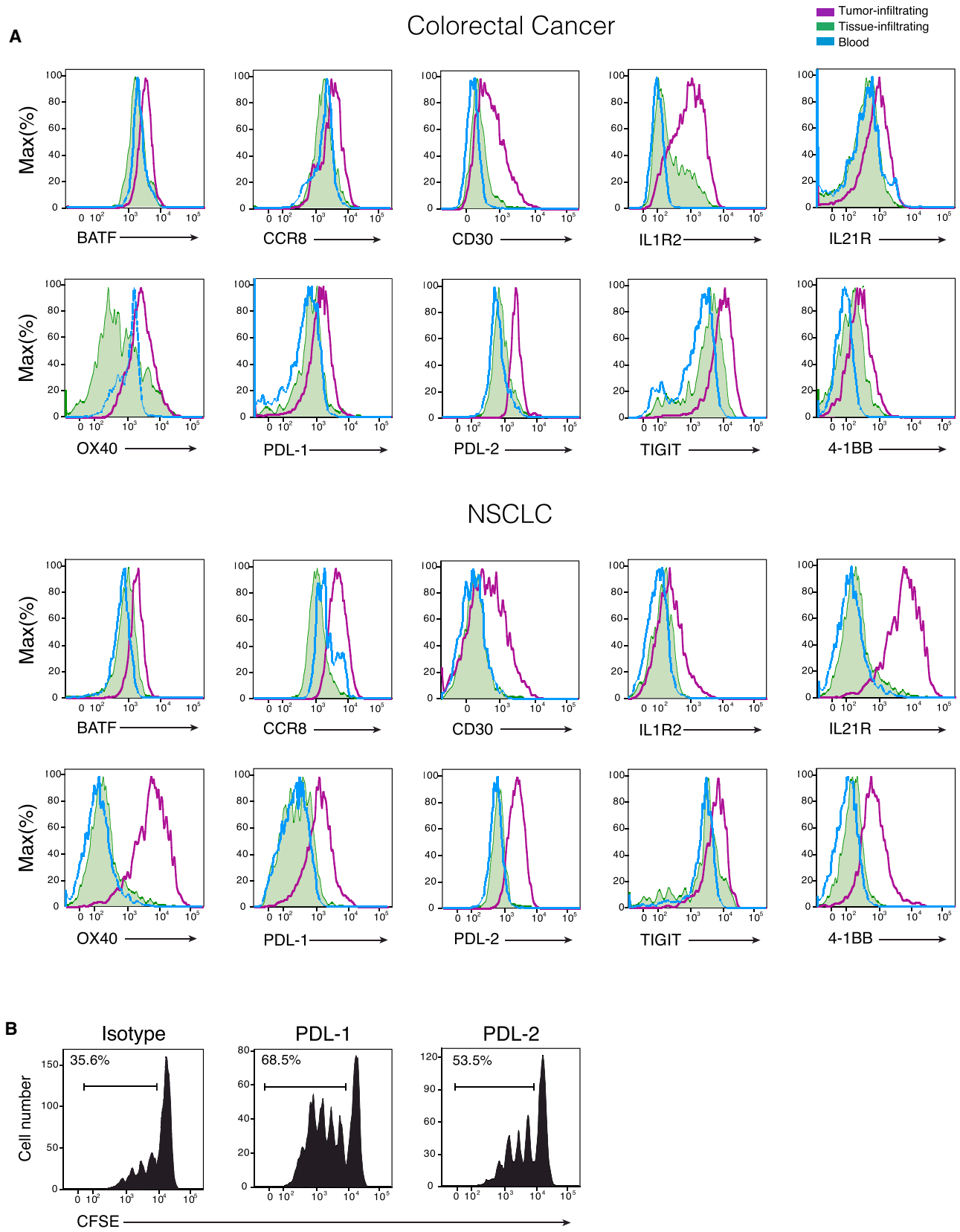
(B) Percentage of co-expression of signature genes with *FOXP3* and *IL2RA* is depicted.

(C) Expression levels of the signature genes classified by the percentage of co-expression are represented as boxplot.

(D) Expression distribution (violin plots) in Treg cells infiltrating CRC, NSCLC, or PB. Plots representing the ontology classes of receptors, signaling and enzymatic activity, cytokine activity, and transcription factors are shown (Wilcoxon Mann Whitney test $p < 0.05$). Color gradient indicates the percentage of cells expressing each gene in Treg cells isolated from the three tissues.

(E) Gene-expression analysis of tumor Treg signature genes in different tumor types. Expression values are expressed as log2 (2^-DCt).

See also Figure S3.




(legend on next page)

T cell subsets to external cues they are exposed to in tumor tissues. We found that tumor-infiltrating-Treg signature genes are not only largely shared between CRC- and NSCLC-infiltrating cells but are also conserved in breast and gastric cancers, as well as in CRC and NSCLC metastatic tumors (in liver and brain, respectively) suggesting that expression of these genes is a common feature of tumor infiltrating Treg cells that might correlate with Treg cell-specific function within the tumor microenvironment.

Although our knowledge on the function of immune checkpoints on lymphocytes is still incomplete, agonist or antagonist monoclonal antibodies targeting checkpoints are in clinical development. We have found that some of these checkpoints (such as GITR, OX40, TIGIT, LAG-3, and TIM-3) and some of their ligands (such as OX40LG, Galectin-9, CD70) are upregulated also in tumor-infiltrating Treg cells, and this fact should be taken into account in interpreting clinical results with checkpoint inhibitors. Indeed, it is likely that assessment of the expression of checkpoints and of their ligands on the various subsets of tumor infiltrating lymphocytes will help to elucidate conflicting results and provide the rationale for combination therapies. Therefore, expression pattern of checkpoints should be evaluated both in tumor-infiltrating lymphocytes and in tumor cells.

Single-cell analysis on selected tumor Treg signature genes confirmed the whole transcriptomic data and provided information on the expression frequency of these genes. Tumor-infiltrating Treg cells express with high frequency genes that are associated with increased suppressor activity, such as the well characterized *OX40*, *CTLA4*, and *GITR*. Moreover, there were a number of interesting and less expected genes the specific expression of which was validated also at the protein level. For example, IL-1R2 upregulation could be another mechanism that tumor resident Treg cells employ to dampen anti-tumor immune responses through the neutralization of IL-1β function on effector cells. PD-L1 and PD-L2 expression has been recently reported on activated T cells or APCs (Boussiotis et al., 2014; Lesterhuis et al., 2011; Messal et al., 2011) but, to the best of our knowledge, neither PD-L2 nor PD-L1 expression has ever been reported in Treg cells, and our finding that they are overexpressed in tumor infiltrating Treg cells adds an additional level of complexity to the PD1 - PD-Ls immunomodulatory axis within the tumor microenvironment. BATF is a transcription factor that has been mainly associated to Th17 development and $CD8^+$ T cells differentiation (Murphy et al., 2013). Our findings revealed that *BATF* transcript is upregulated in tumor-infiltrating Treg cells more than in tumor infiltrating Th17 cells (Figure S4). Expression of BATF in $CD8^+$ T cells is induced by IL-21 (Xin et al., 2015), and we found that IL21R is highly expressed in tumor-infiltrating Treg cells (Figure 4).

We showed that tumor-infiltrating Treg cells express high amounts of 4-1BB (CD137) a marker of TcR-mediated activation (Schoenbrunn et al., 2012) and have shown they display very high suppressor function on effector T cell proliferation. It could be that expression of the signature genes correlated with the enhanced suppressive ability and so contributed to the establishment of a strong immunosuppressive environment at tumor sites.

A corollary to our findings would have that increased number of Treg cells in the tumor environment should associate with a worst clinical outcome. In fact, when *LAYN*, *MAGEH1*, and *CCR8* (which represent three of the most enriched genes in tumor-infiltrating Treg cells) are highly detected in whole-tumor samples there is a significant worsening of the 5-year survival of both CRC and NSCLC patients. Although, the functional roles in Treg cells of LAYN, a transmembrane protein with homology to c-type lectin (Borowsky and Hynes, 1998), and of MAGEH1, a member of the melanoma antigen gene family (Weon and Potts, 2015), are unknown, the high expression of the chemokine receptor CCR8 is instead intriguing. Indeed, CCL18, the ligand of CCR8 (Islam et al., 2013), is highly expressed in different tumors including NSCLC (Chen et al., 2011; Schutyser et al., 2005). The high specificity of CCR8 expression on tumor-infiltrating Treg cells suggests it could be an interesting therapeutic target to inhibit Treg cells trafficking to tumor sites, without disturbing recruitment of other effector T cells that do not express CCR8.

Considerable efforts have been recently put in the development of sophisticated bioinformatics approaches that exploit lymphocyte gene-expression data to understand the immune-modulatory networks at tumor sites, to predict clinical responses to immune-therapies, and to define therapeutic targets (Bindea et al., 2013a; Bindea et al., 2013b; Gentles et al., 2015). The data we present here represent a comprehensive RNA-sequencing analysis performed on tumor-infiltrating human $CD4^+$ Treg, Th1, and Th17 cells. Our findings highlight the relevance of assessing gene-expression patterns of lymphocyte at tumor-sites and suggest that generation of more transcriptomic data of tumor-infiltrating lymphocyte subsets purified from different cancer types might contribute to a better understanding of the dynamics underlying immune modulation in the tumor microenvironment. Moreover, our data represent a resource to generate and validate hypotheses that will increase our knowledge on tumor-infiltrating Treg cell biology and should lead to the identification of therapeutic targets.

## EXPERIMENTAL PROCEDURES

### Human Primary Tissues

Primary human lung or colorectal tumors and non-neoplastic counterparts were obtained from 15 and 14 patients, respectively. Patients' records

**Figure 4. Expression of Tumor-Infiltrating Treg Cells Protein Signatures in CRC and NSCLC Samples**

(A) Representative flow cytometry plots for tumor (purple line) normal (green area) tissue infiltrating Treg cells and peripheral blood Treg cells (blue line) analyzed for the expression of the indicated proteins.

(B) Flow cytometry plots representative of four independent experiments showing suppressive activity of CRC infiltrating Treg cells on proliferation (shown as CFSE dilution) of $CD4^+$ effector T cells. First panel shows the inhibitory effect of Treg cells on the effector T cell proliferation in the presence of an isotype control antibody. The other panels show the inhibitory effect of Treg cells that have been preincubated with anti PD-L1 or PD-L2 antibodies. Percentage of proliferating cells are indicated. The calculated division index is 0.26 in the presence of the control antibody; 0.57 in the presence of anti-PDL-1 and 0.39 in the presence of anti-PDL-2. Data are representative of four independent experiments.

See also Figure S4.

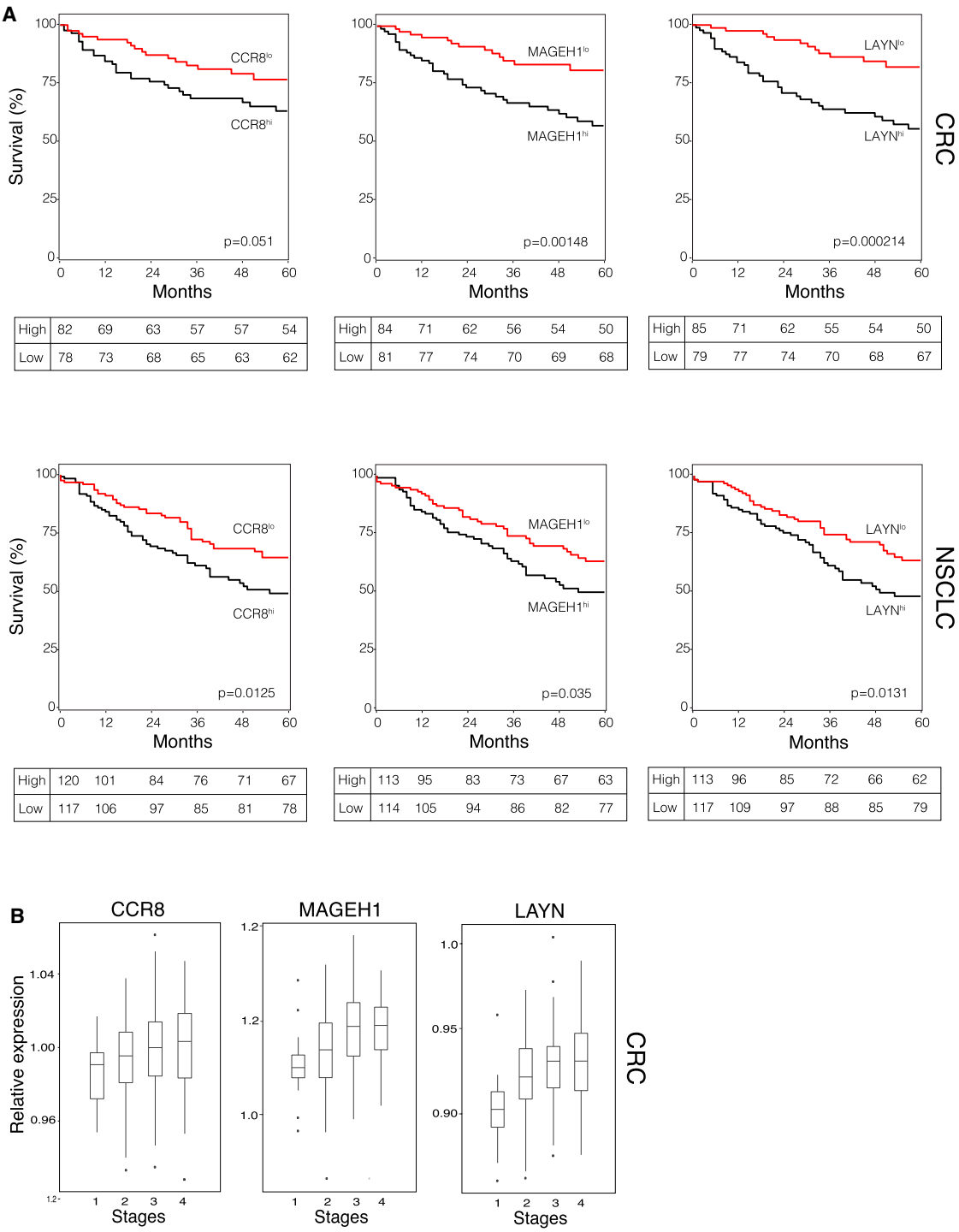



**Figure 5. Prognostic Value of Signature Transcripts of Tumor Infiltrating Treg Cells**

(A) Kaplan-Meier survival curve comparing the high and low expression of the tumor Treg signature transcripts (*CCR8*, *MAGEH1*, *LAYN*) normalized to the *CD3G* for the CRC (n = 177) and NSCLC (n = 263) studies. Univariate analysis confirmed a significant difference in overall survival curve comparing patients with high and low expression. Statistical significance was determined by the log-rank test. (CRC: p = 0.05 for *CCR8*, p = 1.48 × $10^{-3}$ for *MAGEH1*, p = 2.1 × $10^{-4}$ for *LAYN*; NSCLC: p = 0.0125 for *CCR8*, p = 0.035 for *MAGEH1*, p = 0.0131 for *LAYN*.) Each table depicts the Kaplan-Meier estimates at the specified time points.

(B) Expression distributions of *CCR8*, *MAGEH1*, and *LAYN* according to tumor staging at the time of surgery in the cohort of CRC patients.

See also Figure S5.

clinicopathological staging, tumor histotype, and grade are listed in Table S1. Informed consent was obtained from all patients, and the study was approved by the Institutional Review Board of the Fondazione IRCCS Ca' Granda (approval n.30/2014). No patient received palliative surgery or neo-adjuvant chemo- and/or radiotherapy. NSCLC specimens were cut into pieces and single-cell suspensions were prepared by using the Tumor Dissociation Kit, human and the gentleMACS Dissociator (Miltenyi Biotech cat. 130-095-929). Cell suspensions were than isolated by ficoll-hypaque density-gradient centrifugation (Amersham Bioscience). CRC specimens were cut into pieces, incubated in 1 mM EDTA (Sigma-Aldrich) for 50 min at 37°C, and then incubated in type D collagenase solution 0.5 mg/mL (Roche Diagnostic) for 4 hr at 37°C. T cell fractions were recovered after fractionation on a four-step gradient consisting of 100%, 60%, 40%, and 30% Percoll solutions (Pharmacia). See also Supplemental Experimental Procedures.

$CD4^+$ T cell subsets were purified by flow cytomtery sorting using the following fluorochrome conjugated antibodies: anti-CD4 APC/Cy7 (clone OKT4), anti-CD27 Pacific Blue (clone M-T271), anti-IL7R PE (clone MB15-18C9), anti-CD25 PE/Cy7 (clone BC96), anti-CXCR3 PE/Cy5 (clone 1C6/CXCR3), anti-CCR6 APC (clone G034E3), and anti-CCR5 FITC (clone j418F1) using a FACSAria II (BD).

### RNA Isolation and RNA Sequencing

RNA from tumor-infiltrating lymphocytes was isolated using mirVana Isolation Kit. Libraries for Illumina sequencing were constructed from 50 ng of total RNA with the Illumina TruSeq RNA Sample Preparation Kit v2. Paired-end sequencing (2 × 125) was then performed on an Illumina HiSeq 2500. See also Supplemental Experimental Procedures.

### RNA-Seq Data Analysis, Mapping, and Quantification

Raw.fastq files were analyzed using FastQC v0.11.3, and adaptor removal was performed using cutadapt 1.8. Trimming was performed on raw reads using Trimmomatic: standard parameters for phred33 encoding were used. Reads mapping to the reference genome (GRCh38) was performed on quality-checked and trimmed reads using STAR 2.4.1c. The reference annotation is Ensembl v80. The overlap of reads with annotation features found in the reference.gtf was calculated using HT-seq v0.6.1. The output computed for each sample (raw read counts) was then used as input for DESeq2 analysis. Raw counts were normalized using DESeq2's function "rlog," and normalized counts were used to perform and visualize principal component analysis (PCA) results (using DESeq2's "plotPCA" function). See also Supplemental Experimental Procedures.

### Differential Expression Analysis

Differential expression analyses of tumor-infiltrating $CD4^+$ Treg, Th1, and Th17 subsets versus $CD4^+$ Treg, Th1, and Th17 from PBMC were performed using DESeq2. Regulated genes were selected for subsequent analyses if their expression values were found to exceed the threshold of 0.05 FDR (Benjamini-Hochberg correction).

### SOM Analysis

SOM analyses were carried out using the R package oposSOM using default parameters. Expression values of genes selected in the previous differential expression step were *Z*-score normalized and supplied in input to the automated pipeline for SOM training and analysis. Genes from regulated spots in the bidimensional output space were selected according to FDR threshold (< 0.1) at group-level. Expression values of genes assigned to regulated spots extracted from the oposSOM output were subject to correlation analysis using model vectors to further refine the results and genes having expression profiles with $p < 0.05$ were discarded from further analysis and signature definition. See also Supplemental Experimental Procedures.

### GO Analysis

A GO enrichment analysis was performed for biological process terms associated with genes assigned to upregulated spots in the SOM bidimensional space using DAVID. Adjusted p (< 0.05) has been used for terms ranking and selection.

### Capturing of Single Cells, cDNA Preparation, and Single-Cell PCR

Treg cells from CRC and NSCLC were isolated as previously described (see also Table S1). Single cells were captured on a microfluidic chip on the C1 System (Fluidigm) and whole-transcriptome amplified cDNA was prepared on chip using the SMARTer Ultra Low RNA kit (Clontech). For qPCR experiments, harvested cDNA from single cells was pre-amplified using the same pool of TaqMan gene expression assays to be used for qPCR. Single-cell gene expression experiments were performed using the 96 × 96 quantitative PCR (qPCR) DynamicArray microfluidic chips (Fluidigm) on a BioMark real-time PCR reader following manufacturer's instructions. A list of the 78 TaqMan assays used in this study is provided in Supplemental Experimental Procedures.

### Single-Cell Data Analysis

Raw Ct data have been converted to Log2Exp. Co-expression analysis has been performed by considering both CRC and NSCLC samples and genes for which co-expression with *FOXP3* and *IL2RA* was null were discarded for the subsequent analysis. Gene expression was depicted as violin plots after log2 scale transformation. The violin color gradient represents the percentage of cells that are expressing the gene of interest. A non-parametric test (Mann-Whitney $p < 0.05$) has been performed on the selected genes by comparing tumor versus peripheral blood samples (see also Supplemental Experimental Procedures).

### Flow Cytometry Analysis

Surface markers were directly stained with the following fluorochrome-conjugated antibodies and analyzed by flow cytometry: anti-CD4 (OKT4), anti-PD1-LG2 (CL24F.10C12), anti-CD127 (clone RDR5), anti-CD25 (clone 4E3), anti-4-1BB (clone 4B4), anti-CCR8 (Biolegend clone L263G8), anti CD30 (eBioscience, clone Ber-H2), anti-PD-L1 (Biolegend clone 29E.2A3), anti-TIGIT (eBioscience, clone MBSA43), anti-IL1R2 (R and D clone 34141), IL21R (Biolegend clone 2G1-K12), and anti-OX40 (Biolegend clone Ber-ACT35). FOXP3 and BATF intracellular staining was performed with anti-FOXP3 antibody (clone 236A/E7), anti-BATF (clone MBM7C7), and expression analyzed by flow cytometry. See also Supplemental Experimental Procedures.

### Suppression Assay

(CFSE)-labeled responders $CD4^+$ $Naive^+$ T cells from healthy donors were cocultured with different effector to target (E/T) ratios with unlabeled $CD127^-CD25^{low}CD4^+$ T cells sorted from TILs or PBMCs of patients with CRC or NSCLC, using FACS Aria II (BD Biosciences), in the presence of $CD11c^+$ $CD1c^+$dentritic cells as antigen-presenting cells and anti-CD3 (OKT3) mAb. Proliferation of CFSE-labeled cells was assessed after 96 hr. Some suppression assays were also performed with tumor Treg cells that were preincubated with the following antibodies (at a final concentration of 20 μg/ml): anti-human PD-L1 (Biolegend clone 29E.2 A 3), anti-human PD-L2 (Biolegend clone MIH18), and anti-human Functional Grade as isotype control (eBioscience clone MBSA43).

### Kaplan-Meier Analysis

The Kaplan-Meier analysis (KM) was used to compare the high and low expression of the tumor-Treg signature transcripts either CRC (GSE17536, n = 177) and NSCLC (GSE41271, n = 263) patients. See also Supplemental Experimental Procedures.

## ACCESSION NUMBERS

The accession numbers for the data in this paper are as follows: ENA: PRJEB11844 for RNA-seq tumor and tissue infiltrating lymphocytes; ArrayExpress: E-MTAB-2319 for RNA-seq human lymphocytes datasets; ArrayExpress: E-MTAB-513 for Illumina Human BodyMap 2.0 project; GEO: GSE50760 for RNA-seq datasets CRC; GEO: GSE40419 for RNA-seq datasets NSCLC; GEO: GSE17536 for CRC expression profiling by array; and GEO: GSE41271 for NSCLC expression profiling by array.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, six tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.immuni.2016.10.021.

## AUTHOR CONTRIBUTION

M.D., A.A., G.R., and P.G. designed and performed the main experiments, analyzed the data, and contributed to the preparation of the manuscript. V.R. and R.J.P.B. set up all the bioinformatics pipelines, performed the bioinformatics analyses, and contributed to the preparation of the manuscript. S.M., M.M., M.C., E.P., C.P., M.L.S., I.P., and V.V. performed experiments and analyzed the data. S.B., V.V., N.Z., and G.B. coordinated pathology analyses. A.P., L.S., M.T., N.M., P.C.A., O.E., and L.G. coordinated clinical contributions. R.A.P., G.C., R.D.F., H.G.S., and J.G. discussed results, provided advice, and commented on the manuscript. M.D., A.A., G.R., S.A., and M.P. wrote the manuscript. S.A. and M.P. designed the study and supervised research. All authors discussed and interpreted the results.

## ACKNOWLEDGMENTS

We would like to thank S. Biffo and P. Della Bona (San Raffaele Scientific Institute, Milan, Italy) for discussions and critical revision of the manuscript. This study was supported by: the Italian Minister of Health GR2011-02351626 to V.V.; AIRC grant n° IG2015-ID17448 to J.G.; the Italian Minister of University and Research Grant CTN01_00177_817708 "DNA on Disk" to S.B.; the ERC Advanced Grant n° 269022 to S.A.; the Flagship CNR-MIUR grant "EPIGEN," CARIPLO grant n° 2013-0955, AIRC grant n° IG2013-ID14596 and ERC Consolidator Grant n° 617978 to M.P.; and by an unrestricted grant of the "Fondazione Romeo ed Enrica Invernizzi."

Received: March 3, 2016
Revised: September 7, 2016
Accepted: October 4, 2016
Published: November 15, 2016

## REFERENCES

Arpaia, N., Green, J.A., Moltedo, B., Arvey, A., Hemmers, S., Yuan, S., Treuting, P.M., and Rudensky, A.Y. (2015). A Distinct Function of Regulatory T Cells in Tissue Protection. Cell *162*, 1078–1089.

Bindea, G., Galon, J., and Mlecnik, B. (2013a). CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics *29*, 661–663.

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A.C., Angell, H., Fredriksen, T., Lafontaine, L., Berger, A., et al. (2013b). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity *39*, 782–795.

Borowsky, M.L., and Hynes, R.O. (1998). Layilin, a novel talin-binding transmembrane protein homologous with C-type lectins, is localized in membrane ruffles. J. Cell Biol. *143*, 429–442.

Boussiotis, V.A., Chatterjee, P., and Li, L. (2014). Biochemical signaling of PD-1 on T cells and its functional implications. Cancer J. *20*, 265–271.

Burzyn, D., Kuswanto, W., Kolodin, D., Shadrach, J.L., Cerletti, M., Jang, Y., Sefik, E., Tan, T.G., Wagers, A.J., Benoist, C., and Mathis, D. (2013). A special population of regulatory T cells potentiates muscle repair. Cell *155*, 1282–1295.

Campbell, D.J., and Koch, M.A. (2011). Phenotypical and functional specialization of FOXP3+ regulatory T cells. Nat. Rev. Immunol. *11*, 119–130.

Carthon, B.C., Wolchok, J.D., Yuan, J., Kamat, A., Ng Tang, D.S., Sun, J., Ku, G., Troncoso, P., Logothetis, C.J., Allison, J.P., and Sharma, P. (2010). Preoperative CTLA-4 blockade: tolerability and immune monitoring in the setting of a presurgical clinical trial. Clin. Cancer Res. *16*, 2861–2871.

Chen, J., Yao, Y., Gong, C., Yu, F., Su, S., Chen, J., Liu, B., Deng, H., Wang, F., Lin, L., et al. (2011). CCL18 from tumor-associated macrophages promotes breast cancer metastasis via PITPNM3. Cancer Cell *19*, 541–555.

Cipolletta, D., Feuerer, M., Li, A., Kamei, N., Lee, J., Shoelson, S.E., Benoist, C., and Mathis, D. (2012). PPAR-γ is a major driver of the accumulation and phenotype of adipose tissue Treg cells. Nature *486*, 549–553.

Duhen, T., Duhen, R., Lanzavecchia, A., Sallusto, F., and Campbell, D.J. (2012). Functionally distinct subsets of human FOXP3+ Treg cells that phenotypically mirror effector Th cells. Blood *119*, 4430–4440.

Fridman, W.H., Pagès, F., Sautès-Fridman, C., and Galon, J. (2012). The immune contexture in human tumours: impact on clinical outcome. Nat. Rev. Cancer *12*, 298–306.

Galluzzi, L., Buqué, A., Kepp, O., Zitvogel, L., and Kroemer, G. (2015). Immunological Effects of Conventional Chemotherapy and Targeted Anticancer Agents. Cancer Cell *28*, 690–714.

Geginat, J., Paroni, M., Maglie, S., Alfen, J.S., Kastirr, I., Gruarin, P., De Simone, M., Pagani, M., and Abrignani, S. (2014). Plasticity of human CD4 T cell subsets. Front. Immunol. *5*, 630.

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

Gonzalez-Pons, M., and Cruz-Correa, M. (2015). Colorectal Cancer Biomarkers: Where Are We Now? BioMed Res. Int. *2015*, 149014.

Hodi, F.S., O'Day, S.J., McDermott, D.F., Weber, R.W., Sosman, J.A., Haanen, J.B., Gonzalez, R., Robert, C., Schadendorf, D., Hassel, J.C., et al. (2010). Improved survival with ipilimumab in patients with metastatic melanoma. N. Engl. J. Med. *363*, 711–723.

Islam, S.A., Ling, M.F., Leung, J., Shreffler, W.G., and Luster, A.D. (2013). Identification of human CCR8 as a CCL18 receptor. J. Exp. Med. *210*, 1889–1898.

Jacobs, J., Smits, E., Lardon, F., Pauwels, P., and Deschoolmeester, V. (2015). Immune Checkpoint Modulation in Colorectal Cancer: What's New and What to Expect. J. Immunol. Res. *2015*, 158038.

Jamal-Hanjani, M., Thanopoulou, E., Peggs, K.S., Quezada, S.A., and Swanton, C. (2013). Tumour heterogeneity and immune-modulation. Curr. Opin. Pharmacol. *13*, 497–503.

Joller, N., Lozano, E., Burkett, P.R., Patel, B., Xiao, S., Zhu, C., Xia, J., Tan, T.G., Sefik, E., Yajnik, V., et al. (2014). Treg cells expressing the coinhibitory molecule TIGIT selectively inhibit proinflammatory Th1 and Th17 cell responses. Immunity *40*, 569–581.

Josefowicz, S.Z., Lu, L.F., and Rudensky, A.Y. (2012). Regulatory T cells: mechanisms of differentiation and function. Annu. Rev. Immunol. *30*, 531–564.

Kharchenko, P.V., Silberstein, L., and Scadden, D.T. (2014). Bayesian approach to single-cell differential expression analysis. Nat. Methods *11*, 740–742.

Kroemer, G., Galluzzi, L., Zitvogel, L., and Fridman, W.H. (2015). Colorectal cancer: the first neoplasia found to be under immunosurveillance and the last one to respond to immunotherapy? OncoImmunology *4*, e1058597.

Le, D.T., Uram, J.N., Wang, H., Bartlett, B.R., Kemberling, H., Eyring, A.D., Skora, A.D., Luber, B.S., Azad, N.S., Laheru, D., et al. (2015). PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. N. Engl. J. Med. *372*, 2509–2520.

Lesterhuis, W.J., Steer, H., and Lake, R.A. (2011). PD-L2 is predominantly expressed by Th2 cells. Mol. Immunol. *49*, 1–3.

Marabelle, A., Kohrt, H., Sagiv-Barfi, I., Ajami, B., Axtell, R.C., Zhou, G., Rajapaksa, R., Green, M.R., Torchia, J., Brody, J., et al. (2013). Depleting tumor-specific Tregs at a single site eradicates disseminated tumors. J. Clin. Invest. *123*, 2447–2463.

Messal, N., Serriari, N.E., Pastor, S., Nunès, J.A., and Olive, D. (2011). PD-L2 is expressed on activated human T cells and regulates their function. Mol. Immunol. *48*, 2214–2219.

Munn, D.H., and Bronte, V. (2016). Immune suppressive mechanisms in the tumor microenvironment. Curr. Opin. Immunol. *39*, 1–6.

Murphy, T.L., Tussiwand, R., and Murphy, K.M. (2013). Specificity through cooperation: BATF-IRF interactions control immune-regulatory networks. Nat. Rev. Immunol. *13*, 499–509.

Nishikawa, H., and Sakaguchi, S. (2010). Regulatory T cells in tumor immunity. Int. J. Cancer *127*, 759–767.

Panduro, M., Benoist, C., and Mathis, D. (2016). Tissue Tregs. Annu. Rev. Immunol. *34*, 609–633.

Pardoll, D.M. (2012). The blockade of immune checkpoints in cancer immunotherapy. Nat. Rev. Cancer *12*, 252–264.

Peggs, K.S., Quezada, S.A., Chambers, C.A., Korman, A.J., and Allison, J.P. (2009). Blockade of CTLA-4 on both effector and regulatory T cell compartments contributes to the antitumor activity of anti-CTLA-4 antibodies. J. Exp. Med. *206*, 1717–1725.

Ranzani, V., Rossetti, G., Panzeri, I., Arrigoni, A., Bonnal, R.J., Curti, S., Gruarin, P., Provasi, E., Sugliano, E., Marconi, M., et al. (2015). The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. Nat. Immunol. *16*, 318–325.

Sakaguchi, S., Yamaguchi, T., Nomura, T., and Ono, M. (2008). Regulatory T cells and immune tolerance. Cell *133*, 775–787.

Sato, M., Larsen, J.E., Lee, W., Sun, H., Shames, D.S., Dalvi, M.P., Ramirez, R.D., Tang, H., DiMaio, J.M., Gao, B., et al. (2013). Human lung epithelial cells progressed to malignancy through specific oncogenic manipulations. Mol. Cancer Res. *11*, 638–650.

Schoenbrunn, A., Frentsch, M., Kohler, S., Keye, J., Dooms, H., Moewes, B., Dong, J., Loddenkemper, C., Sieper, J., Wu, P., et al. (2012). A converse 4-1BB and CD40 ligand expression pattern delineates activated regulatory T cells (Treg) and conventional T cells enabling direct isolation of alloantigen-reactive natural Foxp3+ Treg. J. Immunol. *189*, 5985–5994.

Schutyser, E., Richmond, A., and Van Damme, J. (2005). Involvement of CC chemokine ligand 18 (CCL18) in normal and pathological processes. J. Leukoc. Biol. *78*, 14–26.

Selby, M.J., Engelhardt, J.J., Quigley, M., Henning, K.A., Chen, T., Srinivasan, M., and Korman, A.J. (2013). Anti-CTLA-4 antibodies of IgG2a isotype enhance antitumor activity through reduction of intratumoral regulatory T cells. Cancer Immunol. Res. *1*, 32–42.

Sharma, P., and Allison, J.P. (2015). Immune checkpoint targeting in cancer therapy: toward combination strategies with curative potential. Cell *161*, 205–214.

Simpson, T.R., Li, F., Montalvo-Ortiz, W., Sepulveda, M.A., Bergerhoff, K., Arce, F., Roddie, C., Henry, J.Y., Yagita, H., Wolchok, J.D., et al. (2013). Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma. J. Exp. Med. *210*, 1695–1710.

Śledzińska, A., Menger, L., Bergerhoff, K., Peggs, K.S., and Quezada, S.A. (2015). Negative immune checkpoints on T lymphocytes and their relevance to cancer immunotherapy. Mol. Oncol. *9*, 1936–1965.

Smith, J.J., Deane, N.G., Wu, F., Merchant, N.B., Zhang, B., Jiang, A., Lu, P., Johnson, J.C., Schmidt, C., Bailey, C.E., et al. (2010). Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. Gastroenterology *138*, 958–968.

Teng, M.W., Ngiow, S.F., von Scheidt, B., McLaughlin, N., Sparwasser, T., and Smyth, M.J. (2010). Conditional regulatory T-cell depletion releases adaptive immunity preventing carcinogenesis and suppressing established tumor growth. Cancer Res. *70*, 7800–7809.

Topalian, S.L., Drake, C.G., and Pardoll, D.M. (2015). Immune checkpoint blockade: a common denominator approach to cancer therapy. Cancer Cell *27*, 450–461.

Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. CA Cancer J. Clin. *65*, 87–108.

Twyman-Saint Victor, C., Rech, A.J., Maity, A., Rengan, R., Pauken, K.E., Stelekati, E., Benci, J.L., Xu, B., Dada, H., Odorizzi, P.M., et al. (2015). Radiation and dual checkpoint blockade activate non-redundant immune mechanisms in cancer. Nature *520*, 373–377.

van den Eertwegh, A.J., Versluis, J., van den Berg, H.P., Santegoets, S.J., van Moorselaar, R.J., van der Sluis, T.M., Gall, H.E., Harding, T.C., Jooss, K., Lowy, I., et al. (2012). Combined immunotherapy with granulocyte-macrophage colony-stimulating factor-transduced allogeneic prostate cancer cells and ipilimumab in patients with metastatic castration-resistant prostate cancer: a phase 1 dose-escalation trial. Lancet Oncol. *13*, 509–517.

Voo, K.S., Bover, L., Harline, M.L., Vien, L.T., Facchinetti, V., Arima, K., Kwak, L.W., and Liu, Y.J. (2013). Antibodies targeting human OX40 expand effector T cells and block inducible and natural regulatory T cell function. J. Immunol. *191*, 3641–3650.

Walter, S., Weinschenk, T., Stenzl, A., Zdrojowy, R., Pluzanska, A., Szczylik, C., Staehler, M., Brugger, W., Dietrich, P.Y., Mendrzyk, R., et al. (2012). Multipeptide immune response to cancer vaccine IMA901 after single-dose cyclophosphamide associates with longer patient survival. Nat. Med. *18*, 1254–1261.

Weon, J.L., and Potts, P.R. (2015). The MAGE protein family and cancer. Curr. Opin. Cell Biol. *37*, 1–8.

Wirth, H., von Bergen, M., and Binder, H. (2012). Mining SOM expression portraits: feature selection and integrating concepts of molecular function. BioData Min. *5*, 18.

Xin, G., Schauder, D.M., Lainez, B., Weinstein, J.S., Dai, Z., Chen, Y., Esplugues, E., Wen, R., Wang, D., Parish, I.A., et al. (2015). A Critical Role of IL-21-Induced BATF in Sustaining CD8-T-Cell-Mediated Chronic Viral Control. Cell Rep. *13*, 1118–1124.

Yang, J.C., Hughes, M., Kammula, U., Royal, R., Sherry, R.M., Topalian, S.L., Suri, K.B., Levy, C., Allen, T., Mavroukakis, S., et al. (2007). Ipilimumab (anti-CTLA4 antibody) causes regression of metastatic renal cell cancer associated with enteritis and hypophysitis. J. Immunother. *30*, 825–830.

Zitvogel, L., Galluzzi, L., Smyth, M.J., and Kroemer, G. (2013). Mechanism of action of conventional and targeted anticancer therapies: reinstating immunosurveillance. Immunity *39*, 74–88.

**Immunity, Volume *45***

## Supplemental Information

# Transcriptional Landscape of Human Tissue Lymphocytes Unveils Uniqueness of Tumor-Infiltrating T Regulatory Cells

**Marco De Simone, Alberto Arrigoni, Grazisa Rossetti, Paola Gruarin, Valeria Ranzani, Claudia Politano, Raoul J.P. Bonnal, Elena Provasi, Maria Lucia Sarnicola, Ilaria Panzeri, Monica Moro, Mariacristina Crosti, Saveria Mazzara, Valentina Vaira, Silvano Bosari, Alessandro Palleschi, Luigi Santambrogio, Giorgio Bovo, Nicola Zucchini, Mauro Totis, Luca Gianotti, Giancarlo Cesana, Roberto A. Perego, Nirvana Maroni, Andrea Pisani Ceretti, Enrico Opocher, Raffaele De Francesco, Jens Geginat, Hendrik G. Stunnenberg, Sergio Abrignani, and Massimiliano Pagani**

Figure S1

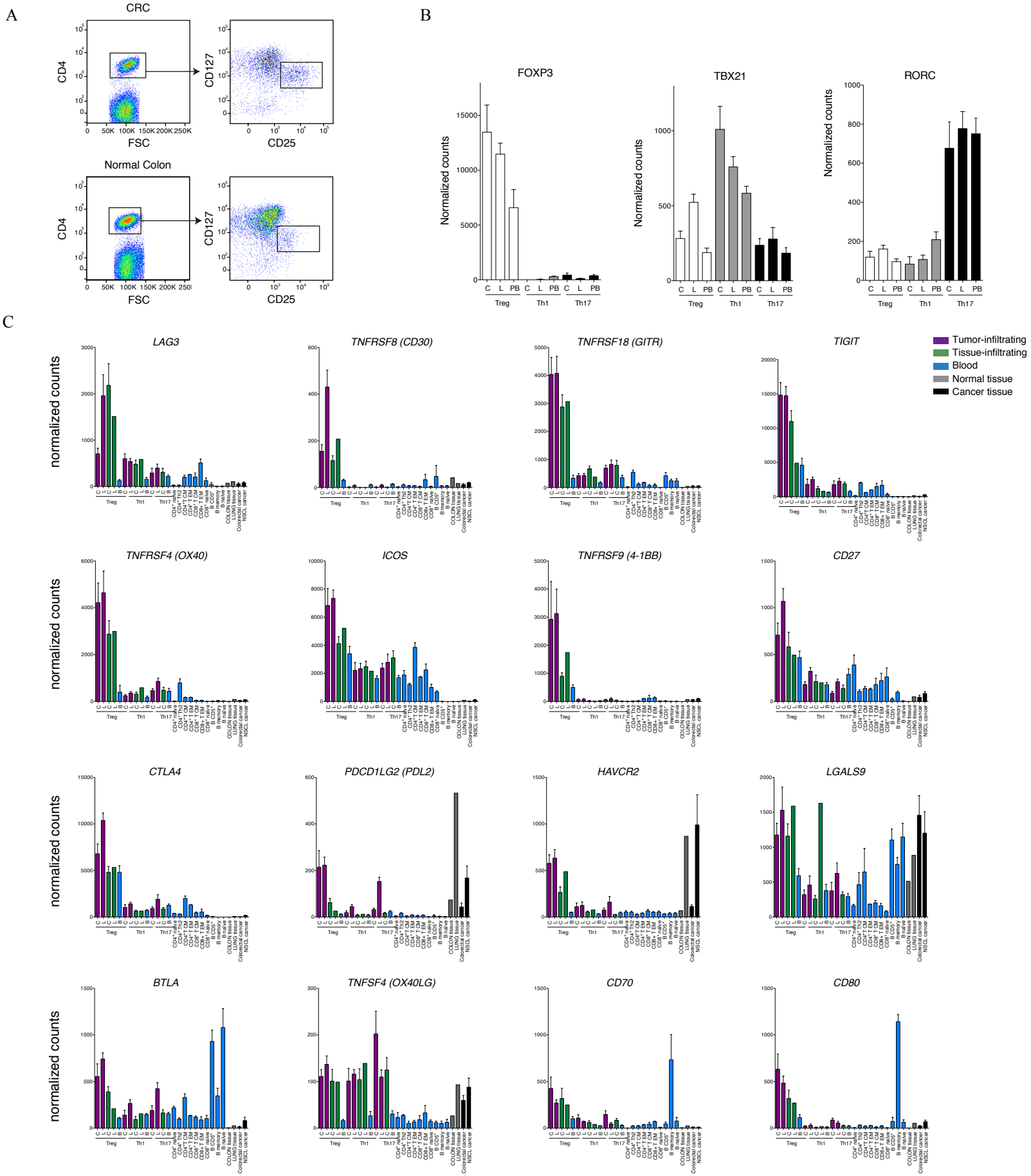



**Figure S1 related to Figure 1. Transcriptome analysis of tumor infiltrating lymphocytes**

(A) Representation of the sorting strategy of Treg cells infiltrating colorectal tumor or normal tissue.

(B) RNA-seq expression values (normalized counts) of FOXP3, TBX21 and RORC in CD4+ Th1, Th17 and Treg cells from CRC (C), NSCLC (L) or peripheral blood (PB) of healthy donors.

(C) RNA-seq normalized counts data for selected immune checkpoints and their ligands are shown as histogram plot. Cell population names are reported in the lower part of each graph, while gene names are shown in the upper part. To distinguish the origin of the different populations a color code has been assigned (upper right part of the figure).

Figure S2

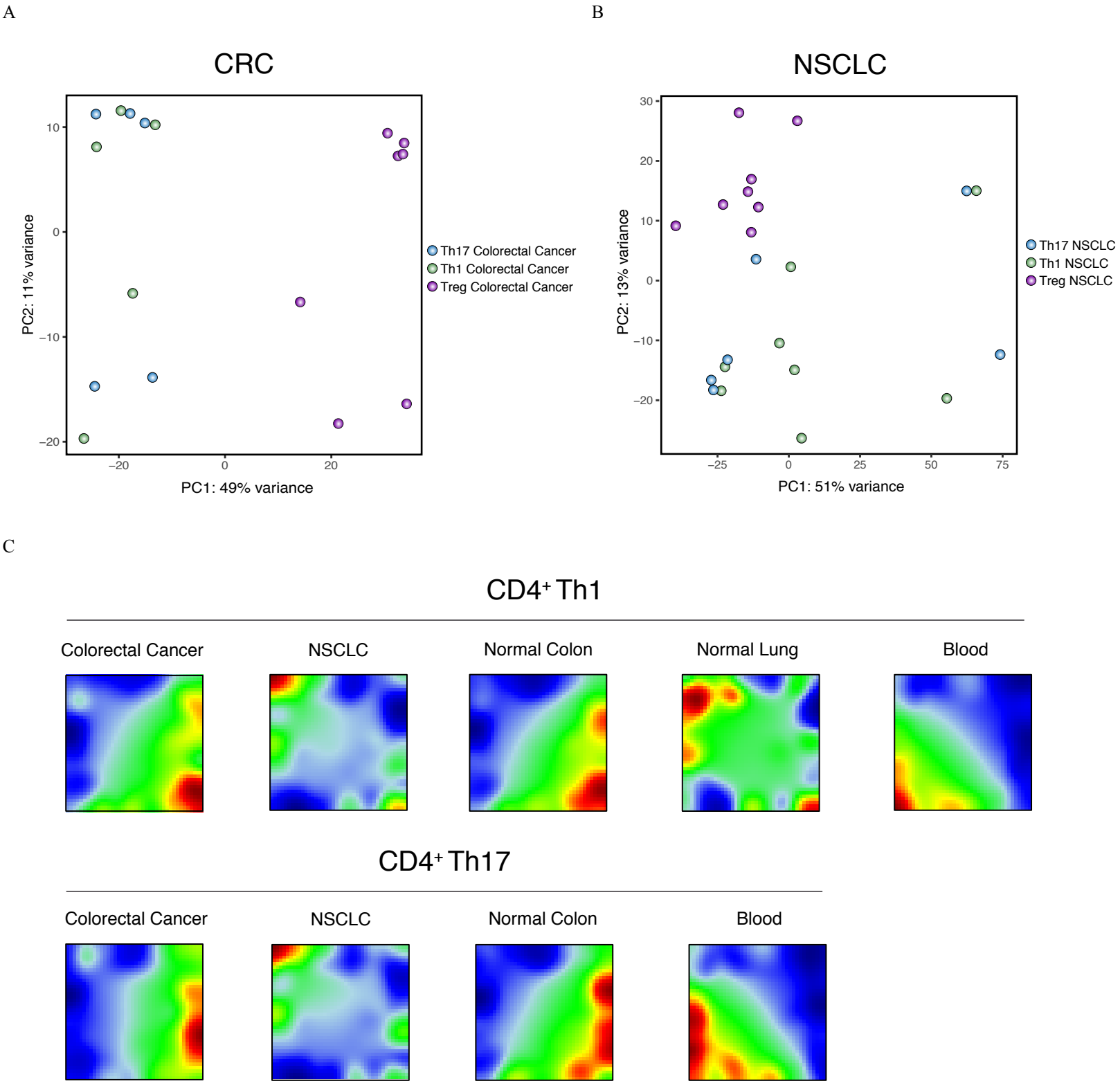
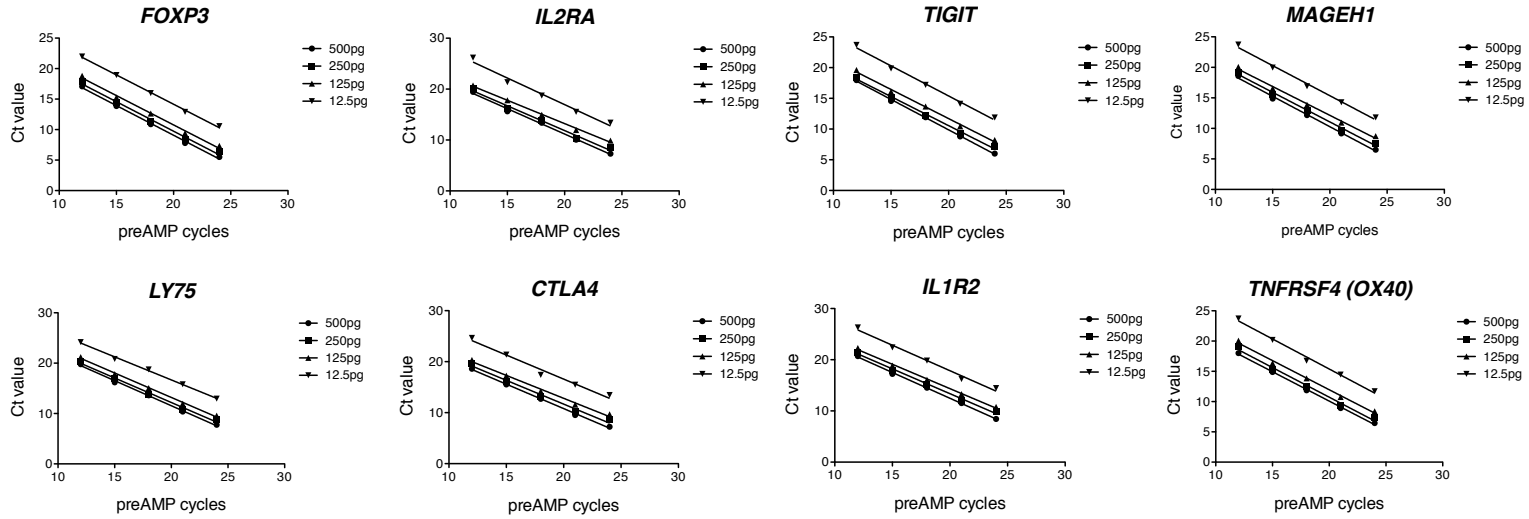



**Figure S2 related to Figure 2. SOM analysis identifies co-regulated genes in tumor infiltrating Treg cells.**
(A-B) Principal Component Analysis (PCA) has been performed on rlog-normalized (DESeq2) counts for tumor infiltrating CD4+ Treg, Th1 and Th17 cells RNA-seq data in CRC (A) and NSCLC (B) samples.
(C) Self-Organizing Maps analysis has been performed on the RNA-seq dataset comprising Th1and Th17 cell subsets. Bidimensional sample-level SOM profiles for different tissues are reported.

Figure S3

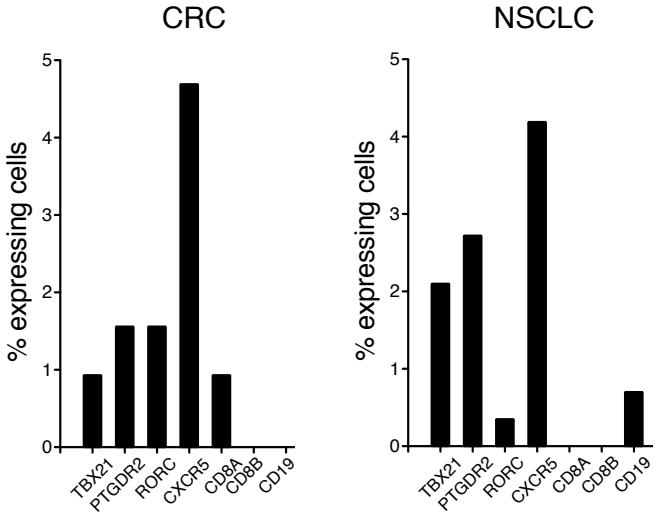
A




B

| $R^2$ | 500pg | 250pg | 125pg | 12.5pg |
|---|---|---|---|---|
| ***FOXP3*** | 0.9967 | 0.9906 | 0.9933 | 0.9978 |
| ***IL2RA*** | 0.9951 | 0.9879 | 0.9966 | 0.9792 |
| ***TIGIT*** | 0.9987 | 0.9948 | 0.9949 | 0.992 |
| ***MAGEH1*** | 0.9965 | 0.9946 | 0.9941 | 0.9922 |
| ***LY75*** | 0.997 | 0.9917 | 0.997 | 0.9963 |
| ***CTLA4*** | 0.9977 | 0.9852 | 0.9918 | 0.9775 |
| ***IL1R2*** | 0.9987 | 0.993 | 0.9969 | 0.987 |
| ***TNFRSF4 (OX40)*** | 0.9982 | 0.9899 | 0.994 | 0.9925 |

C




| CELL TYPE | MARKERS |
|---|---|
| $CD4^+$ Th1 | TBX21 |
| $CD4^+$ Th2 | PTGDR2 |
| $CD4^+$ Th17 | RORC |
| $CD4^+$ Tfh | CXCR5 |
| $CD8^+$ cells | CD8A, CD8B |
| B cells | CD19 |

**Figure S3 related to Figure 3. Single-cell analysis of tumor infiltrating Treg cells.**

(A) Pre-amplification efficiency assessment. 500 pg, 250 pg, 125 pg or 12.5 pg of cDNA (total RNA equivalent), were pre-amplified for 12, 15, 18, 21 or 24 cycles and used as template in individual qPCRs for each gene. For each template condition Ct for each amount of pre-amplified cDNA (Y-axis) was plotted against the number of pre-amplification cycles performed (X-axis). A subset of 8 probes out of 79 is shown.

(B) Fitness of the linear correlation between amplification cycles and Ct values for each template condition was assessed and confirmed by R2.

(C) Assessment of CD4+ Treg, Th1, Th17, Th2, CD8+ T cells and B cell markers expression (percentage of expressing cells) in single Treg cells purified from NSCLC and CRC

Figure S4

A

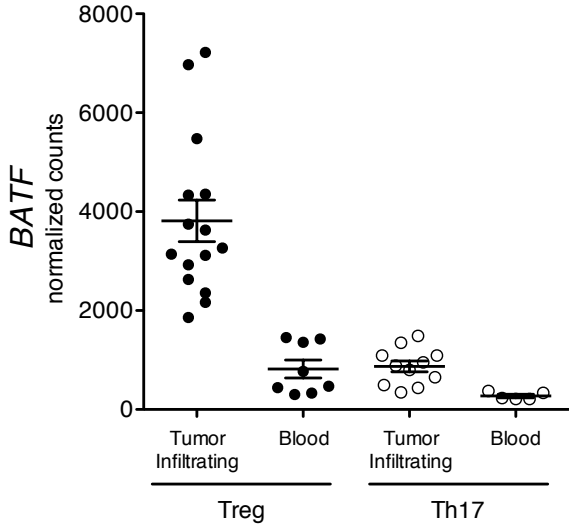



**Figure S4 related to Figure 4. Comparison of BATF expression in CD4+ Treg vs Th17 cells.**
BATF expression levels (RNA-seq normalized counts data) in CD4+ Treg and Th17 subsets isolated from tumor tissue or peripheral blood

Figure S5

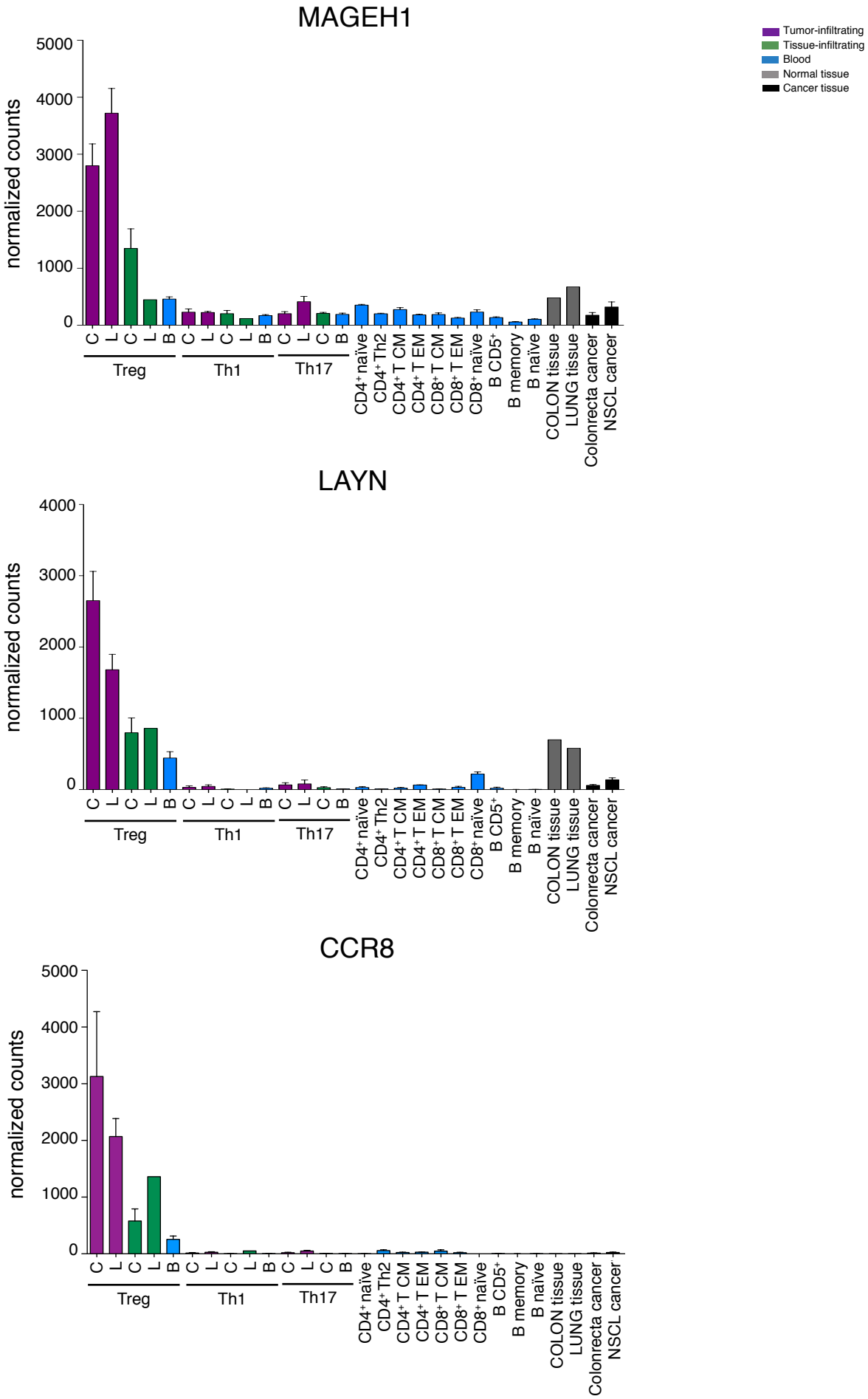



**Figure S5 related to Figure 5. Expression levels of tumour-infiltrating Treg signature genes.**
RNA-seq normalized counts data of three tumour-infiltrating Treg signature genes (MAGEH1, LAYN and CCR8) across listed cell populations. Cell populations are reported as a color code in the upper part of the figure.

**SUPPLEMENTAL TABLES**

**Table S1 related to Table 1. Patients' information and histological analysis**
For each cell subset profiled by RNA-sequencing, patient records are shown including: age at diagnosis, gender, smoking habit (for lung cancer patients), clinicopathological staging (TNM classification) tumor histotype and grade. For Treg cell isolated for qPCR experiment the same information are available, including also the number of live cells captured from each tumor and available for single-cell analysis. CRC: colorectal cancer; NSCLC: non-small cell lung cancer; (T): Tumor Sample; (H): Healthy Tissue; ADC: Adenocarcinoma; SCC: Squamous Cell Carcinoma; MUC ADC: Mucinous Adenocarcinoma.

**Table S2 related to Figure 2. Co-regulated genes in tumor-infiltrating CD4+ Treg cells**
A list of the co-regulated genes in tumor-infiltrating CD4+ Treg cells is reported.

**Table S3 related to Figure 2. Tumor-infiltrating Treg cell Signature: GO enrichment**
All the enriched GO terms (DAVID) for genes assigned to Treg cells regulated spots are reported with corresponding significance p-values.

**Table S4 related to Figure 2. Expression levels of tumor-infiltrating Treg gene signatures in all the subsets analyzed**
Normalized expression values of tumor-infiltrating Treg signature genes across listed cell populations. Cell populations are reported as a color code in the lower part of the table.

**Table S5 related to Figure 1. Expression levels of immune checkpoints genes in all the subsets analyzed.**
RNA-seq normalized counts data for selected immune checkpoints genes and their ligands in all the subsets analyzed. Color code for cell populations is reported in the lower part of the table.

**Supplemental Table S6 Related to Figure 3**.
List of TaqMan Probes and assay number used in RT-qPCR single-cell experiments

| Taqman Assays Numbers | | | |
|---|---|---|---|
| Gene Name | Assay Number | Gene Name | Assay Number |
| BCL2L1 | Hs00236329_m1 | ACP5 | Hs00356261_m1 |
| EOS | Hs00223842_m1 | BATF | Hs00232390_m1 |
| AHCYL1 | Hs00198312_m1 | SLC35F2 | Hs00213850_m1 |
| NFE2L3 | Hs00852569_g1 | LAX1 | Hs00214948_m1 |
| IL12RB2 | Hs00155486_m1 | CCR8 | Hs00174764_m1 |
| CD177 | Hs00360669_m1 | ADPRH | Hs00153890_m1 |
| OX40 | Hs00937194_g1 | IKZF2 | Hs00212361_m1 |
| METTL7A | Hs00204042_m1 | CSF2RB | Hs00166144_m1 |
| ENTPD1 | Hs00969559_m1 | NDFIP2 | Hs00324851_m1 |
| NFAT5 | Hs00232437_m1 | CADM1 | Hs00942508_m1 |
| CTSC | Hs00175188_m1 | ICOS | Hs00359999_m1 |
| SSH1 | Hs00368014_m1 | COL9A2 | Hs00156712_m1 |
| TMEM184C | Hs00217311_m1 | LTA | Hs00236874_m1 |
| HTATIP2 | Hs01091727_m1 | MAGEH1 | Hs00371974_s1 |
| HSDL2 | Hs00953689_m1 | IL21R | Hs00222310_m1 |
| FOXP3 | Hs01085834_m1 | SSTR3 | Hs01066399_m1 |
| IL2RA | Hs00907778_m1 | RNF145 | Hs01099642_m1 |
| LIMA1 | Hs01035646_m1 | LAPTM4B | Hs00363282_m1 |
| NAB1 | Hs00428619_m1 | GRSF1 | Hs00909877_m1 |
| ACSL4 | Hs00244871_m1 | ANKRD10 | Hs00214321_m1 |
| ERI1 | Hs00405251_m1 | NPTN | Hs01033353_m1 |
| FKBP1A | Hs00356621_g1 | HS3ST3B1 | Hs00797512_s1 |
| LEPROT | Hs00956627_s1 | TRAF3 | Hs00936781_m1 |
| NETO2 | Hs00983152_m1 | RRAGB | Hs01099767_m1 |
| VDR | Hs00172113_m1 | ZBTB38 | Hs00257315_s1 |
| CSF1 | Hs00174164_m1 | TIGIT | Hs00545087_m1 |
| GITR | Hs00188346_m1 | TFRC | Hs00951083_m1 |
| IL1R2 | Hs01030384_m1 | JAK1 | Hs01026983_m1 |
| IL1R1 | Hs00991010_m1 | KSR1 | Hs00300134_m1 |
| LAYN | Hs00379511_m1 | ZNF282 | Hs00411965_m1 |
| THADA | Hs00736554_m1 | PTPRJ | Hs01119326_m1 |
| CTLA4 | Hs00175480_m1 | CHRNA6 | Hs02563509_s1 |
| CHST2 | Hs01921028_s1 | IL2RB | Hs01081697_m1 |
| CHST7 | Hs00219871_m1 | TBX21 | Hs00203436_m1 |
| LRBA | Hs01032231_m1 | RORC | Hs01076112_m1 |
| ETV7 | Hs00903229_m1 | CXCR5 | Hs00540548_s1 |
| LY75 | Hs00982383_m1 | CD8A | Hs00233520_m1 |
| ADAT2 | Hs00699339_m1 | CD8B | Hs00174762_m1 |
| GCNT1 | Hs00155243_m1 | PTGDR2 | Hs00173717_m1 |
| CASP1 | Hs00354836_m1 | CD19 | Hs01047410_g1 |

## SUPPLEMENTAL EXPERIMENTAL PROCEDURES

### Human primary tissues

Primary human lung or colorectal tumors and non-neoplastic counterparts were obtained respectively from fifteen and fourteen patients who underwent surgery for therapeutic purposes at Fondazione IRCCS Ca' Granda, Policlinico or San Gerardo Hospitals (Italy). Records were available for all cases and included patients' age at diagnosis, gender, smoking habit (for lung cancer patients), clinicopathological staging (Sobin et al., 2009), tumor histotype and grade (Table S1). No patient received palliative surgery or neoadjuvant chemo- and/or radiotherapy. Informed consent was obtained from all patients, and the study was approved by the Institutional Review Board of the Fondazione IRCCS Ca' Granda (approval n. 30/2014).

Non-small-cell lung cancer (NSCLC) were cut into pieces and single-cell suspensions were prepared by using the Tumor Dissociation Kit, human and the gentleMACS™ Dissociator (Miltenyi Biotech cat. 130-095-929) according to the accompanying standard protocol. Cell suspensions were than isolated by ficoll-hypaque density-gradient centrifugation (Amersham Bioscience). Colorectal cancer (CRC) specimens were cut into pieces and incubated in DTT 0.1 mM (Sigma-Aldrich) for 10 min, then extensively washed in HBSS (Thermo Scientific) and incubated in 1 mM EDTA (Sigma-Aldrich) for 50 min at 37 °C in the presence of 5% CO2. They were then washed and incubated in type D collagenase solution 0.5 mg/mL (Roche Diagnostic) for 4 h at 37°C. Supernatants containing tumor infiltrating lymphocytes were filtered through 100 μm cell strainer, centrifuged and fractionated 1800X g for 30 min at 4°C on a four-step gradient consisting of 100%, 60%, and 40% and 30% Percoll solutions (Pharmacia). The T cell fraction was recovered from the inter- face between the 60% and 40% Percoll layers.

CD4 T cell subsets were purified by FACS sorting using the following fluorochrome conjugated antibodies: anti-CD4 APC/Cy7 (Biolegend clone OKT4), anti-CD27 Pacific Blue (Biolegend, clone M-T271), anti-IL7R PE (Milteniy, clone MB15-18C9), anti-CD25 PE/Cy7 (eBioscience, clone BC96), anti-CXCR3 PE/Cy5 (BD, clone 1C6/CXCR3), anti-CCR6 APC (Biolegend, clone G034E3) and anti-CCR5 FITC (Biolegend, clone j418F1) using a FACSAria II (BD).

### Flow cytometry

To validate surface marker expression cells were directly stained with the following fluorochrome-conjugated antibodies and analyzed by flow cytometry: anti-CD4 (Biolegend, clone OKT4); anti-PD-L2 (Biolegend, Clone CL24F.10C12); anti-CD127 (eBioscience, clone RDR5); anti-BATF (eBioscience, clone MBM7C7), anti-GITR (eBioscience, clone eBIOAITR), anti-CD25 (Miltenyi, clone 4E3) and anti 4-1BB (eBioscience clone 4B4) anti CCR8(Biolegend clone L263G8) anti CD30 (eBioscience, clone Ber-H2) anti PD-L1 (Biolegend clone 29E.2A3) anti TIGIT (eBioscience, clone MBSA43) anti IL1R2 (R and D clone 34141) IL21R (Biolegend clone 2G1-K12) anti OX40 (Biolegend clone Ber-ACT35). Intracellular staining was performed using eBioscience Foxp3 staining kit according to the manufactured's protocol (eBioscience cat 00-5523-00). Briefly cells were harvested and fixed for 30 min in fixation/permeabilization buffer at 4 °C, and than stained with anti-FOXP3 antibody (eBioscience, clone 236A/E7) and anti-BATF (ebioscience clone MBM7C7) in permeabilisation buffer for 30 min at 4 °C. Cells were then washed two times, resuspended in FACS washing buffer and analyzed by flow cytometry.

### Suppression assay.

$4 \times 10^4$ carboxyfluorescein diacetate succinimidyl ester (CFSE)-labeled (1 μM) responders $Naive^+$ T cells from healthy donors were cocultured with different E/T ratio with unlabeled $CD127^-CD25^{low}CD4^+$ T cells sorted from TILs or PBMCs of patients with CRC or NSCLC, using FACS Aria II (BD Biosciences) , in the presence of $CD11c^+CD1c^+$dentritic cells as antigen-presenting cells and 0.5 mg/ml anti-CD3 (OKT3) mAb. Proliferation of CFSE-labeled cells was assessed by flow cytometry after 96 hr culture.

Competition assays with blocking antibodies were performed following the same protocol starting from $2x10^5$ CFSE labeled CD4+ naïve T cells from healthy donors cocultured at 1:32 Responders/Tumor Treg cell ratio. The following antibodies at a final concentration of 20μg/ml were used purified anti-human PDL-1 (biolegend clone 29E.2 A 3); purified anti-human PD-L2 (biolegend clone MIH18); anti-human Functional Grade (ebioscience clone MBSA43)

### RNA isolation and RNA sequencing

RNA from tumor-infiltrating lymphocytes was isolated using mirVana Isolation Kit. Residual contaminating genomic DNA was removed from the total RNA fraction using Turbo DNA-free (Thermo Fisher). The RNA yields were quantified using the QuantiFluor RNA System (Promega) and the RNA quality was assessed by the Agilent 2100 Bioanalyzer (Agilent). Libraries for Illumina sequencing were constructed from 50 ng of

total RNA with the Illumina TruSeq RNA Sample Preparation Kit v2 (Set A). The generated libraries were loaded on to the cBot (Illumina) for clustering on a HiSeq Flow Cell v3. The flow cell was then sequenced using a HiSeq 2500 in High Output mode (Illumina). A paired-end (2×125) run was performed.

**RNA-seq data analysis**
Raw .fastq files were analyzed using FastQC v0.11.3, and adapter removal was performed using cutadapt 1.8. Cutadapt is run both for reverse and forward sequences with default parameters [--anywhere <adapter1> --anywhere <adapter2> --overlap 10 --times 2 --mask-adapter]. Adapter sequences used for libraries preparation are

Adapter1:
AGATCGGAAGAGCACACGTCTGAACTCCAGTCACNNNNNNATCTCGTATGCCGTCTTCTGCTTG

Adapter2:
AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT

Trimming was performed on raw reads using Trimmomatic (Bolger et al., 2014): standard parameters for phred33 encoding were used: ILLUMINACLIP (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15), MINLEN parameter was set to 50.

**Mapping and quantification**: reads mapping to the reference genome (GRCh38) was performed on quality-checked and trimmed reads using STAR 2.4.1c: [STAR --genomeDir <index_star> --runThreadN <cpu_number> --readFilesIn <trimmed>_R1.fastq.gz <trimmed>_R2_P.fastq.gz --readFilesCommand zcat]. The reference annotation is Ensembl v80. The overlap of reads with annotation features found in the reference .gtf was calculated using HT-seq v0.6.1. The output computed for each sample (raw read counts) was then used as input for DESeq2 analysis. Raw counts were normalized using DESeq2's function 'rlog', and normalized counts were used to perform and visualize Principal Component Analysis (PCA) results (using DESeq2's 'plotPCA' function).

**Differential expression analysis**: differential expression analyses of tumor-infiltrating CD4+ Treg/Th1/Th17 subsets vs. CD4+ Treg/Th1/Th17 from PBMC were performed using DESeq2. Upregulated/downregulated genes were selected for subsequent analyses if their expression values were found to exceed the threshold of 0.05 FDR (Benjamini-Hochberg correction).

**SOM (theory)**: SOMs can be thought of as a spatially constrained form of k-means clustering (Ripley, 1996). In this analogy, every unit in the grid corresponds to a cluster to which a certain number of gene expression input vectors are mapped. For each vector during the training phase of the algorithm, the "winning unit" (the one most similar to the current training object) will be updated to become even more similar. A weighted average is used, where the weight of the new object is one of the training parameters of the SOM. The update is not restricted to the winning unit, but it is extended to the units in the immediate neighborhood, so that the structure itself of the map can fit the data. The size of the neighborhood progressively shrinks, so that eventually only the winning units are adapted. In the oposSOM package (Loffler-Wirth et al., 2015), a rectangular SOM topology with a Gaussian neighborhood function is used. Upon training completion, the distance of vectors describing groups of samples (e.g. tumor-infiltrating lymphocytes) is calculated (using the same metric that was used for training) and a map with color-coded distances that is specific for the considered condition is drawn. These expression portraits exhibit characteristic spatial color patterns and serves as fingerprint of the transcriptional activity.

**SOM analysis**: In order to translate RNA-seq expression data into metadata of reduced dimensions and identify genes that are preferentially expressed in tumor-infiltrating lymphocyte subsets, Self-Organizing Maps (SOMs) analysis was performed on our dataset. This method transformed the whole genome expression pattern of about 7,000 differentially expressed genes into a SOM coordinate system, which allowed for intuitive visualization of transcriptional activity of each sample in terms of mosaic portraits. SOM analysis combines strong clustering, dimensionality reduction, multidimensional scaling and visualization capabilities which have been shown to be advantageous compared to alternative methods such as clustering heatmaps and negative matrix factorization when applied to molecular high-throughput data (Wirth et al., 2012). SOM maps constitute fingerprints of the transcriptional activity of the respective cell population sample and allow for direct comparison of the expression of individual samples in a simple and intuitive way. The color gradient is instrumental to visualize over- or underexpression of the nodes for the particular sample compared with the mean expression level of each node in the pool of all samples

studied.Analyses were carried out using the R package oposSOM (Loffler-Wirth et al., 2015) using default parameters. Expression values of genes selected in the previous differential expression step were Z-score normalized and supplied in input to the automated pipeline for SOM training and analysis. Genes from up/downregulated spots in the bidimensional output space were selected according to FDR threshold (<0.1) at group-level. Expression values of genes assigned to regulated spots extracted from the oposSOM output were then subject to correlation analysis using model vectors to further refine the results and genes having expression profiles with P-val<0.05 were discarded from further analysis and signature definition.

**Pearson correlation analyses**: correlation analyses were performed using Pearson correlation metric, and significance p-values were calculated using the *cor.test* function from the WGCNA R package.

**GO analysis**: a GO enrichment analysis was performed for biological process terms associated with genes assigned to up/downregulated spots in the SOM bidimensional space using DAVID (Huang da et al., 2009). Adjusted p-val has been used for terms ranking and selection (<0.05).

**Capturing of single cells, preparation of cDNA and single-cell PCR**
Treg cells from 5 CRC and 5 NSCLC specimens were isolated as previously described (See also Table S1).Single cells were captured on a microfluidic chip on the C1 System (Fluidigm) and whole-transcriptome amplified. cDNA was prepared on chip using the SMARTer Ultra Low RNA kit (Clontech). Cells were loaded onto the chip at a concentration of 3–5E5 cells/ml, stained for viability (LIVE/DEAD cell viability assay; Thermo Fisher) and imaged by phase-contrast and fluorescence microscopy to assess the number and viability of cells per capture site. Only single, live cells were included in the analysis. For qPCR experiments, harvested cDNA was pre-amplified using a 0.2X pool of primers prepared from the same gene expression assays to be used for qPCR. Pre-amplification allows for multiplex sequence-specific amplification 78 targets. In detail, a 1.25 µl aliquot of single cell cDNA was pre-amplified in a final volume of 5 µl using 1 µl of PreAmp Master Mix (Fluidigm) and 1.25 µl pooled TaqMan assay mix (0.2x). cDNA went through amplification by denaturing at 95°C for 15 s, and annealing and amplification at 60°C for 4 min for 20 cycles.After cycling, pre-amplified cDNA was diluted 1:5 by adding 20 µl TE Buffer to the final 5 µl reaction volume for a total volume of 25 µl.
Single-cell gene expression experiments were performed using the 96x96 quantitative PCR (qPCR) DynamicArray microfluidic chips (Fluidigm). A 2.25 µl aliquot of amplified cDNA was mixed with 2.5 µl of TaqMan Fast Adavanced Master Mix (Thermo Fisher) and 0.25 µl of Fluidigm's "sample loading agent," then inserted into one of the chip "sample" inlets. A 2.5 µl aliquot of each 20X TaqMan assay was mixed with 2.5 µl of Fluidigm's "assay loading agent" and individually inserted into one of the chip "assay" inlets. Samples and probes were loaded into 96 x96 chips using an IFC Controller HX (Fluidigm), then transferred to a BioMark real-time PCR reader (Fluidigm) following manufacturer's instructions. A list of the 78 TaqMan assays used in this study is provided below.

**Single-cell data analysis**: The Quality Threshold in the BioMark™ Analysis software is a qualitative tool designed to measure the "quality" of each amplification curve. Basically, each amplification curve is compared to an ideal exponential curve and as the quality score approaches 1 the closer it is to ideal. The further the curve is from ideal, its quality score approaches 0. The default cutoff of 0.65 is an arbitrary value set by Fluidigm. Any curve above 0.65 passes. Any curve below, fails. Baseline correction was set on Linear (Derivative)[default]. Ct Threshold Method was set on Auto (Detectors). This method independently calculates a threshold for each detector on a chip. For clustering and downstream analysis, raw Cts have been converted to Log2Exp by using a Limit of Detection (LOD) of 35, which corresponds to the last PCR cycle. Co-expression analysis has been performed by considering both CRC and NSCLC samples on those genes for which both FOXP3 and IL2RA were co-expressed at least to 2%. Gene's levels above the background were depicted as violin plots after log2 scale transformation by ggplot2 (v. 2.1.10). The violin color gradient is the percentage of cells that are expressing the gene of interest and the upper bound of the color scale is the maximum percentage of cells that express a gene of the whole geneset.

**Procedure for the removal of transcripts whose expression values are affected by the 'dropout' effect.**
Single-cell qPCR data are inherently noisy, and due the limitations of current technologies the expression patterns of a certain number of genes may be afftected by the 'dropout effect'. We performed a gene selection procedure in order to take into account this 'dropout' effect and discard those genes whose expression values cannot be reliably used in a binary comparison (tumor-peripheral vs blood). We fitted a number of parametric distributions to the ratios of detected genes on the total number of tumor cells (both NSCLC and CRC) and selected the reciprocal inverse Gaussian continuous random variable as best fit. We then calculated the median value of the fitted distribution and discarded those genes whose

detection ratio is less than this threshold value (at least 8.4% of detection). We reasoned that these genes are more likely to be affected by the 'dropout' effect. With this threshold we selected 45 genes for which a non-parametric T-test (Wilcoxon Mann Whitney test $p<0.05$) has been performed (by comparing tumor vs. peripheral blood samples)

**Meta analysis Kaplan-Meier and stage correlation**

Statistical analysis was performed by using the R survival package (Therneau T. 2013). Survival times were calculated as the number of days from initial pathological diagnosis to death, or the number of days from initial pathological diagnosis to the last time the patient was reported to be alive. The Kaplan-Meier (KM) was used to compare the high and low expression levels of the tumor-Treg cell signature transcripts in either CRC (GSE17536) and NSCLC (GSE41271) patients. For both studies annotation was normalized to four tumor stages (1,2,3,4). For study GSE41271 five patients were excluded due to incomplete or inaccurate annotation (GSM1012883,GSM1012884,GSM1012885,GSM1013100,GSM1012888), retaining a total of two hundred and sixty three patients. Patients from both studies were labeled as 'High' 'Low' whether or not their relative expression values exceeded a decision boundary (mean of the samples). We define $\ddot{x}_{ij}$ to denote the relative expression of the gene *i* for the *n* samples of the study normalized to the CD3 level:

$$\ddot{x}_{i,j} = \frac{x_{i,j}}{x_{CD3G,j}}\ ; \qquad i = (CCR8, MAGEH1, LAYN) \quad j = 1,2,\ldots\ldots n\ samples$$

To classify a patient, a threshold on the $\ddot{x}_{ij}$ is required and defined as

$$T_{(Upper,Lower)} = median(\ddot{x}_{i,j}) \pm \frac{\sigma(\ddot{x}_{i,j})}{10}$$

where $T_{(Upper,Lower)}$ represent the upper and lower extreme of the decision boundary:

$$\begin{cases} \ddot{x}_{i,j} > T_{Upper}\ High \\ \ddot{x}_{i,j} < T_{Lower}\ Low \\ T_{Upper} \leq \ddot{x}_{i,j} \leq T_{Lower}\ excluded \end{cases}$$

We examined the prognostic significance of tumor Treg cells transcripts by using log-rank statistics; a p-value of less than 0.05 was considered statistically significant.

Since the log-rank test resulted in a p-value of less than 0.05, a post stage comparison by means of box plot representation was performed in order to evaluate the correlation degree between the expression level of the transcripts and tumor stages in the cohort of CRC patients. . The annotation was normalized to four tumor stages (1,2,3,4).

**SUPPLEMENTAL REFERENCES**


Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Loffler-Wirth, H., Kalcher, M., and Binder, H. (2015). oposSOM: R-package for high-dimensional portraying of genome-wide expression landscapes on bioconductor. Bioinformatics *31*, 3225-3227.

Ripley, B. D. (1996) Pattern Recognition and Neural Networks. Cambridge: Cambridge University Press. [i, 2, 8, 36, 42, 59, 69, 73, 79, 91, 92]

Sobin, L.H., Gospodarowicz, M.K., Wittekind, C., International Union against Cancer., and ebrary Inc. (2009). TNM classification of malignant tumours (Chichester, West Sussex, UK ; Hoboken, NJ: Wiley-Blackwell,), pp. xx, 310 p.

Therneau T. 2013. *A package for survival analysis in S.* R package version 2.37-4.

Wirth, H., von Bergen, M., and Binder, H. (2012). Mining SOM expression portraits: feature selection and integrating concepts of molecular function. BioData Min *5*, 18.