In silico characterization of a “universal” Treg signature reveals the pro enkephalin gene as a novel Treg marker

Gilles MARODON, Ph.D

Sorbonne University, INSERM, CNRS, Centre d’Immunologie et de Maladies Infectieuses (CIMI-PARIS), Paris France

correspondence to gilles.marodon@upmc.fr, Sorbonne University, CIMI-PARIS, 91 Bd de l’Hopital, 75013 PARIS

Funding: no extramural financial support was obtained for this work

Abbreviations: Treg, Regulatory T cells; Tconv, Conventional T cells; TNFRSF, Tumor Necrosis Factor Receptor Super Family; MTS, Meta-Treg Signature; TF, Transcription Factors; MENK, Met-Enkephalin; PPI, Protein-Protein Interactions

Keywords: Bioinformatics, Mice, Treg, TNFRSF, IL-2, proenkephalin, met-enkephalin

Abstract

Regulatory T cells (Treg) are crucial in the proper balance of the immune system. A better knowledge of Treg-specific genes may provide novel targets for therapeutic purposes and extend our knowledge on their complex biology. However, to date there is no consensual Treg signature in the literature. Here, we extracted a list of 72 genes differentially expressed in Treg compared to $CD4^+$ conventional T cells across 6 different but comparable publicly available datasets. A simple network analysis confirms that *Foxp3* was a central node of the MTS, linked to many genes of the signature. A third of the genes from the MTS interacted somehow with IL-2, confirming the central role for this cytokine in Treg biology. When projected into a quantitative gene expression database, many genes of the Meta-Treg signature were also expressed by other immune and non-immune cell subsets, with the noticeable exceptions of *Foxp3*, *Ctla4*, *Tnfrsf4* and *Tnfrsf9* and surprisingly, the pro enkephalin (*Penk*) gene. Subsequent bioinformatic analysis of available datasets indicated the molecular mechanisms that could explain specific *Penk* expression in Treg. Altogether, our results show that the “universal” Treg signature concerns a very limited set of genes centered on the IL-2 family, members of the TNF receptor superfamily, and the endogenous opioid pathway.

Introduction

Proper number and function of regulatory T cells (Treg) are essential for a well-balanced immune system: too few of these cells leads to autoimmunity and too much prevents an efficient immune response, with harmful consequences for anti-tumor immunity, for instance. Treg are a subset of $CD4^+$ T cells that express the transcription factor (TF) Foxp3 and the alpha chain of the interleukin-2 receptor CD25, both indispensable for suppressive functions and proper homeostasis. Probably the best example illustrating the crucial role for Foxp3-expressing cells in the homeostasis of the immune system is given by the lethal auto immune syndrome seen in patients bearing mutations in the *FOXP3* gene, the IPEX syndrome [1]. Like most $CD4^+$ T cells, Treg are generated in the thymus upon MHC-driven selection based on affinity of the T cell receptor for self antigens [2]. A sizable proportion of those cells are also induced in the periphery (pTreg) from CD4+ T cells precursors, but those pTreg are reported not to express Helios or Neuropilin-1, contrary to thymic-derived Treg (tTreg) that are positive for those markers [3]. Known pTreg inducers in the periphery can be byproducts of bacterial metabolism [4,5] but it is likely that inflammation per se is a main driver for pTreg differentiation [6]. Thus, finding ways of manipulating Treg for therapeutic purposes in the auto immunity field has become a major endeavor for immunologists worldwide. Moreover, more recent results linking the presence of Treg to a bad prognostic in some cancers extended their potential clinical applications from autoimmunity to cancer immunotherapy [7]. In the case of cancer, one would want to get rid of Treg to wake up a dim immune response to tumors. An example of this powerful approach has been recently illustrated by Treg-depleting CTLA-4 or CD25-specific mAb [8,9]. However, this weak response to tumors is part of a natural tolerance process, preventing the immune system to attack self-tissues. Thus, breaking immune tolerance by removing Treg is not without consequences on the integrity of healthy tissues. This is illustrated by studies showing that Ipilimumab (anti-CTLA-4), a powerful anti-cancer drug affecting Treg, indeed help the immune system to fight tumors at the expense of a generalized auto immunity in treated patients [10]. A better knowledge of Treg biology will be crucial for preserving therapeutic efficacy without severe adverse events.

This knowledge has been mostly collected from mice due to their ease of use, their versatility and the thousands of genetic models available to answer mechanistic questions. To that end, hundreds of investigators have pursued the quest for specific Treg markers, revealing molecules and pathways potentially targetable by monoclonal antibodies or pharmacological compounds. However, there are still confusions about these specific markers, since the comparisons are often made across

unrelated studies employing different technologies. Moreover, a characteristic of Treg is their ability to adapt to the cellular environment in which they are present. This mechanism, referred as to effector class control few years ago [11], deeply affect cell surface phenotype and thus presumably, gene expression patterns as well. In addition, there are evidences that pTreg, contrary to tTreg, might be relatively unstable [12], meaning that pTreg may acquire some effector functions. One can immediately realize that providing a common definition of Treg across this diversity of phenotype and function represents a difficult challenge. Furthermore, Treg specific markers should ideally mark Treg only or should be minimally expressed by other cells of the immune system or other non-immune cells. This has been an overlooked issue since most of the so-called “Treg signatures” are established relative to $CD4^+$ T conventional cells (Tconv), and not looking outside of Tconv (not mentioning major phenotypic differences used to define Treg and Tconv). Safety and efficacy of Treg-based therapies will surely benefit from targeting molecules and/or pathways truly specific to Treg.

In an attempt to resolve some of these issues, we reasoned that digging out a meta-Treg signature (MTS) from available datasets comparing Treg and Tconv should lead to a more robust signature than studies taken separately. Projection of this qualitative signature in quantitative databases reporting normalized gene expression level across multiple cell types should also inform us on the specificity of the MTS beyond Tconv. As a starting point, we decided to focus our investigation in Mus Musculus, because many databases and tools are available in mice. We also focused our analysis in resting Treg freshly isolated ex vivo, taken from non inflammatory settings, to avoid bias due to T cell activation. This naive approach led to the first in silico description of a “universal” Treg signature. The meta-Treg signature compiled here provide a common reference data set that might prove useful to the immunology community and provide new hypotheses to investigate Treg biology further.

Methods

Hardware and software informations

All the analyses in this report were performed using a MacBook Pro mid-2012 (MACOS 10.14.4) with 8Go of RAM and a 2.5Ghz Intel Core i5. The report was written with LibreOffice v6.1.3.2 and Mendeley 1.19.4. Figures were mounted on LibreOffice from panels extracted from the indicated websites. In some instances, Prism v7 (GraphPad Inc) was used.

Extraction of the Meta-Treg signature (MTS)

The datasets used were selected based on a “Treg* AND (Tconv* OR Teff*) AND Mus Musculus” search in the GEO dataset web site (https://www.ncbi.nlm.nih.gov/gds). GEO datasets were manually inspected for inclusion of studies comparing fresh Treg with fresh Tconv from lymphoid organs. Only GSE17580 [13], GSE24210 [14], GSE37532 [15,16], GSE40685 [17], GSE42021 [18], and GSE7852 [19] were selected and a list of genes significantly up or down regulated in Treg compared to Tconv was determined for each dataset using GEO2R with an adjusted p.value cutoff of 0.05 (FalseDiscoveryRate). On average, 687 (range: 338-1581) genes were isolated using this threshold. The commonality within these gene sets was determined using the online tool from the Bioinformatics Evolutionary Genomics department form the Ghent University (http://bioinformatics.psb.ugent.be/webtools/Venn/). All but one dataset (GSE37532) where the [MoGene-1_0-st] Affymetrix Mouse Gene 1.0 ST Array was used, were generated using the [Mouse430_2] Affymetrix Mouse Genome 430 2.0 Array.

Analysis of the MTS

For representing the relations between genes of the MTS, we used Cytoscape v3.7.1 [20] (https://cytoscape.org) with the Spring application and the attached Enrichment analysis for GO terms [21], establishing a network of protein-protein interactions. It should be stressed that these interactions are not necessarily physical but represent regulation or indirect association, manually curated from the literature. The Treg signature from ImmGen (http://immgen.org) was obtained using the Population Comparisons tab from the Data Browsers window. We gathered quantitative data on the expression of the MTS genes in immune and non immune cell subsets using the MyGeneSet tab. Analysis of ATAC-Seq was done using the Chromatin tab. Expression of the *Oprd1* gene was assessed using the Gene Skyline tab with RNA Seq data. Expression of *Penk* in various studies is represented by the graphical tool embedded in GEO2R. Ingenuity Pathway Analysis (Qiagen) was used to investigate upstream cytokine regulators linked to genes of the MTS.

Results

As a starting point, we searched for a common Treg signature (that is a list of genes differentially expressed in Treg compared to conventional T cells (Tconv)) across publicly available datasets retrieved from the GEO web site. We manually selected 6 datasets based on various subjective criteria, and availability of the GEO2R analytical tool for the dataset. We deliberately limited the search to cells isolated from peripheral lymphoid tissues to avoid any bias due to potential tissue specificity of Treg or Tconv genetic profiles. Among the 6 datasets, 4 had been performed in lymph node cells, 1 in spleen and 1 in the bone marrow, 5 had been performed in C56BL/6 (females and males) and one in BALB/c mice. Various ages were present in the datasets (range: 6-36 weeks-old). Various Treg sorting strategies were also used across the datasets, from classical $CD25^+$ sorting to isolation of GFP^+ cells from Foxp3-GFP transgenic mice. So we believe that the 6 chosen datasets were representative of typical Treg profiles found in many studies.

The intersection analysis of differentially expressed genes from these datasets is given as a supplemental table (Table S1): 71 genes were found to be common to all 6 datasets whereas a total of 194 genes were found in at least 5 datasets. Because we were stringent on the initial choice of our 6 datasets, we retained the list of genes common to all 6 datasets, establishing a Meta-Treg signature (MTS). Surprisingly, *Foxp3*, the master TF of Treg, was absent from this list because it was missing from one dataset (Table S1). To be consistent with previous knowledge on Treg, we thus “forced” *Foxp3* into the MTS, establishing a list of 72 genes either enriched or diminished in Treg relative to Tconv (Fig. 1A). Well-known Treg markers, such as *il2ra* (CD25), *Ikzf2* (Helios), *Ikzf4* (Eos), *Il2rb* (CD122), *Socs2*, *Nrp1*, *Ebi3* or *Ctla4* were present in the MTS.

To gain insight into biological functions associated with genes of the MTS, we generated a putative network of proteins-proteins interactions (PPI) from the list of Treg-enriched genes present in the MTS (Fig. 1B). A large number of members of the MTS were not connected to any other member of the list, suggesting that the MTS did not define a particular signaling or metabolic pathway. Those genes were thus excluded from the representation. Significant interactions above randomly expected ones (PPI enrichment p.value $<10^{-16}$) were observed for well established members of the Treg signature (Fig. 1B). A central node of PPI was composed of *Foxp3*, *Ctla4*, *Il2ra*, *il2rb*, *Tnfrsf4* (CD134, OX40), *Tnfrsf9* (CD137, 4-1BB), *Itgae* (CD103), *Ccr6*, *CD83*, *Klrg1*, and *Entpd1* (CD39) gene. A simple network analysis reveals that *Foxp3* was the node most connected to other nodes of the network (degree), as depicted by the size of the node in the figure. The *Foxp3* node had also the higher “betweenness” of the network, showing that most of the nodes directly connected to *Foxp3*.

Those genes constituting the “core” of the MTS were associated with generic GO terms typical of activated cells and linked to regulatory biological processes (Fig. 1B). To establish the robustness of our MTS, we compared it to the one generated in the ImmGen database curated by the Immunological Genome Project Consortium [22]. As expected, most of the MTS genes were found in the list of 1000 DE genes provided by ImmGen using microarray v1 dataset (Fig 1D). Interestingly, 5 genes of our MTS were not present in the ImmGen Treg signature (*Ybx3*, *Ncmap*, *Scamp1*, *Prg4*, and *Gucd1*).

To go further, we searched for possible upstream regulators of the MTS, that would be cytokines with known direct or indirect connections with genes of the MTS. Five cytokines had significant connections with many genes of the MTS (Fig. 2). Not surprisingly, IL-2 was the prime cytokine with connections to 26 genes of the MTS, a highly significant enrichment relative to a random set of genes. More surprising was that IL-4 and IFNg had also connections with genes of the MTS. IL-5 and IL-27 had fewer connections whereas TNF had connections with only 2 members of MTS, the TNFRSF, as expected. Thus, IL-2 might chiefly contribute at establishing the Treg signature but IL-4, IFNg, IL-5, IL-27 and TNF might participate as well.

Because of the good qualitative overlap with our MTS, we used the ImmGen ULI RnA Seq database to determine the expression levels of each genes of the MTS, giving the unique opportunity to quantitatively define “true” vs “shared” Treg markers across multiple immune and non-immune cell types (Fig. 3). There is two subsets of Treg in the Immgen dataset: one is labelled as 4_25hi-Sp and the other is referred as 4_FP3+_Nrplo_Co, isolated from the colon, and most likely representing pTreg. To be coherent with the MTS, mostly obtained from pTreg, we present the heatmap ordered according to expression levels of each gene of the MTS in pTreg, from high on the top to low on the bottom of the figure. Although the general hierarchy of expression was respected, we noticed several differences between the two Treg subsets for some genes of the MTS: *Klrg1* expression was low in CD25+ Treg but high in pTreg whereas it was the contrary for *Gpr83. Ift80* and *Arhgap20* were highly expressed in CD25+ Treg but not in pTreg. However, those genes were poorly specific to Treg since similar levels of expression were observed in Tconv. A second observation was that the MTS extracted from microarrays well defined Treg when projected into the RNA Seq database: about half of the genes of the MTS positively defined Treg relative to Tconv (yellow/red in Treg, white/blue in Tconv) whereas the other half was a mix between genes negatively or not discriminating Treg from Tconv (yellow/red in Tconv, white/blue in Treg). With this representation (Fig. 3), we can then observed which genes truly defined Treg across all the subsets present in ImmGen: only few genes at the top of the list positively defined Treg when compared to the rest of the subsets. Not surprisingly, *Ctla4* was a very specific one, basically expressed at higher levels in Treg than in any other subsets. More

surprising was the presence of *Ncmap* at the top of the list but since it was not part of the overlap with the microarray data (Fig 1D), and since some expression can be seen Tconv (Fig. 3), we cannot firmly conclude on *Ncmap* at this stage. The third gene of the list was *Tnfrsf4* (CD134, OX40), with a pattern of expression resembling *Ctla4* but with detectable expression in a subset of $CD8^+$ dendritic cells. The fourth member of the list was *Penk*, a gene coding for the proenkephalin enzyme, responsible for production of Met-Enkephalins (MENK) peptides, belonging to the endogenous opioid family. However, expression of *Penk* was also detected in stromal cell types. Next was *Foxp3,* highly expressed in Treg, and not detected in any other subsets. *Tnfsrf9* (4-1BB) was next on the list, but its expression was detected in several other subsets, including NK, NKT, and γδ-T cells. The rest of the MTS was much less specific to Treg although their expressions remain enriched in Treg relative to Tconv. At the other end of the spectrum, few genes negatively defined Treg (that are weakly expressed in Treg relative to Tconv) but, as expected, low expression of those genes was not specific to Treg. Most B cells, for instance, were also negative for these markers. Thus, projection of the MTS into the ImmGen database led to expected and unexpected findings: (i) the MTS genes segregated Treg from Tconv but most of the genes were also expressed by other cells than Treg, (ii) some genes were highly specific to Treg, including *Ctla4*, *Foxp3,* members of the TNFRSF family and, more surprisingly *Penk*, a gene belonging to the endogenous opioid pathway.

Following on the surprising observation that *Penk* was a highly specific Treg genetic marker, we wanted to confirm this observation by other means. First, we undertook a survey of studies examining gene expression in Tconv and Treg in various tissues: *Penk* was found to be over expressed by Treg in the thymus [19], in the spleen [23] and LN [15,16] (Fig. 4A) and of several tissues , including fat [15,16], muscle [23] and tumors [24] (Fig. 4B), showing that *Penk* over expression by Treg is consubstantial of their generation and independent of their localisation. We then wanted to know more about the molecular mechanisms that may explain *Penk* over expression in Treg. Using ATAC-Seq data available in ImmGen, we analyzed the chromatin profile of the *Penk* regulatory region in Treg. There was an open chromatin region (OCR) near the TSS in Treg that was not observed in naïve $CD4^+$ T cells, nor in naïve CD8+ T cells (highlighted in Fig. 4C). In contrast, this OCR was observed in stromal cells and in $CD4^+$ T cells stimulated in vitro (Fig. 5C) in which Penk expression was also detected (Fig. 3). Thus, an OCR in the *Penk* promoter region was present specifically in cells permissive for *Penk* expression. We also noticed the presence of a CpG island located upstream of the TSS (the *Penk* gene is on the negative DNA strand). To investigate the possibility that *Penk* expression might be regulated by the methylation status of this region, we looked for *Penk* expression levels in Treg deficient for the DNA-methyl-transferase Dnmt1, the main driver of epigenetic regulation in Treg [25]. Deficiency of Dnmt1 in Treg had no effect on *Penk* expression (Fig.

5A), suggesting that the CpG island of *Penk* is demethylated in wild-type Treg (a methylated state in wild-type Treg would have been associated with an increase in *Penk* gene expression in the absence of Dnmt1). Moreover, deletion of DICER, a master regulator of miRNA function in Treg [26], was associated with a considerable increase in *Penk* expression in Treg (Fig. 5A), suggesting that miRNAs repress *Penk* expression in Treg. Finally, we wanted to know whether the expression of *Penk* by Treg could have functional consequences relevant to the immune system. Indeed, the main ligand of Penk by-products (MENK peptides), the *Oprd1* gene (coding for the opioid delta receptor-1) was expressed in $CD4^+$ Tconv, and highly expressed in $CD4^+$ and $CD8^+$ dendritic cells (DCs) relative to other immune and non immune cell types (Fig. 5B), Interestingly, *Oprd1* was also highly expressed by Vγ2+ T cells, a subset of innate γδ-T cells, known to be regulated by Treg in the intestine [27]. Thus, the main receptor for MENK was expressed by immune subsets known to dialog with Treg.

Finally, we wanted to investigate whether *Penk* over expression by murine Treg would also be observed in humans. The recent availability of the immune cell atlas (http://immunecellatlas.net) focusing on human cells provided a mean to that end although only few cell types and only from the blood are available for now. Remarkably, PENK expression was detected in Mucosal Associated Invariant T cells (MAIT) (not present in the ImmGen subsets) and in activated Treg (Fig. 5C), confirming our observations in mice. An interesting difference with the mouse model was the high expression of the OPDR1 gene in activated Treg, contrary to its almost complete absence in mice Treg (Fig 5B), suggesting that an autocrine loop of MENK/OPDR1 interactions may exist in humans and not in mice.

Discussion

To our knowledge, the present report is the first attempt to define a “universal” Treg signature in mice by a meta analysis of published and available datasets and provide a more solid appreciation of Treg-specific genes. Our unbiased approach confirm that the genes central to Treg identity revolves around well established family of genes belonging to the interleukin-2 family and emphasize the contribution of members of the TNF receptor super-family. Our gene ontology enrichment analysis confirms that the core set of genes of the MTS belongs to biological processes important for Treg function but also show that most of the genes of the MTS did not belong to any particular biological function. One point of caution for the interpretation of this latter observation is that mRNA expression may or may not be correlated with protein expression, and thus with biological functions. For instance, a recent study has found strong discrepancies between mRNA and protein levels in human Treg [28].

Our primary objective was to evaluate whether a true core set of genes really defined Treg across the whole diversity of cell types in the immune system and beyond. Our analysis raises the hypothesis that targeting together the most Treg-restricted set of genes that belongs to the IL-2 and TNFRSF family members, might represent the most effective strategy to affect Treg preferentially over Tconv. But what about other cell types? We found that most of the genes of the MTS were also expressed by other cells of the immune system. Thus, a “universal” Treg signature, reflecting hard wiring of Treg lineage at the gene expression level might not exist at all, a view supported by functional plasticity of Treg (effector class control), and possible acquisition of effector functions by Treg in some inflammatory conditions.

Analysis of the MTS extend the notion that Treg lineage specification (that must be separated from Treg identity [29]) is mostly dictated at the epigenetic rather than at the gene expression level [30]. Nevertheless, a few genes stand out to be very specific to Treg beyond the comparison with Tconv. These were *Ctla4*, *Tnfrsf4* and *Tnfrsf9* and obviously *Foxp3,* all confirmed to be highly expressed in Treg at the protein level. Thus, these molecules should probably concentrate the efforts for therapeutic targeting of Treg. The former is already in the clinics (Ipilimumab) with great efficacy to affect Treg but with possible severe adverse events whereas the others have been the subject of many preclinical studies and are currently being tested in clinical trials. The TF Foxp3 is to our understanding not “druggable”, given the possible catastrophic consequences that this might have on immune tolerance.

A surprising and serendipitous finding in the search for a “universal” Treg signature was that the *Penk* gene was highly specific to Treg in the immune system. Actually, we investigated whether the enrichment in *Penk* at the genetic level would also be seen at the protein level. Deceptively, Penk was not listed in 2 different databases reporting the proteomes of murine Treg or Tconv [31,32]. We believe this observation reflects technical issues related to mass spectrometry rather than a “real” negative result contradicting our analysis. As an evidence for such technical issues, the Tnfrsf1b (TNFR2) protein, a member of the MTS, was also not found in the databases whereas it is clearly expressed by Treg at the cell surface and plays important role in Treg biology [33]. This observation confirms that proteomic analyses are not yet exhaustive.

Additional bioinformatic analysis gave some clues on the molecular mechanisms that might be responsible for *Penk* expression in Treg, namely an OCR close to the *Penk* TSS and a possible demethylated CpG island in the promoter region. We also report a massive up regulation of *Penk* expression in Treg deprived of DICER, the master regulator of miRNA processing and function and crucial for Treg identity and function [26]. This suggest that specific *Penk* expression in Treg may originate from multiple mechanisms, at the epigenetic and post transcriptional levels. Obviously, “wet” experiments are needed to confirm and extend these bioinformatics-generated hypotheses. Nevertheless, we believe that the endogenous opioid pathway should be considered to understand Treg biology further. The impact of opioid peptides produced by innate and adaptive immune cells on immune suppression and pain relief is of little doubts in mice contrary to humans [34]. Moreover, recent studies using *Penk* knock-out mice have unambiguously identified opioid peptides produced by $CD4^+$ T cells as key players in a model of pain control in vivo [35]. Although *Penk* expression by Treg has not been reported before, an influence of MENK peptides on Treg and DC biology, with possible functional impact on tumor growth, has been reported by the group of F. Shan [36,37]. However, the impact of MENK on Treg was rather dull, in contrast to the effects of DC that were more pronounced, in agreement with the high expression of the *Oprd1* gene on DC that we report here. Based on specific *Penk* expression in Treg and the expression pattern of the MENK peptides receptor that we report here, we would like to speculate that peripheral nociception (the neuronal pain sensing pathway) might be intermingled with immune regulation at the Treg level, an hypothesis testable beyond bioinformatics analyses. It will be particularly crucial to analyze mice specifically deficient for *Penk* in Treg to determine the direct vs indirect impact of *Penk* expression on the immune response and how that relates to peripheral nociception.

Acknowledgements

The author wish to thank Dr B Combadière for access to IPA, all members of the ITAC team and all the members of the CIMI-PARIS research center for their encouragements and support, the anonymous curators of GEO and ImmGen for maintaining up-to-date databases.

References

1. Bennett CL, Christie J, Ramsdell F, Brunkow ME, Ferguson PJ, Whitesell L, Kelly TE, Saulsbury FT, Chance PF, Ochs HD. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. Nat Genet. 2001 Jan 1;27(1):20–1.

2. Hsieh CS, Lee HM, Lio CWJ. Selection of regulatory T cells in the thymus. Nature Reviews Immunology. 2012.

3. Yadav M, Louvet C, Davini D, Gardner JM, Martinez-Llordella M, Bailey-Bucktrout S, Anthony BA, Sverdrup FM, Head R, Kuster DJ, Ruminski P, Weiss D, Von Schack D, Bluestone JA. Neuropilin-1 distinguishes natural and inducible regulatory T cells among regulatory T cell subsets in vivo. J Exp Med. 2012;

4. Atarashi K, Tanoue T, Oshima K, Suda W, Nagano Y, Nishikawa H, Fukuda S, Saito T, Narushima S, Hase K, Kim S, Fritz J V., Wilmes P, Ueha S, Matsushima K, Ohno H, Olle B, Sakaguchi S, Taniguchi T, Morita H, Hattori M, Honda K. Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. Nature. 2013 Jul 10;

5. Smith PM, Howitt MR, Panikov N, Michaud M, Gallini CA, Bohlooly-Y M, Glickman JN, Garrett WS. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. Science. 2013 Aug 2;341(6145):569–73.

6. Gagliani N, Amezcua Vesely MC, Iseppon A, Brockmann L, Xu H, Palm NW, De Zoete MR, Licona-Limón P, Paiva RS, Ching T, Weaver C, Zi X, Pan X, Fan R, Garmire LX, Cotton MJ, Drier Y, Bernstein B, Geginat J, Stockinger B, Esplugues E, Huber S, Flavell RA. TH17 cells transdifferentiate into regulatory T cells uring resolution of inflammation. Nature. 2015;

7. Tanaka A, Sakaguchi S. Regulatory T cells in cancer immunotherapy. Cell Research. 2017.

8. Simpson TR, Li F, Montalvo-Ortiz W, Sepulveda MA, Bergerhoff K, Arce F, Roddie C, Henry JY, Yagita H, Wolchok JD, Peggs KS, Ravetch J V, Allison JP, Quezada S a. Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma. J Exp Med. 2013 Jul 29;210(9):1695–710.

9. Arce Vargas F, Furness AJS, Solomon I, Joshi K, Mekkaoui L, Lesko MH, Miranda Rota E, Dahan R, Georgiou A, Sledzinska A, Ben Aissa A, Franz D, Werner Sunderland M, Wong YNS, Henry JY, O'Brien T, Nicol D, Challacombe B, Beers SA, Melanoma TRACERx Consortium, Renal TRACERx Consortium, Lung TRACERx Consortium, Turajlic S, Gore M, Larkin J, Swanton C, Chester KA, Pule M, Ravetch J V., Marafioti T, Peggs KS, Quezada SA. Fc-Optimized Anti-CD25 Depletes Tumor-Infiltrating Regulatory T Cells and Synergizes with PD-1 Blockade to Eradicate Established Tumors. Immunity. 2017;46(4):577–86.

10. Friedman CF, Proverbs-Singh TA, Postow MA. Treatment of the Immune-Related Adverse Effects of Immune Checkpoint Inhibitors: A Review. JAMA oncology. 2016.

11. Chaudhry A, Rudra D, Treuting P, Samstein RM, Liang Y, Kas A, Rudensky AY. CD4+ regulatory T cells control TH17 responses in a Stat3-dependent manner. Science. 2009 Nov 13;326(5955):986–91.

12. Hori S. Regulatory T cell plasticity: beyond the controversies. Trends Immunol. 2011/06/04. 2011;32(7):295–300.

13. Layland LE, Mages J, Loddenkemper C, Hoerauf A, Wagner H, Lang R, Prazeres da Costa CU. Pronounced Phenotype in Activated Regulatory T Cells during a Chronic Helminth Infection. J Immunol. 2010;184(2):713–24.

14. Collison LW, Chaturvedi V, Henderson AL, Giacomin PR, Guy C, Bankoti J, Finkelstein D, Forbes K, Workman CJ, Brown SA, Rehg JE, Jones ML, Ni HT, Artis D, Turk MJ, Vignali DA. IL-35-mediated induction of a potent regulatory T cell population. Nat Immunol. 2010/10/19. 2010;11(12):1093–101.

15. Cipolletta D, Feuerer M, Li A, Kamei N, Lee J, Shoelson SE, Benoist C, Mathis D. PPAR-γ is a major driver of the accumulation and phenotype of adipose tissue T reg cells. Nature. 2012;

16. Cipolletta D, Cohen P, Spiegelman BM, Benoist C, Mathis D. Appearance and disappearance of the mRNA signature characteristic of T_{reg} cells in visceral adipose tissue: Age, diet, and PPARγ effects. Proc Natl Acad Sci. 2015;

17. Samstein RM, Arvey A, Josefowicz SZ, Peng X, Reynolds A, Sandstrom R, Neph S, Sabo P, Kim JM, Liao W, Li MO, Leslie C, Stamatoyannopoulos J a., Rudensky AY. Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. Cell. 2012;151(1):153–66.

18. Toker A, Engelbert D, Garg G, Polansky JK, Floess S, Miyao T, Baron U, D√ºber S, Geffers R, Giehr P, Schallenberg S, Kretschmer K, Olek S, Walter J, Weiss S, Hori S, Hamann A, Huehn J. Active Demethylation of the Foxp3 Locus Leads to the Generation of Stable Regulatory T Cells within the Thymus. J Immunol. 2013;190(7):3180–8.

19. Feuerer M, Herrero L, Cipolletta D, Naaz A, Wong J, Nayer A, Lee J, Goldfine AB, Benoist C, Shoelson S, Mathis D. Lean, but not obese, fat is enriched for a unique population of regulatory T cells that affect metabolic parameters. Nat Med. 2009;15(8):930–9.

20. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003 Nov;13(11):2498–504.

21. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. J Proteome Res. 2019 Feb 1;18(2):623–32.

22. Heng TSP, Painter MW, Elpek K, Lukacs-Kornek V, Mauermann N, Turley SJ, Koller D, Kim FS, Wagers AJ, Asinovski N, Davis S, Fassett M, Feuerer M, Gray DHD, Haxhinasto S, Hill JA,

Hyatt G, Laplace C, Leatherbee K, Mathis D, Benoist C, Jianu R, Laidlaw DH, Best JA, Knell J, Goldrath AW, Jarjoura J, Sun JC, Zhu Y, Lanier LL, Ergun A, Li Z, Collins JJ, Shinton SA, Hardy RR, Friedline R, Sylvia K, Kang J. The Immunological Genome Project: networks of gene expression in immune cells. Nat Immunol. 2008;

23. Burzyn D, Kuswanto W, Kolodin D, Shadrach JL, Cerletti M, Jang Y, Sefik E, Tan TG, Wagers AJ, Benoist C, Mathis D. A special population of regulatory T cells potentiates muscle repair. Cell. 2013 Dec 5;155(6):1282–95.

24. 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 Jun 13;15(6):580–7.

25. Wang L, Liu Y, Beier UH, Han R, Bhatti TR, Akimova T, Hancock WW. Foxp3+ T-regulatory cells require DNA methyltransferase 1 expression to prevent development of lethal autoimmunity. Blood. 2013;

26. Liston A, Lu LF, O'Carroll D, Tarakhovsky A, Rudensky AY. Dicer-dependent microRNA pathway safeguards regulatory T cell function. J Exp Med. 2008/08/30. 2008;205(9):1993–2004.

27. Okeke EB, Uzonna JE. The Pivotal Role of Regulatory T Cells in the Regulation of Innate Immune Cells. Front Immunol. 2019 Apr 9;10:680.

28. Cuadrado E, van den Biggelaar M, de Kivit S, Chen Y-Y, Slot M, Doubal I, Meijer A, van Lier RAW, Borst J, Amsen D. Proteomic Analyses of Human Regulatory T Cells Reveal Adaptations in Signaling Pathways that Protect Cellular Identity. Immunity. 2018 May 15;48(5):1046-1059.e6.

29. Hori S. Lineage stability and phenotypic plasticity of Foxp3+ regulatory T cells. Immunol Rev. 2014;

30. Morikawa H, Ohkura N, Vandenbon A, Itoh M, Nagao-Sato S, Kawaji H, Lassmann T, Carninci P, Hayashizaki Y, Forrest ARR, Standley DM, Date H, Sakaguchi S, FANTOM consortium. Differential roles of epigenetic changes and Foxp3 expression in regulatory T cell-specific transcriptional regulation. Proc Natl Acad Sci. 2014;111(14):5289–94.

31. Barra MM, Richards DM, Hansson J, Hofer A-C, Delacher M, Hettinger J, Krijgsveld J, Feuerer M. Transcription Factor 7 Limits Regulatory T Cell Generation in the Thymus. J Immunol. 2015;195(7):3058–70.

32. Duguet F, Locard-Paulet M, Marcellin M, Chaoui K, Bernard I, Andreoletti O, Lesourne R, Burlet-Schiltz O, Gonzalez de Peredo A, Saoudi A. Proteomic Analysis of Regulatory T Cells Reveals the Importance of Themis1 in the Control of Their Suppressive Function. Mol Cell Proteomics. 2017;16(8):1416–32.

33. Salomon BL, Leclerc M, Tosello J, Ronin E, Piaggio E, Cohen JL. Tumor necrosis factor α and regulatory T cells in oncoimmunology. Front Immunol. 2018;9(MAR):1–12.

34. Plein LM, Rittner HL. Opioids and the immune system – friend or foe. Br J Pharmacol. 2018;175(14):2717–25.

35. Basso L, Boué J, Mahiddine K, Blanpied C, Robiou-du-Pont S, Vergnolle N, Deraison C, Dietrich G. Endogenous analgesia mediated by CD4(+) T lymphocytes is dependent on enkephalins in mice. J Neuroinflammation. 2016 Jun 1;13(1):132.

36. Li X, Meng Y, Plotnikoff NP, Youkilis G, Griffin N, Wang E, Lu C, Shan F. Methionine Enkephalin (MENK) inhibits tumor growth through regulating CD4+Foxp3+ regulatory T cells (Tregs) in mice. Cancer Biol Ther. 2015;16(3):450–9.

37. Gao X, Chen W, Plotnikoff NP, Griffin N, Zhang G, Shan F, Meng Y. Methionine enkephalin (MENK) mounts antitumor effect via regulating dendritic cells (DCs). Int Immunopharmacol. 2017;44:61–71.

Figure legends

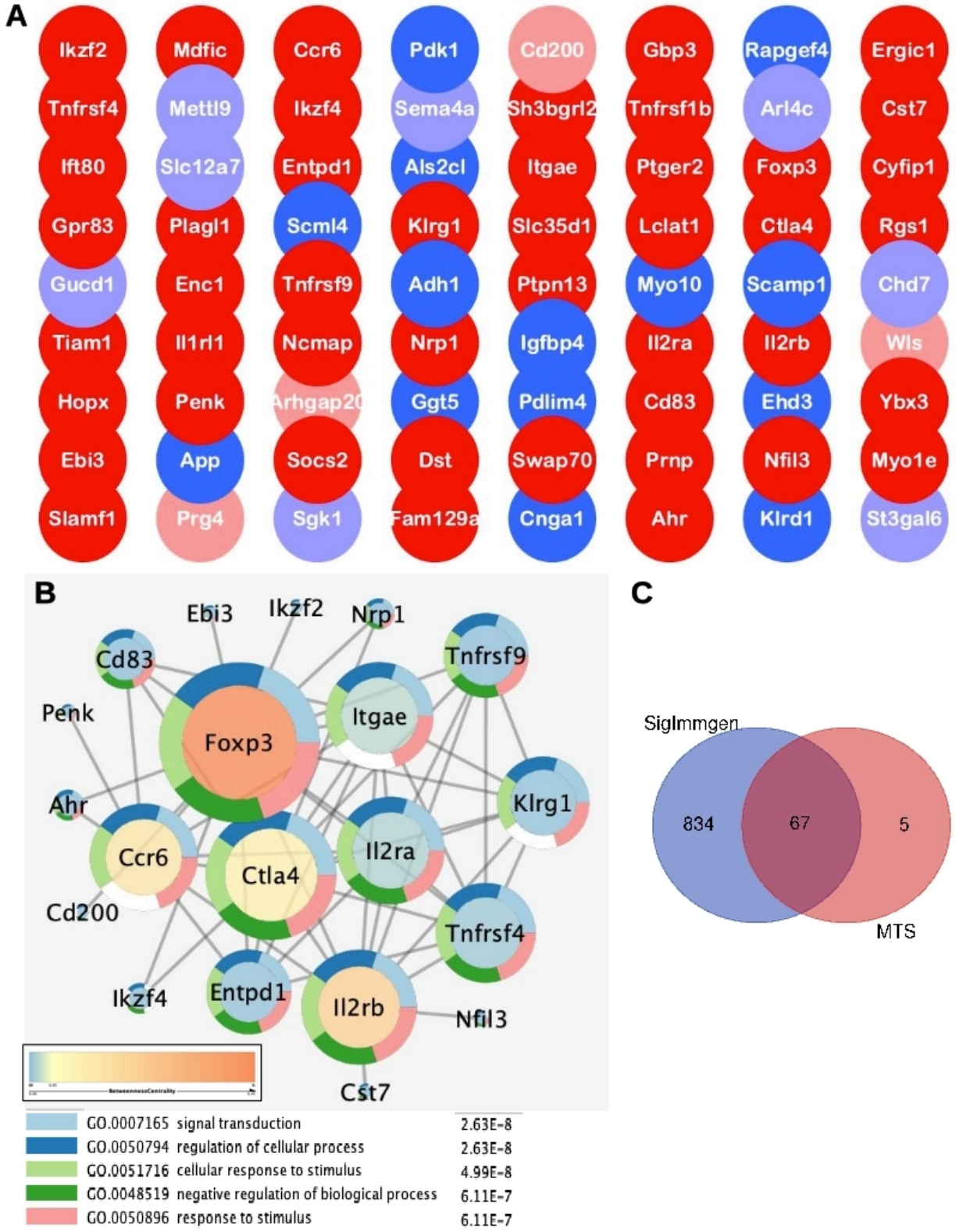
Figure 1: Identification and analysis of the Meta-Treg Signature. **(A)** a list of genes differentially expressed in Treg relative to Tconv common to 6 datasets (see Methods) defines the Meta-treg Signature (MTS)(red= genes up in Treg, blue = genes down in Treg). **(B)** Putative Protein-to-Protein Interaction network of the Treg-enriched genes from the MTS. The size of the node is proportional to the degree (number of connections to other nodes) of each node. The “betweeness” of each node (shortest path to this node) is indicated by a color code (scale is shown in the legend). The GO terms of each node is represented by a color donut around each node (color coded in the legend with the associated FalseDiscoveryRate). **(C)** Intersection of the genes from the MTS and a list of 906 differentially expressed genes from ImmGen (SigImmGen).

Figure 2: Putative cytokines regulators of the MTS according to their cellular localization. Arrows indicate experimentally verified or manually curated interactions from the literature between genes and/or molecules extracted from the IPA databases.

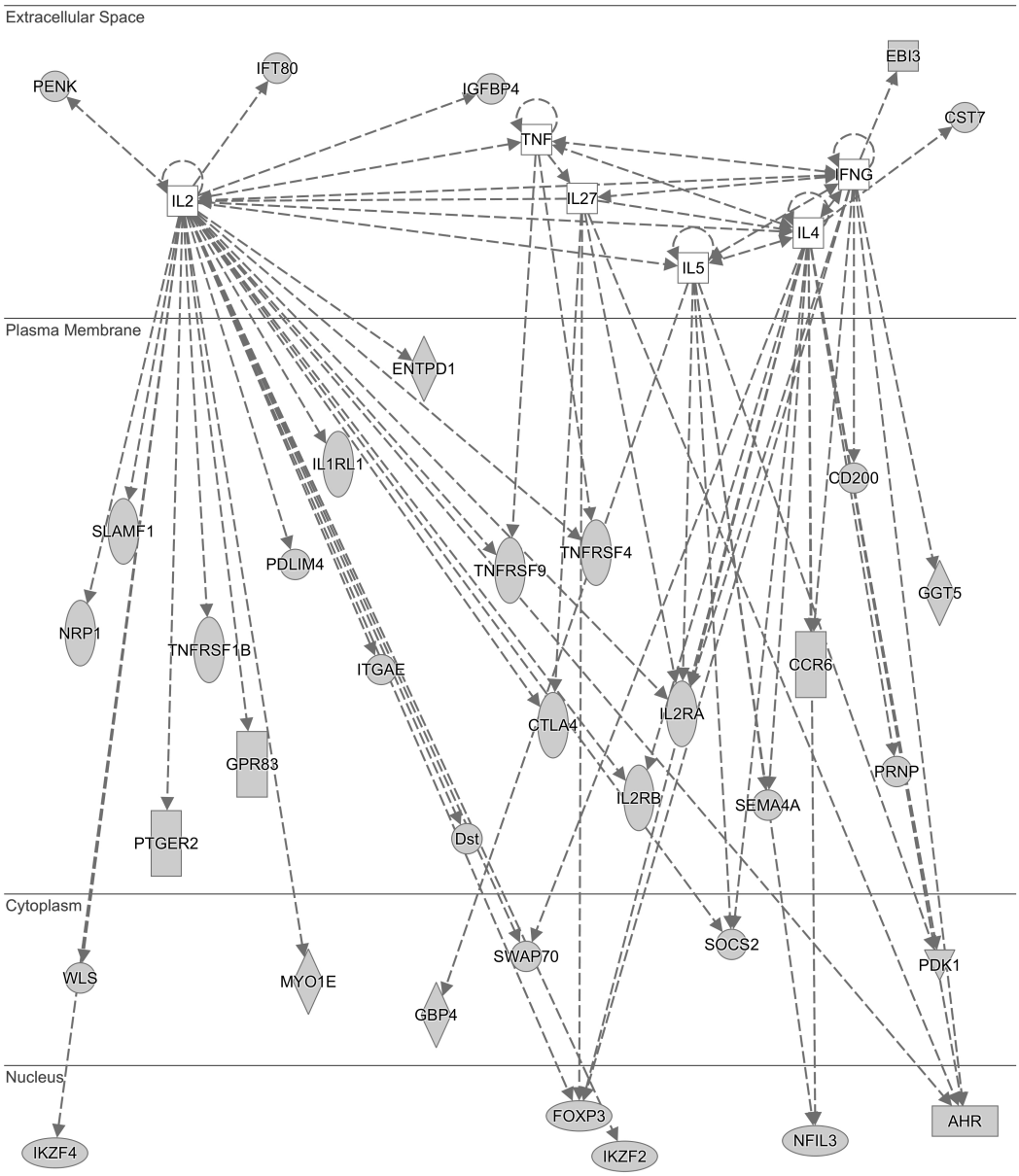
Figure 3: A quantitative assessment of gene expression levels of the MTS across a variety of cell types. Gene list of the MTS was projected into the ImmGen RNA-Seq database and ranked according to gene expression levels in Treg (CD4+FoxP3+Nrp1lo) from the Colon of C57Bl/6 mice. Genes upregulated in Treg relative to Tconv are at the top of the figure (dark red) whereas those down regulated are at the bottom (light blue). Each column is a subset color coded according to its nature (light blue, stem cells; blue, B cells; violet, T cells; light orange, γδ-T cells; purple, NK and Innate Lymphoid cells (ILCs); light green, DCs; green, Macrophages (MF); pink, Monocytes (Mo); red, Granulocytes (GN); orange, stromal cells) and each line is a gene. For sake of clarity, Tconv and Treg columns are highlighted.

Figure 4: **(A)** *Penk* expression in Treg relative to Tconv in lymphoid organs (from left to right: thymus, spleen and lymph nodes) in the indicated GSE datasets. **(B)** Similar analysis in non-lymphoid organs (from left to right: Visceral Adipose Tissue, muscle and B16F10 tumor). Treg are highlighted in light blue, Tconv in light red **(C)** ATAC-Seq profile at the *Penk* locus. Each peak indicates a region accessible to a mutated retrotransposase, revealing an Open Chromatin Region (OCR, highlighted in blue) in activated CD4+ T cells (Act. CD4+), Treg and in Medullary Thymic Epithelial cells (Ep.MEChi) but not in Naive CD4+ T cells (males and females), nor naïve CD8+ T cells. The transcription start site of *Penk* (TSS) is indicated with an arrow, the CpG island is indicated with a green box.

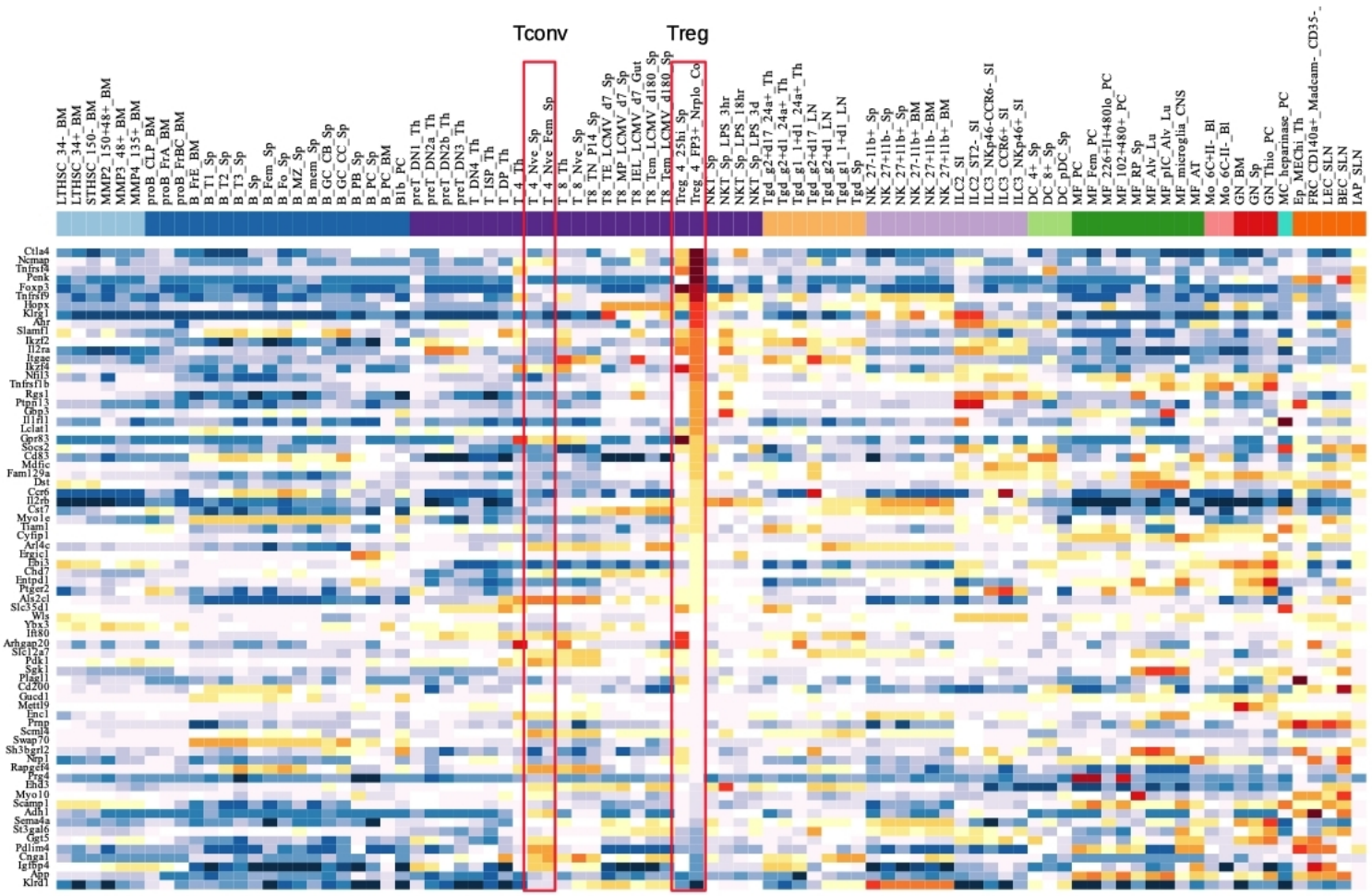
Figure 5: **(A)** *Penk* expression in wild type relative to Dnmt-1-deficient Treg (left panel) or to DICER-deficient Treg (right panel) in the indicated GSE datasets. Wild type Treg are highlighted in light blue, deficient Treg in light red. **(B)** Expression of the *Oprd1* gene (coding for the receptor of MENK peptides produced by Penk) in a variety of cell immune cell types. High expressing cells are labelled in the figure. **(C)** Expression of PENK and OPRD1 genes in human T cell subsets from the blood. Normalized expression value were extracted and represented as columns according to the indicated subsets. For phenotypic definition of the subsets, see http://immunecellatlas.net. Act.=activated


A
Ikzf2
Mdfic
Ccr6
Pdk1
Cd200
Gbp3
Rapgef4
Ergic1
Tnfrsf4
Mettl9
Ikzf4
Sema4a
Sh3bgrl2
Tnfrsf1b
Arl4c
Cst7
Ift80
Slc12a7
Entpd1
Als2cl
Itgae
Ptger2
Foxp3
Cyfip1
Gpr83
Plagl1
Scml4
Klrg1
Slc35d1
Lclat1
Ctla4
Rgs1
Gucd1
Enc1
Tnfrsf9
Adh1
Ptpn13
Myo10
Scamp1
Chd7
Tiam1
Il1rl1
Ncmap
Nrp1
Igfbp4
Il2ra
Il2rb
Wls
Hopx
Penk
Arhgap20
Ggt5
Pdlim4
Cd83
Ehd3
Ybx3
Ebi3
App
Socs2
Dst
Swap70
Prnp
Nfil3
Myo1e
Slamf1
Prg4
Sgk1
Fam129a
Cnga1
Ahr
Klrd1
St3gal6
B
Ebi3
Ikzf2
Nrp1
Cd83
Tnfrsf9
Penk
Itgae
Foxp3
Ahr
Klrg1
Ccr6
Il2ra
Ctla4
Cd200
Tnfrsf4
Ikzf4
Entpd1
Il2rb
Nfil3
Cst7
GO.0007165 signal transduction 2.63E-8
GO.0050794 regulation of cellular process 2.63E-8
GO.0051716 cellular response to stimulus 4.99E-8
GO.0048519 negative regulation of biological process 6.11E-7
GO.0050896 response to stimulus 6.11E-7
C
SigImmgen
834
67
5
MTS

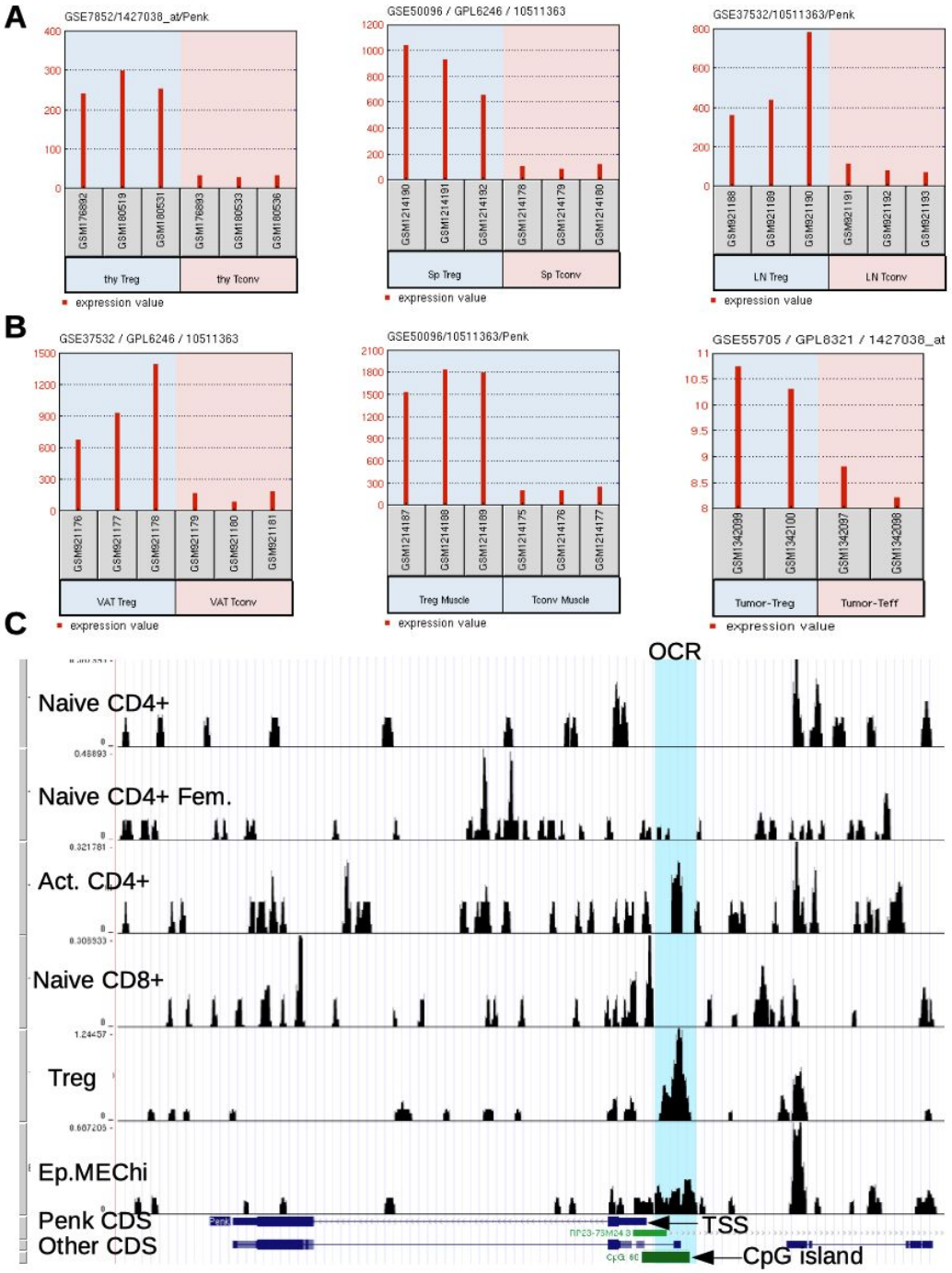
453 **Figure 1**


Extracellular Space
PENK
IFT80
IGFBP4
EBI3
CST7
TNF
IFNG
IL2
IL27
IL4
IL5
Plasma Membrane
ENTPD1
IL1RL1
CD200
SLAMF1
PDLIM4
TNFRSF9
TNFRSF4
GGT5
NRP1
TNFRSF1B
ITGAE
CCR6
CTLA4
IL2RA
GPR83
PRNP
IL2RB
SEMA4A
PTGER2
Dst
Cytoplasm
SWAP70
SOCS2
PDK1
WLS
MYO1E
GBP4
Nucleus
FOXP3
AHR
IKZF4
IKZF2
NFIL3

454 **Figure 2**


Tconv
Treg
LTHSC_34-_BM
LTHSC_34+_BM
STHSC_150-_BM
MMP2_150+48+_BM
MMP3_48+_BM
MMP4_135+_BM
proB_CLP_BM
proB_FrA_BM
proB_FrBC_BM
B_FrE_BM
B_T1_Sp
B_T2_Sp
B_T3_Sp
B_Sp
B_Fem_Sp
B_Fo_Sp
B_MZ_Sp
B_mem_Sp
B_GC_CB_Sp
B_GC_CC_Sp
B_PB_Sp
B_PC_Sp
B_PC_BM
Bl6_PC
preT_DN1_Th
preT_DN2a_Th
preT_DN2b_Th
preT_DN3_Th
T_DN4_Th
T_ISP_Th
T_DP_Th
T_4_Th
T_4_Nve_Sp
T_4_Nve_Fem_Sp
T_8_Th
T_8_Nve_Sp
T8_TN_P14_Sp
T8_TE_LCMV_d7_Sp
T8_MP_LCMV_d7_Sp
T8_IEL_LCMV_d7_Gut
T8_Tcm_LCMV_d180_Sp
T8_Tem_LCMV_d180_Sp
Treg_4_25hi_Sp
Treg_4_FP3+_Nrplo_Co
NKT_Sp
NKT_Sp_LPS_3hr
NKT_Sp_LPS_18hr
NKT_Sp_LPS_3d
Tgd_g2+d17_24a+_Th
Tgd_g2+d1_24a+_Th
Tgd_g1_1+d1_24a+_Th
Tgd_g2+d17_LN
Tgd_g2+d1_LN
Tgd_g1_1+d1_LN
Tgd_Sp
NK_27-11b+_Sp
NK_27+11b-_Sp
NK_27+11b+_Sp
NK_27-11b+_BM
NK_27+11b-_BM
NK_27+11b+_BM
ILC2_SI
ILC2_ST2-_SI
ILC3_NKp46-CCR6-_SI
ILC3_CCR6+_SI
ILC3_NKp46+_SI
DC_4+_Sp
DC_8+_Sp
DC_pDC_Sp
MF_PC
MF_Fem_PC
MF_226+II+480lo_PC
MF_102+480+_PC
MF_RP_Sp
MF_Alv_Lu
MF_pIC_Alv_Lu
MF_microglia_CNS
MF_AT
Mo_6C+II-_Bl
Mo_6C-II-_Bl
GN_BM
GN_Sp
GN_Thio_PC
MC_heparmase_PC
Ep_MECh1_Th
FRC_CD140a+_Madcam-_CD35-
LEC_SLN
BEC_SLN
IAP_SLN
Ctla4
Nemap
Tnfrsf4
Penk
Foxp3
Tnfrsf9
Hopx
Klrg1
Ahr
Slamf1
Ikzf2
Il2ra
Itgae
Ikzf4
Nfil3
Tnfrsf1b
Rgs1
Ptpn13
Gbp3
Il1rl1
Lclat1
Gpr83
Socs2
Cd83
Mdfic
Fam129a
Dst
Ccr6
Il2rb
Cst7
Myo1e
Tiam1
Cyfip1
Arl4c
Ergic1
Ebi3
Chd7
Entpd1
Ptger2
Als2cl
Slc35d1
Wls
Ybx3
Ifit80
Arhgap20
Slc12a7
Pdk1
Sgk1
Plaur1
Cd200
Gucd1
Mettl9
Enc1
Prnp
Scml4
Swap70
Sh3bgrl2
Nrp1
Rapgef4
Pgs4
Ehd3
Myo10
Scamp1
Adh1
Sema4a
St3gal6
Cgt5
Pdlim4
Cngal
Igfbp4
App
Klrd1

455 **Figure 3**


A
GSE7852/1427038_at/Penk
GSM176892
GSM180519
GSM180531
GSM176893
GSM180533
GSM180536
thy Treg
thy Tconv
expression value
GSE50096 / GPL6246 / 10511363
GSM1214190
GSM1214191
GSM1214192
GSM1214178
GSM1214179
GSM1214180
Sp Treg
Sp Tconv
expression value
GSE37532/10511363/Penk
GSM821188
GSM821189
GSM821190
GSM821191
GSM821192
GSM821193
LN Treg
LN Tconv
expression value
B
GSE37532 / GPL6246 / 10511363
GSM821176
GSM821177
GSM821178
GSM821179
GSM821180
GSM821181
VAT Treg
VAT Tconv
expression value
GSE50096/10511363/Penk
GSM1214187
GSM1214188
GSM1214189
GSM1214175
GSM1214176
GSM1214177
Treg Muscle
Tconv Muscle
expression value
GSE55705 / GPL8321 / 1427038_at
GSM1342099
GSM1342100
GSM1342097
GSM1342098
Tumor-Treg
Tumor-Teff
expression value
C
OCR
Naive CD4+
Naive CD4+ Fem.
Act. CD4+
Naive CD8+
Treg
Ep.MEChi
Penk CDS
Other CDS
TSS
CpG Island

456 **Figure 4**

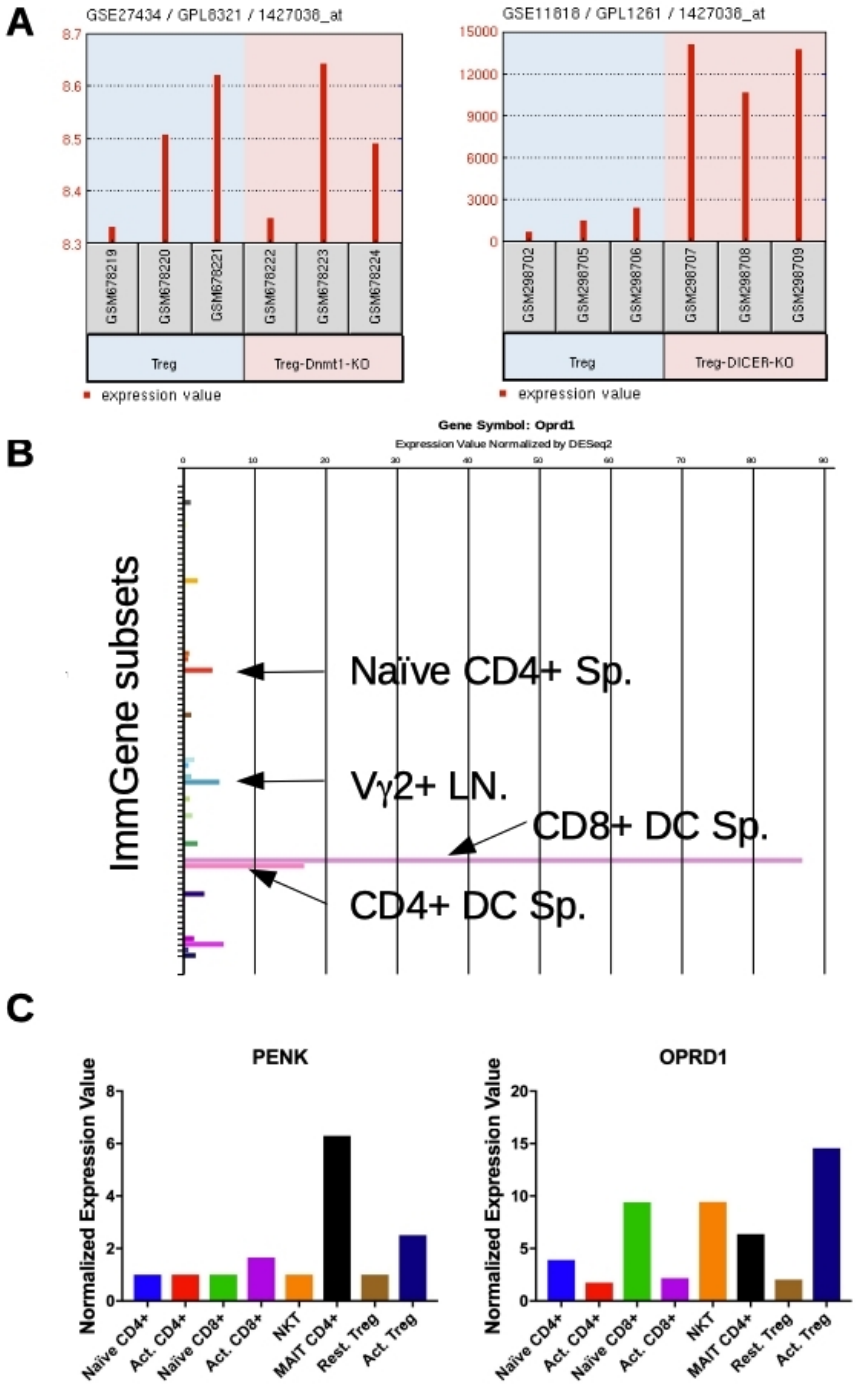

A
GSE27434 / GPL6321 / 1427036_at
8.7
8.6
8.5
8.4
8.3
GSM678219
GSM678220
GSM678221
GSM678222
GSM678223
GSM678224
Treg
Treg-Dnmt1-KO
expression value
GSE11818 / GPL1261 / 1427036_at
15000
12000
9000
6000
3000
0
GSM298702
GSM298705
GSM298706
GSM298707
GSM298708
GSM298709
Treg
Treg-DICER-KO
expression value
B
Gene Symbol: Oprd1
Expression Value Normalized by DESeq2
ImmGene subsets
Naïve CD4+ Sp.
Vγ2+ LN.
CD8+ DC Sp.
CD4+ DC Sp.
C
PENK
Normalized Expression Value
Naïve CD4+
Act. CD4+
Naïve CD8+
Act. CD8+
NKT
MAIT CD4+
Rest. Treg
Act. Treg
OPRD1
Normalized Expression Value
Naïve CD4+
Act. CD4+
Naïve CD8+
Act. CD8+
NKT
MAIT CD4+
Rest. Treg
Act. Treg

Figure 5