Deciphering the mechanisms by which FoxP3 regulates treg identity : an analysis of naturally occurring IPEX mutations

Juliette Leon

Université Paris Cité

École doctorale Hématologie, Oncogenèse et Biothérapies (ED 561)

Laboratoire IHU Imagine-Institut des Maladies Génétiques (UMR-S 1163)

Deciphering the Mechanisms by which FoxP3 Regulates Treg Identity: an Analysis of Naturally Occurring IPEX Mutations

Par Juliette LEON

Thèse de doctorat en biothérapies et biotechnologies

Dirigée par Isabelle ANDRÉ
Et co-dirigée par Christophe BENOIST

Présentée et soutenue publiquement le 09 octobre 2023

Devant un jury composé de :

- Carole GUILLONNEAU, DR, Université de Nantes, Rapportrice
- David KLATZMANN, PU-PH, Sorbonne Université, Rapporteur
- Benoît SALOMON, DR, Sorbonne Université, Examinateur
- Isabelle ANDRÉ, DR, Université Paris Cité, Directrice de thèse
- Christophe BENOIST, Full Professor, Harvard University, Co-directeur de thèse

Titre : Caractérisation des mécanismes par lesquels FoxP3 régule l'identité des Tregs : une analyse structure-fonction des mutations IPEX

Résumé: Le facteur de transcription (FT) FoxP3 orchestre le développement et la fonction des cellules T régulatrices (Tregs), responsables de l'homéostasie immunitaire et du contrôle des réponses immunitaires aberrantes. Les mutations de FoxP3 entraînent une dysfonction des Tregs, à l'origine du syndrome IPEX (Immune dysregulation, Polyendocrinopathy, Enteropathy, X-linked), une maladie auto-immune souvent fatale, dont le phénotype est particulièrement sévère lorsque le domaine de liaison à l'ADN de FoxP3 (FKHD) est muté. Cependant, les mécanismes précis par lesquels FoxP3 maintient l'identité des Tregs restent méconnus. Notamment, il existe une importante variabilité phénotypique interindividuelle, même parmi les membres d'une même famille portant la même mutation. Cette discordance entre génotype et phénotype complique notre compréhension de FoxP3. Ainsi, notre étude cherche à élucider la relation structure-fonction de FoxP3, en analysant les effets moléculaires de mutations "naturelles" associées à l'IPEX, à la fois chez l'humain et dans des conditions contrôlées chez la souris. Tout d'abord, les profils cytométriques et transcriptomiques des cellules T CD4+ sanguines de 12 patients atteints d'IPEX ont été analysés, à la fois en population et en cellule unique. Une signature monomorphe affectant l'ensemble des CD4+ a compliqué l'identification des effets spécifiques à chaque mutation. Cette signature était extrinsèque aux Tregs, absente chez les mères hétérozygotes de patients IPEX, où les Tregs WT exercent une suppression dominante. Secondairement, six mutations faux-sens, provenant de la précédente cohorte IPEX et réparties sur le locus de FoxP3, ont été introduites chez les souris B6. Cette approche permet d'étudier ces mutations dans les conditions hémizygotes et hétérozygotes, avec un nombre significatif de répétitions, et sans facteurs de confusion génétiques et environnementaux. Notre analyse a révélé deux classes de mutations. L'une, R337Q, située dans le FKHD, induit une lymphoprolifération et une infiltration multi-organe chez les mâles hémizygotes, comparables à celle des souris dépourvues de FoxP3 (KO), mais retardée. Les mutations dans d'autres domaines, en revanche, n'ont pas montré de phénotype clair à l'état stable, mais l'induction d'inflammation tissulaire ou le croisement avec des allèles NOD ont révélé des maladies spécifiques (dermatite, colite, diabète). Ces mutations avaient des impacts subtils mais distincts sur la signature ARN dépendante de FoxP3 et sur l'accessibilité de la chromatine, suggérant une interférence dans l'interaction de FoxP3 avec des cofacteurs spécifiques. Enfin, en utilisant un modèle de souris hétérozygotes FoxP3 KO, nous avons analysé l'impact global de FoxP3 sur la chromatine et sur le réseau de FT définissant les Tregs, et découvert que FoxP3 influence ce réseau à travers des rôles répressifs et activateurs. Nous avons aussi identifié une sous-population de Tregs, les Tregs RORγ+, qui semble fonctionner indépendamment de FoxP3. Ainsi, notre étude suggère que la physiopathologie de chaque mutation faux-sens dépend initialement de la position de la mutation, comme illustré par le fait qu'une mutation impactant directement le FKHD est suffisante pour induire la maladie. Cependant, les mutations non-FKHD, probablement perturbant les interactions FoxP3-cofacteurs, nécessitent une combinaison de facteurs génétiques et environnementaux pour se manifester cliniquement. Cela fournit une explication claire pour l'hétérogénéité clinique observée dans l'IPEX.

Mots clefs : **Lymphocytes T régulateurs, FoxP3, Structure-fonction, IPEX**

Title: Deciphering the Mechanisms by Which FoxP3 Regulates Treg Identity: an Analysis of Naturally Occurring IPEX Mutations

Abstract: The transcription factor (TF) FoxP3 plays a crucial role in the immune system by directing the development and function of regulatory T cells (Tregs). These cells maintain immune homeostasis and prevent autoimmunity by suppressing aberrant immune responses. Mutations in FoxP3 can disrupt Treg function, leading to the fatal multisystemic autoimmune disorder IPEX (Immune dysregulation, Polyendocrinopathy, Enteropathy, X-linked). The phenotype is especially severe when the FoxP3 DNA-binding domain (FKHD) is disturbed. However, the exact mechanisms by which FoxP3 maintains Treg identity remain elusive. Significant inter-individual variability is observed in the disease manifestation, even among family members carrying the same mutation. This discrepancy between genotype-and phenotype adds a level of complexity to our understanding of FoxP3's operation. With this in mind, our study aims to unravel the structure-function relationship of FoxP3, by characterizing the molecular effects of naturally occurring FoxP3 mutations associated with IPEX, both in humans and under controlled conditions in mice. First, the flow cytometric and transcriptomic profiles of blood CD4+ T cells from a cohort of 12 IPEX patients, were analyzed at both the population and the single-cell levels. A strong monomorphic signature affecting the whole CD4+ compartment complicated the examination of mutation-specific transcriptional effects in Treg-like cells. This signature appeared to be cell-extrinsic, disappearing in CD4+ from two heterozygous mothers of IPEX patients, where normal Treg cells exerted dominant suppression. Then, as a second strategy, 6 missense mutations from this cohort, spread across the FoxP3 protein, were engineered into B6 mice. This approach allowed us to study these mutations under both the hemizygous (male) and the heterozygous (female) conditions, with a significant number of replicates, and without genetic and environmental confounders. Two classes of mutations emerged from the combined immunologic and genomic analyses. A mutation in the FKHD, R337Q, showed a similar lymphoproliferation and multi-organ infiltration in hemizygous males as complete FoxP3 knockouts (KO), but delayed by months. Conversely, mutations in other domains didn't show clear phenotype at steady state, but immunological challenges or intercrossing with NOD alleles revealed mutation-specific diseases (dermatitis, colitis, diabetes). Our analyses of transcriptomics and chromatin accessibility in a heterozygous setting indicated that these mutations had subtle but distinct intrinsic impacts on the FoxP3-dependent signature and on the chromatin accessibility, in comparison to the strong effect of the FKHD mutation. This suggests that different tissue-specific cofactors might be involved in the function of each mutated version of FoxP3. Finally, using a complete FoxP3 KO heterozygous mouse model, we analyzed single-cell chromatin accessibility in both WT Tregs and Treg-like cells, to understand the overall impact of FoxP3 on Treg TF network. We found that FoxP3 influences Treg chromatin programs through both repressive and activating roles. Interestingly, a subset of Tregs, the RORγ+ Tregs, can function independently of FoxP3. Overall, our research proposes that disease pathogenesis due to specific FoxP3 domain disturbances primarily stems from the mutation's position, illustrated by the fact that mutation directly impacting the DNA-binding domain is sufficient to induce disease. However, non-FKHD mutations, likely interfering with FoxP3-cofactor interactions, also require a combination of genetic and environmental factors to clinically manifest. This provides a clear explanation for the clinical heterogeneity observed in IPEX.

Keywords**: Regulatory T cells, FoxP3, Structure-fonction, IPEX**

Remerciements

……

ABBREVIATIONS

Ag	Antigen
APC	Antigen Presenting Cell
Areg	Amphiregulin
ATAC-seq	Assay for Transposase-Accessible Chromatin Using Sequencing
aTreg	activated or effector regulatory T cells (similar to eTreg)
BACH2	BTB Domain And CNC Homolog 2
BCL6	B-cell lymphoma 6
BMC	Bone marrow chimera
CD25	Interleukin-2 receptor alpha chain
CNS	Conserved Non-coding DNA Sequences
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Cre-ERT	Tamoxifen-dependent Cre recombinase
CTLA-4	Cytotoxic T-Lymphocyte-Associated protein 4
DNA	Deoxyribonucleic Acid
DTR	Diphtheria Toxin Receptor
eTreg	effector or activated regulatory T cells (similar to aTreg)
FC	Fold Change
FKHD	Forkhead DNA binding domain
FOX	Forkhead Box
FOXP3/FoxP3	Forkhead Box P3
GFP	Green Fluorescent Protein
GRN	Genetic Regulatory Network
GVHD	Graft vs Host disease
HD	Healthy donor
HDR	Homology-directed repair
HLA	Human leukocyte antigens
HSCT	Hematopoietic Stem Cell Transplantation
IgE	Immunoglobulin E
IL	Interleukin
INF-γ	Interferon-γ
IPEX	Immune dysregulation, Polyendocrinopathy, Enteropathy, X-linked
IRES	Internal Ribosome Entry Site
IRF4	Interferon Regulatory Factor 4

KI	Knock In
KO	Knock Out
LOF	Loss of function
MHC	Major Histocompatibility Complex
mTOR	Mammalian Target of Rapamycin
NF-κB	Nuclear Factor-kappa B
NFAT	Nuclear Factor of Activated T cells
NGS	Next generation sequencing
NLS	Nuclear localization sequences
NOD	Non-Obese Diabetes mice
OCR	Open chromatin regions
PBMC	Peripheral blood mononuclear cells
PCA	Principal component analysis
PD-1	Programmed cell Death 1
pTreg	peripherally derived regulatory T lymphocytes
RNA	RiboNucleic Acid
RNA-seq	RNA sequencing
ROR-γ	RAR-related Orphan Receptor gamma
RUNX	Runt-related transcription factor
Sc	Single cell (e.g., scRNA-seq or scATAC-seq)
t-SNE	t-Distributed Stochastic Neighbor Embedding
Tconv	Conventional T cell
TCR	T Cell Receptor
TF	Transcription factor
TGF-β	Transforming growth factor-beta
Th	Lymphocytes T helper
TNFRSF	Tumor Necrosis Factor Superfamily of Receptor
Treg	Regulatory T cell
Treg-like	Regulatory T cell that has an active FoxP3 locus without expressing Functional FoxP3 protein (KO)
TSDR	Regulatory T-Cell-Specific Demethylated Region
TSS	Transcription start sites
tTreg	thymus derived Treg
UMAP	Uniform Manifold Approximation and Projection
WT	Wild Type

TABLE OF CONTENTS

Introduction

The Forkhead Box P3 (FoxP3) gene is vital to immune function as it encodes a transcription factor that guides the development and function of regulatory T cells (Tregs) (Fontenot, Gavin and Rudensky, 2003; Hori, Nomura and Sakaguchi, 2003; Fontenot, Rasmussen, Williams, *et al.*, 2005). Tregs, characterized by the expression of CD4, CD25, and the stable and high expression of the transcription factor (TF) FoxP3, play a crucial role in maintaining immune homeostasis and preventing autoimmunity by suppressing abnormal or excessive immune responses (Josefowicz, Lu and Rudensky, 2012).

Mutations in FoxP3 impair Treg function, leading to unregulated immune responses and the development of multisystemic autoimmunity, a condition known as IPEX (Immune dysregulation, Polyendocrinopathy, Enteropathy, X-linked) syndrome (Bennett *et al.*, 2001; Brunkow *et al.*, 2001). It has been proposed that FoxP3 functions through a combination of molecular interactions, epigenetic modifications, and gene regulation mechanisms. However, the exact mechanisms of FoxP3's operation within Tregs are still not fully understood and are subject to ongoing discussion.

In this introduction, we will start with a succinct overview of Tregs and their critical function in the immune system. This will be followed by a description of the FoxP3 structure, and a characterization of the clinical and genetic aspects of IPEX syndrome. Subsequently, we'll delve into the most recent discoveries associated with FoxP3's operation, as per current scientific understanding, and the mouse models that have been instrumental to this knowledge. We'll conclude this introduction by proposing our hypothesis.

Chapter I: General Overview of Treg Pathophysiology

1. Treg definition and differentiation (tTreg/pTreg)

Regulatory T cells (Tregs) are a specialized subpopulation of T cells that play a crucial role in maintaining immune homeostasis and tolerance to self-antigens, thereby preventing autoimmunity. These CD4+ T cells were initially characterized by the stable and high expression of the transcription factor FoxP3, high levels of the interleukin-2 (IL-2) receptor alpha chain (CD25) as well as the expression of the cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) (Josefowicz, Lu and Rudensky, 2012).

Tregs can suppress the function of various immune cells, including CD4+ helper T cells (also called conventional CD4 T cells, Tconv), CD8+ cytotoxic T cells, B cells, and dendritic cells, thus controlling any forms of both innate and adaptive immune responses. For example, they restrain immune responses elicited by microbes or allergens to prevent associated tissue damage (Josefowicz, Lu and Rudensky, 2012); they maintain symbiotic relationships with the microbiota (Russler-Germain, Rengarajan and Hsieh, 2017; Ramanan *et al.*, 2023) and they have been shown to play a key role in tumor progression (Kumagai *et al.*, 2020; Sakaguchi *et al.*, 2020). They achieve this through a variety of mechanisms - that we will describe in more detail below - where their T-cell receptor (TCR) signal as well as FoxP3 expression play important roles (Dikiy and Rudensky, 2023).

Tregs can be classified into two main categories based on their origin **(Figure 1).**

(1) The majority of Tregs come from the thymus as a functionally distinct and mature population from CD4+CD8+ thymocytes. During their development, intermediate-affinity interactions with thymic self-peptide–MHC ligands associated with costimulatory signals from CD28 (Salomon *et al.*, 2000; Lio *et al.*, 2010) select T cells to the Treg lineage. This positive selection leads to the induction of FoxP3 and CD25 expression in a two-step process, recently individualized into two distinct paths (Burchill *et al.*, 2008; Lio and Hsieh, 2008; Owen *et al.*, 2019), and ultimately to the differentiation of functionally distinct and mature antigen-specific Tregs, also called **thymus-derived Tregs (tTregs)**. These tTregs are functionally stable and have broad antigen specificity, mainly to self-antigens (Benoist and Mathis, 2012; Josefowicz, Lu and Rudensky, 2012; Lee *et al.*, 2012). Studies analyzing TCR repertoires in unchallenged mice have demonstrated significant

parallels between the Tregs located in the thymus and those in peripheral lymphoid organs (Hsieh *et al.*, 2006). This finding suggests that thymus-derived Tregs (tTregs) form the majority of these populations.

(2) Some naïve CD4+ T cells differentiate into Treg cells in the periphery, outside the thymus, particularly in the intestinal mucosa(Ramanan *et al.*, 2023). The differentiation is secondary to a variety of conditions, but more typically to a response to antigen stimulation in a particular cytokine environment (transforming growth factor-beta (TGF-β) is particularly important) (Lathrop *et al.*, 2011; Ramanan *et al.*, 2023) .These **peripherally-induced Treg (pTreg)** cells are also functionally stable but have more limited antigen specificity (Lathrop *et al.*, 2011; Solomon and Hsieh, 2016). Coherent with their place of differentiation, they are particularly important in controlling immune responses at mucosal surfaces. In the gut, they have been shown to be crucial to maintain tolerance to food antigens and commensal bacteria (Lathrop *et al.*, 2011; Russler-Germain, Rengarajan and Hsieh, 2017). Two main subsets have been described in the gut depending on the expression of TF: RORγ^{+} or HELIOS^{+} Tregs (Sefik *et al.*, 2015). Outside of the gut, they also play a role in maintaining maternal-fetal tolerance during pregnancy (Samstein, Josefowicz, *et al.*, 2012).

While both tTregs and pTregs express FoxP3 and have regulatory functions, they exhibit some difference in their gene expression profiles (Feuerer *et al.*, 2010) and can preferentially control different types of immune responses (Josefowicz *et al.*, 2012).

Tregs can also be induced *in vitro* from Tconvs in the presence of IL-2 and TGF-β (Chen *et al.*, 2003; Fantini *et al.*, 2004). However, these **induced Treg (iTreg)** cells are not functionally stable, have limited antigen specificity, and their true relevance in vivo are still debated (Benoist and Mathis, 2012).

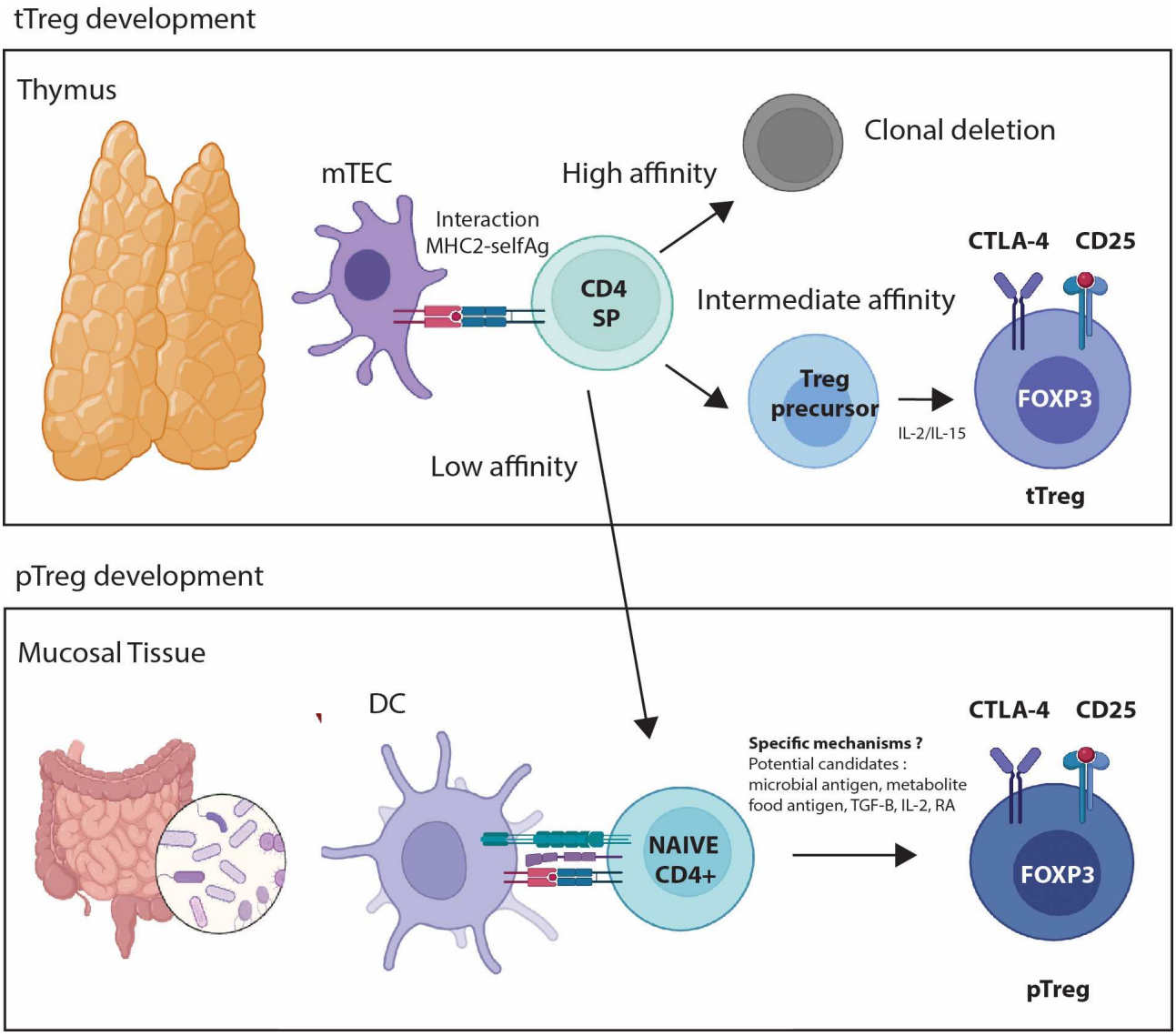


Figure 1: Regulatory T (Treg) development: tTreg and pTreg
DC, dendritic cell; mTEC, Medullary thymic epithelial cells; MHC2, major histocompatibility complex 2; pTreg, peripherally induced Treg; selfAg, self-antigen; tTreg, thymus-derived Treg.

Furthermore, CD8+FOXP3+ regulatory T cells have been described, constituting a much smaller proportion of the T cell compartment (~0.1% and ~0.3% of blood CD8+ T cells in mice in human respectively) (Churlaud *et al.*, 2015; Liston and Aloulou, 2022). They share transcriptional features with CD4+FOXP3+ T cells (Agle *et al.*, 2018; Ménoret *et al.*, 2023) and they have been shown to be suppressive in models of GVHD, lupus and transplantation (Singh *et al.*, 2007; Agle *et al.*, 2018; Picarda *et al.*, 2019). They are out of the scope of this study, so we won't describe them further.

Finally, some studies suggested the existence of regulatory T cell populations that are suppressing primarily through the production of IL-10 and that are FoxP3-independent, called Type 1 regulatory T cells (Tr1) (Groux *et al.*, 1997; Roncarolo *et al.*, 2014). However, these data come from few reports and the reality of their true belonging to the Treg lineage is debated.

2. Resting *versus* activated Treg cells

In opposite to Tconvs that are functionally naïve until they receive antigenic stimulation in the periphery and differentiate into antigen-specific effector T cells, Treg cells directly differentiate within the thymus (or the mucosal barrier) into functionally distinct and mature antigen-specific Tregs (Josefowicz, Lu and Rudensky, 2012). Once in the periphery, Tregs have been classified into two main categories depending on their activation profile. In mice, they are divided into naive-like $CD44^{lo}CD62L^{hi}$ Treg cells (or resting / rTreg) and effector-like $CD44^{hi}CD62L^{lo}$ Treg cells (or activated / aTreg) with a higher degree of activation and suppressive activity. The aTreg compartment has been also called as "effector Tregs". Equivalent states have been described in humans by the Sakaguchi lab: $CD45RA^{+}FoxP3^{lo}$ for the resting Treg cells, $CD45RA^{-}FoxP3^{hi}$ for the activated Treg cells (Miyara *et al.*, 2009).

3. Survival and functional maintenance of Treg cells

Once Tregs leave the thymus and enter the periphery, they still require continuous and balanced TCR signaling, in conjunction with IL-2, for their survival and function (Fontenot, Gavin and Rudensky, 2003; Levine *et al.*, 2014; Chinen *et al.*, 2016).

Whereas TCR signaling is indispensable for Foxp3 expression and for the maintained expression of the full Treg transcriptomic signature (cf. below), TCR signaling also influences the suppressive function of Tregs by enhancing their suppressive activity, facilitating their migration, and allowing them to adapt to the local tissue environment through the modulation of their cytokine production and the expression of tissue-homing receptors (Levine *et al.*, 2014; Tanaka *et al.*, 2023). TCR signaling is key to maintain Tregs in an effector state (aTreg) since it controls expression of key effector genes such as *Irf4, Nfatc1, Nr4a1, Bcl6, Lag3, Ccr8 (Levine et al., 2014)*.

One distinguishing feature of Treg cells is their lack of IL-2 production coupled with the constitutive expression of the high-affinity IL-2 receptor (CD25), setting them apart from activated Tconv cells. FoxP3 represses IL-2 transcription upon TCR stimulation. As a result, Treg cells are highly

dependent on exogenous IL-2 for their survival (D'Cruz and Klein, 2005; Fontenot, Rasmussen, Gavin, *et al.*, 2005) and that CD25-dependent activation of the transcription factor STAT5 had an essential role in the suppressor function.

4. Treg transcriptional signature

One crucial criterion for categorizing a cell as a Treg is its distinct gene signature, often referred to as the "Treg signature," which is intrinsically linked to the expression of FoxP3 (cf. below in the FoxP3 chapter). This signature was initially identified through contrasting Tregs with their conventional T cell (Tconv; Foxp3$^-$CD25$^-$) counterparts in both mice (Fontenot, Rasmussen, Williams, *et al.*, 2005; Sugimoto *et al.*, 2006; Hill *et al.*, 2007) and humans (Ferraro *et al.*, 2014).
The advent of single-cell analysis has recently refined this understanding, leading to the definition of a more specific subset of genes, termed the "core Treg genes". These core genes, which include *FoxP3*, *Ctla4*, *Il2ra* and several members of the TNFR (Tumor Necrosis Factor Receptor) superfamily, provide today the most accurate characterization of a Treg (Zemmour *et al.*, 2018).

5. Treg stability and diversity

Their stability, or the ability to maintain their suppressive phenotype and their specific gene expression profile even under inflammatory conditions, is a vital feature, especially to avoid that they convert into effectors with a self-reactive TCR which might lead to dramatic consequences. This stability is achieved by several epigenetic modifications, including the demethylation of a key intronic element of FoxP3 CNS2 (Conserved non-coding sequence 2) (Zheng *et al.*, 2010; Feng *et al.*, 2014), which lock in the Treg-specific expression profile (cf. the roles of these epigenetic changes will be discussed in the following chapter on FoxP3). On the other hand, the local microenvironment, including the presence of certain cytokines or metabolites, can impact Treg stability and function. In particular, there has actually been a controversy in the early 2010s, about the Treg instability, where one hypothesis proposed that important part of autoreactive T cells in autoimmune disease may consist of "ex-Treg" cells that lose FoxP3 expression and their suppressive activity (Yang *et al.*, 2008; Duarte *et al.*, 2009; Zhou *et al.*, 2009; Bailey-Bucktrout and Bluestone, 2011). However, experiments using a lineage tracing with Tamoxifen-controlled Foxp3-CreERT transgene where the FoxP3+ cells are tagged only during a defined time frame, these tagged cells and their descendants were observed to maintain their stability, even in the presence

of persistent inflammation (Rubtsov *et al.*, 2010; Miyao *et al.*, 2012). Although the conclusion of this subject is that vast majority of Tregs likely remain Tregs, Tregs can lose FoxP3 and suppressive activity when mistreated (e.g., local IL-2 deprivation or dysfunctional FoxP3 as illustrated in this study (Š. Borna *et al.*, 2022)).

Although Tregs are stable, it does not imply that they are inflexible. In response to specific circumstances, such as during an immune reaction or within certain tissue environments, Tregs have the capacity to adjust by expressing unique transcription factors, cytokines, and surface markers (Feuerer, Hill, *et al.*, 2009). Indeed, Treg cells can even secrete proinflammatory cytokines such as IFN-γ in contexts of microbial challenge (Oldenhove *et al.*, 2009; Gocher-Demske *et al.*, 2023). Interestingly, these programs appear to be determined in Treg cells by the same transcription factors that are central to the differentiated functions of the Tconv cells they regulate. For instance, Treg cells expressing T-bet optimally suppress inflammatory Th1 (T helper 1) responses (Koch *et al.*, 2009; Levine *et al.*, 2017) and the ones expressing Irf4, the Th2 responses (Zheng *et al.*, 2009). Then, effector Tregs have been subclassed depending on co-expression of TFs specific to Th cell lineages as shown in the **Figure 2** below (T-bet: Th1, IRF4: Th2, STAT3: Th17, and BCL6 (B-cell lymphoma 6): Tfh)

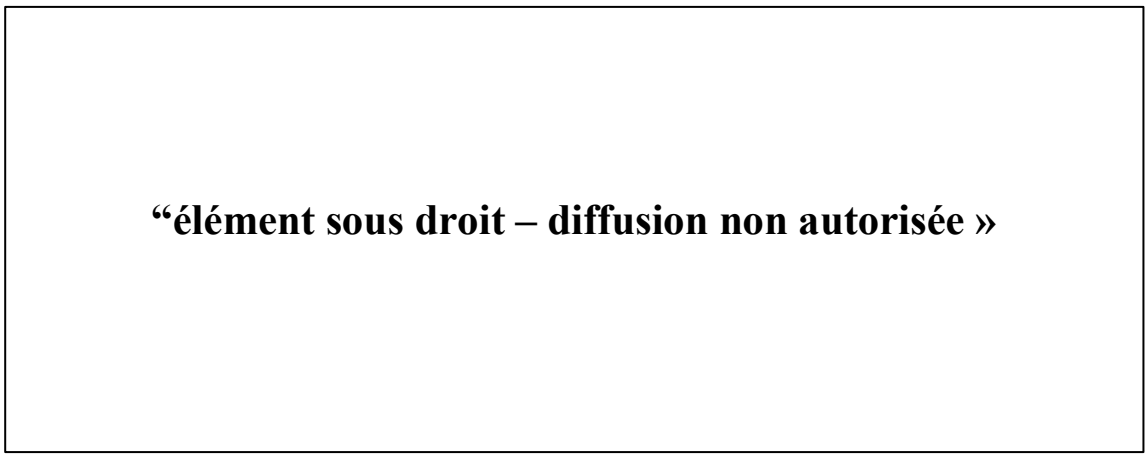


Figure 2. Regulatory T (Treg) cell heterogeneity and suppression of distinct classes of the immune response, from (Josefowicz, Lu and Rudensky, 2012).
Treg upregulate (tissue-specific) transcription factors in response to different environmental stimuli.

Finally, more recently, single cell analysis of lymphoid and non-lymphoid tissues has revealed a more blurry picture than this categorical classification where the TCR signal, associated to specific genes

encoding Treg effector functions, was driving Treg segregation (DiSpirito *et al.*, 2018; Zemmour *et al.*, 2018). Importantly, human and mouse Tregs proved organized along comparable poles.

6. Treg functions

Initially classified as suppressor cells primarily for controlling autoimmunity, Treg cells have since been revealed over the past twenty years to exhibit a more expansive role. They have diverse functions in regulating multiple aspects of immunity and physiological processes, including restraining T helper cell responses, sustaining metabolic equilibrium, safeguarding intestinal stability, and enabling tissue repair (Josefowicz, Lu and Rudensky, 2012; Muñoz-Rojas and Mathis, 2021; Ramanan *et al.*, 2023). Depending on the type of inflammation or affected tissue, Treg cells modulate both adaptive and innate immune responses, and even non-immune cells like adipocytes and fibroblasts.

One crucial insight about Treg function is that their immunosuppressive and tissue-repair capabilities aren't solely reliant on a single effector molecule. Rather, they implement an array of distinct mechanisms targeting a variety of immune and non-immune cells. Interestingly, many effector molecules produced by Treg cells are also found in activated conventional T cells and various innate immune cells. This suggests that the unique, non-redundant immunosuppressive functionality of Treg cells could be more about their specific features rather than their produced molecules. The Rudensky lab's recent study offers an in-depth review of these mechanisms (Dikiy and Rudensky, 2023).

One of these features, for example, is the role of specific self-antigen recognition in guiding Treg cell function. In essence, Tregs use their TCR to find and regulate cells needing suppression. Regardless of whether the effector functions of Treg cells are TCR-dependent or -independent, they require cognate antigen stimulation to be activated and guided towards the correct niches and targets. Interestingly, Treg suppression can occur without the need for matching epitopes on the suppressed T cells, as demonstrated *in vitro (Takahashi et al., 2000)* , or by the dominant protective effect exerted by monospecific Tregs in a polyclonal autoimmune setting (Tarbell *et al.*, 2004).

6.1 Mechanisms of Treg suppression

They employ multiple mechanisms to suppress or downregulate the immune response, and these mechanisms can be broadly categorized into five main types:

1. **Cytokine-Dependent Suppression:** Tregs can produce immunosuppressive cytokines that inhibit the function of effector T cells and other immune cells. These cytokines include interleukin-10 (IL-10) (Rubtsov *et al.*, 2008), transforming growth factor-beta (TGF-β) (Li, Wan and Flavell, 2007), and interleukin-35 (IL-35) (Collison *et al.*, 2007). Tregs have also been shown to secrete granzyme B, which directly triggers the apoptosis of CD8+ T cells, as reported by Grossman et al. in 2004 and Noelle David in 2005. However, this function requires further characterization as there is a lack of additional demonstrations to support it, and more recent literature seems to overlook this aspect.
2. **Cell Contact-Dependent Suppression:** Direct cell-cell contact between Tregs and other immune cells can lead to suppression, in particular through CTLA-4. Tregs express high levels of CTLA-4, which can bind to co-stimulatory molecules (CD80 and CD86) on antigen-presenting cells (APCs) with higher affinity than the co-stimulatory receptor CD28 on effector T cells. By outcompeting CD28 for these ligands, CTLA-4 can inhibit the activation of effector T cells (Wing *et al.*, 2008).
3. **Metabolic Disruption:** Tregs can interfere with the metabolism of other immune cells. For instance, Tregs express high levels of CD25 and can consume IL-2 in the local environment, depriving other immune cells for this crucial molecule (Chinen *et al.*, 2016). In addition, Tregs can express the ectoenzymes CD39 and CD73, which can convert ATP, a pro-inflammatory mediator, into adenosine, an anti-inflammatory molecule, further contributing to immune suppression (Deaglio *et al.*, 2007)**.**
4. **Modulation of Dendritic Cell Function:** Dendritic cells are key antigen-presenting cells that are necessary for the activation of effector T cells. Tregs can interact with dendritic cells (for example through CTLA-4 or LAG-3 (Huang *et al.*, 2004)), and downregulate the expression of co-stimulatory molecules and the production of pro-inflammatory cytokines, thereby suppressing the activation of effector T cells.

6.2. Tissue Tregs

Recent studies have identified a subset of T regulatory cells, known as "Tissue Tregs," that have adapted to thrive and function within their residing non-lymphoid tissues. They play a crucial role in tissue regeneration, maintenance, and repair, and have been identified in various tissues, including the adipose tissue (Feuerer, Herrero, *et al.*, 2009), skeletal muscle (Burzyn *et al.*, 2013), skin (Rosenblum *et al.*, 2011), colonic lamina propria (Schiering *et al.*, 2014), liver (Delacher *et al.*, 2017). Their diverse roles and interactions with the local tissue cells and microenvironment have been excellently reviewed by Munoz et al. (Muñoz-Rojas and Mathis, 2021). Tissue Tregs carry out various non-immunological roles that depend on the specific tissue context. Here are two examples of their pleiotropic tissue functions:

1. **Tissue maintenance and local metabolism:** Tregs can support tissue maintenance by mitigating chronic inflammation and modulating local metabolism. For example, in adipose tissue, Tregs help maintain metabolic function by controlling inflammation caused by obesity (Feuerer, Herrero, *et al.*, 2009).
2. **Tissue repair and regeneration:** Following tissue injury, Tregs can facilitate the repair process by suppressing excessive immune responses that can lead to tissue damage, and by promoting the resolution of inflammation. They can also directly promote tissue repair by producing growth factors like amphiregulin (Areg) (Burzyn *et al.*, 2013). Finally, Tregs have been shown to interact with stem cells in various tissues, potentially influencing tissue regeneration by modulating stem cell function (Ali *et al.*, 2017).

Tissues Tregs display significant heterogeneity, which largely depends on the specific type of tissue they inhabit. However, most of these cells exhibit a common, or "pan-tissue," gene signature (Muñoz-Rojas and Mathis, 2021). This includes genes typically associated with effector Tregs and those likely to be involved in the accumulation and general functionality of tissue Tregs – suggesting that Tregs constantly patrol these locations to maintain homeostasis. Examples of these genes are ST2 (*Il1rl1*), a range of chemokine receptors (*Ccr1, Ccr2, Ccr8, Ccrl2,* and *Cxcr6*), effector molecules (*Il10*, *Ebi3, Areg, Gzmb*), and co-inhibitory molecules (such as *Ctla4* and *Cd274*). Regarding transcription factors, Irf4 and Bach2 have been identified as crucial for entry into the tissue Treg cell precursor pool in lymphoid organs (Delacher *et al.*, 2017, 2020).

Chapter II: FoxP3 and his Structure

1. Forkhead box family

FoxP3 is part of the Forkhead box (FOX) family. This family of transcription factors is characterized by an evolutionarily conserved "forkhead" or "winged-helix" DNA-binding domain (Dai *et al.*, 2021). The name for this family was generated when the first member was discovered in Drosophila (fruit fly) in the late 80s, causing ectopic head structures in the embryos when mutated (Weigel *et al.*, 1989). Since then, 50 forkhead proteins have been identified in humans, with 44 mouse orthologs (Jackson *et al.*, 2010). These are divided into 19 subgroups (FOXA to FOXS) based on sequence homology within and outside the forkhead domain. This forkhead DNA-binding domain (FKHD), highly-conserved, composed of approximately 100 amino acids, includes three alpha-helices (H1-3), three beta-sheets (S1-S3), and two 'wing' (or loops) regions (W1-2) (Jackson *et al.*, 2010; Benayoun, Caburet and Veitia, 2011; Dai *et al.*, 2021) **(Figure 3)**. The domain's arrangement usually follows an H1-S1-H2-H3-S2-W1-S3-W2 pattern. It recognizes a consensus sequence, the forkhead motif, TGTTTAC, primarily via the third helix (Dai *et al.*, 2021). Additionally, the FKHD contains two nuclear localization sequences (NLS) - one in the H1 and the other in W2 (Lopes *et al.*, 2006; Benayoun, Caburet and Veitia, 2011).

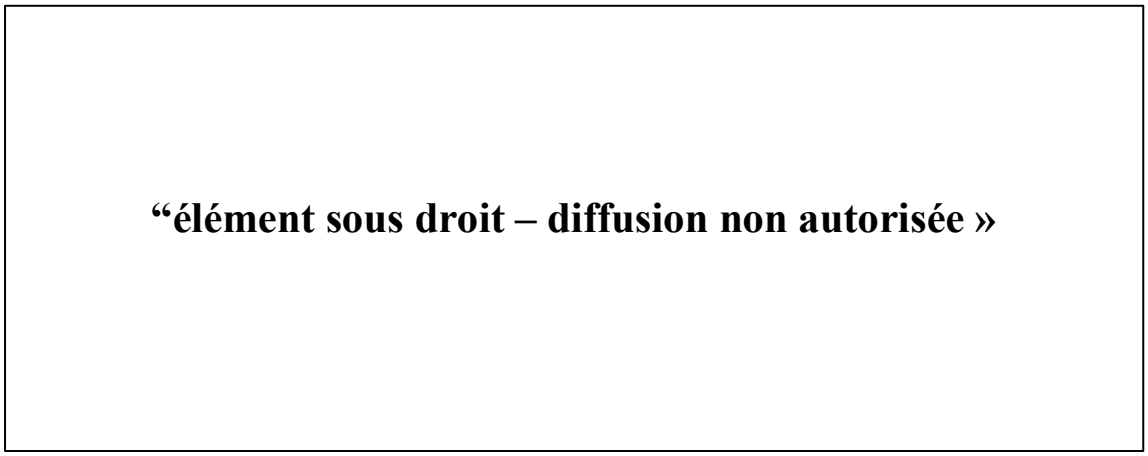


Figure 3. Structure of the Forkhead DNA binding domain from (Dai *et al.*, 2021)
Left, overall structure of forkhead DNA-binding domain (FKHD) where the color-scale showed the degree of conservation among species and among FOX protein. Right, topology of the FKHD, including three alpha-helices (H1-3), three beta-sheets (S1-S3), and two 'wing' regions (W1-2). A fourth helix (H4) is present in certain FOX proteins.

Despite the considerable similarity in their core sequences, FOX transcription factors govern a variety of pleiotropic cellular processes such as cell cycle progression, cell growth, differentiation, metabolism, senescence, survival, and apoptosis (Benayoun, Caburet and Veitia, 2011). They are notable for their ability to orchestrate gene expression in time and space. Interestingly, some FOX proteins like FOXA1 have unconventional ways to regulate gene expression, for instance, they bind to nucleosomes like histones, leading to chromatin de-condensation (Benayoun, Caburet and Veitia, 2011).

2. FoxP3 discovery and its key role in Treg function

FoxP3 discovery comes from the study of the 'scurfy' mouse from Oak Ridge Laboratory, the first X-linked disease reported in mice in 1986 (Russell, Russell and Gower, 1959). These mice exhibited conditions such as splenomegaly, skin lesions, premature death, and extensive lymphoid and myeloid infiltration across multiple organs (Godfrey, Wilkinson and Russell, 1991). These conditions were dependent on $CD4^+$ T cells (Blair *et al.*, 1994). Finally, in 2001, a genetic mapping study identified a 2-base pair insertion into an open reading frame in a critical region of the X Chromosome, later named *Foxp3* (Brunkow *et al.*, 2001). This mutation resulted in a truncated and presumably non-functional protein. Simultaneously, human patients exhibiting a severe, X-linked autoimmune syndrome, known as IPEX (Immune dysregulation, polyendocrinopathy, enteropathy, X-linked), were discovered to carry mutations in the corresponding human gene, *FOXP3* (Chatila et al., 2000; Bennett et al., 2001; Wildin et al., 2001). Hence, it became apparent that FoxP3 mutations led to autoimmune conditions in both mice and humans.

Moreover, few years earlier, in mid-90's, the identification of $CD4^+CD25^+$ T cells with suppressive abilities was the first step in defining Tregs (Sakaguchi *et al.*, 1995), but the challenge remained to differentiate genuine Tregs from effector cells that also upregulated CD25 upon activation. With FoxP3 discovery, researchers swiftly characterized Foxp3 expression in mouse $CD4^+CD25^+$ T cells and found that Foxp3 was specifically expressed in these cells but absent in naive or activated effector T cells (Fontenot, Gavin and Rudensky, 2003; Hori, Nomura and Sakaguchi, 2003; Khattri *et al.*, 2003). Further experiments from the same teams confirmed the central role of Foxp3 in Tregs differentiation and function. Bone marrow chimera experiments showed that Treg differentiation only occurred from Foxp3 WT cells and not from FoxP3 mutant cells (Fontenot, Gavin and Rudensky, 2003). Transferring Foxp3-expressing cells or Tconvs artificially expressing Foxp3 prevented autoimmune activation, signifying

Foxp3's sufficiency for suppressive activity (Fontenot, Gavin and Rudensky, 2003; Hori, Nomura and Sakaguchi, 2003). Autoimmunity was induced when Foxp3 was deleted in adults, proving that Foxp3 expression is required throughout life for Treg function (Kim, Rasmussen and Rudensky, 2007; Williams and Rudensky, 2007). **All these findings together demonstrated the non-redundant role of $CD4^{+}FoxP3^{+}$ Tregs in preventing autoimmunity.**

3. FoxP3 Structure

The *FOXP3* gene is located on the short arm of the X chromosome (Xp11.23) (Chatila *et al.*, 12 2000; Bennett *et al.*, 2001; Wildin *et al.*, 2001). FoxP3 contains 12 exons in the human genome, transcribed in a direction from centromere to telomere. The first exon, the 5' section of the 2nd exon, and the 3' part of the 12th exon are noncoding sequences. The complete open reading frame of FoxP3 in humans translates into a protein composed of 431 amino acids and has a molecular weight of approximately 47.25 kDa. FoxP3 is very well conserved between mice and human, with a sharing of 86.5% of amino acid sequence identity.

FoxP3 contains several functional domains (**Figure 4**) which are all required for optimal function (Zeng, Sollars and Belalcazar, 2011) :

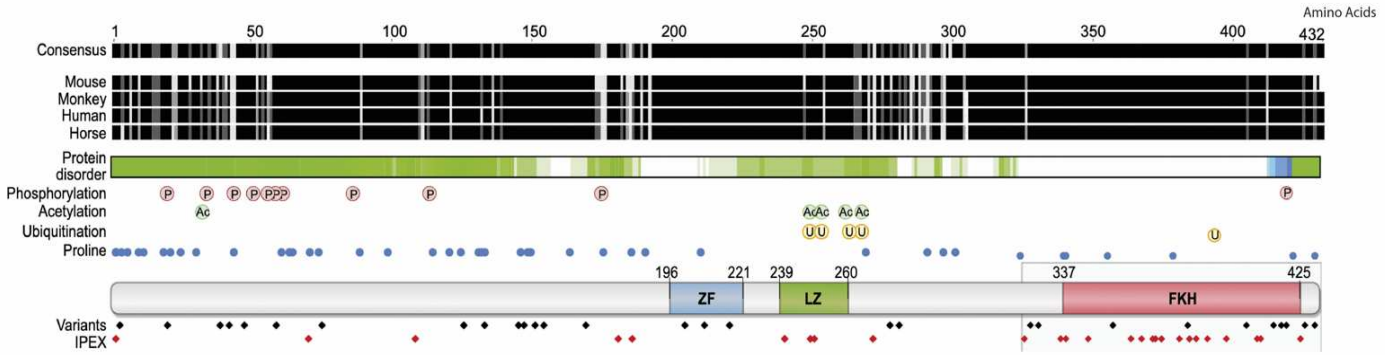


Figure 4. Overall structure of FoxP3 protein from (Kwon *et al.*, 2018)
From the top: conservation of FOXP3 protein across mammals; protein disorder (from D2 P2 browser: d2p2.pro/); posttranslational modifications and proline positions; ***protein domains****; and missense variants in human FOXP3 found in IPEX patients (red) or in population survey (black) (in 2018)*

1. **N-terminal Proline-Rich domain**: The N-terminal of FoxP3 contains a proline-rich region that appears disordered in computational and structural analyses (Andersen, Nissen and Betz, 2012; Kwon *et al.*, 2017), a flexible configuration which facilitates diverse interactions. This domain is responsible for the suppression of some cytokine genes, including the gene encoding IL-2 (Lopes *et al.*, 2006), thus sometimes referred as "the repressor domain"

2. **Zinc Finger domain** : unknown role but might likely play a role in DNA binding - However, his deletion didn't really impact Treg function significantly (Lopes *et al.*, 2006; Zeng, Sollars and Belalcazar, 2011).

3. **Leucine Zipper domain**: also referred to as a coiled-coil (CC), it enables FoxP3 to form anti-parallel homo- and hetero-dimers with other proteins, which has been identified to be in exercising his functions (Chae *et al.*, 2006; Lopes *et al.*, 2006; Song *et al.*, 2012).

4. **Forkhead domain (FKHD)**: already described above. The FKHD is responsible for DNA binding, essential for the transcription regulatory function of FoxP3 (Lopes *et al.*, 2006). It's also known to mediate interactions with other proteins, including nuclear factor of activated T-cells (NFAT) which it redirects from activating pro-inflammatory genes to inducing immune-suppressive genes. One of the distinctions of its FKHD is that it seems to exist predominantly as a domain-swapped dimer (distinct interface from the LZ), which simultaneously binds two distant FoxP3 DNA binding sites (Bandukwala *et al.*, 2011). This suggests that FoxP3 may play a role in long-range gene interactions (Chen *et al.*, 2015). Disruption of domain swapping interferes with FoxP3-mediated suppressor functions (Bandukwala *et al.*, 2011).

Each of these domains contributes to the overall function of FoxP3, providing interfaces for interaction with DNA, other proteins, and potentially RNA. These will be discussed in the following chapter in FoxP3 cofactors. In addition to these domains, post-translational modifications, such as phosphorylation, acetylation, and ubiquitination, can alter the structure of FoxP3, its stability and interaction with partners and subsequently affect its function (Deng *et al.*, 2019). However, we won't discuss further this post-transcriptional part since it is outside of the scope of this manuscript.

4. FoxP3 regulation and the conserved non-coding regions of FoxP3 locus

In both Treg differentiation (t or pTreg), a timely regulation of FoxP3 expression is essential. This is maintained by transcription factors downstream of TCR signaling that bind to the promoter and conserved non-coding sequence regions of the gene locus.

- **The promoter region**, located upstream of the FoxP3 gene, contains binding sites for numerous transcription factors, which play a key role in the regulation of FoxP3 expression, including NFAT, RELA, BACH2, NFIL3, FOXO1 and 3 (Trujillo-Ochoa, Kazemian and Afzali, 2023).
- **Conserved non-coding sequences** (CNS0 to CNS3) which have been found to play vital roles in the regulation of FoxP3 expression (Tone *et al.*, 2008; Zheng *et al.*, 2010).
    - **CNS0**: play a role in the early stages of tTreg cell differentiation, especially in the IL-2-dependent induction and maintenance of Foxp3 expression (Dikiy *et al.*, 2021; Kawakami *et al.*, 2021).
    - **CNS1**: This region is critical for the induction of FoxP3 expression in pTregs and the generation of inducible Tregs (iTregs) in the presence of TGF-β.
    - **CNS2:** Also known as the *Treg-specific demethylated region (TSDR)*, CNS2 is heavily methylated in Tconvs and demethylated in Tregs, which is critical for the **stable expression of FoxP3 and the overall Treg stability** (Feng *et al.*, 2014; Li *et al.*, 2014). It contains binding sites for several transcription factors that regulate FoxP3, including CREB/ATF, NFAT, RUNX1 and STAT5.
    - **CNS3**: This region is thought to act as a pioneer element that initiates the expression of FoxP3 during tTreg differentiation, possibly under the control of transcription factors like RELA and SMAD1 (Feng *et al.*, 2015; Kawakami *et al.*, 2021).

These regions are summarized in **Figure 5**.

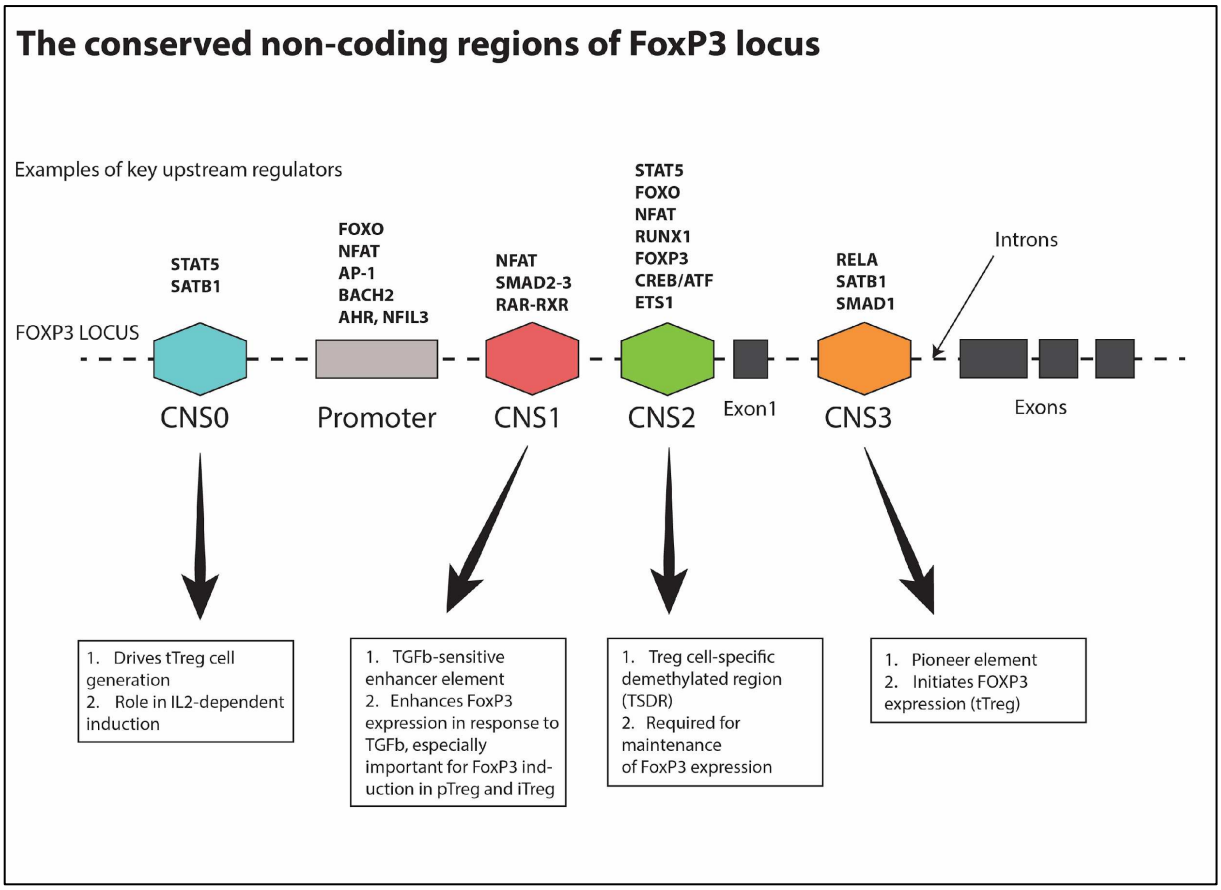


Figure 5: The conserved non-coding regions of FoxP3 locus: position, key upstream binding factor and functions.
Summarized data from a recent review on FoxP3 regulation (Trujillo-Ochoa, Kazemian and Afzali, 2023).

5. Alternative splicing of FoxP3

FoxP3 can undergo alternative splicing, resulting in different isoforms of the protein. To date, three major isoforms of human FoxP3 have been identified:

1. **FOXP3fl:** The full-length FOXP3 isoform is the most extensively studied (cf. data above) with 431 amino acids (AA) in humans and 429 in mice.
2. **FOXP3Δ2**: This isoform is predominant with FOXP3fl, and lacks exon 2 (71-105 AA). It has been shown to support Treg development but to not totally reach the suppressive capacity of FoxP3fl since it was found in IPEX patients (Frith *et al.*, 2019; Du *et al.*, 2022). This might be related to a defect in FoxP3 expression stability (Du *et al.*, 2022; Seitz *et al.*, 2022). FOXP3Δ2

lacks a nuclear export signal, which could impact its relocation into the cytoplasm. However, It has actually been shown recently that FOXP3FL and FOXP3Δ2 are both needed for optimal Treg function (Sato *et al.*, 2021).

3. **FOXP3Δ2/Δ7**: This isoform, less described, lacks exon 2 and the exon 7 (246-272 AA), and is considered to lack suppressive function (Joly *et al.*, 2015).

In contrast to human, only one major full-length isoform of Foxp3 has been described in mice. This difference should be kept in mind in the remainder of our study, since we are comparing Treg biology between human and mice.

Chapter III: Effect of FoxP3 deficiency: focus on IPEX

FoxP3's essential functions in the Treg lineage were discovered by the characterization of the mouse and human models of FoxP3 deficiency. We will first briefly describe the clinical course and genetic characteristics of the human model, the IPEX syndrome. Then we will discuss the genotype-phenotype puzzling relationship and the presence of Treg "wannabe" in heterozygous mothers.

1. Clinical characteristics of IPEX and current treatment

IPEX syndrome is a rare life-threatening X-linked recessive disorder, characterized by an array of clinical manifestations due to the immune dysregulation that results from the defect in Treg functions (Powell, Buist and Stenzel, 1982; Duclaux-Loras *et al.*, 2018; Gambineri *et al.*, 2018).

The primary symptoms typically present in the first weeks to months of life (Barzaghi *et al.*, 2018), but can vary significantly in severity (e.g. late-onset or more mild phenotype) and the systems they affect (e.g. absence of enteropathy) (Consonni, Ciullini Mannurita and Gambineri, 2021). The different clinical manifestations have been recently reviewed by several teams (Barzaghi *et al.*, 2018; Duclaux-Loras *et al.*, 2018; Gambineri *et al.*, 2018; Park *et al.*, 2020; Consonni, Ciullini Mannurita and Gambineri, 2021, **Figure 6**) and can be classified by system, in frequency order :

- **Gastrointestinal Manifestations:** enteropathy is the most frequent feature of IPEX syndrome (occurring in more than 90% of cases in most of series) and often one of the first signs of the disease. It typically presents as severe, intractable diarrhea leading to failure to thrive and malabsorption. The diarrhea is typically chronic and persistent, resistant to dietary changes. Histologically, it's characterized by villous atrophy, crypt hyperplasia, and lymphocytic infiltration in the small intestine, mimicking the changes seen in celiac disease (Barzaghi *et al.*, 2018; Duclaux-Loras *et al.*, 2018; Gambineri *et al.*, 2018; Consonni, Ciullini Mannurita and Gambineri, 2021).
- **Endocrine Manifestations:** type 1 diabetes mellitus is a common endocrine manifestation, often presenting in infancy, unusually early compared to type 1 diabetes in the general population. Other endocrine diseases like thyroiditis and adrenal insufficiency can also occur.

- **Dermatological Manifestations:** patients with IPEX syndrome often present with severe, often chronic, eczema. Other skin conditions such as psoriasis, alopecia, and pemphigoid have been reported.
- **Allergic Manifestations:** food allergy, elevated IgE levels and eosinophilia are common.
- **Hematologic manifestation:** autoimmune hemolytic anemia, thrombocytopenia, neutropenia can be described.
- **Other Manifestations:** other less common but severe manifestations include renal disease, typically presenting as nephrotic syndrome, gastritis, hepatitis and neurological symptoms like seizures and intellectual disability.

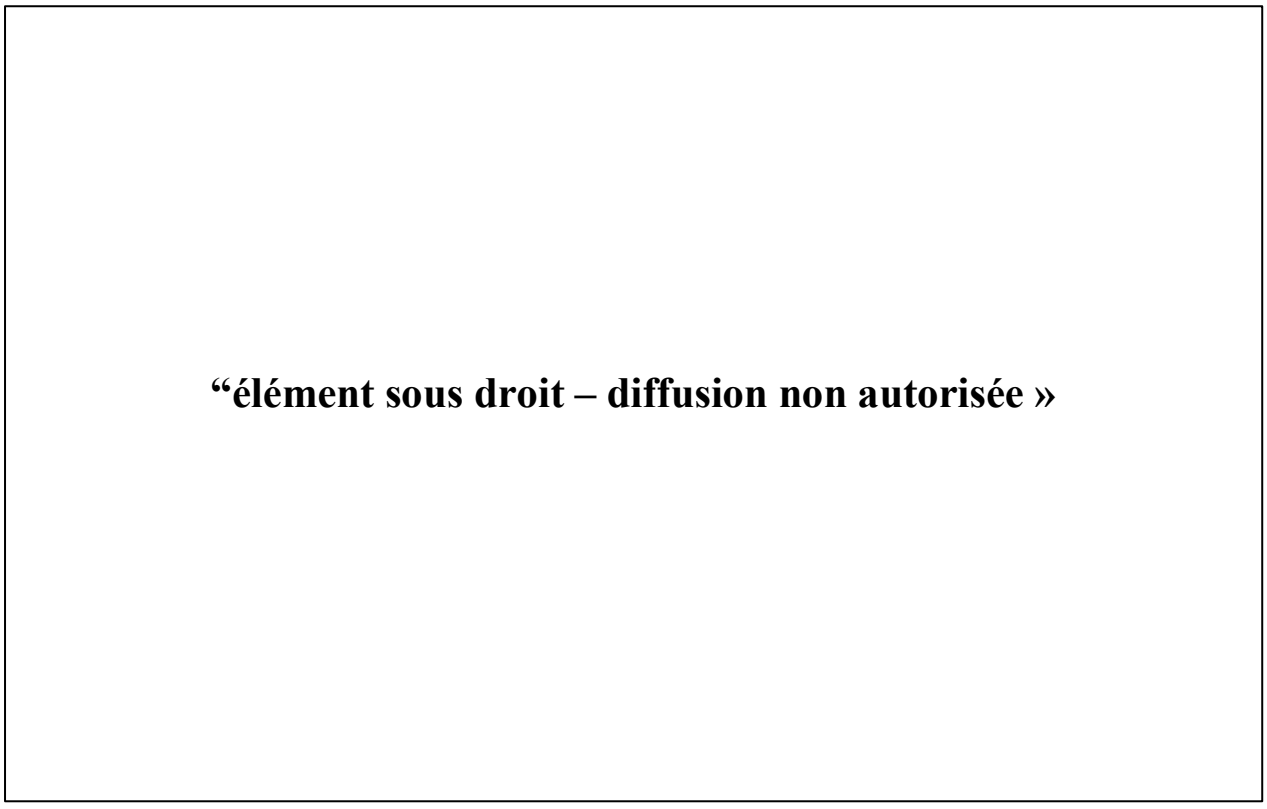


Figure 6: Typical and unusual clinical features in IPEX, from (Consonni, Ciullini Mannurita and Gambineri, 2021).
Percentages for typical features are based on the most recently published IPEX cohort (Barzaghi et al., 2018). Classical triad features are in bold. CIDP, Chronic Inflammatory Demyelinating Polyneuropathy; GI, Gastrointestinal.

With regard to physiopathology, Tregs from IPEX patients have typically been shown to be dysfunctional *in vitro*, which seems to correlate with a decrease in FoxP3 expression (Bacchetta *et al.*, 2006). However, these two characteristics are not always well-defined and some Tregs appear to function near normal Tregs *in vitro (Bacchetta et al., 2006; Bacchetta, Barzaghi and Roncarolo, 2018)*. The subsequent Treg deficiency leads to an overall CD4+ activation but also features like : **(1)** increase of pro-inflammatory cytokines in the plasma, **(2)** T- and B-cells dysfunction, **(3)** hypereosinophilia and hyper IgE, and **(4)** abnormalities in innate immunity not well yet characterized (Bacchetta *et al.*, 2006; Kinnunen *et al.*, 2013; Duclaux-Loras *et al.*, 2018; Gambineri *et al.*, 2018; Narula *et al.*, 2022).

The T-cell compartment exhibits at the protein level global activation and a polarization toward a type-2 inflammatory immune response (Chatila *et al.*, 12 2000; Nieves *et al.*, 2004; Bacchetta *et al.*, 2006). Indeed, numerous reports have outlined the crucial role of the Th2 biased CD4+ compartment in IPEX pathology. This has been corroborated in studies involving FoxP3 deficient mice (Van Gool *et al.*, 2019) and has been further underscored recently by the dramatic clinical effectiveness of anti-IL4/IL13 antibodies in alleviating the symptoms of an IPEX patient (Caruso *et al.*, 2022). Notably, it has recently been shown that IPEX patients present an increase in TSDR demethylated CD4+ T cells, suggesting the possible existence of pathogenic "ex-Tregs" (Narula *et al.*, 2022; Wyatt *et al.*, 2023). These are present even before the disease onset and their presence/proportion does not seem to depend on the phenotype severity.

Several autoantigens targeted by the humoral response have been characterized (Eriksson *et al.*, 2019; Hoshino *et al.*, 2019), including harmonin and villin (intestinal and tubular brush border), GAD-65 (pancreatic islet), IFNa (interferon) and some nuclear receptors (HNF4a, RXRa, and PPARg). Most of them showed predominant expression in tissues that are damaged in patients with IPEX, in particular in the intestinal epithelium, and they seem to be related to the specific clinical phenotype of each patient(Eriksson *et al.*, 2019; Hoshino *et al.*, 2019).

More recently, a serum macrophage derived-inflammatory signature has been found in IPEX patients (Narula *et al.*, 2022), suggesting a dysfunction of innate immunity as well.

Without treatment, the clinical course of IPEX syndrome is severe, with death typically occurring in infancy or early childhood due to severe malnutrition, and/or sepsis (Barzaghi *et al.*, 2018). Immunosuppressive therapy like corticosteroids are often used as first-line treatment to control symptoms, as well as Sirolimus, a mTOR inhibitor (Passerini *et al.*, 2020), and the calcineurin inhibitors

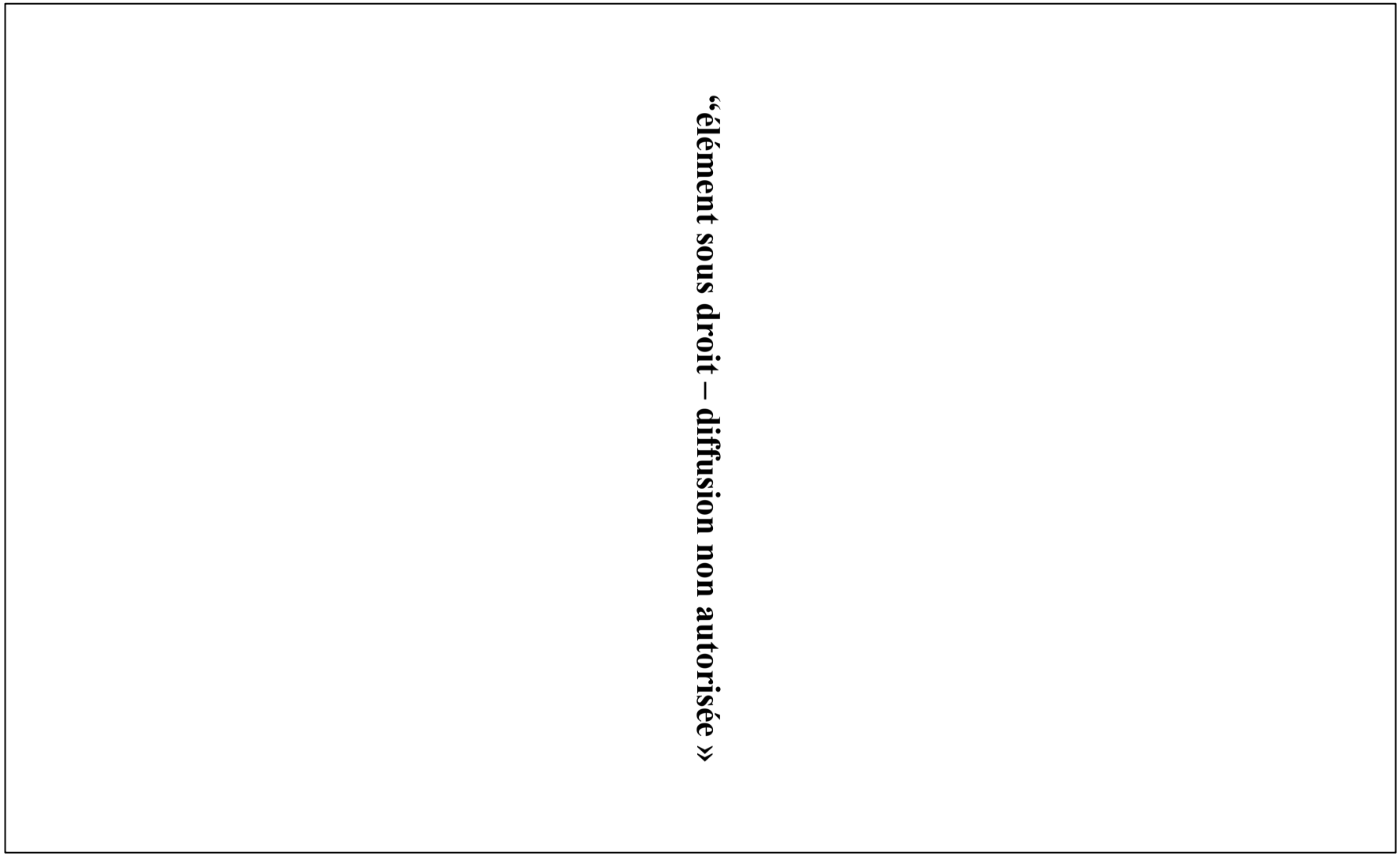
cyclosporine and tacrolimus (Barzaghi *et al.*, 2018; Duclaux-Loras *et al.*, 2018; Gambineri *et al.*, 2018). However, while these therapies may temporarily ameliorate some symptoms, they do not cure the disorder and long-term use can lead to significant side effects. The prognosis has improved substantially over the last decade with allogeneic hematopoietic stem cell transplantation (HSCT), which is the only curative treatment for IPEX syndrome, restoring a Treg compartment (Baud *et al.*, 2001; Barzaghi *et al.*, 2018). Nevertheless, HSCT is at risk and depends on finding a suitable donor. It can also lead to complications such as graft-versus-host disease (GVHD). While still at the experimental stage, gene therapy represents a promising future treatment in which the patient's own stem cells are modified *ex vivo* to correct the *FOXP3* mutation, and then reintroduced into the patient (Masiuk *et al.*, 2019; Goodwin *et al.*, 2020; S. Borna *et al.*, 2022). This would ideally restore normal Treg function without the risk of GVHD.

2. Genetics of IPEX and IPEX-like syndrome

By July 2023, 120 different pathogenic mutations had been described from 154 families, according to the Human Genome Mutation Database (HGMD®). These mutations are distributed throughout the gene and include a wide range of mutation types: missense mutations, nonsense mutations, insertions, deletions, splice site mutations, and complex rearrangements. These mutations can occur in different domains of the FOXP3 protein, but they are predominantly found in the FKHD. For example, of the 55 missense mutations reported so far, 65% of them are in the FKHD (AA 337-421), as stated in the HGMD®. These mutations can lead to a complete loss of FoxP3 function or a partial loss, depending on their nature and location. A comprehensive and up-to-date review of the different IPEX mutations reported in the literature was recently published by Park et al. (Park *et al.*, 2020), and their cohort is presented in the **figure 7** below.

Figure 7: Schematic representation of all reported FOXP3 mutations in IPEX patients from Park et al.

Figure and data from the more recent review from Park et al. Autoimmunity Rev. 2020



There are other genetic defects that can lead to IPEX-like phenotypes, often referred to as IPEX-like syndromes (Gambineri *et al.*, 2018). These conditions share many clinical features with IPEX syndrome - they are most of the time difficult to distinguish without the genotyping results - and are caused by mutations in different genes, which have been proven to be essential to Treg function (Cepika *et al.*, 2018). Mutated proteins include CD25 (IL2RA), CTLA4, LRBA (mutations in LRBA lead to lower levels of CTLA-4), STAT5B, BACH2, and ITCH.

3. An uncertain genotype-phenotype correlation

Efforts to establish specific correlations between FoxP3 mutations and the severity or variety of clinical manifestations, often referred to as genotype-phenotype correlations, have proved challenging. This difficulty is likely due to the disease infrequency and small sample size of studies conducted. Despite these challenges, a few general trends have begun to emerge (Gambineri *et al.*, 2008, 2018; Duclaux-Loras *et al.*, 2018; Park *et al.*, 2020; Consonni, Ciullini Mannurita and Gambineri, 2021):

1. Truncating mutations, including nonsense mutations, frameshifts or deletions that introduce a premature stop codon, generally result in a severe IPEX phenotype. This is probably because these mutations generate a truncated, unstable FOXP3 protein or an unstable *FOXP3* mRNA, or they may lack essential functional domains, notably the FKHD.
2. Missense mutations in the forkhead domain often result in a more severe IPEX phenotype.
3. Missense mutations in other domains: some FoxP3 mutations do not entirely eliminate FoxP3 function but instead cause a partial loss of function. These mutations can lead to a less severe or more varied clinical presentation.

While these trends suggest that more disruptive mutations tend to result in more severe disease, two important points should be highlighted:

- Identical FoxP3 mutations can lead to dramatically different phenotypes in patients. A notable example is the c.1150G>A (A384T) mutation, located in the FKHD, which has been reported in patients who survived past 10 years of age as well as in patients who died prematurely within the first year (Bacchetta, Barzaghi and Roncarolo, 2018).

- Unusual manifestations such as nephropathy (c.210+1G>A splicing mutation (Tsuda *et al.*, 2010; Barzaghi *et al.*, 2018)), hypogammaglobulinemia (c.694T>G (p.C232G) mutation (Okou *et al.*, 2014; Shamriz *et al.*, 2018)), or lymphoproliferation (c.1110G>A (p.M370I) (An *et al.*, 2011; Gambineri *et al.*, 2018)) have been consistently associated with specific FoxP3 mutations in different patients. Some mutations have been consistently associated with early death in multiple patients (c.319–320del and c.1189C>T (R397W)), with neonatal or fetal deaths (Xavier-da-Silva *et al.*, 2015).

The struggle to delineate clear genotype-phenotype correlations implies that many factors beyond the specific FoxP3 mutation may influence the clinical presentation. Environmental influences or the presence of modifier genes could contribute to the clinical phenotype. Moreover, the impact of some mutations may depend on tissue context or interactions with other proteins. However, these factors have yet to be clearly identified. In conclusion, while some genotype-phenotype correlations in IPEX syndrome can be observed, the relationship is intricate and not fully understood.

4. How can heterozygous mothers be a model to study the disease?

A unique aspect of the disease associated with Foxp3 deficiency, observable in both humans and mice, is that it exclusively impacts hemizygous males carrying the mutant X chromosome. The FoxP3 gene is located on the X chromosome. Thus, in heterozygous females, one of the X chromosomes (and the corresponding FoxP3 allele) is randomly inactivated during embryogenesis in each cell, a process known as X-chromosome inactivation. During thymic differentiation, half of the CD4+ progenitors express the WT allele (where the mutant allele is silenced), and the other half express the mutant allele (to the best

of our knowledge, FoxP3 doesn't escape this random X-chromosome inactivation) as shown in **Figure 8**.

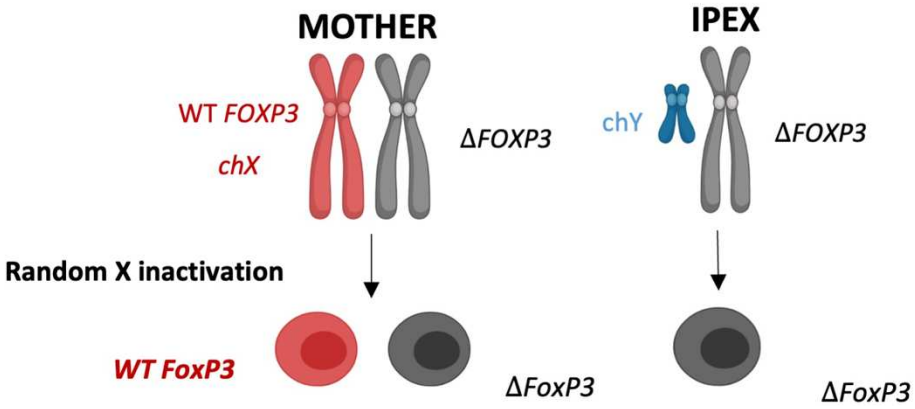


Figure 8. Random X-inactivation leads to two distinct populations of Tregs in heterozygous IPEX mothers.

These heterozygous females remain healthy (Tommasini *et al.*, 2002). One main hypothesis is that the presence of wild-type functional Tregs suppresses autoimmunity by controlling other Treg carrying the Foxp3 mutant allele (called “Treg-like” or “wannabe” Tregs). This hypothesis has been comforted by bone marrow chimera mice (BMC) (half WT stem cells, half FoxP3 KO stem cells) which also stay healthy (Fontenot, Gavin and Rudensky, 2003).

Thus, by using FoxP3-reporter-allele in mouse, it’s possible to study the two populations of Tregs independently (**Figure 9**) and this has led to the discovery that these "wannabe" Foxp3 deficient Tregs can develop almost as well as WT Treg (Gavin *et al.*, 2007). They are mostly in a resting state in the circulation, and they express intermediate levels of CD25, CD44, CTLA4, GITR and ICOS in comparison to WT Treg markers. Yet without functional FoxP3, these cells fail to fully differentiate into effector Tregs and do not acquire suppressive function. They also show instability with loss of FoxP3 transcription during homeostatic proliferation (Gavin *et al.*, 2007).

Thus, Foxp3 presents a unique TF in this regard and allows it to disassociate the genetic program that determines the cells from its functional properties. We will use this heterozygous model in our study to dissect the molecular impact of missense FoxP3 mutations, outside of confounding inflammation seen in hemizygous mice.

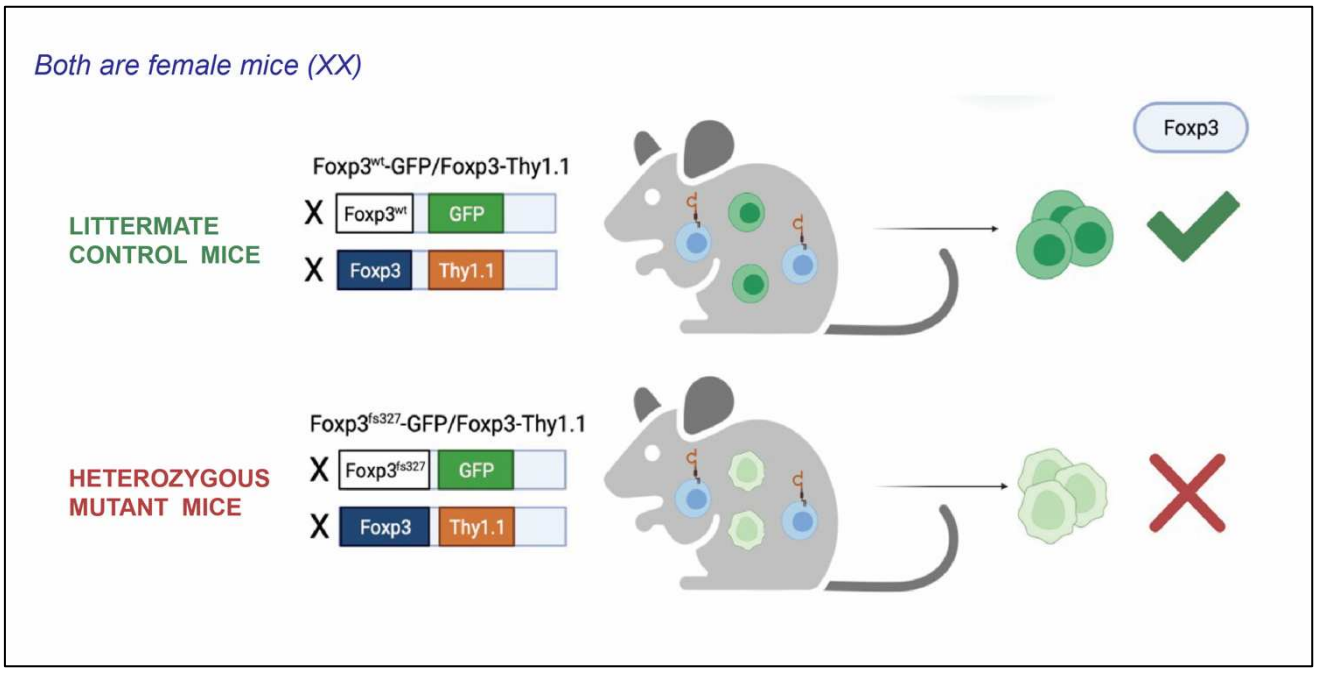


Figure 9. How to analyze *wannabe Treg* in the heterozygous female mice setting?
Both mice (control and mutant) have a similar WT "control" allele flagged by the Thy1.1 reporter (orange). They differ by their second allele, in both cases flagged by GFP (green) which is a KO FoxP3 allele (from Foxp3fs327-GFP mice (cf below in methods)) or WT FoxP3 allele (from Foxp3-ires-GFP mice). In the mutant mice, due to random X-inactivation, one population of Treg cells expresses WT FoxP3 protein (flagged by the Thy1.1 reporter), while another population of wannabe Treg cells expresses a Foxp3 allele with a full loss-of-function frameshift mutation whose expression is reported by GFP. The presence of functional Thy1.1+ Tregs prevents immune dysregulation, thus providing a well-controlled system for investigating FoxP3-intrinsic effects in wannabe Tregs.

5. Brief overview of the corresponding mouse model, Scurfy line

As described above in the FoxP3 discovery part, the **'scurfy' mouse** from Oak Ridge Laboratory (Russell, Russell and Gower, 1959) is the first mouse model of IPEX disease. This naturally occurring mutant strain carries a 2-base pair insertion in exon 9 of the Foxp3 gene leading to a frameshift mutation and a severely truncated protein that does not have FKHD. These mice have a dramatically reduced lifespan (about 16-40 days depending on the genetic background) and suffer from severe autoimmunity, including skin inflammation, lymphoproliferation (splenomegaly and lymph nodes hypertrophy), and multi-organ infiltration (Godfrey, Wilkinson and Russell, 1991).

Although they are the main model of IPEX disease, we can note several differences. The skin disease is more pronounced and diverse in Scurfy than in IPEX with typical scaliness and crusting of the ears and tail, reddening, and swelling of eyelids. The lymphoproliferation is very rare in IPEX, whereas is almost always present in Scurfy, one of the first signs of the disease (**Figure 10**). On the other hand, gut manifestation (such as gastritis, small intestine inflammation, and colitis) is rarely and very mild in Scurfy (Godfrey, Wilkinson and Russell, 1991). Classic pancreas insulitis is also not classically observed in Scurfy. Instead, perivascular inflammation of exocrine pancreas is severe. One reason of these differences might be the early death, which might prevent autoimmunity to develop yet in the gut or in the pancreas.

At the histological level, lymphoid (T and B cells) and myeloid (macrophages and granulocytes) infiltrations are found on the different lymphoid and non-lymphoid organs (Godfrey, Wilkinson and Russell, 1991; Chen, Benoist and Mathis, 2005). In skin and lymphoid organs, the infiltration is not related to the vascularization whereas in all other organs, the infiltration predominates in perivascular localization.

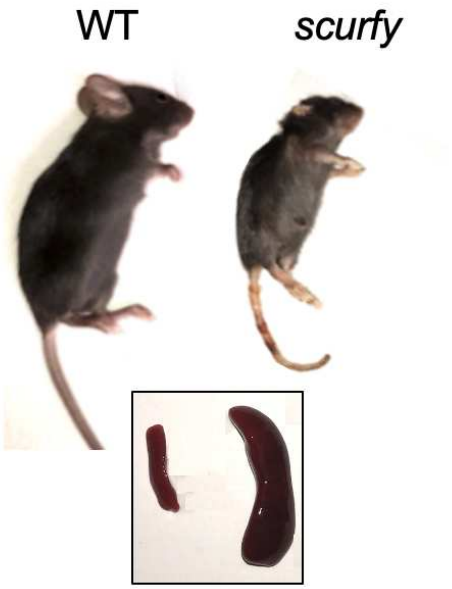


Figure 10. Comparison of WT and FOXP3-KO mice at 24-day-old and their respective spleen.

Chapter IV: How does Foxp3 work in Tregs?

1. FoxP3 a master TF for Tregs, but neither sufficient nor necessary for Treg differentiation

FoxP3 is the lineage-defining TF for Tregs, since it plays a pivotal role in shaping Treg identity. This has been conclusively demonstrated in mouse models lacking FoxP3 and ectopically expression of FoxP3 in Tconvs, which have helped delineate its key functionalities: Treg stability and Treg suppressive function (cf. previous chapter). However, FoxP3's role is complex and not wholly defined since **it is neither sufficient nor necessary for Treg determinism**:

1. Indeed, Tregs with transcriptionally active Foxp3 locus and with a transcriptional signature close to WT Treg, have been shown to develop even in the absence of FoxP3 (first in FoxP3 hemizygous deficiency model), both mouse (Gavin *et al.*, 2007; Lin *et al.*, 2007) and human (Otsubo *et al.*, 2011) *(this subject has already been discussed partly above in heterozygous female)*. This has been explained by the fact that Treg cells exhibit specific epigenetic signatures (such as DNA hypomethylation and histone modification) at functional Treg signature gene loci, including the FoxP3 locus, during the early stages of thymic Treg cell generation, even before FoxP3 and other Treg signature genes are expressed (Samstein, Arvey, *et al.*, 2012; Kitagawa *et al.*, 2017). These modifications are indeed independent of FoxP3 expression, as evidenced in FoxP3-deficient mice (Ohkura *et al.*, 2012). However, these "Treg-like" cells or "Tregs wannabe" have been shown however to have decreased stability and their suppressive function *in vivo* has been questionable. Thus, Treg determinism didn't require FoxP3, but FoxP3 is required to shape the Treg cell epigenetic and transcriptional landscape in order to ensure their stability and function.

2. Simultaneously, the transduction of FoxP3, or its induction by TGF-β, does not suffice to trigger the full *Treg signature* in Tconvs (Gavin *et al.*, 2007; Hill *et al.*, 2007). Several hypotheses have been proposed about the other potential factors involved. First, related to the previous point, the pre-existing epigenetic landscape is needed to induce the full Treg signature. For example, *Ikzf2* which encodes for Helios, a key transcription factor of Tregs, cannot be expressed after FoxP3 overexpression in Tconvs but his locus is opened in Treg-like cells in FoxP3 deficient mice (Ohkura *et al.*, 2012). Second, one

study suggests that Treg cell identity is programmed by a genetic switch that uses multiple redundant TF to maintain lineage commitment (Fu *et al.*, 2012) Furthermore, it has been shown that FoxP3 is engaged with an extensive protein interactome (Rudra *et al.*, 2012), which is a network of more than 300 proteins, and each protein/binding partners within this network might influence Treg identity.

Therefore, these observations lead us to question the specific molecular mechanisms through which FoxP3 imparts Tregs their hallmark suppressive function and stability, and especially his relationship with his interactome.

2. What do we know so far about the molecular mechanisms of Foxp3 functioning

While FoxP3 could adhere to the conventional model of TF activity, involving the activation or repression of target genes through binding to their cis-regulatory elements (promoter and enhancers) and synergizing with other TFs at those sites, the reality appears to be significantly more intricate in the case of FoxP3. We report below several intriguing clues shed light on this complexity.

2.1 FoxP3 is not a pioneer factor

First, some forkhead TFs function as pioneering TFs, meaning that they can directly recognize DNA sequences within the condensed chromatin and open its structure with the help of chromatin modifiers (Zaret, 2020). **FoxP3 is not a pioneer factor**. Indeed, as mentioned above, it utilizes the pre-existing chromatin landscape without significantly altering its structure (Samstein, Arvey, *et al.*, 2012). This pre-existing chromatin landscape is composed of specific epigenetic signatures like DNA hypomethylation and histone modification at functional gene loci, including the FoxP3 locus (Miyao *et al.*, 2012; Samstein, Arvey, *et al.*, 2012; Kitagawa *et al.*, 2017; Herppich *et al.*, 2019; Ohkura *et al.*, 2020; Kawakami *et al.*, 2021), and emerge during the early stages of thymic Treg cell generation, independently of FoxP3 expression. The genome organizer Satb1 is one of the reported orchestrators of the establishment of this landscape (Kitagawa *et al.*, 2017). Thus, the commitment to Treg cell lineage is ensured not by FoxP3 but by DNA demethylation at the FoxP3 locus, irrespectively of ongoing FoxP3 expression. From this, one of the most important modifications is the heavy demethylation of the CNS2 region (Feng *et al.*, 2014), associated with more accessible chromatin and enabling persistent FoxP3 expression. This alteration, referred to as Treg-specific demethylated region (TSDR), appears essential for Treg stability

and function, even amidst inflammatory conditions that typically incite T cell activation. Interestingly, FoxP3 itself has been shown to bind CNS2 and might participate in his own stable expression.

2.2 Most FoxP3 binding sites do not have transcriptional consequences in resting Tregs

Along this line, one thought would have been that FoxP3 exploits this pre-existing landscape, and directly binds these DNA regions as a TF, to upregulate his target genes. However, comparison of the FoxP3-binding sites (~2900 detected by chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq)) with the Treg-specific enhancer/opened regions at steady state (~300-700 detected by DNase sequencing or MeDIP sequencing), in mice (Samstein, Arvey, *et al.*, 2012; Morikawa *et al.*, 2014) has revealed that the two are mainly not overlapping in the genome, except a small subset of exclusively Treg-restricted enhancers, including the Foxp3 conserved non-coding sequence 2 (CNS2). **Treg-specific enhancer had a good correlation with Treg signature, whereas most FoxP3 binding sites do not seem to have transcriptional consequences in resting Tregs** (Morikawa *et al.*, 2014; van der Veeken *et al.*, 2020). However, in activated Treg cells, in several reports, Foxp3 binding regions showed a strong correlation with the downregulated Treg signature (Samstein, Arvey, *et al.*, 2012; Morikawa *et al.*, 2014). Similar analyses in Human Tregs are consistent in that FoxP3 does not appear to control Treg cell-specific DNA hypomethylation but that it could play a major role in controlling activation- associated transcriptional changes in Treg cells (Ohkura *et al.*, 2020). Finally, analysis of DNA sequences at Foxp3 binding sites identified a Forkhead motif only in a small subset of these DNA regions, suggesting cofactor contribution (Samstein, Arvey, *et al.*, 2012).

2.3 FoxP3 interactome

Indeed, FoxP3 does not work alone. First, as we just described, several families of TFs have been implicated in the commitment to the Treg program and bind to cis-regulatory elements within the Foxp3 locus to regulate its expression (cf. the above paragraph and the *FoxP3 regulation* paragraph) (Trujillo-Ochoa, Kazemian and Afzali, 2023). But, adding to this up-stream regulation, FoxP3 is engaged with an extensive protein interactome (Rudra *et al.*, 2012). In this network of more than 300 proteins, including other transcription factors, co-factors, and epigenetic modifiers, each protein/binding partners might influence Treg identity in cooperation with FoxP3 (**Figure 11**). For example, FoxP3 binds DNA regions

already occupied by other TFs (Samstein, Arvey, *et al.*, 2012). FoxP3 has also been shown to form multi-protein complexes that cooperatively regulate gene expression (Fu *et al.*, 2012; Rudra *et al.*, 2012; Kwon *et al.*, 2017). By interacting with different co-factors, FoxP3 can modulate its DNA-binding affinity, regulate its own stability, and influence the chromatin landscape, thereby dictating the transcriptional profile of Tregs. Interestingly, those partners and their complexes form a regulatory network with positive and negative feedback loops (i.e., complex with FoxP3 and Gata3 form a positive feedback loop that increases the transcription of their genes in Treg cells). Among these numerous cofactors, we can cite:

- **NFAT** (Bettelli, Dastrange and Oukka, 2005; Wu *et al.*, 2006), Nuclear Factor of Activated T cells (NFAT), crucial in T cell activation and differentiation. FoxP3 and NFAT form a complex through the FKHD of FoxP3 and the Rel-homology domain of NFAT, that binds to DNA. Notably, FoxP3 appears to repress NFAT's ability to activate specific genes, redirecting it towards an immunosuppressive profile characteristic of Tregs. In a mouse model with targeted mutagenesis at the NFAT-FoxP3 interaction site, FoxP3 lost its ability to bind NFAT, resulting in a failure to promote suppressor function and Treg-like surface properties when expressed in non-regulatory T cells.
- **AML1/Runx1** (Ono *et al.*, 2007; Kitoh *et al.*, 2009) This interaction is vital for the suppressive function of Tregs. Runx1 is a TF that has been shown to act with FoxP3 in molecular complexes associated with the cofactor Cbfβ (Kitoh *et al.*, 2009) and more recently the TF HIC1 (Andrabi *et al.*, 2023). As an example, when FoxP3 is absent, Runx1 activates the transcription of IL-2 and IFN-©, whereas, when FoxP3 is present, the complex prevents their induction. Runx1 has been shown to bind FoxP3 between the LZ and the FKHD (Ono *et al.*, 2007).
- **Histone acetyltransferases (HATs) and deacetylases (HDACs).** Even if not a pioneer factor, FoxP3 has still been shown to induce changes in the chromatin regions that it binds. For example. It interacts with the histone acetyltransferases (HAT), in particular TIP60 and histone deacetylases (HDAC), such as HDAC7 (Li *et al.*, 2007), which might induce change in FoxP3 protein itself (acetylation) (Xiao *et al.*, 2010) but more likely modulates the chromatin structure by adding or removing acetyl groups to histones. For example, there is particular permissive or inhibitory histone H3 modifications within its binding regions (Chen *et al.*, 2006; Zheng *et al.*, 2007; Kwon *et al.*, 2008).

For some of these partners, the approximative (or sometimes) exact binding site with FoxP3 has been elucidated (for example the N-terminal regions for the histones) but for some, it remains largely

unknowns. One recent work from the lab thought to identify possible interaction site by doing a FoxP3 alanine-scan mutagenesis (Kwon *et al.*, 2017) – this work is described in the paragraph below.

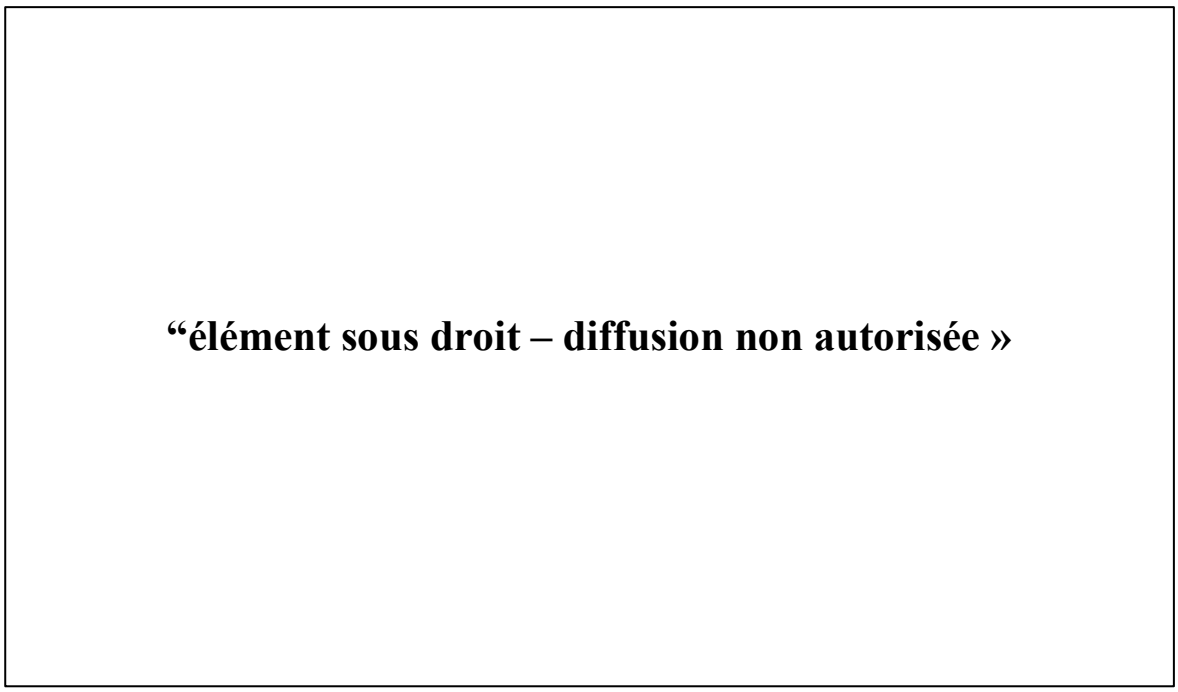


Figure 11. The FoxP3 interactome, figure from (Georgiev, Charbonnier and Chatila, 2019)

Foxp3 may modulate the transcriptome of Treg cells by distinct mechanisms depending on its interaction with diverse binding partners, including chromatin remodelers or various transcriptional co-activators or co-repressors.

2.4 FoxP3 as a transcriptional activator or repressor

Finally, even if not sufficient, FoxP3 is key to determining and maintaining a significant proportion of the Treg signature (Hill *et al.*, 2007; Kwon *et al.*, 2017; van der Veeken *et al.*, 2020). On this subject, FoxP3 shows a functional duality : it acts as a strong repressor and downregulates genes such as *Il2* and *Ifng* but can also act as an activator in upregulating other genes, such as *Il2ra* and *Ctla4*. This aspect of FoxP3 is debated whether it acts as a transcriptional suppressor (Wu *et al.*, 2006; Arvey *et al.*, 2014), activator (Chen *et al.*, 2006; Hill *et al.*, 2007; Kwon *et al.*, 2017) or both.

The “cofactor model” compromised the different views by suggesting that this functional dichotomy depends on the identity of the cofactor(s) with which it interacts in controlling specific targets (Fu *et al.*,

2012; Rudra *et al.*, 2012; Kwon *et al.*, 2017). In particular, Kwon et al. performed an extensive alanine-scan mutagenesis of the FoxP3 protein and generated a set of 130 alanine-replacement mutants of the mouse Foxp3 coding region (Kwon *et al.*, 2017). These mutants were expressed by transduction into CD4+ Tconv cells, for analyzing their FoxP3's interactions with DNA or transcriptional cofactors, as well as their transcriptional activity. This deep analysis showed that FoxP3 existed in distinct multimolecular complexes. It was active and primarily an activator when complexed with the transcriptional factors RELA, IKZF2 and KAT5. In contrast, FoxP3 was inactive when complexed with the histone methyltransferase EZH2 and transcription factors YY1 and IKZF3. Thus, this co-factor hypothesis explains how FoxP3, despite being a single transcription factor, can regulate a diverse array of genes in a context-dependent manner, where FoxP3's impact on transcription is highly dependent on the partners with which it associates.

Another more recent view suggests that FoxP3 mainly functions indirectly through the regulation of the expression of intermediate TFs but this is also still debated (van der Veeken *et al.*, 2020).

3. Using mouse models to study FoxP3+ Tregs and FoxP3 structure-function

Over the last twenty years, most of the discoveries on FoxP3 function come from the use of mouse models of Foxp3 reporter mice and FoxP3 deficient mice. Of course, the initial discoveries come from the study of the ***scurfy*** **mice** (Godfrey, Wilkinson and Russell, 1991; Brunkow *et al.*, 2001), as described above. This naturally occurring mutant strain carries a frameshift mutation in exon 9 of the Foxp3 gene leading to a severely truncated protein that does not have FKHD. These mice have a dramatically reduced lifespan (about 16-25 days) and suffer from severe autoimmunity, including skin inflammation, lymphoproliferation, and multi-organ infiltration. Few years later, **WT FoxP3 with a fluorescent** (Fontenot, Rasmussen, Williams, *et al.*, 2005; Wan and Flavell, 2005; Bettelli *et al.*, 2006) **or a Thy1.1** (Liston *et al.*, 2008) **reporter** alleles were also engineered (as fusion proteins or with the IRES-cassette) in order to easily track FoxP3+ Treg cells *in vivo*. **Foxp3 KO mice** have been then genetically engineered to completely lack the Foxp3 gene (Fontenot, Gavin and Rudensky, 2003; Gavin *et al.*, 2007; Lin *et al.*, 2007) or lack most-of the C-terminal domain (Lin *et al.*, 2005), with or without a reporter cassette. These mice exhibit a phenotype similar to Scurfy mice, although with some variations depending on the specifics of the knockout. Along the same line, **Foxp3 conditional knockout mice (Cre-lox) and**

FoxP3-DTR mice (Kim, Rasmussen and Rudensky, 2007; Lahl *et al.*, 2007; Rubtsov *et al.*, 2010; Tai *et al.*, 2019) were engineered so that Foxp3 can be selectively deleted in specific tissues or/an at different developmental stages or time-points, which allows for more precise study of its function.

These mice have been key to FoxP3 and more broadly Treg function's understanding. However, none of them were designed to analyze the structure-function of FoxP3.

The analysis of structure-function of FoxP3 came first from two main facts:
(1) the awareness of the extend of FoxP3 interactome and the cofactor model (cf. above)
(2) an unexpected finding was made during the investigation of autoimmune phenotypes in a FoxP3 reporter mice, which express a FoxP3 protein fused with GFP at the very N-terminus. The GPF insertion showed to alter FoxP3's interaction with several cofactors, such as Tip60 and Hif1. This modification disturbed the Treg signature genes, leading to dramatically contrasting results - an intensification of autoimmune diabetes, yet offering protection from arthritis (Bettini *et al.*, 2012; Darce *et al.*, 2012).

Given these observations, **mice with IPEX point-mutations,** which carry particular point-mutations similar to those found in human IPEX patients, seem to provide an ideal model for investigating the impact of specific FoxP3 amino acid disruptions on FoxP3-cofactor interactions. While such models are still limited in number, they have contributed significantly to a better understanding of FoxP3's distinct functional domains. The study findings from these mice, primarily from the Hori and Bluestone laboratories, are summarized in the **subsequent table**. All of them are mutations located in the FKHD.

Table 1. Mice engineered with IPEX missense mutations in the literature and their characteristics.

MUTATIONS		Scurfy from our colony (B6)	Scurfy initial description (129)	I363V	M370L	A384T	R397W	"Humanized" IPEX mice (D409A)
Reference		(Chen, Benoist and Mathis, 2005)	(Godfrey, Wilkinson and Russell, 1991)	(Hayatsu *et al.*, 2017)	(Van Gool *et al.*, 2019)	(Hayatsu *et al.*, 2017)	(Hayatsu *et al.*, 2017)	(Goettel *et al.*, 2015)
Clinical course		Scurfy	Scurfy	Milder scurfy-like	Skin, lung inflammation Delayed Scurfy	Skin, lung inflammation Delayed Scurfy	Scurfy-like	Skin, lung inflammation Delayed Scurfy
Analyze at		2 weeks	3-4 weeks	14 weeks	20 weeks	14 weeks	3 weeks	20 weeks
PATHOLOGY	Splenomegaly - ADP	3	3	3		1	3	3
	skin	3	3	2	3	3	3	3
	lung	3	1	3	3	2	3	3
	liver	3	3	1		0	3	1
	pancreas	1	1	1		1	1	0
	colon	1	0	1		0	0	0
	salivary gland	0	1					
	thymus	0	0					
	kidney	1	1					
	small intestine	1	1					1
	heart	1	1					
Biologic features of males				Increase IL4+ Th2, IFNg+ Th1, increased IgE	Increase IL4+ Th2, increased IgE	Increase IL4+IL13+ Th2, IL17+Th17, Increased IgE	Increase IL4+ Th2 , IFNg+ Th1, increased IgE	Increased IgE
Analyse of Heterozygous females				yes	no	yes	yes	no
Discovery				-	Unrestrained TH2 immune response and TH2 Treg-like cells Induce *de novo* Foxp3 binding sites in the TH2 locus	Impaired fitness of A384T tissue-restricted Treg cells Driven by BATF, critical regulator for tissue Treg A384T expands the DNA recognition specificity of FKHD	Impact directly the DNA-Binding site of the FKHD	-

Problematic of the work

Hence, as you may have gathered, our understanding of FoxP3, particularly how its sequence variations influence Treg function and its collaboration with other transcription factors, remains incomplete. Previous structure-function studies, though rare, have demonstrated that localized alterations in FoxP3 can significantly affect various aspects of Treg function and subsequent clinical manifestations. We postulate that a similar link exists between the FoxP3 mutations found in IPEX patients and their array of clinical symptoms. We believe these human mutations provide the most promising model for understanding the functional aspects of FoxP3 and enhancing our comprehension of the disease.

The primary goal of this PhD project is to characterize natural mutations in the transcription factor FoxP3 to unravel its operational mechanisms and its role in the differentiation and functionality of Tregs.

Brief Description of Relevant Experimental Tools

- **(sc) Assay for Transposase-Accessible Chromatin Using Sequencing (ATAC- seq):** Assay for genome-wide chromatin accessibility based on the Tn5 transposase-mediated integration of next generation sequencing adapters selectively within nucleosome-free, accessible chromatin regions. In other words, it allows to identify regulatory regions that have increase accessibility in a certain cell population (opened accessible chromatin regions are tagged and then sequence using next-generation sequencing (NGS)). It has also been used to map regions of TF binding sites and nucleosome positions. sc-ATAC-seq is this assay profiled at the single-cell level - meaning ATAC-seq chromatin accessibility profiles can be assigned to single cells of origin.

- **(sc) RNAseq:** RNA sequencing is a technique that employs Next Generation Sequencing (NGS) to identify and quantify RNA in a biological specimen. This offers a collective view of the cellular transcriptome at any given moment, serving as a dynamic representation of the total pool of RNAs. It can be processed at a population level (called "Bulk RNAseq") or at a single cell level ("scRNAseq").

- **CRISPR/Cas 9 gene editing:** The RNA-guided Cas9 nuclease from the microbial clustered regularly interspaced short palindromic repeats (CRISPR) adaptive immune system can be used to execute efficient genome engineering in eukaryotic cells. It achieves this by merely utilizing a 20-nucleotide targeting sequence within its guide RNA. In this manuscript, we primarily employed CRISPR/Cas9 gene editing within mouse oocytes' pronucleus via Homology-Directed Repair (HDR)

- **Foxp3-ires-GFP and Foxp3-ires-Thy1.1 reporter mice**: These mice have a green fluorescent protein (GFP) or a Thy1.1 tag, inserted into the Foxp3 locus, which allows for the tracking of Foxp3 expression in cells using fluorescence sorting. The IRES sequence allows expression of two independent proteins (FoxP3 and the tag) from a single promoter in a transgenic construct. A single RNA is produced but due to the presence of the IRES, a second translational start is possible on the same RNA. These two models haven't been described to affect the function of Foxp3.

Results

First, we will present an initial investigation, led by David Zemmour, into the transcriptional implications of FoxP3 mutations in an international cohort of IPEX patients (**Chapter 1**). These results have already been published in *Nature Immunology (Zemmour et al., 2021)*. Given its importance in setting the groundwork for subsequent discussions, it is critical to include this research here.

After this groundwork, **the core of this manuscript delves into the engineering and examination of a set of mice that carry the same human mutations as those in Chapter 1,** under a controlled environment facilitating defined perturbations **(Chapter 2)**. This research posits that **the manifestation of IPEX disease arises from the actual mutations, combined with genetic and environmental perturbations**, thereby elucidating the observed intra-familial heterogeneity in IPEX. This study has just been accepted in *Cell Reports* and it is *in press*.

Finally, to further refine our understanding of the potential functions of *wannabe* Tregs in the complete absence of FoxP3 - such as in Tregs from one IPEX patient who bore a null mutation outlined in Chapter 1. My fellow researcher, Kait Chowdhary, and I explored the epigenetic panorama of these Tregs in heterozygous females, leading to the identification of a FoxP3-independent subset of Tregs. These results have been included in a broader manuscript on the Treg regulatory network that was explored by Kait, that is under-review in *Immunity*. For better understanding of the whole context, especially the computational work, I am including the whole paper **but the key results I will be presenting are in Figure 7 and 8**, and the corresponding supplementary figures **(Chapter 3).**

Moreover, the results of these studies have fostered three additional collaborations on FoxP3, the abstracts of which will be appended. Initially, we collaborated with Sun Hur's lab to elucidate the structural mechanism of the R337Q mutant mouse I generated (**Annexe 1**). Subsequently, in collaboration with Ricardo Ramirez, a post-doctoral researcher in our lab, we reevaluated the overall chromatin changes in the absence of FoxP3, highlighting a potential key role of FoxP3 in looping (**Annexe 2**). Lastly, we leveraged these results to investigate the transcriptomic behavior of "ex-" wannabe Tregs from Scurfy mice, in which Marianne Delville, a fellow PhD student, inserted a WT human FoxP3 (**Annexe 3**).

Finally**,** my PhD work was carried out during the COVID-19 pandemic, necessitating an approximately year-long pause in my main mouse-engineering project. During this hiatus, I co-authored two papers unrelated to the central focus of my PhD. These papers probed into the in vivo phenotype of Tregs in severe COVID-19 cases (**Annexe 4**), and an in vitro investigation into the immune-epithelial interaction following SARS-CoV2 infection (**Annexe 5**). This collaborative work also yielded a method paper (**Annexe 6**).

Chapter I: Single-cell analysis of FOXP3 deficiencies in humans and mice unmasks intrinsic and extrinsic $CD4^+$ T cell perturbations.

In the first part of my work, I have been involved in a project led by a former student David Zemmour, that has been published in *Nature Immunology (Zemmour et al., 2021)*.

Starting from our main goal "characterize natural mutations in the transcription factor FoxP3 to unravel its operational mechanisms", we first performed a deep profiling analysis of CD4+ T cells from IPEX patients in order to characterize their Treg cells (i.e., Treg-like cells or "wannabe" Tregs that expressed a dysfunctional FoxP3 gene). This screen associated flow cytometry, single-cell and bulk RNA-seq. We were able to collect an international cohort of 15 IPEX samples, in which the mutations were spread into the FoxP3 locus, targeting different functional domains.

Our findings depict an unanticipated panorama of T cells in IPEX. This landscape features a complex assortment of Treg-like cells, ranging from those that resemble healthy cells to those that appear markedly disrupted. We observed a surprisingly restrained intrinsic signature of FOXP3 but a dominant extrinsic 'IPEX signature' that impacts both Treg and conventional T cells (Tconv). This suggests a feedback-driven exacerbation of T cell perturbations, ultimately leading to clinical disease manifestation.

These observations constitute the foundational work of my main project. I took care of the clinical data collection and of the bulk RNA-seq analysis, especially exploring the extrinsic "IPEX" signature.

For more readability of the whole thesis, the manuscript will be found in Appendix A

Collaboration statement: D.Z., L.M.C., **J.L**. and M.B. performed the experiments.
D.Z., L.M.C., **J.L**., T.A.C., I.A., C. Benoist and D.M. designed the study and analyzed and interpreted the data.
D.Z., **J.L**., C. Benoist and D.M. wrote the manuscript.
E.S., S.K., M.D., S.B., J.Z., K.C., B.N., M.I.G.L., F.R., N.C.B., F.R.L., M.C., I.A., T.A.C., L.M.C. and C. Bruganara provided samples and discussed interpretations.

Chapter II: Mutations from IPEX patients ported to mice reveal different patterns of FoxP3 and Treg dysfunction.

As mentioned earlier, our goal was to identify mutation-specific effects in IPEX patients. However, the challenges that we faced in the previous study, were the prevalent common extrinsic effects that was blurring the more specific intrinsic effect - and the rarity of IPEX, limiting the number of cases for each mutation to one or at most two - a sample size too small to draw any significant conclusions.

To address this, we studied the immunologic and genomic features across a group of six mice, each carrying Foxp3 mutations drawn from the IPEX patients in our prior cohort. This investigation correlated specific alterations in the Foxp3 locus with variations in immunologic status, inflammatory responses, and Treg gene regulatory networks. Notably, we observed significant differences between the FKHD mutation (which had a strong and unspecific effect) and non-FKHD mutations (which showed subtle and specific effects). This research clarified the mechanisms underlying IPEX patient heterogeneity and provided new insights into how FoxP3 collaborates with partner transcription factors to shape Treg identity.

This was my main project during my PhD.

These results have just been accepted in Cell Reports.

For more readability of the whole thesis, the manuscript will be found in Appendix B

Collaboration statement:
JL performed all functional and transcriptional experiments on FoxP3 mutant mice, under the supervision of CB, IA and DM. KC collaborated with **JL** on sample collection for chromatin experiments and KC conducted computational analyses of ATACseq (Figure 7). WZ had performed the DNA binding assay.
All authors contributed to the interpretation, synthesis, and presentation of the results.

Chapter III: An interwoven network of transcription factors, with divergent influences from FoxP3, underlies Treg diversity.

We also asked in a parallel study how wannabe Treg could function without FoxP3. Specifically, we questioned how the transcription factor networks that underpin Treg stability would respond without FoxP3.

Leveraging a Treg network established in Wild Type (WT) Tregs by my peer, Kait Chowdhary, we analyzed the behavior of this network using single-cell ATAC sequencing (scATAC-seq) in FoxP3-deficient mice. For this purpose, I engineered a mouse model with KO FoxP3-ires-GFP. This permits in vivo tagging of *wannabe* Tregs wherein the FoxP3 locus is accessible, especially in heterozygous female devoid of inflammation to study FoxP3 intrinsic effect. We further characterized this population using single-cell RNA sequencing (scRNAseq) in the spleen but also in colon.

These results have been included in a broader manuscript on the Treg regulatory network that was explored by Kait, that is under-review in *Immunity*. For better understanding of the whole context, especially the computational work, I am including the whole paper **but the key results I will be presenting are in Figure 7 and 8**, and the corresponding supplementary figures. I had limited involvement in the first part of the study (figure 1 to 6), more computational, only in discussions.

For more readability of the whole thesis, the manuscript will be found in Appendix C

Collaboration statement: KC and **JL** collaborated on experiments involving FoxP3 KO heterozygote female mice as well RNAseq computational analyses. DR and KC collaborated to adapt the ASAP-seq technology to primary mouse Tregs and on colon/spleen experiments.
All other experiments and all computational analyses were conducted by **KC.**
All experiments and analyses were under the supervision of CB and DM.

Discussion

Overview of findings

FoxP3, the central transcription factor in Tregs, presents intricate molecular functions that have yet to be fully elucidated. Through three distinct studies which utilized both human and mouse models, we have started to establish some novel facets of FoxP3's functionality.

We initially exploited the unique opportunity of collecting samples from patients with IPEX disease and their mothers. This approach allowed us to discern a considerable and surprisingly common transcriptional extrinsic effect that impacted all CD4+ T cells, which are likely attributable to secondary inflammation triggered following Treg dysfunction. This pervasive effect, which initially blurred our precise analysis of mutation-specific effects, was absent in heterozygous mothers where WT Tregs induces a dominant suppression.

Subsequently, we initiated a second project to engineer these IPEX mutations in mice, providing a controlled environment for both genetic and environmental factors and providing the replicates that cannot be achieved in human, both in the hemizygous setting (intrinsic + extrinsic effect) and the heterozygous setting (with only the intrinsic effect remaining). With this strategy, we were able to split mutations depending on their impact on the FKHD. Whereas the FKHD mutation induced systemic immune infiltration at steady-state, the non-FKHD mutation induced a phenotype only after an inflammatory trigger or genetic backcrossing. This underlined the need for additional factors (genetic, epigenetic, environmental) beyond the FoxP3 mutation to reveal the phenotype. Moreover, this phenotype was specific to the mutation location, arguing in favor of a genotype-phenotype correlation.

This argument was emphasized by the fact that in heterozygous females, at the molecular level, mutations were impacting different groups of TF cofactors.

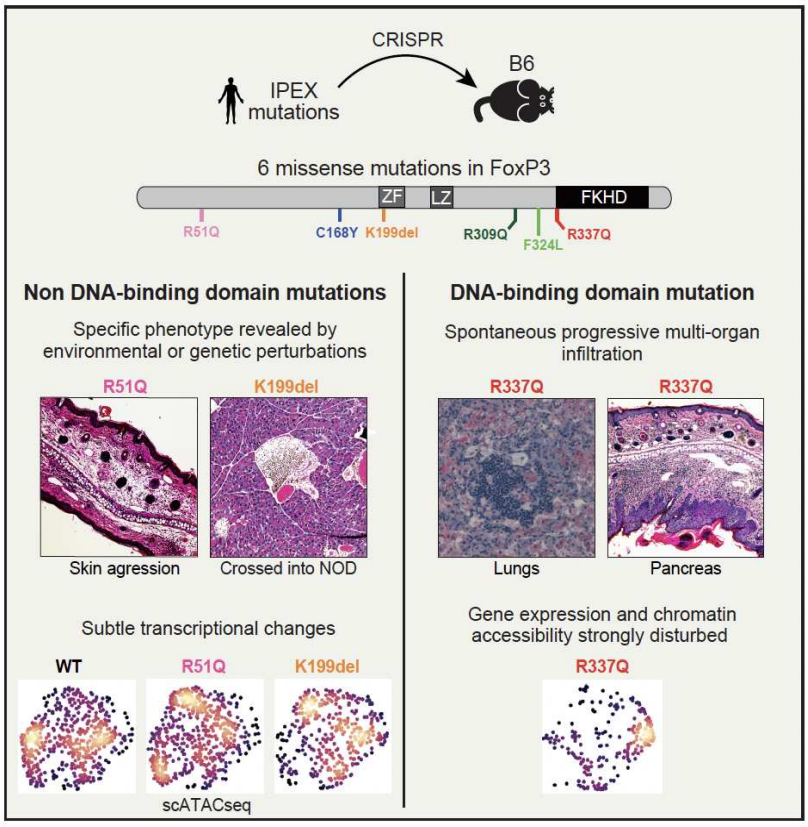


Figure 12. Overview of the results from the molecular study of FoxP3 missense mice

Finally, in collaboration with a colleague who was characterizing the architecture of the Treg regulatory network using scATACseq, we questioned how FoxP3 impacts this Treg network using a model of FoxP3 KO.ires.GFP mice I generated. In the heterozygous context, where we can study the intrinsic effect of FoxP3 KO on Treg-like cells without the confounding influence of inflammation, we found that FoxP3 significantly influences chromatin accessibility, acting as either an activator or repressor, with differing roles in resting and activated states. We also discovered a FoxP3-independent RORγ+ Tregs program.

Then from these observations, I would like to discuss several points.

1) The common extrinsic effect

A notable finding from the initial study was the homogenous transcriptional profile prevalent across all IPEX patients and within the entire CD4+ compartment. What we called the "IPEX" signature or extrinsic effect - which disappeared in the presence of WT Tregs in heterozygous mothers - may represent a systemic response initiated by the general Treg dysfunction caused by FoxP3 mutations. Rather than presenting qualitative differences, this effect manifested as varying intensities among patients, underlining its common and non-specific origin. Interestingly, the implicated genes were not exclusively T cell activation genes, contrary to what one might expect from a suppression loss.

We propose that this effect might be due to one or multiple inductive factors, delivered through cytokine activity or cell-to-cell interaction, usually inhibited by Treg cells but unleashed by FoxP3 deficiency. It may emanate from other T cells or different immunocytes. IL-2, knowing is effect on NK and Th cells activation, might be one of the candidate effectors.

Consequently, future investigations should delve into characterizing other immune cell dysfunctions in Scurfy mice, at different time points (especially the earliest phase before the clinical manifestations) aiming to better target either the cells producing these signals, or the circulating mediators directly involved in this widespread effect. Toward the end of my PhD, I began investigating this phenomenon - research that will be continued by another student in the lab. Preliminary findings show that the extrinsic effect is already present at day 7 in $CD4^+$ before the first clinical disease manifestation (but absent at day

3 of life) and that the first population that increases in proportion along with Treg-like cells is the dendritic cells, suggesting a pivotal role of in the early onset of the disease.

If we consider that the extrinsic effect might play a crucial role in disease pathology, another therapeutic approach for IPEX would involve controlling this effect by targeting the key mediators or cell-types contributing to this effect. Our preliminary clinical-transcriptional correlation in the initial study failed to establish a specific link between treatment types (including steroids, anti-calcineurin, and mTOR inhibitors) and a decrease of this effect - however, due to the limited number of patients and lack of a prospective study, no firm conclusions can be drawn. To note, a recent study tested Dupilumab, a monoclonal antibody that blocks IL-4 and IL-13 (potential key inflammatory mediators in IPEX with pronounced Th2 manifestations) on a patient with IPEX disease, yielding promising results (Caruso *et al.*, 2022).

2) Crucial role of the FKHD

Our CRISPR-engineered mice study distinguishes two classes of *FOXP3* mutations. In the first class, the FKHD mutant, R337Q, showed a "slowed down" version of the full *scurfy* phenotype: skin and lung inflammation at steady-state, activation of many immunocytes, and dysregulated IgE levels. Moreover, R337Q mutation also reproduced most of the transcriptional alterations of full LOF cells (albeit in a muted fashion), with a strong shift in the Treg chromatin architecture, the loss of TCF1 repression and of the positive feedback on the CNS2 element of the *Foxp3* locus. Although the genomic studies were not as extensive in previous studies, this phenotype strongly echoed those previously reported in mice with partial loss-of-function missense mutations in the FKHD (such as A384T, M370I (Hayatsu *et al.*, 2017; Van Gool *et al.*, 2019)), in particular the hyper-IgE and IL4 over-expression, which

we had also noted *in vitro* with strong FKHD mutations (Kwon *et al.*, 2017). This FKHD phenotype is concordant with results from clinical studies in which, generally speaking, FKHD mutations are more severe.

From the molecular standpoint, the DNA-binding domain of FOXP3 might be key to maintaining Treg homeostasis at the resting state (without an inflammatory trigger), likely reinforcing the overall pre-established Treg TFs network (cf. below on the dual role of FoxP3). This hypothesis is concordant with a recent work on 3D chromatin structure in Tregs, led by Ricardo Ramirez in which I was involved **(Appendix - publication 2** (Ramirez *et al.*, 2022))**.** Based on HiChIPseq analysis, this suggests that FoxP3 is associated with a subset of loop structures in Tregs, exhibiting specificity, and that FoxP3 would stabilize the overall structure especially at the resting state.

3) Cofactor model and "Induction" of non-FKHD specific phenotypes

Contrasting with the more monomorphic consequences of FKHD mutations, we observed specific phenotypes in non-FKHD mutations - mutations in other domains of FoxP3 that are not involved in DNA-binding (at least not as directly) - leading us to postulate that each of these mutations might disturb interaction with distinctive co-factors. Interestingly, there was a tissue-specific mutation effect (exacerbation of autoimmune diabetes for K199del, or of skin inflammation for R51Q) - which has also been described in the IPEX literature in patients (Consonni, Ciullini Mannurita and Gambineri, 2021) - meaning that this cofactor-FoxP3 interaction could vary based on tissue location. Additionally, the scATACseq results supported the co-factor model of non-FKHD mutations with a strong shift in case of FKHD mutations and a more subtle but present change in the chromatin accessibility in the other mutations analyzed. This is in accord with the overall "cofactor model", previously applied in several studies (Rudra *et al.*, 2012; Kwon *et al.*, 2017).

This cofactor-FoxP3 interaction might also be dependent on other parameters, since none of the non-FKHD mutations manifested a significant inflammatory phenotype under steady-state conditions in males, contradicting the corresponding human patients who exhibited illness. Moreover, even if the molecular changes are present at steady-state in some non-FKHD mutations, they remain subtle in comparison to the acquired phenotype and, unfortunately, we were not able to identify the exact molecular mechanism through which each mutation triggers a different phenotype. This discrepancy underscores several key considerations:

(1) The potential impact of secondary "hits" or influences stemming from genetic modifiers or environmental triggers is demonstrated by the phenotypic manifestation when backcrossed into another genetic background or after inducing inflammation. Molecular analyses after these environmental/genetic modifiers, in particular if directly at tissue location, might have revealed more significant changes. Moreover, in males, a robust homeostatic drive could aim to restore Treg function, likely helping to restore mutant Tregs function. This homeostatic drive and "phenotypic reversion" might explain why males with the mutations are comparatively well off, similar to the A-Tregs observed in patients from our initial study (Zemmour *et al.*, 2021). Genetic polymorphisms in other key Treg factors, such as those from the CoreTreg genes like CTLA-4 or CD25, could compensate to a certain degree for FoxP3 dysfunction (in particular via feedback mechanisms). These observations reinforce the notion that the impact of FoxP3 on Treg cell function is highly context-dependent, explaining the heterogeneity of onset age and clinical phenotypes displayed by patients bearing the same mutation (Bacchetta, Barzaghi and Roncarolo, 2018; Consonni, Ciullini Mannurita and Gambineri, 2021).

(2) Potential differences between mouse and human FoxP3 and Treg function might be the reason for different clinical manifestations. Despite high protein conservation between the species, only

a fraction of the Treg transcriptional signature is shared (Hill *et al.*, 2007; Ferraro *et al.*, 2014) and disease phenotype is different between *Scurfy* and IPEX (i.e. no diabetes and a more pronounced lymphoproliferation in *Scurfy)*. We hypothesize that, for example, some mutations might affect a co-factor interaction which is not present in mouse physiology (e.g., F324L or C168Y).

A critical question arising from our finding that IPEX is not only related to the actual mutation, centers on how disease triggers/enhancers could be controlled. Based on our results, in addition to genetic strategies that aim to directly correct the mutated FoxP3 gene (Masiuk *et al.*, 2019; Goodwin *et al.*, 2020; S. Borna *et al.*, 2022), treatment design to compensate FoxP3 defect or the overall Treg defect (such as low dose IL2/anti IL2 complex (Zemmour *et al.*, 2021) or CTLA4-Ig (Gerbaux *et al.*, 2023) or mTor inhibitors (Charbonnier *et al.*, 2019)) might allow to control the disease. Some of them (such as mTor inhibitors) are already tested.

4) A dual role for FoxP3 in Treg identity

Lastly, our third investigation regarding FOXP3's impact on the Treg transcription factor network raises a fundamental question: What are the quintessential components necessary for Treg identity and function? This study revisits an established dichotomy between regulators necessary for lineage determination versus those required for effector function. FoxP3 clearly serves a dual role in this context. First, FoxP3 had an important overall impact on chromatin accessibility, acting as an activator or repressor localized to distinct chromatin programs with varied motif enrichments. The loci in each of these programs may define the locations where FoxP3 assembles into different molecular complexes with opposing functions (Kwon *et al.*, 2017). FoxP3 effects were mostly unrelated to its binding, as ascertained

by ChIP-seq, but there was a modest enrichment of FoxP3 binding within FoxP3-repressed regions compared to potentiated regions. Thus, in rTregs, where repressing key cytokines is key, FoxP3 enhances an already existing program, thereby stabilizing Treg identity; this would be mostly performed by the FKHD, mainly acting as a repressor. As we suggested previously, this might occur in the context of stabilizing chromatin loops (Ramirez *et al.*, 2022). Conversely, in aTregs, FoxP3 primarily impacts effector function, prompting specific chromatin programs associated with effector molecule expression and tissue-specific Treg delineation. Along this line, FoxP3 mutation in each IPEX patient might impact these two states differently and we speculate that the subsequent state-specific dysfunctions might also be key in driving disease severity and age of onset. Thus, Tregs should not be defined only by the stable expression of the Treg signature, including Foxp3, at steady state but also by their functional role in suppressing inflammation and maintaining homeostasis, which will be how they react after activation.

5) Why do we see Treg-like cells, and do they have any functional impact?

In the initial results chapter, we employed single-cell resolution to clearly identify Treg-like cells in IPEX patients, including those with a nonsense mutation in exon 1 (Zemmour *et al.*, 2021). This confirmed earlier results, mainly mouse studies, showing that FoxP3 is not essential for Treg differentiation and lineage specification (Gavin *et al.*, 2007; Lin *et al.*, 2007). These wannabe Tregs were present in two different contexts.

The first one is in hemizygous males, where we discovered two types of wannabe Tregs: Treg-like cells closely resembling normal Tregs (Type-A) and Treg-like cells with a more disturbed Treg signature (Type-B). We hypothesize that these Type-B Tregs may originally be Type-A Tregs that have destabilized (due to antigen-TCR activation or due to the inflammatory environment) in a context of

deficient FoxP3. Those Type-B cells would have a significant potential to shift towards T effector cells, lose their suppressive function, downregulate key phenotypic markers (especially in the CoreTreg), and evolve into pathogenic Tregs (that we could called Type-C Tregs). This hypothesis aligns with recent findings from Bacchetta's lab, which documented the presence of two populations of expanded autoreactive T cells in IPEX patients (Š. Borna *et al.*, 2022). The first group originates from the expansion of autoreactive effector T cells, likely resulting from a loss of Treg suppressive function. The second group is derived from Treg cells that lose their phenotypic markers, including CD25. We suggest that in hemizygous males, these Treg-like cells could actually contribute to widespread inflammation, where their self-antigen repertoire would make them particularly pathogenic.

The second context is in heterozygous females, which are protected from inflammation by the Treg population that express WT FoxP3. An intriguing observation from our data, both in mice and in patients, is that post-differentiation, even though a portion of the Treg population expresses a mutant allele, there's no absolute selection of cells carrying the WT allele. While we often see a competitive advantage of WT Tregs over mutant Tregs, it's not a complete out-competition. There was around 70%-30% split in 8-week-old heterozygous females and the two heterozygous IPEX mothers studied in the first chapter showed a similar split, even closer to 50%-50%. The current hypothesis is that WT Tregs retain sufficient suppressive function to maintain immune homeostasis, occupy homeostatic niches, and sequester trophic factors. This equilibrium keeps heterozygous females disease-free. But do these "outcompeted" yet "surviving" mutant Tregs play a functional role in triggering inflammation? Could they potentially be pathogenic, suggesting that mothers with IPEX might be partly symptomatic? Or might they serve a functional or homeostatic role in lymphoid or non-lymphoid tissue?

Here are a few thoughts on this.

1) First, as we proposed for hemizygous males, these Treg-like cells could be pathogenic due to (unstable) expression of a mutant FoxP3 and potential conversion towards an effector Th role with a self-antigen Treg repertoire. In females, the adjacent WT Treg compartment would suppress them, as they suppress other cells. This hypothesis could account for the disappearance of aTreg populations and decreased accessibility of aTreg-specific OCRs, attributed to continuous suppression by WT Tregs. If we entertain the idea that these cells could become pathogenic at any moment, needing constant surveillance by WT Tregs, could we also suggest that under inflammatory triggers, the WT Tregs might be overwhelmed by these populations and heterozygous females might present some pathogenic features, albeit mild (i.e., stronger acute disease, extended recovery time post infection, increased or decreased risk of cancer) ? The long-term phenotype of IPEX mothers is actually not really known due to the lack of epidemiological studies and the need for a high number of index cases for an extremely rare disease - a situation that may be practically unattainable. This could be explored easily in heterozygous female mice.

2) Alternatively, resting Treg-like cells could simply reflect the continuous generation of Tregs in the thymus and periphery, with these populations not enduring over time and not differentiating into aTreg. There would not be a true "suppressing" effect of the WT compartment towards the KO. Indeed, the loss of aTregs in Treg-like cells could occur, because FoxP3 is required for a transition in aTreg differentiation (Gavin *et al.*, 2007). For instance, FoxP3 plays an anti-apoptotic role, enhancing survival of aTregs (Lin *et al.*, 2007). Towards the end of my PhD, I performed some preliminary research supporting this possibility. It appeared that the WT:KO Tregs ratio increased with age, but that this ratio remained steady at 50:50 in the colon throughout life, a location of ongoing Treg differentiation. One counter argument could be the presence of these activated type-B and type-C Tregs in males – refuting the hypothesis that they required FoxP3 for reaching the activated state. However, this discrepancy might come from the

definition of “activation”. The “activated” suppressive Tregs for aTregs, which will require FoxP3, is very different from the “activation” seen in Treg-like cell in males.

3) Finally, should we consider a potential "beneficial" function for these wannabe Tregs, the gut provides an interesting example. On closer examination of the gut, we discovered a population of Treg-like cells that expressed RORγ+, just as WT Tregs would, suggesting that the absolute dependence on FoxP3 for differentiation varies across Treg subpopulations and tissue locations. While our manuscript was under preparation, another group also reported the observation of FoxP3-independent RORγ+ Tregs using a genetic tracing strategy (van der Veeken *et al.*, 2022). On a molecular level, this might suggest that RORγ acts as a pioneer factor here, initiating an alternative regulatory program in these cells, which appears to be favored over the Helios program in the context of FoxP3-deficiency. Although these RORγ+ Treg-like cells were microbe-dependent, like their WT RORγ+ Treg counterparts, their functionality and stability, under both steady-state and inflammatory conditions, warrant further investigation. Very preliminary data related to functionality were generated by conducting a DSS colitis experiment comparing heterozygous females with WT homozygous females. Interestingly, albeit from an experiment that has not yet been replicated, the heterozygous WT/KO females seemed to fare better, hinting at a potential role of these RORγ+ Treg-like cells in maintaining gut homeostasis.

In conclusion, while FoxP3 is still a mysterious TF, using natural arising FoxP3 mutations from IPEX patients, we have been able to parse out two distinct mechanisms of action of FoxP3 that is required for a stable (rTreg) and an effective (aTreg) Treg lineage. Those two seem to be driven by different parts of FoxP3, FKHD being essential for Treg stability.

ANNEXES : Other collaborations during the PhD

In conjunction with my previous work on FoxP3, I actively participated in several collaborative scientific projects. Specifically, the unique challenges posed by the COVID-19 pandemic in the first year of my PhD presented unparalleled opportunities for collaboration, greatly enhancing my scientific knowledge and technical skills. Despite a necessary pause in my primary project, this adversity became an unexpected boon, fostering stimulating discussions and fruitful partnerships across several laboratories. While these projects fall outside of this manuscript's main scope - and thus are not included - they significantly contributed to shaping my scientific mindset and skills. Consequently, I contributed to seven other published works - the abstracts of which are appended. I was first or co-first author in two of these, and second or third author in three.

Published Manuscripts

1. Leng F*, Zhang W*, Ramirez RN, Leon J**,** Zhong Y, Hou L, Yuki K, van der Veeken J, Rudensky AY, Benoist C, Hur S. The transcription factor FoxP3 can fold into two dimerization states with divergent implications for regulatory T cell function and immune homeostasis. **Immunity. 2022** Aug 9;55(8):1354-1369.e8. PMID: 35926508

2. Ramirez RN, Chowdhary K, Leon J, Mathis D, Benoist C. FoxP3 associates with enhancer promoter loops to regulate Treg-specific gene expression. **Sci Immunol. 2022** Jan 14; 7(67):eabj9836. PMID: 35030035.

3. Delville M, Bellier F, Leon J, Klifa R, Lizot S, Vinçon H, Sobrino S, Thouenon R, Marchal A, Garrigue A, Olivré J, Charbonnier S, Lagresle-Peyrou C, Amendola M, Schambach A, Gross D, Lamarthée B, Benoist C, Zuber J, André I, Cavazzana M, Six E. A combination of cyclophosphamide and interleukin-2 allows CD4+ T cells converted to Tregs to control scurfy syndrome. **Blood. 2021** Apr 29;137(17):2326-2336. PMID: 33545713

4. Galván-Peña S*, Leon J*, Chowdhary K, Michelson DA, Vijaykumar B, Yang L, Magnuson AM, Chen F, Manickas-Hill Z, Piechocka-Trocha A, Worrall DP, Hall KE, Ghebremichael M, Walker BD, Li JZ, Yu XG, Mathis D, Benoist C. Profound Treg perturbations correlate with COVID-19 severity. **Proc Natl Acad Sci U S A. 2021** 09 14; 118(37). PMID: 34433692.

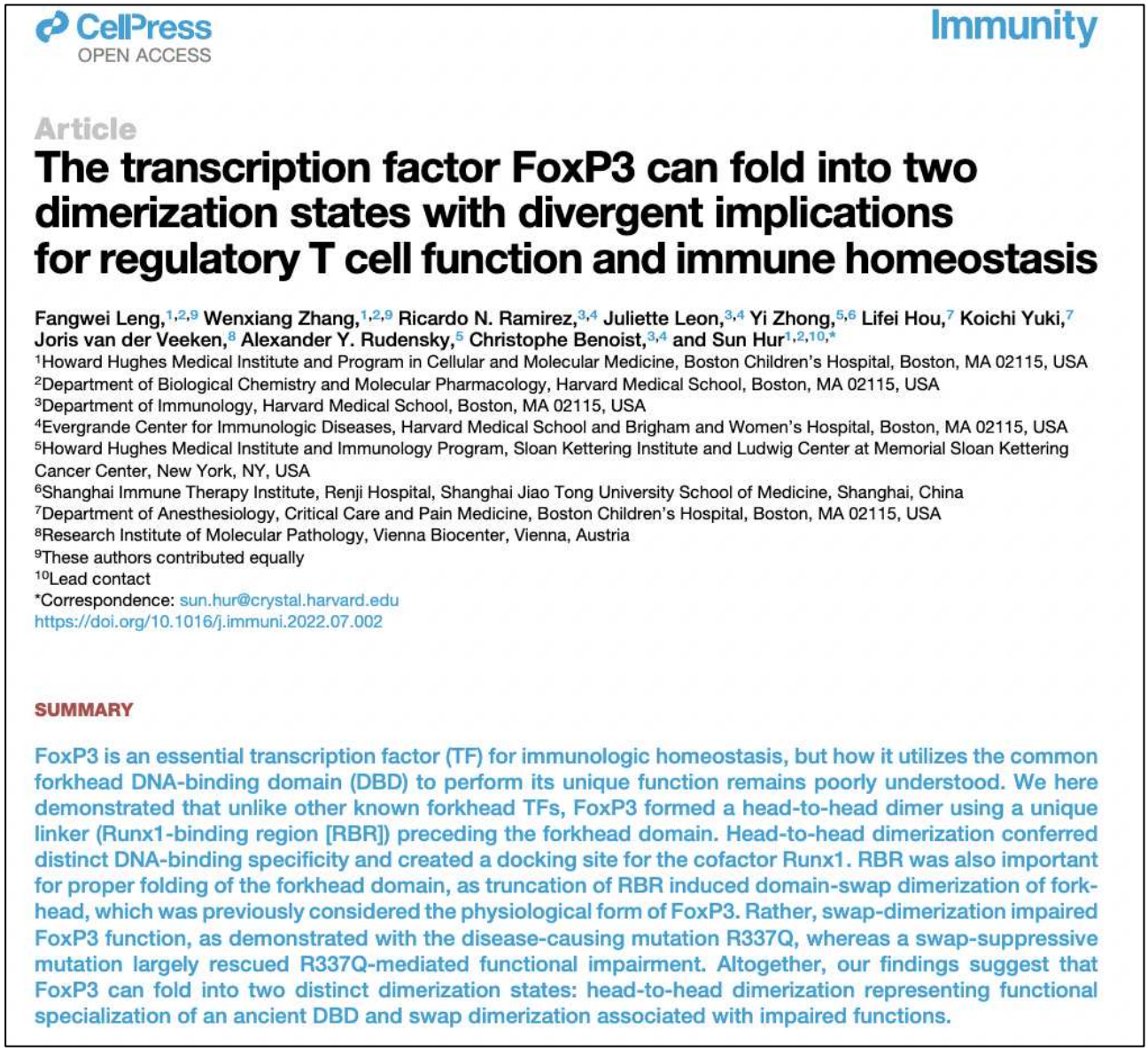
5. Leon J*, Michelson DA*, Olejnik J*, Chowdhary K*, Oh HS*, Hume AJ, Galván-Peña S, Zhu Y, Chen F, Vijaykumar B, Yang L, Crestani E, Yonker LM, Knipe DM, Mühlberger E, Benoist C. A virus specific monocyte inflammatory phenotype is induced by SARS-CoV-2 at the immune-epithelial interface. **Proc Natl Acad Sci U S A. 2022** 01 04; 119(1). PMID: 34969849.

6. Olejnik J, Leon J, Michelson D, Chowdhary K, Galvan-Pena S, Benoist C, Mühlberger E$, Hume AJ$. Establishment of an Inactivation Method for Ebola Virus and SARS-CoV-2 Suitable for Downstream Sequencing of Low Cell Numbers. **Pathogens. 2023** Feb 17;12(2):342. PMID: 36839614.

7. Lamarthée B*, Marchal A*, Charbonnier S*, Blein T*, Leon J, Martin E, Rabaux L, Vogt K, Titeux M, Delville M, Vinçon H, Six E, Pallet N, Michonneau D, Anglicheau D, Legendre C, Taupin JL, Nemazanyy I, Sawitzki B, Latour S, Cavazzana M, André I, Zuber J. Transient mTOR inhibition rescues 4-1BB CAR-Tregs from tonic signal-induced dysfunction. **Nat Commun. 2021** Nov 8;12(1):6446. PMID: 34750385

**= Equal Contribution, $ = Co-supervision*

Publication 1: The transcription factor FoxP3 can fold into two dimerization states with divergent implications for regulatory T cell function and immune homeostasis.



CellPress
OPEN ACCESS

Immunity

Article

The transcription factor FoxP3 can fold into two dimerization states with divergent implications for regulatory T cell function and immune homeostasis

Fangwei Leng,[1,2,9] Wenxiang Zhang,[1,2,9] Ricardo N. Ramirez,[3,4] Juliette Leon,[3,4] Yi Zhong,[5,6] Lifei Hou,[7] Koichi Yuki,[7] Joris van der Veeken,[8] Alexander Y. Rudensky,[5] Christophe Benoist,[3,4] and Sun Hur[1,2,10,*]

[1]Howard Hughes Medical Institute and Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA 02115, USA
[2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA
[3]Department of Immunology, Harvard Medical School, Boston, MA 02115, USA
[4]Evergrande Center for Immunologic Diseases, Harvard Medical School and Brigham and Women's Hospital, Boston, MA 02115, USA
[5]Howard Hughes Medical Institute and Immunology Program, Sloan Kettering Institute and Ludwig Center at Memorial Sloan Kettering Cancer Center, New York, NY, USA
[6]Shanghai Immune Therapy Institute, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[7]Department of Anesthesiology, Critical Care and Pain Medicine, Boston Children's Hospital, Boston, MA 02115, USA
[8]Research Institute of Molecular Pathology, Vienna Biocenter, Vienna, Austria
[9]These authors contributed equally
[10]Lead contact
*Correspondence: sun.hur@crystal.harvard.edu
https://doi.org/10.1016/j.immuni.2022.07.002

SUMMARY

FoxP3 is an essential transcription factor (TF) for immunologic homeostasis, but how it utilizes the common forkhead DNA-binding domain (DBD) to perform its unique function remains poorly understood. We here demonstrated that unlike other known forkhead TFs, FoxP3 formed a head-to-head dimer using a unique linker (Runx1-binding region [RBR]) preceding the forkhead domain. Head-to-head dimerization conferred distinct DNA-binding specificity and created a docking site for the cofactor Runx1. RBR was also important for proper folding of the forkhead domain, as truncation of RBR induced domain-swap dimerization of forkhead, which was previously considered the physiological form of FoxP3. Rather, swap-dimerization impaired FoxP3 function, as demonstrated with the disease-causing mutation R337Q, whereas a swap-suppressive mutation largely rescued R337Q-mediated functional impairment. Altogether, our findings suggest that FoxP3 can fold into two distinct dimerization states: head-to-head dimerization representing functional specialization of an ancient DBD and swap dimerization associated with impaired functions.

Published in *Immunity*, 2022 Aug 9; PMID: 35926508

Here, I generated by CRISPR/Cas9 the mouse model of FoxP3 R337Q mice and took care of the subsequent breeding. I performed the *in vivo* analysis on this line, including histology and flow cytometry. I designed the corresponding panels in Figure 7 and wrote the related part of the manuscript.

***Collaboration statement**: F.L.., W.Z., and S.H. conceived the study, designed experiments, and wrote the manuscript. F.L., W.Z., and **J.L**. performed experiments and analyzed data. R.N.R. and Y.Z. performed computational analysis. L.H., K.Y., and J.v.d.V. assisted experiments and analysis. A.Y.R., C.B., and S.H. supervised experimental design and data analysis.*

Publication 2: FoxP3 associates with enhancer promoter loops to regulate Treg-specific gene expression.

SCIENCE IMMUNOLOGY | RESEARCH ARTICLE

T CELLS

FoxP3 associates with enhancer-promoter loops to regulate T_{reg}-specific gene expression

Ricardo N. Ramirez, Kaitavjeet Chowdhary, Juliette Leon, Diane Mathis*, Christophe Benoist*

Gene expression programs are specified by higher-order chromatin structure and enhancer-promoter loops (EPLs). T regulatory cell (T_{reg}) identity is dominantly specified by the transcription factor (TF) FoxP3, whose mechanism of action is unclear. We applied chromatin conformation capture with immunoprecipitation (HiChIP) in T_{reg} and closely related conventional $CD4^+$ T cells (T_{conv}). EPLs identified by H3K27Ac HiChIP showed a range of connection intensity, with some superconnected genes. TF-specific HiChIP showed that FoxP3 interacts with EPLs at a large number of genes, including some not differentially expressed in T_{reg} versus T_{conv}, but enriched at the core T_{reg} signature loci that it up-regulates. FoxP3 association correlated with heightened H3K27Ac looping, as ascertained by analysis of FoxP3-deficient T_{reg}-like cells. There was marked asymmetry in the loci where FoxP3 associated at the enhancer- or the promoter-side of EPLs, with enrichment for different transcriptional cofactors. FoxP3 EPL intensity distinguished gene clusters identified by single-cell ATAC-seq as covarying between individual T_{regs}, supporting a direct transactivation model for FoxP3 in determining T_{reg} identity.

Published in *Science Immunology*, 2022 Jan 14; PMID: 35030035.

In this study, I generated by CRISPR/Cas9 the mouse model of KO-Foxp3-ires-GFP mice and took care of the subsequent breeding. I performed the flow cytometric experiments, and I collected samples for chromatin and RNAseq experiments on this line. I analyzed the subsequent flow cytometry and bulk RNAseq results. I designed the corresponding panels in the manuscript (included in Figure 3, Figure 4 and Figure S5).

Collaboration statement: *Conceptualization: R.N.R. and C.B. Investigation: R.N.R., K.C., and **J.L.** Supervision: C.B. Funding acquisition: C.B. and D.M. Writing—original draft: R.N.R., K.C., **J.L.,** C.B., and D.M. Writing—review and editing: R.N.R. and C.B.*

Publication 3: A combination of cyclophosphamide and interleukin-2 allows CD4+ T cells converted to Tregs to control scurfy syndrome.

GENE THERAPY

A combination of cyclophosphamide and interleukin-2 allows CD4$^+$ T cells converted to Tregs to control *scurfy* syndrome

Marianne Delville,[1-3] Florence Bellier,[1] Juliette Leon,[1,4] Roman Klifa,[1,5] Sabrina Lizot,[1] Hélène Vinçon,[1] Steicy Sobrino,[1] Romane Thouenon,[1] Armance Marchal,[1] Alexandrine Garrigue,[1] Juliette Olivré,[1] Soëli Charbonnier,[1] Chantal Lagresle-Peyrou,[1,3] Mario Amendola,[6] Axel Schambach,[7] David Gross,[8] Baptiste Lamarthée,[1] Christophe Benoist,[4] Julien Zuber,[1,9] Isabelle André,[1,*] Marina Cavazzana,[1,2,*] and Emmanuelle Six[1,*]

[1]Institut Imagine, Université de Paris, INSERM UMR1163, Laboratory of Human Lymphohematopoiesis, Paris, France; [2]Service de Biothérapie et d'Aphérèse, Hôpital Necker, Groupe Hospitalier Universitaire Ouest, Assistance Publique–Hôpitaux de Paris (AP-HP), Paris, France; [3]Centre d'Investigation Clinique Biothérapie, Groupe Hospitalier Universitaire Ouest, AP-HP, Paris, France; [4]Department of Immunology, Harvard Medical School, Boston, MA; [5]Service d'Immuno-Hématologie Pédiatrique, Hôpital Necker, AP-HP, Paris, France; [6]Genethon, Evry, France; [7]Institute of Experimental Hematology, Hannover Medical School, Hanover, Germany; [8]Institut Necker Enfants-Malades, Université de Paris, U1151, INSERM, Paris, France; and [9]Service de Néphrologie et Transplantation Rénale, Hôpital Necker, Groupe Hospitalier Universitaire Ouest, AP-HP, Paris, France

KEY POINTS

- **The combination of Cy conditioning and IL-2 allows suppressive T cells to rescue *scurfy* mice after disease onset.**
- **Transcriptomic analysis reveals a lasting restoration of Treg identity in FOXP3-transduced *scurfy* cells, even in an inflammatory environment.**

Immunodysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome is caused by mutations in forkhead box P3 (*FOXP3*), which lead to the loss of function of regulatory T cells (Tregs) and the development of autoimmune manifestations early in life. The selective induction of a Treg program in autologous CD4$^+$ T cells by *FOXP3* gene transfer is a promising approach for curing IPEX. We have established a novel in vivo assay of Treg functionality, based on adoptive transfer of these cells into *scurfy* mice (an animal model of IPEX) and a combination of cyclophosphamide (Cy) conditioning and interleukin-2 (IL-2) treatment. This model highlighted the possibility of rescuing *scurfy* disease after the latter's onset. By using this in vivo model and an optimized lentiviral vector expressing human Foxp3 and, as a reporter, a truncated form of the low-affinity nerve growth factor receptor (ΔLNGFR), we demonstrated that the adoptive transfer of FOXP3-transduced *scurfy* CD4$^+$ T cells enabled the long-term rescue of *scurfy* autoimmune disease. The efficiency was similar to that seen with wild-type Tregs. After in vivo expansion, the converted CD4^{FOXP3} cells recapitulated the transcriptomic core signature for Tregs. These findings demonstrate that FOXP3 expression converts CD4$^+$ T cells into functional Tregs capable of controlling severe autoimmune disease. (*Blood*. 2021;137(17):2326-2336)

Published in *Blood*, 2021 Apr 29; PMID: 33545713

In this work, I analyzed the bulk RNAseq of the converted CD4^{FOXP3} cells. I designed the whole related figure (Figure 4) and wrote this part of the manuscript.

***Collaboration statement**: M.D. and **J.L.** conceptualized the study, provided the study methodology, conducted investigations, provided formal analysis, wrote the original draft of the manuscript, and edited and reviewed the manuscript; F.B., R.K., S.L., and H.V. conducted investigations and provided formal analysis; S.S., R.T., A.M., A.G., J.O., and S.C. conducted investigations; C.L.-P. conceptualized the study, provided the study methodology, and edited and reviewed the manuscript; M.A., A.S., and D.G. conceptualized the study and provided the study methodology; B.L. conducted investigations, provided formal analysis, and edited and reviewed the manuscript; C.B., J.Z., I.A., and E.S. conceptualized the study, provided the study methodology, supervised the study, and edited and reviewed the manuscript; and M.C. conceptualized the study, provided the study methodology, acquired funding, supervised the study, and edited and reviewed the manuscript.*

Publication 4: Profound Treg perturbations correlate with COVID-19 severity.

Profound Treg perturbations correlate with COVID-19 severity

Silvia Galván-Peña[a,1], Juliette Leon[a,b,1], Kaitavjeet Chowdhary[a], Daniel A. Michelson[a], Brinda Vijaykumar[a], Liang Yang[a], Angela M. Magnuson[a], Felicia Chen[a], Zachary Manickas-Hill[c,d], Alicja Piechocka-Trocha[c,d], Daniel P. Worrall[c,d], Kathryn E. Hall[c,e], Musie Ghebremichael[c,d], Bruce D. Walker[c,d,f], Jonathan Z. Li[c,g], Xu G. Yu[c,d], MGH COVID-19 Collection & Processing Team[2], Diane Mathis[a], and Christophe Benoist[a,3]

[a]Department of Immunology, Harvard Medical School, Boston, MA 02115; [b]Imagine Institute, INSERM UMR 1163, University of Paris, 75015 Paris, France; [c]Massachusetts Consortium on Pathogen Readiness, Boston, MA 02115; [d]Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology, and Harvard, Cambridge, MA 02139; [e]Department of Medicine, Massachusetts General Hospital, Boston, MA 02114; [f]HHMI, Center for the AIDS Programme of Research in South Africa, Chevy Chase, MD 20815; and [g]Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115

Contributed by Christophe Benoist, July 16, 2021 (sent for review June 18, 2021; reviewed by Alain Fischer and Akiko Iwasaki)

The hallmark of severe COVID-19 is an uncontrolled inflammatory response, resulting from poorly understood immunological dysfunction. We hypothesized that perturbations in FoxP3$^+$ T regulatory cells (Treg), key enforcers of immune homeostasis, contribute to COVID-19 pathology. Cytometric and transcriptomic profiling revealed a distinct Treg phenotype in severe COVID-19 patients, with an increase in Treg proportions and intracellular levels of the lineage-defining transcription factor FoxP3, correlating with poor outcomes. These Tregs showed a distinct transcriptional signature, with overexpression of several suppressive effectors, but also proinflammatory molecules like interleukin (IL)-32, and a striking similarity to tumor-infiltrating Tregs that suppress antitumor responses. Most marked during acute severe disease, these traits persisted somewhat in convalescent patients. A screen for candidate agents revealed that IL-6 and IL-18 may individually contribute different facets of these COVID-19–linked perturbations. These results suggest that Tregs may play nefarious roles in COVID-19, by suppressing antiviral T cell responses during the severe phase of the disease, and by a direct proinflammatory role.

1. Equal Contribution

Published in *PNAS, 2021* Sept 14; PMID: 34433692.

As mentioned below, this work was a true team effort involving scientific and medical teams. In duo with Silvia Galvan-Pena, we performed all the flow-cytometric and transcriptomic experiments of the *in vivo* analysis of Tregs in COVID-19. I then focused on the analyzed of the transcriptome and Silvia focused on the flow cytometry results. We then wrote the manuscript together with Christophe Benoist, with multiple inputs from the co-authors.

Collaboration statement: *This was work done as part of a team effort (involving S.G.-P,* ***J.L****, K.C., and D.A.M.) during the COVID pandemic. S.G.-P.,* ***J.L****., K.C., D.A.M., B.D.W., J.Z.L., X.G.Y., D.M., and C.B. designed research; S.G.-P.,* ***J.L****., K.C., D.A.M., A.M.M., F.C., Z.M.-H., A.P.-T., D.P.W., K.E.H., and M.G. performed research; MGH COVID-19 Collection & Processing Team contributed new reagents/analytic tools; S.G.-P.,* ***J.L****., B.V., L.Y., and C.B. analyzed data; and S.G.-P.,* ***J.L****., K.C., D.A.M., and C.B. wrote the paper*

Publication 5: A virus specific monocyte inflammatory phenotype is induced by SARS-CoV-2 at the immune-epithelial interface.

A virus-specific monocyte inflammatory phenotype is induced by SARS-CoV-2 at the immune–epithelial interface

Juliette Leon[a,b,1], Daniel A. Michelson[a,1], Judith Olejnik[c,d,1], Kaitavjeet Chowdhary[a,1], Hyung Suk Oh[e,1], Adam J. Hume[c,d], Silvia Galván-Peña[a], Yangyang Zhu[a], Felicia Chen[a], Brinda Vijaykumar[a], Liang Yang[a], Elena Crestani[f,g], Lael M. Yonker[h], David M. Knipe[e], Elke Mühlberger[c,d], and Christophe Benoist[a,2]

[a]Department of Immunology, Blavatnik Institute, Harvard Medical School, Boston, MA 02115; [b]INSERM UMR 1163, University of Paris, Institut Imagine, 75015 Paris, France; [c]Department of Microbiology, Boston University School of Medicine, Boston, MA 02118; [d]National Emerging Infectious Diseases Laboratories, Boston University, Boston, MA 02215; [e]Department of Microbiology, Blavatnik Institute, Harvard Medical School, Boston, MA 02115; [f]Division of Immunology, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115; [g]Department of Pediatrics, Harvard Medical School, Boston, MA 02115; and [h]Department of Pediatrics, Massachusetts General Hospital, Boston, MA 02114

Contributed by Christophe Benoist; received September 13, 2021; accepted November 19, 2021; reviewed by Miriam Merad and Tadatsugu Taniguchi

Infection by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) provokes a potentially fatal pneumonia with multiorgan failure, and high systemic inflammation. To gain mechanistic insight and ferret out the root of this immune dysregulation, we modeled, by in vitro coculture, the interactions between infected epithelial cells and immunocytes. A strong response was induced in monocytes and B cells, with a SARS-CoV-2–specific inflammatory gene cluster distinct from that seen in influenza A or Ebola virus-infected cocultures, and which reproduced deviations reported in blood or lung myeloid cells from COVID-19 patients. A substantial fraction of the effect could be reproduced after individual transfection of several SARS-CoV-2 proteins (Spike and some nonstructural proteins), mediated by soluble factors, but not via transcriptional induction. This response was greatly muted in monocytes from healthy children, perhaps a clue to the age dependency of COVID-19. These results suggest that the inflammatory malfunction in COVID-19 is rooted in the earliest perturbations that SARS-CoV-2 induces in epithelia.

1. *Equal Contribution*

Published in *PNAS*, 2022 Jan 04; PMID: 34969849.

As mentioned below, this study was also a true team effort involving three scientific teams. In collaboration with D.A.M., J.O., K.C., H.S.O., A.J.H. and S.G.P., we designed and performed all the in vitro experiments. D.AM. and K.C analyzed the results of the transcriptome while I was focused on the flow cytometry results as well as the gene ontology/ signature overlap analysis. We then wrote the manuscript together with Christophe Benoist.

Collaboration statement: *This was work done as part of a team effort (involving in particular **J.L.**, D.A.M., J.O., K.C., H.S.O., A.J.H. and S.G.P.) during the COVID pandemic; **J.L.**, D.A.M., K.C., S.G.P. and C.B. designed research; **J.L.**, D.A.M., J.O., K.C., H.S.O., A.J.H. and S.G.P. performed research, under the supervision of D.M.K., E.M., and C.B.; **J.L.**, D.A.M., K.C., B.V., L.Y. and C.B. analyzed data; E.C. and L.M.Y. provided controls samples and **J.L.**, D.A.M., K.C., S.G.P. and C.B. wrote the paper. All authors reviewed the manuscript.*

Publication 6: Establishment of an Inactivation Method for Ebola Virus and SARS-CoV-2 Suitable for Downstream Sequencing of Low Cell Numbers.

pathogens

MDPI

Article

Establishment of an Inactivation Method for Ebola Virus and SARS-CoV-2 Suitable for Downstream Sequencing of Low Cell Numbers

Judith Olejnik [1,2], Juliette Leon [3,4], Daniel Michelson [3], Kaitavjeet Chowdhary [3], Silvia Galvan-Pena [3], Christophe Benoist [3], Elke Mühlberger [1,2,*] and Adam J. Hume [1,2,5,*]

[1] Department of Microbiology, Boston University School of Medicine, Boston, MA 02118, USA
[2] National Emerging Infectious Diseases Laboratories, Boston University, Boston, MA 02218, USA
[3] Department of Immunology, Blavatnik Institute, Harvard Medical School, Boston, MA 02115, USA
[4] INSERM UMR 1163, Institut Imagine, University of Paris, 75015 Paris, France
[5] Center for Emerging Infectious Diseases Policy & Research, Boston University, Boston, MA 02118, USA
* Correspondence: muehlber@bu.edu (E.M.); hume@bu.edu (A.J.H.); Tel.: +1-617-358-9153 (E.M.); +1-617-358-9151 (A.J.H.)

Abstract: Technologies that facilitate the bulk sequencing of small numbers of cells as well as single-cell RNA sequencing (scRNA-seq) have aided greatly in the study of viruses as these analyses can be used to differentiate responses from infected versus bystander cells in complex systems, including in organoid or animal studies. While protocols for these analyses are typically developed with biosafety level 2 (BSL-2) considerations in mind, such analyses are equally useful for the study of viruses that require higher biosafety containment levels. Many of these workstreams, however, are not directly compatible with the more stringent biosafety regulations of BSL-3 and BSL-4 laboratories ensuring virus inactivation and must therefore be modified. Here we show that TCL buffer (Qiagen), which was developed for bulk sequencing of small numbers of cells and also facilitates scRNA-seq, inactivates both Ebola virus (EBOV) and SARS-CoV-2, BSL-4 and BSL-3 viruses, respectively. We show that additional heat treatment, necessary for the more stringent biosafety concerns for BSL-4-derived samples, was additionally sufficient to inactivate EBOV-containing samples. Critically, this heat treatment had minimal effects on extracted RNA quality and downstream sequencing results.

Keywords: Ebola virus; SARS-CoV-2; virus inactivation; sequencing

check for updates

Citation: Olejnik, J.; Leon, J.; Michelson, D.; Chowdhary, K.; Galvan-Pena, S.; Benoist, C.; Mühlberger, E.; Hume, A.J. Establishment of an Inactivation Method for Ebola Virus and SARS-CoV-2 Suitable for Downstream Sequencing of Low Cell Numbers. *Pathogens* **2023**, *12*, 342.

Published in *Pathogens*, 2023 Feb 17; PMID: 36839614.

In collaboration with D.M., K.C., and S.G.-P., we performed the *in vitro* work that was not involving BSL3 or 4. We performed the bulk RNAseq and we designed the corresponding panel.

Collaboration statement: *Conceptualization J.O., A.J.H. and E.M.; methodology, J.O., A.J.H. and E.M.; investigation, J.O., A.J.H., **J.L**., D.M., K.C. and S.G.-P.; data curation, J.O., A.J.H., S.G.-P., **J.L.** and E.M.; writing—original draft preparation, J.O., A.J.H., E.M. and C.B.; writing—review and editing, J.O., A.J.H., E.M., **J.L.,** D.M. and C.B.; visualization, J.O., A.J.H., E.M. and **J.L**; supervision, E.M. and C.B.; funding acquisition, E.M. and C.B. All authors have read and agreed to the published version of the manuscript.*

Publication 7: Transient mTOR inhibition rescues 4-1BB CAR-Tregs from tonic signal-induced dysfunction

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-021-26844-1 OPEN

Transient mTOR inhibition rescues 4-1BB CAR-Tregs from tonic signal-induced dysfunction

Baptiste Lamarthée[1,14], Armance Marchal[1,14], Soëli Charbonnier[1,14], Tifanie Blein[1,14], Juliette Leon[2], Emmanuel Martin[3], Lucas Rabaux[1], Katrin Vogt[4], Matthias Titeux[5], Marianne Delville[1,6,7], Hélène Vinçon[1], Emmanuelle Six[1], Nicolas Pallet[8], David Michonneau[6,9], Dany Anglicheau[6,10,11], Christophe Legendre[6,10], Jean-Luc Taupin[6,12], Ivan Nemazanyy[13], Birgit Sawitzki[4], Sylvain Latour[3], Marina Cavazzana[1,6,7], Isabelle André[1] & Julien Zuber[1,6,10]✉

The use of chimeric antigen receptor (CAR)-engineered regulatory T cells (Tregs) has emerged as a promising strategy to promote immune tolerance. However, in conventional T cells (Tconvs), CAR expression is often associated with tonic signaling, which can induce CAR-T cell dysfunction. The extent and effects of CAR tonic signaling vary greatly according to the expression intensity and intrinsic properties of the CAR. Here, we show that the 4-1BB CSD-associated tonic signal yields a more dramatic effect in CAR-Tregs than in CAR-Tconvs with respect to activation and proliferation. Compared to CD28 CAR-Tregs, 4-1BB CAR-Tregs exhibit decreased lineage stability and reduced in vivo suppressive capacities. Transient exposure of 4-1BB CAR-Tregs to a Treg stabilizing cocktail, including an mTOR inhibitor and vitamin C, during ex vivo expansion sharply improves their in vivo function and expansion after adoptive transfer. This study demonstrates that the negative effects of 4-1BB tonic signaling in Tregs can be mitigated by transient mTOR inhibition.

Published in *Nature Communications.* 2021 Nov 8 PMID: 34750385

In this work, I analyzed the RNAseq of the CD28 and the 4-1BB CAR-Tregs, as well as designed the figures and wrote this part of the manuscript (included in Figure 3). All other experiments and analyses were conducted by the other authors.

TABLES AND FIGURES

REFERENCES

Agle, K. *et al.* (2018) 'Bim regulates the survival and suppressive capability of CD8+ FOXP3+ regulatory T cells during murine GVHD', *Blood*, 132(4), pp. 435–447.

Ali, N. *et al.* (2017) 'Regulatory T Cells in Skin Facilitate Epithelial Stem Cell Differentiation', *Cell*, 169(6), pp. 1119–1129.e11.

Andersen, K.G., Nissen, J.K. and Betz, A.G. (2012) 'Comparative Genomics Reveals Key Gain-of-Function Events in Foxp3 during Regulatory T Cell Evolution', *Frontiers in immunology*, 3, p. 113.

Andrabi, S.B.A. *et al.* (2023) 'HIC1 interacts with FOXP3 multi protein complex: a novel mechanism to regulate human regulatory T cell differentiation and function', *bioRxiv*. Available at: https://doi.org/10.1101/2023.05.15.540505.

An, Y.-F. *et al.* (2011) 'Clinical and molecular characteristics of immunodysregulation, polyendocrinopathy, enteropathy, X-linked syndrome in China', *Scandinavian journal of immunology*, 74(3), pp. 304–309.

Arvey, A. *et al.* (2014) 'Inflammation-induced repression of chromatin bound by the transcription factor Foxp3 in regulatory T cells', *Nature immunology*, 15(6), pp. 580–587.

Bacchetta, R. *et al.* (2006) 'Defective regulatory and effector T cell functions in patients with FOXP3 mutations', *The Journal of clinical investigation*, 116(6), pp. 1713–1722.

Bacchetta, R., Barzaghi, F. and Roncarolo, M.-G. (2018) 'From IPEX syndrome to FOXP3 mutation: a lesson on immune dysregulation', *Annals of the New York Academy of Sciences*, 1417(1), pp. 5–22.

Bailey-Bucktrout, S.L. and Bluestone, J.A. (2011) 'Regulatory T cells: stability revisited', *Trends in immunology*, 32(7), pp. 301–306.

Bandukwala, H.S. *et al.* (2011) 'Structure of a domain-swapped FOXP3 dimer on DNA and its function in regulatory T cells', *Immunity*, 34(4), pp. 479–491.

Barzaghi, F. *et al.* (2018) 'Long-term follow-up of IPEX syndrome patients after different therapeutic strategies: An international multicenter retrospective study', *The Journal of allergy and clinical immunology*, 141(3), pp. 1036–1049.e5.

Baud, O. *et al.* (2001) 'Treatment of the immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) by allogeneic bone marrow transplantation', *The New England journal of medicine*, 344(23), pp. 1758–1762.

Benayoun, B.A., Caburet, S. and Veitia, R.A. (2011) 'Forkhead transcription factors: key players in health and disease', *Trends in genetics: TIG*, 27(6), pp. 224–232.

Bennett, C.L. *et al.* (2001) 'The immune dysregulation, polyendocrinopathy, enteropathy, X-linked

syndrome (IPEX) is caused by mutations of FOXP3', *Nature genetics*, 27(1), pp. 20–21.

Benoist, C. and Mathis, D. (2012) 'Treg cells, life history, and diversity', *Cold Spring Harbor perspectives in biology*, 4(9), p. a007021.

Bettelli, E. *et al.* (2006) 'Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells', *Nature*, 441(7090), pp. 235–238.

Bettelli, E., Dastrange, M. and Oukka, M. (2005) 'Foxp3 interacts with nuclear factor of activated T cells and NF-kappa B to repress cytokine gene expression and effector functions of T helper cells', *Proceedings of the National Academy of Sciences of the United States of America*, 102(14), pp. 5138–5143.

Bettini, M.L. *et al.* (2012) 'Loss of epigenetic modification driven by the Foxp3 transcription factor leads to regulatory T cell insufficiency', *Immunity*, 36(5), pp. 717–730.

Blair, P.J. *et al.* (1994) 'CD4+CD8- T cells are the effector cells in disease pathogenesis in the scurfy (sf) mouse', *Journal of immunology* , 153(8), pp. 3764–3774.

Borna, Š. *et al.* (2022) 'Loss of FOXP3 function causes expansion of two pools of autoreactive T cells in patients with IPEX syndrome', *bioRxiv*. Available at: https://doi.org/10.1101/2022.07.10.499494.

Borna, S. *et al.* (2022) 'Towards gene therapy for IPEX syndrome', *European journal of immunology*, 52(5), pp. 705–716.

Brunkow, M.E. *et al.* (2001) 'Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse', *Nature genetics*, 27(1), pp. 68–73.

Burchill, M.A. *et al.* (2008) 'Linked T cell receptor and cytokine signaling govern the development of the regulatory T cell repertoire', *Immunity*, 28(1), pp. 112–121.

Burzyn, D. *et al.* (2013) 'A special population of regulatory T cells potentiates muscle repair', *Cell*, 155(6), pp. 1282–1295.

Caruso, C. *et al.* (2022) 'Case report: Dupilumab treatment improved type 2 disorders in a patient with IPEX syndrome diagnosis', *Frontiers in immunology*, 13, p. 995304.

Cepika, A.-M. *et al.* (2018) 'Tregopathies: Monogenic diseases resulting in regulatory T-cell deficiency', *The Journal of allergy and clinical immunology*, 142(6), pp. 1679–1695.

Chae, W.-J. *et al.* (2006) 'The mutant leucine-zipper domain impairs both dimerization and suppressive function of Foxp3 in T cells', *Proceedings of the National Academy of Sciences of the United States of America*, 103(25), pp. 9631–9636.

Charbonnier, L.-M. *et al.* (2019) 'Functional reprogramming of regulatory T cells in the absence of Foxp3', *Nature immunology*, 20(9), pp. 1208–1219.

Chatila, T.A. *et al.* (12 2000) 'JM2, encoding a fork head–related protein, is mutated in X-linked autoimmunity–allergic disregulation syndrome', *The Journal of clinical investigation*, 106(12), pp. R75–R81.

Chen, C. *et al.* (2006) 'Transcriptional regulation by Foxp3 is associated with direct promoter occupancy and modulation of histone acetylation', *The Journal of biological chemistry*, 281(48), pp. 36828–36834.

Chen, W. *et al.* (2003) 'Conversion of peripheral CD4+CD25- naive T cells to CD4+CD25+ regulatory T cells by TGF-beta induction of transcription factor Foxp3', *The Journal of experimental medicine*, 198(12), pp. 1875–1886.

Chen, Y. *et al.* (2015) 'DNA binding by FOXP3 domain-swapped dimer suggests mechanisms of long-range chromosomal interactions', *Nucleic acids research*, 43(2), pp. 1268–1282.

Chen, Z., Benoist, C. and Mathis, D. (2005) 'How defects in central tolerance impinge on a deficiency in regulatory T cells', *Proceedings of the National Academy of Sciences of the United States of America*, 102(41), pp. 14735–14740.

Chinen, T. *et al.* (2016) 'An essential role for the IL-2 receptor in Treg cell function', *Nature immunology*, 17(11), pp. 1322–1333.

Churlaud, G. *et al.* (2015) 'Human and Mouse CD8(+)CD25(+)FOXP3(+) Regulatory T Cells at Steady State and during Interleukin-2 Therapy', *Frontiers in immunology*, 6, p. 171.

Collison, L.W. *et al.* (2007) 'The inhibitory cytokine IL-35 contributes to regulatory T-cell function', *Nature*, 450(7169), pp. 566–569.

Consonni, F., Ciullini Mannurita, S. and Gambineri, E. (2021) 'Atypical Presentations of IPEX: Expect the Unexpected', *Frontiers in pediatrics*, 9, p. 643094.

Dai, S. *et al.* (2021) 'Toward a mechanistic understanding of DNA binding by forkhead transcription factors and its perturbation by pathogenic mutations', *Nucleic acids research*, 49(18), pp. 10235–10249.

Darce, J. *et al.* (2012) 'An N-terminal mutation of the Foxp3 transcription factor alleviates arthritis but exacerbates diabetes', *Immunity*, 36(5), pp. 731–741.

D'Cruz, L.M. and Klein, L. (2005) 'Development and function of agonist-induced CD25+Foxp3+ regulatory T cells in the absence of interleukin 2 signaling', *Nature immunology*, 6(11), pp. 1152–1159.

Deaglio, S. *et al.* (2007) 'Adenosine generation catalyzed by CD39 and CD73 expressed on regulatory T cells mediates immune suppression', *The Journal of experimental medicine*, 204(6), pp. 1257–1265.

Delacher, M. *et al.* (2017) 'Genome-wide DNA-methylation landscape defines specialization of regulatory T cells in tissues', *Nature immunology*, 18(10), pp. 1160–1172.

Delacher, M. *et al.* (2020) 'Precursors for Nonlymphoid-Tissue Treg Cells Reside in Secondary Lymphoid Organs and Are Programmed by the Transcription Factor BATF', *Immunity*, 52(2), pp. 295–312.e11.

Deng, G. *et al.* (2019) 'Foxp3 Post-translational Modifications and Treg Suppressive Activity', *Frontiers in immunology*, 10, p. 2486.

Dikiy, S. *et al.* (2021) 'A distal Foxp3 enhancer enables interleukin-2 dependent thymic Treg cell lineage commitment for robust immune tolerance', *Immunity*, 54(5), pp. 931–946.e11.

Dikiy, S. and Rudensky, A.Y. (2023) ‘Principles of regulatory T cell function’, *Immunity*, 56(2), pp. 240–255.

DiSpirito, J.R. *et al.* (2018) ‘Molecular diversification of regulatory T cells in nonlymphoid tissues’, *Science immunology*, 3(27). Available at: https://doi.org/10.1126/sciimmunol.aat5861.

Duarte, J.H. *et al.* (2009) ‘Natural Treg cells spontaneously differentiate into pathogenic helper cells in lymphopenic conditions’, *European journal of immunology*, 39(4), pp. 948–955.

Duclaux-Loras, R. *et al.* (2018) ‘Clinical Heterogeneity of Immune Dysregulation, Polyendocrinopathy, Enteropathy, X-Linked Syndrome: A French Multicenter Retrospective Study’, *Clinical and translational gastroenterology*, 9(10), p. 201.

Du, J. *et al.* (2022) ‘FOXP3 exon 2 controls Treg stability and autoimmunity’, *Science immunology*, 7(72), p. eabo5407.

Eriksson, D. *et al.* (2019) ‘The autoimmune targets in IPEX are dominated by gut epithelial proteins’, *The Journal of allergy and clinical immunology*, 144(1), pp. 327–330.e8.

Fantini, M.C. *et al.* (2004) ‘Cutting edge: TGF-beta induces a regulatory phenotype in CD4+CD25- T cells through Foxp3 induction and down-regulation of Smad7’, *Journal of immunology* , 172(9), pp. 5149–5153.

Feng, Y. *et al.* (2014) ‘Control of the inheritance of regulatory T cell identity by a cis element in the Foxp3 locus’, *Cell*, 158(4), pp. 749–763.

Feng, Y. *et al.* (2015) ‘A mechanism for expansion of regulatory T-cell repertoire and its role in self-tolerance’, *Nature*, 528(7580), pp. 132–136.

Ferraro, A. *et al.* (2014) ‘Interindividual variation in human T regulatory cells’, *Proceedings of the National Academy of Sciences of the United States of America*, 111(12), pp. E1111–20.

Feuerer, M., Hill, J.A., *et al.* (2009) ‘Foxp3+ regulatory T cells: differentiation, specification, subphenotypes’, *Nature immunology*, 10(7), pp. 689–695.

Feuerer, M., Herrero, L., *et al.* (2009) ‘Lean, but not obese, fat is enriched for a unique population of regulatory T cells that affect metabolic parameters’, *Nature medicine*, 15(8), pp. 930–939.

Feuerer, M. *et al.* (2010) ‘Genomic definition of multiple ex vivo regulatory T cell subphenotypes’, *Proceedings of the National Academy of Sciences of the United States of America*, 107(13), pp. 5919–5924.

Fontenot, J.D., Rasmussen, J.P., Gavin, M.A., *et al.* (2005) ‘A function for interleukin 2 in Foxp3-expressing regulatory T cells’, *Nature immunology*, 6(11), pp. 1142–1151.

Fontenot, J.D., Rasmussen, J.P., Williams, L.M., *et al.* (2005) ‘Regulatory T cell lineage specification by the forkhead transcription factor foxp3’, *Immunity*, 22(3), pp. 329–341.

Fontenot, J.D., Gavin, M.A. and Rudensky, A.Y. (2003) ‘Foxp3 programs the development and function of CD4+CD25+ regulatory T cells’, *Nature immunology*, 4(4), pp. 330–336.

Frith, K. *et al.* (2019) 'The FOXP3Δ2 isoform supports Treg cell development and protects against severe IPEX syndrome', *The Journal of allergy and clinical immunology*, 144(1), pp. 317–320.e8.

Fu, W. *et al.* (2012) 'A multiply redundant genetic switch "locks in" the transcriptional signature of regulatory T cells', *Nature immunology*, 13(10), pp. 972–980.

Gambineri, E. *et al.* (2008) 'Clinical and molecular profile of a new series of patients with immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome: inconsistent correlation between forkhead box protein 3 expression and disease severity', *The Journal of allergy and clinical immunology*, 122(6), pp. 1105–1112.e1.

Gambineri, E. *et al.* (2018) 'Clinical, Immunological, and Molecular Heterogeneity of 173 Patients With the Phenotype of Immune Dysregulation, Polyendocrinopathy, Enteropathy, X-Linked (IPEX) Syndrome', *Frontiers in immunology*, 9, p. 2411.

Gavin, M.A. *et al.* (2007) 'Foxp3-dependent programme of regulatory T-cell differentiation', *Nature*, 445(7129), pp. 771–775.

Georgiev, P., Charbonnier, L.-M. and Chatila, T.A. (2019) 'Regulatory T Cells: the Many Faces of Foxp3', *Journal of clinical immunology*, 39(7), pp. 623–640.

Gerbaux, M. *et al.* (2023) 'CTLA4-Ig Effectively Controls Clinical Deterioration and Immune Condition in a Murine Model of Foxp3 Deficiency', *Journal of clinical immunology* [Preprint]. Available at: https://doi.org/10.1007/s10875-023-01462-2.

Gocher-Demske, A.M. *et al.* (2023) 'IFNγ-induction of TH1-like regulatory T cells controls antiviral responses', *Nature immunology* [Preprint]. Available at: https://doi.org/10.1038/s41590-023-01453-w.

Godfrey, V.L., Wilkinson, J.E. and Russell, L.B. (1991) 'X-linked lymphoreticular disease in the scurfy (sf) mutant mouse', *The American journal of pathology*, 138(6), pp. 1379–1387.

Goettel, J.A. *et al.* (2015) 'Fatal autoimmunity in mice reconstituted with human hematopoietic stem cells encoding defective FOXP3', *Blood*, 125(25), pp. 3886–3895.

Goodwin, M. *et al.* (2020) 'CRISPR-based gene editing enables *FOXP3* gene repair in IPEX patient cells', *Science advances*, 6(19), p. eaaz0571.

Groux, H. *et al.* (1997) 'A CD4+ T-cell subset inhibits antigen-specific T-cell responses and prevents colitis', *Nature*, 389(6652), pp. 737–742.

Hayatsu, N. *et al.* (2017) 'Analyses of a Mutant Foxp3 Allele Reveal BATF as a Critical Transcription Factor in the Differentiation and Accumulation of Tissue Regulatory T Cells', *Immunity*, 47(2), pp. 268–283.e9.

Herppich, S. *et al.* (2019) 'Dynamic Imprinting of the Treg Cell-Specific Epigenetic Signature in Developing Thymic Regulatory T Cells', *Frontiers in immunology*, 10, p. 2382.

Hill, J.A. *et al.* (2007) 'Foxp3 transcription-factor-dependent and -independent regulation of the regulatory T cell transcriptional signature', *Immunity*, 27(5), pp. 786–800.

Hori, S., Nomura, T. and Sakaguchi, S. (2003) 'Control of regulatory T cell development by the transcription factor Foxp3', *Science*, 299(5609), pp. 1057–1061.

Hoshino, A. *et al.* (2019) 'Identification of autoantibodies using human proteome microarrays in patients with IPEX syndrome', *Clinical immunology* , 203, pp. 9–13.

Hsieh, C.-S. *et al.* (2006) 'An intersection between the self-reactive regulatory and nonregulatory T cell receptor repertoires', *Nature immunology*, 7(4), pp. 401–410.

Huang, C.-T. *et al.* (2004) 'Role of LAG-3 in regulatory T cells', *Immunity*, 21(4), pp. 503–513.

Jackson, B.C. *et al.* (2010) 'Update of human and mouse forkhead box (FOX) gene families', *Human genomics*, 4(5), pp. 345–352.

Joly, A.-L. *et al.* (2015) 'Foxp3 lacking exons 2 and 7 is unable to confer suppressive ability to regulatory T cells in vivo', *Journal of autoimmunity*, 63, pp. 23–30.

Josefowicz, S.Z. *et al.* (2012) 'Extrathymically generated regulatory T cells control mucosal TH2 inflammation', *Nature*, 482(7385), pp. 395–399.

Josefowicz, S.Z., Lu, L.-F. and Rudensky, A.Y. (2012) 'Regulatory T cells: mechanisms of differentiation and function', *Annual review of immunology*, 30, pp. 531–564.

Kawakami, R. *et al.* (2021) 'Distinct Foxp3 enhancer elements coordinate development, maintenance, and function of regulatory T cells', *Immunity*, 54(5), pp. 947–961.e8.

Khattri, R. *et al.* (2003) 'An essential role for Scurfin in CD4+CD25+ T regulatory cells', *Nature immunology*, 4(4), pp. 337–342.

Kim, J.M., Rasmussen, J.P. and Rudensky, A.Y. (2007) 'Regulatory T cells prevent catastrophic autoimmunity throughout the lifespan of mice', *Nature immunology*, 8(2), pp. 191–197.

Kinnunen, T. *et al.* (2013) 'Accumulation of peripheral autoreactive B cells in the absence of functional human regulatory T cells', *Blood*, 121(9), pp. 1595–1603.

Kitagawa, Y. *et al.* (2017) 'Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment', *Nature immunology*, 18(2), pp. 173–183.

Kitoh, A. *et al.* (2009) 'Indispensable role of the Runx1-Cbfβ transcription complex for in vivo-suppressive function of FoxP3+ regulatory T cells', *Immunity*, 31(4), pp. 609–620.

Koch, M.A. *et al.* (2009) 'The transcription factor T-bet controls regulatory T cell homeostasis and function during type 1 inflammation', *Nature immunology*, 10(6), pp. 595–602.

Kumagai, S. *et al.* (2020) 'An Oncogenic Alteration Creates a Microenvironment that Promotes Tumor Progression by Conferring a Metabolic Advantage to Regulatory T Cells', *Immunity*, 53(1), pp. 187–203.e8.

Kwon, H.-K. *et al.* (2008) 'Foxp3 induces IL-4 gene silencing by affecting nuclear translocation of NFkappaB and chromatin structure', *Molecular immunology*, 45(11), pp. 3205–3212.

Kwon, H.-K. *et al.* (2017) 'Different molecular complexes that mediate transcriptional induction and repression by FoxP3', *Nature immunology*, 18(11), pp. 1238–1248.

Kwon, H.-K. *et al.* (2018) 'FoxP3 scanning mutagenesis reveals functional variegation and mild mutations with atypical autoimmune phenotypes', *Proceedings of the National Academy of Sciences of the United States of America*, 115(2), pp. E253–E262.

Lahl, K. *et al.* (2007) 'Selective depletion of Foxp3+ regulatory T cells induces a scurfy-like disease', *The Journal of experimental medicine*, 204(1), pp. 57–63.

Lathrop, S.K. *et al.* (2011) 'Peripheral education of the immune system by colonic commensal microbiota', *Nature*, 478(7368), pp. 250–254.

Lee, H.-M. *et al.* (2012) 'A broad range of self-reactivity drives thymic regulatory T cell selection to limit responses to self', *Immunity*, 37(3), pp. 475–486.

Levine, A.G. *et al.* (2014) 'Continuous requirement for the TCR in regulatory T cell function', *Nature immunology*, 15(11), pp. 1070–1078.

Levine, A.G. *et al.* (2017) 'Stability and function of regulatory T cells expressing the transcription factor T-bet', *Nature*, 546(7658), pp. 421–425.

Li, B. *et al.* (2007) 'FOXP3 interactions with histone acetyltransferase and class II histone deacetylases are required for repression', *Proceedings of the National Academy of Sciences of the United States of America*, 104(11), pp. 4571–4576.

Li, M.O., Wan, Y.Y. and Flavell, R.A. (2007) 'T cell-produced transforming growth factor-beta1 controls T cell tolerance and regulates Th1- and Th17-cell differentiation', *Immunity*, 26(5), pp. 579–591.

Lin, W. *et al.* (2005) 'Allergic dysregulation and hyperimmunoglobulinemia E in Foxp3 mutant mice', *The Journal of allergy and clinical immunology*, 116(5), pp. 1106–1115.

Lin, W. *et al.* (2007) 'Regulatory T cell development in the absence of functional Foxp3', *Nature immunology*, 8(4), pp. 359–368.

Lio, C.-W.J. *et al.* (2010) 'CD28 facilitates the generation of Foxp3(-) cytokine responsive regulatory T cell precursors', *Journal of immunology* , 184(11), pp. 6007–6013.

Lio, C.-W.J. and Hsieh, C.-S. (2008) 'A two-step process for thymic regulatory T cell development', *Immunity*, 28(1), pp. 100–111.

Liston, A. *et al.* (2008) 'Differentiation of regulatory Foxp3+ T cells in the thymic cortex', *Proceedings of the National Academy of Sciences of the United States of America*, 105(33), pp. 11903–11908.

Liston, A. and Aloulou, M. (2022) 'A fresh look at a neglected regulatory lineage: CD8+Foxp3+ Regulatory T cells', *Immunology letters*, 247, pp. 22–26.

Li, X. *et al.* (2014) 'Function of a Foxp3 cis-element in protecting regulatory T cell identity', *Cell*, 158(4), pp. 734–748.

Lopes, J.E. *et al.* (2006) 'Analysis of FOXP3 reveals multiple domains required for its function as a transcriptional repressor', *Journal of immunology* , 177(5), pp. 3133–3142.

Masiuk, K.E. *et al.* (2019) 'Lentiviral Gene Therapy in HSCs Restores Lineage-Specific Foxp3 Expression and Suppresses Autoimmunity in a Mouse Model of IPEX Syndrome', *Cell stem cell*, 24(2), pp. 309–317.e7.

Ménoret, S. *et al.* (2023) 'CD4+ and CD8+ regulatory T cell characterization in the rat using a unique transgenic Foxp3-EGFP model', *BMC biology*, 21(1), p. 8.

Miyao, T. *et al.* (2012) 'Plasticity of Foxp3(+) T cells reflects promiscuous Foxp3 expression in conventional T cells but not reprogramming of regulatory T cells', *Immunity*, 36(2), pp. 262–275.

Miyara, M. *et al.* (2009) 'Functional delineation and differentiation dynamics of human CD4+ T cells expressing the FoxP3 transcription factor', *Immunity*, 30(6), pp. 899–911.

Morikawa, H. *et al.* (2014) 'Differential roles of epigenetic changes and Foxp3 expression in regulatory T cell-specific transcriptional regulation', *Proceedings of the National Academy of Sciences of the United States of America*, 111(14), pp. 5289–5294.

Muñoz-Rojas, A.R. and Mathis, D. (2021) 'Tissue regulatory T cells: regulatory chameleons', *Nature reviews. Immunology*, 21(9), pp. 597–611.

Narula, M. *et al.* (2022) 'Epigenetic and Immunological Indicators of IPEX Disease in subjects with FOXP3 gene mutation', *The Journal of allergy and clinical immunology* [Preprint]. Available at: https://doi.org/10.1016/j.jaci.2022.09.013.

Nieves, D.S. *et al.* (2004) 'Dermatologic and immunologic findings in the immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome', *Archives of dermatology*, 140(4), pp. 466–472.

Ohkura, N. *et al.* (2012) 'T cell receptor stimulation-induced epigenetic changes and Foxp3 expression are independent and complementary events required for Treg cell development', *Immunity*, 37(5), pp. 785–799.

Ohkura, N. *et al.* (2020) 'Regulatory T Cell-Specific Epigenomic Region Variants Are a Key Determinant of Susceptibility to Common Autoimmune Diseases', *Immunity*, 52(6), pp. 1119–1132.e4.

Okou, D.T. *et al.* (2014) 'Exome sequencing identifies a novel FOXP3 mutation in a 2-generation family with inflammatory bowel disease', *Journal of pediatric gastroenterology and nutrition*, 58(5), pp. 561–568.

Oldenhove, G. *et al.* (2009) 'Decrease of Foxp3+ Treg cell number and acquisition of effector cell phenotype during lethal infection', *Immunity*, 31(5), pp. 772–786.

Ono, M. *et al.* (2007) 'Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1', *Nature*, 446(7136), pp. 685–689.

Otsubo, K. *et al.* (2011) 'Identification of FOXP3-negative regulatory T-like (CD4(+)CD25(+)CD127(low)) cells in patients with immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome', *Clinical immunology* , 141(1), pp. 111–120.

Owen, D.L. *et al.* (2019) 'Thymic regulatory T cells arise via two distinct developmental programs', *Nature immunology*, 20(2), pp. 195–205.

Park, J.H. *et al.* (2020) 'Immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome: A systematic review', *Autoimmunity reviews*, p. 102526.

Passerini, L. *et al.* (2020) 'Treatment with rapamycin can restore regulatory T-cell function in IPEX patients', *The Journal of allergy and clinical immunology*, 145(4), pp. 1262–1271.e13.

Picarda, E. *et al.* (2019) 'Cross-Reactive Donor-Specific CD8+ Tregs Efficiently Prevent Transplant Rejection', *Cell reports*, 29(13), pp. 4245–4255.e6.

Powell, B.R., Buist, N.R. and Stenzel, P. (1982) 'An X-linked syndrome of diarrhea, polyendocrinopathy, and fatal infection in infancy', *The Journal of pediatrics*, 100(5), pp. 731–737.

Ramanan, D. *et al.* (2023) 'Regulatory T cells in the face of the intestinal microbiota', *Nature reviews. Immunology* [Preprint]. Available at: https://doi.org/10.1038/s41577-023-00890-w.

Ramirez, R.N. *et al.* (2022) 'FoxP3 associates with enhancer-promoter loops to regulate Treg-specific gene expression', *Science immunology*, 7(67), p. eabj9836.

Roncarolo, M.G. *et al.* (2014) 'Tr1 cells and the counter-regulation of immunity: natural mechanisms and therapeutic applications', *Current topics in microbiology and immunology*, 380, pp. 39–68.

Rosenblum, M.D. *et al.* (2011) 'Response to self antigen imprints regulatory memory in tissues', *Nature*, 480(7378), pp. 538–542.

Rubtsov, Y.P. *et al.* (2008) 'Regulatory T cell-derived interleukin-10 limits inflammation at environmental interfaces', *Immunity*, 28(4), pp. 546–558.

Rubtsov, Y.P. *et al.* (2010) 'Stability of the regulatory T cell lineage in vivo', *Science*, 329(5999), pp. 1667–1671.

Rudra, D. *et al.* (2012) 'Transcription factor Foxp3 and its protein partners form a complex regulatory network', *Nature immunology*, 13(10), pp. 1010–1019.

Russell, W.L., Russell, L.B. and Gower, J.S. (1959) 'EXCEPTIONAL INHERITANCE OF A SEX-LINKED GENE IN THE MOUSE EXPLAINED ON THE BASIS THAT THE X/O SEX-CHROMOSOME CONSTITUTION IS FEMALE', *Proceedings of the National Academy of Sciences of the United States of America*, 45(4), pp. 554–560.

Russler-Germain, E.V., Rengarajan, S. and Hsieh, C.-S. (2017) 'Antigen-specific regulatory T-cell responses to intestinal microbiota', *Mucosal immunology*, 10(6), pp. 1375–1386.

Sakaguchi, S. *et al.* (1995) 'Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases', *Journal of immunology* , 155(3), pp. 1151–1164.

Sakaguchi, S. *et al.* (2020) 'Regulatory T Cells and Human Disease', *Annual review of immunology*, 38, pp. 541–566.

Salomon, B. *et al.* (2000) 'B7/CD28 costimulation is essential for the homeostasis of the CD4+CD25+ immunoregulatory T cells that control autoimmune diabetes', *Immunity*, 12(4), pp. 431–440.

Samstein, R.M., Josefowicz, S.Z., *et al.* (2012) 'Extrathymic generation of regulatory T cells in placental mammals mitigates maternal-fetal conflict', *Cell*, 150(1), pp. 29–38.

Samstein, R.M., Arvey, A., *et al.* (2012) 'Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification', *Cell*, 151(1), pp. 153–166.

Sato, Y. *et al.* (2021) 'Co-Expression of FOXP3FL and FOXP3Δ2 Isoforms Is Required for Optimal Treg-Like Cell Phenotypes and Suppressive Function', *Frontiers in immunology*, 12, p. 752394.

Schiering, C. *et al.* (2014) 'The alarmin IL-33 promotes regulatory T-cell function in the intestine', *Nature*, 513(7519), pp. 564–568.

Sefik, E. *et al.* (2015) 'MUCOSAL IMMUNOLOGY. Individual intestinal symbionts induce a distinct population of RORγ^+ regulatory T cells', *Science*, 349(6251), pp. 993–997.

Seitz, C. *et al.* (2022) 'The FOXP3 full-length isoform controls the lineage-stability of CD4+FOXP3+ regulatory T cells', *Clinical immunology* , 237, p. 108957.

Shamriz, O. *et al.* (2018) 'Hypogammaglobulinemia with decreased class-switched B-cells and dysregulated T-follicular-helper cells in IPEX syndrome', *Clinical immunology* , 197, pp. 219–223.

Singh, R.P. *et al.* (2007) 'CD8+ T cell-mediated suppression of autoimmunity in a murine lupus model of peptide-induced immune tolerance depends on Foxp3 expression', *Journal of immunology* , 178(12), pp. 7649–7657.

Solomon, B.D. and Hsieh, C.-S. (2016) 'Antigen-Specific Development of Mucosal Foxp3+RORγt+ T Cells from Regulatory T Cell Precursors', *Journal of immunology* , 197(9), pp. 3512–3519.

Song, X. *et al.* (2012) 'Structural and biological features of FOXP3 dimerization relevant to regulatory T cell function', *Cell reports*, 1(6), pp. 665–675.

Sugimoto, N. *et al.* (2006) 'Foxp3-dependent and -independent molecules specific for CD25+CD4+ natural regulatory T cells revealed by DNA microarray analysis', *International immunology*, 18(8), pp. 1197–1209.

Tai, Y. *et al.* (2019) 'Dysregulation of humoral immunity in Foxp3 conditional-knockout mice', *Biochemical and biophysical research communications*, 513(4), pp. 787–793.

Takahashi, T. *et al.* (2000) 'Immunologic self-tolerance maintained by CD25(+)CD4(+) regulatory T cells constitutively expressing cytotoxic T lymphocyte-associated antigen 4', *The Journal of experimental medicine*, 192(2), pp. 303–310.

Tanaka, A. *et al.* (2023) 'Construction of a T cell receptor signaling range for spontaneous development of autoimmune disease', *The Journal of experimental medicine*, 220(2). Available at: https://doi.org/10.1084/jem.20220386.

Tarbell, K.V. *et al.* (2004) 'CD25+ CD4+ T cells, expanded with dendritic cells presenting a single

autoantigenic peptide, suppress autoimmune diabetes', *The Journal of experimental medicine*, 199(11), pp. 1467–1477.

Tommasini, A. *et al.* (2002) 'X-chromosome inactivation analysis in a female carrier of FOXP3 mutation', *Clinical and experimental immunology*, 130(1), pp. 127–130.

Tone, Y. *et al.* (2008) 'Smad3 and NFAT cooperate to induce Foxp3 expression through its enhancer', *Nature immunology*, 9(2), pp. 194–202.

Trujillo-Ochoa, J.L., Kazemian, M. and Afzali, B. (2023) 'The role of transcription factors in shaping regulatory T cell identity', *Nature reviews. Immunology* [Preprint]. Available at: https://doi.org/10.1038/s41577-023-00893-7.

Tsuda, M. *et al.* (2010) 'The spectrum of autoantibodies in IPEX syndrome is broad and includes anti-mitochondrial autoantibodies', *Journal of autoimmunity*, 35(3), pp. 265–268.

Van Gool, F. *et al.* (2019) 'A Mutation in the Transcription Factor Foxp3 Drives T Helper 2 Effector Function in Regulatory T Cells', *Immunity*, 50(2), pp. 362–377.e6.

van der Veeken, J. *et al.* (2020) 'The Transcription Factor Foxp3 Shapes Regulatory T Cell Identity by Tuning the Activity of trans-Acting Intermediaries', *Immunity*, 53(5), pp. 971–984.e5.

van der Veeken, J. *et al.* (2022) 'Genetic tracing reveals transcription factor Foxp3-dependent and Foxp3-independent functionality of peripherally induced Treg cells', *Immunity*, 55(7), pp. 1173–1184.e7.

Wan, Y.Y. and Flavell, R.A. (2005) 'Identifying Foxp3-expressing suppressor T cells with a bicistronic reporter', *Proceedings of the National Academy of Sciences of the United States of America*, 102(14), pp. 5126–5131.

Weigel, D. *et al.* (1989) 'The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo', *Cell*, 57(4), pp. 645–658.

Wildin, R.S. *et al.* (2001) 'X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy', *Nature genetics*, 27(1), pp. 18–20.

Williams, L.M. and Rudensky, A.Y. (2007) 'Maintenance of the Foxp3-dependent developmental program in mature regulatory T cells requires continued expression of Foxp3', *Nature immunology*, 8(3), pp. 277–284.

Wing, K. *et al.* (2008) 'CTLA-4 control over Foxp3+ regulatory T cell function', *Science*, 322(5899), pp. 271–275.

Wu, Y. *et al.* (2006) 'FOXP3 controls regulatory T cell function through cooperation with NFAT', *Cell*, 126(2), pp. 375–387.

Wyatt, R.C. *et al.* (2023) 'FOXP3 TSDR Measurement Could Assist Variant Classification and Diagnosis of IPEX Syndrome', *Journal of clinical immunology* [Preprint]. Available at: https://doi.org/10.1007/s10875-022-01428-w.

Xavier-da-Silva, M.M. *et al.* (2015) 'Fetal-onset IPEX: report of two families and review of literature',

Clinical immunology , 156(2), pp. 131–140.

Xiao, Y. *et al.* (2010) 'Histone acetyltransferase mediated regulation of FOXP3 acetylation and Treg function', *Current opinion in immunology*, 22(5), pp. 583–591.

Yang, X.O. *et al.* (2008) 'Molecular antagonism and plasticity of regulatory and inflammatory T cell programs', *Immunity*, 29(1), pp. 44–56.

Zaret, K.S. (2020) 'Pioneer Transcription Factors Initiating Gene Network Changes', *Annual review of genetics*, 54, pp. 367–385.

Zemmour, D. *et al.* (2018) 'Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR', *Nature immunology*, 19(3), pp. 291–301.

Zemmour, D. *et al.* (2021) 'Single-cell analysis of FOXP3 deficiencies in humans and mice unmasks intrinsic and extrinsic CD4+ T cell perturbations', *Nature immunology*, 22(5), pp. 607–619.

Zeng, W.-P., Sollars, V.E. and Belalcazar, A.D.P. (2011) 'Domain requirements for the diverse immune regulatory functions of foxp3', *Molecular immunology*, 48(15-16), pp. 1932–1939.

Zheng, Y. *et al.* (2007) 'Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells', *Nature*, 445(7130), pp. 936–940.

Zheng, Y. *et al.* (2009) 'Regulatory T-cell suppressor program co-opts transcription factor IRF4 to control T(H)2 responses', *Nature*, 458(7236), pp. 351–356.

Zheng, Y. *et al.* (2010) 'Role of conserved non-coding DNA elements in the Foxp3 gene in regulatory T-cell fate', *Nature*, 463(7282), pp. 808–812.

Zhou, X. *et al.* (2009) 'Instability of the transcription factor Foxp3 leads to the generation of pathogenic memory T cells in vivo', *Nature immunology*, 10(9), pp. 1000–1007.

APPENDIX A, *REFERS TO CHAPTER 1*

nature immunology

Articles

Single-cell analysis of FOXP3 deficiencies in humans and mice unmasks intrinsic and extrinsic CD4+ T cell perturbations

David Zemmour[1,2], Louis-Marie Charbonnier [3], Juliette Leon[1,4], Emmanuelle Six [4], Sevgi Keles [5], Marianne Delville [4,6], Mehdi Benamar[3], Safa Baris [7,8], Julien Zuber[4,6], Karin Chen[9,10], Benedicte Neven[4,6], Maria I. Garcia-Lloret[11], Frank M. Ruemmele[4,6], Carlo Brugnara[3], Nadine Cerf-Bensussan[4,6], Frederic Rieux-Laucat[4], Marina Cavazzana[12,13,14], Isabelle André[4,14], Talal A. Chatila [3], Diane Mathis [1]✉ and Christophe Benoist [1]✉

***FOXP3* deficiency in mice and in patients with immune dysregulation polyendocrinopathy enteropathy X-linked (IPEX) syn- drome results in fatal autoimmunity by altering regulatory T (T_{reg}) cells. $CD4^+$ T cells in patients with IPEX syndrome and *Foxp3*-deficient mice were analyzed by single-cell cytometry and RNA-sequencing, revealing heterogeneous T_{reg}-like cells, some very similar to normal T_{reg} cells, others more distant. Conventional T cells showed no widespread activation or helper T cell bias, but a monomorphic disease signature affected all $CD4^+$ T cells. This signature proved to be cell extrinsic since it was extinguished in mixed bone marrow chimeric mice and heterozygous mothers of patients with IPEX syndrome. Normal T_{reg} cells exerted dominant suppression, quenching the disease signature and revealing in mutant T_{reg}-like cells a small cluster of genes regulated cell-intrinsically by FOXP3, including key homeostatic regulators. We propose a two-step pathogenesis model: cell-intrinsic downregulation of core FOXP3-dependent genes destabilizes T_{reg} cells, de-repressing systemic mediators that imprint the disease signature on all T cells, furthering T_{reg} cell dysfunction. Accordingly, interleukin-2 treatment improved the T_{reg}-like compartmenT**

nature immunology

ARTICLES

https://doi.org/10.1038/s41590-021-00910-8

Check for updates

Single-cell analysis of FOXP3 deficiencies in humans and mice unmasks intrinsic and extrinsic CD4$^+$ T cell perturbations

David Zemmour[1,2], Louis-Marie Charbonnier[3], Juliette Leon[1,4], Emmanuelle Six[4], Sevgi Keles[5], Marianne Delville[4,6], Mehdi Benamar[3], Safa Baris[7,8], Julien Zuber[4,6], Karin Chen[9,10], Benedicte Neven[4,6], Maria I. Garcia-Lloret[11], Frank M. Ruemmele[4,6], Carlo Brugnara[3], Nadine Cerf-Bensussan[4,6], Frederic Rieux-Laucat[4], Marina Cavazzana[12,13,14], Isabelle André[4,14], Talal A. Chatila[3], Diane Mathis[1]✉ and Christophe Benoist[1]✉

***FOXP3* deficiency in mice and in patients with immune dysregulation polyendocrinopathy enteropathy X-linked (IPEX) syndrome results in fatal autoimmunity by altering regulatory T (T_{reg}) cells. CD4$^+$ T cells in patients with IPEX syndrome and *Foxp3*-deficient mice were analyzed by single-cell cytometry and RNA-sequencing, revealing heterogeneous T_{reg}-like cells, some very similar to normal T_{reg} cells, others more distant. Conventional T cells showed no widespread activation or helper T cell bias, but a monomorphic disease signature affected all CD4$^+$ T cells. This signature proved to be cell extrinsic since it was extinguished in mixed bone marrow chimeric mice and heterozygous mothers of patients with IPEX syndrome. Normal T_{reg} cells exerted dominant suppression, quenching the disease signature and revealing in mutant T_{reg}-like cells a small cluster of genes regulated cell-intrinsically by FOXP3, including key homeostatic regulators. We propose a two-step pathogenesis model: cell-intrinsic downregulation of core FOXP3-dependent genes destabilizes T_{reg} cells, de-repressing systemic mediators that imprint the disease signature on all T cells, furthering T_{reg} cell dysfunction. Accordingly, interleukin-2 treatment improved the T_{reg}-like compartment and survival.**

T_{reg} cells that express the transcription factor FOXP3 are dominant negative regulators of many facets of the immune system, controlling immune responses and enforcing peripheral tolerance to self, symbiotic commensals and fetal antigens[1,2]. They also have extra-immunological roles in maintaining tissue homeostasis outside of the purely immunological context[3]. T_{reg} cells share a common transcriptional signature, genes differentially expressed relative to their conventional CD4$^+$ counterparts (T_{conv}), in mice and humans[4–6]. Much of the T_{reg} signature is controlled by FOXP3, the lineage's defining transcription factor (TF), although FOXP3 is neither fully necessary nor sufficient to establish the T_{reg} transcriptome, requiring input from transcriptional cofactors[7,8]. Consistent with their pleiotropic functions, T_{reg} cells show a range of phenotypic variation, differing by activation state, effector capabilities and tissue localization[3,9]. T_{reg} heterogeneity has been further refined by single-cell transcriptomics (scRNA-seq)[6,10–12].

IPEX syndrome is perhaps the prototype of monogenic autoimmune syndromes[13], resulting from FOXP3 loss of function (LOF) and thus T_{reg} cell deficiency[1,14]. X-linked, it is transmitted by asymptomatic carrier females to their male progeny who usually present early after birth with a constellation of autoimmune manifestations dominated by severe enteropathy, eczematous dermatitis and type 1 diabetes, with some less common manifestations (nephropathy, hypoparathyroidism and antibody-mediated cytopenias)[15–18]. There are several causes to the variable severity observed in patients with IPEX syndrome. First, complete LOF mutations are generally more deleterious than missense mutations; and mutations in the DNA-binding forkhead or dimerization (leucine zipper) domains are generally more severe than N-terminal mutations, in keeping with results from our alanine-scan dissection of FOXP3 (refs. [8,17]). But manifestations and severity can vary between patients with the same mutation[17,19], suggesting that genetic modifiers, environmental influences or immunological history modify the disease course in each patient. Mice of the spontaneous *scurfy* mutant line, or carrying engineered *Foxp3* LOFs, show a similar rapidly lethal phenotype, dominated by enteropathy, dermatitis and lymphoproliferation[1,20]; partial or slower disease appears in mice with mild *Foxp3* missense mutations[8,21,22].

Also contributing to this variability, the relationship between FOXP3 and T_{reg} cells is now recognized not to be obligate. Several

[1]Department of Immunology, Harvard Medical School, Boston, MA, USA. [2]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [3]Division of Immunology, Boston Children's Hospital, and Department of Pediatrics, Harvard Medical School, Boston, MA, USA. [4]INSERM UMR 1163, University of Paris, Imagine Institute, Paris, France. [5]Division of Pediatric Allergy and Immunology, Meram Medical Faculty, Necmettin Erbakan University, Konya, Turkey. [6]Necker Children's Hospital, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris, France. [7]Division of Pediatric Allergy and Immunology, Marmara University School of Medicine, Istanbul, Turkey. [8]Istanbul Jeffrey Modell Diagnostic and Research Center for Primary Immunodeficiencies, Istanbul, Turkey. [9]Division of Allergy and Immunology, Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, UT, USA. [10]Division of Immunology, Department of Pediatrics, University of Washington and Seattle Children's Research Institute, Seattle, WA, USA. [11]Division of Allergy, Immunology, and Rheumatology, Department of Pediatrics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. [12]Biotherapy Department and Clinical Investigation Center, Assistance Publique Hopitaux Paris, Inserm, Paris, France. [13]Université de Paris, Paris, France. [14]Institut Imagine, Paris, France. ✉e-mail: cbdm@hms.harvard.edu; cbdm@hms.harvard.edu

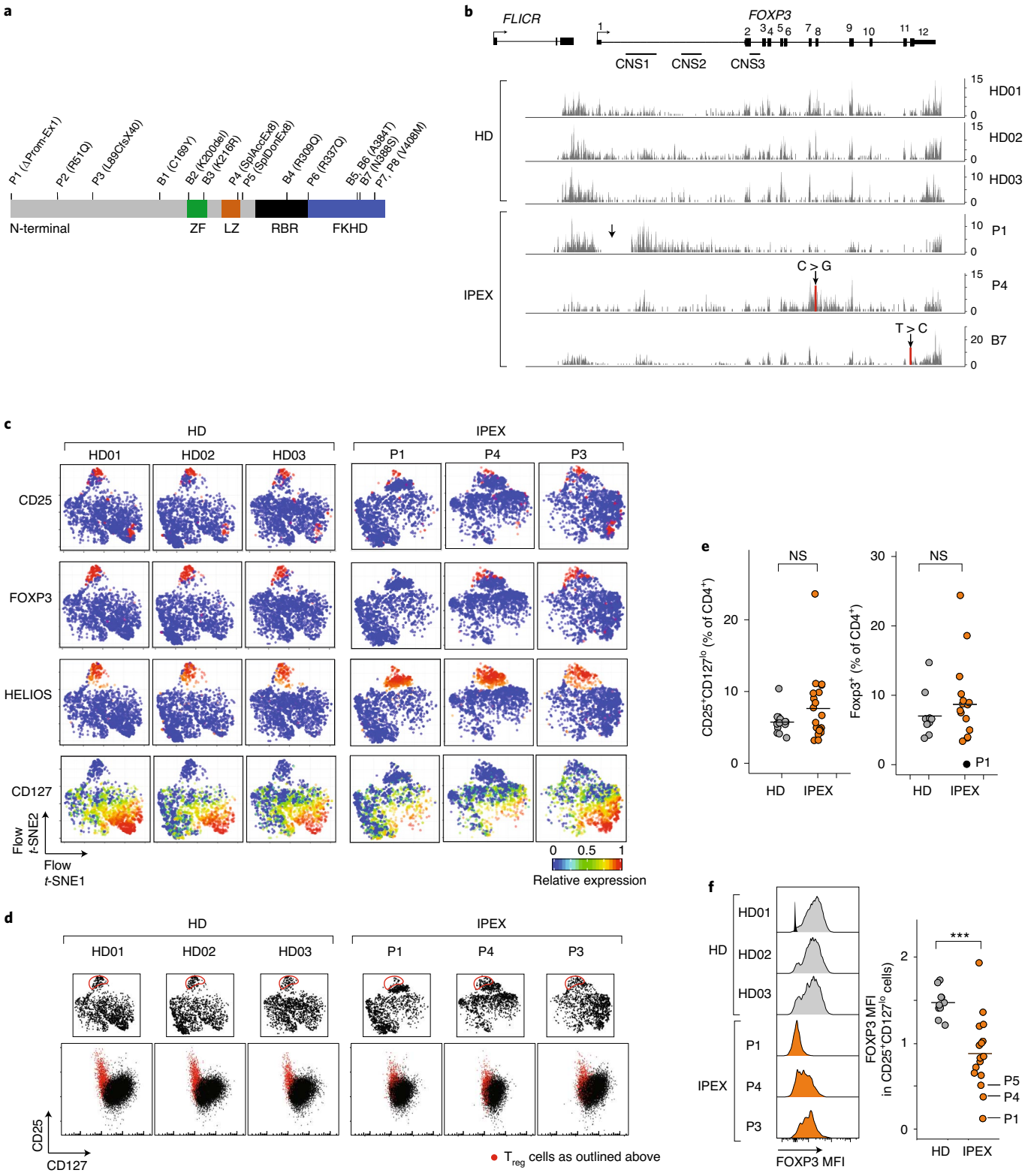


Fig. 1 | Identification of T_{reg}-like cells in IPEX syndrome by flow cytometry. a, Position of FOXP3 mutations in IPEX cohorts ($n=15$ IPEX). ZF, zinc finger; LZ, leucine zipper; RBR, RUNX1-binding region; FKHD, forkhead. **b**, Mapping of RNA-seq reads from CD4$^+$CD25$^+$CD127^{lo} cells to *FOXP3* locus in representative samples (Supplementary Fig. 1). Arrow indicates mutation. **c**, Flow-*t*-SNE plot of CD4$^+$ T cells from anti-CD3, CD4, CD25, CD127, HELIOS, CD45RA and FOXP3 staining in representative samples (full set in Supplementary Fig. 3). Color represents scaled expression. **d**, Flow cytometric anti-CD25/CD127 plots of CD4$^+$ T cells (bottom); red, cells gated in the flow-*t*-SNE plot (top). **e**, Proportion of CD25$^+$CD127^{lo} cells and total FOXP3$^+$ cells in HD and IPEX (each dot is a sample) (Supplementary Fig. 2a), NS, nonsignificant, two-sided Student's *t*-test ($n=15$ IPEX, $n=13$ HD). **f**, FOXP3 expression in CD25$^+$CD127^{lo} HD and IPEX cells. Representative cytometry profiles (left) (Supplementary Fig. 2b for all other samples; one unstained control overlaid with HD01) and quantitation (right) (each dot is a sample; $^{***}P<0.001$, two-sided Student's *t*-test, $n=15$ IPEX, $n=13$ HD).

reports have documented the existence of T_{reg}-like cells in the absence of viable *Foxp3* in *scurfy* mice[23–25] and in some patients with IPEX syndrome[26–28]. In addition, FOXP3 is not exclusive to T_{reg} cells: it is also expressed, albeit at lower levels than in T_{reg} cells, early after T_{conv} cell activation[29–32] and scRNA-seq showed some FOXP3-positive cells otherwise similar to T_{conv} cells[6].

The advent of single-cell transcriptomics opened the potential to revisit the alterations of CD4+ T cells in patients with IPEX syndrome, which were previously difficult to address as the markers potentially used to sort T_{reg} or T_{conv} cells may be themselves perturbed. We have thus performed a deep profiling analysis in patients with IPEX syndrome that associates flow cytometry, single-cell (for resolution) and conventional (for depth) RNA-seq. Complementary analyses in *scurfy* mice assessed the generality of the conclusions, eschewing the inevitable confounders of patient material and permitting experimental manipulations to trace the source of the perturbations. The results paint an unexpected landscape of IPEX T cells, with a complex mix of T_{reg}-like cells, some healthy like and others more perturbed, a surprisingly narrow intrinsic signature of FOXP3 but a dominant 'IPEX signature' that affects both T_{reg} and T_{conv} cells and suggests a feed-forward loop in T cell perturbations culminating in clinical disease.

Results

Study cohorts. This study included a primary and a replication cohort for confirmation and refinement (including some re-sampling of initial patients). Altogether, we analyzed peripheral blood mononuclear cells (PBMCs) from 15 patients with IPEX syndrome and 15 healthy donors (HDs) collected at two reference clinical centers (Necker Hospital, Paris and Children's Hospital, Boston; Supplementary Table 1 and Fig. 1a). HDs had no significant medical history and were recruited during well-child visits or orthopedic follow-ups. IPEX presentation was typical, with the first symptoms appearing neonatally for most (1–8 weeks). As in other cohorts[15–18], enteropathy (12 out of 14), dermatitis (10 out of 14), food allergy (8 out of 13) and diabetes (6 out of 14) were most common, with less frequent kidney, neurological or pulmonary involvement. Four patients were untreated at the time of sampling, but most were managed by immunosuppression (mainly mTOR inhibitors); four of them later received bone marrow transplants. Blood samples were collected at different ages (9 months to 26 years) during routine visits, with no concurrent acute events. Five patients were analyzed at two or three time points (6–12 months apart) to assess the stability of the transcriptional characteristics.

FOXP3 mutations were confirmed by sequencing. Most were missense mutations in different domains, two of them represented twice (V408M and the common A384T). One mutation (P1) was a large deletion spanning the promoter and intron 1; two others affected the exon 8 splice acceptor and donor sites (P4 and P5, respectively). The mutations' impact was seen in the RNA-seq reads at the *FOXP3* locus of sorted CD4+CD25^{hi}CD127^{lo} cells (see below). For HDs, the traces were very reproducible, with reads piling up at the exons, the conserved noncoding sequence (CNS) 1 and the lncRNA *FLICR* (Fig. 1b and Supplementary Fig. 1). Profiles from patients were variable; generally normal for the missense mutations; for P1, aberrantly initiated transcripts piled up around CNS1 with essentially no exonic reads; for P4, the splice mutation in exon 8 disturbed later exons and introns. Notably though, the *FOXP3* locus was active in all IPEX cells, at far greater levels than in T_{conv} cells (Supplementary Fig. 1). *FLICR* transcripts varied markedly in patient cells, unrelated to the proportion of *FOXP3* exonic reads, in line with its independent transcriptional regulation[33].

Identification of T_{reg}-like cells in patients with IPEX syndrome. T_{reg} cells are typically identified by flow cytometry as CD25^{hi}CD127^{lo}, a combination that overlaps well with intracellular FOXP3 expression[34]. Such cells have been reported in patients with IPEX syndrome[26–28], but we first adopted a more global approach to identify T_{reg}-like cells as expression of these two markers could be perturbed by FOXP3 deficiencies. We analyzed CD4+ PBMCs by multiparameter flow cytometry for a larger panel of T_{reg} markers and intracellular staining for FOXP3 and HELIOS (Supplementary Fig. 2). The data from individual staining experiments, including HD and IPEX samples, were then integrated in two-dimensional (2D) space by *t*-distributed stochastic neighbor embedding (*t*-SNE; Fig. 1c and Supplementary Fig. 3). T_{reg} cells from HD donors clustered tightly, with characteristic FOXP3, CD25 and HELIOS expression. For IPEX donors, the same region of the plots contained cells that also expressed CD25 and HELIOS, but variable FOXP3 (predictably absent for P1, more normal for P4). Indeed, 'gating' T_{reg}-like cells on the *t*-SNE showed that they fell in the expected position of the CD25/CD127 plot (Fig. 1d and Supplementary Fig. 2a), although with lower CD25 in IPEX samples. The proportions of T_{reg}-like cells and generally of FOXP3+ cells, were well conserved in patients with IPEX syndrome (Fig. 1e and Supplementary Fig. 2b), although levels of FOXP3 protein were variably reduced (Fig. 1f and Supplementary Fig. 2b). Other than the low or null levels resulting from promoter or splicing mutations, there was no discernible relationship between FOXP3 intensity and domains affected by each mutation. These cytometry results confirmed that cells with surface T_{reg} characteristics, but with low or absent FOXP3, could be identified in every patient with IPEX syndrome.

Widespread perturbations of IPEX CD4+ T cells transcriptomes. Given the presence of T_{reg}-like CD25^{hi}CD127^{lo} cells in PBMCs from patients with IPEX syndrome, we performed population low-input RNA-seq after sorting these and T_{conv} (CD25^{-}CD127^{hi}) cells, as a preliminary to single-cell profiling and to help anchor its interpretation (first and replication cohorts were profiled independently). Several observations emerged. First, the classic T_{reg} signature was well conserved in T_{reg} cells from IPEX donors (Fig. 2a), including prototype transcripts (*IL2RA*, *LRRC32*, *CTLA4*). In ranked fold change (FC) plots that display the T_{reg} signature for individual donors, the distribution of signature genes was tight for HDs but more variable in

Fig. 2 | Transcriptional changes in IPEX T_{reg} and T_{conv} cells by population RNA-seq. Population RNA-seq was performed on sorted T_{conv} and T_{reg}-like cells from HD ($n=12$) and IPEX ($n=10$) donors. **a**, The T_{reg}/T_{conv} FC in HDs (*x* axis) and IPEX (*y* axis); T_{reg} signature[5] genes are highlighted. **b**, Ranked FC plots of T_{reg} signature transcripts for individual donors, ranked according to mean FC in all HD (blue dots). FC values for each donor (black dots) computed from the donor's T_{reg} versus the mean of HD T_{conv} cells. **c**, Index and CV of T_{reg} up signature transcripts (each dot is a sample). ***$P<0.001$, two-sided Student's *t*-test . **d**, FC versus *P* value (volcano) plot comparing normalized expression in all IPEX to all HD samples. Genes with differential expression (two-sided Student's *t*-test $P<0.01$, FC > 2) are highlighted (452 up genes and 1,032 down genes) and numbers are shown. **e**, Heat map of the expression ratio for IPEX signature genes defined in **d** and in T_{reg} cells from each donor, computed against mean expression in HD T_{reg} cells (each cohort was computed against its own HD T_{reg} set). **f**, Comparison of mean IPEX effect (all IPEX versus all HD samples) in T_{reg} (*x* axis) versus T_{conv} (*y* axis) cells. **g**, Same FC/FC plot as in **f**, but highlighted with representative signatures of tumor-infiltrating T_{reg} cells, activated T_{reg} versus resting T_{reg} cells and IFN-I induced genes. **h**, Same volcano plot as in **d**, but with highlights from a representative CD4+ T activation signature (NS, hypergeometric test). **i**, Heat map, for the IPEX signature genes defined in **d**, of their overlap with the pathways and signatures that significantly overlap with IPEX signature (hypergeometric test, $P<0.001$). Present genes are shown by tick marks, color coded by their IPEX/HD FC. Up, upregulated; down (DN), downregulated.

IPEX T_{reg} cells, ranging from quasi-normal (P7) to markedly affected (P6, P8; Fig. 2b and Supplementary Fig. 4a), perhaps surprisingly, P1 with the complete LOF mutation did not show the strongest reduction in T_{reg} signature genes. The T_{reg} signature intensity score seemed better preserved in T_{reg} cells from IPEX donors than its coefficient of variation (CV) (Fig. 2b,c), suggesting an instability of the T_{reg}

signature in absence of a fully functional FOXP3. These indices did not correlate significantly with clinical outcomes, although patients under immunosuppression trended to a lower index (Supplementary Fig. 4b); this correlation is confounded, however, as patients with more severe *FOXP3* mutations and thus a lower T_{reg} up index, are clinically more affected and more likely to receive strong treatment. The IPEX mutations did not affect all signature genes equally; some T_{reg} signature genes were actually over-expressed in IPEX T_{reg} cells (*DUSP4*, *LRRC32*, *CTLA4*), whereas most were downregulated as expected (Supplementary Fig. 4c).

It is worth mentioning that no patient's T_{reg} or T_{conv} cells showed unusual expression of *IL4* or *IL5*, transcripts, which have been found to be paradoxically upregulated in response to forced expression of mutant FOXP3 (refs. [8,22]). More generally, there was no specific induction of cytokine genes in T_{conv} cells from patients with IPEX syndrome, which might have denoted a loss of T_{reg} cell control (Supplementary Fig. 4d).

We then assessed more generally the impact of IPEX mutations on transcriptomes of T_{reg}-like cells. Widespread differences were observed (Fig. 2d), reproducibly in the two cohorts (Supplementary Fig. 5a and Supplementary Table 2). This 'IPEX signature' was consistent in every patient, involving all the same genes, albeit at variable overall intensity (Fig. 2e; one HD did show partial bias, perhaps because of unrecognized pathology). It did not correlate with the T_{reg} signature intensity score (Supplementary Fig. 5b) and was strikingly similar in T_{reg} and T_{conv} cells (Fig. 2f), indicating a global impact on CD4$^+$ T cells that transcended the sole effect of FOXP3 in T_{reg} cells (only 17 genes of the T_{reg} signature belonged to this IPEX signature; Supplementary Table 2). There was no marked relationship between the main clinical parameters, including current corticosteroid or rapamycin treatment and the IPEX signature score (Supplementary Fig. 5c), which was also present in untreated patients with IPEX syndrome. The IPEX signature was not accompanied by activation of endogenous retroviruses, which might plausibly be reactivated (Supplementary Fig. 5d). Enrichment analysis showed significant overlap between the IPEX signature and several gene expression signatures of CD4$^+$ T cells (Fig. 2g–i and Supplementary Table 3), but notably not with the signatures of T cell activation (Fig. 2h), again indicating that the absence of T_{reg} suppression did not result in wholesale T cell activation. If anything, a T_{reg} activation signature[35] was downregulated in IPEX cells (Fig. 2g). Transcripts differentially expressed in tumor-infiltrating T_{reg} cells (two independent studies) were counter-regulated in IPEX CD4$^+$ T cells, as were interferon (IFN)-stimulated genes (Fig. 2g,i). This broad change was consistent with the upregulation of major response regulators such as *RELB*, *KRAS* or *HIF1*. Overall, our results indicated that the T_{reg} signature was in large part maintained in T_{reg} cells from patients with IPEX syndrome, albeit with a notably high degree of variability. More unexpected was the peculiar transcriptomic footprint shared by IPEX T_{reg} and T_{conv} cells, which might result from integration of extracellular cues, but not from a generic T cell activation state.

Different types of T_{reg} cells in patients with IPEX syndrome. From these preliminary indications, we proceeded to the crux of the work, applying scRNA-seq to identify with unbiased profiling the actual types of FOXP3-deficient T_{reg}-like cells, which might be blurred by the averaging inherent in population profiling or have escaped recognition because of shifts in CD25 or CD127 markers. We sorted total CD4$^+$ T cells to yield granular information on both T_{reg} and T_{conv} pools and performed scRNA-seq on the two cohorts as above (a total of 52,776 cells passing quality control (Supplementary Table 4) from 11 IPEX and 11 HD donors). All results were observed in both the initial and replication cohorts, but are combined below for simplicity. Experimental confounders were minimized by multiplexing IPEX and HD samples in the same scRNA-seq libraries, using DNA-coded tags[36]. Residual intercohort and interexperiment effects were corrected using the canonical correlation analysis (CCA) and *k*-nearest neighbor-based integration methodology[37] (Fig. 3a).

In the uniform manifold approximation and projection (UMAP) of the integrated data, CD4$^+$ T cells partitioned sharply according to disease; cells from HD donors clustered together, whereas those from patients with IPEX syndrome were much more dispersed (Fig. 3a), each tending to form an island distant to various degrees from the HD group (Fig. 3b). These distances were not batch artifacts (similar HD/IPEX partitions were seen in each batch; Supplementary Fig. 6a). Replicate samples from three patients drawn >1 year apart mapped to the same regions, as did samples from patient P7 collected before and after immunosuppressive treatment (Supplementary Fig. 6b), indicating that the different locations were patient-specific and not consequences of environmental or treatment variables. Conventional and scRNA-seq reflected the same perturbation, as UMAP1 scores axis that partitioned IPEX and HD samples corresponded with the IPEX signature above (Supplementary Fig. 6c).

We then sought to deconvolute, at single-cell resolution, T_{reg} cells among these CD4$^+$ T cells. T_{reg} cells were identified in HD and IPEX samples in an unsupervised approach, using community detection in the reciprocal principal-component analysis (PCA) integration network (implemented in Seurat v.3 (ref. [37])), also supported by activity at the *FOXP3* locus. Three distinct populations of CD4$^+$ T cells could be identified (Supplementary Fig. 6d): two T_{conv} populations (resting and activated) and one T_{reg} population (confirmed by expression of T_{reg} signature genes *FOXP3*, *IKZF2*, *CTLA4*, *IL2RA* and *TIGIT*; Supplementary Fig. 7a). When T_{reg} cells thus identified were displayed onto the UMAP space, HD T_{reg} cells formed a tight cluster for all HD samples (Fig. 3c), which colocalized with *FOXP3* expression (yellow dots on Fig. 3c). In IPEX samples, T_{reg}-like cells were similarly identified, in proportions equivalent to controls, even for fully FOXP3-deficient P1 (Supplementary Fig. 7b), but they split into two groups (Fig. 3c and Supplementary Fig. 8a); some (hereafter referred to as type A IPEX T_{reg} cells) mapped in the same cluster as HD T_{reg} cells, whereas others were clustered away in their respective patient-specific 'island' (type B). Both were truly T_{reg} cells, expressing *FOXP3* mRNA (Fig. 3c), core T_{reg} transcripts (*IL2RA*, *CTLA4*, *IKZF2* and low *IL7R*) and the T_{reg} signature (Fig. 3d). Type A and type B T_{reg} cells did differ, with higher levels of T_{reg} signature genes in type A T_{reg} cells and higher representation in type B cells of the IPEX signature (Fig. 3e and Supplementary Fig. 8b). Finally, proportions varied between patients with IPEX syndrome (Fig. 3f); type A T_{reg} cells dominated in some, almost to HD levels, but were far less abundant in others. This proportion was stable in independent

Fig. 3 | scRNA-seq reveals the heterogeneous effect of FOXP3 deficiency in IPEX T_{reg} cells. scRNA-seq was performed on peripheral blood CD4$^+$ T cells from IPEX ($n=11$) and HDs ($n=11$). **a**, 2D UMAP plot of all CD4$^+$ cells from IPEX and HD samples (52,776 cells altogether, merged with CCA and knn-graph; Methods). **b**, Same UMAP as **a**, color-coded by individual donor (see also Supplementary Fig. 6d). **c**, Same UMAP as **a**, with four representative HD and IPEX donors (see Supplementary Fig. 8 for other donors). Blue, green and red cells are resting T_{conv}, activated T_{conv} and T_{reg} cells, respectively. *FOXP3*-expressing cells (RNA) are yellow. Type A and B T_{reg} cells in IPEX donors are indicated by an arrow. **d**, Normalized counts expression of *IL2RA*, *CTLA4*, *IKZF2*, *IL7R* and the T_{reg} signature in T_{reg} cells from HD and IPEX (type A and B) and T_{conv} cells; each dot is a sample. $^*P<0.05$, Student's *t*-test. **e**, Average expression of the IPEX signature in type A and B IPEX T_{reg} cells and HD T_{reg} cells (normalized counts). $^{***}P<10^{-3}$, $^{****}P<10^{-4}$, two-sided Student's *t*-test. **f**, Proportion of type A T_{reg} cells in total T_{reg} cells for each sample. **g**, Proportion of type A T_{reg} cells plotted against the average UMAP1 dimension for each sample.

samplings of the same patient and related to the intensity of the IPEX signature in T_{reg} and T_{conv} cells; patients with the highest index had the lowest fraction of type A T_{reg} cells (Fig. 3g).

Thus, scRNA-seq analysis revealed a subset of T_{reg} cells that closely resemble normal T_{reg} cells and another with more extensive perturbations. The origin of these two distinct T_{reg} populations in

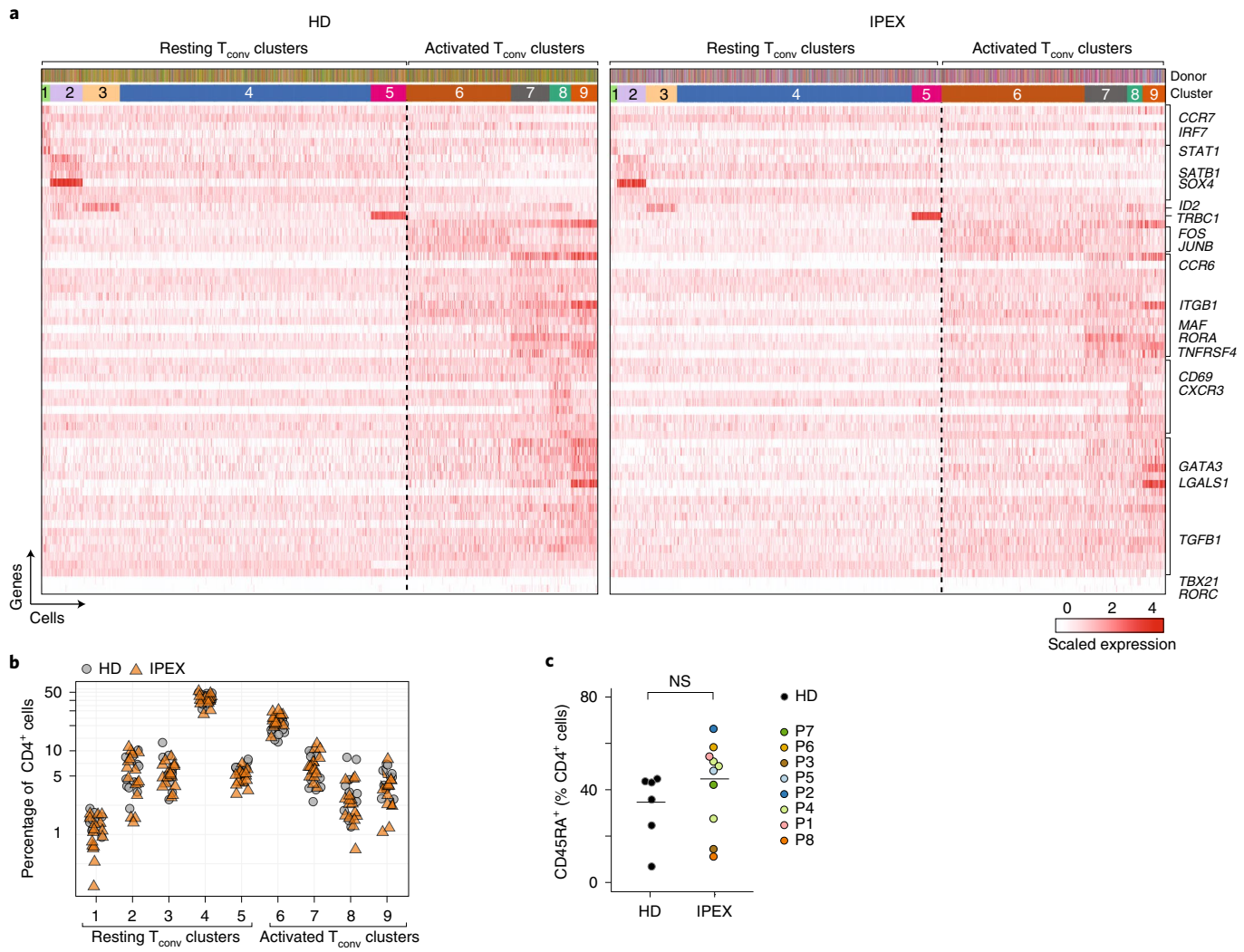


Fig. 4 | Normal T_{conv} phenotypes in patients with IPEX syndrome. a, From the scRNA-seq T_{conv} datasets (HDs and IPEX combined), the IPEX signature was regressed out (Methods) and biclustering was performed to define clusters in resting and activated T_{conv} cells, plotted as a heat map for expression of the most characteristic genes for each T_{conv} cluster (each vertical line represents one cell). Donor origin and cluster annotations for every cell (top). **b**, Proportion of resting and activated T_{conv} clusters in total CD4$^+$ cells in HD and IPEX in scRNA-seq data. **c**, Proportion of CD45RA$^+$ (resting) CD4$^+$ T cells determined by flow cytometry in IPEX and HDs. NS, two-sided Student's *t*-test ($n = 11$ HDs, 11 IPEX).

patients with IPEX syndrome was unclear. To exclude maternal microchimerism (wild-type (WT) T_{reg} cells from the mother could have a competitive advantage in the IPEX offspring), we checked female-specific *XIST* transcripts; no *XIST*-positive cells were found among any patient's T_{reg} cells. A striking feature of type A T_{reg} cells was their marked downregulation of the IPEX signature.

IPEX does not affect T_{conv} phenotypes. T_{reg} cells can affect T_{conv} polarization in many ways[9] and it is generally assumed that T_{reg} cell deficiency in patients with IPEX syndrome leads to excessive T_{conv} cell activation and differentiation and an over-representation of their effector states[38,39]. Unsupervised clustering of T_{conv} cells performed in the shared reciprocal PCA space, after regressing out the IPEX effect, distinguished nine clusters (five resting and four activated/memory types, judging from the expression of characteristic markers and transcription factors such as *CCR7*, *SATB1*, *CD69*, *TBX21* and *GATA3*; Fig. 4a). Their distribution and the expression of defining transcripts, was strikingly similar for HD and IPEX samples (Fig. 4a,b). This observation was confirmed by flow cytometry, which showed similar ranges of CD45RA$^+$ cells among CD4$^+$ T cells in IPEX and HD PBMCs (Fig. 4c). Thus, beyond the shared signature, IPEX disease seemed to have limited impact on other phenotypic aspects of circulating T_{conv} cells.

Mixed populations of *Foxp3*-active cells in *Foxp3*-deficient mice. *FOXP3* deficiency in patients with IPEX syndrome thus led to a global change affecting all CD4$^+$ T cells and to a diversity of T_{reg}-like cells. To assess mutational impacts in a setting devoid of genetic or therapeutic variables, we performed parallel analyses on *Foxp3*-deficient mice (the previously described *Foxp3*$^{\Delta EGFPcre}$ inactivating reporter line[25], hereafter denoted Δ*Foxp3*), in which FOXP3 protein is absent, but *Foxp3* locus activity can be detected.

We used scRNA-seq to analyze the diversity of CD4$^+$ T splenocytes in Δ*Foxp3* mice and WT controls (four mice per group, 18,000 cells altogether, again multiplexing samples in the same libraries; Supplementary Table 4). As in human patients, cells from WT and Δ*Foxp3* mice clustered separately on the UMAP, here again reflecting a generic ΔFoxp3 signature (Fig. 5a) present in every mutant mouse and affecting T_{reg} as well as T_{conv} cells (Supplementary Fig. 9a). This cross-cutting '*ΔFoxp3* signature' was confirmed by

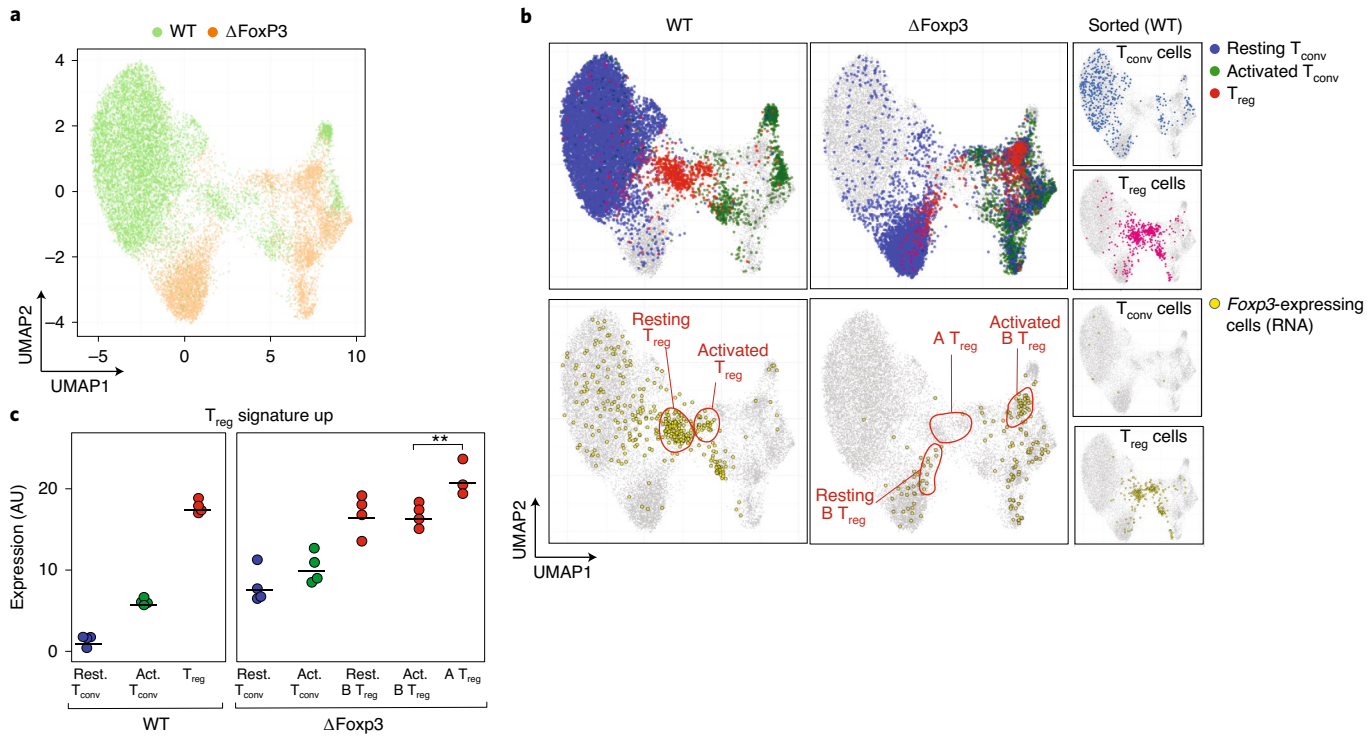


Fig. 5 | scRNA-seq reveals heterogeneous effects of *Foxp3* ablation in Δ*Foxp3* mice. scRNA-seq was performed on CD4$^+$ T cells from WT ($n=4$) and Δ*Foxp3* ($n=4$) mice (18,097 cells altogether). **a**, 2D UMAP plot of CD4$^+$ single-cell transcriptomes for WT (green) and ΔFoxP3 (orange). **b**, Same UMAP as **a**. Resting T_{conv} (blue), activated T_{conv} (green) and T_{reg} cells (red) (assigned by graph-based clustering on the merged dataset after regressing out the Δ*Foxp3* signature; Methods) are highlighted (top); cells with an active *Foxp3* locus (*GFP* or *Cre* transcripts detected) are colored yellow (bottom). Small inserts (right) show sorted control WT T_{conv} and WT T_{reg} cells, sorted and included as spike-in controls. **c**, T_{reg} signature expression in resting T_{conv}, activated T_{conv} and T_{reg} cells in WT and Δ*Foxp3* mice (normalized counts). $^{**}P<10^{-2}$, two-sided Student's *t*-test. GFP, green fluorescent protein.

population RNA-seq of Δ*Foxp3* T_{reg} cells. The same perturbation was found in Δ*Foxp3* T_{reg} cells from spleen and lung, showing that it extended to parenchymal locations, specifically in a site relevant to scurfy disease (Supplementary Fig. 9b).

Unsupervised clustering in the shared CCA space identified, for WT CD4$^+$ cells, resting and activated T_{conv} cells and a tight group of T_{reg} cells (blue, green and red, respectively; Fig. 5b). These assignments were confirmed by the location of sorted T_{conv} or T_{reg} cells (Fig. 5b), by differential expression of canonical genes (for example *Ccr7*, *Sell*, *Cd44*, *Tbx21*; Supplementary Fig. 9c and Supplementary Table 5) and at the *Foxp3* locus (as *Cre* and *Foxp3* transcripts; Fig. 5b). Note how T_{conv} clusters were similarly structured in Δ*Foxp3* and WT mice, as they had been in human patients (Supplementary Fig. 9c).

As for patients with IPEX syndrome, T_{reg}-like cells of Δ*Foxp3* mice were multiform (Fig. 5b); a minor fraction of type A T_{reg} cells closely resembled normal T_{reg} cells, in the same small proportions as in the most complete IPEX deficiencies; a larger proportion of type B T_{reg} cells that mapped into resting and active areas of the UMAP (Fig. 5b), also expressing *Foxp3* (Fig. 5b). Both expressed the T_{reg} signature, highest for type A T_{reg} cells (Fig. 5c). Together, these T_{reg}-like populations accounted for 7% of total CD4$^+$ T cells, as in WT T_{reg} cells (Supplementary Fig. 9d).

Thus, T_{reg}-like cells in *Foxp3*-deficient mice showed the same heterogeneity as in human patients with IPEX syndrome; a mix of healthy and altered T_{reg} cells, but with a dominant transcriptional signature that cut across both T_{reg} and T_{conv} CD4$^+$ T cells.

Cell-intrinsic and extrinsic impact of *Foxp3* deficiency. In humans and mice, FOXP3 deficiency resulted in a heterogenous mix of T_{reg}-like cells and a strong disease-specific signature unexpectedly shared by T_{reg} and T_{conv} cells. The latter suggested cell-extrinsic influences on the transcriptomes of CD4$^+$ T cells, which we investigated in two ways.

First, we analyzed FOXP3-deficient T_{reg} cells present in the same environment as WT cells, in blood CD4$^+$ T cells from two mothers of patients with IPEX syndrome. X inactivation in females occurs with equal probability, such that half of the T cells in carrier mothers inactivate the mutant X chromosome and half the WT X chromosome. In the presence of T_{reg} cells expressing WT FOXP3, cell-extrinsic signals should be repressed, unmasking the true cell-intrinsic effect of *FOXP3* mutations. We performed scRNA-seq on blood CD4$^+$ T cells from mothers of patients P5 and B2, analyzing X chromosome reads that overlapped known single-nucleotide polymorphism variants to assign the active X chromosome of each cell to either the patient's mutant haplotype or the nontransmitted WT (Fig. 6a and Methods). T_{reg} cells in which the mutant chromosome was active were present in roughly equal proportions to their WT counterparts, as was the case for T_{conv} cells. In the UMAP, cells with the mutant and WT ChrX active mapped to the same areas for both T_{reg} and T_{conv} cells (Fig. 6b) and all T_{reg} cells were type A, irrespective of the active ChrX. The IPEX signature induced in mutant T_{conv} cells was fully extinguished (Fig. 6c), with no significant gene expression differences between T_{conv} with WT or mutant ChrX, highlighting the absence of a cell-intrinsic effect of FOXP3 in unchallenged T_{conv} pools. The T_{reg} up signature genes that were dysregulated in P5 or B2 T_{reg} cells generally normalized in mother's T_{reg} cells expressing the same mutant FoxP3 (Fig. 6d). These results indicate that the signals generating the IPEX signature are cell extrinsic.

Second, we constructed bone marrow chimeras (BMCs) by transfer of equal proportions of congenically marked stem cells

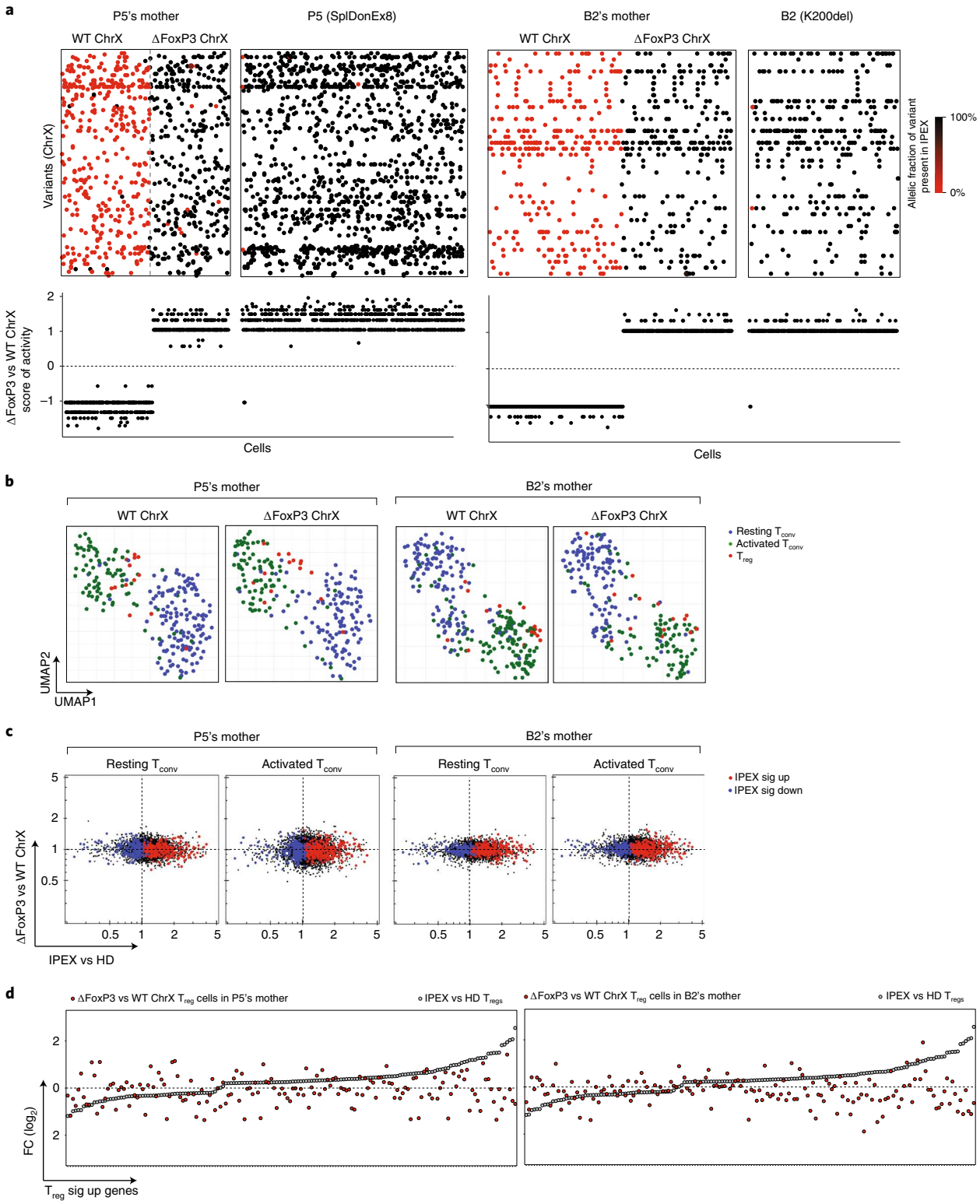


Fig. 6 | Cell-intrinsic and extrinsic effects of *FOXP3* deficiency in T_{reg} and T_{conv} cells of IPEX female carriers (mothers). scRNA-seq was performed on sorted and hashtagged blood CD4$^+$ T cells from two IPEX mothers: P5's and B2's mothers (2,880 cells altogether). **a**, Because of X chromosome (ChrX) inactivation in females, a mix of FOXP3-deficient (Δ*FOXP3*) and -proficient (WT) T cells are present in IPEX mothers. Biclustering heat map, showing for each single cell (columns) the allelic fraction of X chromosome variants carrying the FOXP3 mutation (present in P5 or B2) (top). FOXP3-deficient versus -proficient X chromosome score of activity for each single cell (bottom). **b**, 2D UMAP plot of all CD4$^+$ cells from P5's and B2's mothers split by active X chromosome (FOXP3-proficient (left); FOXP3-deficient (right)). Resting T_{conv}, activated T_{conv} and T_{reg} cells are highlighted in blue, green and red, respectively. **c**, scRNA-seq data for T_{conv} cells were collapsed and expression ratios between FOXP3-deficient and -proficient T_{conv} cells (resting or activated) were calculated for IPEX (*x* axis) and in P5's and B2's mothers (*y* axis). Upregulated and downregulated IPEX signature genes are indicated in red and blue, respectively. **d**, FC plots of FOXP3-deficient versus -proficient T_{reg} cells in P5's and B2's mothers, T_{reg} up signature genes, ranked according to FC in IPEX versus HDs (gray dots).

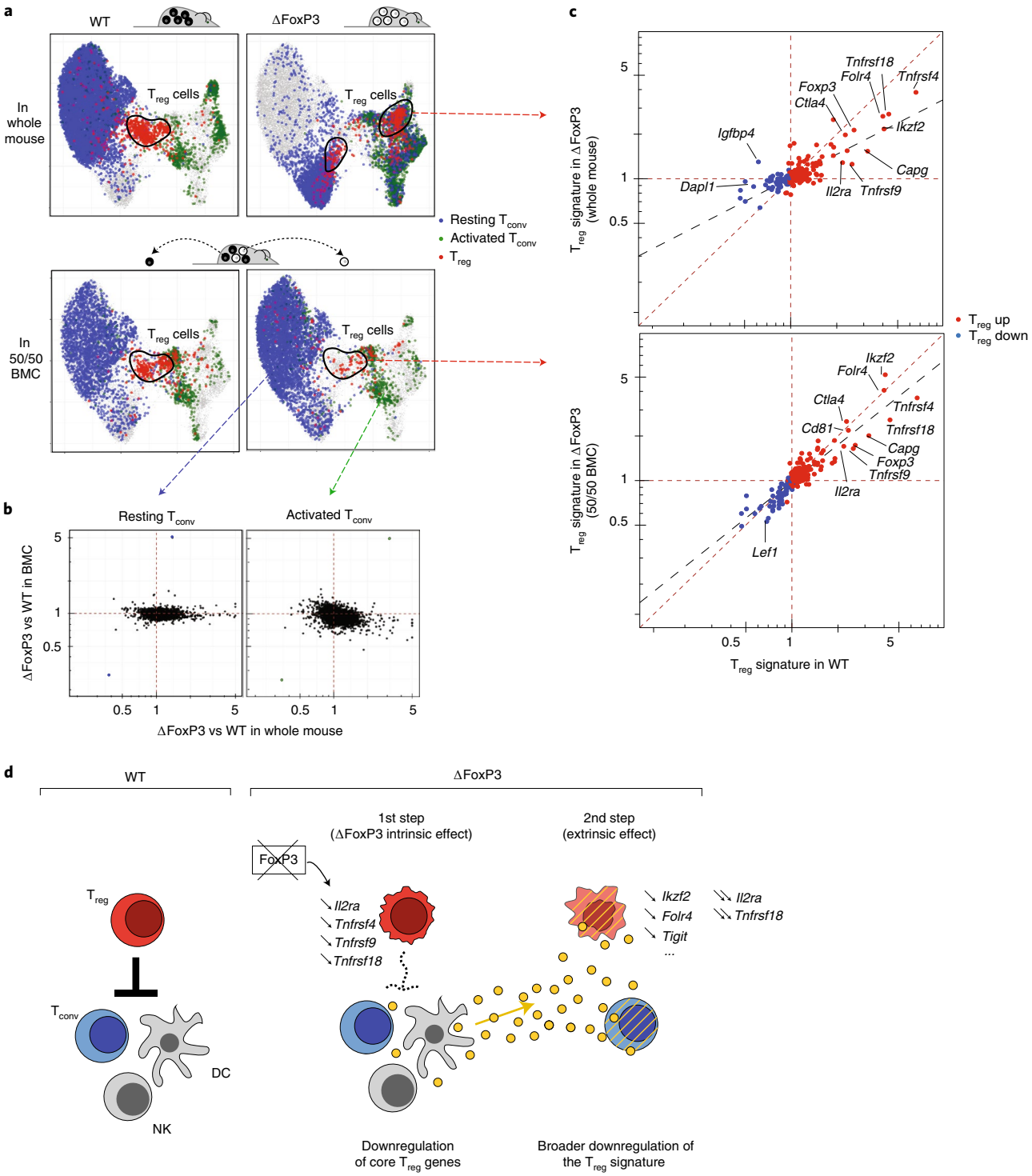


Fig. 7 | Cell-intrinsic and extrinsic effects of *Foxp3* deficiency in T_{reg} and T_{conv} cells in mice. scRNA-seq was performed on sorted and hashtagged CD4$^+$ T cells from 50/50 mixed BMC mice (WT and Δ*Foxp3* hematopoietic stem cells, congenically marked) (8,556 cells altogether, $n=3$ mice). **a**, Same UMAP as shown in Fig. 5b, integrating WT and Δ*Foxp3* CD4$^+$ cells from whole mice (top) or from chimeras (bottom). Resting T_{conv} (blue), activated T_{conv} (green) and T_{reg} cells (red) (identified by graph-based clustering on the merged dataset after regressing out the Δ*Foxp3* signature; Methods). **b**, scRNA-seq data for T_{conv} cells were collapsed and expression ratios between WT and Δ*Foxp3* T_{conv} were calculated for whole mice (*x* axis) or for 50/50 BMCs (*y* axis); for the latter, WT and Δ*Foxp3* cells originated from the same host. **c**, T_{reg}/T_{conv} ratio in Δ*Foxp3* cells (*y* axis) in whole mice (top) and in 50/50 BMCs (bottom), both plotted against the same T_{reg}/T_{conv} expression ratio in WT cells (*x* axis). T_{reg} signature genes are highlighted. Dashed lines (linear regression) represent the slope of the fit. **d**, Two-step model of the impact of FOXP3 deficiency in mice and humans. First, the absence of FOXP3 downregulates a few core T_{reg} signature genes in T_{reg} cells (cell-intrinsic effect), leading to perturbed control of other immune cells (which might include known T_{reg} targets such as T_{conv} cells, dendritic cells (DCs), natural killer (NK) cells, or others), unleashing a secondary intercellular response (for example, cytokines and cell surface modulators), which imprints the dominant IPEX signature on T_{reg} and T_{conv} cells and further affects T_{reg} functionality, amplifying disease pathology.

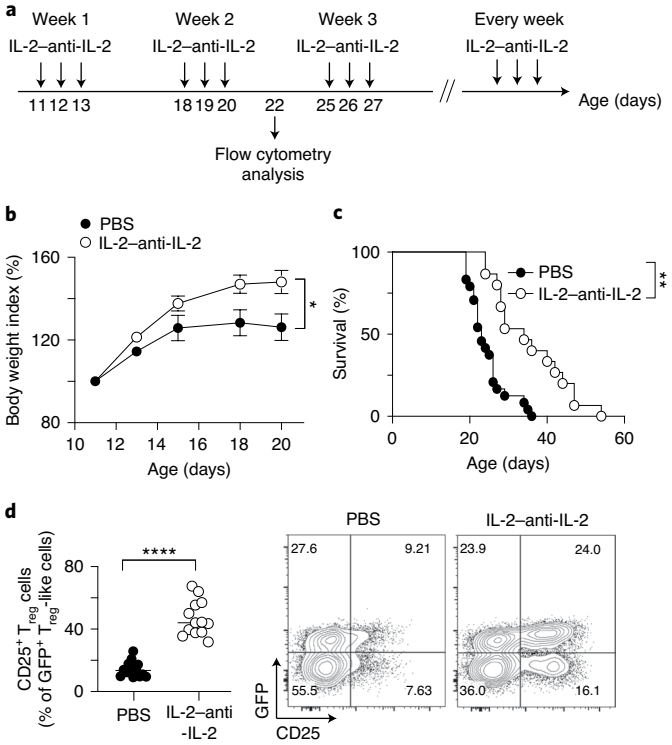


Fig. 8 | IL-2 therapy mitigates *Foxp3*-deficiency disease. a, IL-2 treatment protocol: daily injection of IL-2–anti-IL-2 complex for 3 d consecutively every week, starting 11 d after birth. **b**, Body weight index growth curves of vehicle (PBS) and IL-2 (IL-2–anti-IL-2)-treated Δ*Foxp3* mice (mean ± s.e.m.). $^{*}P < 0.05$, one-way analysis of variance, $n = 10$ and 19 (PBS and IL-2 groups, respectively). **c**, Survival curves of vehicle (PBS) and IL-2 (IL-2–anti-IL-2)-treated Δ*Foxp3* mice. $^{**}P < 10^{-4}$, log-rank test, $n = 24$ and 15 (PBS and IL-2 groups, respectively). **d**, Proportion of CD25^{+}GFP^{+} T_{reg} cells (of GFP^{+} T_{reg}-like cells) in vehicle (PBS) and IL-2 (IL-2–anti-IL-2)-treated Δ*Foxp3* mice (day 22). Representative flow cytometric CD25/GFP plots of CD4^{+} T cells (right). $^{****}P < 10^{-4}$, two-sided Student's *t*-test, $n = 14$ and 13 (PBS and IL-2 groups, respectively).

from WT and Δ*Foxp3* mice, a protocol that prevents the scurfy-like disease that appears after reconstitution with *Foxp3*-deficient stem cells[40]. As in hemizygous human mothers, cell-extrinsic transcriptional hallmarks should be reverted in genetically *Foxp3*-deficient cells by the presence of WT cells in the same mouse. Ten weeks after reconstitution, 5,556 CD4^{+} T cells from either Δ*Foxp3* or WT compartments of BMC mice were profiled by multiplexed scRNA-seq (three mice per group, with data aligned to the same UMAP space as in Fig. 5).

In this mixed setting, many of the characteristics of Δ*Foxp3* CD4^{+} T cells essentially disappeared (Fig. 7a), whereas WT cells were unchanged. First, on the UMAP plot, resting and activated T_{conv} cells from Δ*Foxp3* compartments shifted and overlapped with WT cells from the same mice. Accordingly, the Δ*Foxp3* signature was flattened out for mutant cells in the BMC setting (Fig. 7b and Supplementary Fig. 10a). These results demonstrated the cell-extrinsic origin of the Δ*Foxp3* signature in *Foxp3*-deficient T_{conv} cells. Δ*Foxp3* cells were partially outcompeted by WT T_{reg} cells (1:3) in the BMC mice (Supplementary Fig. 10b), but their phenotypes were also corrected; they were all type A T_{reg} cells that overlapped with WT T_{reg} cells on the UMAP (Fig. 7a) and their expression of the T_{reg} signature was almost completely normalized (Fig. 7c and Supplementary Table 5). Only 13 genes remained as under-expressed in the Δ*Foxp3* T_{reg} compartment, including classic FOXP3 target loci (*Il2ra*, *Tnfrsf4*, *Tnfrsf9*, *Tnfrsf18* and *Capg*); no genes were upregulated, confirming that FOXP3 is mainly a transcriptional activator. The other T_{reg} signature genes were otherwise normally expressed, contrasting with their general downregulation in Δ*Foxp3* mice.

Complementing FoxP3 deficiency with IL-2. Thus, the dominant suppressive effect of WT T_{reg} cells curtailed the Δ*Foxp3* signature in T_{reg} and T_{conv} cells, revealing a narrower intrinsic effect of *Foxp3* on a minor proportion of the T_{reg} signature, highlighting both intrinsic and extrinsic consequences of the *Foxp3* deficiency in the disease context. We thus propose a 'two-step' model of the IPEX/scurfy disease (Fig. 7d). First, the intrinsic effect of *Foxp3* in T_{reg} cells dysregulates a few core T_{reg} genes, which unleashes a systemic response. Secondarily, this milieu imprints broad changes on both T_{reg} and T_{conv} cells (IPEX/Δ*Foxp3* signatures) and further destabilizes T_{reg} signature expression and dampens T_{reg} effector function.

To provide a mechanistic proof of concept, we reasoned that enhancing an essential core T_{reg} gene in Δ*Foxp3* mice might restore T_{reg} functionality. *Il2ra* is one of the genes intrinsically regulated by FOXP3 above (see elsewhere[8,40,41]) and is preferentially expressed in type A T_{reg} cells (Fig. 7c). Interleukin (IL)-2, its ligand, is the key trophic cytokine for T_{reg} cells and induces Il2ra expression in a positive feed-forward loop[42]. IL-2 administration might thus complement the deficit due to the missing FoxP3 and, in part, revert the T_{reg} deficiency. Therefore, we evaluated the potential of serial IL-2 injections (stabilized in a complex with an anti-IL-2 antibody[43]) into Δ*Foxp3* pups (three consecutive daily injections per week starting at 11 d of age; Fig. 8a). Compared to the phosphate-buffered saline (PBS) control, IL-2–anti-IL-2 complexes partially prevented weight loss (Fig. 8b) and prolonged survival (Fig. 8c) of Δ*Foxp3* mice, at least for a period of time. The treatment also expanded the CD25^{+} fraction within T_{reg}-like cells, thus presumably growing the type A fraction (Fig. 8d). These results show that reverting one of the core deficiencies resulting from the Δ*Foxp3* mutation does improve T_{reg} function, at least for some time and additionally suggest that IL-2 therapy might be used to mitigate Δ*Foxp3* disease.

Discussion

These results provide unprecedented cellular and genomic resolution of a primary immunodeficiency and reveal unexpectedly multifarious molecular and cellular consequences of FOXP3 deficiency. There were surprisingly limited cell-intrinsic perturbations in T_{reg} cells, associated with a dominant and monomorphic signature of cell-extrinsic origin, which cuts across the transcriptome of all CD4^{+} T cells and amplifies T_{reg} perturbations.

We thus propose a two-step model of the disease's molecular pathogenesis. FOXP3 is actually important for only very few T_{reg} genes; when the milieu was normalized by the presence of WT T_{reg} cells the cell-intrinsic impact of the Δ*Foxp3* deficiency was exerted on only a handful of genes (*Il2ra*, *Tnfrsf4*, *Tnfrsf9*, *Tnfrsf18*, *Capg*, *Ikzf2* and *Ctla4*). This surprisingly short list corresponds to a 'core set' of genes that are expressed by all T_{reg} cells[6] and are directly transactivated by and bind FOXP3 (refs. [8,44]). These genes encode the major homeostatic regulator of T_{reg} cells (*IL2RA*) and several members of the tumor necrosis factor (TNF) receptor superfamily, which are also connected to T_{reg} homeostasis and function[45]. We propose that their downregulation is the first step (accordingly, *IL2RA* deficiency also causes an IPEX-like disease). This initiates a systemic reaction which constitutes the second step of the FOXP3 deficiency syndrome; a broadly shared signature that marks both T_{reg} and T_{conv} cells, further perturbs T_{reg} signature transcripts and amplifies in a vicious circle the defect in T_{reg}-suppressive activity.

What, then, is this IPEX signature, the hundreds of genes that were impacted equivalently in T_{reg} and T_{conv} cells? It was shared among all patients with IPEX syndrome, with different intensities

rather than qualitative differences, stressing its common etiology. It proved stable over several years in each multiply tested patient, before and after (and unrelated to) treatment. The genes involved were not simply T cell-activation genes, as might have been expected from a loss of suppression (activation-induced transcripts were actually repressed). We hypothesize that it is due to an inductive element(s), delivered via cytokine or cell–cell contact, normally repressed by T_{reg} cells but unleashed by FOXP3 deficiency. This signal might originate from other T cells, or from other immunocytes, much as dendritic and natural killer cells are the first responders to acute T_{reg} ablation in diphthera toxin receptor models[46,47]. The presence of WT T_{reg} cells in mixed BMC mice or in heterozygous females would restore the negative feedback, thus reinstating better T_{reg} function, evoking the 'infectious tolerance' concept[48], with the presence of normal suppressors improving the tone and function of defective ones. Pathway and ontology analysis revealed no clear match, except for effects on a few IFN- and TNF-induced genes. There was, on the other hand, an intriguing anti-correlation for genes over-expressed in tumor-infiltrating T_{reg} cells[49], where $CD4^+$ T cells from patients with IPEX syndrome turned off much of the tumor T_{reg} signature.

This principal effect cutting across both T_{reg} and T_{conv} cells evokes the debated hypothesis that FOXP3 deficiency has an intrinsic effect in T_{conv} cells, associated with the transient induction of FOXP3 in activated T_{conv} cells[29–32]. The mixed chimeras and the mother–son pairs demonstrated that FOXP3 deficiency affected T_{conv} cells extrinsically. However, it remains an open question whether this disease signature contributes to pathology via further dampening of T_{reg} function or through T_{conv} defects.

That FOXP3 is not absolutely required for T_{reg} differentiation and homeostasis is now well established, with the existence of T_{reg}-like cells described repeatedly in mice and humans lacking FOXP3 (refs. [23–28]). These 'T_{reg} wannabes' were reported to maintain some T_{reg} features (self-reactivity, partial T_{reg} signature, activity at the *FOXP3* locus) while acquiring some T_{conv} characteristics (no in vitro anergy, cytokine expression). The unexpected insight emerging from our single-cell analysis of deficient humans and mice was the wide array of cells with T_{reg}-like characteristics and/or an active *FOXP3* locus. Rather than a single population of T_{reg} wannabes, several distinct populations were present: first, a large component of type B T_{reg} cells, with many T_{reg} transcriptional characteristics, but perturbed by the IPEX signature; and second (and most mysterious), the subset of type A T_{reg} cells that seemed almost unaffected relative to WT T_{reg} cells (full T_{reg} signature, active *FOXP3* and no IPEX signature), as if these cells had somehow become nonresponsive to the systemic influence, generating escape variants, for example by dampening receptors or signaling or because they reside in protected niches. But then, why was this the case only for T_{reg} cells?

These results show that it can be misleading to infer the transcriptional footprint of a TF and its mechanistic causality in disease, solely from the transcriptome of deficient cells, as it can be perturbed as here by dominant cell-extrinsic effects. When the extrinsic effects were blocked in the chimeras, the true core signature of *Foxp3* proved much smaller than the first analysis of the deficiency had suggested. IPEX syndrome is a rare disease and our cohorts were not powered to robustly detect clinical correlates of the patients' cellular and transcriptional characteristics. On the other hand, there was no obvious correlation between clinical severity indicators and the integrated gene expression metrics. This finding is congruent with the notion that molecular severity of the FOXP3 mutations in patients with IPEX syndrome only loosely correlates with clinical severity[19,50]; here, the null mutation in patient P1 did not determine the most severe disease.

There are some implications of our findings for therapeutic strategies in IPEX syndrome. Current management, when bone marrow transplantation is not an option, involves immunologic dampening via immunosuppressants. Our results might suggest harnessing those type A T_{reg} cells that are present in the patients by sustaining their homeostasis or complementing the small set of primary FOXP3 targets identified here. Indeed, the effect of the IL-2–anti-IL-2 treatment validates this proposition (it would be interesting to see whether combination of IL-2 with other activators of the core gene set, for example TNF family members, would further improve efficacy). Alternatively, damage might be avoided in patients by blocking the signaling mechanism that imparts the dominant IPEX signature to all T cells.

In conclusion, the new landscape of T_{reg} and T_{conv} cells revealed in patients with IPEX syndrome by single-cell analysis and their correction in mixed cell contexts, have opened a new perspective on the disease and on the role of FOXP3 and T_{reg} cells in immune homeostasis.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-021-00910-8.

Received: 22 July 2020; Accepted: 26 February 2021;
Published online: 8 April 2021

References

1. Josefowicz, S. Z., Lu, L. F. & Rudensky, A. Y. Regulatory T cells: mechanisms of differentiation and function. *Annu. Rev. Immunol.* **30**, 531–564 (2012).
2. Wing, J. B., Tanaka, A. & Sakaguchi, S. Human $FOXP3^+$ regulatory T cell heterogeneity and function in autoimmunity and cancer. *Immunity* **50**, 302–316 (2019).
3. Panduro, M., Benoist, C. & Mathis, D. Tissue Tregs. *Annu. Rev. Immunol.* **34**, 609–633 (2016).
4. Hill, J. A. et al. Foxp3 transcription-factor-dependent and -independent regulation of the regulatory T cell transcriptional signature. *Immunity* **27**, 786–800 (2007).
5. Ferraro, A. et al. Interindividual variation in human T regulatory cells. *Proc. Natl Acad. Sci. USA* **111**, E1111–E1120 (2014).
6. Zemmour, D. et al. Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nat. Immunol.* **19**, 291–301 (2018).
7. Ono, M. Control of regulatory T-cell differentiation and function by T-cell receptor signalling and Foxp3 transcription factor complexes. *Immunology* **160**, 24–37 (2020).
8. Kwon, H. K., Chen, H. M., Mathis, D. & Benoist, C. Different molecular complexes that mediate transcriptional induction and repression by FoxP3. *Nat. Immunol.* **18**, 1238–1248 (2017).
9. Campbell, D. J. & Koch, M. A. Phenotypical and functional specialization of $FOXP3^+$ regulatory T cells. *Nat. Rev. Immunol.* **11**, 119–130 (2011).
10. Li, C. et al. TCR transgenic mice reveal stepwise, multi-site acquisition of the distinctive fat-Treg phenotype. *Cell* **174**, 285–299 (2018).
11. Dispirito, J. R. et al. Molecular diversification of regulatory T cells in nonlymphoid tissues. *Sci. Immunol.* **3**, eaat5861 (2018).
12. Miragaia, R. J. et al. Single-cell transcriptomics of regulatory T cells reveals trajectories of tissue adaptation. *Immunity* **50**, 493–504 (2019).
13. Powell, B. R., Buist, N. R. & Stenzel, P. An X-linked syndrome of diarrhea, polyendocrinopathy, and fatal infection in infancy. *J. Pediatr.* **100**, 731–737 (1982).
14. Ramsdell, F. & Ziegler, S. F. FOXP3 and scurfy: how it all began. *Nat. Rev. Immunol.* **14**, 343–349 (2014).
15. Barzaghi, F., Passerini, L. & Bacchetta, R. Immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome: a paradigm of immunodeficiency with autoimmunity. *Front. Immunol.* **3**, 211 (2012).
16. d'Hennezel, E., Bin, D. K., Torgerson, T. & Piccirillo, C. A. The immunogenetics of immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome. *J. Med. Genet.* **49**, 291–302 (2012).
17. Duclaux-Loras, R. et al. Clinical heterogeneity of immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome: a French multicenter retrospective study. *Clin. Transl. Gastroenterol.* **9**, 201 (2018).
18. Gambineri, E. et al. Clinical, immunological, and molecular heterogeneity of 173 patients with the phenotype of immune dysregulation,

polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome. *Front. Immunol.* **9**, 2411 (2018).
19. Barzaghi, F. et al. Long-term follow-up of IPEX syndrome patients after different therapeutic strategies: an international multicenter retrospective study. *J. Allergy Clin. Immunol.* **141**, 1036–1049 (2018).
20. Godfrey, V. L., Wilkinson, J. E. & Russell, L. B. X-linked lymphoreticular disease in the scurfy (*sf*) mutant mouse. *Am. J. Pathol.* **138**, 1379–1387 (1991).
21. Wan, Y. Y. & Flavell, R. A. Regulatory T-cell functions are subverted and converted owing to attenuated Foxp3 expression. *Nature* **445**, 766–770 (2007).
22. Van Gool, F. et al. A mutation in the transcription factor *Foxp3* drives T helper 2 effector function in regulatory T cells. *Immunity* **50**, 362–377 (2019).
23. Lin, W. et al. Regulatory T cell development in the absence of functional Foxp3. *Nat. Immunol.* **8**, 359–368 (2007).
24. Gavin, M. A. et al. Foxp3-dependent programme of regulatory T-cell differentiation. *Nature* **445**, 771–775 (2007).
25. Charbonnier, L. M. et al. Functional reprogramming of regulatory T cells in the absence of Foxp3. *Nat. Immunol.* **20**, 1208–1219 (2019).
26. Bacchetta, R. et al. Defective regulatory and effector T cell functions in patients with FOXP3 mutations. *J. Clin. Invest.* **116**, 1713–1722 (2006).
27. Otsubo, K. et al. Identification of FOXP3-negative regulatory T-like ($CD4^+CD25^+CD127^{low}$) cells in patients with immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome. *Clin. Immunol.* **141**, 111–120 (2011).
28. Boldt, A. et al. Differences in FOXP3 and CD127 expression in Treg-like cells in patients with IPEX syndrome. *Clin. Immunol.* **153**, 109–111 (2014).
29. Walker, M. R. et al. Induction of FoxP3 and acquisition of T regulatory activity by stimulated human $CD4^+$. *J. Clin. Invest.* **112**, 1437–1443 (2003).
30. Gavin, M. A. et al. Single-cell analysis of normal and FOXP3-mutant human T cells: FOXP3 expression without regulatory T cell development. *Proc. Natl Acad. Sci. USA* **103**, 6659–6664 (2006).
31. Allan, S. E. et al. Activation-induced FOXP3 in human T effector cells does not suppress proliferation or cytokine production. *Int. Immunol.* https://doi.org/10.1093/intimm/dxm014 (2007).
32. McMurchy, A. N. et al. A novel function for FOXP3 in humans: intrinsic regulation of conventional T cells. *Blood* **121**, 1265–1275 (2013).
33. Zemmour, D. et al. *Flicr*, a long noncoding RNA, modulates Foxp3 expression and autoimmunity. *Proc. Natl Acad. Sci. USA* **114**, E3472–E3480 (2017).
34. Seddiki, N. et al. Expression of interleukin (IL)-2 and IL-7 receptors discriminates between human regulatory and activated T cells. *J. Exp. Med.* **203**, 1693–1700 (2006).
35. Pesenacker, A. M. et al. A regulatory T-cell gene signature is a specific and sensitive biomarker to identify children with new-onset type 1 diabetes. *Diabetes* **65**, 1031–1039 (2016).
36. Stoeckius, M. et al. Cell hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018).
37. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
38. Bakke, A. C., Purtzer, M. Z. & Wildin, R. S. Prospective immunological profiling in a case of immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX). *Clin. Exp. Immunol.* **137**, 373–378 (2004).
39. Ziegler, S. F. FOXP3: of mice and men. *Annu. Rev. Immunol.* **24**, 209–226 (2006).
40. Fontenot, J. D., Gavin, M. A. & Rudensky, A. Y. Foxp3 programs the development and function of $CD4^+CD25^+$ regulatory T cells. *Nat. Immunol.* **4**, 330–336 (2003).
41. Hori, S., Nomura, T. & Sakaguchi, S. Control of regulatory T cell development by the transcription factor Foxp3. *Science* **299**, 1057–1061 (2003).
42. Malek, T. R. & Ashwell, J. D. Interleukin 2 upregulates expression of its receptor on a T cell clone. *J. Exp. Med.* **161**, 1575–1580 (1985).
43. Boyman, O. et al. Selective stimulation of T cell subsets with antibody-cytokine immune complexes. *Science* **311**, 1924–1927 (2006).
44. Samstein, R. M. et al. Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell* **151**, 153–166 (2012).
45. Remedios, K. A. et al. The TNFRSF members CD27 and OX40 coordinately limit T_H17 differentiation in regulatory T cells. *Sci. Immunol.* **3**, eaau2042 (2018).
46. Kim, J. M., Rasmussen, J. P. & Rudensky, A. Y. Regulatory T cells prevent catastrophic autoimmunity throughout the lifespan of mice. *Nat. Immunol.* **8**, 191–197 (2007).
47. Sitrin, J. et al. Regulatory T cells control NK cells in an insulitic lesion by depriving them of IL-2. *J. Exp. Med.* **210**, 1153–1165 (2013).
48. Cobbold, S. & Waldmann, H. Infectious tolerance. *Curr. Opin. Immunol.* **10**, 518–524 (1998).
49. Plitas, G. et al. Regulatory T cells exhibit distinct features in human breast cancer. *Immunity* **45**, 1122–1134 (2016).
50. Gambineri, E. et al. Clinical and molecular profile of a new series of patients with immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome: inconsistent correlation between forkhead box protein 3 expression and disease severity. *J. Allergy Clin. Immunol.* **122**, 1105–1112 (2008).

Methods

Mice. *Foxp3*$^{IRES\text{-}GFP}$/B6 (ref. [51]), *Foxp3*$^{\Delta EGFPiCre}$*xR26*^{YFP} (Δ*Foxp3*) and *Foxp3*$^{\Delta IRES\text{-}GFP}$/B6 (Δ*Foxp3*) mice on the C57Bl/6J background were maintained in our colony. Except when specified, 3-week-old male mice were used in this study (Supplementary Table 4). Mice were housed under specific-pathogen-free conditions and all experimentation was performed following the animal protocol guidelines of Harvard Medical School and Boston Children's Hospital (protocol 02954).

Human cohorts. Male IPEX and healthy donor whole-blood samples were obtained under protocols reviewed and approved by the local Institutional Review Boards at each center (Boston Children's Hospital 04-09-113R, Necker/Imagine C 15-13_CODECOH_AR, HMS IRB15-0504). Anonymized clinical data included age at onset and at blood sampling, clinical symptoms (enteropathy, diabetes, eczema, other autoimmune diseases and allergy), ancillary testing, treatments and long-term outcomes (Supplementary Table 1). Two independent cohorts of samples were profiled (Supplementary Table 1). Each cohort contains samples from both centers. Cohort 1 samples were processed in two different scRNA-seq runs and cohort 2 samples (replication cohort) were processed in three scRNA-seq runs. Samples from each cohort were sorted and sequenced as one batch for population RNA-seq (Supplementary Table 4).

Peripheral blood mononuclear cell isolation. Whole blood was collected in K_2 EDTA tubes and processed within a few hours. An equal volume of room-temperature PBS/2 mM EDTA was mixed into 15 ml of blood and carefully layered over 14 ml Ficoll Hypaque solution (GE Healthcare). After centrifugation for 30 min at 900*g* (with no break), at room temperature, the mononuclear cell layer was washed three times with excess HBSS (Gibco) for 10 min at 400*g* and resuspended in 2 ml of HBSS. The pellet was resuspended in 90% FBS-10% dimethylsulfoxide, 20 million cells ml^{-1}, 1 ml per vial, cooled progressively in isopropyl alcohol (Mr Frosty Freezing Container, Thermo Fisher Scientific) for 24 h and stored long term in liquid nitrogen.

Vials were thawed in 10 ml 10% FBS RPMI and cells were washed (500*g* for 5 min) and resuspended in FACS buffer (phenol red-free RPMI, 2% FBS, 0.1% azide and 10 mM HEPES, pH 7.9). After cell counting, samples were allocated for flow cytometry, population RNA-seq and scRNA-seq.

Flow cytometric profiling of IPEX and HDs. Cells were stained for flow cytometry in 100 µl of FACS buffer (phenol red-free DMEM, 2% FBS, 0.1% azide and 10 mM HEPES, pH 7.9) for 10 min with 10 µl FcBlock (supernatant of clone 2.4G2, ATCC HB-197, hybridoma cultures) and the following antibodies: CD3 BV605 or A700 (2 µl, OKT3, BioLegend), CD4 PerCP-Cy5.5 (2 µl, OKT4, BioLegend), CD25 PE-Cy7 (3 µl, BC96, BioLegend), CD127 A488 (3 µl, A019D5, BioLegend) and CD45RA PB (3 µl, HI100, BioLegend). After permeabilization/fixation for 2 h on ice, FOXP3 (APC, clone PCH101; BioLegend; 2 µl) and anti-HELIOS PE (PE, clone 22F6, BioLegend, 2 µl) staining was performed overnight at 4 °C in the dark in accordance with the manufacturer's instructions (eBioscience FOXP3/Transcription Factor Staining Buffer set). Data were recorded on an LSRII flow cytometer (BD Biosciences) and analyzed using FlowJo v.10. FOXP3 mean fluorescence intensity was normalized to the mean of HDs. To generate *t*-SNE plots from flow cytometry data, compensated, scaled data in the lymphocyte/singlet/CD3$^+$/CD4$^+$ gate were exported in .csv format from FlowJo. In R, the matrix containing CD3, CD4, FOXP3, HELIOS, CD25, CD127 and CD45RA expression was centered and scaled (by marker) before performing PCA using the prcomp function and *t*-SNE projections were calculated with package Rtsne (dims = 2, perplexity = 50, check_duplicates = F, pca = F, max_iter = 500).

Low-input RNA-seq of human and mouse samples. *RNA-seq libraries.* RNA-seq was performed with the standard ImmGen low-input protocol (www.immgen.org). Human T_{reg} and T_{conv} cells were sorted as DAPI$^-$CD4$^+$CD25^{hi}CD127^{lo} and CD4$^+$CD25$^-$CD127^{hi}, respectively on a Moflo Astrios Cell Sorter (Beckman Coulter). Mouse T_{reg} and T_{conv} cells were sorted as DAPI$^-$TCRβ$^+$CD4$^+$G/YFP$^+$CD25$^+$ and DAPI$^-$TCR$^+$CD4$^+$G/YFP$^-$, respectively. A total of 1,000 cells were double-sorted directly into 5 µl of lysis buffer (TCL Buffer (QIAGEN) supplemented with 1% 2-mercaptoethanol). Smart-seq2 libraries were prepared as previously described[52] with slight modifications. Briefly, total RNA was captured and purified on RNAClean XP beads (Beckman Coulter). Polyadenylated messenger RNA was then selected using an anchored oligonucleotide(dT) primer (50 –AAGCAGTGGTATCAACGCAGAGTACT30VN-30) and converted to complementary DNA via reverse transcription. First-strand cDNA was subjected to limited PCR amplification followed by Tn5 transposon-based fragmentation using the Nextera XT DNA Library Preparation kit (Illumina). Samples were then PCR amplified for 12 cycles using barcoded primers such that each sample carried a specific combination of eight base Illumina P5 and P7 barcodes for subsequent pooling and sequencing. Paired-end sequencing was performed on an Illumina NextSeq 500 using 2 × 38-bp reads with no further trimming.

Reads were aligned to the human genome (GENCODE GRCh38 primary assembly and gene annotations v.27) or to the mouse genome (mm10) with STAR 2.5.4a (https://github.com/alexdobin/STAR/releases). Ribosomal RNA gene annotations were removed from a .gtf (general transfer format) file. Gene-level quantification was calculated by featureCounts (http://subread.sourceforge.net/). Raw read counts tables were normalized by the median of ratios method with the DESeq2 package from Bioconductor (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) and then converted to .gct and .cls format. Samples with fewer than 1 million uniquely mapped reads were excluded from normalization to mitigate the effect of poor quality samples on normalized counts. Genes with a minimum read count of five in all replicates of a population (31,448 human genes) were retained. A pseudo count of one was added and $\log_2$-transformed before quantile normalization. Quantile-normalized counts were converted back to a linear scale.

Signatures and indexes. IPEX signature genes were identified by computing the ratio of expression in T_{reg} or T_{conv} cells of all patients with IPEX syndrome versus all HDs (*P* value as a simple uncorrected Student's *t*-test). IPEX-up indices for each individual were calculated by selecting the 100 transcripts most over-expressed in IPEX T_{reg} cells overall (FC > 3), calculating their differential expression in T_{reg} cells of each patient relative to the mean of all HD T_{reg} cells and averaging the $\log_2$ of these FCs (and similarly for IPEX-down indices from transcripts with IPEX / HD < 0.37). The T_{reg} up indices were similarly calculated (expression in T_{reg} cells of each individual over mean expression in all HDs, average the $\log_2$ of these FCs).

Gene set enrichment analysis. Gene signatures were curated from published datasets (references in the signature name)[6]. Human and mouse T_{reg} signatures have been reported[4,5]. Data were downloaded from Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/). Only datasets containing replicates were used. To reduce noise, genes with a CV between biological replicates <0.6 in either comparison groups were selected. Up- and downregulated transcripts were defined as having an FC in gene expression >1.5 or <0.6 and a Student's *t*-test *P* value <0.05. Other signatures were obtained from MSigDB C7 Immunologic signatures collection[53]. Gene set enrichment analysis with IPEX signature was performed using the hypergeometric distribution and type I error was controlled using false discovery rate. Signatures with *P* value <0.001 (all with false discovery rate < 6%) are reported.

scRNA-seq analysis of human PBMC samples. scRNA-seq was performed in several batches (different experiment dates in Supplementary Table 1 and Supplementary Fig. 4): two for cohort 1 and three for cohort 2. Cohort 1 samples were profiled with the 10X Genomics Single Cell 3′ Reagent kit (V2 chemistry), cohort 2 samples with the 10X Genomics Single Cell 3′ Reagent kit (V3 chemistry) and sample barcoding with DNA-tagged antibodies ('hashtagging')[36]. See also Supplementary Table 4.

Cell sorting and pooling using cell hashtagging. Cells were stained in 100 µl of FACS buffer (phenol red-free RPMI, 2% FBS, 0.1% azide and 10 mM HEPES, pH 7.9) for 10 min with 10 µl FcBlock (homemade) and the following antibodies: CD3 A700 (2 µl, OKT3, BioLegend), CD4 PerCP-Cy5.5 (2 µl, OKT4, BioLegend), CD25 PE-Cy7 (3 µl, BC96, BioLegend), CD127 A488 (3 µl, A019D5, BioLegend), 5 µl (2.5 µg) of a unique hashtag antibody (TotalSeq-A0251 to A0258 anti-human hashtag 1 to 8 antibody). A total of 8,000 DAPI$^-$CD3$^+$CD4$^+$ T cells were single sorted using the Moflo Astrios Cell Sorter (70-µm nozzle, Beckman Coulter) in 30 µl of PBS–BSA 0.1%. Samples with different hashtag antibodies were sorted in the same tube and the total volume was adjusted to 30 µl.

scRNA-seq libraries. Cells were encapsulated in one channel per sample (cohort 1 runs) or in one channel per pool (cohort 2 samples) of a 10X Genomics Chromium instrument and libraries were constructed with the Single Cell 3′ Reagent kit (V2 for cohort 1 and V3 chemistry for cohort 2) (https://support.10xgenomics.com/single-cell-gene-expression/library-prep/). Libraries were sequenced on the NextSeq 500 platform (28/8/0/91, Read1/i7/i5/Read2). Gene counts were obtained by aligning reads to the hg38 transcriptome using CellRanger software (v.3.0.2) (10X Genomics) using default parameters.

Hashtag libraries. Hashtag libraries were made separately as described by Stoeckius et al.[36] (https://citeseq.files.wordpress.com/2019/02/cell_hashing_protocol_190213.pdf). In brief, at the cDNA amplification step in the Single Cell 3′ Reagent kit protocol, the yield of HTO (Hashtag Oligo) products was increased using an 'additive' primer to cDNA PCR. Hashtag-derived cDNAs (<180 bp) and mRNA-derived cDNAs (>300 bp) were then separated using 0.6× SPRI bead selection. The supernatant contained the hashtag-derived cDNA that was purified with two rounds of 2× SPRI beads. The sequencing oligonucleotides were added by PCR which also amplified the Hashtag library. Libraries were sequenced on the NextSeq 500 platform (28/8/0/91, Read1/i7/i5/Read2).

Hashtag count matrices were obtained from CITE-Seq-Count package (https://zenodo.org/record/2590196) using default parameters. Each droplet from the gene count matrix was matched to a hashtag using the HTODemux function from the Seurat v.3.1.2 package. Doublets (droplets with two hashtags) were excluded and cells were assigned to the max hashtag signal. The hashtag count data were also analyzed by *t*-SNE for a visual check (clear separated clusters for each hashtag). All single cells from the gene count matrix were matched unambiguously to a single hashtag (and therefore their original donor).

Quality control. Cells with fewer than 500 (V2) or 1,000 (V3) counts were excluded. Dead cells with more than 10% of counts mapping to mitochondrial genes were excluded. Doublets were excluded using scrublet (doublet score > 0.3)[54]. Finally, cells were annotated using singleR[55] with the BluePrintEncode reference data. Cells that were not annotated as $CD4^+$ cells were excluded.

Batch correction. Batch correction was performed using the Integration method in Seurat V3 as described by Stuart et al.[37] and used only to visualize the whole dataset with UMAP. Cohort 1 and cohort 2 samples were first normalized independently using the SCTransform function in Seurat V3 with parameters to regress out the following variables: experiment date, percent of mitochondrial gene mapping and the number of counts of each single cell (SCTransform(vars.to.regress = c('experiment_date','percent_mito', 'nCount_RNA'), variable.features.*n* = 500). Integration of both cohorts together was then performed using the top 500 variable genes in both cohorts. Each cohort was projected into a common CCA space. Anchors (robust pairwise cell correspondence between datasets) were found by knn and snn graphs (FindIntegrationAnchors function) and they were used to transform the data in an integrated space. We used this integrated space, reduced to the top 40 principal components (PCs) by PCA to visualize with UMAP[56]. The number of significant PCs was determined by comparison to PCA over a randomized matrix, as described by Klein et al.[57]. Of note, another batch correction method, scAlign[58] produced similar results.

Clustering. We found shared clusters across HD and IPEX samples ('regressing out the IPEX effect') using the Integration method in Seurat V3 as described by Stuart et al.[37]. This time, each sample was first normalized independently using the SCTransform function in Seurat V3 with parameters to regress out the following variables: percent of mitochondrial gene mapping and the number of counts of each single cell (SCTransform(vars.to.regress = c('percent_mito', 'nCount_RNA'), variable.features.*n* = 500). Integration of samples together was then performed using the top 500 variable genes. Each cohort was projected into a common reciprocal PCA space. Anchors (robust pairwise cell correspondence between samples) were found by knn and snn graphs (FindIntegrationAnchors) and they were used to transform the data in a shared space. We used this shared space and reduced it to the top 32 most significant PCs by PCA for clustering[56]. In this space, a shared nearest neighbor graph was constructed from a *k*-nearest-neighbor graph ($k = 20$) by pruning cell–cell edges with less than 1/15 neighbor overlap. Community detection using the Louvain algorithm at a resolution of 0.5 found 11 clusters. Automated annotation using singleR with the BluePrintEncode reference data and manual annotation using canonical markers (see Figs. 3 and 4 and Supplementary Fig. 7) clearly distinguished resting T_{conv}, activated T_{conv} and T_{reg} cells. Type A IPEX T_{reg} cells were defined as T_{reg} cells in IPEX samples that overlapped with the HD T_{reg} cells (top left quadrant in UMAP). Type B IPEX T_{reg} cells were all the others.

Differential gene expression taking into account technical and biological variables. We used limma-trend[59], as benchmarked elsewhere[60]. Briefly, default trimmed mean of *m*-value normalization from edgeR was applied to the SCTransform-corrected count matrix. A linear model was fitted to the data using a contrast matrix including confounding variables (lmFit function). To find cluster-defining markers shared across HD and IPEX samples (regressing out the IPEX effect) or markers between type A and B IPEX T_{reg} cells, the contrast matrix contained the sample origin as a confounding variable. Empirical Bayes method was then used to estimate the overall trend of gene expression variance and adjust the genewise residual variances toward this global trend (less variance for genes trending high and more for those trending low) (eBayes(trend = TRUE, robust = TRUE)). The TopTable function was then used to extract the differential statistics corresponding to the contrast of interest. Benjamini–Hochberg correction was applied to control type I error. Adjusted *P* values <0.05 were deemed significant.

Analysis of active and inactive X chromosomes carrying FOXP3 mutation in IPEX mother cells. scRNA-seq was performed as described in the previous paragraph of peripheral blood $CD4^+$ cells from two mothers of patients with IPEX syndrome: P5's and B2's mothers (see also Supplementary Table 4). A total of 919 cells from P5's mother and 1,961 cells from B2's mother were analyzed after quality control. Because X inactivation in females occurs randomly during embryogenesis, a mix of FOXP3-deficient (Δ*FOXP3*) and -proficient (WT) T cells are present in IPEX mothers (female carriers). We sought to identify them using X chromosome variants in patients with IPEX syndrome (P5 and B2) and their mother. For each sample, we started from the CellRanger bam file (possorted_genome_bam.bam), converted it back to fastq (bamtofastq-1.2.0 --nthreads 32 possorted_genome_bam.bam fastq) and aligned the reads using STAR v.020201 ($STAR_PATH/STAR --genomeDir GRCh38_fasta --readFilesIn $READ2 --runThreadN 32 --outSAMtype BAM SortedByCoordinate) to the GRCh38 human genome to be able to run bcftools. Variants on the X chromosome were called using bcftools v.1.10.2 (bcftools mpileup -Q 30 -A -x -Ou --threads 32 -r X -f $FASTA_REF Aligned.sortedByCoord.out.bam | bcftools call -mv -Ov -o calls.vcf). We then used scAlleleCount (https://github.com/barkasn/scAlleleCount) to obtain for each single cell the reference and alternate allele count for the X chromosome variants (cell × variant matrices). We then constructed the IPEX (ΔFoxp3) X chromosome haplotype (present in sons and carrying the FOXP3 mutation) and the complementary haplotype (present in the mother and carrying the 'WT' FOXP3). To do so, we only focused on high confidence variants that were present both in the mother and in the son with a quality >10. We then filtered out the uninformative variants (variants with 100% allelic fraction in the mothers, present in both X chromosome). Using the son as a reference, we constructed the IPEX (ΔFoxp3) X chromosome haplotype and the complementary WT haplotype. For each single cell the distances to the ΔFoxp3 and the WT haplotypes were computed and we used the ratio as a score to identify whether the active X chromosome harbored WT or mutant FOXP3 (Fig. 6). There was no evidence of X chromosome recombination in mothers. In IPEX, 3 out of 815 were assigned to an active WT chromosome in P5 and 2 out of 1,612 in B2 (misclassification or microchimerism). Overall, 469 P5's mother's cells and 513 B2's mother's cells were confidently assigned to a ΔFoxp3 or WT active X chromosome (Fig. 6).

Bone marrow chimeras. C57BL/6J mice were irradiated with 10 Gy and reconstituted with equal proportions of congenically labeled T cell-depleted bone marrow cells from WT (*Foxp3*$^{ires\text{-}gfp}$x*CD45.1*) and Δ*Foxp3* (*Foxp3*$^{\Delta EGFPiCre}$*xR26*$^{-YFP}$*xCD45.1/CD45.2*) mice. Bone marrow cells were collected from femurs, tibias and hip bones from two WT and two Δ*Foxp3* male mice. After red blood cell lysis with ACK for 1 min at 4 °C, T cells were depleted; bone marrow single-cell suspensions were incubated with 20 µl of biotinylated anti-CD3e antibodies (OKT3, BioLegend) for 10 min in 2 ml of MACS buffer (PBS, FBS 0.5%, EDTA 2 mM), washed and then incubated with 200 µl of streptavidin beads (Dynabeads Biotin Binder, 11047; Thermo Fisher Scientific) for 20 min in 5 ml of MACS buffer. Isolation of the $CD3^-$ population was performed after three magnet incubations for 2 min. A total of 4 million cells (2 million WT and 2 million Δ*Foxp3*) were injected intravenously in each mouse. Mice were treated for 2 weeks with trimethoprim–sulfamethoxazol and analyzed 10 weeks later.

scRNA-seq profiling of $CD4^+$ T splenocytes cells in mice and bone marrow chimeras. scRNA-seq profiling of mouse samples was performed in two experiments (see also Supplementary Table 4). In the first experiment, two WT mice and two Δ*Foxp3* mice were profiled with the 10X Genomics Single Cell 3′ Reagent kit (V2 chemistry) at one channel per mouse. In the second experiment, $CD4^+$ T cells from three BMCs, two WT mice and two Δ*Foxp3* mice were pooled with sorted WT T_{reg} and T_{conv} cells as controls and profiled with the 10X Genomics Single Cell 3′ Reagent kit (V3 chemistry). Both experiments were analyzed together, after batch correction.

scRNA-seq libraries. Spleens were collected. After red blood cell lysis with ACK for 1 min at 4 °C, 30% of splenocytes (~30 million cells) were stained for sorting by flow cytometry in 200 µl of FACS buffer (PBS, 2% BSA) for 20 min at 4 °C in the dark, with the following antibodies (1:100 dilution): 100 µl of FcBlock (homemade), TCRβ PE-Cy7 (H57–597; BioLegend), CD4 PerCP-Cy5.5 (GK1.5; BioLegend), DAPI, CD45.1 PE-Cy7 (clone A20, BioLegend) and CD45.2 A700 (clone 104, BioLegend). Whole Δ*Foxp3* and WT mouse $CD4^+$ T cells were sorted as $DAPI^-TCR\beta^+CD4^+$. For BMC mice, Δ*Foxp3* $CD4^+$ T cells were sorted as $DAPI^-TCR\beta^+CD4^+CD45.1^+CD45.2^+$. WT $CD4^+$ T cells were sorted as $DAPI^-TCR\beta^+CD4^+CD45.1^+CD45.2^-$. A total of 40,000–50,000 cells were single sorted in 50 µl of PBS–BSA 0.05% and the total volume was adjusted to a concentration of 1,000 cells μl^{-1} after cell counting with a hemocytometer. In the first experiment, cells from each tube were encapsulated in one channel of a 10X Chromium instrument and libraries were constructed with a Single Cell 3′ Reagent kit (V2 chemistry) (https://support.10xgenomics.com/single-cell-gene-expression/library-prep/). In the second experiment, to pool samples for cell hashtagging, cells were sorted in 50 µl of PBS–BSA 0.05% in two different tubes (5,000 to 10,000 sorted cells for each sample) for sequencing in two 10X lanes and libraries were constructed with a Single Cell 3′ Reagent kit (V3 chemistry) (see also Supplementary Table 4).

Libraries were sequenced on the NextSeq 500 platform (28/8/0/91, Read1/i7/i5/Read2). Gene counts were obtained by aligning reads to the mm10 transcriptome using CellRanger software (v.3.0.2) (10X Genomics) using default parameters. The mm10 transcriptome was complemented with the transgene sequences of *Ires–Gfp*, *Yfp* and *Cre* to map reads to the *Foxp3* locus in *Foxp3*$^{ires\text{-}gfp}$/B6 and *Foxp3*$^{\Delta GFPiCre}$*xR26*$^{-YFP}$/B6 mice. In *Foxp3*$^{ires\text{-}gfp}$/B6 mice, *Foxp3* locus expression was calculated as the sum of reads mapping to *Foxp3* and *Ires–Gfp*. In *Foxp3*$^{\Delta GFPiCre}$*xR26*$^{-YFP}$/B6 mice, *Foxp3* locus expression was calculated as the sum of reads of mapping to *Foxp3* and *Cre* (not GFP because the read 2 length was not long enough to reach the *GFP* sequence).

Hashtag libraries. See similar section in the human analysis (scRNA-seq analysis of human PBMC samples)

Quality control. See similar section in the human analysis (scRNA-seq analysis of human PBMC samples). SingleR was used with the ImmGen reference data.

Batch correction. Experiment 1 and 2 were profiled with 10X V2 and V3, respectively. Batch correction was performed using the Integration method in

Seurat V3 as described by Stuart et al.[37] and only used to visualize the whole mouse dataset with UMAP. Experiment 1 (10X V2) and experiment 2 (10X V3) samples were first normalized independently using the SCTransform function in Seurat V3 with parameters to regress out the following variables: percent of mitochondrial gene mapping and the number of counts of each single cell (SCTransform(vars.to.regress = c('percent_mito', 'nCount_RNA'), variable.features.n = 500). Integration of both experiments together was then performed using the top 500 variable genes in both cohorts. Each cohort was projected into a common CCA space. Anchors were found by knn and snn graphs (FindIntegrationAnchors) and they were used to transform the data in an integrated space. We used this integrated space, reduced to the top 50 PCs by PCA to visualize with UMAP. The number of significant PCs was determined by comparison to PCA over a randomized matrix, as described by Klein et al.[57].

Clustering. We found shared clusters across samples using the Integration method in Seurat V3 as described by Stuart et al.[37]. This time, each sample was first normalized independently using the SCTransform function in Seurat V3 with parameters to regress out the following variables: percent of mitochondrial gene mapping and the number of counts of each single cell (SCTransform(vars.to.regress = c('percent_mito', 'nCount_RNA'), variable.features.n = 500). Integration of samples together was then performed using the top 500 variable genes. Each cohort was projected into a common CCA space. Anchors were found by knn and snn graphs (FindIntegrationAnchors) and they were used to transform the data in a shared space. We used this shared space and reduced it to the top 28 most significant PCs by PCA for clustering. In this space, a shared nearest neighbor graph was constructed from a k-nearest-neighbor graph (k = 20) by pruning cell–cell edges with less than 1/15 neighbor overlap. Community detection using the Louvain algorithm at a resolution of 1.5 found 11 clusters. Automated annotation using singleR with the ImmGen reference data and manual annotation using canonical markers (Supplementary Fig. 9) clearly distinguished resting T_{conv}, activated T_{conv}, resting T_{reg} and activated T_{reg} cells. Type A IPEX T_{reg} cells were defined as T_{reg} cells in IPEX samples that overlapped with the HD T_{reg} cells. Type B IPEX T_{reg} cells were all the others.

Differential gene expression. See similar section in the human analysis (scRNA-seq analysis of human PBMC samples).

Flow cytometry analysis of mouse cells. A single-cell suspension was obtained from murine splenocytes after mechanical dissociation through a 40-µm strainer using a syringe plunger. Red blood cell lysis was performed using 1 ml of ACK (Lonza) for 1 min on ice. After careful perfusion with 5 ml of sterile PBS, the right lung was collected, minced and enzymatically dissociated in RPMI containing 0.5 mg ml^{-1} DNase I (Sigma), 1.5 mg ml^{-1} collagenase IV (Sigma), 10% FBS and 2 mM EDTA with constant stirring at 37 °C for 35 min. Single-cell suspensions were then filtered through a 70-µm strainer and washed twice with 10% FBS + 2 mM EDTA RPMI buffer. Lung CD4$^+$ T cells were enriched before staining and sorting, using negative magnetic selection (Stem Cell).

After Fc blocking, antibody staining was performed in ice-cold buffer (DMEM without phenol red, 2% FBS, 1 mM EDTA) for 15 min at a dilution of 1:100 with antibodies against CD45 (30-F11; BioLegend), CD4 (GK1.5; BioLegend), CD25 (PC61; BioLegend), TCRβ (H57-597; BioLegend), CD19 (SA011F11; BioLegend) and DAPI. Stained samples were then analyzed by flow cytometry on a BD Fortessa or Symphony analyzers and processed using FlowJo v.10 (Tree Star).

IL-2 treatment. The IL-2–anti-IL-2 monoclonal antibody complex was formed by incubating 1 µg of recombinant mouse IL-2 (BioLegend) and 9 µg of purified anti-mouse IL-2 monoclonal antibody (clone JES6-1A12) (Bio-X-cell) for 20 min at 37 °C in sterile PBS. This complex was administered for 3 d consecutively per week starting at day 11 after birth on the basis of 40 ng of IL-2 per gram weight of treated mice.

Statistics. Analysis was conducted using R-3.6.2. Heat maps were made with Morpheus (https://software.broadinstitute.org/morpheus) or the pheatmap package in R (https://cran.r-project.org/web/packages/pheatmap/index.html). All other plots were made with ggplot2 (ref. [61]). Statistical tests are described in their respective method section.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data reported in this paper have been deposited in the Gene Expression Omnibus database under SuperSeries accession no. GSE168492: GSE166866 (human population RNA-seq), GSE166860 (mouse population RNA-seq), GSE167976 (human scRNA-seq) and GSE167575 (mouse scRNA-seq).

References

51. Bettelli, E. et al. Reciprocal developmental pathways for the generation of pathogenic effector T_H17 and regulatory T cells. *Nature* **441**, 235–238 (2006).
52. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
53. Godec, J. et al. Compendium of immune signatures identifies conserved and species-specific biology in response to inflammation. *Immunity* **44**, 194–206 (2016).
54. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291 (2019).
55. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
56. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* https://doi.org/10.1038/nbt.4314 (2018).
57. Klein, A. M. et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
58. Johansen, N. & Quon, G. scAlign: a tool for alignment, integration, and rare cell identification from scRNA-seq data. *Genome Biol.* **20**, 166 (2019).
59. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
60. Soneson, C. & Robinson, M. D. Bias, robustness and scalability in single-cell differential expression analysis. *Nat. Methods* **15**, 255–261 (2018).
61. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, 2009).

Acknowledgements

We thank M. Levings and A. Rudensky for insightful discussions and K. Hattori, C. Araneo, K. Seddu and the Klarman Cell Observatory team for help with mice, cell sorting and single-cell profiling. This work was funded by grants from the National Institutes of Health to C. Benoist and D.M. (AI116834 and AI150686), T.A.C. (AI085090) and L.M.C. (AI153174); the Institut National de la Santé et Recherche Médicale; the European Union Seventh Framework (269037 and 261387) and Horizon 2020 (693762); and the Agence Nationale pour la Recherche (Investissement d'Avenir ANR-10-IAHU-01) to I.A., E.S., M.D., J.L., M.C., B.N., F.R.L., F.R. and N.C.B. J.L. was supported by an INSERM Poste d'Accueil and an Arthur Sachs scholarship.

Author contributions

D.Z., L.M.C., J.L. and M.B. performed the experiments; E.S., S.K., M.D., S.B., J.Z., K.C., B.N., M.I.G.L., F.R., N.C.B., F.R.L., M.C., I.A., T.A.C., L.M.C. and C. Bruganara provided samples and discussed interpretations; D.Z., L.M.C., J.L., T.A.C., I.A., C. Benoist and D.M. designed the study and analyzed and interpreted the data; D.Z., J.L., C. Benoist and D.M. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-021-00910-8.

Correspondence and requests for materials should be addressed to D.M. or C.B.

Peer review information *Nature Immunology* thanks Fred Ramsdell and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Zoltan Fehervari was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

SUPPLEMENTARY FIGURE LEGENDS

Fig. S1. *FOXP3* locus activity in IPEX Treg-like cells

Read mapping to the FOXP3 locus from sorted CD4+ CD25+ CD127low cells in HD and IPEX samples from both cohorts (population RNAseq). Arrow indicates mutation. Mutated bases are colored. All samples from both cohorts are shown as well as control traces from representative Tconvs samples (sorted CD4+ CD25- CD127+ cells)

Fig. S2. Flow cytometric analysis of Treg-like cells in IPEX

a, b. Flow cytometric analysis of HD and IPEX CD4+ cells: CD25 and CD127 (a), FOXP3 and HELIOS (b). Gated cells are HD Tregs and IPEX Treg-like cells. All samples from both cohorts are shown. See also Fig. 1.

Fig. S3. Identification of Treg-like cells in IPEX by flow-tSNE

Flow-tSNE plots of CD3+ CD4+ cells using flow cytometric expression of CD3, CD4, CD25, CD127, HELIOS, CD45RA and FOXP3. Color represents scaled expression of CD25, FOXP3, HELIOS and CD127. All samples from both cohorts are shown.

Fig. S4. IPEX Treg-like cells maintain expression of the Treg signature, but with increased noise

a. Same ranked FC plots as 2b. showing the distribution of each Treg signature gene expression in IPEX and HD donors. The y-axis displays the gene expression foldchange in each donor over the average expression in HD Tconvs. Genes are ranked by the average Treg over Tconv

foldchange in HD. Regression lines are shown in blue. Age (in years) and summary statistics (intensity (index) and variability (coefficient of variation, CV) of expression the Treg signature), are shown for each sample.

b. Clinical correlations with the Treg UP index (computed in a). RAPA, rapamycin, CS, corticosteroids, CNI, calcineurin inhibitor, IV Ig, intravenous immunoglobulins. * Mann Whitney test $p < 0.05$.

c. Ranked FC plot showing the distribution of expression of the Treg signature gene in IPEX (orange) and HD (green) Treg samples. The y-axis corresponds to the expression ratio in IPEX Treg-like cells relative to the mean in HD Tregs. Genes are ranked by their average shift in IPEX samples.

d. Average expression of cytokine-encoding transcripts across Treg and Tconvs in IPEX (mean and SEM, normalized to mean of HD Tregs or Tconvs). * two-sided t-test $p < 0.05$.

Fig. S5. IPEX signature: reproducibility, independence from the Treg Up index, and clinical correlates.

a. IPEX/HD expression ratio in cohort 1 vs cohort 2 showing the reproducibility of the IPEX effect between the two cohorts.

b. Absence of correlation between the IPEX index (x-axis) and Treg UP index (y-axis)

c. Clinical correlations with the IPEX index. RAPA, rapamycin, CS, corticosteroids, CNI, calcineurin inhibitor, IV Ig, intravenous immunoglobulins. All comparison showed no significant differences (two-sided t.test).

d. Proportion of human endogenous retroviruses (HERV) mapped reads in IPEX and HD Treg and Tconv samples.

Fig. S6. scRNAseq analysis of CD4+ cells in IPEX and HD reveals a stable IPEX signature that affect all CD4+ cells (resting, activated Tconvs and Tregs)

a. Same UMAP plots as 3a., showing the reproducible segregation of HD and IPEX samples: plots are split by experiment and cohort. Individual HD and IPEX donors are highlighted in different colors.

b. Same UMAP as 3a. showing the stability of the IPEX transcriptomic signature: before and after treatment initiation (P7), over several days (B4a and B4b, four days apart) and or years (> 2 years apart, P4), and across two different experiments (B5, technical replicates)

c. UMAP1 correlates with the IPEX signature defined by population RNAseq (Pearson correlation r = 0.75). y-axis shows the IPEX/HD expression ratio of the IPEX signature genes (population RNAseq). The x axis shows their correlation with UMAP1 in scRNAseq.

d. Same UMAP as 3a. showing each HD and IPEX donor individually. Blue, green, and red cells represent resting Tconvs, activated Tconvs, and Tregs, respectively.

Fig. S7. Identification of resting Tconvs, activated Tconvs and Tregs in IPEX by scRNAseq

a. Single-cell biclustering heatmap of canonical resting Tconv, activated Tconv and Treg genes. Top ribbons indicate donor origin and annotations for every single cell.

b. Similar proportions of Tregs, resting and activated Tconvs in total CD4+ cells in HD and IPEX (scRNAseq).

Fig. S8. Heterogeneous Treg-like cells in IPEX (A and B types) identified by scRNAseq

a. Same UMAP as 3a. showing each HD and IPEX donor individually. Blue, green, and red cells represent resting Tconvs, activated Tconvs, and Tregs, respectively. *FOXP3*-expressing cells (RNA) are in yellow. An arrow indicates type-A IPEX Tregs, overlapping with HD Tregs.

b. Down-tuning of the IPEX signature in type A IPEX Tregs vs. type B IPEX Tregs. Volcano plot comparing the gene expression profiles of type A versus type B IPEX Tregs (p values from two-sided t.test). Up- and downregulated signature genes are highlighted (red and blue, respectively). $\chi 2$ -test p values.

Fig. S9. Identification of resting Tconvs, activated Tconvs and Tregs in Δ*Foxp3* mice by scRNAseq

a. Same UMAP as 5a showing the individual samples in both experiments and the absence of batch effect.

b. Shared Δ*Foxp3* signature in spleen and lung Tregs. Top: Δ*Foxp3* vs WT expression ratio in spleen vs lung Tregs. Bottom: Volcano plot comparing the expression profile of Δ*Foxp3* vs WT lungs Tregs. Up- and downregulated Δ*Foxp3* signature genes are highlighted (red and blue, respectively). $\chi 2$ -test p values, n =3 mice per group, population RNAseq.

c. Single-cell biclustering heatmap of the expression of canonical genes in WT and Δ*Foxp3* Tregs , resting Tconvs, and activated Tconvs. Top ribbons indicate mouse origin and cluster annotations for each single cell.

d. Proportion of Tregs, resting and activated Tconvs in total CD4+ cells in WT and Δ*Foxp3* mice (scRNAseq). * two-sided t.test $p < 0.05$ (n = 4 WT, 4 Δ*Foxp3* mice).

Fig. S10. Lower proportion of Δ*Foxp3* Tregs and absence of the Δ*Foxp3* signature in Δ*Foxp3* Tconvs and Tregs in mixed bone marrow chimera with WT cells.

a. Δ*Foxp3* signature expression in whole mice and mixed bone marrow chimera (50/50 WT and Δ*Foxp3*) showing the downregulation of the Δ*Foxp3* signature expression in resting, activated Δ*Foxp3* Tconv and Δ*Foxp3* Tregs in 50/50 BMC. **** two-sided t.test $p < 10^{-4}$ (n = 4 WT mice, 4 Δ*Foxp3* mice, 3 BMC mice).

b. WT Tregs outcompete Δ*Foxp3* Tregs in in 50/50 BMC. Proportion of WT and Δ*Foxp3* Tregs, resting and activated Tconvs in total CD4+ cells. ** two-sided t.test $p < 10^{-2}$ (n = 3 mice).

Figure S1

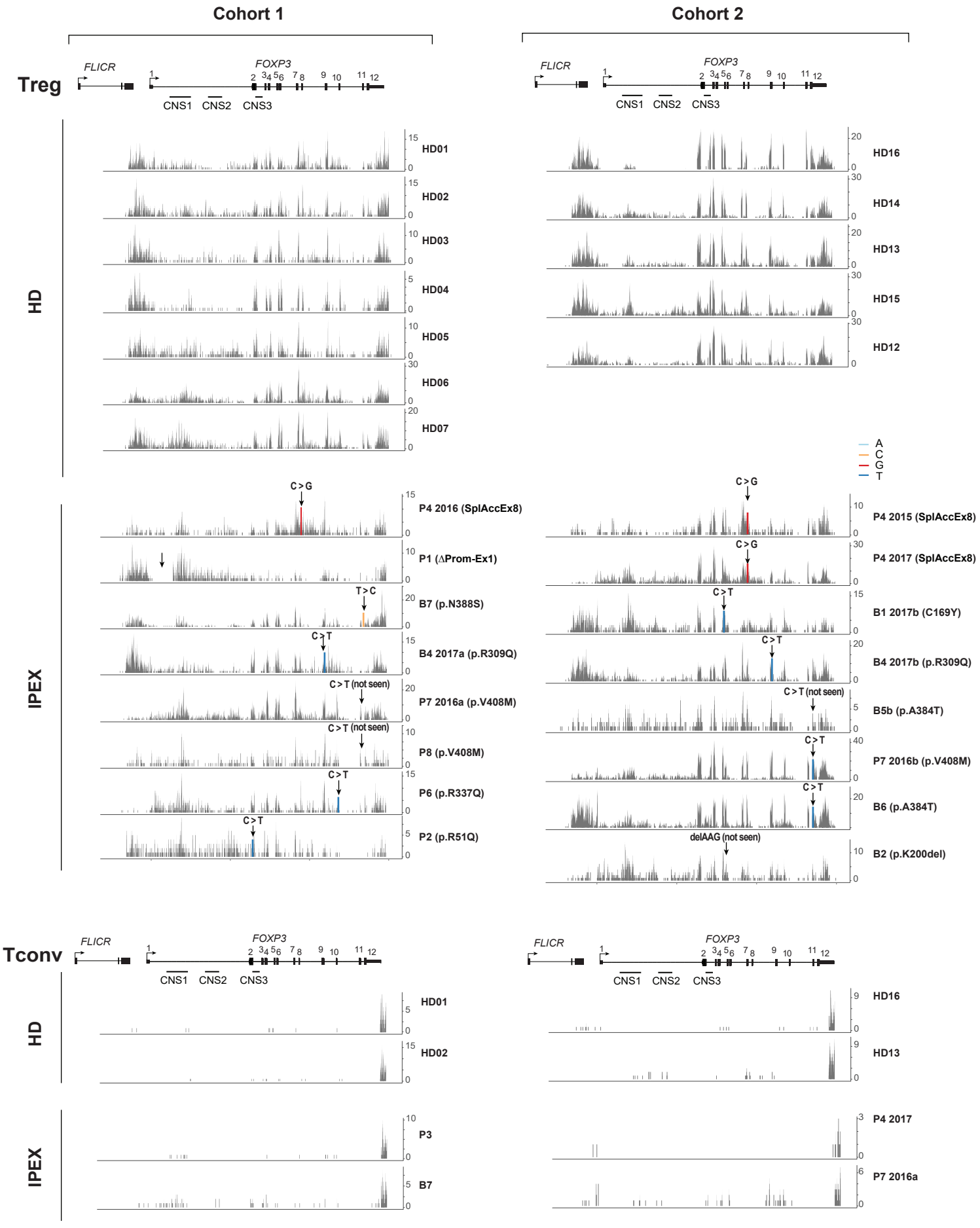

Cohort 1
Cohort 2
Treg
FLICR
FOXP3
CNS1
CNS2
CNS3
HD
HD01
HD02
HD03
HD04
HD05
HD06
HD07
HD16
HD14
HD13
HD15
HD12
IPEX
A
C
G
T
C>G
P4 2016 (SplAccEx8)
P1 (ΔProm-Ex1)
T>C
B7 (p.N388S)
C>T
B4 2017a (p.R309Q)
C>T (not seen)
P7 2016a (p.V408M)
P8 (p.V408M)
P6 (p.R337Q)
P2 (p.R51Q)
P4 2015 (SplAccEx8)
P4 2017 (SplAccEx8)
B1 2017b (C169Y)
B4 2017b (p.R309Q)
B5b (p.A384T)
P7 2016b (p.V408M)
B6 (p.A384T)
delAAG (not seen)
B2 (p.K200del)
Tconv
P3
B7
P4 2017
P7 2016a

Figure S2

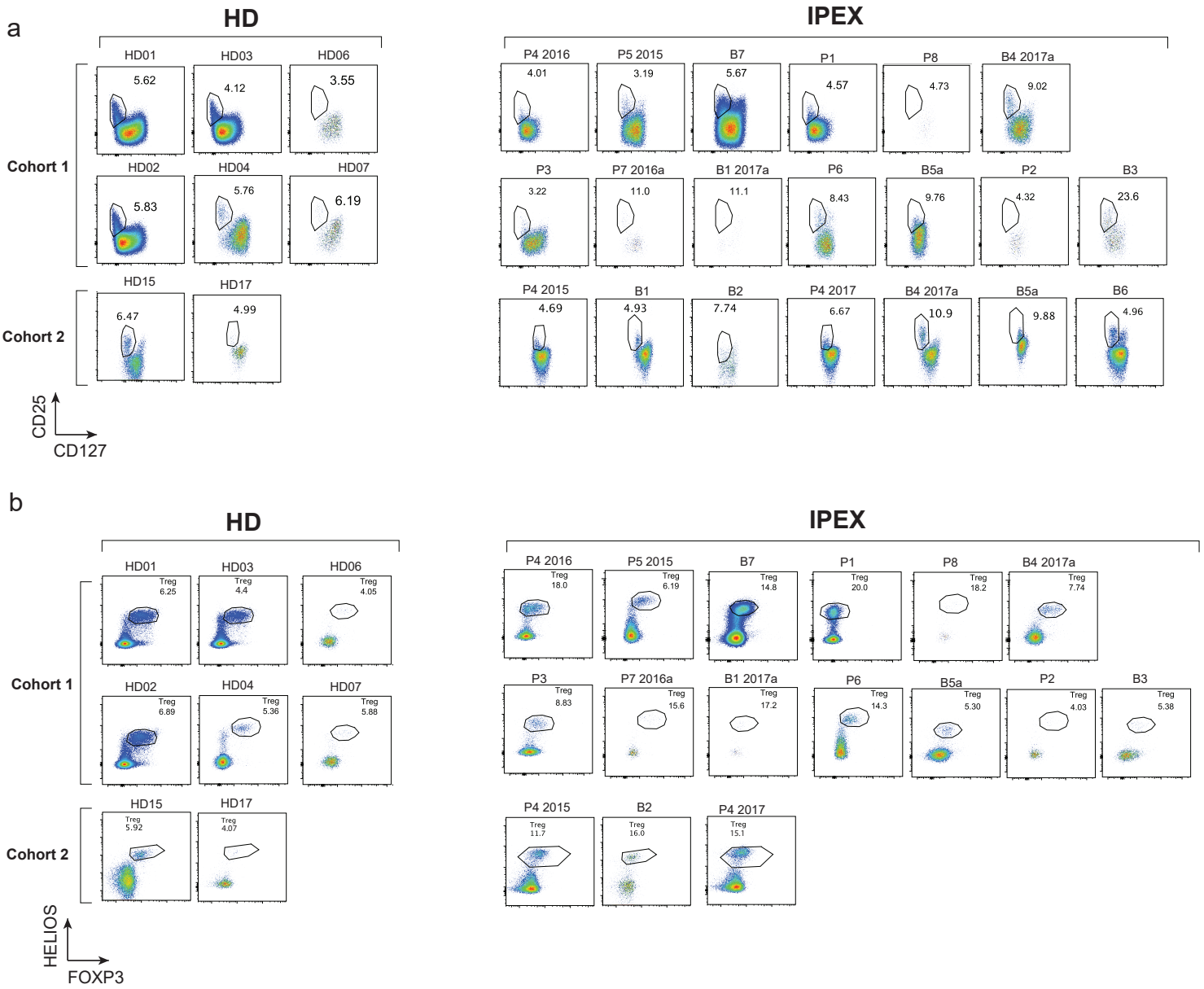

a
HD
IPEX
HD01
HD03
HD06
5.62
4.12
3.55
Cohort 1
HD02
HD04
HD07
5.83
5.76
6.19
HD15
HD17
6.47
4.99
Cohort 2
CD25
CD127
P4 2016
P5 2015
B7
P1
P8
B4 2017a
4.01
3.19
5.67
4.57
4.73
9.02
P3
P7 2016a
B1 2017a
P6
B5a
P2
B3
3.22
11.0
11.1
8.43
9.76
4.32
23.6
P4 2015
B1
B2
P4 2017
B4 2017a
B5a
B6
4.69
4.93
7.74
6.67
10.9
9.88
4.96
b
HD
IPEX
HD01
HD03
HD06
Treg 6.25
Treg 4.4
Treg 4.05
Cohort 1
HD02
HD04
HD07
Treg 6.89
Treg 5.36
Treg 5.88
HD15
HD17
Treg 5.92
Treg 4.07
Cohort 2
HELIOS
FOXP3
P4 2016
P5 2015
B7
P1
P8
B4 2017a
Treg 18.0
Treg 6.19
Treg 14.8
Treg 20.0
Treg 18.2
Treg 7.74
P3
P7 2016a
B1 2017a
P6
B5a
P2
B3
Treg 8.83
Treg 15.6
Treg 17.2
Treg 14.3
Treg 5.30
Treg 4.03
Treg 5.38
P4 2015
B2
P4 2017
Treg 11.7
Treg 16.0
Treg 15.1

Figure S3

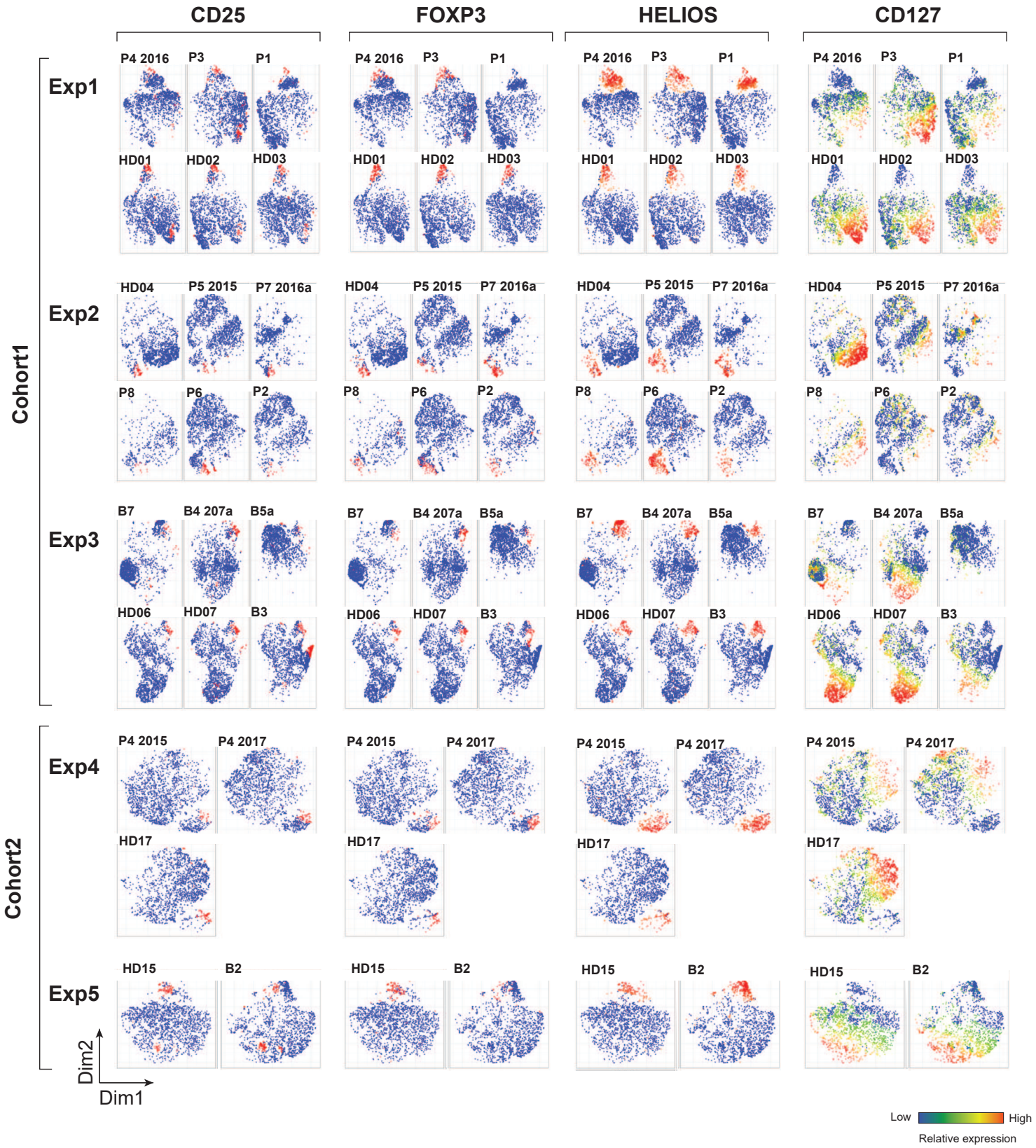
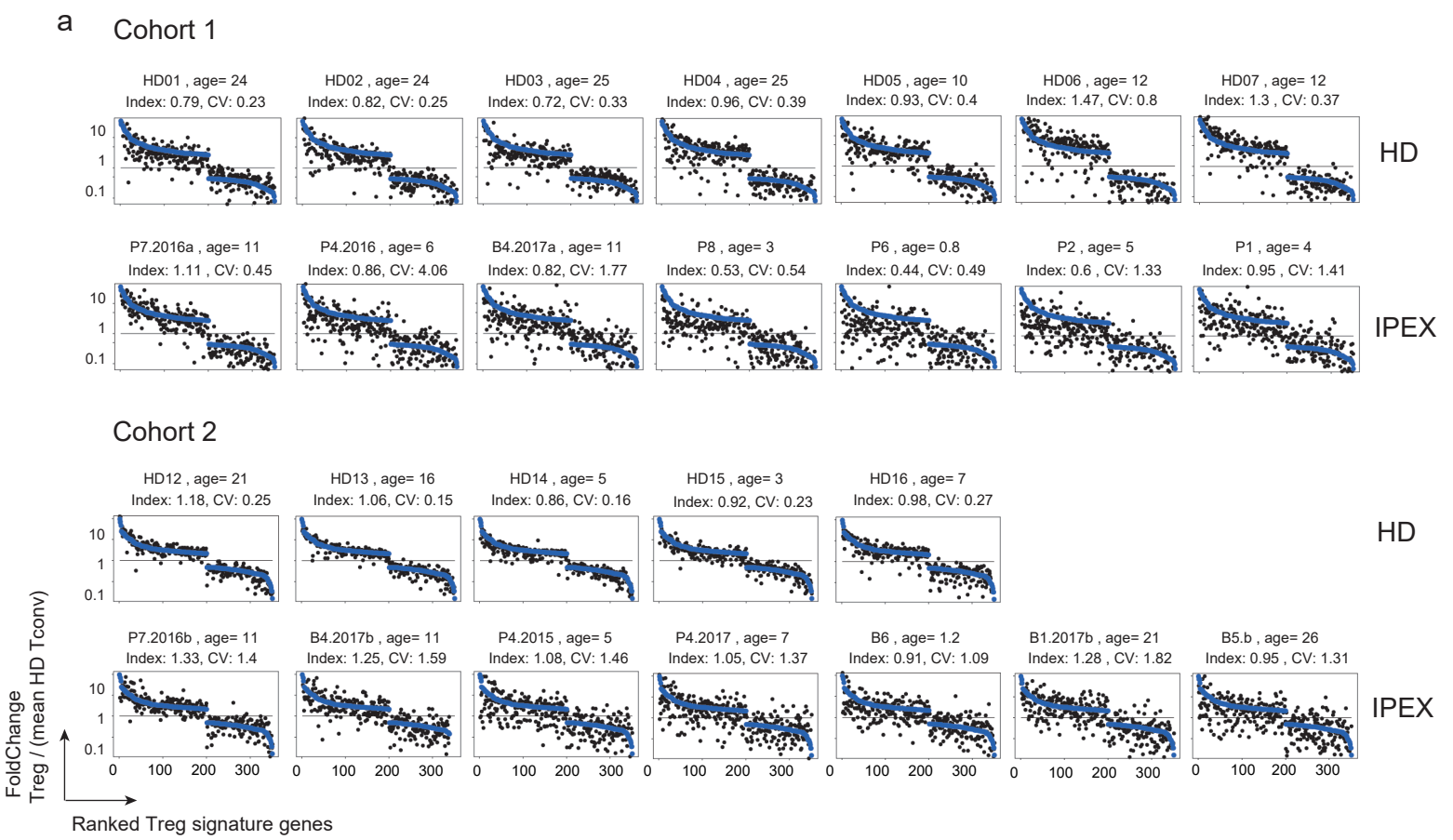
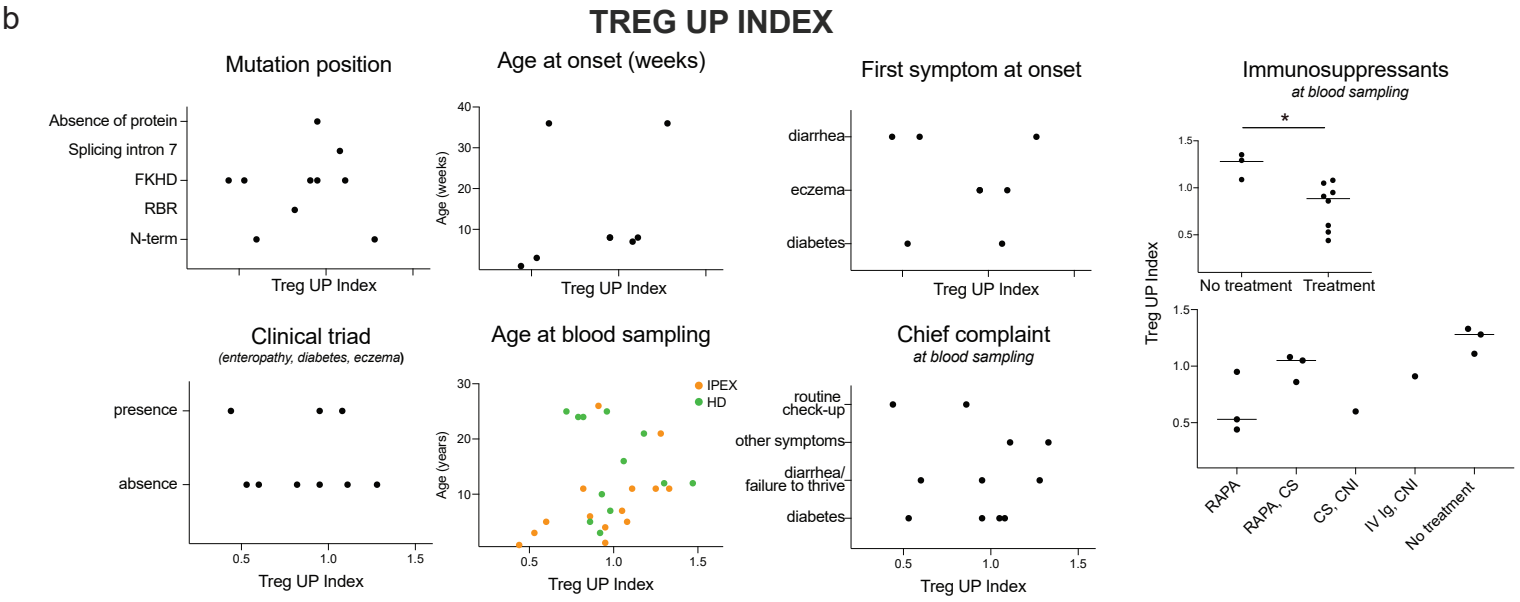
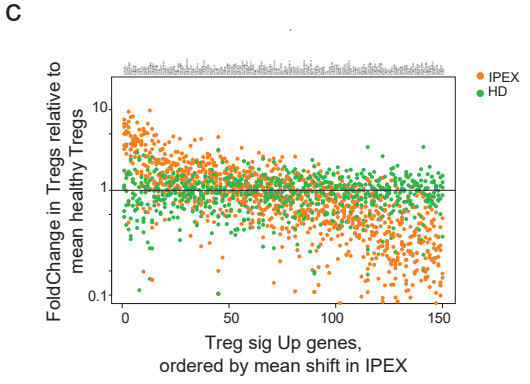

CD25
FOXP3
HELIOS
CD127
Cohort1
Exp1
P4 2016
P3
P1
HD01
HD02
HD03
Exp2
HD04
P5 2015
P7 2016a
P8
P6
P2
Exp3
B7
B4 207a
B5a
HD06
HD07
B3
Cohort2
Exp4
P4 2015
P4 2017
HD17
Exp5
HD15
B2
Dim2
Dim1
Low
High
Relative expression

Figure S4


a
Cohort 1
HD01 , age= 24
Index: 0.79, CV: 0.23
HD02 , age= 24
Index: 0.82, CV: 0.25
HD03 , age= 25
Index: 0.72, CV: 0.33
HD04 , age= 25
Index: 0.96, CV: 0.39
HD05 , age= 10
Index: 0.93, CV: 0.4
HD06 , age= 12
Index: 1.47, CV: 0.8
HD07 , age= 12
Index: 1.3 , CV: 0.37
HD
P7.2016a , age= 11
Index: 1.11 , CV: 0.45
P4.2016 , age= 6
Index: 0.86, CV: 4.06
B4.2017a , age= 11
Index: 0.82, CV: 1.77
P8 , age= 3
Index: 0.53, CV: 0.54
P6 , age= 0.8
Index: 0.44, CV: 0.49
P2 , age= 5
Index: 0.6 , CV: 1.33
P1 , age= 4
Index: 0.95 , CV: 1.41
IPEX
Cohort 2
HD12 , age= 21
Index: 1.18, CV: 0.25
HD13 , age= 16
Index: 1.06, CV: 0.15
HD14 , age= 5
Index: 0.86, CV: 0.16
HD15 , age= 3
Index: 0.92, CV: 0.23
HD16 , age= 7
Index: 0.98, CV: 0.27
HD
P7.2016b , age= 11
Index: 1.33, CV: 1.4
B4.2017b , age= 11
Index: 1.25, CV: 1.59
P4.2015 , age= 5
Index: 1.08, CV: 1.46
P4.2017 , age= 7
Index: 1.05, CV: 1.37
B6 , age= 1.2
Index: 0.91, CV: 1.09
B1.2017b , age= 21
Index: 1.28 , CV: 1.82
B5.b , age= 26
Index: 0.95 , CV: 1.31
IPEX
FoldChange
Treg / (mean HD Tconv)
Ranked Treg signature genes


b
TREG UP INDEX
Mutation position
Absence of protein
Splicing intron 7
FKHD
RBR
N-term
Treg UP Index
Age at onset (weeks)
Age (weeks)
Treg UP Index
First symptom at onset
diarrhea
eczema
diabetes
Treg UP Index
Immunosuppressants
at blood sampling
Treg UP Index
No treatment
Treatment
Clinical triad
(enteropathy, diabetes, eczema)
presence
absence
Treg UP Index
Age at blood sampling
Age (years)
IPEX
HD
Treg UP Index
Chief complaint
at blood sampling
routine check-up
other symptoms
diarrhea/ failure to thrive
diabetes
Treg UP Index
RAPA
RAPA, CS
CS, CNI
IV Ig, CNI
No treatment


c
FoldChange in Tregs relative to mean healthy Tregs
IPEX
HD
Treg sig Up genes,
ordered by mean shift in IPEX

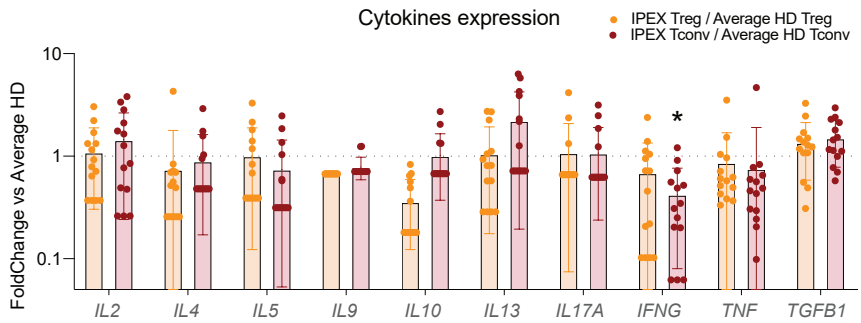
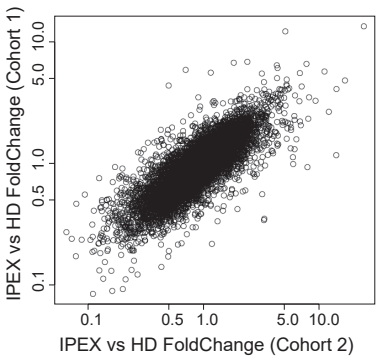
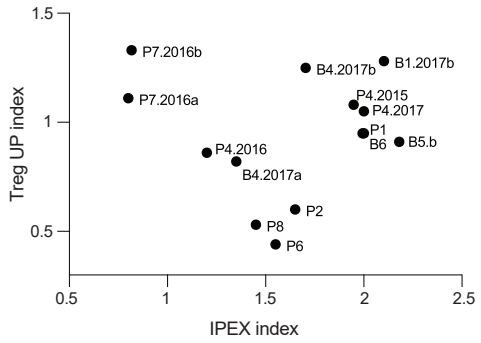

d
Cytokines expression
IPEX Treg / Average HD Treg
IPEX Tconv / Average HD Tconv
FoldChange vs Average HD
IL2
IL4
IL5
IL9
IL10
IL13
IL17A
IFNG
TNF
TGFB1

Figure S5

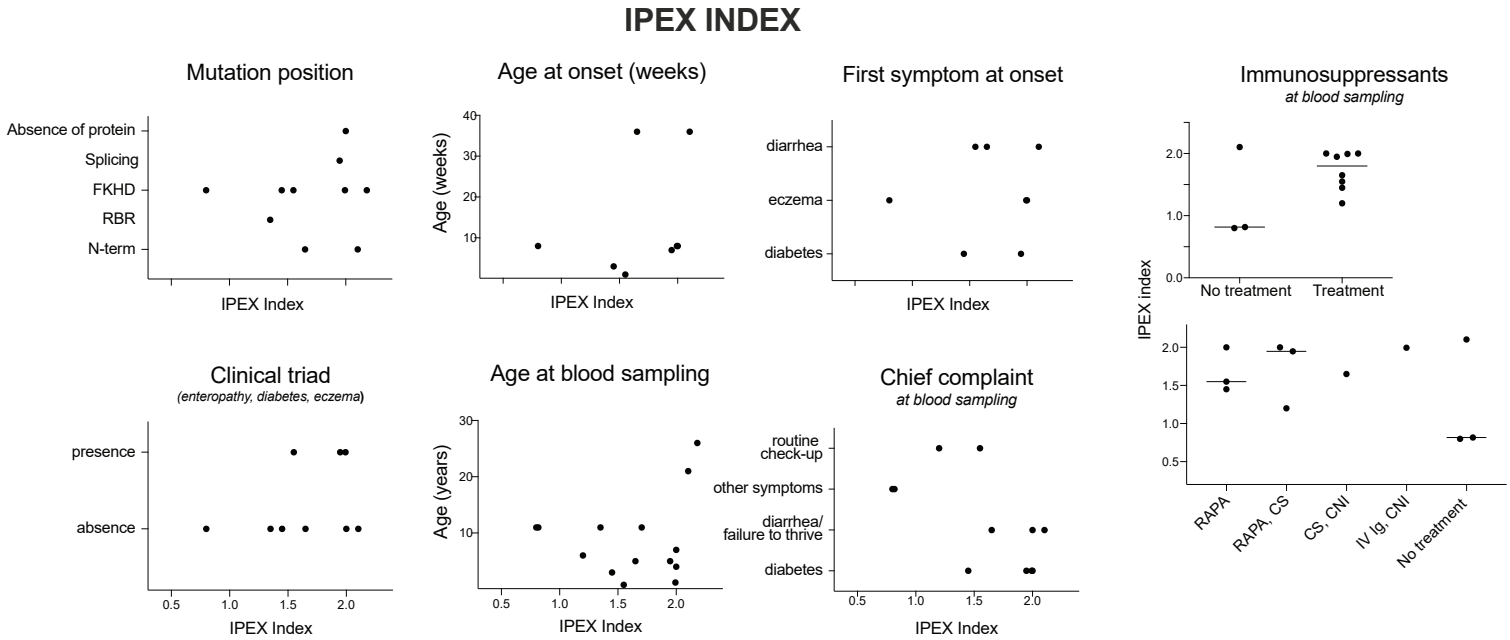
a



b



c



d

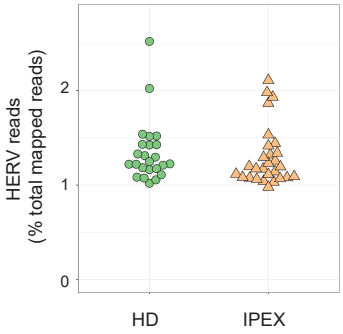
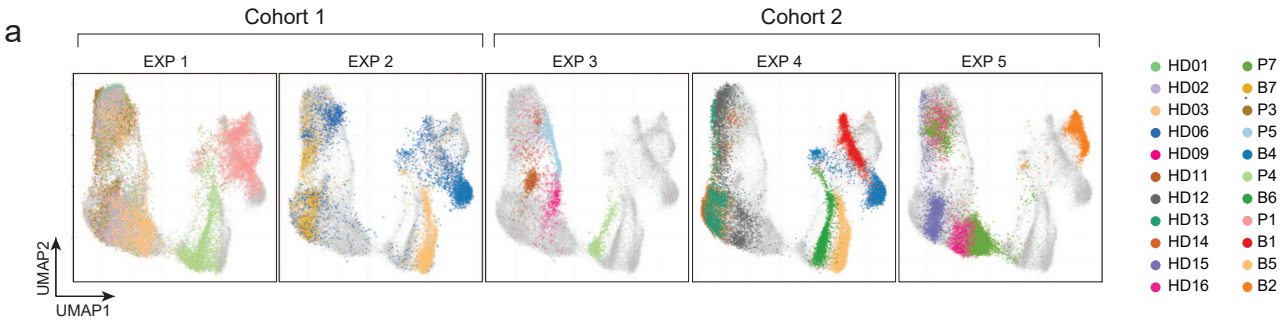
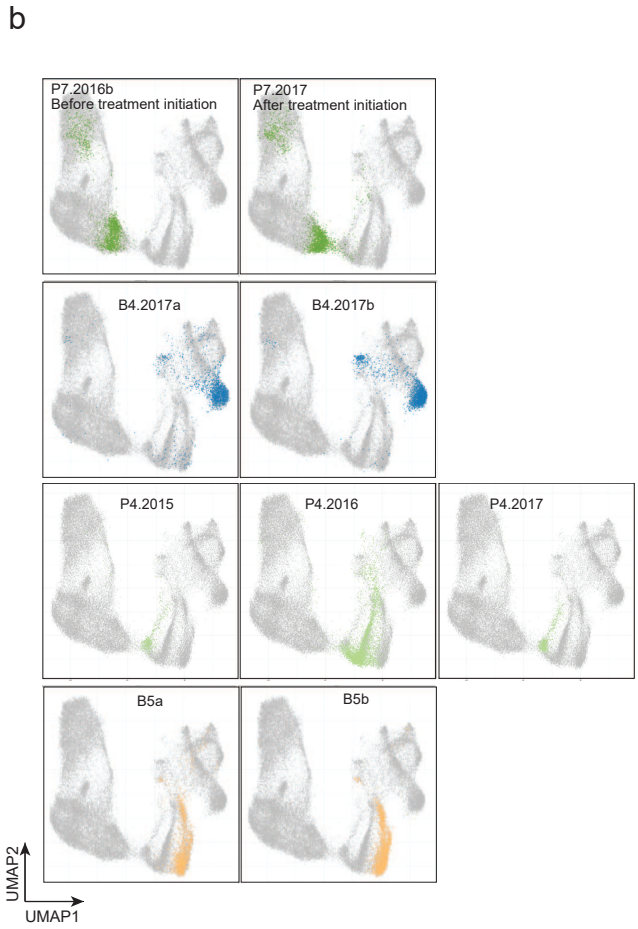
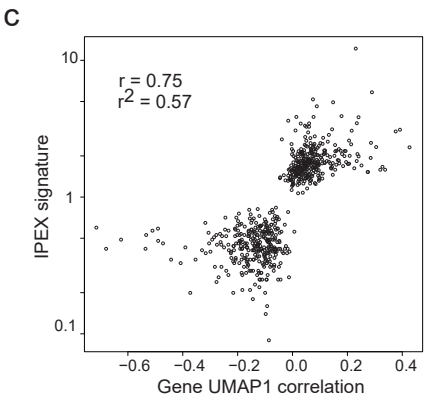


Figure S6


a
Cohort 1
Cohort 2
EXP 1
EXP 2
EXP 3
EXP 4
EXP 5
UMAP2
UMAP1
HD01
HD02
HD03
HD06
HD09
HD11
HD12
HD13
HD14
HD15
HD16
P7
B7
P3
P5
B4
P4
B6
P1
B1
B5
B2


b
P7.2016b
Before treatment initiation
P7.2017
After treatment initiation
B4.2017a
B4.2017b
P4.2015
P4.2016
P4.2017
B5a
B5b
UMAP2
UMAP1


c
r = 0.75
r2 = 0.57
IPEX signature
10
1
0.1
−0.6
−0.4
−0.2
0.0
0.2
0.4
Gene UMAP1 correlation

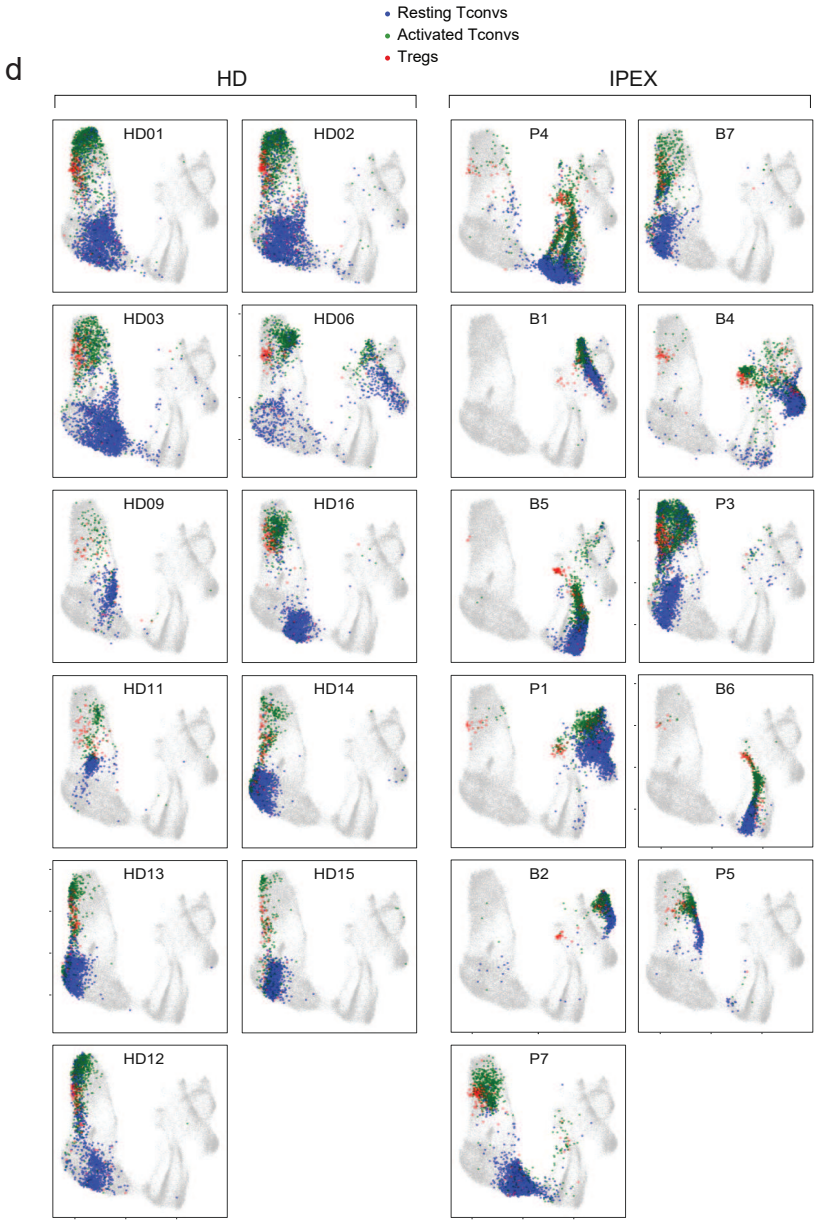

d
Resting Tconvs
Activated Tconvs
Tregs
HD
IPEX
HD01
HD02
HD03
HD06
HD09
HD16
HD11
HD14
HD13
HD15
HD12
P4
B7
B1
B4
B5
P3
P1
B6
B2
P5
P7

Figure S7


a
HD
IPEX
Resting Tconv
Activated Tconv
Treg
Donor
Cluster
CCR7
SELL
TCF7
GATA3
GZMA
GZMB
ITGB1
KLRG1
KLRB1
MAF
PFN1
S100A6
S100A10
IL7R
FOXP3
IL2RA
IKZF2
CTLA4
TIGIT
Genes
Cells
0
2
4
Scaled expression

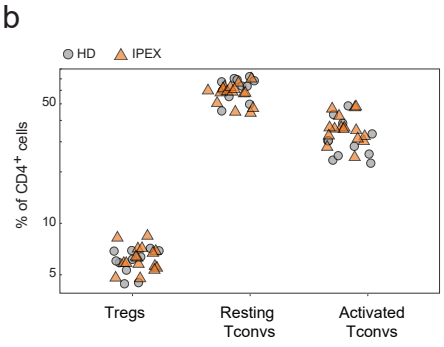

b
HD
IPEX
% of CD4+ cells
50
10
5
Tregs
Resting Tconvs
Activated Tconvs

Figure S8

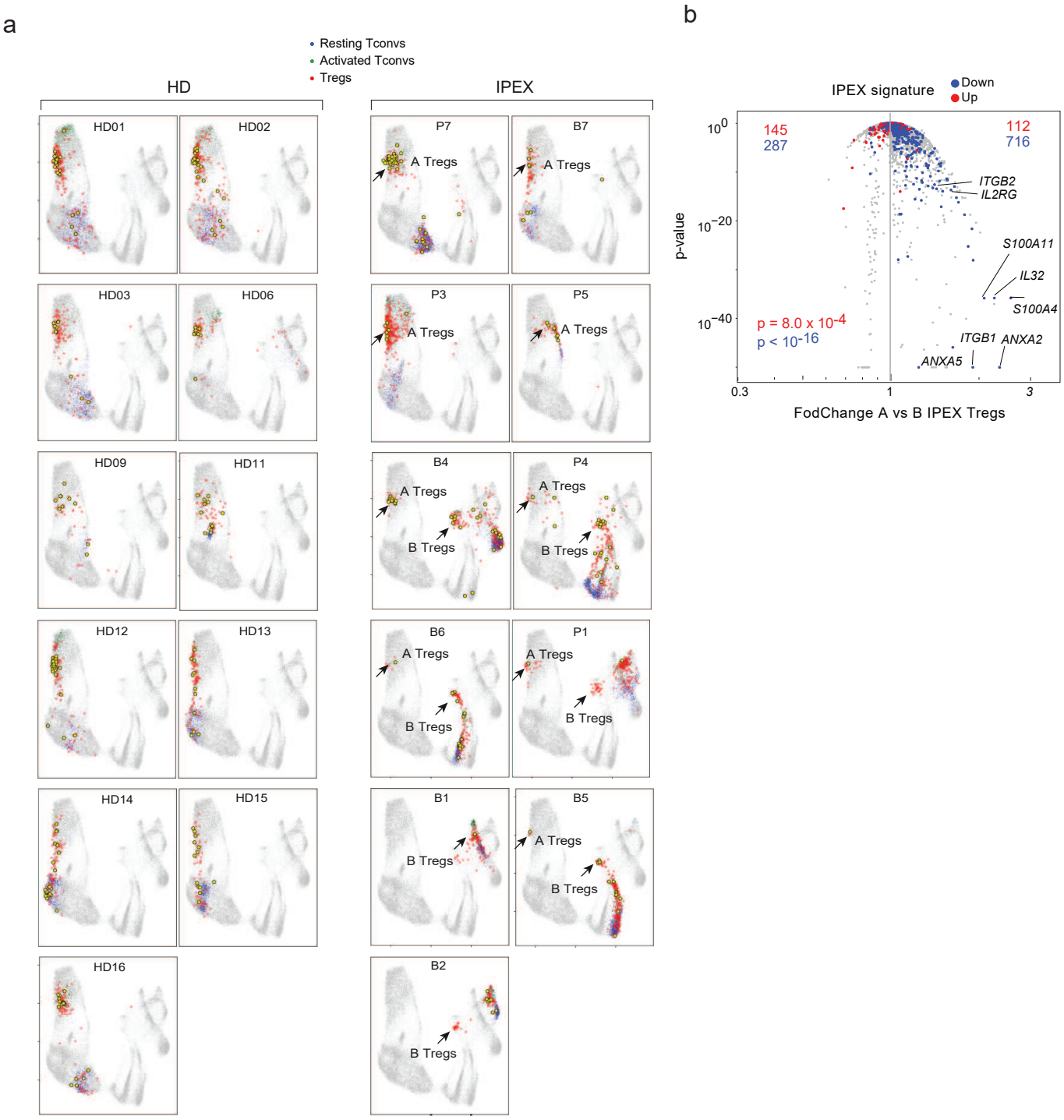

a
Resting Tconvs
Activated Tconvs
Tregs
HD
IPEX
HD01
HD02
HD03
HD06
HD09
HD11
HD12
HD13
HD14
HD15
HD16
P7
B7
P3
P5
B4
P4
B6
P1
B1
B5
B2
A Tregs
B Tregs
b
IPEX signature
Down
Up
145
287
112
716
ITGB2
IL2RG
S100A11
IL32
S100A4
ITGB1
ANXA2
ANXA5
p = 8.0 x 10-4
p < 10-16
p-value
FodChange A vs B IPEX Tregs

Figure S9

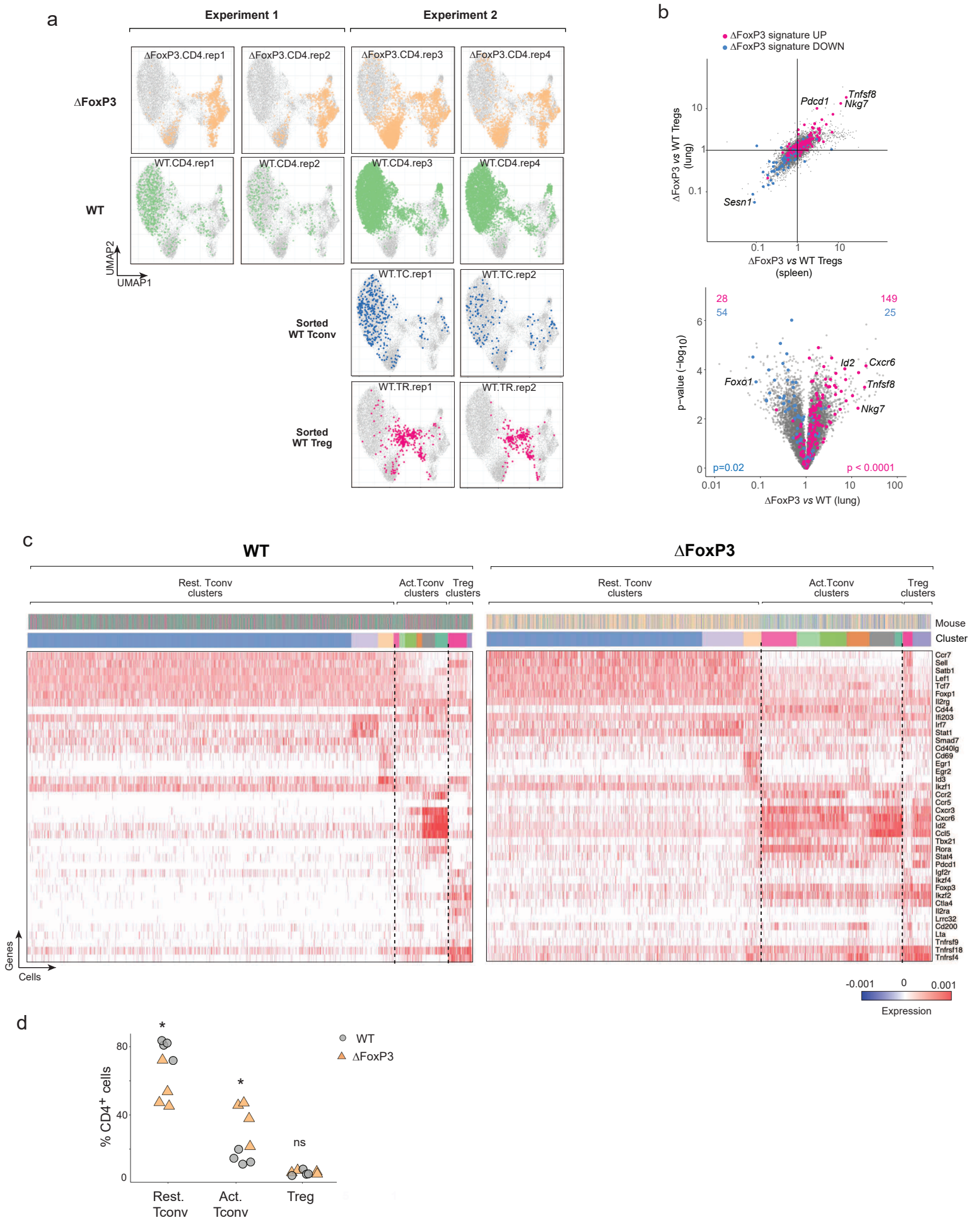
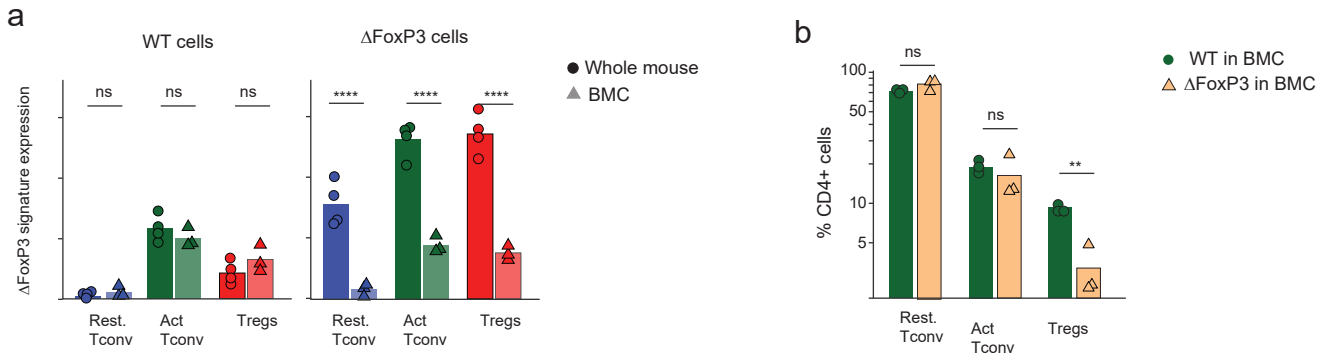

a
Experiment 1
Experiment 2
ΔFoxP3.CD4.rep1
ΔFoxP3.CD4.rep2
ΔFoxP3.CD4.rep3
ΔFoxP3.CD4.rep4
ΔFoxP3
WT.CD4.rep1
WT.CD4.rep2
WT.CD4.rep3
WT.CD4.rep4
WT
UMAP2
UMAP1
WT.TC.rep1
WT.TC.rep2
Sorted WT Tconv
WT.TR.rep1
WT.TR.rep2
Sorted WT Treg
b
ΔFoxP3 signature UP
ΔFoxP3 signature DOWN
Pdcd1
Tnfsf8
Nkg7
Sesn1
ΔFoxP3 vs WT Tregs (lung)
ΔFoxP3 vs WT Tregs (spleen)
28
149
54
25
Id2
Cxcr6
Foxo1
Tnfsf8
Nkg7
p=0.02
p < 0.0001
p−value (−log10)
ΔFoxP3 vs WT (lung)
c
WT
ΔFoxP3
Rest. Tconv clusters
Act.Tconv clusters
Treg clusters
Mouse
Cluster
Genes
Cells
Ccr7
Sell
Satb1
Lef1
Tcf7
Foxp1
Il2rg
Cd44
Ifi203
Irf7
Stat1
Smad7
Cd40lg
Cd69
Egr1
Egr2
Id3
Ikzf1
Ccr2
Ccr5
Cxcr3
Cxcr6
Id2
Ccl5
Tbx21
Rora
Stat4
Pdcd1
Igf2r
Ikzf4
Foxp3
Ikzf2
Ctla4
Il2ra
Lrrc32
Cd200
Lta
Tnfrsf9
Tnfrsf18
Tnfrsf4
-0.001
0
0.001
Expression
d
% CD4+ cells
WT
ΔFoxP3
*
*
ns
Rest. Tconv
Act. Tconv
Treg

Figure S10


a
WT cells
ΔFoxP3 cells
Whole mouse
BMC
ΔFoxP3 signature expression
ns
ns
ns
****
****
****
Rest. Tconv
Act Tconv
Tregs
Rest. Tconv
Act Tconv
Tregs
b
WT in BMC
ΔFoxP3 in BMC
% CD4+ cells
100
50
10
5
ns
ns
**
Rest. Tconv
Act Tconv
Tregs

APPENDIX B, *REFERS TO CHAPTER 2*

Mutations from patients with IPEX ported to mice reveal different patterns of FoxP3 and Treg dysfunction.

Juliette Leon[1,2], Kaitavjeet Chowdhary[1], Wenxiang Zhang[3], Ricardo N. Ramirez[1], Isabelle André[2], Sun Hur[3], Diane Mathis[1], and Christophe Benoist[1,4,+]

[1]Department of Immunology, Harvard Medical School, Boston, USA

[2] INSERM UMR 1163, University of Paris, Imagine Institute, Paris, France

[3] Howard Hughes Medical Institute, Program in Cellular and Molecular Medicine, Boston Children's Hospital, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, USA

[4]Lead Contact

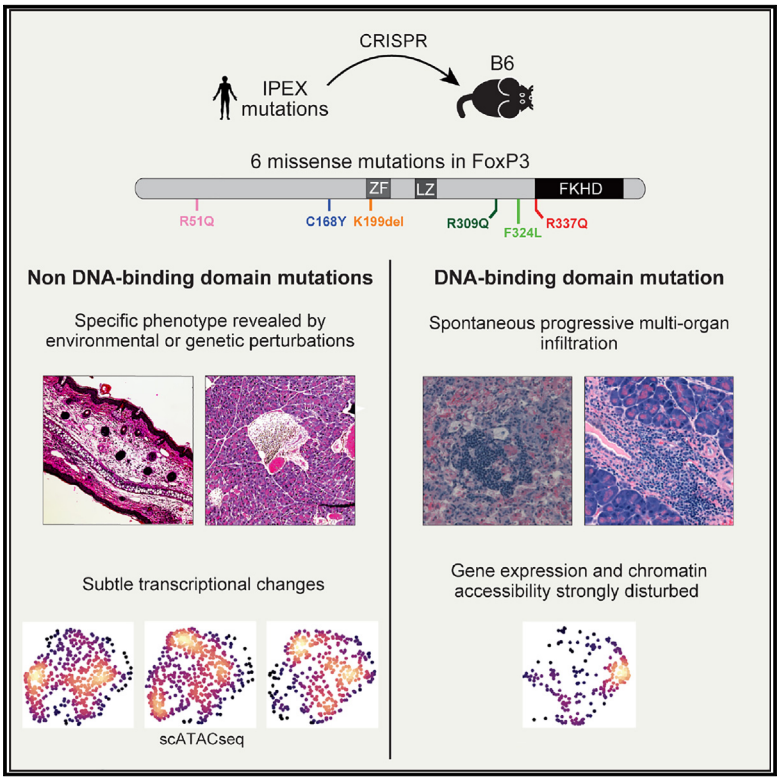
[+]Correspondence to: Christophe Benoist
Department of Immunology
Harvard Medical School
77 Avenue Louis Pasteur, Boston, MA 02115
e-mail: cbdm@hms.harvard.edu

Cell Reports

Article

Mutations from patients with IPEX ported to mice reveal different patterns of FoxP3 and Treg dysfunction

Graphical abstract



Highlights

- A panel of 6 FOXP3 mutations from patients with IPEX CRISPRed into the mouse germline
- Normal Tregs profoundly alter phenotypes of mutant Tregs by rescue and/or competition
- DNA-binding mutation induces a spontaneous multiorgan autoimmunity and high IgE
- Perturbations reveal organ-specific autoimmunity from N-terminal IPEX mutations

Authors

Juliette Leon, Kaitavjeet Chowdhary, Wenxiang Zhang, ..., Sun Hur, Diane Mathis, Christophe Benoist

Correspondence

cbdm@hms.harvard.edu

In brief

Leon et al. show that IPEX FoxP3 mutations can split into two. A Forkhead domain mutation induces spontaneous multiorgan inflammation, while inflammatory or genetic challenges reveal organ-specific impacts of mutations outside the DNA-binding domain. IPEX heterogeneity results from the location of FoxP3 mutations as well as from genetic and environmental perturbations.

Leon et al., 2023, Cell Reports *42*, 113018
August 29, 2023
https://doi.org/10.1016/j.celrep.2023.113018

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Mutations from patients with IPEX ported to mice reveal different patterns of FoxP3 and Treg dysfunction

Juliette Leon,[1,2] Kaitavjeet Chowdhary,[1] Wenxiang Zhang,[3] Ricardo N. Ramirez,[1] Isabelle André,[2] Sun Hur,[3] Diane Mathis,[1] and Christophe Benoist[1,4,*]

[1]Department of Immunology, Harvard Medical School, Boston, MA, USA
[2]INSERM UMR 1163, University of Paris, Imagine Institute, Paris, France
[3]Howard Hughes Medical Institute, Program in Cellular and Molecular Medicine, Boston Children's Hospital, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA
[4]Lead contact
*Correspondence: cbdm@hms.harvard.edu
https://doi.org/10.1016/j.celrep.2023.113018

SUMMARY

Mutations of the transcription factor FoxP3 in patients with "IPEX" (immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome) disrupt regulatory T cells (Treg), causing an array of multiorgan autoimmunity. To understand the functional impact of mutations across FoxP3 domains, without genetic and environmental confounders, six human *FOXP3* missense mutations are engineered into mice. Two classes of mutations emerge from combined immunologic and genomic analyses. A mutation in the DNA-binding domain shows the same lymphoproliferation and multiorgan infiltration as complete FoxP3 knockouts but delayed by months. Tregs expressing this mutant FoxP3 are destabilized by normal Tregs in heterozygous females compared with hemizygous males. Mutations in other domains affect chromatin opening differently, involving different cofactors and provoking more specific autoimmune pathology (dermatitis, colitis, diabetes), unmasked by immunological challenges or incrossing NOD autoimmune-susceptibility alleles. This work establishes that IPEX disease heterogeneity results from the actual mutations, combined with genetic and environmental perturbations, explaining then the intra-familial variation in IPEX.

INTRODUCTION

Immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is a rare and severe autoimmune disorder caused by mutations in the *FOXP3* gene, which perturb the homeostasis and function of regulatory T cells (Tregs). Patients with IPEX present with severe immune dysregulation and develop autoimmune diseases at a very young age, as do *Foxp3*-deficient *scurfy* mice. As *FOXP3* is located on chromosome X, patients with IPEX are hemizygous males, while heterozygous carrier females are protected by normal Tregs that coexist with cells that express the mutant allele.[1–3] Many IPEX-causing *FOXP3* mutations have been reported,[4,5] but IPEX remains a rare disease.

Our understanding of the mode of operation of FoxP3, a transcription factor (TF) of the Forkhead family, is unresolved at present. While it contributes to Treg function, it is neither necessary nor sufficient to establish Treg identity: Treg-like cells develop in the absence of FoxP3, and Treg-specific gene expression signatures comprise FoxP3-independent and -dependent modules.[3,6–10] Some have argued that FoxP3 is primarily a transcriptional repressor[11–15]; indeed, the repression of cytokines produced upon activation of conventional $CD4^+$ (Tconv) T cells, especially interleukin-2 (IL-2), is a well-established function of FoxP3. Conversely, several studies suggest that FoxP3 activates the majority of its target genes, with the difference between activation and repression being determined by the regulatory partners it coopts at each locus[7,16–19]. Further, it has been proposed that FoxP3 mostly functions indirectly, by regulating the expression of intermediate TFs like TCF1,[20] but we argued elsewhere that indirect control applies only to a minor segment of the Treg signature.[21]

IPEX includes enteropathy, dermatitis, and variable endocrine autoimmunity (primarily type 1 diabetes or thyroiditis) and, more rarely, autoimmune hepatitis, nephropathy, and cytopenias.[22–26] Disease typically begins in very young infants, prenatally in some cases, but other patients are diagnosed as adolescents. The range in IPEX clinical severity stems in part from the actual mutation, with complete loss-of-function (LOF) alleles being generally most deleterious while missense and small deletions are better tolerated.[27,28] But clinical manifestations and severity also vary widely between patients with the same mutation.[25] The variance may result from genetic modifiers that modulate the perturbed peripheral tolerance and its inflammatory consequences (severity in *scurfy* mice also varies with inbred backgrounds). However, immunological challenges and infectious history may also influence IPEX by tuning alternative tolerance pathways

and/or the degree of effector cell activation. Genetic modifiers and environmental elements are near impossible to ascertain in IPEX families, as the rarity of the disease precludes realistically powered genetic association or microbiome studies.

We recently used single-cell transcriptomics to analyze Treg dysregulation in a set of patients with IPEX and their relatives.[3] Every patient hosted Treg-like cells expressing the mutant FOXP3 protein. A monomorphic disease signature affected all $CD4^+$ T cells, whether Treg or Tconvs. This signature was cell extrinsic because it was extinguished by the presence of normal Treg cells in mixed bone marrow chimeric mice and in heterozygous mothers of patients with IPEX. We proposed that both aspects contribute to the pathology by compounding Treg dysregulation. The cell-intrinsic effects of each IPEX-causing mutation proved difficult to assess given the dominant cell-extrinsic perturbation, as well as the genetic and immunologic variables in each patient.

To circumvent these difficulties, and to analyze in a controlled manner the influence of genetic and environmental variables on the functional consequences of FOXP3 mutations, we used CRISPR germline editing to port into mice a panel of six *FOXP3* IPEX mutations. The results identify 2 different classes of IPEX mutations, objectify the importance of perturbations (genetic or inflammatory), and reveal an association between mutation location and specific disease manifestations.

RESULTS

Engineering mouse lines with FoxP3 missense mutations from patients with IPEX

The primary goal of this study was to examine, in a setting where genetic and environmental confounders are controlled, the genomic and immunologic consequences of IPEX-causing *FOXP3* mutations by introducing them into the genome of inbred C57BL/6J (B6) mice. We selected a panel of 6 IPEX missense mutations, mainly derived from our previous study of Treg genomics in patients with IPEX.[3] We eschewed complete LOF mutations, which cause full-blown *scurfy*-like disease and mutations in the DNA-binding domain that have been extensively analyzed,[29–32] and instead chose missense mutations that give rise to a range of pathologic and transcriptomic perturbations in patients[3] at positions conserved in the human and mouse FoxP3 proteins (Figure S1A). The mutations were spread across different domains of FoxP3 (Figure 1A): R51Q and C168Y in the N-terminal domain; K199del in the zinc finger domain; R309Q and F324L between the leucine zipper and Forkhead (FKHD) domains; and R337Q, located at the beginning of the FKHD and predicted to distinguish different conformations.[33] With regard to their clinical phenotype, patients from whom the mutations originated suffered from characteristic enteropathy, but several of them presented milder forms of the disease. All patients were still alive at last follow-ups, with ages ranging from 3 to 21 years, with two (K199del, R337Q) having required bone marrow transplantation[3,34] (Figure 1A; Table S1).

The mutations were introduced into the mouse genome by CRISPR-based mutagenesis via microinjection of editing complexes (Cas9 protein, gRNA, and ssDNA oligonucleotide for homology-driven repair) into the male pronucleus of fertilized oocytes. These oocytes derived from B6 females crossed to a Foxp3-GFP[36] male, aiming to introduce the mutations into a GFP-tagged *Foxp3* locus. However, this was achieved in only half the instances (Figure S1B), likely because the editing complex leaked into the cytoplasm and edited the B6-derived female pronucleus. For reference, we used a similarly generated mouse line bearing a frameshift mutation that eliminates the FKHD[19] (*Foxp3fs327-gfp*, hereafter KO [knockout]).

Founder animals with the desired mutations were identified by PCR and sequencing and were bred to expand and establish the lines. We have previously reported a preliminary assessment of R337Q.[33] All *Foxp3* exons were sequenced to verify the absence of adventitious mutations (Figure S1B). For experiments, the breeding strategy generated (1) heterozygous females in which the mutant allele was balanced by a wild-type (WT) allele (provided by a *Foxp3-Thy1.1* reporter); because of random chromosome X (ChrX) inactivation, such females contained both Treg-like cells expressing the mutant FoxP3 and normal Tregs expressing the WT allele; the latter ensured normal immunologic homeostasis (all heterozygous females healthy and fertile) and provided internal control cells in each mouse; and (2) hemizygous males expressing only the mutant FoxP3 and thus exposed to autoimmune manifestations, the equivalent of patients with IPEX (Figure S1C). Breedings were tailored to obtain experimental animals and sex-matched WT controls from the same litter (see STAR Methods; lines with mutations in the GFP-tagged or untagged alleles were controlled correspondingly; Figure S1C).

We first evaluated the expression of the mutant FoxP3 proteins *in vivo* in Treg-like cells from heterozygous females devoid of inflammation (Figure 1B). All mutant proteins were readily detected at levels equivalent to those of WT FoxP3, except for R337Q and K199del, which showed partially reduced levels (Figure 1C). These decreases did not result from interference with the epitope recognized by the anti-Foxp3 monoclonal antibody (mAb) used for detection (which binds a different sequence) and were confirmed with an alternative anti-FoxP3 mAb (Figure S1D). Tregs in all mutant lines showed largely normal levels of *Foxp3* mRNA (Figure 1D).

We also tested whether these mutations impacted FoxP3's DNA-binding ability. Proteins with these mutations were expressed *in vitro* (HEK-293T cells), with a C-terminal hemagglutinin (HA) tag. Extracts from these cells were used in pull-down assays, with a double-stranded DNA (dsDNA) oligonucleotide that contains an inverted repeat of the canonical TGTTTAC FKHD motif, which we recently showed to form a high-efficacy binding site for FoxP3 dimers.[33] The R337Q mutation led to a partial reduction in binding (Figure 1E), consistent with its position in the FKHD and with our previous results,[33] as did the neighboring F324L mutation, which might affect the domain-swapped conformation of FoxP3.[37] We also noted modest but reproducible reductions in DNA-binding activity for some N-terminal mutations. Overall, the mouse lines expressed mutant FoxP3 proteins at or near normal levels.

Differentiation and homeostasis of Tregs expressing the mutant FoxP3 proteins

To explore the functional impact of these FoxP3 IPEX mutations, we first assessed by flow cytometry their influence on Treg differentiation and homeostasis. To reveal the mutation's cell-intrinsic

CellPress
OPEN ACCESS

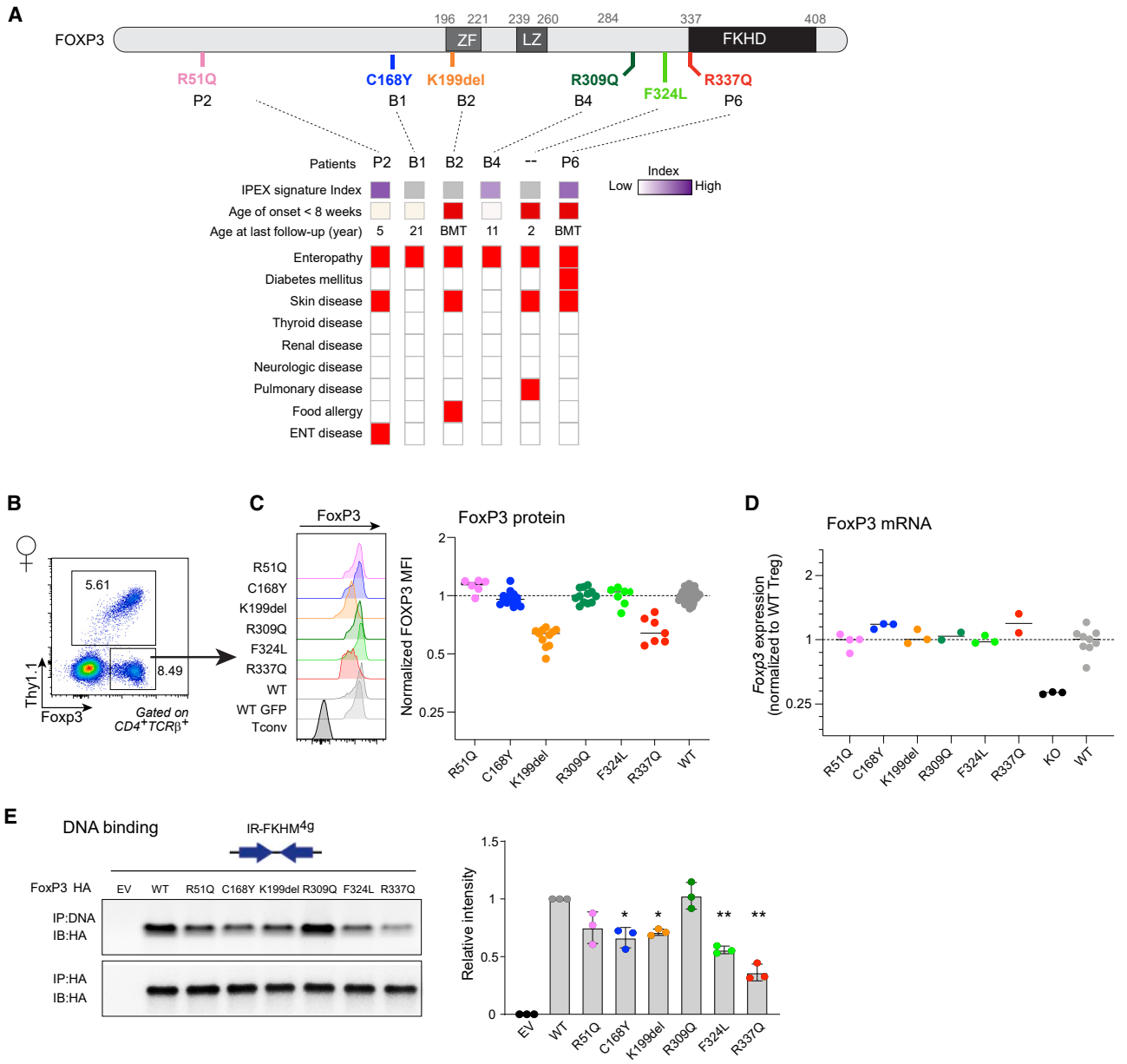


Figure 1. Selection and expression of mutant FoxP3 molecules

(A) Position of the engineered mutations into the mouse FOXP3 protein, and clinical characteristics of the human patients with IPEX from which they originate (coded as in Zemmour et al.,[35] except F324L, which is from Bacchetta et al.[34]). BMT, treated by bone marrow transplantation after sampling. Domains: ZF, zinc finger; LZ, leucine-zipper; FKHD, Forkhead. The "IPEX signature index" denoted severity by quantifying in Zemmour et al.[35] the transcriptomic disease footprint in CD4$^+$ T cells.

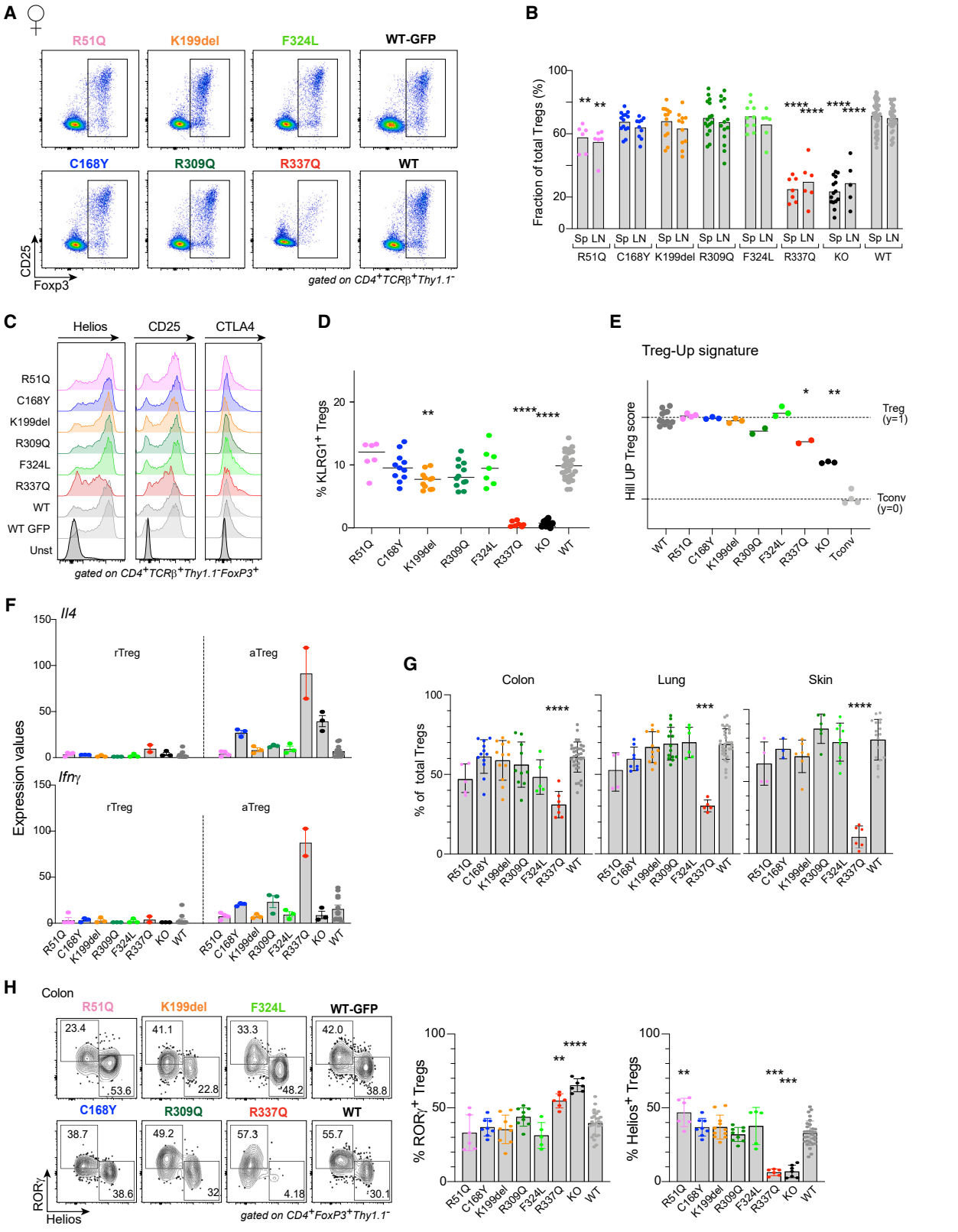
(B) Gating strategy for Tregs in heterozygous females to distinguish mutant Treg (Thy1.1$^-$Foxp3$^+$) from WT Treg (Thy1.1$^+$Foxp3$^+$).

(C) Foxp3 staining in splenic mutant Tregs, gated as in (B), in heterozygous females, and its quantification in a dot plot. MFI quantification at right was normalized vs. the mean MFI of WT littermates from the same experiment and background (B6 or B6.*Foxp3-ires-gfp*). Each dot is an individual mouse.

(D) Foxp3 mRNA expression in mutant Tregs, from RNA-seq profiling, also normalized to the mean of matched WT controls. Each dot is an individual mouse.

(E) DNA-binding capacity of the different mutations in HEK293T cells. HA-tagged FOXP3 molecules (WT or mutant) were transiently expressed in 293T cells, pulled down by anti-HA beads, and incubated with FOXP3-binding ds-oligonucleotide IR-FKHM4g.[33] Left: representative blot; right: quantitation from 3 independent experiments; t test *p < 0.05, **p < 0.01.

effects, we studied heterozygous females, which are protected from inflammation by the presence of WT Tregs,[1,3,20] mutant Tregs competing with normal Tregs driven by the WT *Foxp3-Thy1.1* allele (Figures 1B and S1C; for clarity, the presentation below only refers to Thy1.1-negative mutant Tregs). All lines, except for R337Q, had normal proportions of Tregs expressing



(legend on next page)

the mutant FoxP3 proteins in lymphoid organs, as shown by Foxp3/CD25 plots (Figures 2A and 2B). The R337Q mutation induced a significant decrease in mutant Treg proportions to a level similar to that of the full KO (Figure 2B), accompanied by a drop in FoxP3 and CD25 mean fluorescence intensity (MFI). K199del Tregs also showed reductions in FoxP3 and CD25, but these were more muted (Figures 2A and 2C).

We then analyzed more broadly Treg phenotypes in FoxP3 mutant mice. Except for R337Q, typical Treg markers like Helios, CD25, and CTLA-4 were normally expressed in Tregs from all mice (Figure 2C). Tregs in lymphoid organs exhibit significant diversity but are most simply grouped as "effector" or "activated" (aTreg) and "resting" (rTreg) Treg populations. Since aTregs have higher anti-inflammatory potential,[38–40] we asked if the mutations impacted the ability to mature into aTregs. Screening a panel of effector markers showed that the mutations did not impair aTreg maturation or *in vivo* proliferative capacity reflected by Ki67. However, R337Q (and, more subtly, K199del) led to the total absence of the KLRG1$^+$ Treg subset and to a significant decrease in CD44^{hi}CD62^{lo} Tregs (Figures S2A, S2B, and 2D).

A possible loss of Treg identity was investigated by RNA sequencing (RNA-seq) profiling of purified CD4$^+$TCRb$^+$CD25^{hi} Thy1.1$^-$ Tregs, profiling separately the CD44$^-$CD62L^{hi} (rTreg) and CD44^{hi}CD62L$^{lo/-}$ (aTreg) populations (normal Foxp3-Thy1.1$^+$ Tregs from some of the same mice were also profiled and showed no deviation). With a "Treg score" calculated from canonical "Treg signature" transcripts,[7] these FoxP3$^+$ cells were indeed *bona fide* Tregs, with a strong bias for expression of Treg signature genes (Figure 2E). R337Q was the exception, with a score intermediate between normal and FoxP3-deficient Treg "wannabes" in full KO mice.[3,8,10,41] Normal expression of the Treg signature was confirmed on volcano plots (Figure S2C) and ranked signature plots (Figure S2D), which highlighted the marked shift in R337Q Tregs but also uncovered deviations in signature genes in some mutant Tregs (R51Q, K199del, R309Q), further detailed below. An essential FoxP3 function is cytokine-encoding gene repression, like *Ifng*. Among mutants, derepression of cytokine genes was evident for R337Q Tregs, as illustrated for *Ifng* and *Il4* (Figure 2F); other mutant lines closely mirrored WT controls (with perhaps the exception of *Il4* in C168Y).

To finish this survey of Treg populations, we verified their ability to migrate and reside in non-lymphoid tissues (Figure 2G). Mutant Tregs were found in similar proportions as WT Tregs in the colon, lung, and skin (Figure 2G), again with the exception of R337Q. R337Q Tregs were outcompeted to different degrees, especially in the skin (Figure 2G). Colonic Tregs include a balance of Helios$^+$ and RORg$^+$ subsets.[42] Only R337Q perturbed this ratio, with a dramatic dearth of Helios$^+$ colonic Tregs (Figure 2H), which is of interest in light of recent reports that Helios$^+$ Tregs are far more dependent on FoxP3 than RORg$^+$ Tregs.[20,43]

Overall, most of these IPEX-causing mutations allowed quasi-normal Treg differentiation and homeostasis in heterozygous females, with R337Q being the notable exception.

Impact of the missense mutation in hemizygous males

We then studied the effects of FoxP3 mutations in hemizygous males, a context similar to that of IPEX patients, without the protection from WT Tregs found in heterozygous females. At 8 weeks old, these mice were devoid of overt signs of pathology. Tregs were present in lymphoid organs of mutant mice in the same proportion as in WT males, except for R337Q, which showed a 2-fold increase, the opposite of the change noted in females (Figures 3A and 3B). R337Q Tregs also recovered FoxP3 and CD25 expression to levels closer to those of WT Tregs (Figure S3A); R337Q mice also included a sizable proportion of FoxP3^{int}CD25$^-$ cells, as was already described in full FoxP3 deficiencies[3,8,10,41] (Figure 3A). In contrast, the mild decrease in FoxP3 MFI noted in K199del females was also present in males (Figures 3A and S3A). R337Q Tregs were characterized by a high activation state (Figures 3C and S3B)—although less than in the full KO—and upregulated various aTreg markers like PD-1, KLRG1, and CXCR6 (Figures 3D and S3B). In non-lymphoid tissues, mutant males had normal Treg proportions, except for R337Q, which had an increase in gut, lungs, and skin (Figure 3E), a mirror image of the paucity of tissue Tregs observed in R337Q females. Finally, and perhaps most striking, was the subset distribution among colonic Tregs of R337Q males: a high proportion of Helios$^+$ Tregs and fewer RORg$^+$ Tregs (Figure 3F)—again, the exact opposite of the ratio observed in females (Figure 2G).

In short, most mutations had little impact on Treg homeostasis at baseline in hemizygous males, except for R337Q. The contrast in cell phenotypes between males and females highlighted how the interaction between intrinsic Treg dysfunction and the cellular environment (including competition from WT Tregs) can influence the expression and impact of the mutations.

Pathological effects of missense mutations

The next step was to investigate the mutation's impact on immunological tolerance in hemizygous males. We first explored it in

Figure 2. Treg phenotypes in heterozygous females

(A) Representative flow cytometric CD25/FoxP3 plots of gated CD4$^+$TCRB$^+$Thy1.1$^-$ splenocytes from heterozygous females; the two WT littermates shown are from the B6 and B6.*Foxp3-ires-gfp* backgrounds.

(B) Proportions of Tregs expressing the mutant FoxP3 among total Tregs in spleen from heterozygous females.

(C) Expression of Helios, CD25, and CTLA-4 in FoxP3$^+$ Tregs (representative of at least 3 experiments).

(D) Proportions of KLRG1$^+$ cells among splenic Tregs from heterozygous females.

(E) rTregs were sorted as CD19$^-$TCRB$^+$CD4$^+$CD25$^+$Thy1.1$^-$ and their RNA-seq profiles determined. The Treg-Up gene expression signature from Hill et al.[7] was used to compute a Treg score (scaled to WT Treg = 1 and Tconv = 0).

(F) Expression of *Il4* and *Ifng* transcripts in Tregs in the RNA-seq data from (E).

(G) Proportions of Tregs expressing the mutant FoxP3 among total Tregs in different tissues.

(H) Representative RORg/Helios plots of colon Tregs expressing the different mutant FoxP3 or WT control littermates (quantification of multiple experiments at bottom).

All results are from 2–4 independent experiments; each dot is an individual mice; error bars indicate mean ± SD; *p < 0.05, **p < 0.01, ***p < 0.0001, ****p < 0.0001, from Mann-Whitney test.

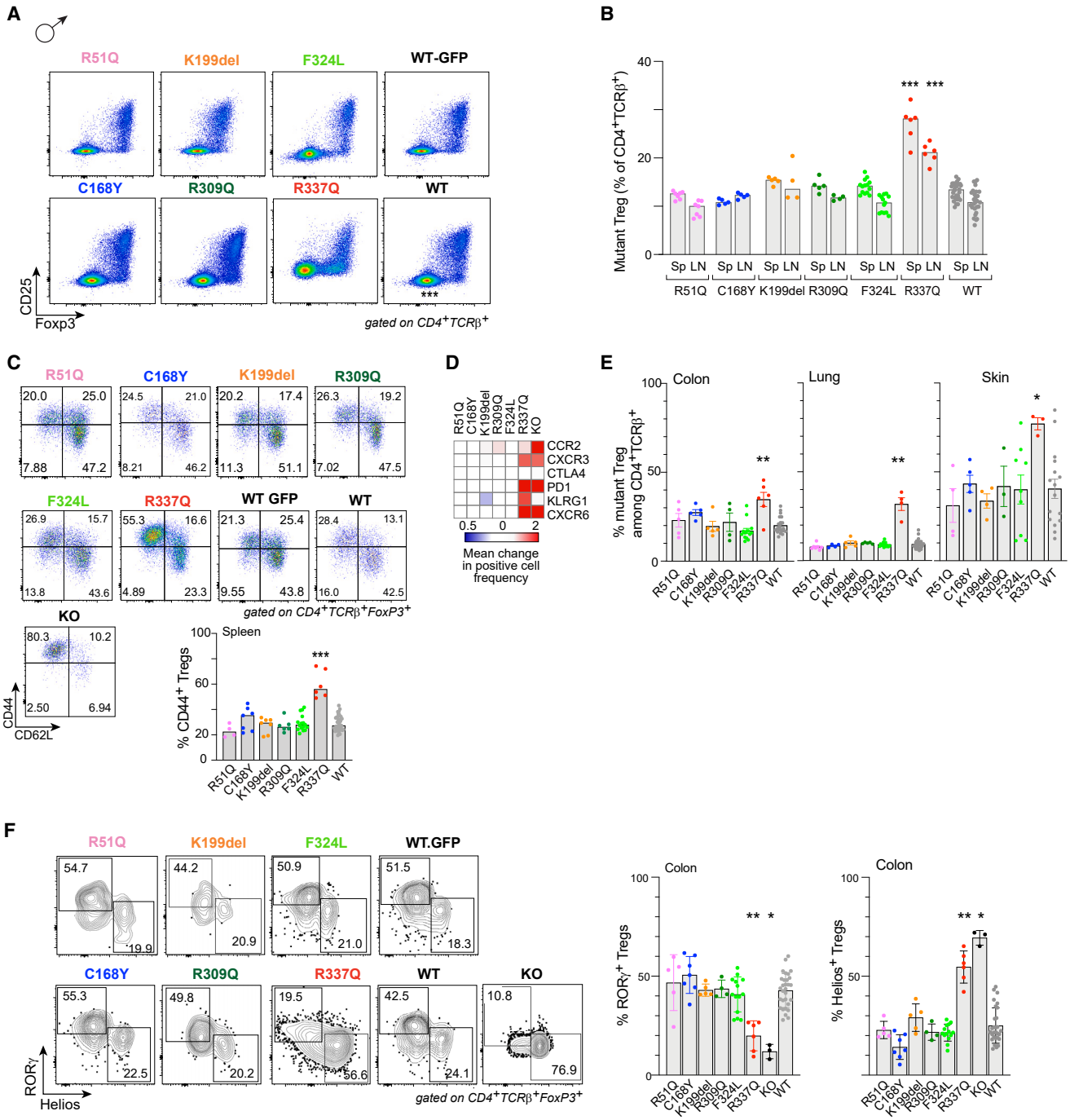


Figure 3. Treg phenotypes in hemizygous males

(A) Representative CD25/FoxP3 plots of gated CD4$^+$TCRB$^+$Thy1.1$^-$ T splenocytes from hemizygous males; the two WT littermates shown are from B6 and B6.*Foxp3-ires-gfp* backgrounds.

(B) Proportions of Tregs expressing the mutant FoxP3 among CD4$^+$TCRβ$^+$ cells in spleen from hemizygous males.

(C) Representative CD44/CD62L plots of splenic Tregs from the different mutants and two WT control littermates (quantified at bottom).

(D) Cumulative heatmap of the change in proportion of different markers in the splenic Treg from the mutants, each normalized to its WT littermate.

(E) Proportions of Tregs expressing the mutant FoxP3 in different tissues.

(F) Representative RORg/Helios plots of colonic Tregs expressing the different mutant FoxP3s or from WT control littermates (quantification from multiple experiments at right).

All results are from 2–3 independent experiments; each dot is an individual mouse; error bars indicate mean ± SD; *p < 0.05, **p < 0.01, ***p < 0.0001, from Mann-Whitney test.

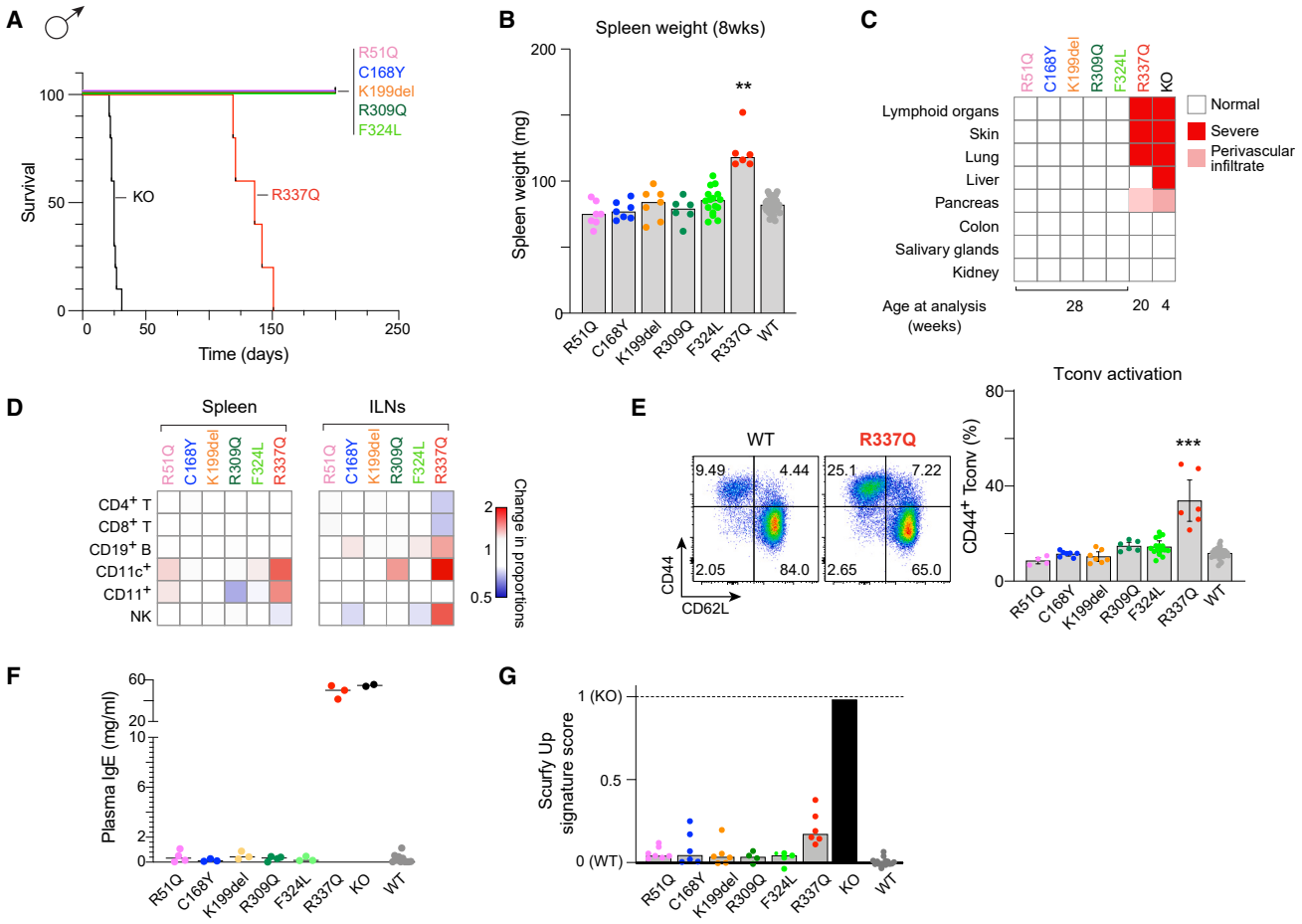


Figure 4. Autoimmune disease manifestations in hemizygous mutant males at baseline

(A) Survival curve of mutant males.

(B) Spleen weight at 8 weeks of age.

(C) Representative heatmap of the pathological changes in male mutants, at 28 weeks of age for all mice except KO (4 weeks) and R337Q (20 weeks).

(D) Heatmap of averaged changes in immunocyte proportions in spleen and inguinal lymph nodes (ILNs) for each mutant line relative to its WT littermates.

(E) Representative flow cytometry CD44/CD62L plots from R337Q and WT littermate Tregs. (quantification from multiple experiments at right).

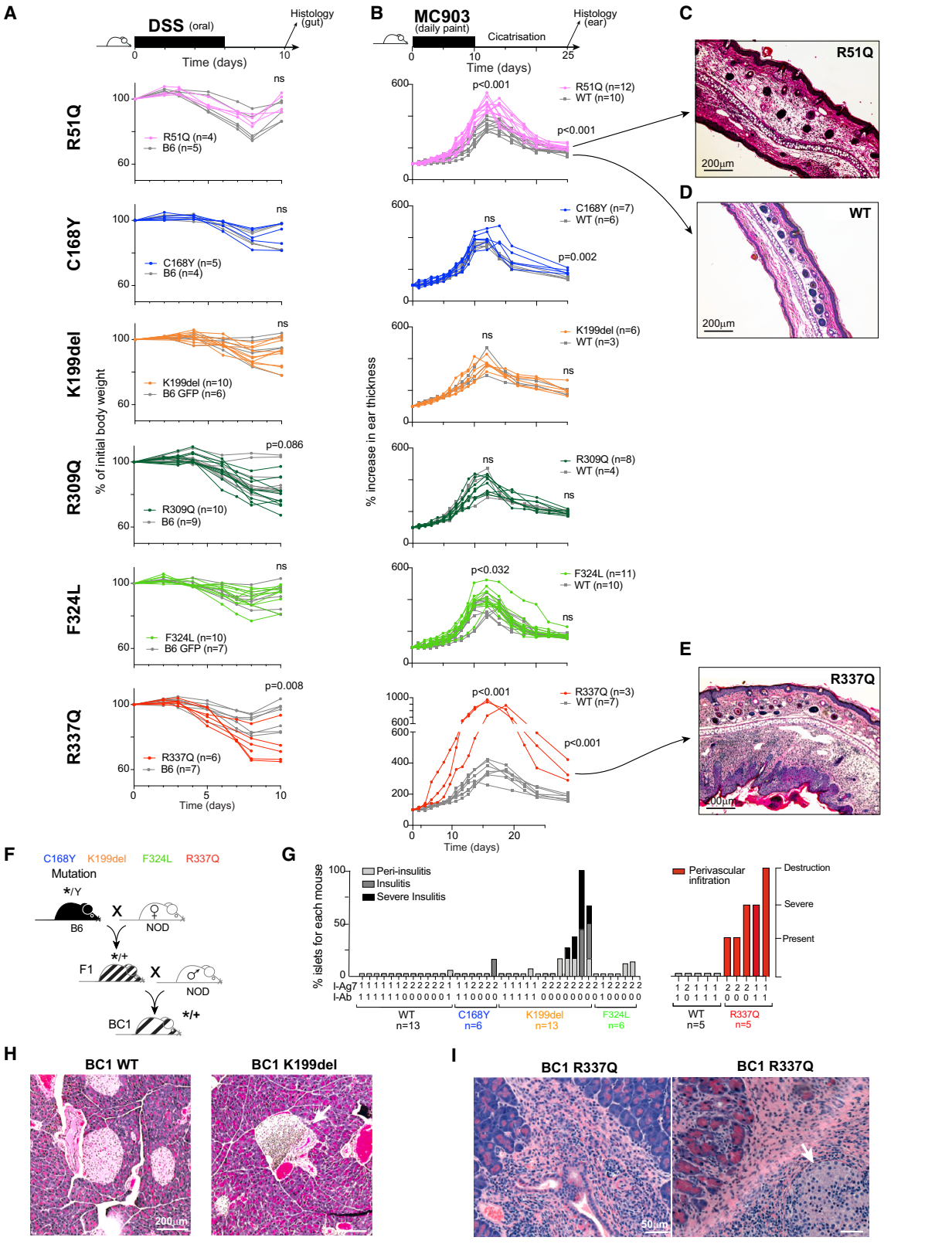
(F) Plasma IgE concentration, quantified by ELISA.

(G) Scurfy-Up transcriptomic score among resting splenic Tregs (scaled from WT Treg = 0 and *FoxP3* KO Treg = 1).

unchallenged mice. None of the mutant mice showed the rapidly lethal wasting disease of mice with full FoxP3 deficiency (Figure 4A), and most lines thrived during a 6 month follow-up (Figure S4A). However, R337Q began to develop skin lesions around 10–12 weeks of age, associated with growth stagnation, leading to death around 22 weeks of age (Figures 4A and S4A). Lymphoproliferation was already present at 8 weeks in R337Q males with a splenomegaly and diffuse adenopathies (Figure 4B) and was far more pronounced at 20 weeks (Figure S4B). In older mice, while severe leukocyte infiltrates were found in the skin and lungs of R337Q mice, but not in the liver (Figures 4C and S4C), all the other mutants were free of inflammation in all parenchymal tissues screened (skin, lungs, colon, liver, pancreas, kidney, salivary gland; Figure 4C).

Dysregulation of other immunocyte populations is a hallmark of Treg dysfunction.[44] We thus examined other cell types in lymphoid organs in these mice. Only R337Q showed changes: proportions of dendritic, myeloid, and natural killer (NK) cells were increased in the spleen and lymph nodes (LNs) (Figure 4D), along with a stronger activation of Tconvs (Figure 4E). Additionally, R337Q showed elevated plasma immunoglobulin E (IgE) levels (Figure 4F), as has been described in most patients with IPEX and in Treg-deficient mice[5,30,32,45–47] (Figure 4F). Accordingly, high frequencies of IL-4-producing cells were observed in R337Q and KO mice (Figure S4D).

In FOXP3-deficient males (human or mouse), a characteristic "IPEX signature" or "scurfy signature" is present in the transcriptome of all CD4+ T cells and Tconvs as well as Treg-like cells, which reflects the immunologic dysregulation and/or the homeostatic drive that attempts to restore Treg function.[3] As another measure of disease, we generated RNA-seq profiles from Treg-like cells in mutant males, and a biased expression of this signature was observed in R337Q Tregs, albeit not as marked as in KO Tregs, but not for any of the other missense mutations



(legend on next page)

(Figures 4G and S4E). This bias was concordant with the overall pathology of the mutants.

Pathology revealed by immunologic challenges

We continued to be surprised by the paucity of effects resulting from these mutations, with the exception of R337Q, even though they had been identified in patients with IPEX with clear disease. Tregs in these mice appeared in normal numbers and of normal phenotypes and maintained self-tolerance at baseline. This discrepancy could be due to environmental triggers and/or genetic cofactors revealing Treg dysfunction in patients with IPEX given the variability seen in patients bearing the same *FOXP3* mutation. Accordingly, we exposed 8-week-old mutant males to various challenges related to the IPEX pathological triad, enteropathy with dextran sodium sulfate (DSS) colitis,[48] eczema with the MC903 model of atopic dermatitis,[49] and type 1 diabetes (T1D), by introducing susceptibility variants from diabetes-prone NOD/ShyLtJ mice,[50,51] all being models known to be affected by Tregs.

DSS colitis induces a strong Treg response, and its course is impacted by Treg fitness.[52,53] We followed the standard protocol (6 day DSS treatment, 4 day recovery; Figure 5A). R337Q males exhibited a stronger response and retained pathology at day 10. Other mutants responded as the controls, except for an inconstant trend in R309Q.

Second, we induced atopic dermatitis-like lesions in the ears with topical MC903, a vitamin D analog.[49] R51Q males showed significantly increased responses, a stronger acute phase, and persistent inflammation (Figure 5B), with residual ulceration and acanthosis in the ears at day 25 (Figure 5C). Control littermates were essentially clear of sequalae (Figure 5D). As in the DSS model, responses to MC903 were explosive in R337Q males (Figure 5B), with uncleared inflammation 2 weeks later (Figure 5E). None of the mutant ears displayed fibrosis at day 25 (Figures 5C–5E and S5).

Diabetes in NOD mice is controlled dominantly by alleles of major histocompatibility complex (MHC) class II molecules (H2-A^{g7}) and by a collection of other loci that collectively contribute to dysfunctional tolerance.[51] NOD males are also far less susceptible to T1D than females. We reasoned that focal defects in FoxP3 and Treg function might be revealed in the context of a partial complement of NOD susceptibility alleles that would be insufficient alone to engender autoimmune diabetes. Males from intercrosses or first-generation backcrosses between non-susceptible inbred strains (like our B6 mutants) and the NOD strain never present with diabetes, and show only sporadic and tardy insulitis,[54,55] because genetic susceptibility alleles needs several generations to be established. Thus, we bred first-generation backcross (BC1) mice onto NOD from four mutant lines (Figure 5F), monitoring these BC1 for diabetes and scoring insulitis at 18–20 weeks of age. No overt diabetes was observed except in one of the R337Q BC1 mice. Clear peri-insulitis and insulitis, which resembled insulitis in NOD mice but spread to many islets, was observed in K199del BC1s (Figures 5G and 5H) but not in the other mutant or control mice. Insulitis in these mice was "respectful"[56] in that it coexisted with healthy-looking beta cells in the same islets (Figure 5H). Importantly, no insulitis was observed in K199del BC1 mice that carried one copy of the B6-derived H2-Ab allele, which is normally protective,[57] indicating that autoimmunity in K199del males was still under MHC control. Backcrossed R337Q mice showed a very different phenotype: there was no insulitis, but perivascular/ductal spaces showed extensive vasculitis that damaged exocrine tissue (Figures 5G and 5I; with complete destruction in some cases) but respected the islets (Figure 5I, right panel). This peri-vascular infiltrate was similar to pancreatic infiltration observed in full *Foxp3* KO mice and was independent of MHC control, as it occurred equally in $A^{g7/g7}$ and $A^{g7/b}$ mice (Figure 5G).

Thus, challenging the immune system revealed covert susceptibilities in FoxP3 mutant mice. This susceptibility was revealed in different locations as a function of the particular mutation (skin for R51Q, islets for K199del), the latter influenced by the MHC and autoimmune susceptibility genes, unlike the *scurfy*-like pancreatic vasculitis of R337Q mice.

Mutation-associated transcriptional changes

Given the range and specificity of phenotypes resulting from the different FoxP3 mutations, we sought to identify transcriptional modules specifically affected by these mutations that would give us clues about underlying mechanisms of pathogenesis. We analyzed heterozygous females in order to determine the intrinsic effects of each mutation and purified separately rTregs and aTregs for population RNA-seq to differentially analyze the impact of FoxP3 in the two states (Figure 6A). Each mutation was analyzed (in 2–4 biological replicates) in comparison with sex- and background-matched WT littermates.

Most of the mutations induced very few transcriptomic changes relative to WT Tregs, with 15–77 differential transcripts

Figure 5. Genetic or inflammatory challenges reveal hidden phenotypes in FoxP3 mutant males

(A) Induced colitis: weight loss in DSS-treated mutant males and their control littermates.

(B) Induced dermatitis: ear thickness in MC903-treated mutant males and their control littermates.

(C–E) Representative images of H&E staining of ear sections from MC903-treated mice at the resolution phase of dermatitis (day 25): R51Q (C), WT (D), and R337Q (E).

(F) Breeding of first backcross (BC1) onto the NOD background.

(G) Pancreatic autoimmunity. Left: insulitis score in hemizygous BC1 male pancreata of the 3 mutant lines shown (0: normal; 1: peri-insulitis; 2; insulitis; 3: severe insulitis); each bar is an individual mouse (>15 islets scored/mouse), with WT littermates of the three lines combined for simplicity. Right: scoring of perivascular infiltrates in R337Q BC1 pancreata (present: several perivascular infiltrates, intact exocrine acini; severe: abundant perivascular infiltrates, with destruction of surrounding acini in some regions, intact islets; destruction: quasi-complete destruction of the exocrine pancreas).

(H) Representative severe insulitis (arrow) in a K199del BC1 pancreas and normal islet from a WT pancreas.

(I) Representative perivascular infiltrates R337Q BC1 pancreata; note how the severe inflammation at right dilacerates the acinar tissue but leaves the islet (arrow) comparatively intact.

All results are from hemizygous males and from 2–5 independent experiments; p values from Mann-Whitney test.

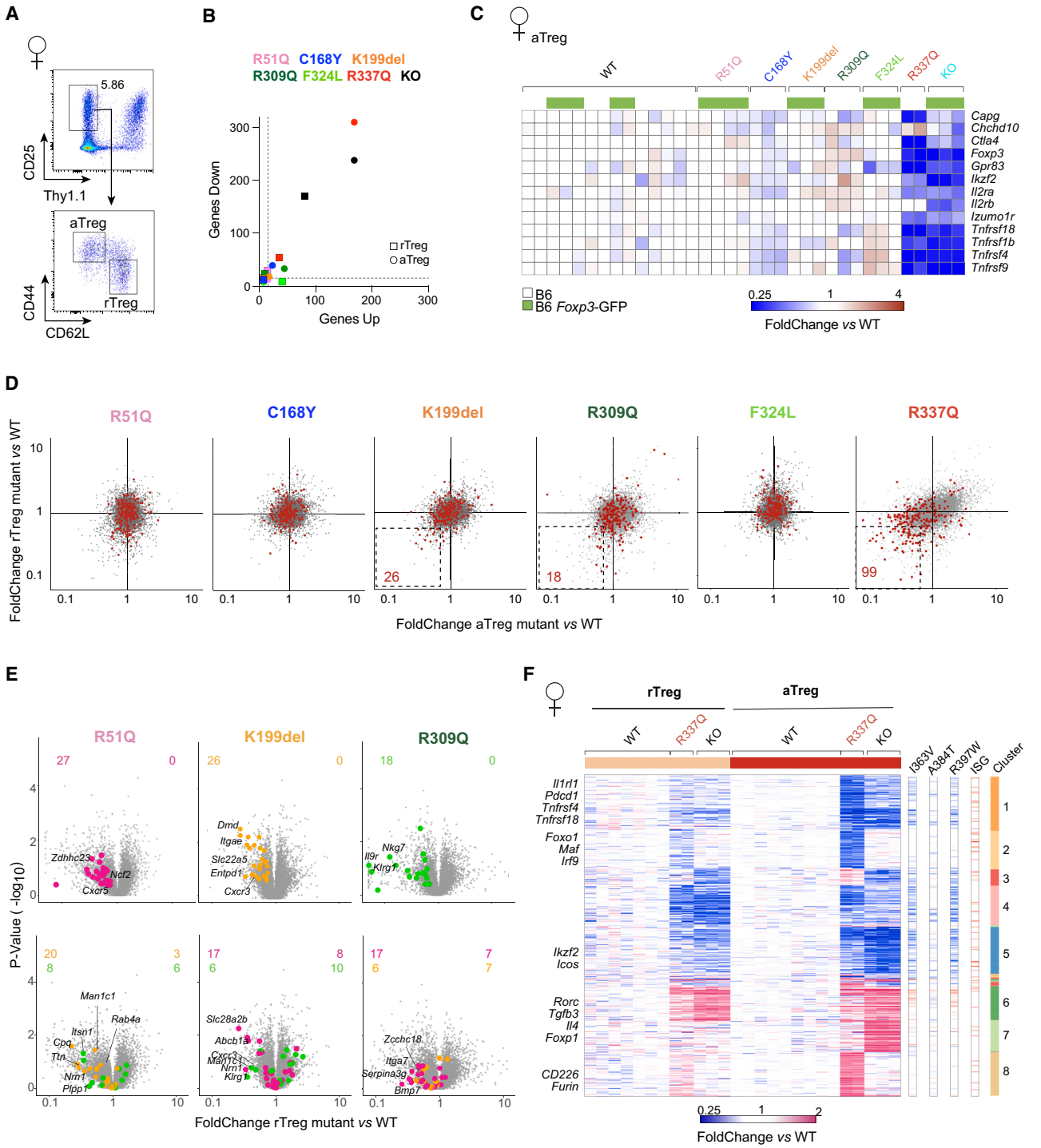


Figure 6. Transcriptional changes resulting from the mutations

Low-input RNA-seq was performed on Tregs expressing the FoxP3 mutants in heterozygous females and matching WT littermates.

(A) Representative gating to isolate Treg (Thy1.1-CD25^{high}) from mutant or WT littermates. Bottom: gating of rTreg and aTreg subsets.

(B) Numbers of differential expressed genes (at arbitrary thresholds of fold change <0.5 or >2, and t test $p < 0.01$) in aTreg (round) or rTreg (square) in the various mutants vs. averaged values from their matched WT littermates. Experimental noise in these datasets was estimated by recomputing apparent differentially expressed genes (DEGs) after permuting the data labels.

(C) Heatmap of the mutant/WT fold change, for core Treg genes (defined in Zemmour et al.[35]), calculated for each donor mouse against the averaged expression in all WT Tregs from matching backgrounds (B6 or B6.*Foxp3-ires-gfp*) (indicated on top); each column is an independent mouse.

(legend continued on next page)

(at arbitrary thresholds of 2-fold and t test nominal $p < 0.01$; Figure 6B; corresponding volcano plots are shown in Figure S6A). These values were not much higher than the noise estimated by permutation analysis (Figure 6B). The exception was R337Q, with 578 differentially expressed genes in the aTreg dataset. Lists of FoxP3-dependent genes have been defined, mostly from comparative profiling of Tregs in heterozygous FoxP3-deficient females,[3,19,20] including a narrow "core Treg signature" of genes that appear to be intrinsic targets of FoxP3.[3,35] There was very limited impact for R51Q, C168Y, K199del, R309Q, and F324L on these core Treg transcripts (Figure 6C), while R337Q aTregs showed an almost total downregulation, with the exception of *Chchd10* and *Il2rb*, both of which were well downregulated in the KO control (Figure 6C).

These results denoted a very narrow and mild imprint of most of the Foxp3 mutations. Thus, we plotted together the changes observed in rTregs and aTregs, reasoning that concordance in both datasets would increase the reliability and power to detect subtle effects. We also focused on an external list of *Foxp3*-dependent genes[20] (Figure 6D). As a proof of principle, most downregulations in R337Q Tregs were clearly present in both Treg states, albeit with quantitative differences (Figure 6D, right). K199del and R309Q showed a few genes with reproducible downregulation, although the FoxP3-dependent signature as a whole was not significantly shifted. No trend was noted for C168Y and F324L, while R51Q induced a downregulation of part of the signature, mostly in rTregs. To substantiate the latter, we generated data from other R51Q mice that derived from an independent founder line, which confirmed this subtle downregulation in R51Q rTregs (Figure S6B). No such signal was found in aTregs of either R51Q line. We then asked if these subtle changes in R51Q ($n = 27$), K199del ($n = 26$), and R309Q ($n = 18$) were specific for each of the N-terminal mutations or were shared between the three non-FKHD mutations (Figure 6E; Table S3). There was some sharing of the genes affected by the N-terminal mutations R51Q and K199del, whereas R309Q had its own footprint. As expected given the small number of genes affected, Gene Ontology and pathway analysis did not reveal any significant enrichment, nor could we find clues explaining the phenotypic specificity. There was also no significant overlap with the changes observed in the gene expression profiles from the original patients with IPEX (not unexpected given the weak changes in mice and $n = 1$ patient). We also searched genome-wide association study (GWAS) databases for loci implicated in human autoimmune diseases. Although not a statistically significant overlap, two of the reproducible targets revealed in the R51Q mutant (*NCF2* and *ZDHHC23*) had been associated with autoimmune diseases with strong skin manifestations, systemic lupus erythematosus, and psoriasis, respectively (Table S3).

As evidenced in Figure 6C, the impact of R337Q on Treg core signature genes was similar but not identical to that of a complete FoxP3 LOF. Broader investigation (Figures 6F and S6C) showed that many of the changes seen in Treg-like "wannabes" from KO females were also present in R337Q Tregs, albeit milder (e.g., clusters 5, 6, and 7 in Figure 6F). In addition, two gene clusters were uniquely induced or repressed in R337Q but not in KO aTregs (clusters 2 and 8 in Figure 6F; Table S4). To further assess the specificity of these R337Q effects, we retrieved gene expression profiles from previous studies of FKHD mutations in mice.[30] The R337Q-specific clusters were not observed with these mutations either (Figure 6F, right), nor with other mutations in our present panel (Figure S6D). Interferon-response pathway and nuclear factor κB (NF-κB) signaling genes were specifically enriched in cluster 2. The R337Q mutation hinders the recently discovered head-to-head dimerization of FoxP3,[33] and we speculate that the consequences of R337Q that are not shared with complete FoxP3 LOF or other FKHD mutants may stem from its ability to form some, but not other, DNA-binding conformations.

Finally, we compared the effect of the mutations at 8 weeks of age in female and male Tregs (the latter without competition from WT Tregs but in a partially destabilizing milieu). We focused on R337Q since the consequences of the other mutations were too subtle to validly allow such a comparison. In rTegs, the effects were very similar (Figure S6E), but in aTregs, the comparison proved more complex: some transcripts were equally affected in both contexts, but many were affected only in heterozygous female aTregs. These disparities likely underlie the paradoxically opposite phenotypes of R337Q-expresssing Tregs in hemizygous males and heterozygous females noted in Figures 2 and 3.

Effects of FoxP3 mutations at the chromatin level

The mutant mouse lines thus seemed to identify two classes of mutations from patients with IPEX. R337Q yielded a "*scurfy*-like" phenotype, with transcriptional changes and immunological manifestations that resemble those of full LOF or mutations that strongly interfere with DNA binding[29,30,32,58] and others, exemplified by R51Q or K199del, whose transcriptional consequences were more discrete, were mostly distinct, and led to focused Treg defects different from those of scurfy disease. Because chromatin structure can yield a more robust landscape of gene regulatory programs than RNA-seq (i.e., not affected by transcriptional bursting, mRNA stability), we performed single-cell assay for transposase-accessible chromatin sequencing (scATAC-seq) on purified Tregs (Figures S7A and S7B). We focused on three mutants that exemplified these classes, sorting splenic Tregs from heterozygous females (and their corresponding WT littermates)

(D) Comparison of mutant/WT fold changes in aTreg (x axis) versus rTreg (y axis), for each mutant, highlighted with FoxP3-dependent genes (genes downregulated in the absence of FoxP3 in van der Veeken et al.[20]).
(E) Fold change vs. p value plots comparing normalized expression in mutant Tregs from heterozygous females with WT Tregs from the same background (B6 or B6.*Foxp3-ires-gfp*). Highlights correspond to the "concordant" FoxP3-dependent genes between rTreg and aTreg comparisons (for K199del or R309Q, defined in D) or between the two independent founder lines for R51Q (per Figure S6B).
(F) Heatmap of transcripts differentially expressed in R337Q and KO relative to matched littermates (selected on arbitrary threshold as in A) shown is expression is normalized to the mean of matched control littermates; each column is an independent mouse. Overlap with DEGs described in previous FoxP3 mutations,[30] and interferon-sensitive genes are shown at right.

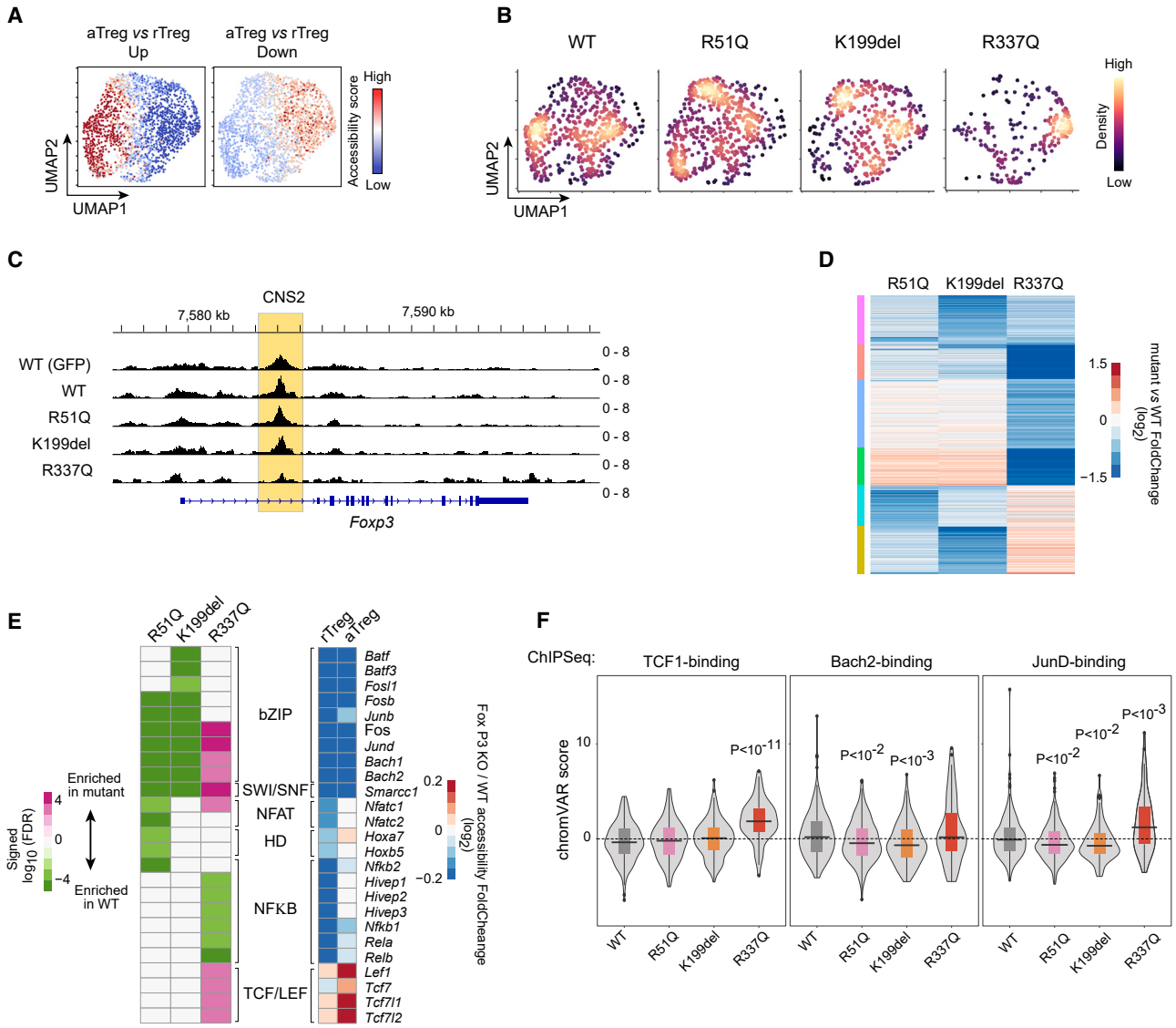


Figure 7. Changes in chromatin architecture in Tregs from heterozygous mutant females

Single-cell ATAC-seq was performed in multiplex mode on Tregs from heterozygous mutant females and their WT littermates.

(A) UMAP representation of the ATAC signal, grouping cells from all mice. The relative accessibility (chromVAR scores) in each cell was computed for OCRs that distinguish aTreg and rTreg populations.

(B) Density of Tregs from WT and FoxP3 mutant mice, overlaid onto the same UMAP as in (A).

(C) Aggregated accessibility profiles at the *Foxp3* locus for Tregs expressing the mutant or WT FoxP3; *Foxp3* CNS2 region is highlighted.

(D) Differential accessibility: heatmap of mutant/WT fold change (log2) in aggregated chromatin accessibility of Tregs from each mutant vs. its respective WT littermate (ordered by k-means).

(E) TF involvement: TF binding motifs enriched (at FDR < 1×10^{-3}) in OCRs differentially accessible in mutant vs. WT aTregs (pink and green, enriched in mutant or WT, respectively). Right: for reference, the log2 fold change in accessibility KO Tregs vs. WT for OCRs that contain the corresponding TF motifs (data from Chowdhary et al.[43]), split between rTregs and aTregs.

(F) Distributions of per-cell relative accessibility (chromVAR score) of OCRs within genomic regions that bind TCF1, Bach2, or JunD determined by ChIP-seq (data from van der Veeken et al.[20] and Grant et al.[63]) among aTregs from each genotype. Nominal t test p values.

and hash tagging cells from all genotypes in the same run[59] for optimal comparability. Relative accessibility (chromVAR score[60]) of open chromatin regions (OCRs) that characterize Treg states[43] readily distinguished aTregs and rTregs on a 2D uniform manifold approximation and projection (UMAP) visualization of the combined data (Figure 7A). Tregs from each genotype occupied distinct regions of the UMAP space (Figure 7B). Whereas R51Q and K199del mutant cells largely comingled with WT cells, albeit with different preferential densities, R337Q Tregs shifted away from other WT and mutant cells (Figures 7B and S7C). Most of these R337Q cells were in a resting-like chromatin state, consistent with the flow cytometry (Figure S2). This dichotomy was

also observed at the *Foxp3* locus itself. While K199 and R51Q Tregs showed no notable alterations, R337Q Tregs had decreased accessibility at CNS2, a *cis*-regulatory element bound by FoxP3 itself, locking in the stability of *Foxp3* expression[61,62] (Figure 7C). We confirmed this change in accessibility at CNS2 by bulk ATAC-seq (Figure S7D). Thus, R337Q, but not the N-terminal mutations R51Q and K199del, impaired epigenetic feedback that controls *Foxp3* expression. The differences in these two classes of mutants extended to a global level: a small subset of differential OCRs showed consistently decreased accessibility across all mutants, but OCRs with increased accessibility in R337Q cells had opposite effects in K199 and R51Q cells, and vice versa (Figure 7D).

FoxP3 collaborates with other TFs via direct protein-protein interaction and colocalization at regulatory regions.[18,64,65] How did each FoxP3 mutation affect the activity of other TFs? To address this question, we searched for TF binding motifs enriched (at FDR $<1 \times 10^{-3}$) in OCRs differentially accessible in mutant vs. WT Tregs. To avoid confounding effects resulting from changes in cell population abundance, we stratified comparisons between WT and mutant cells by cell state (rTregs or aTregs, per Figure 7A). The results portrayed very different effects of the mutations (Figure 7E, left). A number of OCRs were more accessible in R337Q mutant aTregs relative to WT aTregs. These were enriched in motifs that bind key Treg regulators, including those of the TCF/LEF and bZIP families, while OCRs with decreased accessibility in R337Q relative to WT aTregs were enriched for NF-κB motifs (Figure 7E, left panel; for reference, the average change in accessibility of OCRs that contain these motifs in fully FoxP3-deficient Tregs[43] are shown at right). In contrast, OCRs whose accessibility was reduced by R51Q and K199del mutations were enriched in a distinct set of motifs, partially overlapping between the two (in particular, bZIP-binding motifs for Bach or Fos and also NF-AT motifs). There were very few changes among rTregs, with only increased TCF/LEF motif enrichment among R337Q mutants, consistent with the aTreg effects. While TCF/LEF motif-containing sites increased in accessibility in both the R337Q mutant and full KO, bZIP motif-containing sites increased in accessibility among R337Q Tregs but had diminished accessibility in the full KO (Figure 7E). While only a subset of full KO changes was present in the non-FKHD R51Q and K199del mutations, those that were present were consistent with the direction of effect elicited by complete FoxP3 deficiency (Figure 7E). To orthogonally validate these findings using *bona fide* TF binding data, we examined the per-cell relative accessibility (chromVAR score) of OCRs known to bind TCF1, Bach2, and JunD from independent chromatin immunoprecipitation (ChIP)-seq data[20,63] and containing the corresponding motifs. Consistent with the motif-enrichment results, R337Q Tregs showed increased accessibility at JunD- and TCF1-bound sites; the latter likely corresponded to increased expression of its mRNA (Figure S7E). These effects are consistent with the notion that FoxP3 represses TCF-1 expression, accounting for the indirect part of its action.[20] On the other hand, R51Q and K199del cells had decreased accessibility at Bach2- and JunD-bound sites (Figure 7F). Thus, the divergence between mutations inside vs. outside the FKHD extended to global differences in the activities of FoxP3 partner TFs.

DISCUSSION

This porting of mutations from patients with IPEX into mice, allowing comparative analyses in robust numbers (beyond N = 1 in patients) and with control of genetic and immunologic variables, uncovered an unexpectedly multifaceted aspect of IPEX-causing *FOXP3* mutations and thus of FoxP3 function. It exposed the distinction between two classes of IPEX-causing mutations, that different mutations elicited preferential effects on chromatin that involved different sets of cofactors, and that immunologic or genetic perturbations were necessary to manifest the specific disease facets tied to each mutation. The comparison of Treg-like cells that express the same mutant FoxP3 protein in mutant males and carrier females also showed that the behavior varied strikingly as a function of the organismal environment, emphasizing the importance of cell-extrinsic regulatory loops in modulating Treg function.

This panel distinguishes two classes of *FOXP3* mutations in which we can also integrate the isolated mutations analyzed previously by others and ourselves. In the first class, the FKHD mutant R337Q echoed strongly the phenotypes already described in various missense FKHD mutations in mice (A384T, M370I).[30,32] Their phenotypes appeared like a "slowed-down" version of the full *scurfy* LOF phenotype: a number of manifestations appeared in unperturbed mice, including skin and lung inflammation, activation of many immunocytes, and dysregulated IgE levels. The R337Q mutation also reproduced most of the transcriptional alterations of full LOF cells (albeit in a muted fashion), with a reshuffling of the Treg chromatin architecture, the loss of repression by TCF1, and diminished positive feedback by FoxP3 at the CNS2 element of the *Foxp3* locus. Although the genomic studies were not as extensive in previous studies, many of the phenotypic traits were the same as those previously reported in mice with partial LOF mutations in the FKHD (A384T, M370I[30,32]), in particular the hyper-IgE and IL-4 over-expression, which we had also noted *in vitro* with strong FKHD mutations.[18] The second class encompasses mutations in other domains of FoxP3, i.e., are not involved in DNA binding (at least not as directly). These mutations preserved the general distribution of chromatin architecture across Treg diversity, maintained the ability of FoxP3 to repress TCF-1 and LEF1, and only induced very subtle transcriptional changes; affected males showed essentially no alterations in baseline immunoregulation, with no Th2 deviation or hyper-IgE production. The chromatin accessibility data suggest that mutations of the second class impacted FoxP3-controlled chromatin regions that involved a different segment of the cofactors that partake in the Treg regulatory network. For instance, they diminished accessibility at Bach2- and JunD-binding sites, which were actually boosted by the R337Q mutation. This dichotomy concords with results from clinical studies, in which, generally speaking, FKHD mutations are generally more severe.

Importantly, the full physiological importance and specificity of these mutations was revealed only by additional perturbations: inflammatory for R51Q and R309Q (and R337Q) and genetic for K199del. These observations bolster the hypothesis that the heterogeneity of onset age and clinical phenotypes displayed by patients bearing the same mutation results from

genetic or immunologic covariates.[26,66–68] In these contexts, the facets of autoimmune dysfunction also differed according to the mutation: exacerbation of autoimmune diabetes for K199del or of skin inflammation for R51Q. This diversity, which contrasts with the more monomorphic consequences of FKHD mutations, evokes the phenotype of a mouse with a GFP fusion into the N terminus of FoxP3, which protected from arthritis but accelerated autoimmune diabetes.[69,70] Mechanistically, insertion of the bulky GFP domain disrupted a number of interactions by FoxP3 (TIP-60, Ikzf4, HIF-1α, Irf4), thus causing the bias Treg function.[69,70] In the same vein, we suggest that the R51Q and K199del mutations affect particular interactions in which FoxP3 engages, and those conditioned the differential Treg dysfunction in the present mice. It is interesting that the transcriptional consequences of R51Q were most marked in rTregs, while those of K199del were equally seen in rTegs and aTregs. Unfortunately, the transcriptional differences between R51Q- and K199del-expressing Tregs proved too subtle to rigorously nominate the candidate transcripts and functions that underpin these outcomes.

The diametrically opposed phenotypes of Tregs expressing R337Q in male and female settings were quite striking. This outcome very likely reflected the balance of cell-extrinsic and -intrinsic consequences of the mutations observed in cells from patients with IPEX and their heterozygous female relatives,[3] and in earlier observations in FoxP3-deficient mice,[3,41] but the more standardized setting in mice made the comparison more striking. In females, R337Q Tregs were sharply reduced in numbers and showed low FoxP3 and CD25 expression, a block in aTreg maturation, and poor ability to colonize non-lymphoid tissues. In males, R337Q Tregs were present in increased numbers, with substantially restored FoxP3 and CD25, increased cell activation, and increased representation in tissues (also with a complete swap in the relative representation of RORg$^+$ and Helios$^+$ subsets of colonic Tregs, in line with the demonstration by us and the Rudensky lab that these subsets have different dependence on FoxP3 function[43,71]). Two core differences between the male and female contexts may explain the divergent outcomes: (1) the presence in heterozygous females of normal Tregs expressing a WT FoxP3, which can outcompete the mutant Tregs for homeostatic niches and trophic factors (in particular IL-2,[3] with a vicious circle where lower access to IL-2 decreased CD25 expression, further reducing IL-2 capture by mutant Tregs), and (2) in males, a strong homeostatic drive that aims to restore Treg function, akin to the strong proliferation observed after acute depletion,[72,73] likely helps to restore mutant Tregs and their maturation and tissue-homing potential (IL-2 produced by dendritic cells or other T cells may be involved here). Is this homeostatic drive and "phenotypic reversion" why males with the mutations are comparatively well off, akin to the aTregs observed in patients?[3] One might also speculate that true sex-specific differences in immune function are at play. Either way, these observations reinforce the notion that the impact of FoxP3 on Treg function is highly context dependent.

How did the mice reproduce the human disease? Some observations did fit: C168Y and F324L showed essentially no phenotype at baseline or when challenged, and we were unable to define any immunologic or transcriptional Treg defect. Patients with these mutations had the mildest symptoms, with late onset and relatively well-tolerated disease. Overall, though, the disease in mice tended to be less severe than in patients, all of which did show the characteristic diagnostic triad. One reason may simply be time (we followed the mice for up to 40 weeks only, shorter than the age at diagnosis for several patients). It is also conceivable that B6 mice are generally more refractory to autoimmunity, and it was interesting that the pancreatic vasculitis of R337Q mice was far more pronounced on the NOD BC1 than B6 backgrounds. The mutation-specific propensity to additional autoimmune diseases like diabetes did not fit (K199del led to pancreatic autoimmunity in mice, but the corresponding patient was not diabetic; the R337Q patient was diabetic). Here, the difference plausibly rests in the complement of autoimmunity susceptibility loci present in the index patients—importantly, islet-specific autoimmunity in K199del BC1 mice remained under the control of MHC alleles, demonstrating the epistasis between *FOXP3* mutations and genetic susceptibility variants.

Limitation of study

The molecular mechanism through which each mutation triggers a different phenotype remains unknown, and we were not able to pinpoint specific transcriptional changes in Tregs induced by the non-FKHR mutations that would connect to the specific autoimmunity they elicit. This may be because the effects are buffered by network regulation and/or only become manifest after challenge (then parsing cause and consequence is difficult). The connection between disease manifestations in humans and mice was not perfect (e.g., most patients presented with enteropathy), possibly due to species differences or other genetic variants at play in these patients.

In conclusion, this work has uncovered unexpected nuances concerning the *in vivo* import of FoxP3's domains in Tregs and how these partake in the interplay between Treg deficiencies, genetic variation, and immunologic challenges that conspire to determine disease manifestations in each patient with IPEX.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - IPEX human mutations
  - Mice
  - Cell lines
- METHOD DETAILS
  - Mutant mice generation by CRISPR/Cas9
  - Inflammatory challenges
  - Histology
  - ELISA IgE
  - Plasmids

- FoxP3-DNA pulldown assay
- Isolation of lymphocytes from lymphoid and non-lymphoid tissues
- Flow cytometry
- Cell sorting for bulk RNAseq and scATACseq
- Bulk RNA-seq library preparation
- Bulk RNA-seq pre-processing and analysis
- scATAC-seq library preparation
- scATAC hashtagging (ASAP-seq)
- scATAC-seq preprocessing and visualization
- scATAC-seq analysis
- Peak sets
- Motif accessibility analysis
- Gene scores
- Pseudobulk track visualization
- Differential accessibility
- Motif enrichment
- OCR signature relative accessibility

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.113018.

ACKNOWLEDGMENTS

We are grateful to L. Du, Dr. R. Bronson, J. Lee, I. Magill, A. Ciulla, N. Patel, L. Yang, M. Sleeper, C. Araneo, J. Nelson, K. Hattori, V. Piekarsa, and A. Ortiz-Lopez for help with breeding and experimentation and Drs. T. Chatila, L.M. Charbonnier, B. Salomon, and F. Rieux-Laucat for insightful discussions. This work was funded by NIH grants AI-165697 and AI-150686 to C.B. and D.M. and by the JPB Foundation to C.B. and D.M. S.H. is an HHMI investigator, and W.Z. was supported by a Modell fellowship, J.L. by an IMAGINE MD/PhD grant and an Arthur Sachs scholarship, and K.C. by NIGMS fellowships T32GM007753 and T32GM144273 and an HSCI MD/PhD Training Fellowship.

AUTHOR CONTRIBUTIONS

J.L., I.A., and C.B. conceptualized the study. J.L., K.C., W.Z., and R.N.R. performed and analyzed experiments. C.B., D.M., and S.H. supervised. J.L., K.C., and C.B. drafted the manuscript, and all authors edited it.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 27, 2023
Revised: July 26, 2023
Accepted: August 4, 2023

REFERENCES

1. Gavin, M.A., Torgerson, T.R., Houston, E., Deroos, P., Ho, W.Y., Stray-Pedersen, A., Ocheltree, E.L., Greenberg, P.D., Ochs, H.D., and Rudensky, A.Y. (2006). Single-cell analysis of normal and FOXP3-mutant human T cells: FOXP3 expression without regulatory T cell development. Proc. Natl. Acad. Sci. USA *103*, 6659–6664.
2. Di Nunzio, S., Cecconi, M., Passerini, L., McMurchy, A.N., Baron, U., Turbachova, I., Vignola, S., Valencic, E., Tommasini, A., Junker, A., et al. (2009). Wild-type *FOXP3* is selectively active in $CD4^{+}CD25^{hi}$ regulatory T cells of healthy female carriers of different *FOXP3* mutations. Blood *114*, 4138–4141.
3. Zemmour, D., Charbonnier, L.M., Leon, J., Six, E., Keles, S., Delville, M., Benamar, M., Baris, S., Zuber, J., Chen, K., et al. (2021). Single-cell analysis of FOXP3 deficiencies in humans and mice unmasks intrinsic and extrinsic $CD4^{+}$ T cell perturbations. Nat. Immunol. *22*, 607–619.
4. Barzaghi, F., Amaya Hernandez, L.C., Neven, B., Ricci, S., Kucuk, Z.Y., Bleesing, J.J., Nademi, Z., Slatter, M.A., Ulloa, E.R., Shcherbina, A., et al. (2018). Long-term follow-up of IPEX syndrome patients after different therapeutic strategies: An international multicenter retrospective study. J. Allergy Clin. Immunol. *141*, 1036–1049.e5.
5. Jamee, M., Zaki-Dizaji, M., Lo, B., Abolhassani, H., Aghamahdi, F., Mosavian, M., Nademi, Z., Mohammadi, H., Jadidi-Niaragh, F., Rojas, M., et al. (2020). Clinical, Immunological, and Genetic Features in Patients with Immune Dysregulation, Polyendocrinopathy, Enteropathy, X-linked (IPEX) and IPEX-like Syndrome. J. Allergy Clin. Immunol. Pract. *8*, 2747–2760.e7.
6. Sugimoto, N., Oida, T., Hirota, K., Nakamura, K., Nomura, T., Uchiyama, T., and Sakaguchi, S. (2006). Foxp3-dependent and -independent molecules specific for CD25+CD4+ natural regulatory T cells revealed by DNA microarray analysis. Int. Immunol. *18*, 1197–1209.
7. Hill, J.A., Feuerer, M., Tash, K., Haxhinasto, S., Perez, J., Melamed, R., Mathis, D., and Benoist, C. (2007). Foxp3 transcription-factor-dependent and -independent regulation of the regulatory T cell transcriptional signature. Immunity *27*, 786–800.
8. Lin, W., Haribhai, D., Relland, L.M., Truong, N., Carlson, M.R., Williams, C.B., and Chatila, T.A. (2007). Regulatory T cell development in the absence of functional Foxp3. Nat. Immunol. *8*, 359–368.
9. Ohkura, N., Hamaguchi, M., Morikawa, H., Sugimura, K., Tanaka, A., Ito, Y., Osaki, M., Tanaka, Y., Yamashita, R., Nakano, N., et al. (2012). T cell receptor stimulation-induced epigenetic changes and Foxp3 expression are independent and complementary events required for Treg cell development. Immunity *37*, 785–799.
10. Charbonnier, L.M., Cui, Y., Stephen-Victor, E., Harb, H., Lopez, D., Bleesing, J.J., Garcia-Lloret, M.I., Chen, K., Ozen, A., Carmeliet, P., et al. (2019). Functional reprogramming of regulatory T cells in the absence of Foxp3. Nat. Immunol. *20*, 1208–1219.
11. Schubert, L.A., Jeffery, E., Zhang, Y., Ramsdell, F., and Ziegler, S.F. (2001). Scurfin (FOXP3) acts as a repressor of transcription and regulates T cell activation. J. Biol. Chem. *276*, 37672–37679.
12. Bettelli, E., Dastrange, M., and Oukka, M. (2005). Foxp3 interacts with nuclear factor of activated T cells and NF-kappa B to repress cytokine gene expression and effector functions of T helper cells. Proc. Natl. Acad. Sci. USA *102*, 5138–5143.
13. Li, B., Samanta, A., Song, X., Iacono, K.T., Bembas, K., Tao, R., Basu, S., Riley, J.L., Hancock, W.W., Shen, Y., et al. (2007). FOXP3 interactions with histone acetyltransferase and class II histone deacetylases are required for repression. Proc. Natl. Acad. Sci. USA *104*, 4571–4576.
14. Wu, Y., Borde, M., Heissmeyer, V., Feuerer, M., Lapan, A.D., Stroud, J.C., Bates, D.L., Guo, L., Han, A., Ziegler, S.F., et al. (2006). FOXP3 controls regulatory T cell function through cooperation with NFAT. Cell *126*, 375–387.
15. Arvey, A., van der Veeken, J., Samstein, R.M., Feng, Y., Stamatoyannopoulos, J.A., and Rudensky, A.Y. (2014). Inflammation-induced repression of chromatin bound by the transcription factor Foxp3 in regulatory T cells. Nat. Immunol. *15*, 580–587.
16. Chen, C., Rowell, E.A., Thomas, R.M., Hancock, W.W., and Wells, A.D. (2006). Transcriptional regulation by Foxp3 is associated with direct promoter occupancy and modulation of histone acetylation. J. Biol. Chem. *281*, 36828–36834.
17. Ferraro, A., D'Alise, A.M., Raj, T., Asinovski, N., Phillips, R., Ergun, A., Replogle, J.M., Bernier, A., Laffel, L., Stranger, B.E., et al. (2014). Interindividual variation in human T regulatory cells. Proc. Natl. Acad. Sci. USA *111*, E1111–E1120.

18. Kwon, H.K., Chen, H.M., Mathis, D., and Benoist, C. (2017). Different molecular complexes that mediate transcriptional induction and repression by FoxP3. Nat. Immunol. *18*, 1238–1248.

19. Ramirez, R.N., Chowdhary, K., Leon, J., Mathis, D., and Benoist, C. (2022). FoxP3 associates with enhancer-promoter loops to regulate Treg-specific gene expression. Sci. Immunol. *7*, eabj9836.

20. van der Veeken, J., Glasner, A., Zhong, Y., Hu, W., Wang, Z.M., Bou-Puerto, R., Charbonnier, L.M., Chatila, T.A., Leslie, C.S., and Rudensky, A.Y. (2020). The transcription factor Foxp3 shapes regulatory T cell identity by tuning the activity of trans-acting intermediaries. Immunity *53*, 971–984.e5.

21. Ramirez, R.N., Chowdhary, K., and Benoist, C. (2021). Indirect regulation is not the dominant mode of action for the transcription factor FoxP3 in Treg cells. Submitted.

22. Powell, B.R., Buist, N.R., and Stenzel, P. (1982). An X-linked syndrome of diarrhea, polyendocrinopathy, and fatal infection in infancy. J. Pediatr. *100*, 731–737.

23. Wildin, R.S., and Freitas, A. (2005). IPEX and FOXP3: clinical and research perspectives. J. Autoimmun. *25*, 56–62.

24. Barzaghi, F., Passerini, L., and Bacchetta, R. (2012). Immune dysregulation, polyendocrinopathy, enteropathy, x-linked syndrome: a paradigm of immunodeficiency with autoimmunity. Front. Immunol. *3*, 211.

25. Duclaux-Loras, R., Charbit-Henrion, F., Neven, B., Nowak, J., Collardeau-Frachon, S., Malcus, C., Ray, P.F., Moshous, D., Beltrand, J., Goulet, O., et al. (2018). Clinical Heterogeneity of Immune Dysregulation, Polyendocrinopathy, Enteropathy, X-Linked Syndrome: A French Multicenter Retrospective Study. Clin. Transl. Gastroenterol. *9*, 201.

26. Gambineri, E., Ciullini Mannurita, S., Hagin, D., Vignoli, M., Anover-Sombke, S., DeBoer, S., Segundo, G.R.S., Allenspach, E.J., Favre, C., Ochs, H.D., and Torgerson, T.R. (2018). Clinical, immunological, and molecular heterogeneity of 173 patients with the phenotype of Immune Dysregulation, Polyendocrinopathy, Enteropathy, X-Linked (IPEX) syndrome. Front. Immunol. *9*, 2411.

27. McMurchy, A.N., Gillies, J., Allan, S.E., Passerini, L., Gambineri, E., Roncarolo, M.G., Bacchetta, R., and Levings, M.K. (2010). Point mutants of forkhead box P3 that cause immune dysregulation, polyendocrinopathy, enteropathy, X-linked have diverse abilities to reprogram T cells into regulatory T cells. J. Allergy Clin. Immunol. *126*, 1242–1251.

28. Wiesmann, C., Katschke, K.J., Yin, J., Helmy, K.Y., Steffek, M., Fairbrother, W.J., McCallum, S.A., Embuscado, L., DeForge, L., Hass, P.E., and van Lookeren Campagne, M. (2006). Structure of C3b in complex with CRIg gives insights into regulation of complement activation. Nature *444*, 217–220.

29. Lin, W., Truong, N., Grossman, W.J., Haribhai, D., Williams, C.B., Wang, J., Martín, M.G., and Chatila, T.A. (2005). Allergic dysregulation and hyperimmunoglobulinemia E in Foxp3 mutant mice. J. Allergy Clin. Immunol. *116*, 1106–1115.

30. Hayatsu, N., Miyao, T., Tachibana, M., Murakami, R., Kimura, A., Kato, T., Kawakami, E., Endo, T.A., Setoguchi, R., Watarai, H., et al. (2017). Analyses of a mutant Foxp3 allele reveal BATF as a critical transcription factor in the differentiation and accumulation of tissue regulatory T cells. Immunity *47*, 268–283.e9.

31. Kwon, H.K., Chen, H.M., Mathis, D., and Benoist, C. (2018). FoxP3 scanning mutagenesis reveals functional variegation and mild mutations with atypical autoimmune phenotypes. Proc. Natl. Acad. Sci. USA *115*, E253–E262.

32. Van Gool, F., Nguyen, M.L.T., Mumbach, M.R., Satpathy, A.T., Rosenthal, W.L., Giacometti, S., Le, D.T., Liu, W., Brusko, T.M., Anderson, M.S., et al. (2019). A mutation in the transcription factor Foxp3 drives T helper 2 effector function in regulatory T cells. Immunity *50*, 362–377.e6.

33. Leng, F., Zhang, W., Ramirez, R.N., Leon, J., Zhong, Y., Hou, L., Yuki, K., van der Veeken, J., Benoist, C., Hur, S., Rudensky, A.Y., and Benoist, C. (2022). The transcription factor FoxP3 can fold into two dimerization states with divergent implications for regulatory T cell function and immune homeostasis. Immunity *55*, 1354–1369.e8.

34. Bacchetta, R., Passerini, L., Gambineri, E., Dai, M., Allan, S.E., Perroni, L., Dagna-Bricarelli, F., Sartirana, C., Matthes-Martin, S., et al. (2006). Defective regulatory and effector T cell functions in patients with FOXP3 mutations. J. Clin. Invest. *116*, 1713–1722.

35. Zemmour, D., Zilionis, R., Kiner, E., Klein, A.M., Mathis, D., and Benoist, C. (2018). Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. Nat. Immunol. *19*, 291–301.

36. Bettelli, E., Carrier, Y., Gao, W., Korn, T., Strom, T.B., Oukka, M., Weiner, H.L., and Kuchroo, V.K. (2006). Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. Nature *441*, 235–238.

37. Bandukwala, H.S., Wu, Y., Feuerer, M., Chen, Y., Barboza, B., Ghosh, S., Stroud, J.C., Benoist, C., Mathis, D., Rao, A., and Chen, L. (2011). Structure of a domain-swapped FOXP3 dimer on DNA and its function in regulatory T cells. Immunity *34*, 479–491.

38. Herman, A.E., Freeman, G.J., Mathis, D., and Benoist, C. (2004). CD4+CD25+ T regulatory cells dependent on ICOS promote regulation of effector cells in the prediabetic lesion. J. Exp. Med. *199*, 1479–1489.

39. 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.

40. 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.

41. Gavin, M.A., Rasmussen, J.P., Fontenot, J.D., Vasta, V., Manganiello, V.C., Beavo, J.A., and Rudensky, A.Y. (2007). Foxp3-dependent programme of regulatory T-cell differentiation. Nature *445*, 771–775.

42. Traxinger, B.R., Richert-Spuhler, L.E., and Lund, J.M. (2022). Mucosal tissue regulatory T cells are integral in balancing immunity and tolerance at portals of antigen entry. Mucosal Immunol. *15*, 398–407.

43. Chowdhary, K., Léon, J., Ramanan, D., Mathis, D., and Benoist, C. (2023). An interwoven network of transcription factors, with divergent influences from FoxP3, underlies Treg diversity. Preprint at bioRxiv. https://doi.org/10.1101/2023.05.18.541358.

44. Benoist, C., and Mathis, D. (2013). Treg cells, life history, and diversity. In Immune Tolerance, D. Mathis and A. Rudensky, eds. (Cold Spring Harbor Press), pp. 31–44.

45. Wildin, R.S., Smyk-Pearson, S., and Filipovich, A.H. (2002). Clinical and molecular features of the immunodysregulation, polyendocrinopathy, enteropathy, X linked (IPEX) syndrome. J. Med. Genet. *39*, 537–545.

46. Singh, N., Chandler, P.R., Seki, Y., Baban, B., Takezaki, M., Kahler, D.J., Munn, D.H., Larsen, C.P., Mellor, A.L., and Iwashima, M. (2007). Role of CD28 in fatal autoimmune disorder in *scurfy* mice. Blood *110*, 1199–1206.

47. Sharma, R., Sung, S.S.J., Gaskin, F., Fu, S.M., and Ju, S.T. (2012). A novel function of IL-2: chemokine/chemoattractant/retention receptor genes induction in Th subsets for skin and lung inflammation. J. Autoimmun. *38*, 322–331.

48. Chassaing, B., Aitken, J.D., Malleshappa, M., and Vijay-Kumar, M. (2014). Dextran sulfate sodium (DSS)-induced colitis in mice. Curr. Protoc. Im. *104*. Unit.

49. Li, M., Hener, P., Zhang, Z., Kato, S., Metzger, D., and Chambon, P. (2006). Topical vitamin D3 and low-calcemic analogs induce thymic stromal lymphopoietin in mouse keratinocytes and trigger an atopic dermatitis. Proc. Natl. Acad. Sci. USA *103*, 11736–11741.

50. Acha-Orbea, H., and McDevitt, H.O. (1987). The first external domain of the nonobese diabetic mouse class II I-A beta chain is unique. Proc. Natl. Acad. Sci. USA *84*, 2435–2439.

51. Pearson, J.A., Wong, F.S., and Wen, L. (2016). The importance of the Non Obese Diabetic (NOD) mouse model in autoimmune diabetes. J. Autoimmun. *66*, 76–88.

52. Wang, L., Ray, A., Jiang, X., Wang, J.Y., Basu, S., Liu, X., Qian, T., He, R., Dittel, B.N., and Chu, Y. (2015). T regulatory cells and B cells cooperate to form a regulatory loop that maintains gut homeostasis and suppresses dextran sulfate sodium-induced colitis. Mucosal Immunol. *8*, 1297–1312.

53. Sefik, E., Geva-Zatorsky, N., Oh, S., Konnikova, L., Zemmour, D., McGuire, A.M., Burzyn, D., Ortiz-Lopez, A., Lobera, M., Yang, J., et al. (2015). Individual intestinal symbionts induce a distinct population of RORγ^{+} regulatory T cells. Science *349*, 993–997.

54. Makino, S., Muraoka, Y., Kishimoto, Y., and Hayashi, Y. (1985). Genetic analysis for insulitis in NOD mice. Jikken Dobutsu *34*, 425–431.

55. Wicker, L.S., Miller, B.J., Coker, L.Z., McNally, S.E., Scott, S., Mullen, Y., and Appel, M.C. (1987). Genetic control of diabetes and insulitis in the non-obese diabetic (NOD) mouse. J. Exp. Med. *165*, 1639–1654.

56. André, I., Gonzalez, A., Wang, B., Katz, J., Benoist, C., and Mathis, D. (1996). Checkpoints in the progression of autoimmune disease: lessons from diabetes models. Proc. Natl. Acad. Sci. USA *93*, 2260–2263.

57. McDevitt, H.O. (1998). The role of MHC class II molecules in susceptibility and resistance to autoimmunity. Curr. Opin. Immunol. *10*, 677–681.

58. Wildin, R.S., Ramsdell, F., Peake, J., Faravelli, F., Casanova, J.L., Buist, N., Levy-Lahad, E., Mazzella, M., Goulet, O., Perroni, L., et al. (2001). X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. Nat. Genet. *27*, 18–20.

59. Mimitou, E.P., Lareau, C.A., Chen, K.Y., Zorzetto-Fernandes, A.L., Hao, Y., Takeshima, Y., Luo, W., Huang, T.S., Yeung, B.Z., Papalexi, E., et al. (2021). Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. Nat. Biotechnol. *39*, 1246–1258.

60. Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). ChromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat. Methods *14*, 975–978.

61. Feng, Y., Arvey, A., Chinen, T., van der Veeken, J., Rudensky, A.Y., Gasteiger, G., and Rudensky, A.Y. (2014). Control of the inheritance of regulatory T cell identity by a cis element in the Foxp3 locus. Cell *158*, 749–763.

62. Li, X., Liang, Y., LeBlanc, M., Benner, C., and Zheng, Y. (2014). Function of a Foxp3 cis-element in protecting regulatory T cell identity. Cell *158*, 734–748.

63. Grant, F.M., Yang, J., Nasrallah, R., Clarke, J., Sadiyah, F., Whiteside, S.K., Imianowski, C.J., Kuo, P., Vardaka, P., Todorov, T., et al. (2020). BACH2 drives quiescence and maintenance of resting Treg cells to promote homeostasis and cancer immunosuppression. J. Exp. Med. *217*, e20190711.

64. Rudra, D., Deroos, P., Chaudhry, A., Niec, R.E., Arvey, A., Samstein, R.M., Leslie, C., Shaffer, S.A., Goodlett, D.R., and Rudensky, A.Y. (2012). Transcription factor Foxp3 and its protein partners form a complex regulatory network. Nat. Immunol. *13*, 1010–1019.

65. Samstein, R.M., Arvey, A., Josefowicz, S.Z., Peng, X., Reynolds, A., Sandstrom, R., Neph, S., Sabo, P., Kim, J.M., Liao, W., et al. (2012). Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. Cell *151*, 153–166.

66. Tsuda, M., Torgerson, T.R., Selmi, C., Gambineri, E., Carneiro-Sampaio, M., Mannurita, S.C., Leung, P.S.C., Norman, G.L., and Gershwin, M.E. (2010). The spectrum of autoantibodies in IPEX syndrome is broad and includes anti-mitochondrial autoantibodies. J. Autoimmun. *35*, 265–268.

67. Ge, T., Wang, Y., Che, Y., Xiao, Y., and Zhang, T. (2017). Atypical Late-Onset Immune Dysregulation, Polyendocrinopathy, Enteropathy, X-Linked Syndrome with Intractable Diarrhea: A Case Report. Front. Pediatr. *5*, 267.

68. Al Maawali, A., Derfalvi, B., Van Limbergen, J., Issekutz, A., Issekutz, T., Ghandourah, H., and Rashid, M. (2020). IPEX Syndrome with Normal FOXP3 Protein Expression in Treg Cells in an Infant Presenting with Intractable Diarrhea as a Single Symptom. Case Reports Immunol. *2020*, 9860863.

69. Darce, J., Rudra, D., Li, L., Nishio, J., Cipolletta, D., Rudensky, A.Y., Mathis, D., and Benoist, C. (2012). An N-terminal mutation of the Foxp3 transcription factor alleviates arthritis but exacerbates diabetes. Immunity *36*, 731–741.

70. Bettini, M.L., Pan, F., Bettini, M., Finkelstein, D., Rehg, J.E., Floess, S., Bell, B.D., Ziegler, S.F., Huehn, J., Pardoll, D.M., and Vignali, D.A.A. (2012). Loss of epigenetic modification driven by the Foxp3 transcription factor leads to regulatory T cell insufficiency. Immunity *36*, 717–730.

71. van der Veeken, J., Campbell, C., Pritykin, Y., Schizas, M., Verter, J., Hu, W., Wang, Z.M., Matheis, F., Mucida, D., Charbonnier, L.M., et al. (2022). Genetic tracing reveals transcription factor Foxp3-dependent and Foxp3-independent functionality of peripherally induced Treg cells. Immunity *55*, 1173–1184.e7.

72. Kim, J.M., Rasmussen, J.P., and Rudensky, A.Y. (2007). Regulatory T cells prevent catastrophic autoimmunity throughout the lifespan of mice. Nat. Immunol. *8*, 191–197.

73. Feuerer, M., Shen, Y., Littman, D.R., Benoist, C., and Mathis, D. (2009). How punctual ablation of regulatory T cells unleashes an autoimmune lesion within the pancreatic islets. Immunity *31*, 654–664.

74. Makino, S., Kunimoto, K., Muraoka, Y., Mizushima, Y., Katagiri, K., and Tochino, Y. (1980). Breeding of a non-obese, diabetic strain of mice. Exp. Anim. *29*, 1–13.

75. Liston, A., Nutsch, K.M., Farr, A.G., Lund, J.M., Rasmussen, J.P., Koni, P.A., and Rudensky, A.Y. (2008). Differentiation of regulatory Foxp3^{+} T cells in the thymic cortex. Proc. Natl. Acad. Sci. USA *105*, 11903–11908.

76. Granja, J.M., Corces, M.R., Pierce, S.E., Bagdatli, S.T., Choudhry, H., Chang, H.Y., and Greenleaf, W.J. (2021). ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. Nat. Genet. *53*, 403–411.

77. Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. Nat. Methods *18*, 1333–1341.

78. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., III, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29.

79. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165.

80. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

81. Zemmour, D., Pratama, A., Loughhead, S.M., Mathis, D., and Benoist, C. (2017). Flicr, a long noncoding RNA, modulates Foxp3 expression and autoimmunity. Proc. Natl. Acad. Sci. USA *114*, E3472–E3480.

82. Cusanovich, D.A., Hill, A.J., Aghamirzaie, D., Daza, R.M., Pliner, H.A., Berletch, J.B., Filippova, G.N., Huang, X., Christiansen, L., DeWitt, W.S., et al. (2018). A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility. Cell *174*, 1309–1324.e18.

83. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., III, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21.

84. McInnes, L., Healy, J., and Mellvile, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Preprint at arXiv. https://doi.org/10.48550/arXiv.1802.03426.

85. Melsted, P., Booeshaghi, A.S., Liu, L., Gao, F., Lu, L., Min, K.H.J., da Veiga Beltrame, E., Hjörleifsson, K.E., Gehring, J., and Pachter, L. (2021).

Modular, efficient and constant-memory single-cell RNA-seq preprocessing. Nat. Biotechnol. *39*, 813–818.

86. Booeshaghi, A.S., Min, K.H., Gehring, J., and Pachter, L. (2022). Quantifying orthogonal barcodes for sequence census assays. Preprint at bioRxiv. https://doi.org/10.1101/2022.10.09.511501.

87. Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nat. Methods *14*, 865–868.

88. Yoshida, H., Lareau, C.A., Ramirez, R.N., Rose, S.A., Maier, B., Wroblewska, A., Desland, F., Chudnovskiy, A., Mortha, A., Dominguez, C., et al. (2019). The cis-Regulatory Atlas of the Mouse Immune System. Cell *176*, 897–912.e20.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
BV605 anti-mouse CD45 (clone 30-F11)	Biolegend	Cat# 103140
AF700 anti-Mouse TCRβ (clone H57-597)	Biolegend	Cat# 109224
BUV737 anti-Mouse TCRβ (clone H57-597)	BD	Cat# 612821
BUV737 anti-Mouse NK1.1 (clone PK136)	BD	Cat# 741715
PE-eFluor™ 610 anti-mouse CD4 (clone RM4-5)	Invitrogen	Cat# 61-0042-82
PE/Cyanine7 anti-mouse CD4 (clone RM4-5)	Biolegend	Cat# 100528
AF700 anti-mouse CD8a (clone 53–5.8)	Biolegend	Cat# 100730
BV605 anti-mouse CD19 (clone 6D5)	Biolegend	Cat# 115540
APC anti-rat CD90/mouse CD90.1 (clone OX-7)	Biolegend	Cat# 202526
PE/Cyanine7 anti-rat CD90/mouse CD90.1 (clone OX-7)	Biolegend	Cat# 202518
PerCP/Cyanine5.5 anti-mouse/human CD44 (clone IM7)	Biolegend	Cat# 103032
BV510 anti-mouse/human CD44 (clone IM7)	Biolegend	Cat# 103044
PE/Cyanine7 anti-mouse CD62L (clone MEL-14)	Biolegend	Cat# 104418
BV785 anti-mouse CD62L (clone MEL-14)	Biolegend	Cat# 104440
BV785 anti-mouse CCR2 (cloneSA203G11)	Biolegend	Cat# 150621
BV510 anti-mouse CXCR3 (clone CXCR3-173)	Biolegend	Cat# 126528
BV785 anti-mouse CXCR5 (clone L138D7)	Biolegend	Cat# 145523
APC anti-mouse CXCR6 (clone SA051D1)	Biolegend	Cat# 151106
PE anti-mouse CD11b (clone M1/70)	Biolegend	Cat# 101208
PE-eFluor™ 610 anti-mouse CD11c (clone N418)	Invitrogen	Cat# 61-0114-82
Pacific Blue anti-mouse CD3 (clone 17AE)	Biolegend	Cat# 100214
PE-eFluor™ 610 anti-mouse KLRG1 (clone 2F1)	Invitrogen	Cat# 61-5893-82
PE/Cyanine7 anti-mouse PD-1 (clone 29F.1A12)	Biolegend	Cat# 135216
PE anti-mouse CD25 (clone PC-61)	Biolegend	Cat# 102008
Pacific Blue anti-mouse CD103 (clone 2E7)	Biolegend	Cat# 121418
BV605 anti-mouse CTLA-4 (clone UC10-4B9)	Biolegend	Cat# 106323
Pacific Blue anti-mouse Helios (clone 22F6)	Biolegend	Cat# 137220
APC anti-Foxp3 (clone FJK-16s)	eBioscience	Cat# 17-5773-82
AF488 anti-Foxp3 (clone FJK-16s)	eBioscience	Cat# 53-5773-82
PE anti-RORγt (clone AFKJS-9)	eBioscience	Cat# 12-6988-82
APC anti-RORγt (clone AFKJS-9)	eBioscience	Cat# 17-6988-82
PE-eFluor™ 610 anti-Gata3 (clone TWAJ)	eBioscience	Cat# 61-9966-42
AF700 anti-Ki-67 (clone 16A8)	Biolegend	Cat# 652419

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
TotalSeq-A anti-mouse hashtags (1,6,9,10,12)	Biolegend	Cat# 155801 Cat# 155811 Cat# 155817 Cat# 155819 Cat# 155823
Chemicals, peptides, and recombinant proteins		
Alt-R S.p. HiFi Caspase 9 Nuclease V3	IDT	Cat# 1081060
Phusion® High-Fidelity DNA Polymerase	NEB	Cat# M0530L
DreamTaq DNA Polymerase	Thermo Fisher	EP0702
DreamTaq PCR Master Mix	Thermo Fisher	K1071
PstI restriction endonuclease	NEB	Cat# R0140S
SmaI restriction endonuclease	NEB	Cat# R0141S
AvaI restriction endonuclease	NEB	Cat# R0152S
BglII restriction endonuclease	NEB	Cat# R0144S
NdeI restriction endonuclease	NEB	Cat# R0111S
Collagenase type II (Gibco)	Thermo Fisher	Cat# 17101015
Collagenase type II	Sigma-Aldrich	Cat# C6885-5G
Collagenase type IV	Sigma-Aldrich	Cat# C5138-1G
Hyaluronidase from bovine testes	Sigma-Aldrich	Cat# H3884
Dispase (Gibco)	Thermo Fisher	Cat# 17105041
DNase I, from bovine pancreas	Sigma-Aldrich	Cat# D4527
Benzonase	Millipore	Cat# 70664
Anti-HA Magnetic Beads	Thermo Fisher	Cat#88837
Proteinase K	NEB	Cat# P8107S
Dextran Sulfate Sodium (DSS)	Thermo Fisher	Cat# J62101.22
Calcipotriol (MC903)	MedChemExpress	Cat# HY-10001
10% Neutral Buffered Formalin	eBioscience	Cat# 00-4980-03
Bouin's solution	VWR	Cat# 15990-01
2-Mercaptoethanol	Sigma-Aldrich	Cat# M7522
TCL RNA lysis buffer	Qiagen	Cat# 1031576
10% Novex TBE Gels	Invitrogen	Cat#EC62755BOX
TE Buffer	Invitrogen	Cat# 12090015
ACK lysis buffer	Gibco	Cat# A10492-01
Critical commercial assays		
Foxp3/Transcription Factor Staining Buffer Set	eBioscience	Cat# 00-5523-00
LIVE/DEAD™ Fixable Near IR (780) Viability Kit	Invitrogen	Cat# L34994
IgE Mouse Uncoated ELISA Kit	Thermo Fisher	Cat # 88-50460-22
QIAquick Gel Extraction Kit	QIAGEN	Cat # 28706
QIAquick Nucleotide Removal kit	QIAGEN	Cat# 28306
Chromium Next GEM Single Cell ATAC	10x Genomics	Cat# 1000176
Deposited data		
scATAC-seq of splenic mutant Tregs	This manuscript	GEO: GSE237198
Bulk RNA-seq of splenic mutant Tregs (R51Q,C168Y,K199del,R309Q, F324L,R337Q)	This manuscript	GEO: GSE225891
Bulk RNA-seq of Foxp3^{+} regulatory T cells and Foxp3 reporter-null cells	van der Veeken et al. [20]	GEO: GSE154680
Gene expression by array of splenic mutant Tregs (I363V, A384T, R397W)	Hayatsu et al.[30]	GEO: GSE89654

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Cell lines		
HEK293T	ATCC	CRL-11268
Experimental models: Organisms/strains		
Mouse: C57BL/6J	Jackson Laboratory	000664
Mouse: NOD/LtJ.DOI	Makino et al.[74]	From our Colony
Mouse: B6.Foxp3IRES-GFP	Bettelli et al.[36]	From our Colony
Mouse: B6.Foxp3Thy1.1	Liston et al.[75]	From our Colony
Mouse: B6.Foxp3fs327-GFP	Ricardo et al.[19]	From our Colony
Mouse: B6. Foxp3R51Q	This manuscript	From our Colony
Mouse: B6. Foxp3C168Y	This manuscript	From our Colony
Mouse: B6. Foxp3K199del	This manuscript	From our Colony
Mouse: B6. Foxp3R309Q	This manuscript	From our Colony
Mouse: B6. Foxp3F324L	This manuscript	From our Colony
Mouse: B6. Foxp3R337Q	This manuscript	From our Colony
Oligonucleotides		
pcDNA-HA-FoxP3	Leng et al. 2022[33]	N/A
pcDNA-HA-FoxP3-R51Q	This manuscript	N/A
pcDNA-HA-FoxP3-C168Y	This manuscript	N/A
pcDNA-HA-FoxP3-K199del	This manuscript	N/A
pcDNA-HA-FoxP3-R309Q	This manuscript	N/A
pcDNA-HA-FoxP3-F324L	This manuscript	N/A
pcDNA-HA-FoxP3-R337Q	This manuscript	N/A
Customed Alt-R™ HDR Donor Oligo	IDT	N/A
Customed Alt-R® CRISPR-Cas9 crRNA	IDT	N/A
Alt-R® CRISPR-Cas9 tracrRNA, 20 nmol	IDT	Cat# 1072533
Software and algorithms		
R studio software v2022.12.0 + 353	The R Foundation	https://www.r-project.org/
Python software v3.9.7	Python	https://www.python.org/
Cell Ranger ATAC software v1.2	10x Genomics	N/A
ArchR v1.0.1	Granja et al.[76]	https://www.archrproject.com/
Signac v1.4	Stuart et al.[77]	https://stuartlab.org/signac/
ASAP-seq	Mimitou et al.[59]	https://github.com/caleblareau/asap_to_kite
Seurat package (v4.0.2)	Hao et al.[78]	https://satijalab.org/seurat/
ChromVar v1.4.1	Schep et al.[60]	N/A
deeptools v3.0.2	Ramirez et al.[79]	N/A
IGV v2.4.14	Robinson et al.[80]	N/A
GenePattern software package	Broad Institute	http://software.broadinstitute.org/cancer/software/genepattern/
PRISM v9.5.0	GraphPad	https://www.graphpad.com/
FlowJo 10.7	BD Biosciences	https://www.flowjo.com/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Christophe Benoist (cb@hms.harvard.edu).

Materials availability

Mouse lines generated in this study have been cryopreserved and will be made available to qualified investigators upon justified request to the lead contact.

Data and code availability

Data newly reported in this manuscript, BulkRNAseq and scATAC-seq data, have been deposited at the Gene Expression Omnibus and are publicly available as of the date of publication. This paper also analyzes existing, publicly available data. Accession numbers are listed in the key resources table.

This paper does not report original code.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

IPEX human mutations

The 6 mutations from IPEX patients (Table S1) were previously reported in Zemmour et al.[3]: R51Q, C168Y (C169Y in human FOXP3), K199del (K200del in human FOXP3), R309Q and R337Q; or in Bacchetta et al[34].

Mice

C57Bl/6J (B6) mice were purchased from the Jackson Laboratory. *Foxp3*$^{IRES-GFP}$,[36] Foxp3$^{Thy1.1}$[75] and *Foxp3*$^{fs327-GFP}$[19] mice on the B6 background, and NOD/LtJ.DOI (NOD) mice were maintained in our laboratory. Except when specified, 6 to 10-wk-old male and female mice were used in experiments. They originated from breeding of heterozygous Foxp3$^{Mutant/WT}$ females with Foxp3$^{WT-Thy1.1}$ males, to generate both mutants and WT littermate controls from the same cage, female or male. For backcross to the NOD background, mutant males were crossed with NOD females to generate heterozygous Foxp3$^{mutant/WT}$ F1 females, which were then crossed to NOD male. All mice were housed under specific pathogen-free conditions and all experiments were performed in accordance with guidelines from the Institutional Animal Care Committee of Harvard Medical School (protocol #IS00001257).

Cell lines

For transfection, HEK293T cells were purchased from ATCC (CRL-11268), and Cells were maintained in DMEM (High glucose, L-glutamine, Pyruvate) with 10% fetal bovine serum, 1% penicillin/streptomycin.

METHOD DETAILS

Mutant mice generation by CRISPR/Cas9

Mutant mice were generated by CRISPR mutagenesis, essentially as described,[81] except that the micro-injected complex included recombinant Cas9 (0.13μM Alt-R S.p. HiFi Caspase 9; IDT), 0.6μM sgRNA (IDT), and 0.3μM of a single stranded recombination template carrying the desired mutation with 60 bp of flanking homology arms, mixed in 0.2μm-filtered 0.1X TE Buffer). This mix was microinjected into the male pronucleus of fertilized mouse oocytes, which were then implanted into pseudo-pregnant females.

To generate fertilized oocytes, we used males from the *Foxp3*$^{IRES-GFP}$[36] knockin in order to insert the mutation directly into this tagged locus. However, this strategy was successful only in half of the mutation generation, for R51Q, K199del and F324L. The failures were, a posteriori, explained by the fact that the mix likely leaked into the cytoplasm and the mutation was ultimately inserted into the non-tagged female genetic material. Two independent founders carrying the R51Q mutation were obtained, one in the normal allele, one in the GFP-tagged allele; the former was used for most experiments.

To facilitate further mouse genotyping, we added a silent mutation in each mutated template, introducing or deleting a restriction enzyme site. The mutation and the remainder of the FoxP3-coding region were verified by PCR amplification and Sanger sequencing after using a DNA gel extraction kit (Qiagen). Genotyping of each strain was performed by PCR amplification around the mutated and subsequent restriction enzyme digestion was done. If there were doubt about the digestion, the genotype was further verified by Sanger sequencing of the amplicon. The PCR primers as well as the enzyme restriction used for each line for genotyping are described in Table S1. Genotyping of BC1 mice at MHC-II loci (H2-A$^{g7/g7}$ or H2-A$^{b'g7}$) was done by PCR amplification and Sanger sequencing, based on sequences from ref. [50].

Inflammatory challenges

For DSS treatment, 2.5% DSS (Thermo Scientific) was provided in the drinking water from day 0 to day 6 and mice were followed until day 10 (weight and clinical state).

For MC903 treatment, 3 nmol of calcitriol dissolved in 70% ethanol (10mg stock, MedChemExpress – Cat# HY-10001) was administered topically to the dorsal and ventral aspects of the skin of one adult mice ear, daily, for 10 days and then mice were followed until day 25. Mice were monitored once per day for ear local inflammation, which was scored based on the ear thickness, measured using a caliper – as a proportion of increase (%) from the non-injured D0 measure.

Histology

All tissues were collected and fixed in 10% formalin at least 24 h, except the colon which was fixed in Bouin solution. Tissues were embedded in paraffin and processed at the HMS Rodent Histopathology for hematoxylin and eosin (H&E) staining. Images were acquired with a Nikon Ti inverted microscope at ×10 magnification. For general grading tissue immune infiltration, evaluation from H&E sections was performed in a blinded fashion by an independent pathologist at the HMS Rodent Histopathology Core. For insulitis, several islets (>15) were scored for each mouse (where 0: normal; 1: peri-insulitis, accumulation around the islet but no breach of the islet capsule; 2; clear insulitis, inflammatory cells invading and in contact with β-cells; 3: severe insulitis, >50% of the islet taken over or destroyed) Perivascular inflammation in the connective/ductal spaces of the pancreas (observed in R337Q mice) was scored as 0: no inflammation; 1: clear perivascular infiltrates, intact acinar or islet tissue; 2: abundant perivascular infiltrates, with destruction of the surrounding acinar tissue, but large areas of the pancreas still present; 3: severe, quasi-complete destruction of the exocrine pancreas.

ELISA IgE

IgE concentrations were measured by commercial ELISA kit according to the manufacturer's recommendations (Thermo Fisher Scientific), after plasma dilution at 1:50 to 1:100.

Plasmids

For Mammalian expression plasmids, HA-tagged mouse FoxP3 CDS was inserted into pcDNA3.1+ vector between KpnI and BamHI sites. All FoxP3 mutations including R51Q, C168Y, K199del, R309Q, F324L and R337Q were generated by site-directed mutagenesis using Phusion High Fidelity (New England Biolabs) DNA polymerases.

FoxP3-DNA pulldown assay

HEK293T cells were transfected with pcDNA encoding HA-tagged FoxP3 (wild-type or mutants). After 48 h, cells were lysed using RIPA buffer (10mM Tris-HCl, pH 8.0, 1mM EDTA, 1% Triton X-100, 0.1% Sodium Deoxycholate, 0.1% SDS, 140 mM NaCl and 1x proteinase inhibitor) and treated with Benzonase (Millipore) for 30 min. The lysate was then incubated with Anti-HA Magnetic Beads (Thermo Fisher) for 1 h. Beads were washed three times using RIPA buffer and incubated with IR-FKHM^{4g} (TAGGAAAATT**TGTTTAC** TCGA**GTAAACA** TC) for 20 min at room temperature. Bound DNA was recovered using proteinase K (New England Biolabs), purified using QIAquick Nucleotide Removal kit (QIAGEN) and analyzed on 10% Novex TBE Gels (Invitrogen).

Isolation of lymphocytes from lymphoid and non-lymphoid tissues

Spleen and lymph nodes

Immunocytes were released using mechanical disruption followed by filtering and washes in FACS Buffer (phenol red-free DMEM (Gibco) containing 2% fetal calf serum (FCS)). Additional step for red blood cell lysis in spleen samples was performed using ACK lysis buffer (Gibco, ref A10492-01).

Colon

intestinal tissues were measured, cleaned, and incubated in RPMI (Gibco) containing 1 mM dithiothreitol, 20 mM EDTA, and 2% FBS at 37°C for 15 min to remove epithelial cells. The colon was then minced and dissociated in RPMI containing 1.5 mg/mL collagenase II (Gibco), 0.5 mg/mL Dispase (Gibco), and 1% FCS, at 37°C for 45 min with constant stirring. The digested materials were filtered through a 40-μm cell strainer, washed with 2% FBS-RPMI and resuspended in FACS buffer.

Lungs

the right lobe of the lungs was collected, minced and dissociated in collagenase solution (1 mg/mL collagenase type IV (Sigma), 150 μg/mL DNase I (Sigma) and 1% FCS in RPMI) and incubated in a water bath at 37°C with constant shaking for 30 min. Digested tissues were filtered through a 40-μm cell strainer and washed in 2% FCS. Red blood cells were lysed using ACK lysis buffer and the pellet was then resuspended in FACS buffer.

Skin

The hairless part of the mouse ears was cut and the dorsal and ventral side were separated by pulling them apart with forceps. Any remaining cartilage was gently removed. The two sides were then minced and digested in RPMI containing 2 mg/ml collagenase type II (Sigma), 150 μg/ml DNase I (Sigma), and 0.5 mg/ml hyaluronidase (Sigma) for 50 min – with frequent vortexing (every 15 min). The dissociation lysat was filtered through a 40-μm cell strainer, washed with 2% FBS-RPMI and resuspended in FACS buffer.

Flow cytometry

Surface staining of the single cell suspensions was performed for 30 min at 4°C, and viability was assessed using LIVE/DEAD Fixable viability dye as per the manufacturer's instructions (Thermo Fisher Scientific).

The following surface markers antibodies from Biolegend were used: anti-CD3 (clone: 145-2C11), anti-CD4 (RM4-5), anti-CD8a (53–5.8), anti-CD45 (30-F11), anti-CD11b (M1/70), anti-CD11c (N418), anti-CD19 (6D5), anti-TCRβ (H57-597), anti-NK1.1 (PK136), anti-KLRG1 (2F1), anti-PD-1 (29F.1A12), anti-CD25 (PC-61), anti-CD44 (IM7), anti-CD62L (MEL-14), anti-Thy1.1 (OX-7), anti-CCR2 (SA203G11), anti-CXCR3 (CXCR3-173), anti-CXCR5 (L138D7), anti-CXCR6 (SA051D1), anti-CD103 (2E7). Samples were

then fixed overnight at 4°C using the using 100 μL of Foxp3 Fix/Perm buffer (eBioscience). After membrane permeabilization using 1X permeabilization buffer (eBioscience) for 5 min, intracellular staining was performed for 120 min at room temperature using the following antibodies: anti-Ctla-4 (UC10-4B9, Biolegend), anti-Helios (22F6, Biolegend), anti-Foxp3 (FJK16, Thermofisher), anti-Gata3 (TWAJ, Thermofisher), anti-RORγ (AFKJS-9, Thermofisher), anti-Ki-67 (16A8, Biolegend). Cells were acquired with an Aurora flow cytometer (Cytek Biosciences) or a FACSymphony flow cytometer (BD Biosciences). Data were analyzed using FlowJo software version 10 (TreeStar, BD LifeSciences).

Cell sorting for bulk RNAseq and scATACseq

Cells were sorted using BD MoFlo Astrios EQ, FACSAria-561, or FACSAria-594 machines.

For Bulk RNAseq, male mutant Tregs were sorted as aTreg (DAPI$^-$ CD19$^-$ TCRb$^+$ CD4$^+$ CD25^{hi} CD44^{hi} CD62L^{lo}) or rTreg ((DAPI$^-$ CD19$^-$ TCRb$^+$ CD4$^+$ CD25^{hi} CD44$^-$ CD62L$^+$). Female mutant Tregs were sorted as aTreg (DAPI$^-$ CD19$^-$ TCRb$^+$ CD4$^+$ Thy1.1$^-$ CD25^{hi} CD44^{hi} CD62L^{lo}) or rTreg ((DAPI$^-$ CD19$^-$ TCRb$^+$ CD4$^+$ Thy1.1$^-$ CD25^{hi} CD44$^-$ CD62L$^+$). Tconv (when applicable) were sorted as DAPI$^-$ CD19$^-$ TCRb$^+$ CD4$^+$ CD25$^-$. For mutations inserted into the Foxp3$^{ires.GFP}$ background, the Treg purity was checking using the GFP channel: 99–100% for the rTreg compartment and 94–96% for the aTreg compartment. For scATAC seq, we used the same gating strategy, except that for the mutations inserted into the Foxp3$^{ires.GFP}$ background, we also added a "GFP+" gate, increasing the purity (Figure S7A).

Bulk RNA-seq library preparation

BulkRNA-seq analyses were perfomed in mutant males and heterozygous females (cf. above the sorting strategy). 1000 cells of each population were directly double-sorted into 5 μL TCL buffer (Qiagen) supplemented with 1% 2-mercaptoethanol (Sigma) for cell lysis. Samples were then processed into Smart-seq2 RNA-seq library preparation and sequencing by the Broad Genomics Platform, following the standard ImmGen ultra-low-input RNA-seq protocol (immgen.org).

Bulk RNA-seq pre-processing and analysis

Reads were aligned to the GRCm38 - mm10 genome by STAR and counts quantified using featureCounts (Subread). Samples with fewer than 8,000 genes with more than ten reads, high contamination by hematopoietic-cell-specific transcripts, median transcript integrity number for housekeeping genes below 45, or poor intra-replicate correlation were excluded from downstream analyses.

Software

All transcriptomic analyses were performed using R studio software (v2022.12.0 + 353), and visualizations were generated using R studio software (v2022.12.0 + 353) or GraphPad Prism software (v9.5.0).

Filtering data

Genes with a minimum read count of 15 in all replicates for each specific population (i.e., rTreg male C168Y) were retained for the downstream comparisons.

Differentially expressed genes

We used an uncorrected t test to estimate the significance of differential gene expression between the different groups from the normalized read counts dataset. Genes with a FoldChange >2 or <0.5 and p value <0.01 were selected for further analysis. For each comparison, a population was compared to a pool of WT from the same state (rTreg or aTreg), the same sex (male or female) and the same background (straight B6 or B6.Foxp3$^{ires.GFP}$)

Selection of R51Q, K199del and R309Q specific genes

These genesets were extracted from the FoxP3-dependent gene signature from,[20] displayed in Table S2. Then they were further refined based on the concordance between the rTreg and aTreg datasets for K199del and R309Q, and between the reTreg from B6 and B6.Foxp3$^{ires.GFP}$ for R51Q (FoldChange in both setting less than 0.75 and at least in one of them less than 0.66)

Transcriptomic scores

The *Hill* Up and *Scurfy* Up scores were similarly calculated. Briefly, the average FoldChange of the up-signature genes from the respective signature was calculated for each replicate. The values were then normalized into a specific scale, where 0 corresponded to the average value in Tconv for the Hill Up and in WT Treg for Scurfy Up, and 1 corresponded to the average value in WT Treg for the Hill Up and in Scurfy Treg for Scurfy Up.

Comparison to published datasets

For published datasets, count matrices were downloaded directly through GEO and used for differential expression analyses. We used the following dataset GSE154680[20], GSE89654[30] (Table S2).

scATAC-seq library preparation

Nuclei isolation, transposition, GEM generation, and library construction targeting capture of 10000 cells were carried out as detailed in the Chromium Next GEM Single Cell ATAC manual (10x Genomics) with modifications to allow sample hashtagging (see "scATAC Hashtagging (ASAP-seq)" below). Libraries were pooled and sequenced on an Illumina NovaSeq 6000 to a final median depth of approximately 20–30,000 paired-end reads per cell. Sequencing data were converted to fastq files, aligned to the mm10 reference genome, and quantified per cell using Cell Ranger ATAC software (10x Genomics, v1.2).

scATAC hashtagging (ASAP-seq)

To include multiple experimental conditions in the same scATAC run, we hashtagged cells using a modification of the ASAP-seq strategy[59] for low cell input primary cell samples. Before sorting, cells were hashtagged with mouse TotalSeqA DNA-barcoded hashtag antibodies at the same time as staining with fluorophore-conjugated antibodies (BioLegend). Hashtags used are provided in Table S5. Cells were sorted into DMEM +5% FCS in DNA Lo-Bind tubes (Eppendorf, cat # 022431021). After spinning down for 5 min at 500g in a refrigerated centrifuge at 4°C, cells were resuspended in 100 μL chilled 0.1x Omni Lysis buffer (1x Omni Lysis buffer (10mM Tris-HCl, 10mM NaCl, 3 mM $MgCl_2$, 0.1% Tween 20, 0.1% NP40 substitute/IGEPAL, 0.01% Digitonin, 1% BSA in nuclease free water) diluted 1:10 in Wash/Lysis Dilution Buffer (10mM Tris-HCl, 10mM NaCl, 3 mM $MgCl_2$, 1% BSA in nuclease free water), gently mixed by pipetting, and incubated on ice for 6.5 min. Following lysis, 100 μL chilled wash buffer was added and gently mixed by pipetting. Cells were spun down for 5 min at 500 g at 4°C, all but 5 μL of supernatant was removed, and 45 μL of chilled 1x nuclei buffer (10x Genomics) was added without mixing. After one more centrifugation step at 500g, 4°C for 5 min, supernatant was removed, and samples were resuspended in 7ul 1x nuclei buffer for cell counting and input into transposition, barcoding, and library preparation according to the Chromium Next GEM Single Cell ATAC manual (10x Genomics).

Modifications to the original 10X protocol were made as described in the original ASAP-seq publication and as detailed at https://citeseq.files.wordpress.com/2020/09/asap_protocol_20200908.pdf. Briefly, 0.5 μL of 1uM BOA bridge oligo was spiked into the barcoding reaction. During GEM incubation, an additional 5 min incubation at 40°C was added to the beginning of the protocol. 43.5 instead of 40.5 μL of Elution Solution I was added during silane bead elution to recover 43 μL. 40 μL was used for SPRI clean up as indicated in the protocol, while 3 μL was set aside. During SPRI cleanup, the supernatant was saved. The bead bound fraction was processed as in the protocol, while for the supernatant fraction, 32 μL SPRI was added for 5 min. Beads were collected on a magnet, washed twice with 80% ethanol, and eluted in 42 μL EB. This 42 μL was combined with the 3 μL set aside from the previous step as input into the HTO indexing reaction. HTO Indexing PCR was run with partial P5 and indexed Rpxx primers (https://citeseq.files.wordpress.com/2020/09/asap_protocol_20200908.pdf) as: 95°C 3 min, 12–14 cycles of (95°C 20 s, 60°C 30 s, 72°C 20 s), 72°C 5 min. The PCR product was cleaned up with 1.6X SPRI purification for quantification and sequencing alongside ATAC libraries.

scATAC-seq preprocessing and visualization

Data analysis was performed using Signac v1.4[77]. For quality control, only cells with at least 1×10^3 fragments per cell (depending on sequencing depth of experiment), greater than 50 percent reads in peaks, TSS enrichment score greater than 2, nucleosome signal less than 10, and ratio of blacklist-region reads less than 0.05 were retained for further analysis. Putative doublets identified by ArchR v1.0.1[76] and non-Treg, non-Tconv contaminant cells were also removed. We used the latent semantic indexing approach as previously described[82,83]. Binarized count matrices were normalized using the term frequency-inverse document frequency (TF-IDF) transformation and reduced in dimensionality by singular value decomposition (SVD). As the first component was highly correlated with sequencing depth, SVD components 2–30 were used to generate a shared nearest neighbor (SNN) graph for clustering and as input into UMAP[84] with cosine distance metric for visualization.

scATAC-seq analysis

Hashtag counts + assignments

Hashtag processing followed the original recommendations of the ASAP-seq paper[59], using asap_to_kite (https://github.com/caleblareau/asap_to_kite) to process FASTQs files for downstream quantification by the bustools and kite workflows.[85,86] We used the HTODemux()[87] function in the Seurat package (v4.0.2)[78] to remove doublets and call hashtag identities.

Peak sets

To enable comparisons across conditions and datasets, we used a common set of Treg-specific open chromatin regions, defined previously[43] by supplementing pan-immunocyte OCRs from the ImmGen consortium[88] with additional peaks from Treg scATAC-seq data.

Motif accessibility analysis

Bias-corrected relative motif accessibility was calculated using chromVAR.[60] We used motifmatchr (https://github.com/GreenleafLab/motifmatchr) to scan OCRs in our refence set from the curated set of mouse motif PWMs from the Buenrostro lab (https://github.com/buenrostrolab/chromVARmotifs/tree/master/data/mouse_pwms_v2.rda).

Gene scores

Gene scores were calculated with Archr v1.0.1, using an exponentially weighted function that accounts for the activity of distal OCRs in a distance-dependent manner[76] and provides an approximate proxy for gene expression.

Pseudobulk track visualization

To visualize pseudobulk profiles, BAM files containing reads for each group of cells were extracted using Sinto (https://github.com/timoast/sinto), shifted to account for Tn5 cut-sites, and converted to bigwigs using deeptools[79] for display in the Integrative Genomics Viewer[80].

Differential accessibility

We calculated differential accessibility between FoxP3 mutant and WT cells (GFP WT for R51Q and K199del and B6 WT sorted on $CD25^+$ population for R337Q) by using a logistic regression per OCR with number of fragments per cell included as a latent variable. To avoid effects driven by cell composition, we computed differentials separately between clusters corresponding to rTregs and aTregs. OCRs with average $|\log_2$ Fold Change$| > 0.1$ and p value <0.05 were designated as differential. In the case of the full KO comparison, we computed log2FC between quantile-normalized aggregated scATAC pseudobulk profiles from GFP^+ FoxP3 KO ($Foxp3^{fs327}$-GFP/Foxp3-Thy1.1) or GFP^+ WT ($Foxp3^{wt}$-GFP/Foxp3-Thy1.1) cells from a previous study[43].

Motif enrichment

To calculate enrichment of motifs within differential OCRs, we used a permutation testing framework. We compared the number of observed motif matches within each OCR set to the number of matches among a set of 100 background OCRs matched for GC content and accessibility (chosen using the chromVAR getBackgroundPeaks() function). Significance was assessed using a two-sided Z test, with Benjamini-Hochberg false discovery rate correction. We kept motif enrichments with FDR $< 1 \times 10^{-3}$.

OCR signature relative accessibility

Per-cell relative accessibility of OCR sets, including signature OCRs, TF binding sites, etc, was calculated using the chromVAR computeDeviations() function[60].

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were performed using R studio software (v2022.12.0 + 353), Python software (v3.9.7) or GraphPad Prism (v9.5.0) software. If not stated otherwise, data were presented as mean ± SD, statistical significance was calculated by a Mann-Whitney test. Except for selection of differential expressed genes (cf. above), $p < 0.05$ was considered significant. P-values for gene signature enrichment either up or down in volcano plots were determined using the $\chi 2$ test, relative to a 50/50 null distribution (behavior of the signature in the selected dataset versus the signature in a dataset where 50% of the genes would be represented in either side of the volcano plot).

Cell Reports, Volume *42*

Supplemental information

Mutations from patients with IPEX ported to mice reveal different patterns of FoxP3 and Treg dysfunction

Juliette Leon, Kaitavjeet Chowdhary, Wenxiang Zhang, Ricardo N. Ramirez, Isabelle André, Sun Hur, Diane Mathis, and Christophe Benoist

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. FoxP3 mutations' position and breeding strategy. Related to Figure 1.

A. Amino-acid conservation between mouse (top line, NP_473380.1) and human (bottom line, NP_054728.2) FOXP3 proteins. Positions of the engineered missenses mutations are highlighted in a square.

B. Mutation positions at the FoxP3 locus: (left) Mapping of RNA-seq reads from Mutant Treg cells to FoxP3 locus in representative samples. Square indicates mutations. Mutations obtained into the Foxp3.ires.GFP background are labelled at the right; (right) Sanger traces of the engineered mutations – base changes corresponding to the princeps mutation and to the silent mutation associated to the enzyme restriction site are highlighted in colors.

C. Breeding strategy to obtain experimental mice, both male and female, both mutant and control, from the same breeders.

D. K199del Foxp3 MFI (orange line) was tested with two different FoxP3 antibodies, for which the targeted epitope was supposedly not impacted by the mutation position (Clone FJK-16S targeting an epitope between the amino acid (AA) 75 and 125 of FoxP3; Clone NRRF-30 targeting an epitope between the AA 1 and 75 of FoxP3). The K199del Foxp3 MFI was lower than WT Foxp3 MFI from WT *Foxp3.ires.GFP* mice (grey line), but higher than the negative control (Tconv, black line). However, GFP MFI, surrogate marker of the Foxp3 locus expression, didn't change in K199del Treg versus WT Treg.

Figure S2. Flow cytometric and transcriptomic Treg profiles in heterozygous females. Related to Figure 2.

A. Representative flow cytometry plots of several Treg markers in gated $CD4^{+}TCRB^{+}Thy1.1^{-}$ Foxp3 mutant Tregs from heterozygous females: CXCR3, CXCR5, PD-1, CD103, Ki67,

CTLA-4, Helios, KLRG1. The two WT littermates shown are from the B6 (WT) and *Foxp3-ires-gfp* backgrounds (WT GFP).

B. Representative flow cytometry CD44/CD62L plots of gated $CD4^+TCRB^+Thy1.1^-$Foxp3 mutant Tregs from heterozygous females; the two WT littermates shown are from the B6 (WT) and *Foxp3-ires-gfp* backgrounds (WT GFP).

C. FoldChange vs p-value (volcano) plots comparing normalized expression in mutant Tregs to WT Tregs from the same background (B6 or *Foxp3-ires-gfp*), in rTreg (top line), aTreg (bottom line) of heterozygous females – highlighting the Hill Treg signature and their respective numbers of genes up/down.

D. Ranked FC plots of the up Treg signature transcripts for each mutant Treg, ranked according to mean FC a pool of Tconv cells versus WT Tregs (black dots). FC values for mutants (colors dots) are computed from the mutant Tregs versus a pool of matched-background WT Tregs in heterozygous females.

Figure S3. Impact of FoxP3 mutants in hemizygous males, with a focus on R337Q. Related to Figure 3.

A. FoxP3 MFI quantification in mutant male Tregs. MFI was normalized vs the mean MFI of WT littermates from the same experiment and background (B6 or *Foxp3-ires-gfp*). Each dot is an individual mouse.

B. Representative flow cytometry plots PD-1/KLRG1 of gated $CD4^+TCRB^+Foxp3^+$ male R337Q mutant Tregs or WT Tregs at the left. At the right, percentage of $KLRG1^+$ or PD1+ Tregs in all mutants. Each dot is an individual mouse.

Figure S4. Pathological manifestations and transcriptomic changes in mutant males at baseline. Related to Figure 4.

A. Weight follow-up of male mutant mice from 5 to 26 weeks old (6 months) or to 20 weeks old for R337Q (dead or euthanatized before 26 weeks old), color-coded by mutation. Each follow-up point shows the average weight and its SD (at least 3 mice from two independent litters for each mutation). The WT curve is the average weight of a pool of WT.

B. Spleen weight in older hemizygous male mice, at 26-28 weeks-old for all except R337Q which shows the spleen's weight at 18-22 weeks. Each dot is an individual mouse.

C. Representative pictures of H&E staining of lungs from R337Q or WT littermate at 20 weeks-old, showing mononuclear infiltrate in the lungs.

D. *Il4* gene expression in hemizygous male mutant Tregs, computed against the mean expression in a pool of WT Tregs from the same background (B6 or *Foxp3-ires-gfp*). Each dot is an individual mouse.

E. FoldChange vs p-value (volcano) plots comparing normalized expression in male mutant aTregs to male WT aTregs from the same background (B6 or *Foxp3-ires-gfp*) – highlighting the Scurfy signature and their respective numbers of genes up/down.

Figure S5. Pathological ear manifestations from MC903-treated mice at the resolution phase of dermatitis (day 25) in FoxP3 mutant males. Related to Fig5

Representative pictures of H&E staining of ear skin from MC903 treated mice at Day 25: C168Y, K199del, R309Q, F324L.

Figure S6. Transcriptional changes in mutant rTreg and mutant aTreg. Related to Figure 6.

A. FoldChange vs p-value (volcano) plots comparing normalized expression of each mutant Tregs to WT Tregs from the same background (B6 or *Foxp3-ires-gfp*) in heterozygous females. The top row shows the rTreg dataset and the bottom row, the aTreg dataset. The Foxp3 dependent gene signature up (defined from the dataset GSE154680) is highlighted in red and the numbers of up/down genes from this signature are shown. P-values for the gene signature enrichment are determined using the $\chi 2$ test (cf. Methods).

B. FoldChange vs FoldChange plots comparing R51Q Treg to WT Tregs in two different lines: main line (B6.Foxp3.GFP) in y-axis and alternative line (B6.Foxp3) in x-axis. Left plot shows the rTreg dataset and the right plot, the aTreg dataset. The Foxp3 dependent gene signature up (defined from the dataset GSE154680) is highlighted in red. The square is isolating the concordant genes between the main line and the alternative line in the rTreg dataset.

C. FoldChange vs FoldChange plot comparing R337Q rTregs to WT rTregs (y-axis) from KO rTregs to WT rTregs (x-axis). The Foxp3 dependent gene signature (defined from the dataset GSE154680) is highlighted in red (up) and blue (down). The green dots are the genes highly concordant between the two lines and the numbers are shown.

D. Similar Heatmap than Fig. 6F but also showing the behavior of the other mutants (R51Q, C168Y, K199del, R309Q, F324L).

E. FoldChange vs FoldChange plots comparing R337Q Treg to WT Tregs in two different sexes: male in y-axis and female in x-axis. Left plot shows the rTreg dataset and the right plot, the aTreg dataset. The Foxp3 dependent gene signature (defined from the dataset GSE154680) is highlighted in red (up) and blue (down).

Figure S7. Changes in chromatin accessibility in mutant Tregs from heterozygous females. Related to Figure 7.

A. Flow cytometry gating strategy for isolate the mutant Treg population for scATAC seq.

B. scATAC Quality control 2D plot showing in y-axis the percentage of reads in peaks and in x-axis the number of fragments in logarithmic scale (log10).

C. Density of cells from WT $CD25^{+}$ overlaid onto UMAP from Fig. 7A

D. Aggregated accessibility profiles of cells from each genotype at the *Foxp3* locus in bulk ATACseq dataset; *Foxp3* CNS2 region is highlighted.

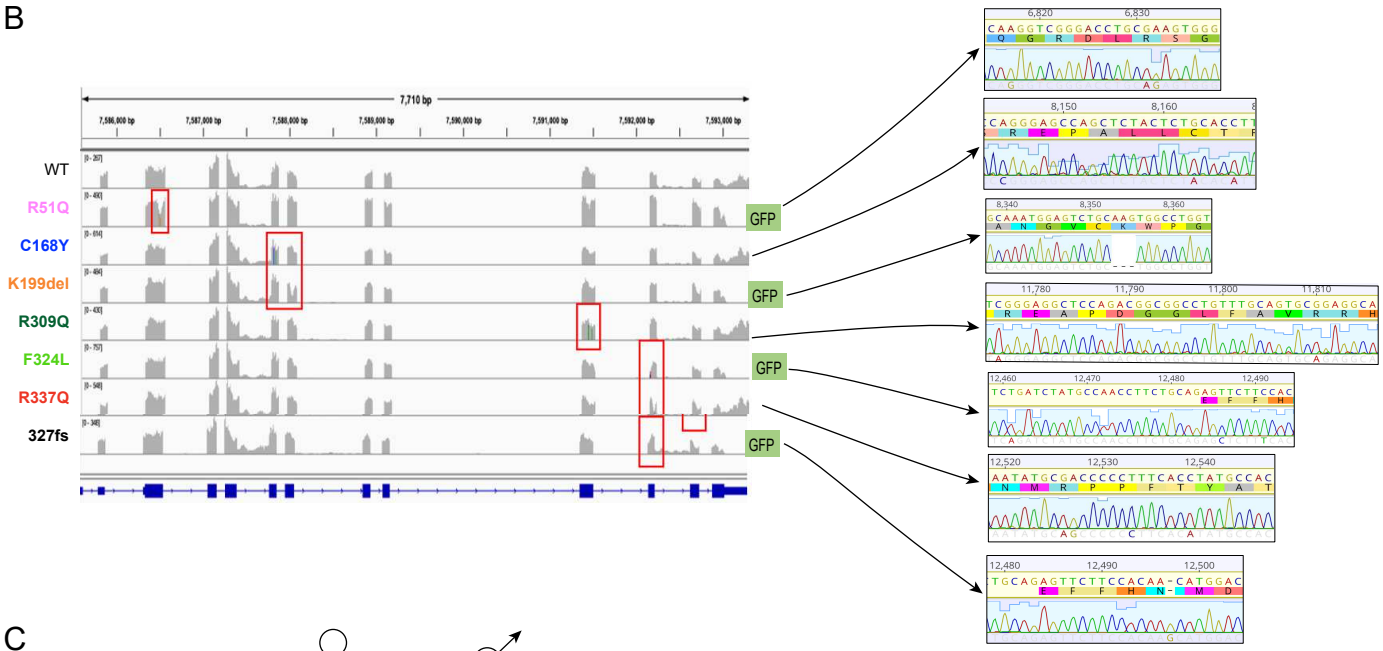
E. Heatmap of fold Change (log2) in aggregated accessibility of cells from each mutant vs its respective WT comparator for specific transcription factors: *Batf, Jund, Bach2, Smarcc1, Lef1, Tcf7*

Fig. S1

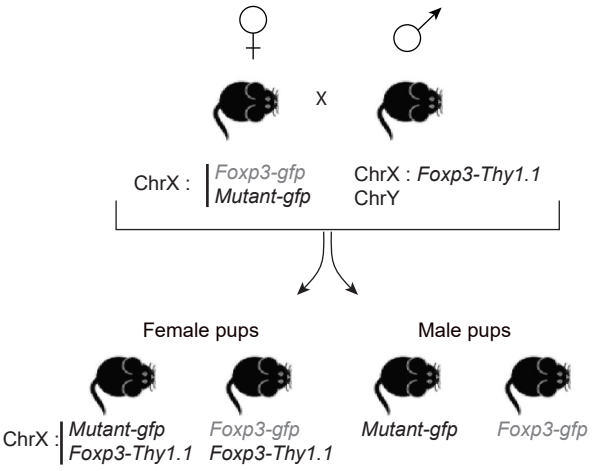
A


p.R51Q
MOUSE NP_473380.1 1 MPNPRPAKPMAPSLALGPSPGVLPSWKTAPKGSELLGTRGSGGPFQGRDLRSGAH-TSSSLNPLPPSQLQLPTVPLVMVA 79
HUMAN NP_054728.2 1 MPNPRPGKPSAPSLALGPSPGASPSWRAAPKASDLLGARGPGGTFQGRDLRGGAHASSSSLNPMPPSQLQLPTLPLVMVA 80
p.R51Q
NP_473380.1 80 PSGARLGPSPHLQALLQDRPHFMHQLSTVDAHAQTPVLQVRPLDNPAMISLPPPSAATGVFSLKARPGLPPGINVASLEW 159
NP_054728.2 81 PSGARLGPLPHLQALLQDRPHFMHQLSTVDAHARTPVLQVHPLESPAMISLTPPTTATGVFSLKARPGLPPGINVASLEW 160
p.C168Y
p.Lys199del
NP_473380.1 160 VSREPALLCTFPRSGTPRKDSNLLAAPQGSYPLLANGVCKWPGCEKVFEEPEEFLKHCQADHLLDEKGKAQCLLQREVVQ 239
NP_054728.2 161 VSREPALLCTFPNPSAPRKDSTLSAVPQSSYPLLANGVCKWPGCEKVFEEPEDFLKHCQADHLLDEKGRAQCLLQREMVQ 240
p.C169Y
p.Lys200del
p.R309Q
NP_473380.1 240 SLEQQLELEKEKLGAMQAHLAGKMALAKAPSVASMDKSSCCIVATSTQGSVLPAWSAPREAPDGGLFAVRRHLWGSHGNS 319
NP_054728.2 241 SLEQQLVLEKEKLSAMQAHLAGKMALTKASSVASSDKGSCCIVAAGSQGPVVPAWSGPREAPD-SLFAVRRHLWGSHGNS 319
p.F324L
p.R337Q
p.R309Q
NP_473380.1 320 SFPEFFHNMDYFKYHNMRPPFTYATLIRWAILEAPERQRTLNEIYHWFTRMFAYFRNHPATWKNAIRHNLSLHKCFVRVE 399
NP_054728.2 320 TFPEFLHNMDYFKFHNMRPPFTYATLIRWAILEAPEKQRTLNEIYHWFTRMFAFFRNHPATWKNAIRHNLSLHKCFVRVE 399
p.F324L
p.R337Q
NP_473380.1 400 SEKGAVWTVDEFEFRKKRSQRPNKCSNPCP-- 429
NP_054728.2 400 SEKGAVWTVDELEFRKKRSQRPSRCSNPTPGP 431

B


7,710 bp
WT
R51Q
C168Y
K199del
R309Q
F324L
R337Q
327fs
GFP
GFP
GFP
GFP

C


X
ChrX : Foxp3-gfp Mutant-gfp
ChrX : Foxp3-Thy1.1 ChrY
Female pups
Male pups
ChrX : Mutant-gfp Foxp3-Thy1.1
Foxp3-gfp Foxp3-Thy1.1
Mutant-gfp
Foxp3-gfp

D

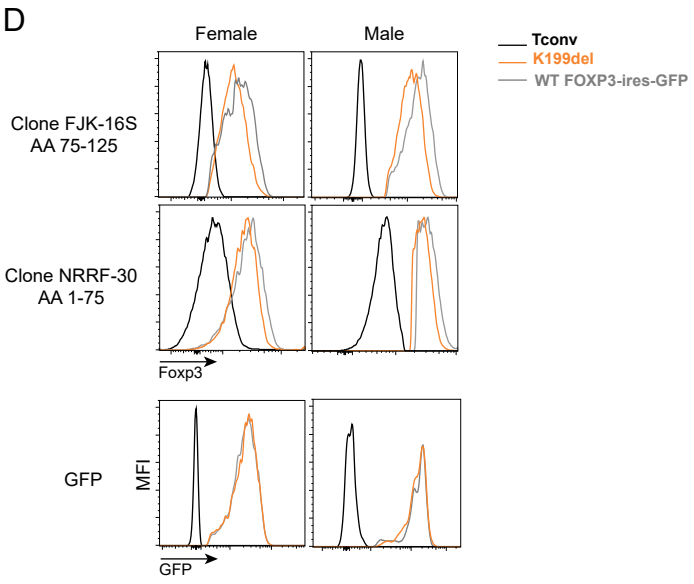

Female
Male
Tconv
K199del
WT FOXP3-ires-GFP
Clone FJK-16S AA 75-125
Clone NRRF-30 AA 1-75
Foxp3
GFP
MFI
GFP

Fig. S2

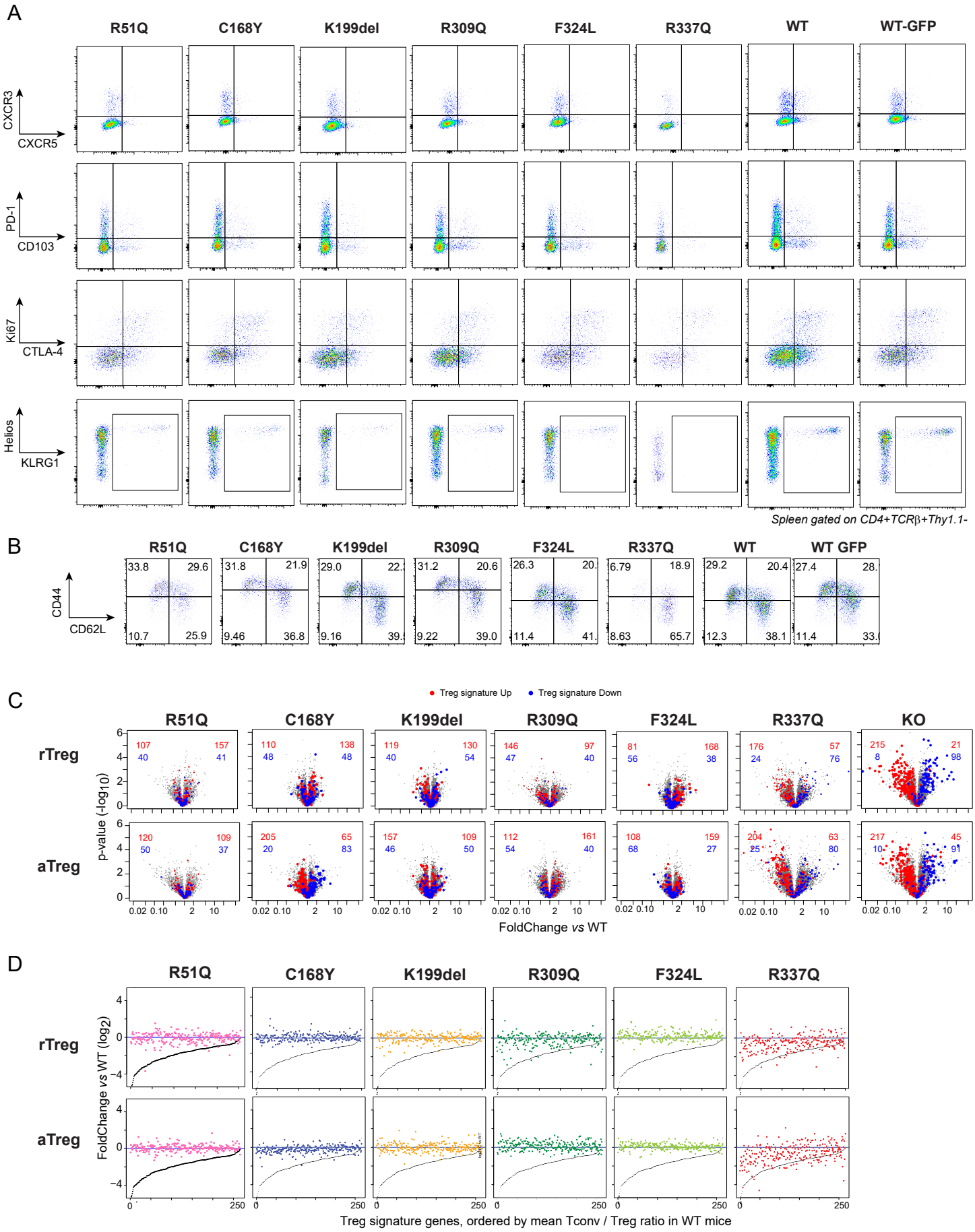

A
R51Q
C168Y
K199del
R309Q
F324L
R337Q
WT
WT-GFP
CXCR3
CXCR5
PD-1
CD103
Ki67
CTLA-4
Helios
KLRG1
Spleen gated on CD4+TCRβ+Thy1.1-
B
R51Q
C168Y
K199del
R309Q
F324L
R337Q
WT
WT GFP
CD44
CD62L
33.8
29.6
10.7
25.9
31.8
21.9
9.46
36.8
29.0
31.2
20.6
9.22
39.0
26.3
11.4
6.79
18.9
8.63
65.7
29.2
20.4
12.3
38.1
27.4
11.4
C
Treg signature Up
Treg signature Down
R51Q
C168Y
K199del
R309Q
F324L
R337Q
KO
rTreg
aTreg
p-value (-log10)
FoldChange vs WT
D
R51Q
C168Y
K199del
R309Q
F324L
R337Q
rTreg
aTreg
FoldChange vs WT (log2)
Treg signature genes, ordered by mean Tconv / Treg ratio in WT mice

Fig. S3

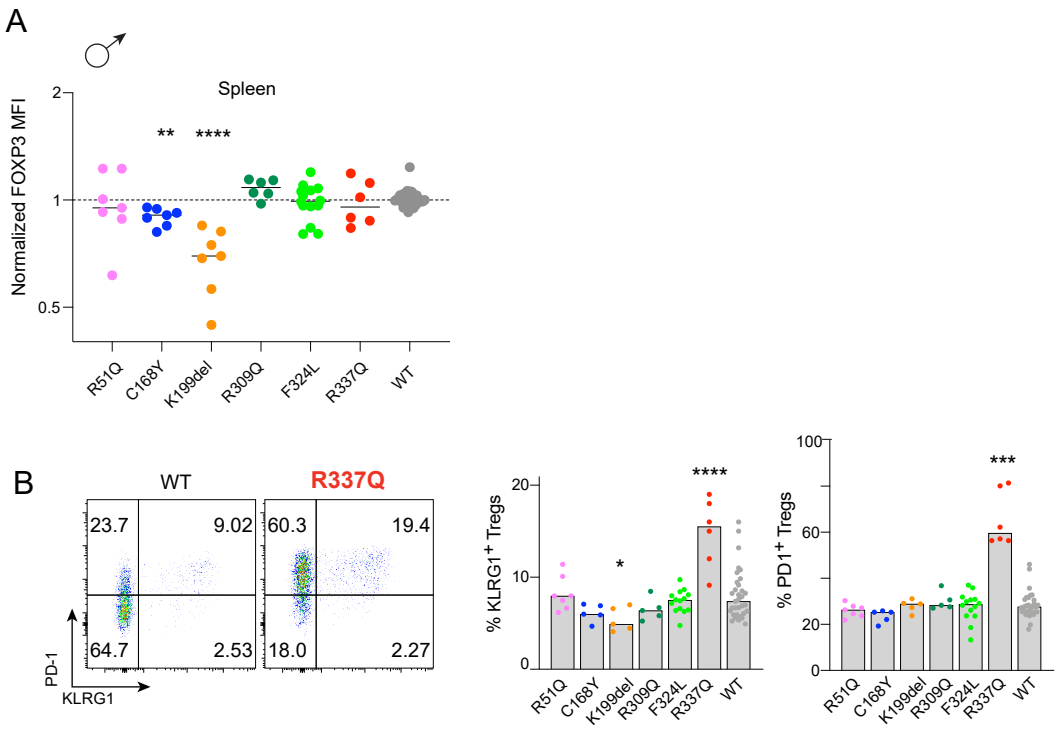

A
Spleen
Normalized FOXP3 MFI
**
****
R51Q
C168Y
K199del
R309Q
F324L
R337Q
WT
B
WT
R337Q
23.7
9.02
64.7
2.53
60.3
19.4
18.0
2.27
PD-1
KLRG1
% KLRG1+ Tregs
*
****
% PD1+ Tregs
***

Fig. S4

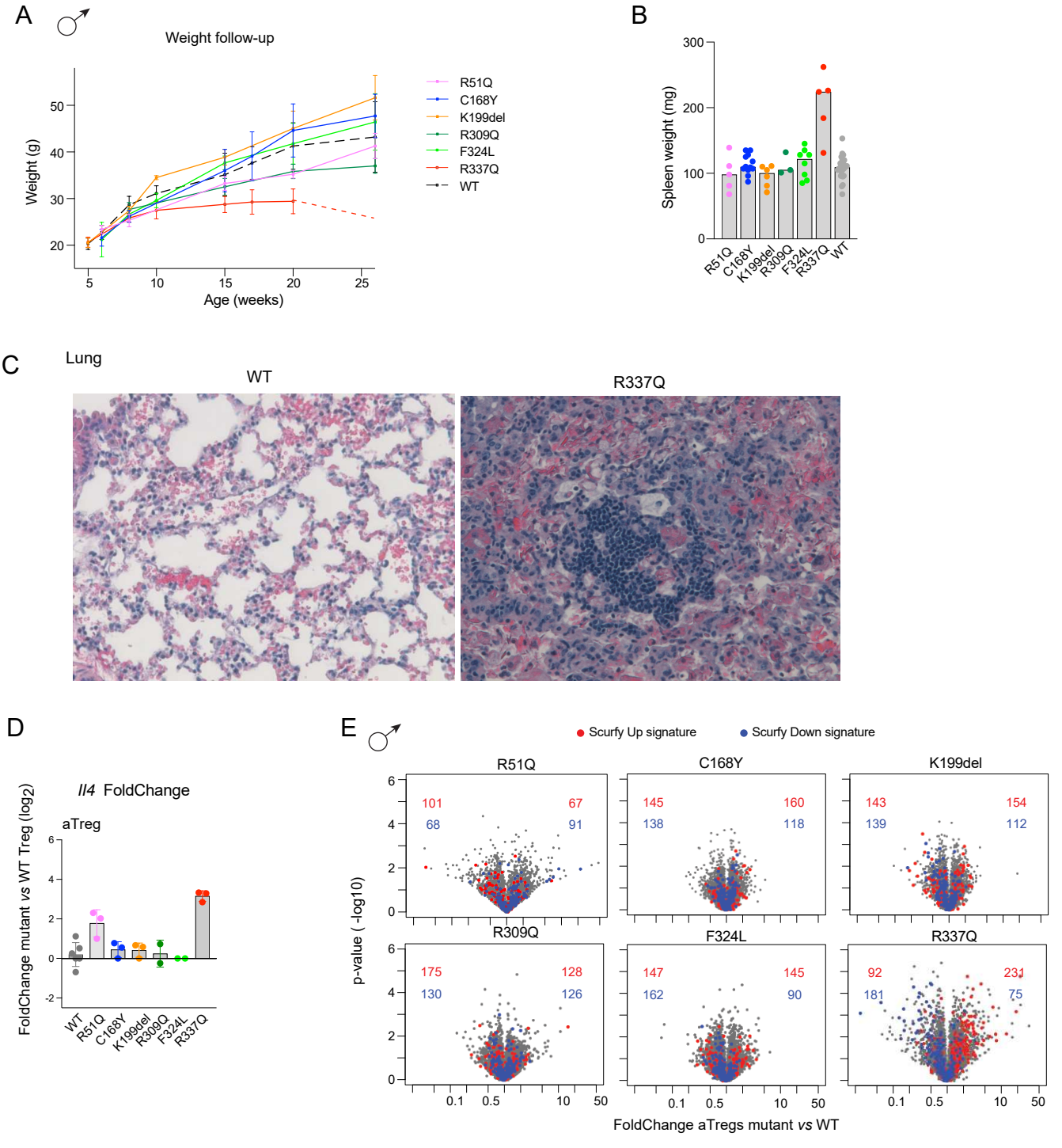

A
Weight follow-up
Weight (g)
Age (weeks)
R51Q
C168Y
K199del
R309Q
F324L
R337Q
WT
B
Spleen weight (mg)
C
Lung
WT
R337Q
D
Il4 FoldChange
aTreg
FoldChange mutant vs WT Treg (log2)
E
Scurfy Up signature
Scurfy Down signature
p-value (-log10)
FoldChange aTregs mutant vs WT

Fig. S5

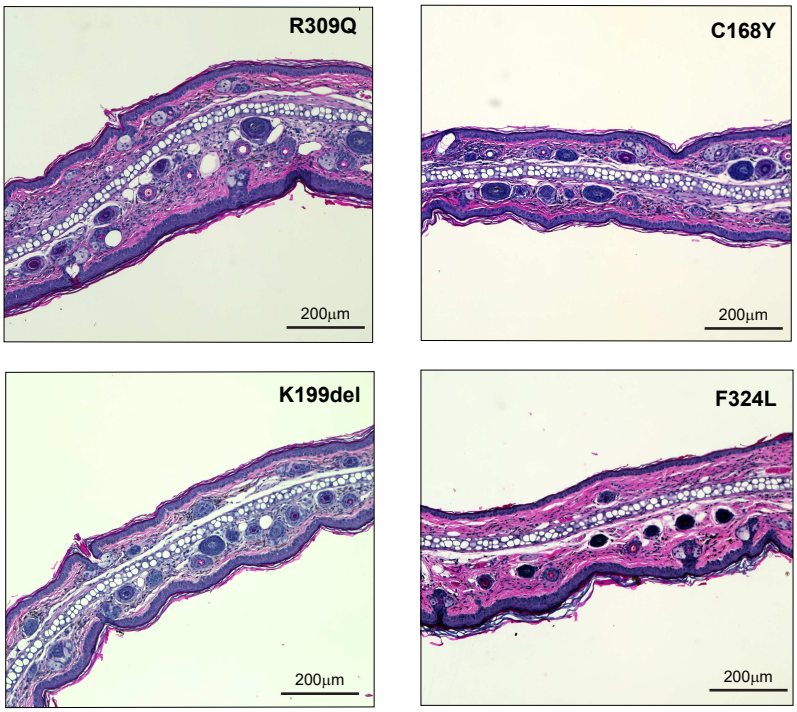


Fig. S6

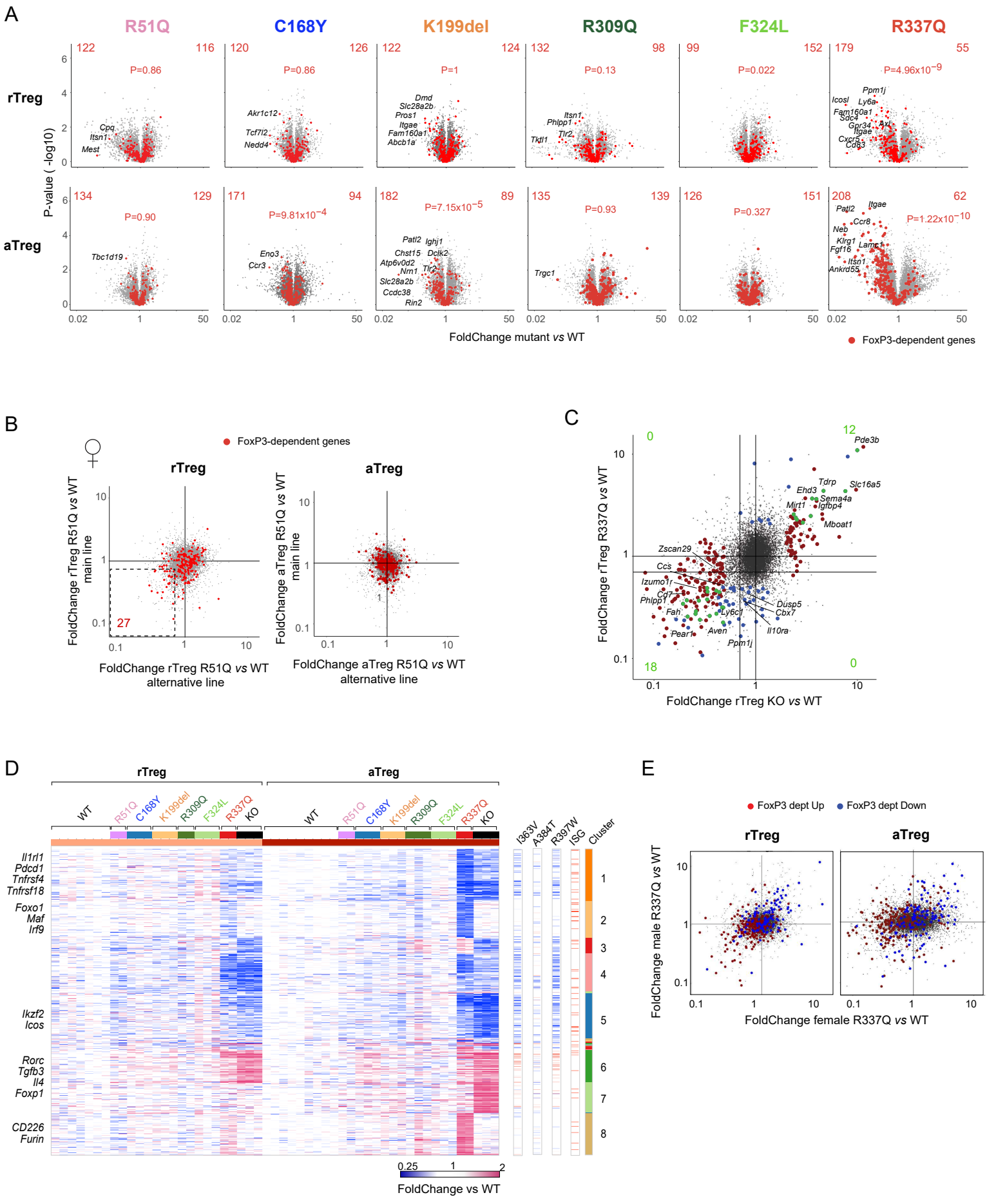
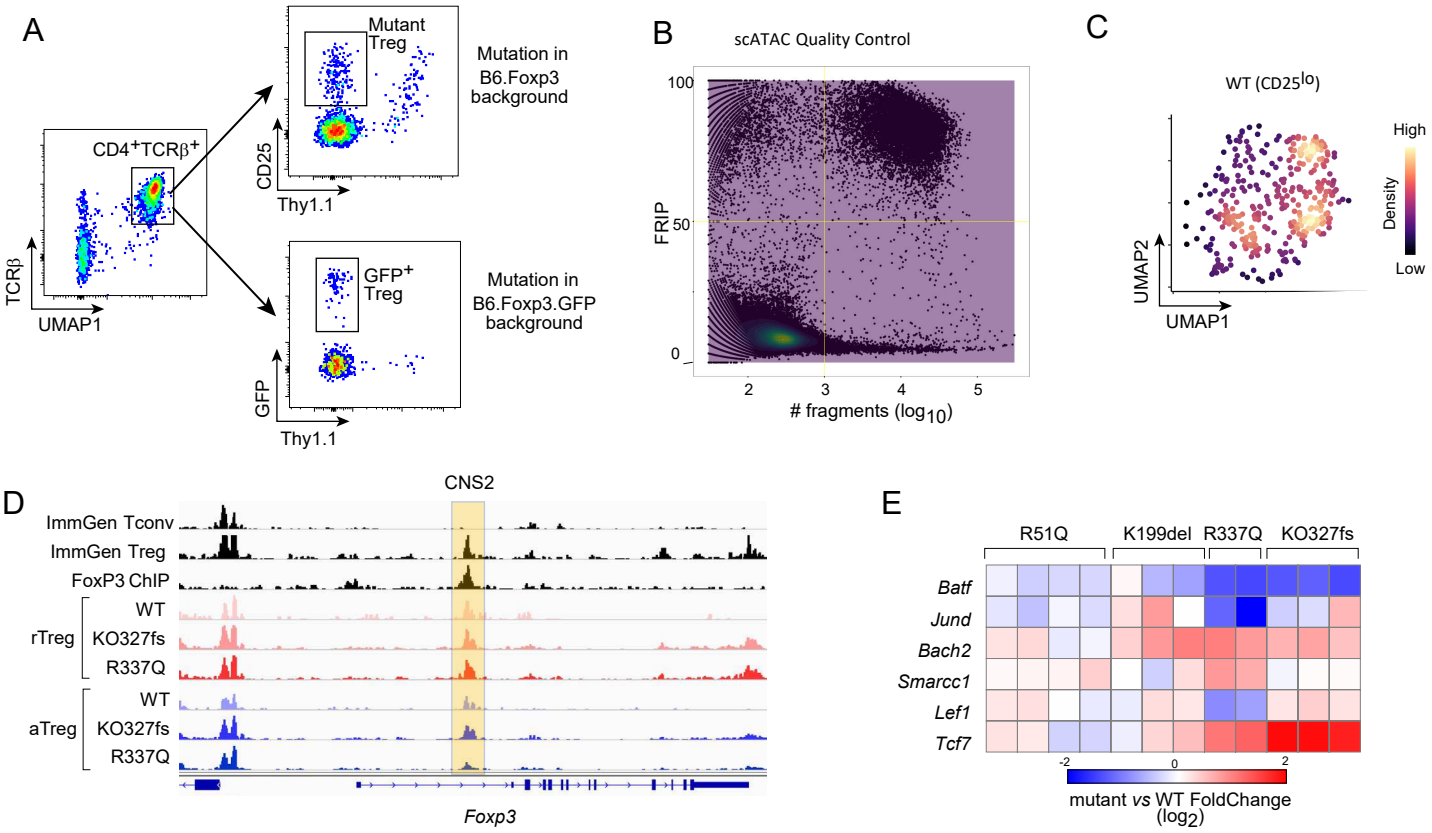

A
R51Q
C168Y
K199del
R309Q
F324L
R337Q
rTreg
aTreg
P-value (-log10)
FoldChange mutant vs WT
FoxP3-dependent genes
B
FoxP3-dependent genes
rTreg
aTreg
FoldChange rTreg R51Q vs WT main line
FoldChange rTreg R51Q vs WT alternative line
FoldChange aTreg R51Q vs WT main line
FoldChange aTreg R51Q vs WT alternative line
C
FoldChange rTreg R337Q vs WT
FoldChange rTreg KO vs WT
D
rTreg
aTreg
WT
R51Q
C168Y
K199del
R309Q
F324L
R337Q
KO
I363V
A384T
R397W
ISG
Cluster
FoldChange vs WT
E
FoxP3 dept Up
FoxP3 dept Down
rTreg
aTreg
FoldChange male R337Q vs WT
FoldChange female R337Q vs WT

Fig. S7


A
CD4+TCRβ+
TCRβ
UMAP1
Mutant Treg
CD25
Thy1.1
Mutation in B6.Foxp3 background
GFP+ Treg
GFP
Thy1.1
Mutation in B6.Foxp3.GFP background
B
scATAC Quality Control
FRIP
100
50
0
2
3
4
5
fragments (log10)
C
WT (CD25lo)
High
Density
Low
UMAP2
UMAP1
D
CNS2
ImmGen Tconv
ImmGen Treg
FoxP3 ChIP
rTreg
WT
KO327fs
R337Q
aTreg
WT
KO327fs
R337Q
Foxp3
E
R51Q
K199del
R337Q
KO327fs
Batf
Jund
Bach2
Smarcc1
Lef1
Tcf7
-2
0
2
mutant vs WT FoldChange (log2)

APPENDIX C, *REFERS TO CHAPTER 3*

An interwoven network of transcription factors, with divergent influences from FoxP3, underlies Treg diversity

Kaitavjeet Chowdhary[1], *Juliette Leon*[1,2], *Deepshika Ramanan*[1], *Diane Mathis*[1] *and Christophe Benoist*[1],+

[1] Department of Immunology, Harvard Medical School, Boston, MA, USA

[2] INSERM UMR 1163, University of Paris, Imagine Institute, Paris, France

[+]Correspondence to: Christophe Benoist
Department of Immunology
Harvard Medical School
77 Avenue Louis Pasteur, Boston, MA 02115
e-mail: cbdm@hms.harvard.edu

An interwoven network of transcription factors, with divergent influences from FoxP3, underlies Treg diversity

Kaitavjeet Chowdhary[1], *Juliette Léon*[1,2], *Deepshika Ramanan*[1], *Diane Mathis*[1] *and Christophe Benoist*[1,+]

[1]Department of Immunology, Harvard Medical School, Boston, MA, USA
[2] INSERM UMR 1163, University of Paris, Imagine Institute, Paris, France

+Address correspondence to: Christophe Benoist
Department of Immunology
Harvard Medical School
77 Avenue Louis Pasteur, Boston, MA 02115
e-mail: cbdm@hms.harvard.edu
Phone: (617) 432-7741

ABSTRACT

FoxP3+CD4+ regulatory T cells (Tregs), essential for immunologic and organismal homeostasis, have diverse functions and corresponding gene expression programs. How the many controlling transcription factors (TFs) organize to determine Treg identity and diversity remains unclear. We combined single-cell chromatin accessibility profiling, machine learning, and high-density natural genetic variation, validated with TF knockout, **CRISPR-editing**, and binding data, to define the Treg regulatory network. Distal enhancers proved driven by imbricated multi-TF inputs, employing strategies different from promoter regions. Topic modeling resolved a framework of chromatin programs shaped by distinct TF motifs. **This framework anchored surprisingly heterogenous responses to IL2. It identified an unrecognized role for the Smarcc1 remodeler.** FoxP3 impacted only some segments of this framework, either activating or repressing programs, **amplifying a core Treg identity defined independently. Its absence in Treg-like cells unleashed cytokine expression, but not Th de-differentiation.** This work provides a unifying scaffold to understand and manipulate Treg states.

INTRODUCTION

Cell identities are defined by characteristic gene expression programs. Each program is controlled by the action of specific transcription factors (TFs), which bind and regulate target *cis*-regulatory elements to effect changes in gene expression[1]. TFs act combinatorially, by binding to common regulatory elements, assembling into complexes, or organizing into transcriptional networks. In some systems, so-called "master TFs" sit at the apex of such regulatory networks[2-4]: their expression is both necessary and sufficient to initiate cell-type-specific programs, by launching feedback and feed-forward loops or antagonizing factors that promote alternative fates[1,5]. Other TFs are required not for cell-type specification but rather for responses to environmental signals[6]. While many studies of how TFs regulate cell identity have focused on transitions between differentiated cell types, how diversification is achieved within a single cell type has received less attention.

Regulatory T cells (Treg) are a subset of CD4+ T lymphocytes that act as dominant controllers of immunologic and organismal homeostasis. Humans and mice with dysfunctional Tregs due to loss-of-function mutations in Foxp3, the Treg lineage-defining TF, develop early-onset, uncontrolled autoimmunity and lymphoproliferation[7]. Tregs have diverse functions, distributed across varied phenotypic poles[8]. For example, distinct Treg programs marked by T-bet+CXCR3+ or IRF4+ phenotypes preferentially restrain Th1 or Th2 inflammation, respectively[9,10]. Non-lymphoid tissues harbor unique Treg populations[11], which enforce tolerance to commensal microbes[12], facilitate tissue regeneration[13], or control extra-immunologic consequences of inflammation[14]. Treg phenotypic specialization is undergirded by characteristic molecular programs[9,10,15-20]. This specialization is often considered in terms of a "one TF-one state" model[7], in which the expression of single context-specific TFs (e.g. T-bet, PPARγ, RORγ, cMAF, BATF), along with that of FoxP3, mediates differentiation of each Treg subpopulation[9,19,21-25], **although more combinatorial models have been considered as well[26,27].** However, how the many TFs expressed in Tregs are systematically organized to determine Treg identity and diversity remains unclear.

Although Foxp3 expression defines Treg identity, its mechanism of action has not been resolved. FoxP3 does not act as a pioneer factor, instead opportunistically binding to regions opened earlier in development[28,29]. Unlike traditional lineage-defining master regulators, FoxP3 is neither fully necessary nor sufficient to establish Treg identity: Treg-like cells ("Treg wannabes") can develop in the absence of FoxP3, and FoxP3-independent and -dependent modules characterize Treg-specific gene expression signatures[30-35]. FoxP3 interacts with a large array of

transcriptional regulators, chromatin remodelers, and TFs to program Treg identity[36-39]. However, the active structure of FoxP3, and even its recognition motif(s) in DNA are in question[40,41], and there is unsettled debate as to whether FoxP3 acts directly as an activator[31,42-44], a repressor[36,37,45,46], or both, depending on its interacting cofactors[39], or indirectly by tuning the expression of other TFs[47]. An integrative view of how FoxP3 affects TF control across diverse Treg states has so far been elusive.

Here, we integrate several orthogonal strategies to systematically connect TFs to their target Treg programs (Fig 1A). To study the most proximal effects of TF action without confounders from transcriptional bursting or stability, we focused on the role of TFs in modulating chromatin accessibility. Using single-cell chromatin-accessibility profiling, we found that diverse Treg programs were shaped not by individual master TFs, but rather by imbricated multi-TF inputs. Combining machine learning approaches, natural genetic variation, and Treg-specific TF knockouts (KOs), we parsed this combinatorial complexity to resolve the organization of the Treg genetic regulatory network (GRN). FoxP3 had profound and varied effects on this network, with differential influences on distinct chromatin programs. Collectively, these results offer a holistic and clarifying perspective on how one cell type coordinates combinations of TFs and *cis*-regulatory elements to achieve phenotypic diversification.

RESULTS

Imbricated transcription factor activities underlie Treg diversity

To gain a broad view of how accessibility of regulatory regions varies across diverse Treg subpopulations, we generated single-cell Assay for Transposase-Accessible Chromatin using sequencing (scATAC-seq) profiles from splenic Treg (TCRβ+CD4+GFP+) and T conventional (Tconv; TCRβ+CD4+GFP-) cells sorted from a male *Foxp3*$^{IRES\text{-}GFP}$ reporter mouse (Fig S1A)[48]. Data were of high quality (median 3x10^4 fragments per cell, Table S1) and after filtering, we retained profiles for 5,810 Treg and 1,654 Tconv cells (Fig S1B-E). Aggregation of reads from Treg and Tconv single cells recapitulated known patterns of cell-type specific chromatin accessibility at the *Foxp3* locus (Fig 1B). To enable comparisons, we mapped reads to a common peak set consisting of open chromatin regions (OCRs) from a pan-immune cell atlas[29] supplemented with new peaks identified in this dataset (Table S2).

Within the Treg pool, scATAC-seq profiles captured splenic Treg heterogeneity at multiple levels. Tregs separated broadly by activation status in a 2D uniform manifold approximation and projection (UMAP) visualization, as indicated by relative accessibility of OCR signatures that distinguish activated (aTreg) and resting (rTreg) populations (Fig 1C, Table S3)[47]. To more granularly annotate Treg diversity, we computed 'gene scores,' chromatin-based proxies for gene expression[49]. Confirming the chromatin signatures, rTreg and aTreg populations were reciprocally marked by high gene scores for *Ccr7* and *Sell* versus *Il10* (Fig 1D). Importantly, aTregs could be further delineated by gene scores for Treg functional molecules such as *Cxcr3*, *Klrg1*, and *Pdcd1*, each representing previously defined markers of distinct Treg poles[9,18,19,50]. We also noted a small but distinct cluster of cells with high *Rorc* (encoding RORγ) gene scores, corresponding to a Treg subset that dominates in the colon, but is also present at low levels in the spleen[24,25]. **Aggregated accessibility tracks per cell state confirmed the robust differential signals at these loci (Fig S2).**

With this landscape of Treg heterogeneity in our scATAC-seq data, we returned to our driving question, how TF activity relates to these Treg states. We examined the relative accessibility per cell of OCRs that contain known TF motifs, grouped into "archetypes"[51] to reduce redundancy (Fig 1E, S1F). Activation-related motifs had the greatest variability across single cells, cleanly partitioning the Treg pool (Fig S1G). For OCRs with AP-1, NF-AT, or NF-κB motifs, **accessibility was highest in aTregs, as expected since these are generically activation-related TFs. However, closer examination revealed imbricated arrangements: each motif was preferentially active in a slightly different region of the aTreg space, and each region**

included interlaced accessibility of multiple motifs. This perspective synthesized disparate results about individual TFs and their relevance to Treg physiology. For instance, although both c-Maf and RORγ have been reported to be active within the same population, Treg-specific c-Maf knockouts have broader phenotypes[52]. Accordingly, preferential accessibility of the NR/19 motif (corresponding to RORγ) was restricted to only a portion of the large swath of aTregs in which the MAF motif was preferentially accessible. Several studies have suggested BATF to be a main driver of tissue-Treg programs, but its ablation in Tregs affects only a fraction of tissue-Treg-related OCRs[19,53]. Here, relative accessibility of the BATF motif did not stand out from several other motifs with similar patterns. The activity of OCRs containing TCF/LEF was also curtailed in aTregs, consistent with the dampened expression of TCF-1 and Lef1, but not with the notion that TCF-1 would be a general FoxP3-controlled mediator of the Treg chromatin program across Treg states[47]. A "one TF-one state" model of Treg diversification would predict that individual TF motifs would be confined to discrete, mutually exclusive cell states. Instead, we found that imbricated and overlapping combinations of multiple factors constitute the diversity of Treg programs.

OCR usage in Treg single cells

How did this phenotypic variance at the cell level arise from the activity of individual OCRs? To understand the organization of regulatory element activity, we computed similarities in OCR usage across Treg and Tconv single cells, using cosine similarity between outputs from latent semantic indexing (LSI) of OCRs, and visualized them in a 2D UMAP projection (Fig 2A, Table S2). OCR activity patterns did not resolve into clearly defined classes, instead forming graded continua. Most OCRs overlapping or near transcription start sites (TSS) separated strikingly from those mapping further away ('distal' OCRs, mostly in enhancers) (Fig 2B, Fig S3A). Consistent with prior observations[29,54-56], TSS OCRs had the least variable accessibility (Fig 2C), confirming that distal OCRs contributed most to Treg diversity. Cell- and state-specific OCRs grouped together on the OCR UMAP visualization (Fig 2D, Table S3). This analysis therefore demonstrated that individual OCRs had unique and independent activity patterns and were not simply co-induced as discrete regulatory blocks.

Next, we asked how TFs interfaced with this organization of Treg OCRs. As noted in other contexts[51,57], individual OCRs contained motifs for several TFs, with a median of 8 motifs per OCR (Fig S3B). To connect TF binding to OCR activity, we positioned their motifs onto the OCR UMAP (Fig 2E). Motifs were not present at random, but demarcated distinct (STAT, NF-κB, and NR/19 (RORγ)) yet overlapping (NF-AT, AP-1, BATF) accessibility patterns in the OCR space. We also observed variegated configurations when projecting the binding sites of several TFs, deduced

from ChIP-seq and CUT&RUN experiments in Tregs[28,47,58-61] (Fig 2F, Table S4). TFs bound preferentially but not exclusively to distinct groups of state-specific OCRs, with different TFs occupying rTreg- (Lef1, TCF-1) or aTreg- (Bach2, JunD) biased loci. Binding distributions were dynamic: in unstimulated Tregs, the NF-κB component p65 bound primarily to TSS regions, but in stimulated cells shifted to occupy diverse distal regions (Fig 2F), including a cluster of OCRs enriched for NF-κB motifs (Fig 2F). FoxP3-binding had a narrow footprint on the OCR landscape (Fig 2F), with greatest enrichment among TSS and a group of rTreg-preferential distal OCRs, a concentration among clusters of STAT and NF-κB motif-containing OCRs, and diffuse minor representation among aTreg-specific loci. Thus, mirroring results at the cell level, each group of OCRs was occupied by overlapping combinations of TFs.

How did organization of OCR usage relate to gene expression? Most genes are controlled by multiple *cis*-regulatory elements[62-64]. This multiplicity confers molecular and evolutionary robustness, but also enables the expression of one gene in different differentiated cell types[65-67]. Our OCR UMAP allowed us to ask, within a single cell type, how OCRs linked to the same gene varied in their patterns of accessibility, and hence shared regulatory drivers. Several techniques have been proposed to connect OCRs to their target genes[29,64,68-71]. We formed OCR-gene links by using covariation of OCR accessibility with gene expression from paired splenic Treg scATAC and single cell transcriptomic (scRNA) datasets (FDR < 0.05)[71], also annotating OCRs within 15kb of each TSS although not meeting correlation criteria (Fig 2G, Fig S4, Table S5). Visualization of gene-linked OCRs on the OCR UMAP revealed divergent regulatory strategies. While some genes had OCRs with homogeneous activity patterns (e.g., *Rorc, Klrg1*), others had more varied distributions (e.g., *Ccr7*) or even multiple sub-patterns (e.g., *Tigit, Pdcd1*, *Ctla4*), with disparate accessibility profiles among loci previously shown to control *Foxp3* expression[72]. Thus, genes vary in how flexibly their associated OCRs are used across Treg states.

Topic modeling learns Treg chromatin programs

If individual TFs were insufficient to parse the imbricated organization of the Treg GRN, could one instead group patterns of OCR usage into co-regulated modules, and then determine the TFs that drove them? Each cell state within the Treg continuum could be conceptualized as the integration of multiple discrete regulatory programs. To learn such programs, we employed a machine learning approach, probabilistic topic modeling, derived from text mining, to categorize co-varying OCRs into "topics"[73]. Topic modeling, which has been applied to single cell genomic data[74,75], is well-suited for sparse and incomplete data and, unlike conventional partition clustering approaches, allows for the attribution of more than one program to each OCR, thus better

accommodating the continuous structure of the Treg OCR space. In practice, we modified a previous method for topic modeling of scATACseq data[74] for robustness by using an ensemble approach to create a consensus set of reproducible topics. The optimal solution parsed 17 such topics (Computational Note 1). Each topic learned by this strategy captured a different pattern of coordinated accessibility across Treg single cells. OCRs could belong to multiple topics, each topic being strongly represented in 1.5 to 3.9x10^4 OCRs (Computational Note 1, Table S6). OCRs from each topic occupied distinct poles of the OCR UMAP space (Fig 3A) and had diverse patterns of accessibility across Treg single cells (Fig 3B). Some topics, which included an overrepresentation of TSS OCRs (e.g., 6 and 8), were broadly accessible, while others captured highly specific gradations in accessibility. For example, Topics 3, 10, and 14 learned groupings of OCRs modulated in different facets of the aTreg pool, separate from Topic 9, which included OCRs active in RORγ+ cells.

Thus, topic modeling provided a quantitative approach to summarize patterns of Treg OCR accessibility and a tractable entry point to relate Treg chromatin programs to TF activity. To nominate candidate regulators of topic OCR accessibility, we computed the enrichment of TF motifs within each topic, yielding very different assignments for TSS vs distal OCRs (Fig 3C, Table S7). Distal OCR enrichments highlighted state-specific TF connections. While topics more accessible in aTregs were enriched in motifs for AP-1, NF-κB, and nuclear receptor factors, Tconv- and rTreg-preferential topics instead included KLF/SP, ETS, and TCF/LEF motifs (consistent with[47]). Some enrichments were highly specific: Gata3 only in Topic 2, RORγ only in Topic 9. Importantly, no topic was defined by enrichment of any one TF or TF family, indicating that this modular description of the Treg regulatory program itself required the combinatorial activity of different TFs. Consistent with their less variable accessibility, TSS OCRs were enriched in motifs represented in most topics, including ETS and KLF/SP family members.

To validate these links between TFs and topics, we made use of independent experimental TF-binding data. Using TFs for which high quality ChIP-seq or CUT&RUN profiles in Tregs were available, we compared the predicted topic-specific motif enrichments with the distribution of biochemical TF binding sites (Fig 3D). This overlap was highly significant for almost all TFs examined. Thus, enrichment of TF motifs within topics delineated the structure by which combinatorial TF inputs connected to diverse Treg epigenomic programs.

We used GREAT[76] to connect OCRs to genes, and identify functional pathways enriched in the regulatory regions defined by each topic (Fig 3E,Table S8). We will refrain from cherry-picking the results, but the over-arching conclusion was that the Topics did not merely bring up scattered functional annotations (which they might have) but different

biological programs – for instance, it is not as if all aTreg-associated Topics flagged the same activation-related gene sets. Topic 14 had distinctive associations related to regulation of other immunocytes.

Overall, these topics now provide a framework that can be re-applied to compare the state of the Treg GRN in different datasets and experiments. This is illustrated for several experiments below, and we have implemented a web application (https://cbdm.connect.hms.harvard.edu/Topic_Plotting/) that allows any investigator to upload a single-cell dataset and receive its decomposition in terms of these Topics.

Topics across tissues

Tregs traffic to and take residence in nonlymphoid tissues (tissue Tregs), where they elaborate specialized regulatory programs[11]. As the multi-step acquisition of tissue-Treg modules begins in lymphoid organs[16,17,19,20,50,77], it was important to know whether topics learned in spleen Tregs would capture tissue-Treg programs. To start, we looked for the distribution of a set of OCRs that we had previously reported as being induced in tissue Tregs (colon, visceral adipose tissue, muscle) relative to splenic Tregs[16] (Table S3). These pan-tissue-Treg OCRs were enriched specifically within Topics 3 and 14 (Fig 4A), with 40% of this OCR set overlapping with these two topics. To further this comparison, we generated new scATAC-seq profiles of Tregs isolated from spleen and colonic lamina propria from the same *Foxp3*$^{\text{IRES-GFP}}$ reporter mice (Fig S5A-B), multiplexed by hashtagging in the same run[78]. We computed the variance explained in this new dataset by the topics determined from our initial data. Highlighting the robustness of these topics, there was excellent concordance (Pearson r=0.99) between the relative variance explained across the new and previous spleen Treg scATAC datasets (Fig 4B). Spleen and colon Tregs differed in this regard, however. Topics 3 and 14 explained more variance (Fig 4C) and had increased accessibility within colonic Treg cells (Fig 4D), confirming the earlier enrichment results. Accessibility of Topics 3 and 14 was uniformly increased across colon Treg cells (Fig 4E). Several other topics, mostly biased towards rTreg (e.g., Topics 11, 12, 5), were decreased in accessibility and relative variance explained (Fig 4D, Fig S5D). Topics also captured more specific patterns of accessibility: for example, Topics 9 and 10 were preferentially active in RORγ+ and Helios+ Tregs, respectively (Fig 4F, Fig S5C-D, Table S3). Thus, topics were robust in their explanatory power, and captured tissue-specific specializations, confirming the notion that tissue-Treg programs represent amplifications of patterns already present in Tregs from lymphoid tissues[16,17,19,20,50,77]. Together with Fig. 3C, which indicates that Topics 3 and 14 were not driven by any one TF but

by an ensemble**, these results indicate that tissue-Treg chromatin programs are contained within the topic framework** and result from combinatorial opening.

Topics and the response to IL2

Having established the Treg chromatin programs at baseline in lymphoid and nonlymphoid tissues at steady state, it was of interest to assess how this network would adapt to acute stimulus. IL2 and the STAT family TFs that it predominantly activates, are canonical controllers of Treg cell differentiation in the thymus, and of the maintenance of FoxP3 expression and Treg numbers in the periphery[79-82], in a homeostatic negative feedback loop[83-86]. There have been detailed studies of IL2's transcriptional signature in Tregs[87,88].

We generated scATAC-seq profiles of splenic Tregs from mice treated acutely (2 hrs prior to favor direct effects without secondary confounders) with IL2 (Fig. S6A). Underscoring the potent effects of IL2, this short exposure led to a marked shift in Treg chromatin states (Fig 5A). After classifying cells as rTreg and aTregs based on their relative accessibility of OCR signatures from Fig 1 (Fig 5B, S6B), the UMAP visualization (Fig 5A,B, S6C) showed that rTreg populations shifted more than did aTreg cells in response to IL2 (median Local Inverse Simpson's Index[89] between treated and untreated 1.13 in rTreg vs 1.60 in aTreg, $p<2.2x10^{-16}$; Fig S6D). Visualization of the closest cells in treated vs untreated pools in high-dimensional OCR space (nearest neighbor in an LSI embedding with IL2 effect removed[89]) confirmed that all rTregs responded sharply but only a fraction of aTregs did (Fig. 5C). This lesser response of aTregs was also reflected by lower STAT5 phosphorylation after exposure to IL2 in culture (Fig. 5D), and explainable at least in part by a lower presence of the high affinity receptor for IL2, IL2RA, on the surface of aTregs (Fig. 5E), with lower accessibility of enhancers upstream and within the gene body of the *Il2ra* locus in aTregs (Fig. S6E). Induction of STAT motif-containing OCRs was seen across rTreg and aTreg responses (Fig S6F,G). aTregs had lower levels of STAT target accessibility at baseline and following stimulation (Fig S6G). However, while rTregs homogeneously increased their relative chromatin accessibility of STAT5-containing OCRs in response to IL2, aTregs responded heterogeneously (Fig. 5F). Smigiel et al previously reported a lower responsiveness to IL2 among $CD44^{hi}CCR7^{lo}$Tregs[90], attributing this low reactivity to an inability of those cells to use CCR7 to home to locales of high IL2 concentration. Since the injected IL2 in our experiments is diffusible and not constrained by local production, the present results indicate an intrinsically low response of aTregs to

IL2, relative to rTregs, and suggest that they may be less sensitive to IL2-centered therapeutic interventions.

As might be expected, no topic uniquely encompassed the response to IL2, and only a minority of OCRs were affected within each topic. Selecting robust responses (|log2FoldChange|>1.5 in either rTregs or aTregs; 860 OCRs), we grouped the different patterns of changes in accessibility (Fig. 5G, Table S9), and determined the match between Topics and these clusters (Fig. 5G, left panel and Table S9). Induced OCRs were specifically enriched for Topics 4 and 12. Topic 12 contains an interferon-responsive component (Fig 3E, Table S8), and its induction was consistent with recent data indicating that IL2 elicits a significant Interferon-Stimulated Gene response in Tregs[88], most likely because IL2 signals via STAT1 as well as STAT5. Motif enrichment showed these OCRs to be enriched for STAT5 and BCL6 motifs (Fig 5G, right panel and Table S9). Repressed OCRs, on the other hand, were enriched for Topic 16 regions across all cells, confirmed by an integrated analysis that also revealed a weaker but more general effect on Topic 17 OCRs (Fig. 5H). Topics 16 and 17 are chromatin programs preferentially active in Tconv, which suggests that IL2 bolsters Treg identity by suppressing Tconv-specific features, a notion of interest given IL2's role in supporting Treg differentiation in the thymus[85]. OCRs with greater reduction in accessibility among rTregs were enriched in aTreg-specific topics (Topics 3, 10, 14), and KLF/SP or RUNX motifs, whereas those with greater reduction in aTregs were enriched in Forkhead or ZF motifs. Thus, chromatin topics captured the specific response to IL2, described differences across cell states, and highlighted connections to different inductive and repressive mechanisms of IL2 action.

Natural genetic variation causally parses state- and OCR-specific TF effects

To bolster and validate these patterns derived from machine learning, we exploited naturally occurring genetic variation as an orthogonal approach to functionally establish causal relationships. Wild-derived Cast/Eij (Cast) mice differ from reference C57BL/6 (B6) mice by approximately 20 million variants[91]. As elegantly established in previous work[47,92-96], B6xCast F1 offspring can be used to causally link sequence variation to chromatin features: because the two genomes are present within the same cell, controlling for any changes in *trans* effects (i.e., TF expression), allelic skews in chromatin accessibility can be causally attributed to *cis*-regulatory alterations (i.e., disrupted TF motifs), the equivalent of a genome-wide mutagenesis experiment.

We generated scATAC profiles of Tregs sorted from B6xCast F1 mice (Fig S7A-B; n=5,980 Tregs). We avoided known biases in reference-specific read mapping by adapting a

published pipeline[97] to align reads to a common coordinate system, assigning informative reads to their allele of origin (Fig S7C). We adapted analytical frameworks established in previous studies ("mean diff" calculations[47,95]) to link changes in TF motifs with corresponding shifts in allele-specific chromatin accessibility ("allelic motif effect (AME)"), with an algorithmic modification to apply to the topic modeling context (Computational Note 2). For each topic, we identified the cells with accessibility of topic OCRs containing each candidate motif and computed the F1 AME for motifs in each topic only in these relevant cells (Fig 6A). Thus, in contrast to bulk analyses of F1 data that identify *average* effects of TFs genome-wide and in all cells, the resulting topic-specific AME quantified the causal contribution of TF motifs to each individual topic, only in the relevant cells in which the topic is active.

The results provided an unprecedented view of regulators of Treg chromatin programs, both in breadth and context-specificity. To identify the strongest modulators of accessibility, we highlighted effects detected in both motif enrichment and AME analyses (Fig 6B, Table S7). However, motifs with significant AMEs without corresponding enrichment may still reflect causal regulatory function (all AMEs in Fig S7D). The vast majority of AMEs corresponded to positive contributions of TF motifs to chromatin accessibility, very few motifs (e.g., Foxj1, Nfe2l1) having repressive effects, with members of the same TF family at times having opposing actions. While ETS, KLF/SP, and TCF/LEF motif families had detectable AMEs in several topics, their genetic effects overlapped with motif enrichments primarily in rTreg- and Tconv-biased topics. AMEs sharpened the scope of AP-1 and NF-κB effects, narrowing their multi-topic enrichment to a cluster of AMEs specific to aTreg-preferential Topics 3, 4, and 10. AMEs also identified restricted effects: for example, in accordance with its accessibility in the small group of RORγ+ cells, Topic 9 had a significant AME for the RORγ motif. Thus, integrating genetic effects with topic motif enrichment pierced through the combinatorial imbrication of the Treg regulatory network to refine links between TFs and Treg chromatin programs.

We validated this Treg TF network by testing with orthogonal *trans*-regulatory perturbations. First, we evaluated the impact of Treg-specific Gata3 ablation by generating scATAC profiles of Tregs from *Foxp3-cre*×*Gata3*$^{fl/fl}$ mice and *Foxp3-cre*×*Gata3*$^{+/+}$ littermates. Gata3-dependent OCRs defined from these data (GATA motif-containing OCRs with a greater than two-fold decrease in accessibility in the knockout; Table S3) were enriched only in Topic 2 (FDR < 0.01), the topic predicted to be under Gata3 control based on both AME and motif enrichment (Fig 6C). Secondly, we used published ATAC-seq data from Tregs sufficient or deficient in c-Maf[98]. In our model, MAF family motifs had significant AMEs in **Topics 9, 3, and 14** but enrichment only in Topics 3 and 14. Accordingly, c-Maf-dependent OCRs (Table S3) from the

knockout analysis were enriched only in Topics 3 and 14 (FDR < 0.01; Fig 6D). Thus, TF elimination in Tregs validated and matched predictions of TF-dependent control from our Treg TF network.

Thirdly, we tested the importance of a regulatory factor predicted by the model but not previously associated with Treg biology. Smarcc1 (BAF155) is a core subunit of all mammalian SWI/SNF chromatin remodeling complexes[99,100]. Some SWI/SNF components have been associated with changes during T cell activation[96,101] or regulation of *Foxp3* expression[102], but Smarcc1 has not been previously implicated in control of state-specific Treg chromatin. Our network predicted a specific effect of Smarcc1 in aTreg-skewed topics. To test its relevance, we delivered CRISPR-Cas9 ribonucleoprotein complexes[103] carrying *Smarcc1*-targeting or control gRNAs to CD4+ T cells isolated from *Foxp3*$^{IRES-GFP}$ mice, parked *in vivo* the edited cells in Treg-depleted *Foxp3*^{DTR} hosts, and sorted GFP+ Tregs 1 week later for bulk ATAC-seq analysis. OCRs which lost accessibility in Smarcc1 KO cells ($p<0.05$, Table S3) were specifically enriched (FDR<0.01) for Topic 3 OCRs, as predicted by our network (Fig 6E). Topic 10 was not represented, potentially reflecting redundancy between Smarcc1 and Smarcc2 paralogs[99,104]. Thus, based on the network predictions, we experimentally validated an unrecognized role for Smarcc1 in selectively controlling aTreg-specific chromatin programs.

FoxP3-independent and -dependent control of the Treg TF network

While the above provided an integrative view of TF control of Treg chromatin programs, the elephant in the room was FoxP3, the Treg lineage-defining TF, which did not appear in our network. This result is consistent with the uncertainties that surround the DNA motif(s) actually recognized by FoxP3[40,41] and the notion that FoxP3 is not a pioneer factor that modifies chromatin accessibility[28,29]. To understand the intrinsic role of FoxP3 in a setting unconfounded by systemic inflammation, we made use of female mice heterozygous for a Foxp3 loss-of-function allele (Foxp3^{fs327}-GFP/Foxp3-Thy1.1 mice, "KO" in Fig 7A)[44]. As Foxp3 is encoded on the X-chromosome, due to random X-inactivation, one population of Treg cells expresses wild-type FoxP3 protein (flagged by the Thy1.1 reporter), while another population of Treg-like cells expresses a Foxp3 allele with a full loss-of-function frameshift mutation whose expression is reported by GFP. The presence of functional Thy1.1+ Tregs prevents immune dysregulation, thus providing a well-controlled system for investigating FoxP3-intrinsic effects. Control mice (WT in Fig 7A) are similarly constructed, but with a functional FoxP3 encoded upstream of the GFP reporter.

We sorted GFP+ Treg and GFP- Tconv from both WT and KO heterozygous mice for scATAC-seq, hashtagging all cells into the same run by genotype and bins of CD25 expression (Fig 7A-B, Fig S8A-C). UMAP visualization of these data showed profound changes in the chromatin states of FoxP3-deficient GFP+ cells (Fig 7B, Fig S8D). While Tconv from WT and KO mice mostly co-mingled, as expected from the unperturbed environments, FoxP3-deficient Treg-like cells occupied a region of the embedding distinct from that occupied by WT Tregs (Fig 7B). For scale, the Local Inverse Simpson's Index was significantly lower ($p<2.2x10^{-16}$) for FoxP3 KO and WT cells (median=1.21) than for Gata3 KO and WT cells (median = 1.73) (Fig S8E). FoxP3-deficient Treg-like cells also showed diminished accessibility of aTreg-specific OCRs, most KO Tregs being in a resting-like chromatin state (Fig 7C). This inability of FoxP3-deficient Tregs to progress to activated states was confirmed by flow cytometry (Fig 7D, Fig S8F).

To better understand the OCRs driving this shift, we computed differential accessibility between Treg-like cells and WT Tregs, performing comparisons separately in rTreg and aTreg populations to avoid cell composition-driven effects (Table S3). We overlaid the differential accessibility from each comparison onto the OCR UMAP from Fig 2. FoxP3-repressed OCRs (lower accessibility in WT) were concentrated in rTreg- or Tconv-preferential loci, while FoxP3-potentiated OCRs (higher accessibility in WT) were in aTreg-specific regions (Fig 7E). Both at the level of individual OCRs (Fig 7E, S9A) and topics (Fig 7F), FoxP3-dependent changes were more pronounced in comparisons within aTregs than in rTregs**,** especially for FoxP3-potentiated OCRs. While FoxP3 binding was somewhat more enriched in repressed OCRs, most FoxP3-dependent changes were unrelated to FoxP3 binding (Fig 7E, Fig S9A), supporting previous results[28,47]. Compared with FoxP3-negative loci, FoxP3-bound OCRs were more widely accessible (Fig S9B) and less variable (Fig S9C) across Treg single cells.

How did FoxP3 influence the Treg GRN? For a comprehensive view, we looked for differential accessibility within motif-topic connections defined above (Fig 7G, Table S7). FoxP3 effects split into two groups: FoxP3 repressed ETS, KLF/SP, TCF/LEF, RUNX, and FOX motif accessibility in Tconv- and rTreg- preferential topics, but boosted accessibility of EGR, NF-κB, AP-1, and MAF motifs in aTreg-biased topics (Fig 7G, Fig S9D). Notably, some motifs with causal effects in several topics (e.g., CTCF, ETS, KLF/SP) were affected in opposite directions in different programs, highlighting that FoxP3 did not influence TFs homogenously genome-wide, and underscoring the power of parsing context-specificity with topic modeling. In short, this analysis suggested that FoxP3 had two roles: (1) repressing Tconv- and rTreg-like programs and (2) promoting activation-related chromatin features.

If FoxP3 does not entirely define Treg identity[31-35,105], what are the "Treg wannabes" that develop in its absence? We compared, in the Topic framework, KO Treg-like cells and Tconv. KO Treg-like cells most strongly induced Topic 4 (controlled by NF-κB, known to be important to Treg identity[59,106]) and repressed Tconv-preferential Topics 16 and 17 (Fig 7H). Strikingly, these FoxP3-independent effects (Fig. 7I, S9E) seemed a carbon-copy of the FoxP3-dependent effects observed when comparing WT and KO Tregs (Fig. 7G, Fig. S9F). Together, these results provide an integrated vista of Treg identity and FoxP3 function. Core features of Treg identity are established independently of FoxP3 and subsequently amplified by its expression. FoxP3 is then required for Treg activation, where FoxP3 protects Treg identity and enables the induction of aTreg-specific chromatin programs, which underly Treg suppressive and effector functions.

FoxP3 deficiency differentially affects Treg subsets in vivo

We noted an overrepresentation of cells with high *Rorc* gene scores and NR/19 motif accessibility among the FoxP3-deficient Treg-like population (Fig 8A, Fig S10A-B), which suggested a gut connection for these cells, since RORγ+ Tregs dominate in the colon[24,25]. The comparison of RORγ+ Treg proportions in KO Treg-like cells with Tregs from WT littermates showed a modest increase in RORγ+ cells across several organs, but a major shift in the colon, where almost all were RORγ+ (Fig 8B). In contrast, a drop in Helios+ Tregs was observed in the colon (Fig 8B-C), consistent with a drop in *Ikzf2* accessibility in the genomic data (Fig S10A-B). These results confirm recent observations from van der Veeken et al[107].

In normal mice, RORγ+ Tregs are microbiota-dependent[24,25], and we asked whether RORγ+ Tregs of KO mice were also regulated by bacterial inputs. After treatment with a broad-spectrum antibiotic cocktail (VNMA, 4 weeks), the fraction of total GFP+ Tregs decreased strongly in the KO colon, but not in WT (Fig. 8D) where the drop in RORγ+ Tregs was balanced by an increase of the Helios+ pool (Fig. 8E). The persisting Treg-like cells in KO mice remained RORγ$^+$, and no Helios+ Tregs emerged (Fig. 8E). Thus, confirming and extending the single-cell genomics, this analysis added another layer to the variegated function of FoxP3: RORγ$^+$ and Helios+ Treg populations depend differentially on FoxP3.

FoxP3-deficient RORγ+ Tregs in the colon produced IL17A (Fig. 8F), as also recently observed[107]. This was not the case in the spleen, indicating that the expression of Teff cytokines is not merely unleashed by the absence of FoxP3, but requires an active driver, present in the gut but not in the spleen (Fig 8F). **Did these FoxP3-deficient RORγ+ cells maintain their Treg identity, or were they turning into Th17? We performed scRNAseq on sorted GFP+ Treg**

and GFP- Tconv from colons of WT and KO heterozygous females (Fig 8G). In accordance with the scATACseq, colonic KO Tregs were shifted in the UMAP relative to WT Tregs (Fig 8G, S10C) and had high *Il17a* expression (Fig 8H). However, these KO Tregs clearly remained in the Treg space on the UMAP, far from the *Rorc+Il17a+* Th17 cells in the same dataset (Fig 8G-H, S10C). They retained the ability to express IL10 (Fig. 8J), maintained Treg signature scores closer to WT Treg than to Tconv (Fig 8I), and matched WT Tregs in their *Foxp3* locus accessibility pattern (Fig. S10D). Thus, KO Tregs did not become Th17 cells but maintained several facets of Treg identity while simply de-repressing cytokine production. Notably, another population of *Gata3+* KO Treg-like cells downregulated *Ikzf2* expression and expressed type 2 cytokine transcripts (*Il4, Il5, Il13*; Fig S10C), indicating parallel de-repression in *Rorc+* and *Gata3+* populations. In summary, while not required for RORγ+ Treg differentiation, FoxP3 was required to repress Teff cytokines induced in the gut environment.

DISCUSSION

A growing body of work has described a panoply of functions performed by Tregs and characterized their associated molecular programs[11,108]. Some transcriptional modules that accompany this specialization have been defined, but a systematic view of how TFs are integrated to control these programs has been missing. Here, we combined the informative power of single-cell genomics, machine learning, and high-density genetic variation, with validation by gene knockout and TF-binding datasets, to resolve the architecture of the Treg regulatory network. By avoiding confounders that result from averaging across all cells or enhancers, and by causally linking TF activities to specific chromatin subprograms, the results greatly simplify the nebulous complexities of Treg control and resolve some of the contradictions. We found that the diversity of Treg chromatin states arose from combinatorial, imbricated gradients of TF inputs. Against this background, FoxP3 had a profound impact, with a diversity of influences on different programs.

Treg enhancers were organized into a continuum, indicating the unique activity of each individual regulatory element across single cells (Fig 2). This was somewhat surprising, as one might have expected that enhancers operational within a defined cell-type would tend to fall into a limited number of classes. This continuum was structured, however: OCRs that bound a given TF largely congregated in particular sections of the OCR space. While the sparsity inherent to scATAC-seq may certainly contribute to the apparent continuity, one might also speculate that it represents cell-to-cell fluctuations in genomic activity (i.e. biological noise[109]) or dynamic responses to environmental signals[6]. Seemingly subtle modulation of T cell enhancers can lead to dramatically divergent immunological effects[110]. The fine-grained diversity in enhancer usage may enable Tregs to flexibly respond to varied and unpredictable immunological perturbations.

Conceptually or experimentally handling such a diversity of enhancer elements is challenging. Individual TFs, with their overlapping binding patterns, provide insufficient resolution. Topic modeling, although an approximation of these continuous activities, delivered a tractable abstraction through which to learn the drivers of major Treg regulatory programs, and single-cell resolution enabled identification of programs active only in rare subpopulations (e.g., RORγ+ spleen Tregs). Many motifs had inconstant effects across topics, indicating that TFs do not contribute homogeneously to chromatin activity across Treg cells. Such variation in TF operation indicated that Treg TF activity was highly context-specific, perhaps mediated by availability or activity of different cofactors. Parsing TF actions by topic should help resolve the ambiguities and contradictions around the function of individual factors, whose specific actions might otherwise have been obscured by averaging across cells or OCRs in population-level chromatin studies.

Topics provided a quantitative summary of Treg chromatin state that proved portable across replication studies and conditions. Topic activities were robustly reproducible and, although learned from baseline spleen data, captured programs amplified in non-lymphoid tissues and parsed heterogeneous responses to IL2 stimulation.

Genetic variation analysis enabled interrogation of causal relationships between TF motifs and Treg topics. Notably, each topic consisted of contributions from some motifs which were important for several topics and other motifs which were highly specific in their activity. This may suggest that the constellation of Treg chromatin programs is driven by combinations of a common set of TFs required for Treg identity overlaid by a specific set of TFs that enable Treg diversity, as observed in other contexts[96]. **The power of genetic variation to identify novel regulators was highlighted by our validation of a previously unrecognized role for Smarcc1in the control of a specific aTreg chromatin program.**

Motif enrichments were markedly different at OCRs in promoter regions vs distal enhancers. TSS OCRs involved a limited set of motifs, with topic diversification that might occur through coupling with distal sites in the same program, via physical enhancer-promoter interaction. Indeed, analysis of FoxP3-mediated enhancer-promoter loops previously showed distinct motif localization to either side of enhancer-promoter loop structures[44].

The combinatorial organization of TFs that control each program matched findings from several other fields and organismal settings[57,65,111] but challenged the persistent notion specific to the immunology literature that phenotypic specialization is mediated by single master TFs. For example, several reports have proposed that a particular TF critically controls programs relevant to tissue Tregs (e.g., BATF or BLIMP-1[19,53,112,113]), or controls discrete components (e.g., T-bet[9]). Tissue Treg programs did prove dominated by a pair of topics (3 and 14), which do include AP-1 and PRDM family motifs among their controlling regulators, but importantly not in a unique manner, several other TFs conspiring to control these tissue Treg topics. Accordingly, although relative accessibility of the BATF or PRDM1 motifs was highest in aTregs, their distributions overlapped with several other motifs (e.g., other AP-1, MAF, or NF-κB factors) with similar patterns. **We speculate that, by extension, the framework developed here might explain why factors identified as controlling pan-tissue-Treg chromatin programs affect only a portion of those OCRs[19,20].**

What is the mechanism by which FoxP3 controls Treg regulatory programs? Even in a setting unconfounded by inflammation, FoxP3 deficiency had a major impact on chromatin accessibility. While previous studies have debated whether FoxP3 acts as an activator or repressor, the present results highlight that it acts as both. Importantly, FoxP3's repressive and

activating functions affected distinct chromatin topics with different motif enrichments (Fig 6). The OCR membership of each of these topics provides genomic definition to the locations where FoxP3 might assemble into different nuclearly segregated molecular complexes with distinct co-factor composition and opposing functions[39].

The results also offer an integrated model for the relationship between FoxP3 and Treg identity. In rTregs, FoxP3 amplified a program of Treg identity present in "Treg wannabes", with a striking superimposition between the pre-existing program and what FoxP3 amplifies, down to the finest topic and motif details. How this concordance is achieved molecularly is an open question. One interpretation is that FoxP3 is uniquely attuned to cooperate with the TF ensemble that supports the Treg-like wannabe status, or that it replaces a similar Forkhead factor that would anchor the Tconv-wannabe distinction. Which factors specify the pre-FoxP3 Treg program will be an important area of future study, for which these data provides candidates. FoxP3 affected a greater number of OCRs within aTregs, suggesting an accentuated role for FoxP3 following Treg maturation. FoxP3 deficiency led to disappearance of aTreg populations and decreased accessibility of aTreg-specific OCRs, in accordance with[30,32,53]. In principle, this loss of aTregs could be because FoxP3 is required for aTreg differentiation[30], or because of a diversion of aTregs to an alternative fate (including death) in the absence of FoxP3[32,114]. These widened FoxP3 effects may depend on the availability of new cofactors in aTregs. Notably, the OCRs most strongly potentiated by FoxP3 belonged to tissue-Treg topics (3 and 14), consistent with the notion that FoxP3 is required for VAT Treg specification[17], but not colon RORγ+ Tregs[107].

The discovery of a population of FoxP3-independent RORγ+ Tregs added another layer to the relationship between FoxP3 and Treg differentiation. While this manuscript was in preparation, van der Veeken et al also reported FoxP3-independent RORγ+ Tregs[107]. We found FoxP3-deficient RORγ+ Treg-like cells to be microbe-dependent, as are normal FoxP3-sufficient counterparts. They produced IL17A but only in the colon**, and without diversion to Th17,** suggesting that cytokine de-repression in the absence of FoxP3 requires a positive environmental input, possibly via TCR or Wnt/β-catenin signaling, both of which have previously been associated with RORγ+IL17+ phenotypes[115-117]. However, extrinsic inputs cannot be the sole determinants of this program, as an intrinsic bias towards an increased RORγ/Helios ratio remained even in the context of antibiotic treatment. Thus, RORγ may partially compensate for the absent FoxP3 in these cells, ensuring their competitive survival over Helios+ counterparts, but unable to fulfill all of FoxP3's functions, like repression of cytokine genes upon stimulation.

Taken together, this study provides a model where TF function is conditioned by cell state, where specific combinations of TFs integrate input signals to enable Treg diversification. FoxP3 plays multimodal functions to maintain Treg identity and enable aTreg and consequently tissue-Treg-specific differentiation. In the colon, Tregs amplify pre-existing TF networks to adapt to new environments. Tissue-specific stimulation presents a unique challenge to self-reactive Tregs, where FoxP3 is important for balancing proportions of RORγ+ and Helios+ Tregs and preventing production of inflammatory cytokines. Overall, these results offer a quantitative and clarifying picture of the TF control of Treg diversity and provide a roadmap for its manipulation.

ACKNOWLEDGEMENTS

We thank R. Ramirez, S. Mostafavi, J. Buenrostro, A. Rudensky, J. van der Veeken, and Y. Zhong for helpful discussions; N. Ramirez, C. Gerhardinger at the Harvard Bauer core for scATAC-seq; the HMS Biopolymers Sequencing Facility; M. Sleeper, J. Nelson, D. Ischiu at the HMS Immunology Flow core; K. Hattori for assistance with mouse strains; C. Laplace for assistance with graphics; D. Mallah for assistance with web tool development. This work was funded by grants from the NIH to C.B. and D.M. (AI150686, AI165697, AI125603) and a grant from the JPB Foundation to C.B. and D.M. K.C was supported by NIGMS grants T32GM007753 and T32GM144273 and a Harvard Stem Cell Institute MD/PhD Training Fellowship, J.L. by an IMAGINE MD/PhD grant and an Arthur Sachs scholarship, and D.R. by the Damon Runyon Cancer Research Foundation (DRG 2300-17, National Mah Jongg League Fellow).

REFERENCES

1. Lee, T.I. & Young, R.A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237-1251 (2013).

2. Davis, R.L., Weintraub, H., & Lassar, A.B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987-1000 (1987).

3. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676 (2006).

4. Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035-1041 (2010).

5. Sorrells, T.R. & Johnson, A.D. Making sense of transcription networks. *Cell* **161**, 714-723 (2015).

6. Heinz, S., Romanoski, C.E., Benner, C., & Glass, C.K. The selection and function of cell type-specific enhancers. *Nat Rev Mol. Cell Biol* **16**, 144-154 (2015).

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

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

9. Koch, M.A. *et al.* The transcription factor T-bet controls regulatory T cell homeostasis and function during type 1 inflammation. *Nat. Immunol.* **10**, 595-602 (2009).

10. Zheng, Y. *et al.* Regulatory T-cell suppressor program co-opts transcription factor IRF4 to control T(H)2 responses. *Nature* **458**, 351-356 (2009).

11. Muñoz-Rojas, A.R. & Mathis, D. Tissue regulatory T cells: regulatory chameleons. *Nat Rev Immunol* **21**, 597-611 (2021).

12. Lathrop, S.K. *et al.* Peripheral education of the immune system by colonic commensal microbiota. *Nature* **478**, 250-254 (2011).

13. Burzyn, D. *et al.* A special population of regulatory T cells potentiates muscle repair. *Cell* **155**, 1282-1295 (2013).

14. Feuerer, M. *et al.* Lean, but not obese, fat is enriched for a unique population of regulatory T cells that affect metabolic parameters. *Nat Med* **15**, 930-939 (2009).

15. Feuerer, M. *et al.* Genomic definition of multiple ex vivo regulatory T cell subphenotypes. *Proc Natl Acad Sci U S A* **107**, 5919-5924 (2010).

16. Dispirito, J.R. *et al.* Molecular diversification of regulatory T cells in nonlymphoid tissues. *Sci Immunol* **3**, eaat5861 (2018).

17. Li, C. *et al.* TCR transgenic mice reveal stepwise, multi-site acquisition of the distinctive fat-Treg phenotype. *Cell* **174**, 285-299 (2018).

18. Zemmour, D. *et al.* Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nat Immunol* **19**, 291-301 (2018).

19. Delacher, M. *et al.* Precursors for nonlymphoid-tissue Treg cells reside in secondary lymphoid organs and are programmed by the transcription factor BATF. *Immunity.* **52**, 295-312 (2020).

20. Delacher, M. *et al.* Single-cell chromatin accessibility landscape identifies tissue repair program in human regulatory T cells. *Immunity.* **54**, 702-720 (2021).

21. Chung, Y. *et al.* Follicular regulatory T cells expressing Foxp3 and Bcl-6 suppress germinal center reactions. *Nat Med* **17**, 983-988 (2011).

22. Linterman, M.A. *et al.* Foxp3$^+$ follicular regulatory T cells control the germinal center response. *Nat Med* **17**, 975-982 (2011).

23. Wang, Y., Su, M.A., & Wan, Y.Y. An essential role of the transcription factor GATA-3 for the function of regulatory T cells. *Immunity* **35**, 337-348 (2011).

24. Sefik, E. *et al.* Individual intestinal symbionts induce a distinct population of RORγ^+ regulatory T cells. *Science* **349**, 993-997 (2015).

25. Ohnmacht, C. *et al.* The microbiota regulates type 2 immunity through RORγ^+ T cells. *Science* **349**, 989-993 (2015).

26. Fu, W. *et al.* A multiply redundant genetic switch 'locks in' the transcriptional signature of regulatory T cells. *Nat Immunol.* **13**, 972-980 (2012).

27. Schumann, K. *et al.* Functional CRISPR dissection of gene networks controlling human regulatory T cell identity. *Nat Immunol* **21**, 1456-1466 (2020).

28. Samstein, R.M. *et al.* Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell* **151**, 153-166 (2012).

29. Yoshida, H. *et al.* The cis-Regulatory Atlas of the Mouse Immune System. *Cell* **176**, 897-912 (2019).

30. Gavin, M.A. *et al.* Foxp3-dependent programme of regulatory T-cell differentiation. *Nature.* **445**, 771-775 (2007).

31. Hill, J.A. *et al.* Foxp3 transcription-factor-dependent and -independent regulation of the regulatory T cell transcriptional signature. *Immunity.* **27**, 786-800 (2007).

32. Lin, W. *et al.* Regulatory T cell development in the absence of functional Foxp3. *Nat. Immunol* **8**, 359-368 (2007).

33. Ohkura, N. *et al.* T cell receptor stimulation-induced epigenetic changes and Foxp3 expression are independent and complementary events required for Treg cell development. *Immunity* **37**, 785-799 (2012).

34. Charbonnier, L.M. *et al.* Functional reprogramming of regulatory T cells in the absence of Foxp3. *Nat. Immunol.* **20**, 1208-1219 (2019).

35. Zemmour, D. *et al.* Single-cell analysis of FOXP3 deficiencies in humans and mice unmasks intrinsic and extrinsic $CD4^+$ T cell perturbations. *Nat Immunol* **22**, 607-619 (2021).

36. Wu, Y. *et al.* FOXP3 controls regulatory T cell function through cooperation with NFAT. *Cell* **126**, 375-387 (2006).

37. Ono, M. *et al.* Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature* **446**, 685-689 (2007).

38. Rudra, D. *et al.* Transcription factor Foxp3 and its protein partners form a complex regulatory network. *Nat Immunol* **13**, 1010-1019 (2012).

39. Kwon, H.K., Chen, H.M., Mathis, D., & Benoist, C. Different molecular complexes that mediate transcriptional induction and repression by FoxP3. *Nat Immunol* **18**, 1238-1248 (2017).

40. Bandukwala, H.S. *et al.* Structure of a domain-swapped FOXP3 dimer on DNA and its function in regulatory T cells. *Immunity.* **34**, 479-491 (2011).

41. Leng, F. *et al.* The transcription factor FoxP3 can fold into two dimerization states with divergent implications for regulatory T cell function and immune homeostasis. *Immunity* **55**, 1354-1369 (2022).

42. Chen, C. *et al.* Transcriptional regulation by Foxp3 is associated with direct promoter occupancy and modulation of histone acetylation. *J Biol. Chem.* **281**, 36828-36834 (2006).

43. Li, B. *et al.* FOXP3 interactions with histone acetyltransferase and class II histone deacetylases are required for repression. *Proc Natl Acad Sci U S A* **104**, 4571-4576 (2007).

44. Ramirez, R.N. *et al.* FoxP3 associates with enhancer-promoter loops to regulate Treg-specific gene expression. *Sci Immunol* **7**, eabj9836 (2022).

45. Schubert, L.A. *et al.* Scurfin (FOXP3) acts as a repressor of transcription and regulates T cell activation. *J. Biol. Chem.* **276**, 37672-37679 (2001).

46. Arvey, A. *et al.* Inflammation-induced repression of chromatin bound by the transcription factor Foxp3 in regulatory T cells. *Nat Immunol* **15**, 580-587 (2014).

47. van der Veeken, J. *et al.* The transcription factor Foxp3 shapes regulatory T cell identity by tuning the activity of trans-acting intermediaries. *Immunity.* **53**, 971-984 (2020).

48. Bettelli, E. *et al.* Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. *Nature* **441**, 235-238 (2006).

49. Granja, J.M. *et al.* ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet* **53**, 403-411 (2021).

50. Li, C. *et al.* PPARγ marks splenic precursors of multiple nonlymphoid-tissue Treg compartments. *Proc Natl Acad Sci U S A* **118**, e2025197118 (2021).

51. Vierstra, J. *et al.* Global reference mapping of human transcription factor footprints. *Nature* **583**, 729-736 (2020).

52. Xu, M. *et al.* c-MAF-dependent regulatory T cells mediate immunological tolerance to a gut pathobiont. *Nature* **554**, 373-377 (2018).

53. Hayatsu, N. *et al.* Analyses of a mutant Foxp3 allele reveal BATF as a critical transcription factor in the differentiation and accumulation of tissue regulatory T cells. *Immunity* **47**, 268-283 (2017).

54. Heintzman, N.D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108-112 (2009).

55. Shen, Y. *et al.* A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116-120 (2012).

56. Corces, M.R. *et al.* Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat Genet.* **48**, 1193-1203 (2016).

57. Gerstein, M.B. *et al.* Architecture of the human regulatory network derived from ENCODE data. *Nature* **489**, 91-100 (2012).

58. Kitagawa, Y. *et al.* Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment. *Nat Immunol* **18**, 173-183 (2017).

59. Oh, H. *et al.* An NF-κB Transcription-Factor-Dependent Lineage-Specific Transcriptional Program Promotes Regulatory T Cell Identity and Function. *Immunity* **47**, 450-465 (2017).

60. Grant, F.M. *et al.* BACH2 drives quiescence and maintenance of resting Treg cells to promote homeostasis and cancer immunosuppression. *J. Exp. Med.* **217**, (2020).

61. Sidwell, T. *et al.* Attenuation of TCR-induced transcription by Bach2 controls regulatory T cell differentiation and homeostasis. *Nat. Commun.* **11**, 252 (2020).

62. ENCODE Project Consortium An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).

63. Osterwalder, M. *et al.* Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239-243 (2018).

64. Gasperini, M. *et al.* A Genome-wide Framework for Mapping Gene Regulation via Cellular Genetic Screens. *Cell* **176**, 377-390 (2019).

65. Spitz, F. & Furlong, E.E. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* **13**, 613-626 (2012).

66. Gasperini, M., Tome, J.M., & Shendure, J. Towards a comprehensive catalogue of validated and target-linked human enhancers. *Nat Rev Genet* **21**, 292-310 (2020).

67. Panigrahi, A. & O'Malley, B.W. Mechanisms of enhancer action: the known and the unknown. *Genome Biol* **22**, 108 (2021).

68. Fulco, C.P. *et al.* Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664-1669 (2019).

69. Ma, S. *et al.* Chromatin potential identified by shared single-cell profiling of RNA and chromatin. *Cell* **183**, 1103-1116 (2020).

70. Avsec, Z. *et al.* Effective gene expression prediction from sequence by integrating long-range interactions. *Nat Methods* **18**, 1196-1203 (2021).

71. Kartha, V.K. *et al.* Functional inference of gene regulation using single-cell multi-omics. *Cell Genom.* **2**, (2022).

72. Wang, K. & Fu, W. Transcriptional regulation of Treg homeostasis and functional specification. *Cell Mol. Life Sci* **77**, 4269-4287 (2020).

73. Blei, D.M., Ng, A.Y., & Jordan, M.I. Latent Dirichlet Allocation. *Journal of Machine Learning Research* **3**, 993-1022 (2003).

74. Bravo Gonzalez-Blas, C. *et al.* cisTopic: cis-regulatory topic modeling on single-cell ATAC-seq data. *Nat Methods* **16**, 397-400 (2019).

75. Bielecki, P. *et al.* Skin-resident innate lymphoid cells converge on a pathogenic effector state. *Nature* **592**, 128-132 (2021).

76. McLean, C.Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol.* **28**, 495-501 (2010).

77. Miragaia, R.J. *et al.* Single-cell transcriptomics of regulatory T cells reveals trajectories of tissue adaptation. *Immunity* **50**, 493-504 (2019).

78. Mimitou, E.P. *et al.* Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat Biotechnol.* **39**, 1246-1258 (2021).

79. Furtado, G.C., Curotto de Lafaille, M.A., Kutchukhidze, N., & Lafaille, J.J. Interleukin 2 signaling is required for CD4(+) regulatory T cell function. *J Exp Med* **196**, 851-857 (2002).

80. Almeida, A.R., Legrand, N., Papiernik, M., & Freitas, A.A. Homeostasis of peripheral CD4+ T cells: IL-2R alpha and IL-2 shape a population of regulatory cells that controls CD4+ T cell numbers. *J Immunol* **169**, 4850-4860 (2002).

81. Setoguchi, R., Hori, S., Takahashi, T., & Sakaguchi, S. Homeostatic maintenance of natural Foxp3(+) CD25(+) CD4(+) regulatory T cells by interleukin (IL)-2 and induction of autoimmune disease by IL-2 neutralization. *J Exp Med* **201**, 723-735 (2005).

82. Feng, Y. *et al.* Control of the inheritance of regulatory T cell identity by a cis element in the Foxp3 locus. *Cell* **158**, 749-763 (2014).

83. Fontenot, J.D. & Rudensky, A.Y. A well adapted regulatory contrivance: regulatory T cell development and the forkhead family transcription factor Foxp3. *Nat. Immunol* **6**, 331-337 (2005).

84. Malek, T.R. & Bayer, A.L. Tolerance, not immunity, crucially depends on IL-2. *Nat. Rev. Immunol* **4**, 665-674 (2004).

85. Hemmers, S. *et al.* IL-2 production by self-reactive CD4 thymocytes scales regulatory T cell generation in the thymus. *J Exp Med* **216**, 2466-2478 (2019).

86. Wong, H.S. *et al.* A local regulatory T cell feedback circuit maintains immune homeostasis by pruning self-activated T cells. *Cell* **184**, 3981-3997 (2021).

87. Moro, A. *et al.* Dynamic transcriptional activity and chromatin remodeling of regulatory T cells after varied duration of interleukin-2 receptor signaling. *Nat Immunol*(2022).

88. Baysoy, A. *et al.* The interweaved signatures of common-gamma-chain cytokines across immunologic lineages. *J Exp Med* **220**, (2023).

89. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289-1296 (2019).

90. Smigiel, K.S. *et al.* CCR7 provides localized access to IL-2 and defines homeostatically distinct regulatory T cell subsets. *J Exp. Med* **211**, 121-136 (2014).

91. Keane, T.M. *et al.* Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289-294 (2011).

92. Heinz, S. *et al.* Effect of natural genetic variation on enhancer selection and function. *Nature* **503**, 487-492 (2013).

93. Vierbuchen, T. *et al.* AP-1 Transcription Factors and the BAF Complex Mediate Signal-Dependent Enhancer Selection. *Mol. Cell* **68**, 1067-1082 (2017).

94. Link, V.M. *et al.* Analysis of genetically diverse macrophages reveals local and domain-wide mechanisms that control transcription factor binding and function. *Cell* **173**, 1796-1809 (2018).

95. van der Veeken, J. *et al.* Natural genetic variation reveals key features of epigenetic and transcriptional memory in virus-specific CD8 T cells. *Immunity.* **50**, 1202-1217 (2019).

96. Zhong, Y. *et al.* Hierarchical regulation of the resting and activated T cell epigenome by major transcription factor families. *Nat Immunol* **23**, 122-134 (2022).

97. Huang, S. *et al.* A novel multi-alignment pipeline for high-throughput sequencing data. *Database. (Oxford)* **2014**, (2014).

98. Neumann, C. *et al.* c-Maf-dependent Treg cell control of intestinal TH17 cells and IgA establishes host-microbiota homeostasis. *Nat Immunol* **20**, 471-481 (2019).

99. Mashtalir, N. *et al.* Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. *Cell* **175**, 1272-1288 (2018).

100. Centore, R.C. *et al.* Mammalian SWI/SNF Chromatin Remodeling Complexes: Emerging Mechanisms and Therapeutic Strategies. *Trends Genet* **36**, 936-950 (2020).

101. Jeong, S.M. *et al.* The SWI/SNF chromatin-remodeling complex modulates peripheral T cell activation and proliferation by controlling AP-1 expression. *J Biol Chem.* **285**, 2340-2350 (2010).

102. Loo, C.S. *et al.* A Genome-wide CRISPR Screen Reveals a Role for the Non-canonical Nucleosome-Remodeling BAF Complex in Foxp3 Expression and Regulatory T Cell Function. *Immunity* **53**, 143-157 (2020).

103. Seki, A. & Rutz, S. Optimized RNP transfection for highly efficient CRISPR/Cas9-mediated gene knockout in primary T cells. *J Exp Med* **215**, 985-997 (2018).

104. Otto, J.E. *et al.* Structural and functional properties of mSWI/SNF chromatin remodeling complexes revealed through single-cell perturbation screens. *Mol. Cell* **83**, 1350-1367 (2023).

105. Sugimoto, N. *et al.* Foxp3-dependent and -independent molecules specific for CD25+CD4+ natural regulatory T cells revealed by DNA microarray analysis. *Int. Immunol* **18**, 1197-1209 (2006).

106. Owen, D.L. *et al.* Thymic regulatory T cells arise via two distinct developmental programs. *Nat. Immunol.* **20**, 195-205 (2019).

107. van der Veeken, J. *et al.* Genetic tracing reveals transcription factor Foxp3-dependent and Foxp3-independent functionality of peripherally induced Treg cells. *Immunity* **55**, 1173-1184 (2022).

108. Georgiev, P., Charbonnier, L.M., & Chatila, T.A. Regulatory T Cells: the Many Faces of Foxp3. *J Clin. Immunol* **39**, 623-640 (2019).

109. Spudich, J.L. & Koshland, D.E., Jr. Non-genetic individuality: chance in the single cell. *Nature.* **262**, 467-471 (1976).

110. Simeonov, D.R. *et al.* T cell subset-selective *IL2RA* enhancers shape autoimmune diabetes risk. *bioRxiv*2020 (20 A.D.).

111. Mitchell, P.J. & Tjian, R. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371-378 (1989).

112. Cretney, E. *et al.* The transcription factors Blimp-1 and IRF4 jointly control the differentiation and function of effector regulatory T cells. *Nat Immunol.* **12**, 304-311 (2011).

113. Vasanthakumar, A. *et al.* The transcriptional regulators IRF4, BATF and IL-33 orchestrate development and maintenance of adipose tissue-resident regulatory T cells. *Nat. Immunol.* **16**, 276-285 (2015).

114. Tai, X. *et al.* Foxp3 transcription factor is proapoptotic and lethal to developing regulatory T cells unless counterbalanced by cytokine survival signals. *Immunity* **38**, 1116-1128 (2013).

115. Keerthivasan, S. *et al.* β-Catenin promotes colitis and colon cancer through imprinting of proinflammatory properties in T cells. *Sci Transl. Med* **6**, 225ra28 (2014).

116. Osman, A. *et al.* TCF-1 controls T_{reg} cell functions that regulate inflammation, $CD8^+$ T cell cytotoxicity and severity of colon cancer. *Nat Immunol* **22**, 1152-1162 (2021).

117. Quandt, J. *et al.* Wnt-β-catenin activation epigenetically reprograms Treg cells in inflammatory bowel disease and dysplastic progression. *Nat. Immunol.* **22**, 471-484 (2021).

118. Rubtsov, Y.P. *et al.* Regulatory T cell-derived interleukin-10 limits inflammation at environmental interfaces. *Immunity* **28**, 546-558 (2008).

119. Zhu, J. *et al.* Conditional deletion of *Gata3* shows its essential function in T_H1-T_H2 responses. *Nat Immunol* **5**, 1157-1165 (2004).

120. Ramanan, D. *et al.* An immunologic mode of multigenerational transmission governs a gut Treg setpoint. *Cell* **181**, 1276-1290 (2020).

121. Tuncel, J., Benoist, C., & Mathis, D. T cell anergy in perinatal mice is promoted by T reg cells and prevented by IL-33. *J Exp. Med* **216**, 1328-1344 (2019).

122. Stuart, T. *et al.* Single-cell chromatin state analysis with Signac. *Nat Methods* **18**, 1333-1341 (2021).

123. Cusanovich, D.A. *et al.* A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility. *Cell* **174**, 1309-1324 (2018).

124. Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902 (2019).

125. McInnes, L., Healy, J., & Mellvile, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv. org* **arXiv:1802.03426v2**, (2018).

126. Melsted, P. *et al.* Modular, efficient and constant-memory single-cell RNA-seq preprocessing. *Nat Biotechnol.* **39**, 813-818 (2021).

127. Booeshaghi, A.S., Min, K.H., Gehring, J., & Pachter, L. Quantifying orthogonal barcodes for sequence census assays. *bioRxiv*2022 (2022).

128. Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* **14**, 865-868 (2017).

129. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 (2021).

130. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

131. Schep, A.N., Wu, B., Buenrostro, J.D., & Greenleaf, W.J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975-978 (2017).

132. Sassone-Corsi, M. *et al.* Sequestration of gut pathobionts in intraluminal casts, a mechanism to avoid dysregulated T cell activation by pathobionts. *Proc Natl Acad Sci U S A* **119**, e2209624119 (2022).

133. Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160-W165 (2016).

134. Robinson, J.T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24-26 (2011).

135. Li, D. *et al.* WashU Epigenome Browser update 2022. *Nucleic Acids Res.*(2022).

136. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576-589 (2010).

137. Traag, V.A., Waltman, L., & van Eck, N.J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* **9**, 5233 (2019).

138. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).

139. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* **25**, 2078-2079 (2009).

140. Li, Q., Brown, J.B., Huang, H., & Bickel, P.J. Measuring reproducibility of high-throughput experiments. *The Annals of Applied Statistics* **5**, 1752-1779 (2011).

141. Yang, M.G. *et al.* Characterization of sequence determinants of enhancer function using natural genetic variation. *eLife* **11**, (2022).

142. Gu, Z. & Hubschmann, D. rGREAT: an R/bioconductor package for functional enrichment on genomic regions. *Bioinformatics.* **39**, (2023).

143. Love, M.I., Huber, W., & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

144. R Development Core Team R: a language and environment for statistical computing. *Vienna, Austria: the R Foundation for Statistical Computing.*Available online at http://www.R-project.org (2017).

145. Wickham,H. *ggplot2: Elegant Graphics for Data Analysis*(Springer-Verlag, New York, 2009).

146. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**, 90-95 (2007).

MATERIALS & METHODS

Mice

Foxp3$^{\text{IRES-GFP}}$ mice[48] and *Foxp3*$^{fs327\text{-GFP}}$ x *Foxp3*$^{tm10.1(\text{Casp9,-Thy1})\text{Ayr}}$ (or Foxp3$^{\text{Thy1.1}}$)[44] heterozygous female mice were bred on the C57Bl/6J background and maintained in our colony in the specific pathogen–free facility at Harvard Medical School (HMS). Gata3 conditional knockout mice were generated by crossing Cre+ *Foxp3-cre* mice[118] with heterozygous Gata3$^{fl/+}$ mice[119]. To generate B6/Cast F1 mice, Cast/EiJ males (Jackson Labs, strain #000928) were crossed with C57BL/6J females (Jackson Labs, strain #000664). All experimentation was performed following animal protocols approved by the HMS Institutional Animal Use and Care Committee (protocols IS00000054 and IS00001257). Except when specified, 6-10-week-old male mice were used throughout this study. For heterozygous FoxP3 KO experiments, 5-10-week-old female mice were used. Experiments involving knockout mice always used WT littermate controls for comparisons.

Isolation, analysis, and sorting of T lymphocytes

Spleen and lymph nodes: Immunocytes were released using mechanical disruption followed by filtering and washes in phenol red-free DMEM containing 10mM HEPES (Gibco) and 2% fetal calf serum (FCS). Red blood cells in spleen samples were lysed before filtering using ACK lysis buffer (Gibco, ref A10492-01). In some samples, CD4+ T cells were enriched using negative magnetic selection using the Dynabeads Untouched Mouse CD4 Cells Kit (ThermoFisher Scientific, cat#11415D).

Colon: Immunocytes were isolated as previously described[120]. Briefly, intestines were cleaned (Peyer's patches removed in the case of the small intestine), and treated with RPMI containing 1 mM DTT, 20 mM EDTA and 2% FCS at 37°C for 15 min to remove epithelial cells. They were then minced and dissociated in collagenase solution (1.5 mg/mL collagenase II (Gibco), 0.5mg/mL dispase (Gibco) and 1% FCS in RPMI) with constant stirring at 37°C for 40 min. Single cell suspensions were filtered and washed with RPMI containing 5% FCS.

Lungs: Immunocytes were isolated as previously described[121]. Briefly, mice were first perfused with 5mL of ice-cold PBS through the heart's right ventricle. Tissues were minced and dissociated in collagenase solution (1 mg/mL collagenase IV (Gibco), 150μg/mL DNase I (Sigma) and 1% FCS in DMEM) and incubated in a water bath at 37°C with constant shaking for 30 min. Digested tissues were filtered and washed in 2% FCS. For lungs, red blood cells were lysed using ACK lysis buffer.

Flow Cytometry: After Fc blocking and Live/Dead staining (Zombie Fixable Viability Kit, Biolegend), extracellular staining was done in ice cold phenol red-free DMEM containing 2% FBS for 30 min using antibodies against surface markers (see antibodies below). In order to maintain GFP signal, cells were first pre-fixed in 1% formaldehyde for 15 min at room temperature and then fixed overnight at 4°C using 100 µL of Fix/Perm buffer (eBioscience FoxP3/TF Staining Buffer Set, cat# 00-5523-00). After membrane permeabilization using 1X permeabilization buffer (eBioscience FoxP3/TF Staining Buffer Set, cat# 00-5523-00) for 5 min, intracellular staining was performed for 2 hours at room temperature (see antibodies below). Data was recorded on a FACSymphonyTM (BD Biosciences) or on an Aurora (Cytek Biosciences) flow cytometer and analyzed using FlowJo 10 software.

Sorting: Cells were sorted using BD MoFlo Astrios EQ, FACSAria-561, or FACSAria-594 machines. For experiments with *Foxp3*$^{IRES\text{-}GFP}$ mice, Tregs were sorted as DAPI- Dump (CD19, Cd11c, Cd8)- CD4+ TCRb+ GFP+ and Tconv (where applicable) as DAPI- Dump (CD19, Cd11c, Cd8)- CD4+ TCRb+ GFP-. For experiments using mice without Foxp3 reporters, Tregs were sorted as DAPI- Dump (CD19, CD11c, CD8)- CD4+ TCRb+ CD25hi and Tconv (where applicable) as DAPI- Dump (CD19, CD11c, CD8)- CD4+ TCRb+ CD25lo. For FoxP3 heterozygous female scATAC-seq experiments, KO or WT Treg cells were sorted as DAPI- Dump (CD19, CD8)- CD4+ TCRb+ Thy1.1- GFP+, separating cells by bins of CD25 expression (2 bins for KO and 3 bins for WT cells see Fig S6A-C).

Antibodies: For sorting or flow analyses of lymphoid and non-lymphoid tissues, the following antibodies were used (BioLegend): CD4-APC (clone GK1.5) or CD4-BV605 (clone RM4-5), TCRb-PECy7/BUV737/AF700 (clone H57-597), CD25-PE/APC (clone PC-61), CD19-FITC/PB (clone 6D5), CD8a-FITC/PB (clone 53-6.7), CD11b-FITC (clone M1/70), CD44-BV510 (clone IM7), CD62L-BV785 (clone MEL-14), CD45.2-AF700 (clone 104), Thy1.1-PE/PECy7/APC (clone OX-7), RORγ PE/APC/BV785 (clone AFKJS-9), Helios-PB (clone 22F6), IL17A-PE/APC (clone TC11-18H10.1).

Antibiotic Treatment

For antibiotic treatment, vancomycin (0.5 g/L, VWR Life Science), metronidazole (1 g/L, Sigma), neomycin (1 g/L, Fisher), and ampicillin (1g/L, Sigma) were dissolved in drinking water with 0 calorie sweetener and given to mice for 4 weeks, with replenishment of fresh antibiotics every 7-10 days.

Cytokine Staining

For *ex-vivo* cytokine staining, single-cell suspensions of colon and spleen were stimulated for 3 h at 37°C with 50 ng/ml PMA (Sigma-Aldrich) and 1 µM ionomycin (Sigma-Aldrich) in the presence of protein transport inhibitor cocktail (eBioscience) in complete RPMI 1640 supplemented with 10% FBS (Thermo Fisher Scientific). Extracellular surface antigen staining and intracellular transcription factor and cytokine staining were performed as described above.

IL2 Experiments

IL2 injections

For scATAC-seq of IL2 treated Tregs, 6-week-old *Foxp3*$^{IRES\text{-}GFP}$ mice were injected intravenously with 10 µg of recombinant mouse IL2 (PeproTech, 212-12) or PBS vehicle control (in 110 µl volume). Mice were euthanized exactly 2 hours after injection for isolation and processing of splenic Tregs, as described in previous sections.

pSTAT5 Flow Cytometry

For *ex vivo* quantification of STAT5 phosphorylation, after mechanical dissociation and ACK red blood cell lysis, single-cell suspensions of spleens from 10-week-old *Foxp3*$^{Thy1.1}$ mice were stimulated for 15 minutes at 37°C with the indicated concentration of recombinant mouse IL2 (PeproTech, 212-12) in serum-free RPMI 1640. Stimulation was quenched by cell fixation at a final concentration of 2% PFA on ice for 30 minutes. Afterwards, cells were washed twice with ice-cold PBS and permeabilized with pre-chilled 90% methanol on ice for 30 minutes. Cells were washed and stained with pSTAT5 (Tyr694) antibody (1:20, BioLegend; clone A17016B.Rec) at room temperature for 40 minutes, at which point, surface marker antibodies were added to samples for additional 25 minutes staining, prior to washing, resuspension, and flow cytometry data collection. Cells were distinguished as aTreg vs rTreg using CD44 and CD62L staining.

CRISPR/Cas9 RNP Editing and Transfer

We adapted a previous protocol for gene editing using CRISPR/Cas9 ribonucleoprotein (RNP) complex nucleofection[103] for application to primary mouse Tregs. We isolated CD4 T cells from pooled spleen and lymph nodes of 6-8-week-old *Foxp3*$^{IRES\text{-}GFP}$ mice as described above and using magnetic negative selection using the Dynabeads Untouched Mouse CD4 Cells Kit (ThermoFisher Scientific, cat#11415D), resting cells for at least 30 min at 4°C prior to electroporation. We freshly prepared gRNA duplexes by combining crRNA (pre-designed validated crRNA, IDT) and tracrRNA oligos (IDT) at equimolar

concentrations to a final duplex concentration of 40 µM, heating the duplex at 95°C for 5 min in a PCR thermocycler and cooling down at RT for 15 minutes before keeping on ice. For every reaction, we used 12x10^6 CD4 T cells, washed with Ca/Mg free PBS and spun down at 200G for 10 min at RT to remove FBS. During this spin, immediately prior to nucleofection, we assembled RNP complexes by thoroughly mixing 2.25 µl of 5 µg/µl TrueCut Cas9 Protein V2 (Thermo Fisher) with 5ul 40µM gRNA duplex and incubating at RT for 15 minutes. For each target, we used a pool of two RNPs carrying different gRNAs duplexes (assembled separately). Cells were resuspended in 100 ul of Electroporation Solution P4 (Lonza) with the pool of both targeting RNPs or with RNPs carrying control non-targeting gRNAs for 2 min at RT and transferred into Large Lonza cuvettes for nucleofection with a Lonza Amaxa 4D Nucleofector (Program DS137). 150 µl of pre-warmed complete T cell medium (RPMI-1640, 10% FBS, supplemented with L-Glutamine, NEAA, beta-mercaptoethanol, and 10mM HEPES) with pre-equilibrated cytokine (10ng/mL recombinant mouse IL7 and 2000 IU/mL recombinant human IL2) was added immediately afterwards and cells were rested for 1h at 37°C. Cells from each cuvette reaction were subsequently transferred to separate wells of a 12 well plate containing 1mL of pre-warmed complete medium, adding another 750 µl of media used to wash residual cells from cuvettes. After 2 additional hours of rest, cells were expanded with anti-CD3/anti-CD28 beads (Thermo Fisher) at a 1:1 ratio for 3 days. All cells in culture were transferred into separate Treg-depleted *Foxp3*^{DTR} hosts, injected with DT (Sigma) at 20 ng/g mouse body weight for two consecutive days prior to transfer and again on the day following transfer. Cells were parked *in vivo* for 1 week prior to sorting 10,000 GFP+ Tregs from pooled spleen and lymph nodes for bulk ATAC-seq. 3 biological replicates passing quality-control were used for targeting conditions and 2 biological replicates passing quality-control for non-targeting conditions (see Table S1).

Bulk ATAC-seq library preparation

Bulk ATAC-seq libraries were prepared using the ImmGen ATAC-seq protocol as described in[29].

scATAC-seq library preparation

For non-hashtagged experiments, nuclei isolation, transposition, GEM generation, and library construction targeting capture of 10000 cells were carried out as detailed in the Chromium Next GEM Single Cell ATAC manual (10x Genomics). Libraries were pooled and sequenced on an

Illumina NovaSeq 6000 to a final median depth of approximately 20-30,000 paired-end reads per cell. Sequencing data were converted to fastq files, aligned to the mm10 reference genome, and quantified per cell using Cell Ranger ATAC software (10x Genomics, v1.2).

ASAP-seq

For experiments with multiple conditions per scATAC run, we hashtagged cells using a modification of the ASAP-seq strategy[78] for low cell input primary cell samples. Before sorting, cells were hashtagged with mouse TotalSeqA DNA-barcoded hashtag antibodies at the same time as staining with fluorophore-conjugated antibodies (BioLegend). Hashtags used in each experiment are provided in Table S1. Cells were sorted into DMEM + 5% FCS in DNA Lo-Bind tubes (Eppendorf, cat # 022431021). After spinning down for 5 min at 500g in a refrigerated centrifuge at 4°C, cells were resuspended in 100 µl chilled 0.1x Omni Lysis buffer (1x Omni Lysis buffer (10mM Tris-HCl, 10mM NaCl, 3 mM $MgCl_2$, 0.1% Tween-20, 0.1% NP40 substitute/IGEPAL, 0.01% Digitonin, 1% BSA in nuclease free water) diluted 1:10 in Wash/Lysis Dilution Buffer (10mM Tris-HCl, 10mM NaCl, 3 mM $MgCl_2$, 1% BSA in nuclease free water)), gently mixed by pipetting and incubated on ice for 6.5 min. Following lysis, 100 µl chilled wash buffer was added and gently mixed by pipetting. Cells were spun down for 5 min at 500g at 4°C, all but 5 µl of supernatant was removed, and 45 µl of chilled 1x nuclei buffer (10x Genomics) was added without mixing. After one more centrifugation step at 500g, 4°C for 5 min, supernatant was removed, and samples were resuspended in 7ul 1x nuclei buffer for cell counting and input into transposition, barcoding, and library preparation according to the Chromium Next GEM Single Cell ATAC manual (10x Genomics).

Modifications to the original 10X protocol were made as described in the original ASAP-seq publication and as detailed at https://citeseq.files.wordpress.com/2020/09/asap_protocol_20200908.pdf. Briefly, 0.5 µl of 1uM BOA bridge oligo was spiked into the barcoding reaction. During GEM incubation, an additional 5 min incubation at 40°C was added to the beginning of the protocol. 43.5 instead of 40.5 µl of Elution Solution I was added during silane bead elution to recover 43 µl. 40 µl was used for SPRI clean up as indicated in the protocol, while 3 µl was set aside. During SPRI cleanup, the supernatant was saved. The bead bound fraction was processed as in the protocol, while for the supernatant fraction, 32 µl SPRI was added for 5 min. Beads were collected on a magnet, washed twice with 80% ethanol, and eluted in 42 µl EB. This 42 µl was combined with the 3 µl set aside from the previous step as input into the HTO indexing reaction. HTO Indexing PCR was run with partial P5 and indexed Rpxx primers

(https://citeseq.files.wordpress.com/2020/09/asap_protocol_20200908.pdf) as: 95°C 3 min, 12-14 cycles of (95°C 20 sec, 60°C 30 sec, 72°C 20 sec), 72°C 5 min. The PCR product was cleaned up with 1.6X SPRI purification for quantification and sequencing alongside ATAC libraries.

scRNA-seq Library Preparation

We sorted GFP+ (WT or KO) or Thy1.1+ (WT) Tregs and reporter negative Tconv from spleens and colonic lamina propria of heterozygote female mice for scRNA-seq. Cell encapsulation and library generation for scRNA-seq was carried out using the 10x Genomics Chromium Single Cell 5′ v2 and V(D)J platform with Feature Barcoding. Data were processed using the standard CellRanger pipeline (10x Genomics). Cells from each condition were hashed with DNA-barcoded TotalSeqC Hashtag and ADT antibodies (BioLegend). Complete hashtag information is provided in Table S1.

scATAC-seq preprocessing and visualization

Data analysis was performed using Signac v1.4[122]. For quality control, only cells with at least 1-4x10^3 fragments per cell (depending on sequencing depth of experiment), greater than 50 percent reads in peaks (relaxed to 30 percent in colon samples due to presence of colon-specific peaks), TSS enrichment score greater than 2, nucleosome signal less than 10, and ratio of blacklist-region reads less than 0.05 were retained for further analysis. See Table S1 for quality-control metrics for all datasets used in this study. Putative doublets identified by ArchR v1.0.1[49] and non-Treg, non-Tconv contaminant cells were also removed. We used the latent semantic indexing approach as previously described[123,124]. Binarized count matrices were normalized using the term frequency-inverse document frequency (TF-IDF) transformation and reduced in dimensionality by singular value decomposition (SVD). As the first component was highly correlated with sequencing depth, SVD components 2-30 were used to generate a shared nearest neighbor (SNN) graph for clustering and as input into UMAP[125] with cosine distance metric for visualization. For visualization of B6/Cast F1 datasets, we created 2 features per OCR, one corresponding to B6-assigned reads and one corresponding to Cast-specific reads, using both versions as independent features for visualization and clustering.

scATAC-seq Analysis

Hashtag counts + assignments

Hashtag processing followed the original recommendations of the ASAP-seq paper[78], using asap_to_kite (https://github.com/caleblareau/asap_to_kite) to process FASTQs files for

downstream quantification by the bustools and kite workflows[126,127]. We used the HTODemux()[128] function in the Seurat package (v4.0.2)[129] to remove doublets and call hashtag identities.

Peak Sets

To enable comparisons across conditions and datasets, we used a common set of open chromatin regions throughout the study. After initially processing data mapped to a set of pan-immune cell OCRs defined by the ImmGen consortium[29], we used iterative sub-clustering at increasing resolution, guided by the silhouette score, to group cells into "metacells" with median size of 350 cells each. We used Archr v1.0.1[49] and MACS2 v2.2.7.1[130] to call fixed-width peaks of 250bp within each of these subgroups of cells, merged into a final set of high-resolution peaks. We compared *de novo* peak calls with peaks from the ImmGen *cis*-regulatory atlas, keeping only those ImmGen peaks that overlapped at least one peak in the *de novo* merged peak set. In cases where new Treg-specific peaks overlapped with an existing ImmGen peak, we retained the ImmGen version for cross-reference compatibility. In cases where a Treg peak did not match an ImmGen peak, it was added to the reference peak set. In total, our final ImmGen+Treg reference OCR set consisted of 216419 peaks and is provided in Table S2. Unless otherwise stated, this OCR set was used for all datasets throughout the study, with two exceptions. In the spleen and colon Treg comparisons, we re-called peaks in colon Tregs and added any missing regions to the reference peak set for that analysis. In the B6/Cast F1 analysis, to avoid missing Cast-specific peaks, we called peaks on Cast allele-specific reads and added any new regions to the peak set for F1 analyses. New peaks were also called to capture any additional IL2 response elements. All additional peaks are provided in Table S10.

Motif Accessibility Analysis

Bias-corrected relative motif accessibility was calculated using chromVAR[131]. We used motifmatchr (https://github.com/GreenleafLab/motifmatchr) to scan OCRs in our refence set from the curated set of mouse motif PWMs from the Buenrostro lab (https://github.com/buenrostrolab/chromVARmotifs/tree/master/data/mouse_pwms_v2.rda) or for all individual human and mouse motifs corresponding to ENCODE motif archetypes (v1.0)[51] (https://github.com/jvierstra/motif-clustering), creating a merged MEME file from all included motif databases. Individual motif models and their assignment to archetypes were obtained from the information provided in[51]. For display of motif accessibility at the level of archetypes, accessibility of individual motif models per archetype were averaged. **Only motifs whose corresponding TF(s) are expressed in Treg cells were displayed.**

Number archetype motifs per OCR

To calculate the number of unique motif archetypes per OCR, we defined an OCR to be positive for a motif archetype if at least 50% of the motifs belonging to the archetype of interest had a match in that OCR.

Motif Variability

Motif Variability was calculated as the standard deviation of the chromVAR bias-corrected motif accessibility scores.

Gene Scores

Gene scores were calculated with Archr v1.0.1, using an exponentially weighted function that accounts for the activity of distal OCRs in a distance-dependent manner[49] and provides an approximate proxy for gene expression. Gene modules scores were calculated using the AddModuleScore() function on gene scores in Seurat, using signatures corresponding to RORγ or Helios Tregs defined from[132].

OCR Signature Generation

OCR signatures distinguishing aTreg and rTreg populations for annotation in Figure 1 were derived from tables provided in[47], using OCRs with a greater than 2-fold increase in accessibility in aTregs. OCR signatures (i.e. aTreg vs rTreg, Treg vs Tconv) derived from the scATAC-seq data used OCRs with |average $\log_2$FC| > 0.25 using the FoldChange() function in Signac.

OCR Signature Relative Accessibility

Relative accessibility of OCR sets, including topics, signature OCRs, etc, was calculated using the chromVAR computeDeviations() function[131].

Pseudobulk Track Visualization

To visualize pseudobulk profiles, BAM files containing reads for each group of cells were extracted using Sinto (https://github.com/timoast/sinto), shifted to account for Tn5 cut-sites, and converted to bigwigs using deeptools[133] for display in the Integrative Genomics Viewer[134] or the WashU Epigenome Browser[135].

OCR Variance

OCR variance was calculated from log-transformed, quantile-normalized metacell counts from the scATAC-seq data as described in the peak calling section. Variance was calculated across these normalized metacell values using the rowVars() function in the matrixStats package (https://github.com/HenrikBengtsson/matrixStats).

IL2 Cell State Assignment

To assign cells to rTreg, aTreg, or rorTreg state in the IL2 scATAC-seq analysis, we first computed per-cell relative accessibility (chromVAR scores) of aTreg vs rTreg and aTreg

vs rorTreg OCR signatures. We then fit a 3-component Gaussian Mixture Model to this matrix of per-cell signature accessibility scores to classify cells as rTreg, aTreg, or rorTreg.

HOMER motif enrichment

We ran motif enrichment in OCRs from each cluster of differential IL2 responses using HOMER[136]. We used the findMotifsGenome.pl with a background set of all Treg OCRs defined in this study, motif length of 10, scan size of 'given', and -S set to 25. Top motifs with p value < 10^{-5} and at least 5% hits in each OCR cluster were kept as significant.

IL2 Cell Matching Visualization

To pair cells with the most similar chromatin cell states in the IL2 treated vs untreated conditions, we used Harmony[89] to remove the IL2 effect from the LSI embedding of the scATAC-seq data. In the treatment effect-removed LSI embedding, closest cells were identified as the nearest neighbor on the k-nearest-neighbor graph constructed from Harmony corrected LSI dimensions 2 to 30 using the FindNeighbors() function in Seurat. A subsample of nearest neighbor pairs was visualized on the original UMAP embedding, coloring the visualization by the starting cell state (rTreg vs aTreg) of the untreated cell in each pair.

OCR UMAP Visualization

To visualize the global structure of OCR usage across Treg and Tconv single cells, we used latent semantic indexing of OCRs instead of cells. After running TF-IDF normalization of the Cell x OCR matrix, we reduced dimensionality with SVD, retaining components 2-30. We used cosine similarity between individual OCRs as input into UMAP[125] to generate the final visualization. SVD components 2-30 were used for Leiden clustering[137] of OCRs to identify the TSS-enriched cluster. To visualize subpopulation-specific OCR signatures, per OCR |average $\log_2$FC| calculated using the FoldChange() function in Signac was overlaid onto the OCR UMAP.

Public TF Binding and Chromatin Accessibility Data

FASTQ files for published TF-binding and chromatin accessibility datasets were downloaded from the NIH Sequence Read Archive. Reads were trimmed using Trimgalore v0.6.6 (https://github.com/FelixKrueger/TrimGalore), aligned using Bowtie2[138], filtered to retain high quality, singly-mapped reads with Samtools[139], and duplicates removed using Picard (http://broadinstitute.github.io/picard/). Peaks were called using MACS2[130]. For TF-binding datasets, peaks were called in comparison to corresponding IgG and/or TF knockout controls. Peaks with irreproducible discovery rate[140] < 0.05 were kept as high-quality binding annotations.

In cases where binding sites were provided in the reference publication, we used the provided sites, lifting over coordinates to the mm10 genome when necessary. For FoxP3 binding analysis, we used a previously defined set of robust FoxP3 binding sites[29,39], derived from intersecting FoxP3-bound regions from 2 different datasets[28,58]. **Ets1 binding data were obtained from[58], Lef1 and TCF-1 from[47], Bach2 from[61], JunD from[60], and p65 from[59].**

OCR-Gene Correlation Analysis

We used FigR v0.1.0 to determine OCR to gene correlations[71]. We used our splenic Treg scATAC-seq and splenic WT Treg profiles from the FoxP3 KO/WT heterozygote female scRNA-seq dataset for paired analysis. Briefly, scRNA-seq and scATAC-seq profiles were first integrated using CCA[124] on variable genes (in RNA) and gene scores (in ATAC). Integrated embeddings were used to match cells across datasets using the scOptMatch algorithm. After pairing, correlations between gene expression and OCR accessibility for regions within 100kb of each TSS were determined and compared to a background null model. For gene-OCR correlations, we kept OCRs with correlation $p < 0.05$ and additionally annotated any other OCRs within 15kb of the gene TSS that did not meet these significance criteria.

B6/Cast F1 Allele-Specific Read Alignment

We adapted the diploid pseudogenome alignment strategy implemented in the lapels/suspenders pipeline[97] and used in recent studies[47,95,96,141] for allele-specific mapping. We obtained B6 and Cast pseudogenomes, MOD files, and variant vcf files from the UNC collaborative cross project (http://csbio.unc.edu/CCstatus/index.py?run=Pseudo) and Mouse Genome Project[91]. Reads were aligned to both B6 and Cast pseudogenomes, shifted to a common set of B6-based reference coordinates, and assigned to an allele of origin based on which allele has stronger mapping. Non-specific reads with equally strong mapping to both alleles were randomly split in half into B6 and Cast groups to obtain the final reads used for allele-specific analyses. Allele-specific BAMs were converted to fragments files using Sinto for input into downstream scATAC analysis.

Ensemble Topic Modeling

Further details on ensemble topic modeling are provided in Computational Note 1.

TSS vs Distal OCRs

We defined TSS OCRs as those OCRs which overlapped an annotated TSS as retrieved from UCSC mm10 annotations

(http://hgdownload.cse.ucsc.edu/goldenPath/mm10/database/refFlat.txt.gz). Distal OCRs were defined as those OCRs that did not overlap annotated TSS. Because OCRs not overlapping TSS in the TSS-predominant cluster (Fig S2A) on the OCR UMAP had similar patterns of accessibility to TSS-overlapping OCRs and mapped closer to TSS than OCRs outside of this cluster, this cluster likely contained promoter OCRs not within 250bp of the TSS. Thus, any OCRs in the TSS cluster were removed from the distal OCR set for analysis.

Topic Motif Enrichment

To calculate enrichment of motifs within topic OCRs, we used a permutation testing framework. We compared the number of observed motif matches within each topic (separately by TSS or distal designations) to the number of matches among a set of 100 background OCRs matched for GC content and accessibility (chosen using the chromVAR getBackgroundPeaks() function). Significance was assessed using a two-sided Z test, with Benjamini-Hochberg false discovery rate correction. We kept motif enrichments with FDR < $1x10^{-10}$. To filter out noise, we considered only motifs of TFs expressed in Tregs and present in at least 3% percent of queried OCRs (TSS or distal) in that topic.

Topic OCR enrichment (Tissue)

We defined a set of pan-tissue Treg OCRs as OCRs with at least a 2-fold increase in accessibility relative to spleen Tregs across all comparisons of Tregs from muscle, colon, and visceral adipose tissue in a previously published dataset[16]**.** We then used the same permutation framework described above to look for enrichment of these OCRs within each topic.

TF Motif, Binding Overlap Analysis

To quantify significance of overlaps in enrichments of TF binding sites and motif within topics, we first computed separate enrichments for TF and matched motifs in each topic using the permutation framework described above, keeping enrichments with FDR < $1x10^{-10}$ in both comparisons. Overlap between the two sets of enrichments was quantified using a binomial test where success was defined as concordance (for presence or absence of enrichment).

Topic Variance Explained

We used the reconstruction error from a linear regression to calculate the variance explained by each topic. For each topic, we fit a linear regression model with centered, Treg metacell pseudobulk accessibility as the response variable and an indicator variable of whether or not each OCR was assigned to a topic as the predictor variable (with 0 intercept). The proportion of total variance in the metacell accessibility explained by the fit regression model was used to estimate the percentage of variance in accessibility explained by each topic. Note that this will be an underestimate in actual variance explained, given that 1) this uses a linear regression framework,

2) looks at variance across single cells collapsed to the metacell level, and 3) uses binarized topic OCR assignments. However, the approach is still a useful approximation for comparing relative variance explained across different conditions.

Differential Topic Accessibility

To calculate differential topic accessibility, we calculated the $\log_2$ Fold Change in average accessibility of OCRs from each topic between the two conditions of interest. To assess significance, we used a two-sided t-test between the log transformed values from each condition, correcting p values for multiple hypothesis testing using the Benjamini-Hochberg false discovery rate procedure.

GREAT Analysis

To link regulatory regions in each topic to gene sets, we used GREAT[76], as implemented by the rGREAT package[142]. We used the rGREAT::great() function with default parameters and 'txdb:mm10' as the TSS source to look for enrichment of gene ontology term annotations in topic regions. For interpretability and to avoid an unreasonably large number of terms, we restricted enrichment to gene ontology terms from the Biological Process category, with "immune system process", "response to stimulus", or "signaling" annotations. Enrichments with p-value < $1x10^{-10}$, fold enrichment > 4, fraction genes in gene set > 50%, and observed region hits > 20 were kept for display.

Topic AME

Further details on the approach for quantifying topic-specific AMEs are provided in Computational Note 2.

TF KO Analysis

In the case of *Foxp3-cre*×*Gata3*$^{fl/fl}$ and *Foxp3-cre*×*Gata3*$^{+/+}$ littermate comparisons, cells corresponding to each genotype were aggregated into pseudobulks, filtering out OCRs with fewer than a mean of 5 reads across samples. In the case of c-Maf knockout data, we used bulk ATAC-seq counts from[98] intersected with our reference peak set, once again filtering out OCRs with fewer a mean of 5 reads across samples. Samples were quantile normalized prior to computing the fold change between KO and WT conditions. To narrow in on likely direct effects, we identified distal (non-TSS) OCRs with a greater than two-fold decrease in accessibility in the knockout vs wildtype that contained GATA or MAF archetype motifs, respectively. We then computed enrichment of each of these OCR sets within topic OCRs using a permutation test versus matched

background OCRs as described above. **For Smarcc1 analysis, we used DESeq2[143] to calculate differential accessibility between bulk ATAC-seq samples from *Smarcc1* edited vs control cell biological replicates. Distal (non-TSS) OCRs with mean base accessibility greater than 5 and loss of accessibility in the *Smarcc1* edited condition with p value < 0.05 were selected as Smarcc1-dependent OCRs. We computed topic OCR enrichment using a permutation test within this OCR set as described for other TF KO analyses.**

FoxP3 heterozygous female analysis

Differential OCRs

We calculated differential accessibility using a logistic regression per OCR with number of fragments per cell included as a latent variable. OCRs with average $|\log_2$ Fold Change$| > 0.1$ and p value < 0.05 were designated as differential**. To avoid effects driven by cell composition, we computed differentials separately subclustered on cells corresponding to rTregs and aTregs from each genotype, based on chromVAR scores of aTreg vs rTreg signature OCRs. For FoxP3-independent OCR comparisons, we compared rTreg KO cells with Tconv.**

Differential accessibility of motifs within each topic

To calculate differential accessibility of motifs within each topic, we calculated the $\log_2$ Fold Change in average accessibility of topic OCRs containing the motif of interest between FoxP3 WT and KO cells. To avoid effects driven by cell composition, we computed differentials separately between clusters corresponding to rTregs and aTregs, as per above. To assess significance, we used a two-sided t-test between the log transformed values from each condition, correcting p values for multiple hypothesis testing using the Benjamini-Hochberg false discovery rate procedure. Effects with FDR < 0.05 were displayed on the heatmap. **FoxP3-independent OCR comparisons were done in the same manner but instead comparing rTreg KO cells with Tconv.**

Local Inverse Simpson's Index

The Local Inverse Simpson's Index (LISI) is a measure originally developed for evaluating efficacy of batch integration methods in single cell genomics, but which provides a general metric for overlap between populations[89]. To use LISI to compare overlap between FoxP3 KO and WT vs Gata3 KO and WT cells, we first integrated scATAC profiles from both datasets into a common embedding using Seurat[124]. We computed diversity in local neighborhoods using the LISI metric from the LISI package for both FoxP3 KO and WT vs Gata3 KO and WT cells, using the LSI components 2-30 in this common, integrated embedding

(https://github.com/immunogenomics/LISI). **For IL2 comparisons, we used the original LSI components 2-30 for LISI computations, as these cells were already in a common embedding within the same experiment.**

scRNA-seq Analysis

scRNA-seq data analysis was performed in Seurat (v4.0.2)[129]. Cells with greater than 500 UMIs and fewer than 10% mitochondrial reads were kept for downstream analysis, with contaminant cells removed using known marker genes. We used the HTODemux()[128] function to remove doublets and call hashtag identities. The top 2500 variable genes and first 35 principal components were used for nearest-neighbor graph construction and UMAP visualization. Tregs were classified as RORγ+ using signatures from[24]. Treg signature scores were calculated as the output from the AddModuleScore() function for the Treg Up signature minus the output for the Treg Down signature[31].

Data Visualization

Graphs and visualizations were generated in R (v4.1.0[144]) using ggplot2 (v2.3.3.5[145]) and in Python (v3.9.7, http://www.python.org) using matplotlib v3.5.0[146]. Flow cytometry data was analyzed using FlowJo v10 (BD LifeSciences), with corresponding plots and statistical analyses done in GraphPad Prism (www.graphpad.com).

Data Availability

Raw and processed data files have been deposited at the Gene Expression Omnibus (GEO: GSE216910 and GSM5712663). Reviewer token: snazigsixhkdvul

FIGURE LEGENDS

Figure 1: Single-cell ATAC-seq reveals imbricated transcription factor activities across diverse Treg cell states A) Experimental Overview. Single cell chromatin accessibility profiling was used to link TF activities to diverse Treg chromatin states. Continuous Treg cell states were first annotated using TF motif enrichment. Topic modeling was used to learn groups of co-varying OCRs that formed discrete regulatory modules underlying observed cell states. *Cis*-regulatory variation in B6/Cast F1 hybrid scATAC-seq data enabled identification of causal regulators of Treg chromatin programs. The resulting Treg regulatory network was validated using TF binding (ChIP-seq, CUT&RUN) and knockout datasets. **All generated datasets and associated metrics are described in Table S1.** B) Aggregated accessibility profiles of Treg and Tconv single cells at the *Foxp3* locus from scATAC-seq data of **splenic Tregs generated from a *Foxp3*$^{IRES-GFP}$ reporter mouse;** highlights indicate conserved non-coding sequence (CNS) loci previously described to control *Foxp3* expression. C) Relative accessibility (chromVAR scores) across Treg single cells of OCRs increased in accessibility in aTreg vs rTreg populations (FoldChange>2 in data from ref from[47]) visualized on UMAP of splenic Treg scATAC-seq data. D) Gene scores, chromatin-based proxies for gene expression, visualized for select genes on Treg UMAP from (C). E) Relative accessibility (chromVAR motif scores) of OCRs containing indicated TF motifs; motifs averaged within 'archetypes'[51] to reduce redundancy. **Only motifs whose corresponding TF(s) are expressed in Treg cells are shown. Motif logos are representative of TFs from each archetype.**

Figure 2: Global usage of open chromatin regions across Treg single cells A) UMAP visualization of OCR usage across Treg and Tconv single cells **from *Foxp3*$^{IRES-GFP}$ spleen scATAC-seq data in Fig 1.** B) OCR UMAP from (A) with OCRs overlapping annotated TSS highlighted; color indicates density of TSS OCRs. **C) OCR UMAP from (A) colored by variance in accessibility across Tregs.** D) UMAP from (A) colored by $\log_2$ Fold Change between indicated populations from Treg scATAC-seq data in Figure 1. E) Annotation of OCRs overlapping indicated TF motif archetypes on OCR UMAP from (A); color indicates density of motif-containing OCRs. F) Annotation of OCRs overlapping indicated TF binding sites on OCR UMAP from (A); color indicates density of bound OCRs. G) Annotation on OCR UMAP from (A) of OCRs with accessibility correlated (FigR[71] $p < 0.05$) with expression of indicated genes (in red) or other OCRs within 15 kb of gene TSS not meeting this correlation significance threshold (in green). Locations of *Foxp3* conserved non-coding sequences (CNS) and promoter loci are indicated.

Figure 3: Topic modeling learns coordinated accessibility programs across Treg single cells A) Annotation of OCRs assigned to each topic on OCR UMAP from Fig 2; color indicates density of OCRs. B) Relative accessibility (chromVAR scores) of OCRs from each topic across Treg single cells, visualized on UMAP of Treg scATAC-seq data from Fig 1. C) Enrichment of TF motifs within distal or TSS OCRs from each topic; black indicates enrichment FDR < $1x10^{-10}$. Motifs are organized by TF family (row annotation). D) Overlap between enrichment of TF-binding data (top) or corresponding motifs (bottom). Green indicates concordant enrichments in both motif and TF-binding analyses, whereas blue indicates enrichment only in TF-binding data and yellow enrichment only using motif data. P value indicates significance of overlapping enrichments between motif and binding data (binomial test). **E) Gene Ontology gene sets significantly enriched among regulatory regions in each topic (using GREAT[76] analysis). Heatmap indicates fold change of enrichment relative to background. Full table of enrichments and pathway names is provided in Table S8.**

Figure 4: Topics across tissues A) Proportion of pan-tissue Treg OCRs (from[16]) overlapping each Treg topic; color indicates significance of enrichment (permutation test). B) Proportion of variance in accessibility explained by each topic in **spleen scATAC data from Fig 1 (spleen 1) or in independently generated *Foxp3*$^{IRES-GFP}$ spleen Treg scATAC data in Figure 4 (spleen 2).** C) Proportion of variance in accessibility explained by each topic **in colon or spleen Treg scATAC data (from the same *Foxp3*$^{IRES-GFP}$ reporter mouse).** D) Differential accessibility per topic between aggregated colon and spleen Treg scATAC profiles. E) Relative accessibility (chromVAR scores) of OCRs from tissue- (Topics 3, 14) or spleen- (Topic 11) specific topics across spleen and colon Treg single cells visualized on scATAC UMAP, separated by organ. F) Relative accessibility (chromVAR scores) of OCRs from Helios- (Topic 10) or RORγ- (Topic 9) specific topics across spleen and colon Treg single cells visualized on scATAC UMAP, separated by organ. Right panel indicates gene module scores for genes from a Helios vs RORγ specific gene expression signature visualized on the same UMAP.

Figure 5: Topics and the response to IL2 A) UMAP of scATAC-seq of splenic Tregs from *Foxp3*$^{IRES-GFP}$ mice treated with 10 µg of recombinant mouse IL2 (IV) or PBS vehicle control for 2 hours, colored by density of cells from each condition. B) Relative accessibility (chromVAR scores) of OCRs increased in accessibility in aTreg vs rTreg populations, visualized on UMAP of scATAC-seq data from (A) and split by treatment group. C) Visualization of pairs identifying the closest cells in IL2 -untreated and -treated conditions

in high-dimensional OCR space, overlaid onto UMAP from (A). Color indicates the cell state (rTreg or aTreg) of the untreated, control cell in each pair. Random samples of 40 pairs of rTreg control and 40 pairs of aTreg control cells shown for visual clarity. D) Fold Change of rTreg or aTreg pSTAT5 mean fluorescence intensity, relative to untreated control, after *ex vivo* stimulation with indicated concentration of recombinant mouse IL2 for 15 minutes.*,p<0.05, t-test. N=3 mice. E) IL2RA (CD25) mean fluorescence intensity among splenic rTreg or aTreg. p value from paired t-test. N=9 mice. F) Relative accessibility of OCRs containing Stat5 motif (chromVAR scores), visualized on UMAP from A and split by cell state and treatment status. G) Heatmap of OCRs with $|\log_2\text{FoldChange}|>1.5$ in accessibility between treated and untreated conditions in either rTreg or aTreg comparisons (860 OCRs). Displayed values indicate $\log_2$Fold Change of aggregated accessibility in treated versus untreated cells, matched for cell state (rTreg or aTreg). Each column indicates a different biological replicate. Heatmap to left shows significance of enrichment ($\log_{10}$(FDR), permutation test) of Topic OCRs in each OCR cluster (enrichments with FDR<0.05 shown). TF motifs (HOMER[136], $p < 10^{-5}$) enriched in each OCR cluster are indicated on right. H) Differential accessibility per topic between aggregated IL2 treated vs untreated spleen Treg scATAC profiles, separated by cell state.

Figure 6: Genetic variation identifies causal regulators of Treg chromatin programs A) Overview of Topic AME calculation: to quantify the contribution of motifs to the accessibility of each topic, AME was calculated in reads aggregated from cells with high accessibility of topic OCRs containing each candidate motif in spleen Treg scATAC-seq profiles from B6/Cast F1 hybrid mice. B) Topic-specific AME (FDR<0.10) of motifs in each Treg topic. Heatmap in same order as in Figure 3 and shows overlap between motif enrichment and significant AME scores. Positive AMEs (pink) indicate positive effect on chromatin accessibility and negative AMEs (green) indicate negative effect on chromatin accessibility. Motifs are ordered by TF family. C) Enrichment (signed $\log_{10}$(FDR), permutation test) in each topic of Gata3-dependent OCRs (GATA-motif containing OCRs decreased in accessibility > two-fold in ***Foxp3-cre×Gata3$^{fl/fl}$*** **Treg-specific Gata3 KO vs *Foxp3-cre×Gata3$^{+/+}$* WT**). Panel above indicates predicted topic effect based on topic AME (top bar) or GATA family motif enrichment (bottom bar). D) Enrichment (signed $\log_{10}$(FDR), permutation test) in each topic of c-Maf-dependent OCRs (MAF-motif containing OCRs decreased in accessibility > 2-fold in ***Foxp3-cre×Maf$^{fl/fl}$*** **Treg-specific c-Maf KO vs *Foxp3-cre×Maf$^{+/+}$* WT**[98]). Panel above indicates predicted topic effect based on MAF family topic AME (top bar) or motif enrichment (bottom bar**). E) Enrichment (signed $\log_{10}$(FDR),**

permutation test) in each topic of Smarcc1-dependent OCRs (OCRs with loss of accessibility at p<0.05 in *Foxp3*$^{IRES-GFP}$ Tregs electroporated with CRISPR/Cas9 ribonucleoprotein complexes carrying *Smarcc1*-targeting versus control gRNAs and transferred for 1 week into Treg-depleted *Foxp3*^{DTR} hosts). Panel above indicates predicted topic effect based on overlap of topic AME and motif enrichment, as in 6B.

Figure 7: FoxP3 effects on the Treg regulatory network A) Experimental Scheme. FoxP3-deficient (KO) and -sufficient (WT) GFP+ Tregs were sorted along with GFP- Tconv from **Foxp3^{fs327}-GFP/Foxp3-Thy1.1 or Foxp3^{wt}-GFP/Foxp3-Thy1.1 heterozygous female mice for** scATAC-seq. B) UMAP of scATAC of Treg and Tconv from FoxP3 WT or KO heterozygous female mice, separated by genotype and cell type. C) Relative accessibility (chromVAR scores) across WT and KO Treg single cells of OCRs increased in accessibility in aTreg vs rTreg populations, visualized on UMAP of scATAC-seq data from (B). D) Proportion of rTreg and aTreg populations in FoxP3 WT or KO populations by CD44 and CD62L flow cytometry. E) OCR UMAP from Fig 2 colored by $\log_2$ FoldChange in chromatin accessibility in FoxP3 WT vs KO cells in aTreg or rTreg comparisons. Right panel shows annotation of OCRs overlapping FoxP3 binding sites; color indicates density of FoxP3-bound OCRs. F) Differential accessibility per topic between WT Treg and KO Treg-like cells in rTreg or aTreg comparisons. G) Differential accessibility per motif in each topic (distal OCRs, FDR<0.05) between WT and KO Treg cells in rTreg or aTreg comparisons for motif to topic connections from Figure 6B. Grey indicates significant motif to topic connection from topic AME in Figure 6B but no significant change in accessibility across FoxP3 comparisons. **H) Differential accessibility per topic between KO rTreg and Tconv cells. I) Differential accessibility per motif in each topic (distal OCRs, FDR<0.05) between KO rTreg and Tconv cells for motif to topic connections from Figure 6B. Grey indicates significant motif to topic connection from topic AME in Figure 6B but no significant change in accessibility across differential comparisons.**

Figure 8: FoxP3-independent RORγ+ Treg-like cells A) *Rorc* gene scores and NR/19 (contains RORγ) motif accessibility (chromVAR score) in WT and KO Treg populations in scATAC-seq data from Fig 7. B) Proportion of RORγ+ vs Helios+ Tregs among WT or KO Tregs in heterozyous female mice across organs. C) Quantification of (A). D) Proportion of GFP+ Tregs among WT or KO Tregs with or without VNMA antibiotic treatment. E) Proportion of RORγ+ vs Helios+ Tregs among WT or KO colon Tregs before or after VNMA antibiotic treatment. F) Proportion of IL17A+ RORγ+ cells among stimulated FoxP3 KO or WT Tregs from spleen or colon. **G) UMAP of**

scRNA-seq of colonic lamina propria Treg and Tconv from FoxP3 WT or KO heterozygous female mice, separated by genotype and cell type. Pink indicates *Rorc*+ reporter+ WT or KO Tregs, blue indicates all other reporter+ WT or KO Tregs, and orange indicates reporter- Tconv. H) *Il17a* expression overlaid onto UMAP from (G). I) Distribution of Treg gene signature expression scores across indicated populations in colon scRNA-seq data from (G) J) *Il10* expression overlaid onto UMAP from (G).

Figure 1

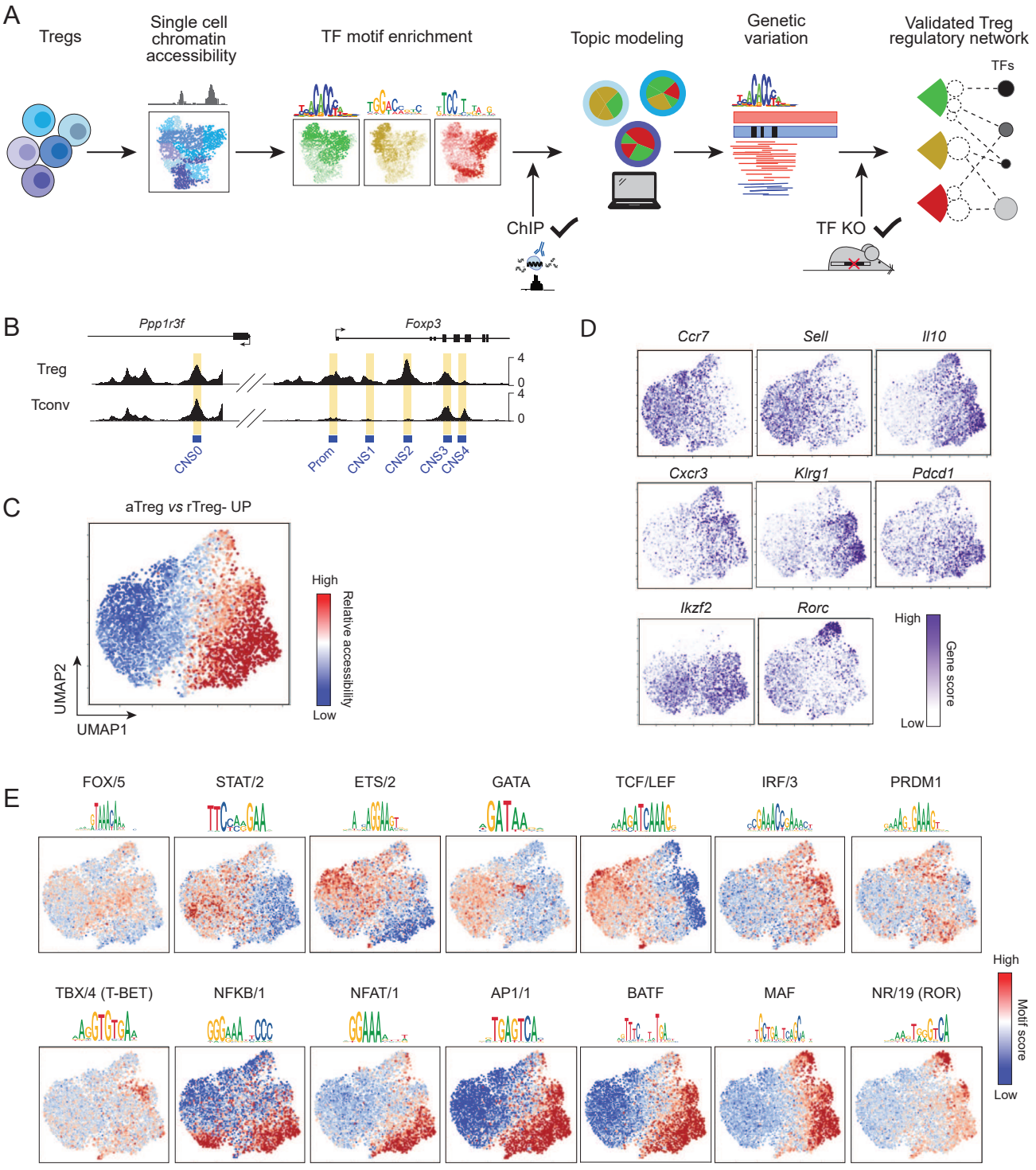
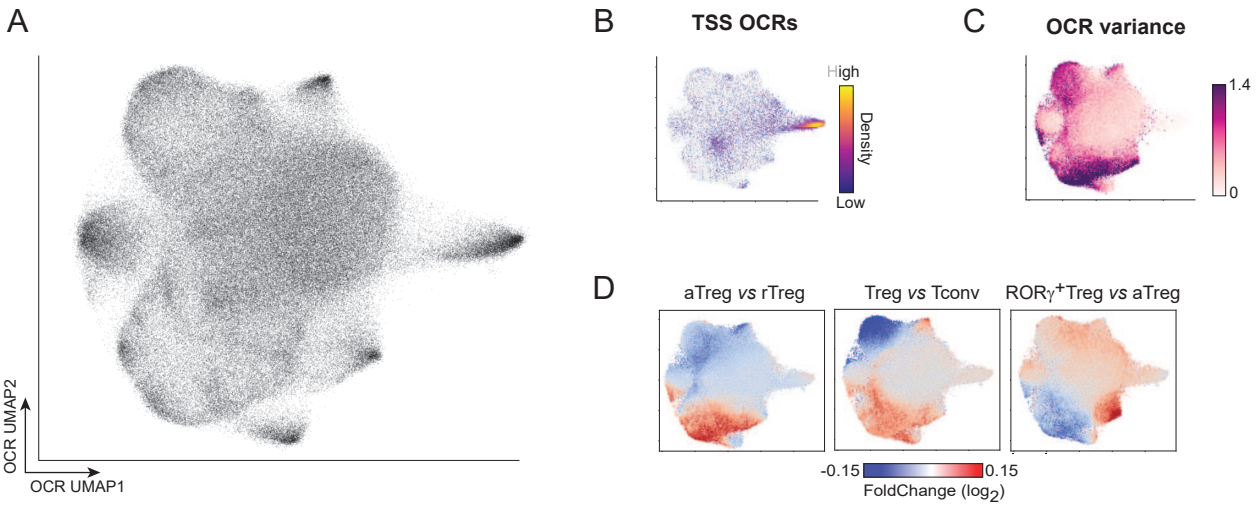
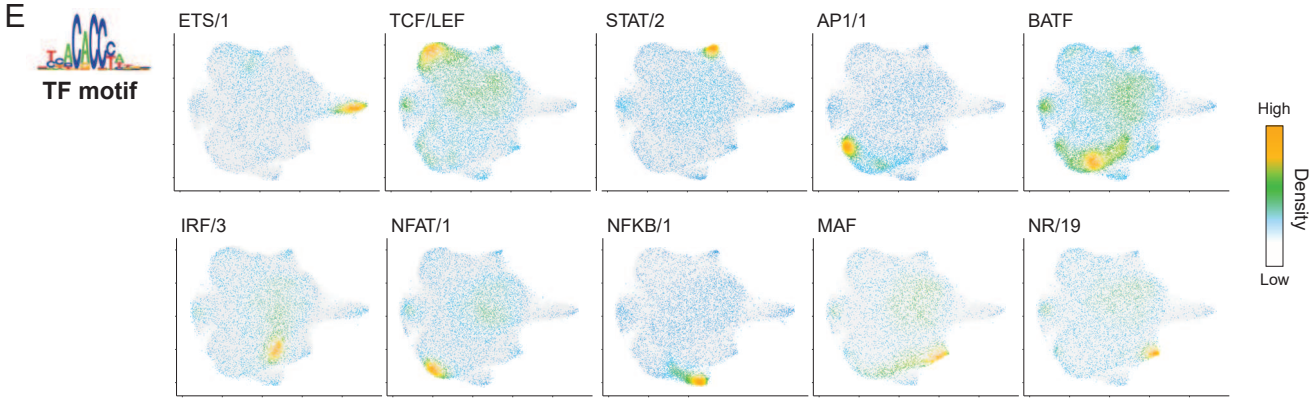
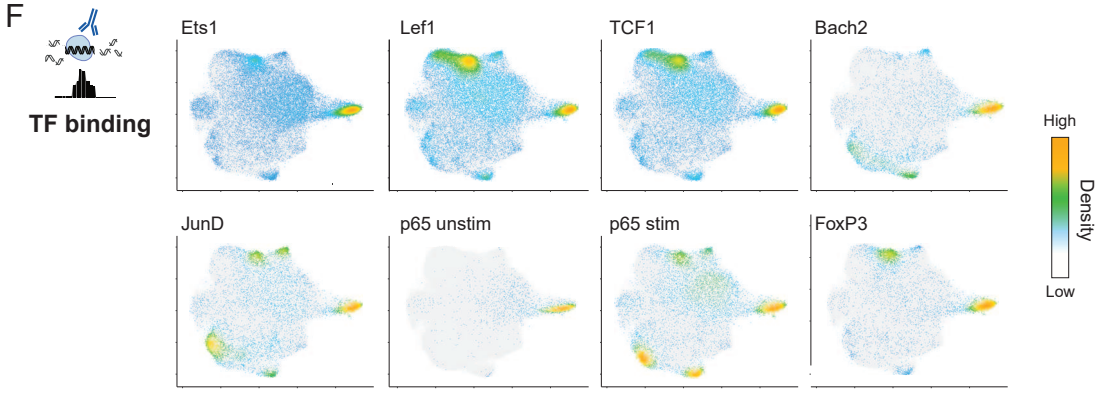

A
Tregs
Single cell chromatin accessibility
TF motif enrichment
Topic modeling
Genetic variation
Validated Treg regulatory network
TFs
ChIP
TF KO
B
Ppp1r3f
Foxp3
Treg
Tconv
CNS0
Prom
CNS1
CNS2
CNS3
CNS4
C
aTreg vs rTreg- UP
High
Relative accessibility
Low
UMAP2
UMAP1
D
Ccr7
Sell
Il10
Cxcr3
Klrg1
Pdcd1
Ikzf2
Rorc
High
Gene score
Low
E
FOX/5
STAT/2
ETS/2
GATA
TCF/LEF
IRF/3
PRDM1
TBX/4 (T-BET)
NFKB/1
NFAT/1
AP1/1
BATF
MAF
NR/19 (ROR)
High
Motif score
Low

Figure 2


A
OCR UMAP2
OCR UMAP1
B
TSS OCRs
High
Density
Low
C
OCR variance
1.4
0
D
aTreg vs rTreg
Treg vs Tconv
RORγ+Treg vs aTreg
-0.15
0.15
FoldChange (log2)


E
TF motif
ETS/1
TCF/LEF
STAT/2
AP1/1
BATF
IRF/3
NFAT/1
NFKB/1
MAF
NR/19
High
Density
Low


F
TF binding
Ets1
Lef1
TCF1
Bach2
JunD
p65 unstim
p65 stim
FoxP3
High
Density
Low

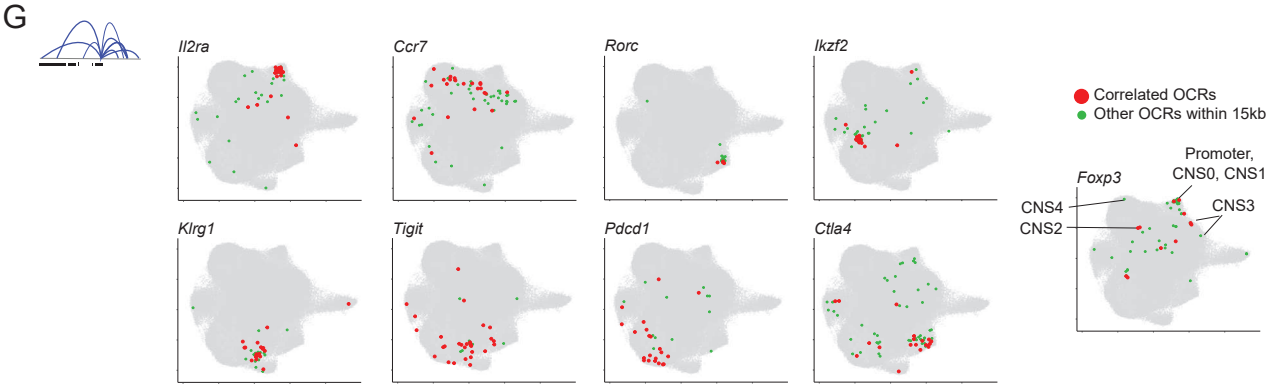

G
Il2ra
Ccr7
Rorc
Ikzf2
Klrg1
Tigit
Pdcd1
Ctla4
Correlated OCRs
Other OCRs within 15kb
Foxp3
Promoter, CNS0, CNS1
CNS4
CNS2
CNS3

Figure 3

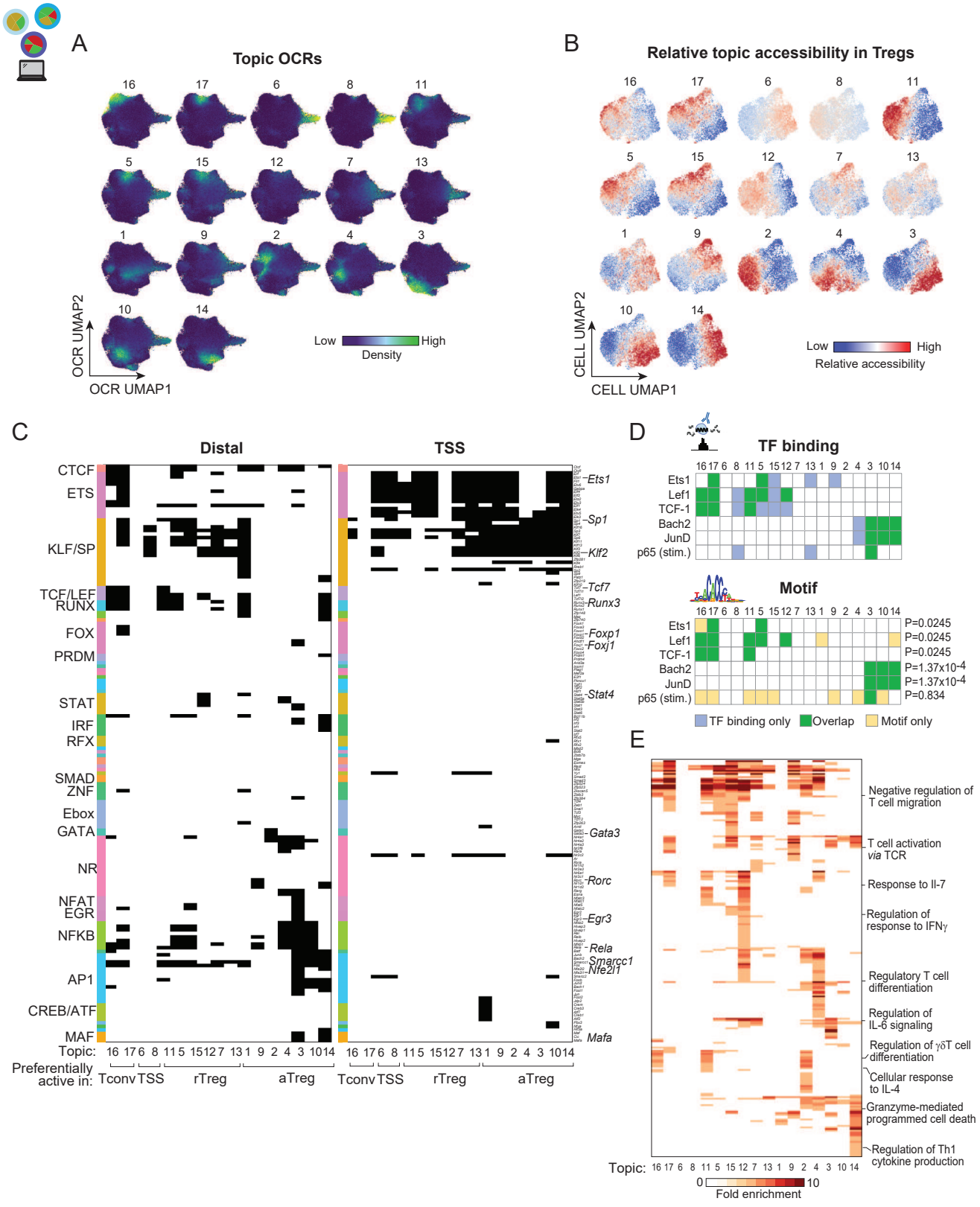

A
Topic OCRs
16 17 6 8 11
5 15 12 7 13
1 9 2 4 3
10 14
OCR UMAP2
OCR UMAP1
Low High
Density
B
Relative topic accessibility in Tregs
CELL UMAP2
CELL UMAP1
Low High
Relative accessibility
C
Distal
TSS
CTCF
ETS
KLF/SP
TCF/LEF
RUNX
FOX
PRDM
STAT
IRF
RFX
SMAD
ZNF
Ebox
GATA
NR
NFAT
EGR
NFKB
AP1
CREB/ATF
MAF
Ets1
Sp1
Klf2
Tcf7
Runx3
Foxp1
Foxj1
Stat4
Gata3
Rorc
Egr3
Rela
Smarcc1
Nfe2l1
Mafa
Topic: 16 17 6 8 11 5 15 12 7 13 1 9 2 4 3 10 14
Preferentially active in: Tconv TSS rTreg aTreg Tconv TSS rTreg aTreg
D
TF binding
Ets1
Lef1
TCF-1
Bach2
JunD
p65 (stim.)
Motif
P=0.0245
P=0.0245
P=0.0245
P=1.37x10-4
P=1.37x10-4
P=0.834
TF binding only
Overlap
Motif only
E
Negative regulation of T cell migration
T cell activation via TCR
Response to Il-7
Regulation of response to IFNγ
Regulatory T cell differentiation
Regulation of IL-6 signaling
Regulation of γδT cell differentiation
Cellular response to IL-4
Granzyme-mediated programmed cell death
Regulation of Th1 cytokine production
Topic: 16 17 6 8 11 5 15 12 7 13 1 9 2 4 3 10 14
0 10
Fold enrichment

Figure 4

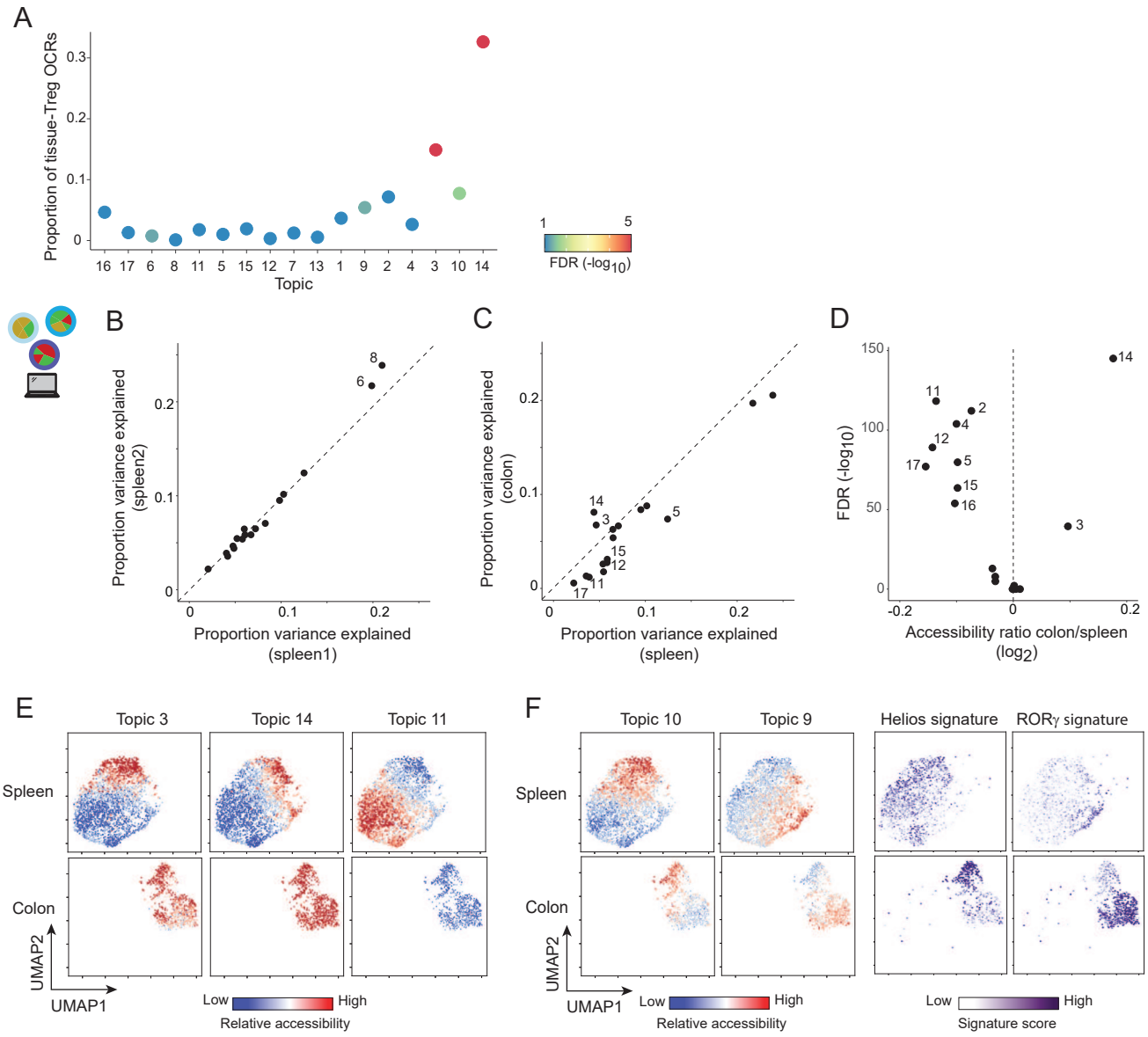

A
Proportion of tissue-Treg OCRs
0.3
0.2
0.1
0
16 17 6 8 11 5 15 12 7 13 1 9 2 4 3 10 14
Topic
1
5
FDR (-log10)
B
Proportion variance explained (spleen2)
Proportion variance explained (spleen1)
C
Proportion variance explained (colon)
Proportion variance explained (spleen)
D
FDR (-log10)
Accessibility ratio colon/spleen (log2)
E
Topic 3
Topic 14
Topic 11
Spleen
Colon
UMAP2
UMAP1
Low
High
Relative accessibility
F
Topic 10
Topic 9
Helios signature
RORγ signature
Signature score

Figure 5

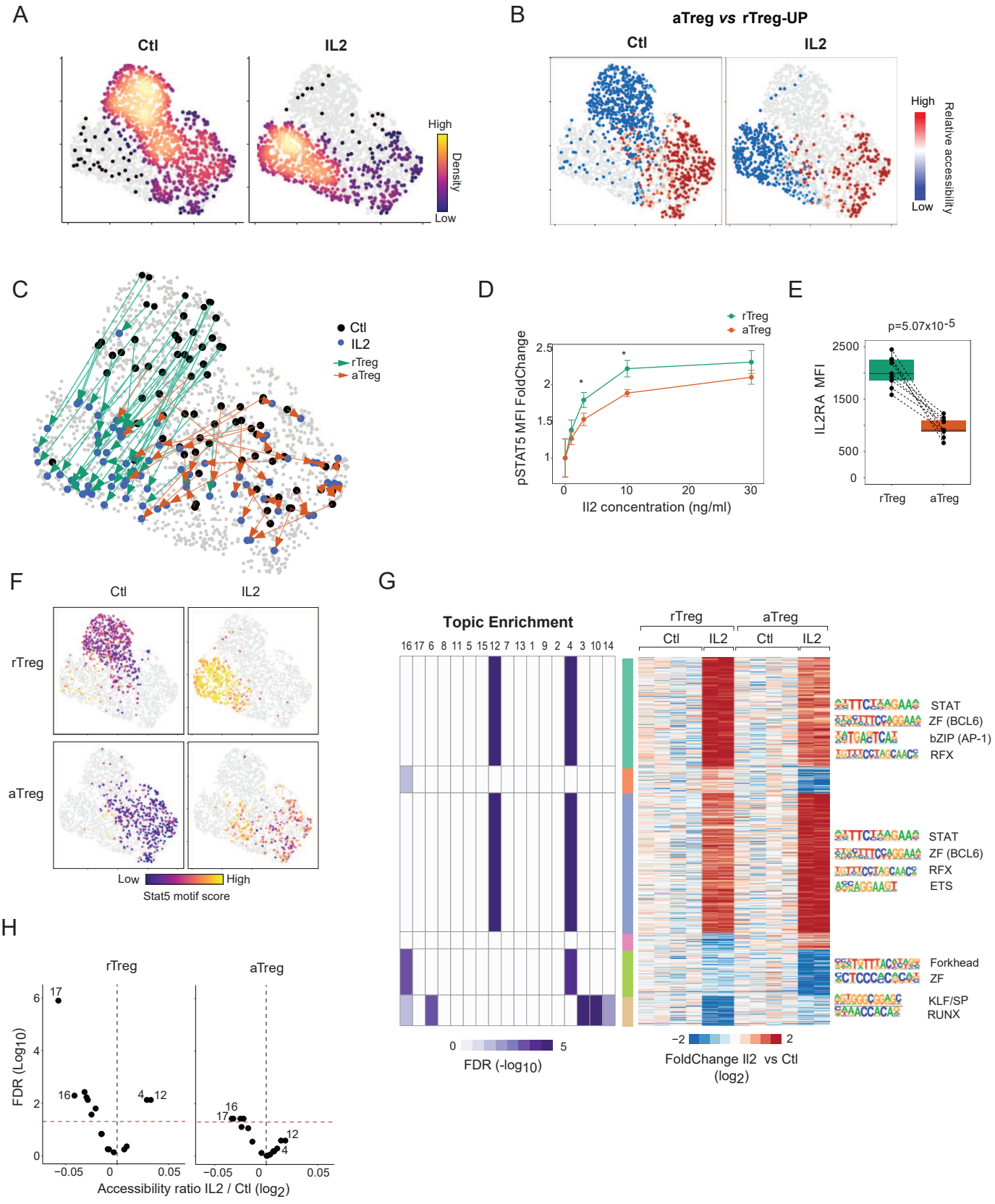

A
Ctl
IL2
High
Density
Low
B
aTreg vs rTreg-UP
Ctl
IL2
High
Relative accessibility
Low
C
Ctl
IL2
rTreg
aTreg
D
rTreg
aTreg
pSTAT5 MFI FoldChange
Il2 concentration (ng/ml)
E
p=5.07x10-5
IL2RA MFI
rTreg
aTreg
F
Ctl
IL2
rTreg
aTreg
Low
High
Stat5 motif score
G
Topic Enrichment
16 17 6 8 11 5 15 12 7 13 1 9 2 4 3 10 14
rTreg
aTreg
Ctl
IL2
Ctl
IL2
STAT
ZF (BCL6)
bZIP (AP-1)
RFX
STAT
ZF (BCL6)
RFX
ETS
Forkhead
ZF
KLF/SP
RUNX
FDR (-log10)
FoldChange Il2 vs Ctl
(log2)
H
rTreg
aTreg
FDR (Log10)
Accessibility ratio IL2 / Ctl (log2)

Figure 6

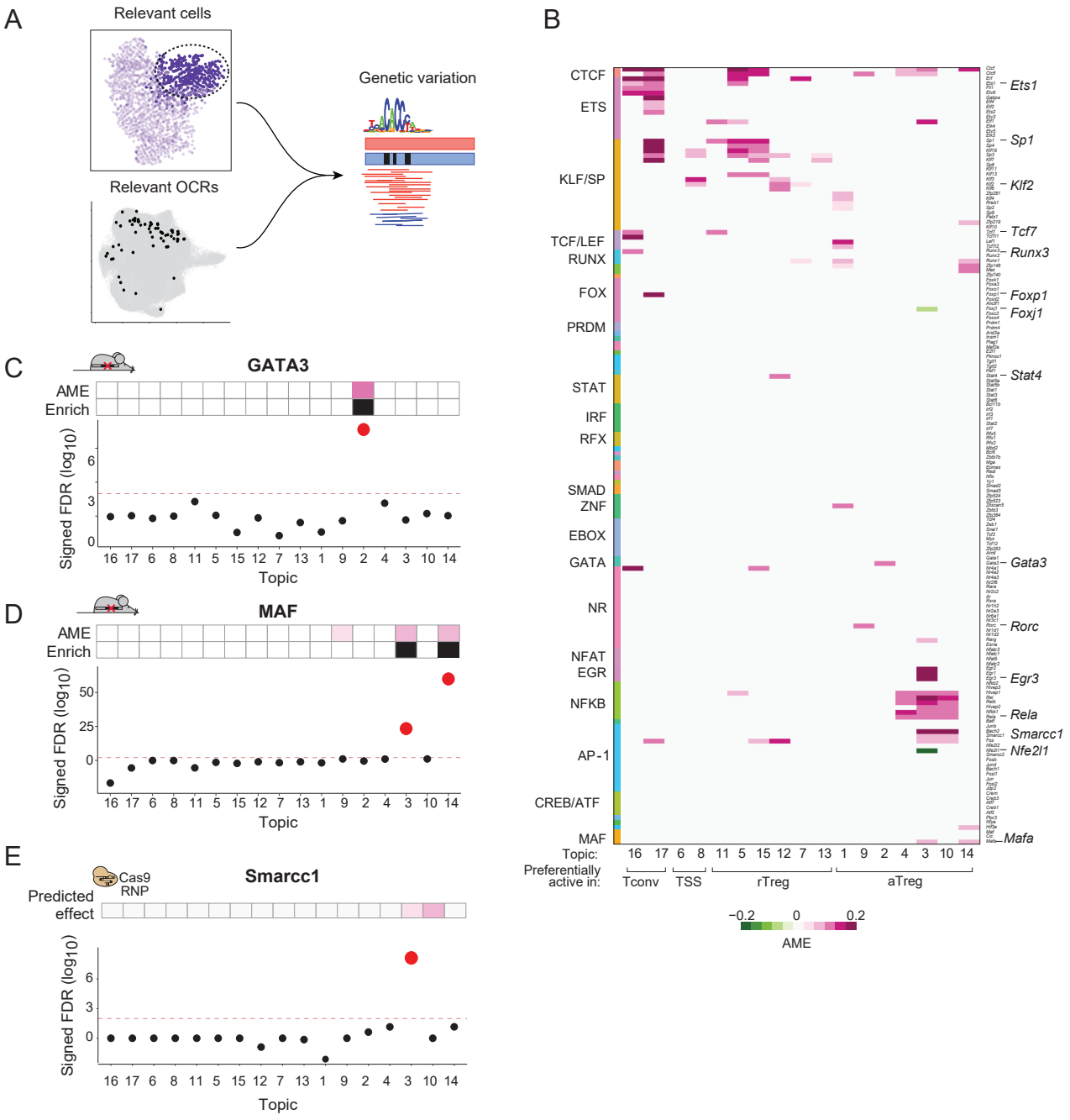

A
Relevant cells
Relevant OCRs
Genetic variation
B
CTCF
ETS
KLF/SP
TCF/LEF
RUNX
FOX
PRDM
STAT
IRF
RFX
SMAD
ZNF
EBOX
GATA
NR
NFAT
EGR
NFKB
AP-1
CREB/ATF
MAF
Ets1
Sp1
Klf2
Tcf7
Runx3
Foxp1
Foxj1
Stat4
Gata3
Rorc
Egr3
Rela
Smarcc1
Nfe2l1
Mafa
Topic:
16 17 6 8 11 5 15 12 7 13 1 9 2 4 3 10 14
Preferentially active in:
Tconv
TSS
rTreg
aTreg
−0.2
0
0.2
AME
C
GATA3
AME
Enrich
Signed FDR (log10)
Topic
D
MAF
E
Cas9 RNP
Predicted effect
Smarcc1

Figure 7

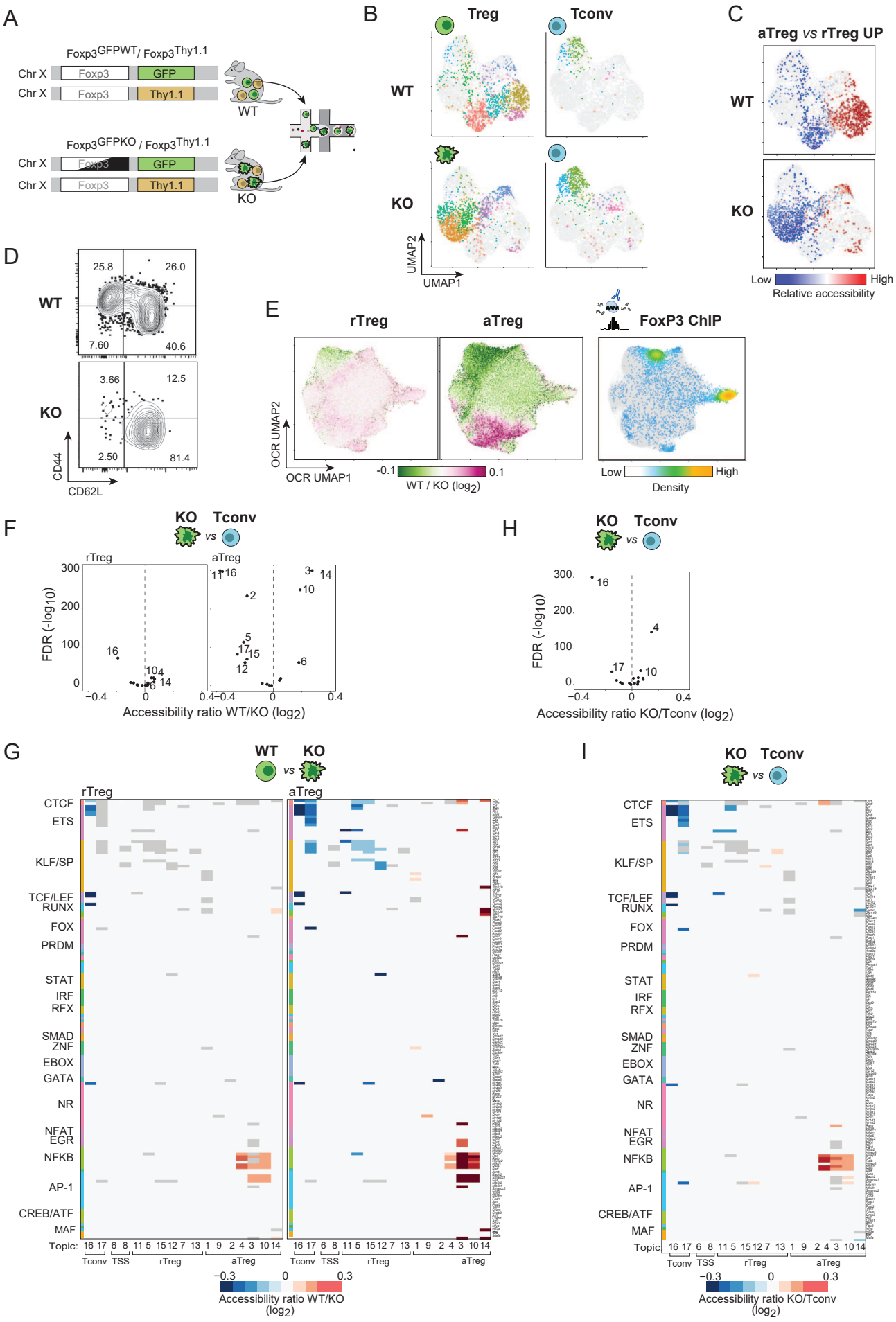
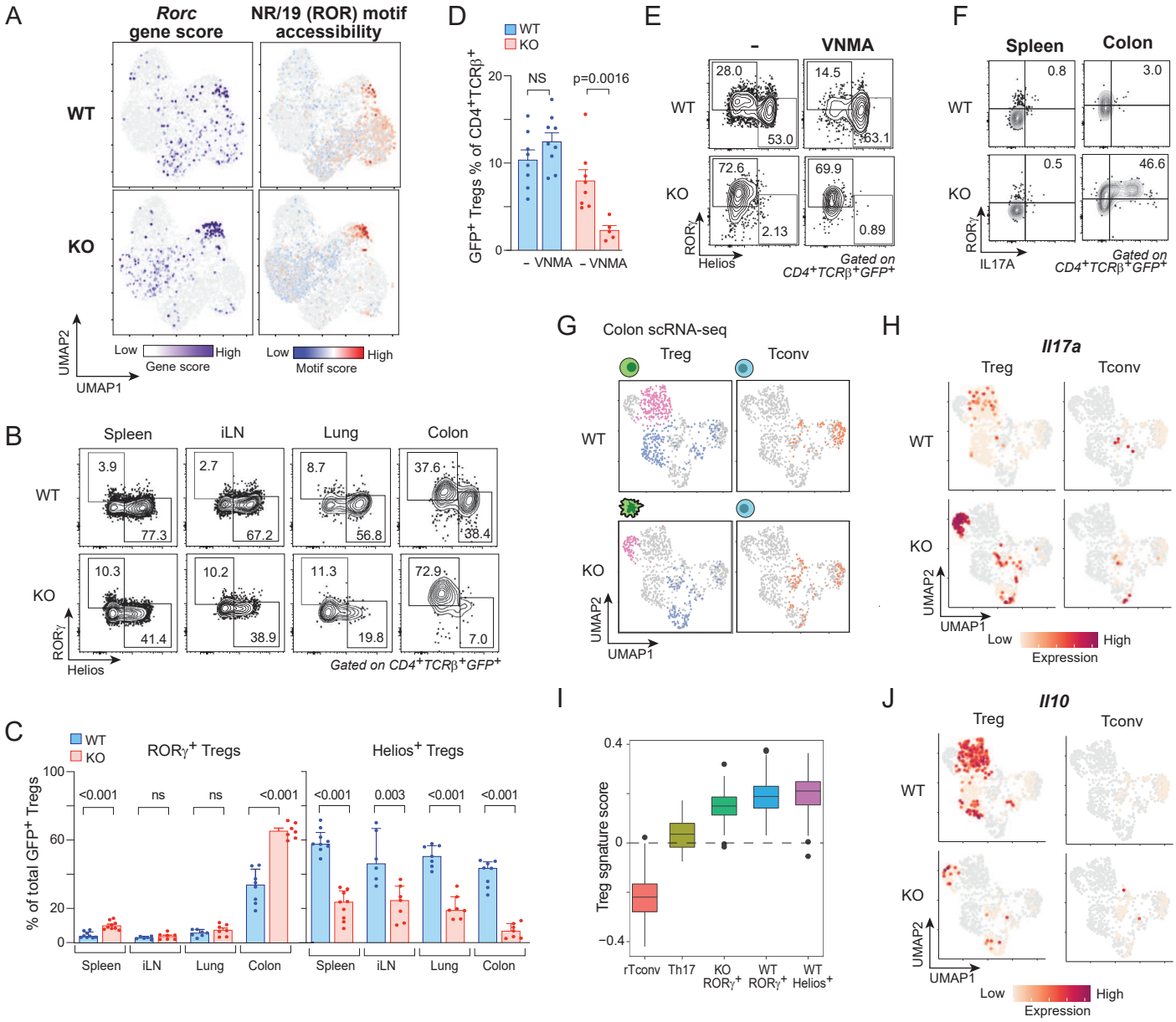

A
Foxp3GFPWT/ Foxp3Thy1.1
Chr X
Foxp3
GFP
Chr X
Foxp3
Thy1.1
WT
Foxp3GFPKO / Foxp3Thy1.1
Chr X
Foxp3
GFP
Chr X
Foxp3
Thy1.1
KO
B
Treg
Tconv
WT
KO
UMAP2
UMAP1
C
aTreg vs rTreg UP
WT
KO
Low
High
Relative accessibility
D
WT
25.8
26.0
7.60
40.6
KO
3.66
12.5
2.50
81.4
CD44
CD62L
E
rTreg
aTreg
FoxP3 ChIP
OCR UMAP2
OCR UMAP1
-0.1
0.1
WT / KO (log2)
Low
High
Density
F
KO
Tconv
vs
rTreg
aTreg
FDR (-log10)
Accessibility ratio WT/KO (log2)
H
KO
Tconv
vs
FDR (-log10)
Accessibility ratio KO/Tconv (log2)
G
WT
KO
vs
rTreg
aTreg
CTCF
ETS
KLF/SP
TCF/LEF
RUNX
FOX
PRDM
STAT
IRF
RFX
SMAD
ZNF
EBOX
GATA
NR
NFAT
EGR
NFKB
AP-1
CREB/ATF
MAF
Topic:
Tconv
TSS
rTreg
aTreg
−0.3
0
0.3
Accessibility ratio WT/KO (log2)
I
KO
Tconv
vs
CTCF
ETS
KLF/SP
TCF/LEF
RUNX
FOX
PRDM
STAT
IRF
RFX
SMAD
ZNF
EBOX
GATA
NR
NFAT
EGR
NFKB
AP-1
CREB/ATF
MAF
Topic:
Tconv
TSS
rTreg
aTreg
−0.3
0
0.3
Accessibility ratio KO/Tconv (log2)

Figure 8


A
Rorc gene score
NR/19 (ROR) motif accessibility
WT
KO
UMAP2
UMAP1
Low
High
Gene score
Motif score
B
Spleen
iLN
Lung
Colon
3.9
77.3
2.7
67.2
8.7
56.8
37.6
38.4
10.3
41.4
10.2
38.9
11.3
19.8
72.9
7.0
RORγ
Helios
Gated on CD4+TCRβ+GFP+
C
WT
KO
RORγ+ Tregs
Helios+ Tregs
% of total GFP+ Tregs
<0.001
ns
ns
<0.001
<0.001
0.003
<0.001
<0.001
Spleen
iLN
Lung
Colon
D
GFP+ Tregs % of CD4+TCRβ+
NS
p=0.0016
– VNMA
– VNMA
E
–
VNMA
28.0
53.0
14.5
63.1
72.6
2.13
69.9
0.89
RORγ
Helios
Gated on CD4+TCRβ+GFP+
F
Spleen
Colon
0.8
3.0
0.5
46.6
RORγ
IL17A
Gated on CD4+TCRβ+GFP+
G
Colon scRNA-seq
Treg
Tconv
UMAP2
UMAP1
H
Il17a
Treg
Tconv
Low
High
Expression
I
Treg signature score
rTconv
Th17
KO RORγ+
WT RORγ+
WT Helios+
J
Il10
Treg
Tconv
Low
High
Expression

SUPPLEMENTARY FIGURE LEGENDS

Figure S1: Additional analysis of Treg scATAC-seq data A) Sort strategy for isolation of Treg and Tconv cells **from *Foxp3*$^{IRES-GFP}$ reporter mouse**. B) Fraction of reads in peaks vs number of unique reads per cell plot for quality control of scATAC-seq data. Dotted lines indicate thresholds used to select cells for analysis. C) UMAP of scATAC-seq data of splenic Treg and Tconv from ***Foxp3*$^{IRES-GFP}$ reporter,** colored by subpopulation annotation. D) UMAP of Treg scATAC-seq from Figure 1 colored by number of unique reads per cell. E) UMAP of Treg and Tconv scATAC-seq from (C) colored by number of unique reads per cell. F) Relative accessibility (chromVAR scores) of OCRs containing indicated TF motifs for most variable TF motifs across Treg scATAC-seq data; motifs averaged within 'archetypes' to reduce redundancy. **Only motifs whose corresponding TF(s) are expressed in Treg cells are shown.** G) Variability of motif accessibility (chromVAR scores) across Treg scATAC-seq data.

Figure S2: Genome browser views of state-specific loci. Aggregated accessibility of cells from rTreg, aTreg, rorTreg, or Tconv cell states at state-specific loci from scATAC-seq data in Figure 1.

Figure S3: Additional analysis of OCR distributions A) OCR UMAP from Fig 2, colored by Leiden clustering of OCRs, with cluster enriched for TSS indicated. B) Distribution of number of unique archetype motifs per OCR

Figure S4: Additional gene-OCR correlations Annotation on OCR UMAP from Fig 2 of OCRs with accessibility correlated (FigR $p < 0.05$) with expression of indicated genes (in red) or other OCRs within 15 kb of gene TSS not meeting this correlation significance threshold (in green).

Figure S5: Additional comparisons of spleen and colon Treg scATAC profiles A) Fraction reads in peaks vs number of unique reads per cell plot for quality control of **scATAC-seq data from Figure 4.** Dotted lines indicate thresholds used to select cells for analysis. B) UMAP of Treg scATAC-seq from Figure 4 colored by number of unique reads per cell. Dotted lines separate spleen and colon Tregs. C) Relative accessibility (chromVAR scores) of OCRs containing indicated TF motifs across spleen and colon Treg single cells visualized on scATAC UMAP from Figure 4, separated by organ. D) Relative accessibility (chromVAR scores) across spleen and

colon Treg single cells of OCRs from all topics visualized on scATAC UMAP from Figure 4, separated by organ.

Figure S6: Additional analysis of IL2 scATAC-seq data. A) Fraction reads in peaks vs number of unique reads per cell plot for quality control of scATAC-seq data from Fig 5. Dotted lines indicate thresholds used to select cells for analysis. B) Classification of cells in scATAC-seq data from Fig 5 by cell state based on relative accessibility of OCR signatures distinguishing these populations (derived from data in Fig 1). C) UMAP of scATAC-seq of IL2 -treated or -untreated splenic Tregs from Fig 5, colored by density of cells from each condition in each biological replicate. D) Overlap between IL2 treated and untreated cells within rTreg or aTreg pools, as measured by the Local Inverse Simpson's Index (LISI). LISI values range from 1 for complete separation to 2 for complete overlap (p value from Wilcoxon Rank Sum Test). E) Aggregated accessibility profiles at *Il2ra* locus of IL2 treated or untreated Treg single cells from scATAC-seq data in Figure 5, separated by cell state (rTreg or aTreg). F) Fold Change-Fold Change plot showing log_2Fold Change in accessibility of each OCR after IL2 treatment in aTreg vs rTreg conditions. Stat5 motif-containing OCRs are indicated in red. G) Distribution of per-cell raw accessibility (counts per million) of Stat5 motif-containing OCRs in control or IL2 treated cells compared across rTreg and aTreg cell states.

Figure S7: Additional F1 genetic variation analysis A) Fraction reads in peaks vs number of unique reads per cell plot for quality control of scATAC-seq data, separated by reads assigned to B6 or Cast alleles. Dotted lines indicate thresholds used to select cells for analysis. B) UMAP of F1 Treg scATAC-seq colored by number of unique reads per cell. C) log_2 (B6/Cast) Allelic Ratio for all OCRs aggregated across all F1 Treg single cells from Figure 6. Red line indicates median. D) Motif enrichment (left, grey indicates FDR<10^{-10}) and significant Topic AME (right, FDR<0.10) for distal OCRs (not filtered on intersection of analyses), with heatmap order as in Figure 6B, with motifs grouped by TF family.

Figure S8: Additional analysis of FoxP3 heterozygous female scATAC-seq data A) Sort strategy used for cell hashtagging in **WT and KO FoxP3 heterozygote scATAC-seq experiment in Fig 7.** B) Combined UMAP of Tconv, KO Treg, and WT Treg as in Figure 7, colored by Leiden clusters. C) scATAC-seq UMAP from Figure 7, colored by number of unique reads per cell. D) UMAP from Fig S8B separated by hashtag assignment, with colored boxes corresponding to sort

gates from S8A. E) Overlap between FoxP3 KO and WT populations (left) vs between Gata3 KO and WT populations (right), as measured by the Local Inverse Simpson's Index (LISI). LISI values range from 1 for complete separation to 2 for complete overlap (p value from Wilcoxon Rank Sum Test). F) Quantification of CD44+ Tregs from Fig 7D across biological replicates.

Figure S9: Analysis of FoxP3-bound regions and FoxP3-dependent chromatin accessibility changes A) Number of differential (average |log2FC|>0.1, p< 0.05) OCRs in rTreg or aTreg comparisons in data from Fig 7, colored by FoxP3 binding status. B) Fraction of WT Treg cells in which OCRs are accessible based on FoxP3 binding status (p value from Wilcoxon Rank Sum Test). C) OCR variance across WT Tregs based on FoxP3 binding status (p value from Wilcoxon Rank Sum Test). D) Differential accessibility per motif in each topic (distal OCRs) between WT and KO Treg cells in rTreg or aTreg comparisons for all motif to topic connections with significant AME as in Fig S7D. Grey indicates significant topic AME in Figure S7D but no significant (FDR < 0.05) change in accessibility across FoxP3 comparisons. **E) Differential accessibility per motif in each topic (distal OCRs) between KO rTreg and Tconv cells for all motif to topic connections with significant AME as in Fig S7D. Grey indicates significant topic AME in Figure S7D but no significant (FDR < 0.05) change in accessibility. F) Comparison of differential accessibility per topic in rTreg KO vs Tconv cells (y-axis) vs in rTreg WT vs KO (x-axis) cells.**

Figure S10: Additional analysis of FoxP3-independent RORγ+ Treg-like cells A) *Ikzf2* gene scores in WT and KO Tregs in heterozygous female scATAC-seq data from Fig 7 (p value from Wilcoxon Rank Sum Test). **B) Aggregated accessibility at *Ikzf2* and *Rorc* loci of single cells from WT or KO Tregs in heterozygous female scATAC-seq data from Fig 7. C) Expression of TF and cytokine transcripts overlaid onto colon scRNA-seq UMAP from Fig 8G. D) Aggregated accessibility at the *Foxp3* locus of single cells from WT Treg, KO Treg, and Tconv from heterozygous female scATAC-seq data from Fig 7, separated by cell state.**

Fig. S1

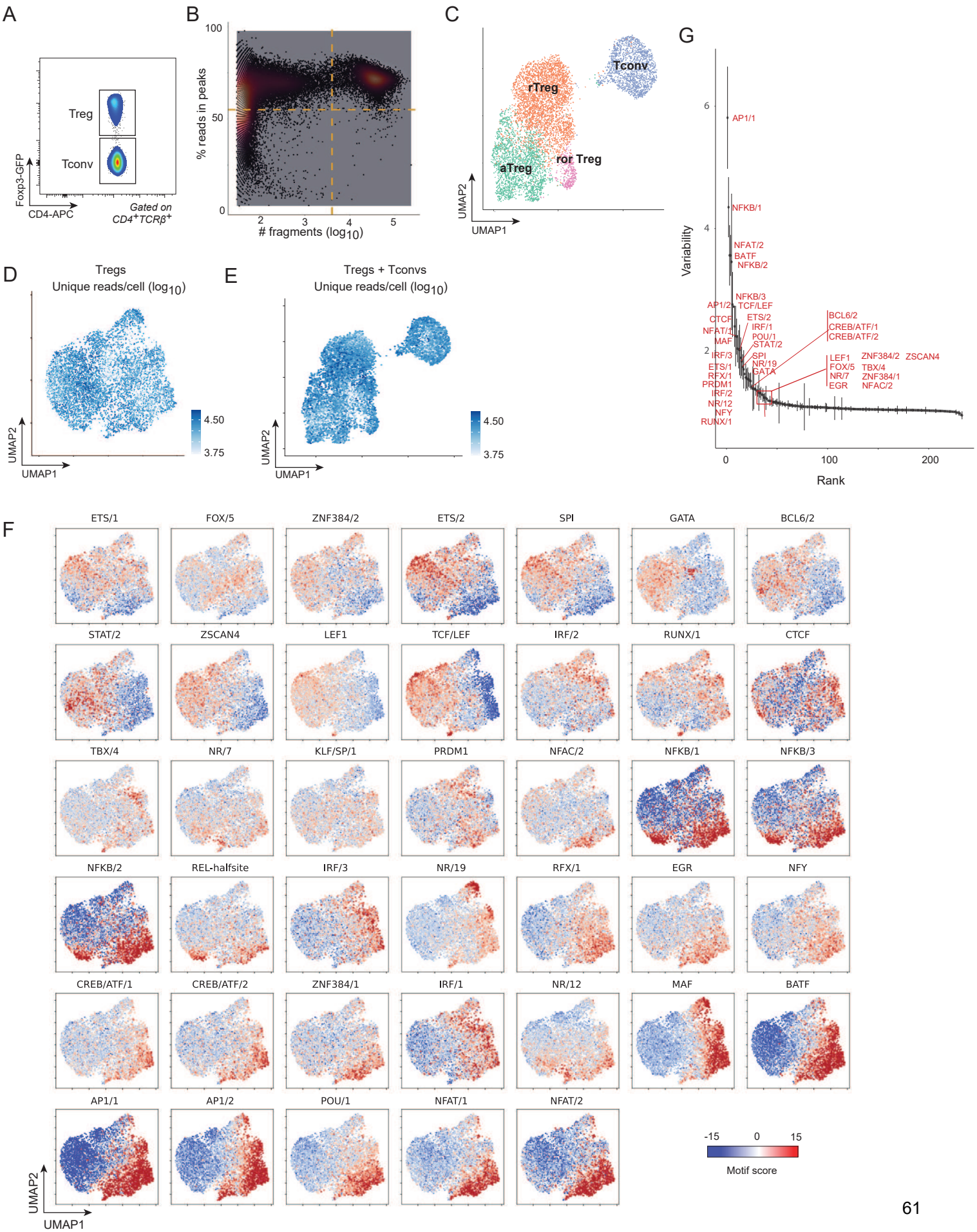
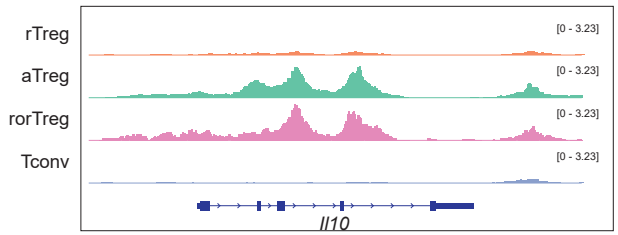
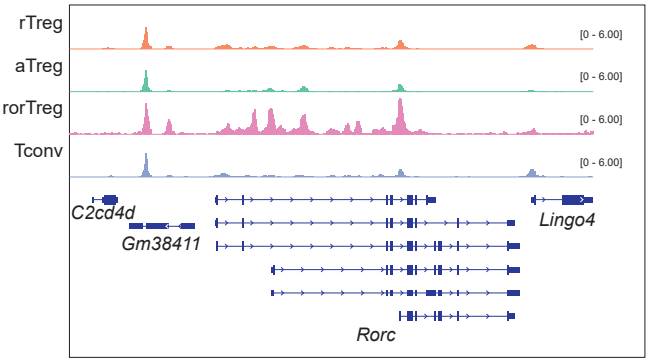
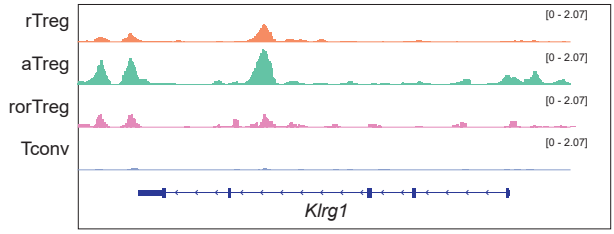
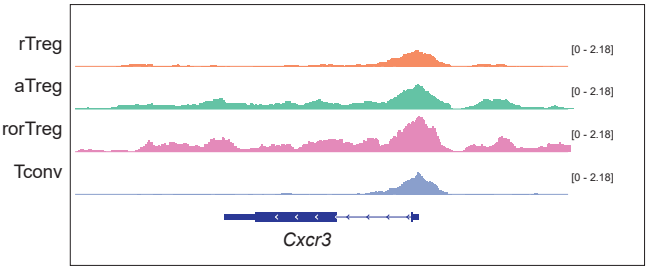

A
Treg
Tconv
Foxp3-GFP
CD4-APC
Gated on CD4+TCRβ+
B
% reads in peaks
# fragments (log10)
C
Tconv
rTreg
aTreg
ror Treg
UMAP2
UMAP1
D
Tregs
Unique reads/cell (log10)
4.50
3.75
E
Tregs + Tconvs
Unique reads/cell (log10)
G
Variability
Rank
AP1/1
NFKB/1
NFAT/2
BATF
NFKB/2
NFKB/3
AP1/2
TCF/LEF
CTCF
NFAT/1
MAF
IRF/3
ETS/1
RFX/1
PRDM1
IRF/2
NR/12
NFY
RUNX/1
ETS/2
IRF/1
POU/1
STAT/2
SPI
NR/19
GATA
BCL6/2
CREB/ATF/1
CREB/ATF/2
LEF1
ZNF384/2
ZSCAN4
FOX/5
TBX/4
NR/7
ZNF384/1
EGR
NFAC/2
F
ETS/1
FOX/5
ZNF384/2
ETS/2
SPI
GATA
BCL6/2
STAT/2
ZSCAN4
LEF1
TCF/LEF
IRF/2
RUNX/1
CTCF
TBX/4
NR/7
KLF/SP/1
PRDM1
NFAC/2
NFKB/1
NFKB/3
NFKB/2
REL-halfsite
IRF/3
NR/19
RFX/1
EGR
NFY
CREB/ATF/1
CREB/ATF/2
ZNF384/1
IRF/1
NR/12
MAF
BATF
AP1/1
AP1/2
POU/1
NFAT/1
NFAT/2
-15
0
15
Motif score
UMAP2
UMAP1

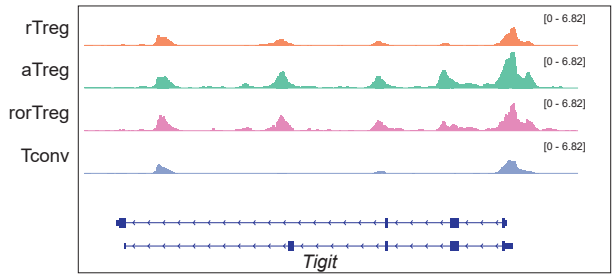
Fig. S2

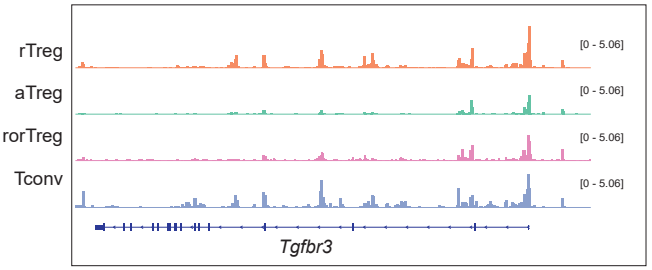

rTreg
aTreg
rorTreg
Tconv
[0 - 3.23]
[0 - 3.23]
[0 - 3.23]
[0 - 3.23]
Il10


rTreg
aTreg
rorTreg
Tconv
[0 - 6.00]
[0 - 6.00]
[0 - 6.00]
[0 - 6.00]
C2cd4d
Gm38411
Lingo4
Rorc


rTreg
aTreg
rorTreg
Tconv
[0 - 2.07]
[0 - 2.07]
[0 - 2.07]
[0 - 2.07]
Klrg1


rTreg
aTreg
rorTreg
Tconv
[0 - 2.18]
[0 - 2.18]
[0 - 2.18]
[0 - 2.18]
Cxcr3


rTreg
aTreg
rorTreg
Tconv
[0 - 6.82]
[0 - 6.82]
[0 - 6.82]
[0 - 6.82]
Tigit


rTreg
aTreg
rorTreg
Tconv
[0 - 5.06]
[0 - 5.06]
[0 - 5.06]
[0 - 5.06]
Tgfbr3

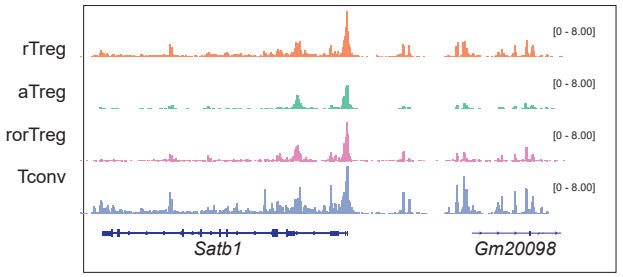
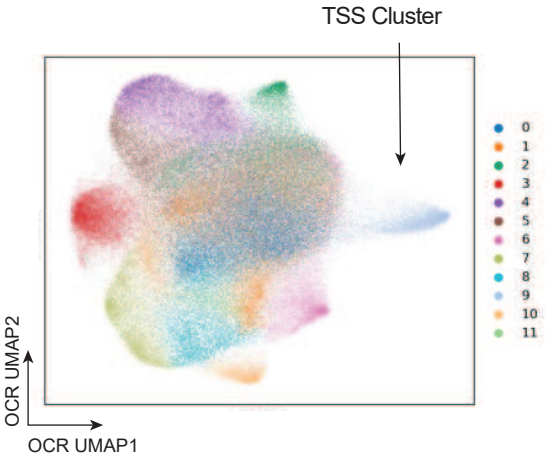

rTreg
aTreg
rorTreg
Tconv
[0 - 8.00]
[0 - 8.00]
[0 - 8.00]
[0 - 8.00]
Satb1
Gm20098

Fig. S3

A


TSS Cluster
0
1
2
3
4
5
6
7
8
9
10
11
OCR UMAP2
OCR UMAP1

B

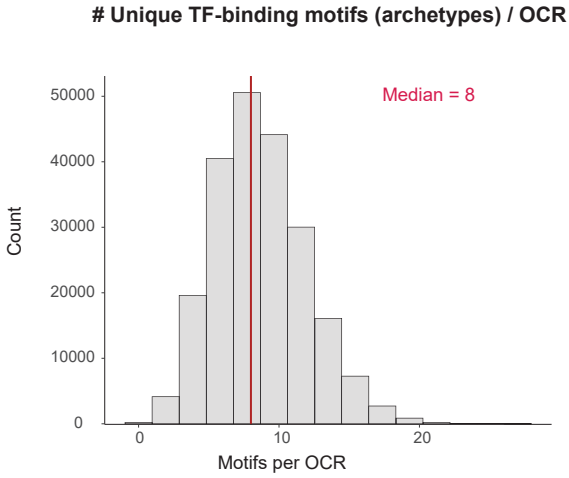

# Unique TF-binding motifs (archetypes) / OCR
Median = 8
Count
50000
40000
30000
20000
10000
0
0
10
20
Motifs per OCR

Fig. S4


Transcription Factors
Batf
Foxo1
Foxp1
Id2
Id3
Ikzf4
Lef1
Maf
Nfil3
Nfkb1
Nr4a1
Nr4a3
Rela
Satb1
Tbx21
Tcf7
Activation
Cd44
Cd69
Icos
S1pr1
Sell
Correlated OCRs
Other OCRs within 15kb
Treg effector molecules
Cd83
Ebi3
Gzmb
Il10
Il1rl1
Il2rb
Lag3
Lrrc32
Nt5e
Tgfbr3
Chemokine / Cytokine receptors
Ccr4
Ccr6
Ccr8
Cxcr3
Ifgae
Itgb1
Itgb8
TNF Receptor superfamily
Cd27
Tnfrsf18
Tnfrsf4
Tnfrsf8
Tnfrsf9
S100 family
S100a10
S100a11
S100a6
Interferon-related
Dusp2
Ifit1
Socs2
Other Treg genes
Igfbp4
Il18r1
Izumo1r
Nrp1
Pde3b
Ramp1
Spn
OCR UMAP2
OCR UMAP1

Fig. S5

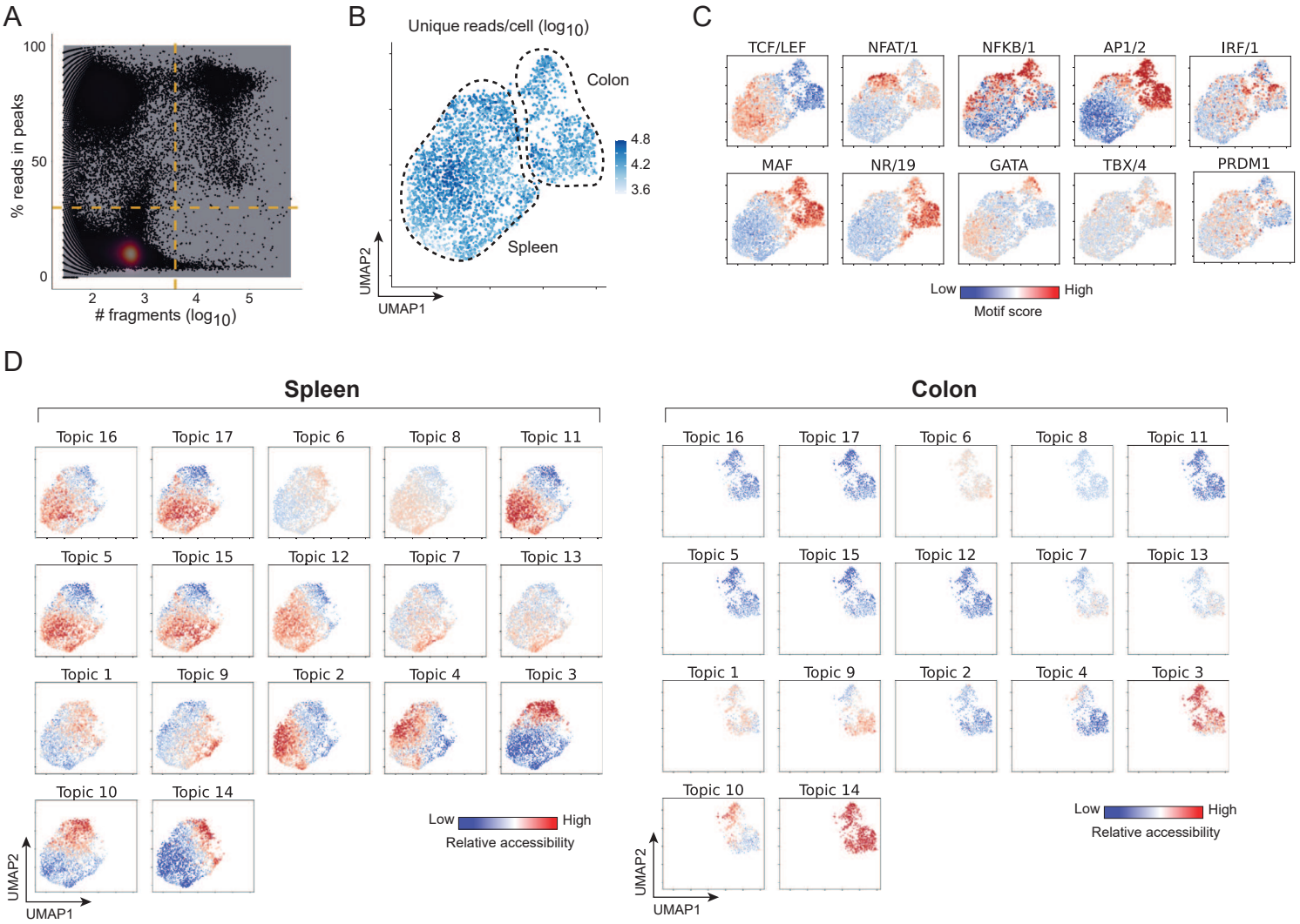

A
100
50
0
% reads in peaks
2
3
4
5
fragments (log10)
B
Unique reads/cell (log10)
Colon
Spleen
4.8
4.2
3.6
UMAP2
UMAP1
C
TCF/LEF
NFAT/1
NFKB/1
AP1/2
IRF/1
MAF
NR/19
GATA
TBX/4
PRDM1
Low
High
Motif score
D
Spleen
Topic 16
Topic 17
Topic 6
Topic 8
Topic 11
Topic 5
Topic 15
Topic 12
Topic 7
Topic 13
Topic 1
Topic 9
Topic 2
Topic 4
Topic 3
Topic 10
Topic 14
Low
High
Relative accessibility
UMAP2
UMAP1
Colon
Topic 16
Topic 17
Topic 6
Topic 8
Topic 11
Topic 5
Topic 15
Topic 12
Topic 7
Topic 13
Topic 1
Topic 9
Topic 2
Topic 4
Topic 3
Topic 10
Topic 14
Low
High
Relative accessibility
UMAP2
UMAP1

Fig. S6

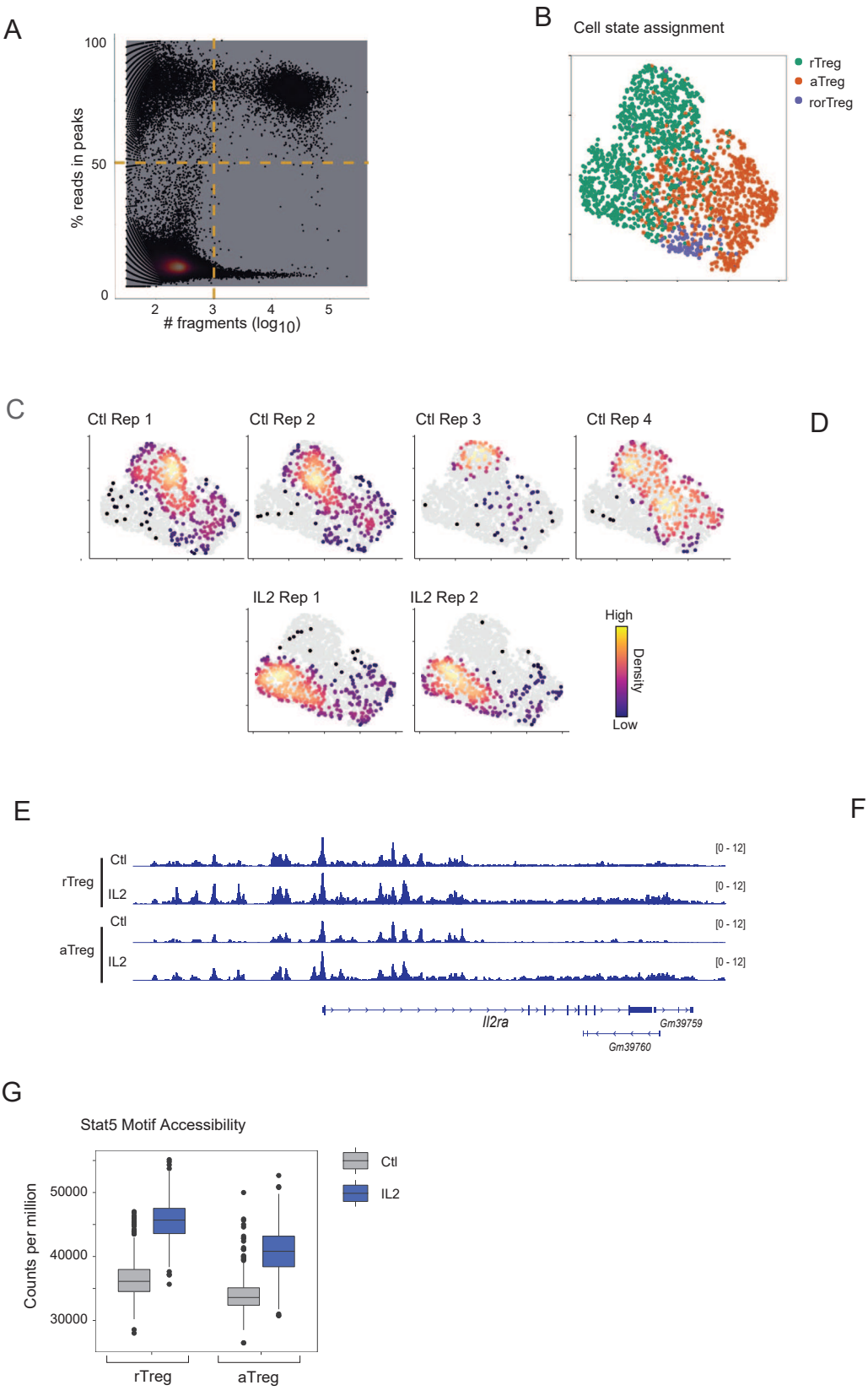

A
100
50
0
% reads in peaks
2
3
4
5
fragments (log10)
B
Cell state assignment
rTreg
aTreg
rorTreg
C
Ctl Rep 1
Ctl Rep 2
Ctl Rep 3
Ctl Rep 4
IL2 Rep 1
IL2 Rep 2
High
Density
Low
D
p < 2.2x10-6
Local Inverse Simpson's Index
2
1.5
1
rTreg
aTreg
E
rTreg
aTreg
Ctl
IL2
[0 - 12]
Il2ra
Gm39759
Gm39760
F
Stat5 motif
FoldChange aTreg (log2)
2.5
0
-2.5
FoldChange aTreg (log2)
G
Stat5 Motif Accessibility
Counts per million
50000
40000
30000
rTreg
aTreg
Ctl
IL2

Fig. S7

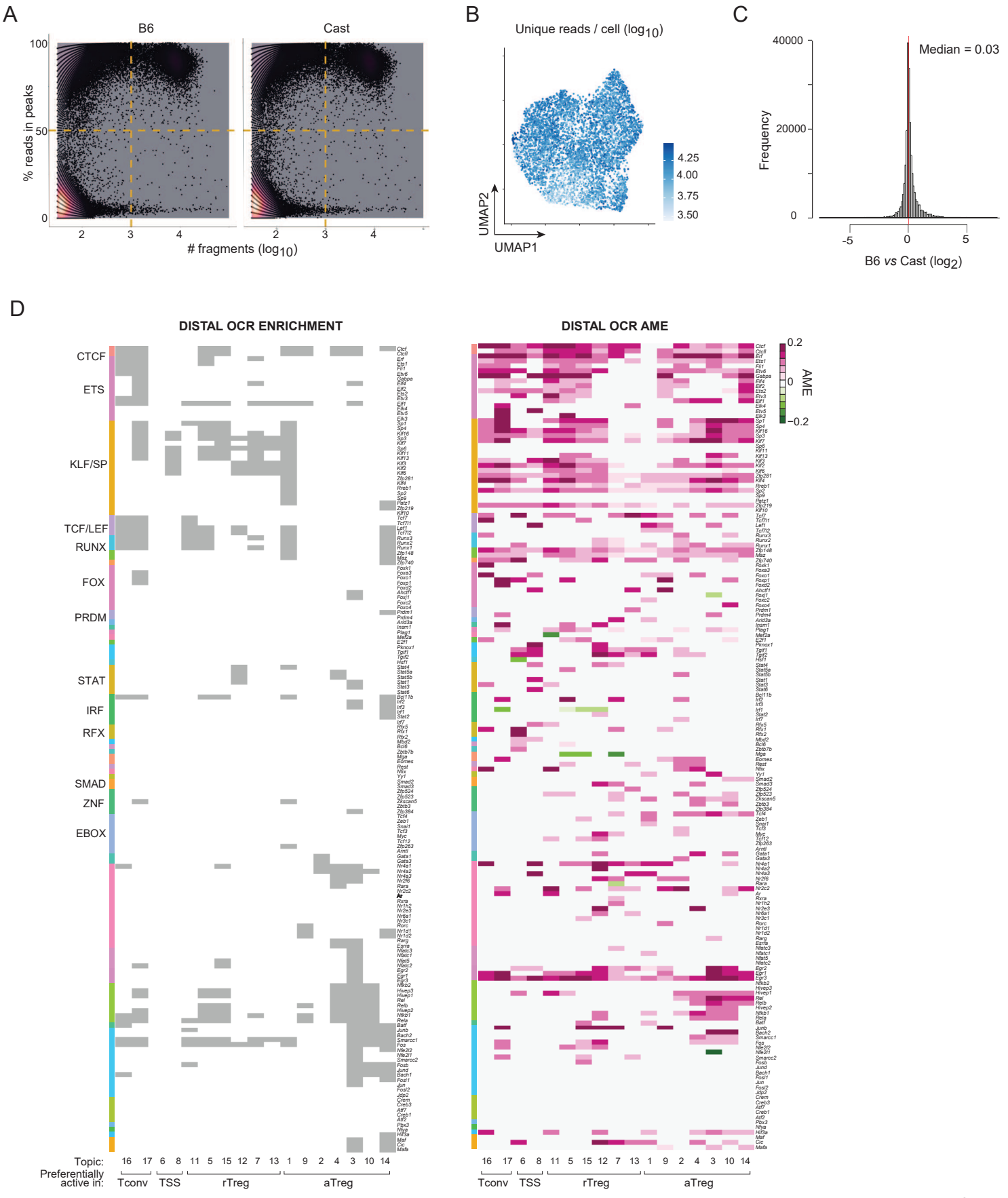

A
B6
Cast
% reads in peaks
# fragments (log10)
B
Unique reads / cell (log10)
UMAP2
UMAP1
C
Median = 0.03
Frequency
B6 vs Cast (log2)
D
DISTAL OCR ENRICHMENT
DISTAL OCR AME
CTCF
ETS
KLF/SP
TCF/LEF
RUNX
FOX
PRDM
STAT
IRF
RFX
SMAD
ZNF
EBOX
AME
Topic:
Preferentially active in:
Tconv
TSS
rTreg
aTreg

Fig. S8

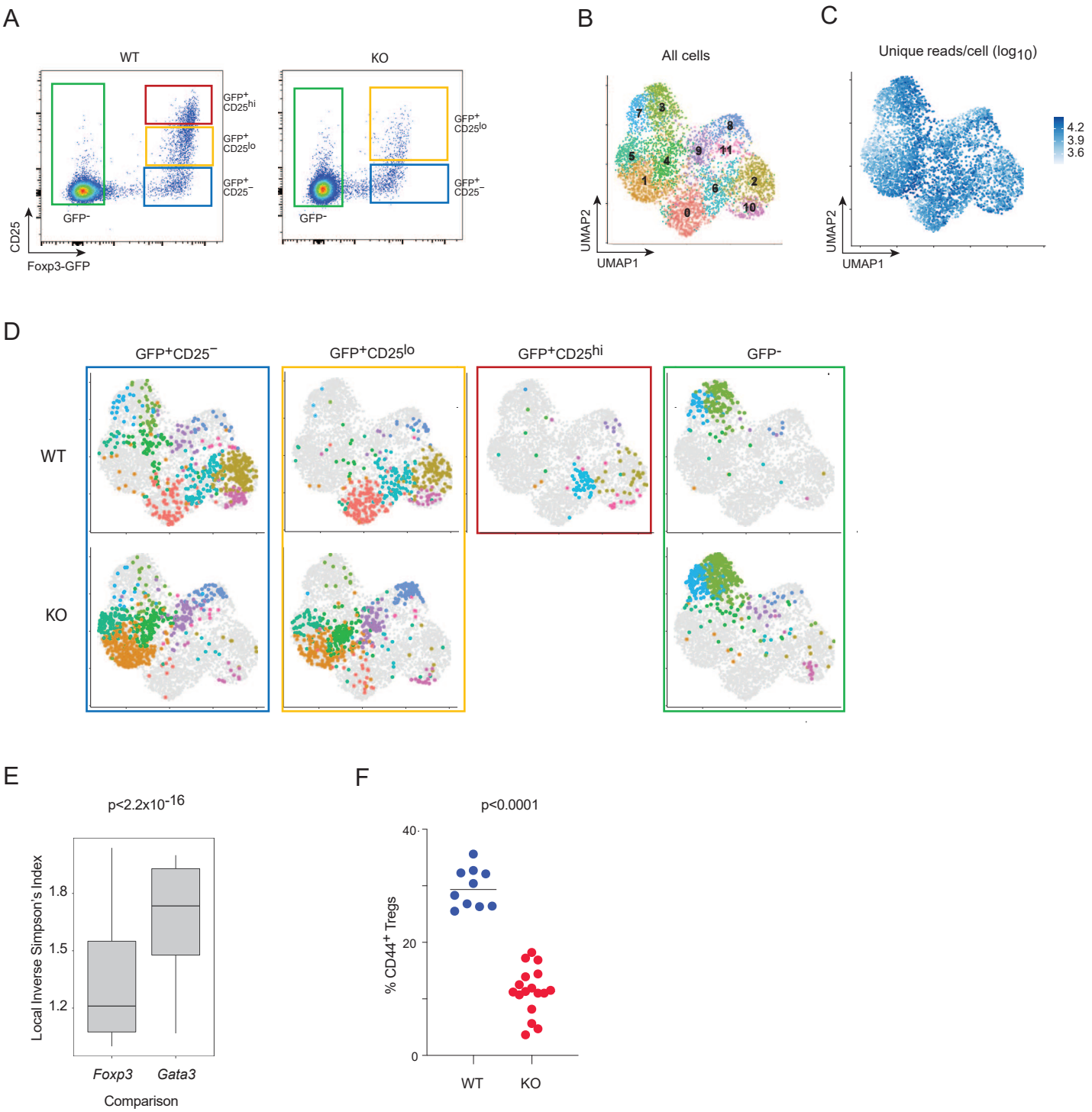

A
WT
KO
GFP+ CD25hi
GFP+ CD25lo
GFP+ CD25−
GFP-
CD25
Foxp3-GFP
B
All cells
UMAP2
UMAP1
C
Unique reads/cell (log10)
4.2
3.9
3.6
D
GFP+CD25−
GFP+CD25lo
GFP+CD25hi
GFP-
WT
KO
E
p<2.2x10-16
Local Inverse Simpson's Index
Foxp3
Gata3
Comparison
F
p<0.0001
% CD44+ Tregs
WT
KO

Fig. S9

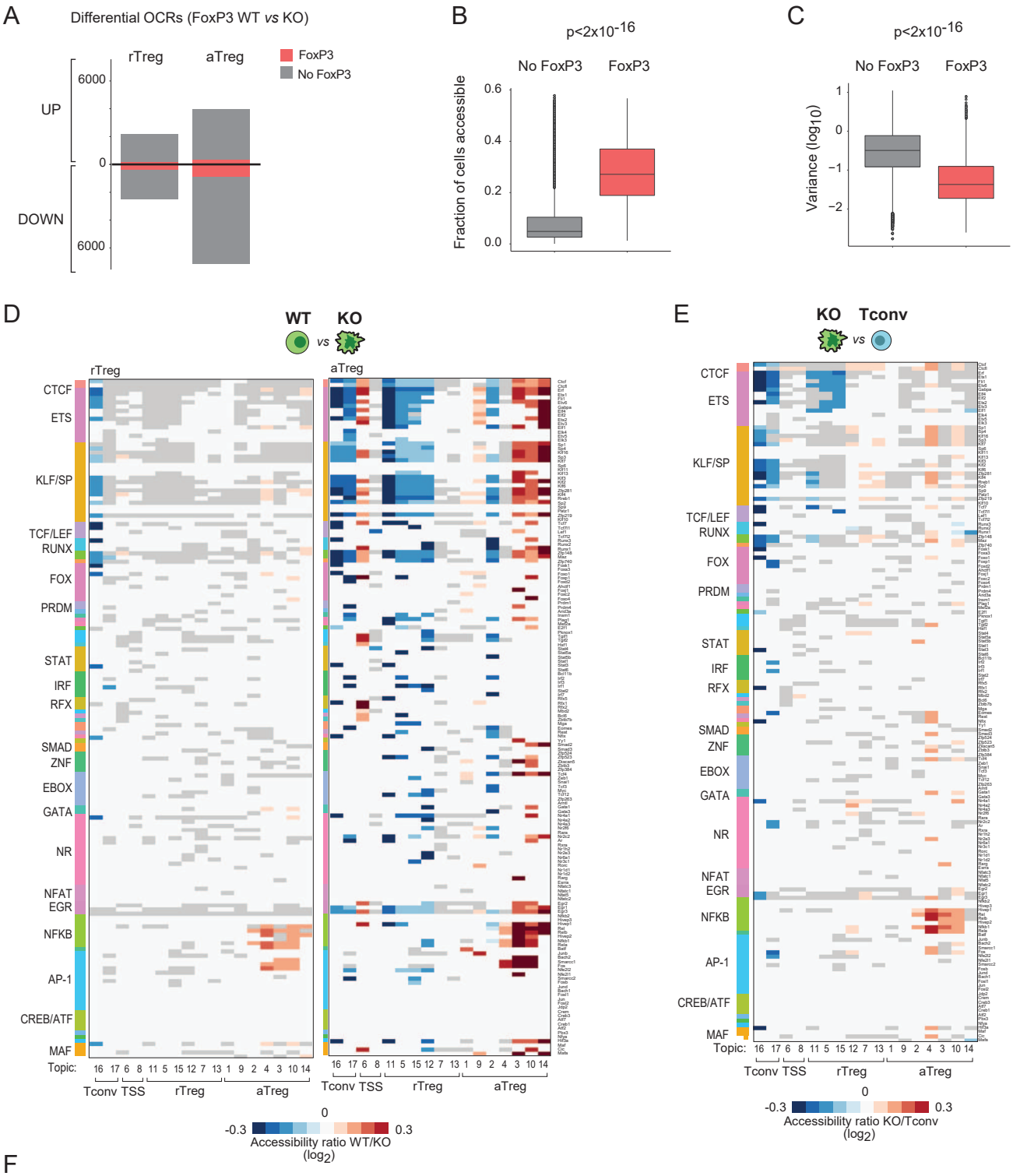
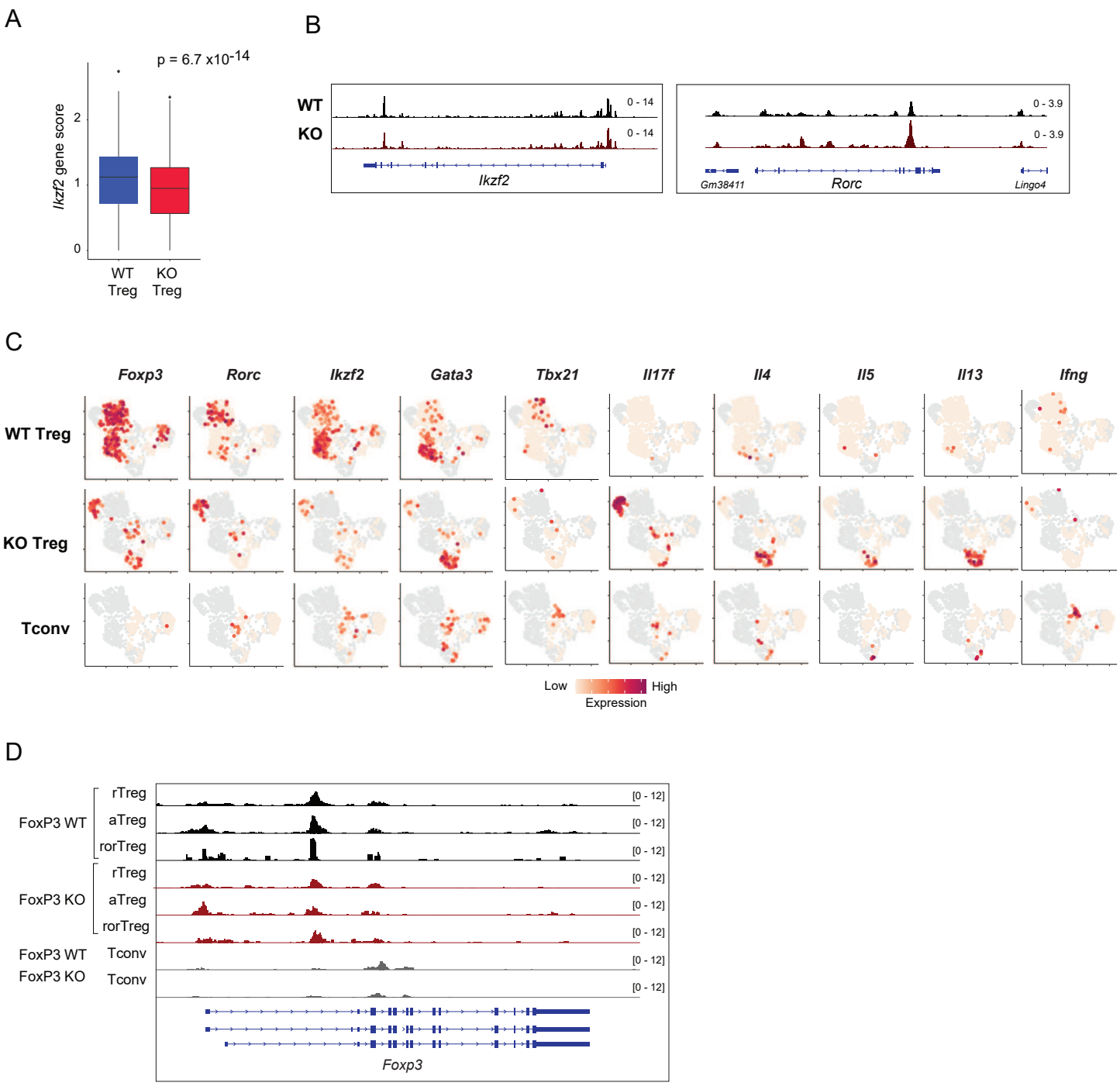

A
Differential OCRs (FoxP3 WT vs KO)
rTreg
aTreg
FoxP3
No FoxP3
UP
DOWN
6000
0
6000
B
p<2x10-16
No FoxP3
FoxP3
Fraction of cells accessible
0.6
0.4
0.2
0.0
C
p<2x10-16
No FoxP3
FoxP3
Variance (log10)
1
0
−1
−2
D
WT
vs
KO
rTreg
aTreg
CTCF
ETS
KLF/SP
TCF/LEF
RUNX
FOX
PRDM
STAT
IRF
RFX
SMAD
ZNF
EBOX
GATA
NR
NFAT
EGR
NFKB
AP-1
CREB/ATF
MAF
Topic:
16 17 6 8 11 5 15 12 7 13 1 9 2 4 3 10 14
Tconv TSS rTreg aTreg
Tconv TSS rTreg aTreg
0
-0.3
0.3
Accessibility ratio WT/KO
(log2)
E
KO
vs
Tconv
CTCF
ETS
KLF/SP
TCF/LEF
RUNX
FOX
PRDM
STAT
IRF
RFX
SMAD
ZNF
EBOX
GATA
NR
NFAT
EGR
NFKB
AP-1
CREB/ATF
MAF
Topic:
16 17 6 8 11 5 15 12 7 13 1 9 2 4 3 10 14
Tconv TSS rTreg aTreg
0
-0.3
0.3
Accessibility ratio KO/Tconv
(log2)
F
KO vs Tconv
rWT vs rKO
0.4
0
-0.4
-0.4
0
0.4
4
7
13
10
12
2
6
8
14
9
5
15
11
17
16

Fig. S10


A
p = 6.7 x10-14
Ikzf2 gene score
2
1
0
WT Treg
KO Treg
B
WT
KO
0 - 14
0 - 14
Ikzf2
0 - 3.9
0 - 3.9
Gm38411
Rorc
Lingo4
C
Foxp3
Rorc
Ikzf2
Gata3
Tbx21
Il17f
Il4
Il5
Il13
Ifng
WT Treg
KO Treg
Tconv
Low
High
Expression
D
FoxP3 WT
rTreg
aTreg
rorTreg
FoxP3 KO
rTreg
aTreg
rorTreg
FoxP3 WT
Tconv
FoxP3 KO
Tconv
[0 - 12]
Foxp3

SUPPLEMENTARY TABLE LEGENDS

TABLE S1: Sample characteristic for all datasets included in this study. A) scATAC, scRNA, and bulk ATAC dataset metrics. B) Hashtags used for ASAP-seq or scRNA-seq hashtagging in each experiment.

TABLE S2: Treg OCRs. Information on ImmGen+Treg OCRs used throughout this work, as well as additional information related to OCR UMAP and condition-specific changes in accessibility

TABLE S3: Chromatin Accessibility and Gene Expression Signatures. OCRs and genes used as differential signatures. In cases where external signatures were used, the provided table indicates their overlap with the reference OCRs from this study. A) aTreg vs rTreg OCRs (from[47]). B) aTreg vs rTreg OCRs (from scATAC-seq, this study). C) Treg vs Tconv OCRs (from scATAC-seq, this study). D) Spleen RORγ+ Treg vs aTreg OCRs (from scATAC-seq, this study). E) FoxP3 WT vs KO OCRs (from scATAC-seq, this study). F) Pan-tissue Treg OCRs (from[16]). G) Gata3- (from scATAC-seq, this study) and cMaf-dependent (from[98]) OCRs. H) RORγ+ and Helios+ Treg-specific gene signatures (derived from[132]). **I) Smarcc1-dependent OCRs (from bulk ATAC-seq, this study).**

TABLE S4: TF-Binding Table. Overlap of TF-binding datasets with reference OCR set from this study

TABLE S5: Gene-OCR associations. Gene-OCR correlations from FigR ($p<0.05$) (Association = 2 in table) or other OCRs with 15kb of gene TSS (Association = 1 in table)

TABLE S6: Topic Assignments. OCR membership in each topic

TABLE S7: Treg Regulatory Network A) NonTSS Topic Motif Enrichment. B) TSS Topic Motif Enrichment. C) NonTSS Topic AME, unfiltered. D) NonTSS Topic AME, overlap with enrichment. E) FoxP3-dependent (WT vs KO) differential accessibility in topics for rTreg, all AME. F) FoxP3-dependent (WT vs KO) differential accessibility in topics for rTreg, enrichment-AME overlap. G) FoxP3-dependent (WT vs KO) differential accessibility in topics for aTreg, all AME. H) FoxP3-dependent (WT vs KO) differential accessibility in topics for aTreg, enrichment-AME overlap**. I) FoxP3-independent (KO vs Tconv) differential accessibility in topics, all AME. J) FoxP3-**

independent (KO vs Tconv) differential accessibility in topics, enrichment-AME overlap. K) Motif Grouping Annotation Order. Note: for E-J, a value of 300 corresponds to grey cells in Fig 6 and S7, where there is a significant topic AME (FDR<0.10) but no significant differential accessibility (FDR<0.05).

TABLE S8: Topic GREAT Gene Set Enrichment Results related to Figure 3E. Gene Ontology gene sets (from GREAT[76]) significantly enriched among regulatory regions in each topic. Values in table indicate fold change of enrichment relative to background.

TABLE S9: IL2-dependent OCRs Results related to Figure 5G. A) Fold Change (log_2) of aggregated accessibility in treated versus untreated cells, matched for cell state (rTreg or aTreg) for OCRs in Fig 5G. Each column is a different biological replicate. B) HOMER TF motif enrichment statistics in each IL2 response cluster (from Fig 5G). C) -log_{10}(FDR) of Topic OCR enrichments in each IL2 response cluster (from Fig 5G).

TABLE S10: Additional OCRs. Information on additional OCRs used in F1, colon/spleen, and **IL2 analyses.**

APPENDIX D, *Résumé détaillé*

Introduction :
Le facteur de transcription FoxP3 orchestre le développement et la fonction des cellules T régulatrices (Tregs), acteurs clés du maintien de l'homéostasie immunitaire et du contrôle des réponses immunitaires anormales et excessives. En effet, les mutations de FoxP3 entraînent une dysfonction des Tregs, à l'origine d'une maladie auto-immune multisystémique, s'illustrant par le syndrome IPEX (Immune dysregulation, Polyendocrinopathy, Enteropathy, X-linked) chez l'humain et par le phénotype Scurfy chez la souris. FoxP3 fait partie de la grande famille des facteurs de transcription possédant un domaine de liaison à l'ADN de type Forkhead (la famille FOX), et comprends plusieurs domaines fonctionnels : la partie N-terminale qui semple être le lieu de l'interaction de nombreux cofacteurs, le domaine en doigt de zinc de fonction indéterminée, le domaine zipper à leucines qui joue un rôle dans la dimérisation et enfin le forkhead domain (FKHD) qui est le domaine de liaison à l'ADN. Il a été suggéré que FoxP3 fonctionne grâce à une combinaison d'interactions moléculaires, de modifications épigénétiques et de mécanismes de régulation génique. Notamment, il joue un rôle clès pour l'induction d'une partie de la signature transcriptomique spécifique des Tregs appelée la « signature Treg». Cependant, les mécanismes précis par lesquels FoxP3 opère au sein des Tregs ne sont toujours pas entièrement compris. FoxP3 n'est pas un facteur « Pioneer ». En effet, il se fixe sur des séquences d'ADN déjà accessibles. Il a été montré qu'ainsi FoxP3 n'est ni suffisant, ni nécessaire pour le développement d'un Treg.
En particulier, la manière dont les variations de séquences de FoxP3 influencent la fonction des Tregs, et son mode de collaboration avec d'autres facteurs de transcription, reste élusif. Il existe une importante variabilité phénotypique interindividuelle dans le syndrome IPEX, même parmi les membres d'une même famille portant la même mutation. Cependant, le phénotype semble particulièrement sévère lorsque le FKHD est muté, et certaines atteintes d'organes, comme par exemple l'atteinte rénale, semble spécifique de certaines mutations faux-sens. Ainsi, notre étude cherche à élucider la structure-fonction de FoxP3, en analysant les effets moléculaires de mutations "naturelles" associées au syndrome IPEX, à la fois chez l'humain et dans des conditions contrôlées chez la souris. Nous postulons qu'il existe une explication moléculaire entre les mutations de FoxP3 trouvées chez les patients atteints d'IPEX et leur éventail de symptômes cliniques.

Résultats :
Les résultats de cette thèse sont présentés dans trois articles différents, les deux premiers déjà publiés (*Nature Immunology* et *Cell Reports*) et le dernier en révision dans le journal *Immunity*.

Le premier article s'est intéressé directement aux Tregs humains des patients présentant le syndrome IPEX. Les profils cytométriques et transcriptomiques de cellules T CD4+ sanguines de 15 patients atteints du syndrome IPEX ont été analysés, à la fois en population et en cellule unique. Les mutations des patients étaient réparties sur l'ensemble du gène *FOXP3*, affectant différents domaines. Ces patients présentaient des atteintes cliniques diverses, dont la distribution était semblable aux autres cohortes décrites. Plus de 90% d'entre eux présentaient une atteinte digestive, et la maladie s'était déclenchée pour la majorité avant deux mois. De cette analyse sont revenues plusieurs observations. La première est que tous les patients atteints du syndrome IPEX ont des Tregs « wannabe » (des cellules exprimant FoxP3 mutant). Ces Tregs wannabe apparaissent pour certains d'entre eux quasiment sains alors que d'autres très dysfonctionnants. La deuxième est que l'identification des effets propres à chaque mutation a été compliquée par la présence d'une signature transcriptomique monomorphe affectant

l'ensemble des $CD4^+$. Cette signature était extrinsèque aux Tregs : elle était absente des cellules $CD4^+$, Tregs et Tconvs, des mères hétérozygotes de patients atteints du syndrome IPEX, où les Tregs WT exercent une suppression dominante. L'analyse d'ontologie génique et de recherche de signature transcriptomique similaire a montré que cette signature ne ressemblait au signature habituelle de l'activation T. Ceci suggère que cet effet est lié au défaut de suppression des Tregs wannabe, et exacerbe les perturbations du compartiment T CD4+, à l'origine des signes cliniques. Finalement, en utilisant un modèle de souris chimérique (Transplantation de moelle osseuse avec reconstitution d'une moelle 50% FoxP3 WT et 50% déficiente en FoxP3), nous avons pu montrer qu'en cas d'absence complète de FoxP3 et sans effet confondant de l'inflammation, l'effet « intrinsèque » de FoxP3 était limité à la régulation d'une dizaine de gènes seulement, incluant *Il2ra* et *Ctla4.*

Dans la deuxième partie de la thèse, nous avons décidé d'utiliser uniquement des modèles murins. Cette approche permettait d'étudier les mutations de FoxP3 dans les conditions hémizygotes et hétérozygotes (sans le facteur confondant de l'effet extrinsèque), avec un nombre significatif de répétitions, et sans facteurs de confusion génétiques et environnementaux lorsque les souris étaient étudiées à l'état basal. Dans ce contexte, six mutations faux-sens, provenant de la cohorte IPEX précédente et réparties sur l'ensemble du locus FoxP3, ont été introduites chez la souris B6.FoxP3.ires.GFP. Ceci a été réalisé par la technique de CRISPR/Cas9. Ces mutations comprenaient : R51Q et C168Y localisées dans le domaine N-Terminal; K199del située dans le domaine en doigts de zinc; R309Q et F324L entre le domaine zipper à leucines et le domaine de fixation à l'ADN, et enfin R337Q, qui correspond au premier acide aminé du domaine de liaison à l'ADN, le FKHD. Notre analyse a été exhaustive, engloabnt l'étude de ces souris et de leur Tregs à la fois sur le plan clinique, sur le plan protéique (cytométrie en flux), sur le plan transcriptomique (RNAseq) et sur le plan épigénétique (scATACseq). Comme attendu, l'ensemble de ces mutations faux-sens a conduit à la production d'une protéine FOXP3 mutante, à des niveaux proches du FOXP3 de type sauvage (WT). Seule la mutation située dans le FKHD, domaine de liaison à l'ADN, engendrait une réduction significative de la capacité de FoxP3 à se lier à l'ADN. Notre analyse a par la suite révélé deux classes distinctes de mutations. La première représente les mutations situées dans le FKHD, illustrée dans notre étude par R337Q. Cette mutation a induit une lymphoprolifération et une infiltration multi-organe chez les mâles hémizygotes, qui étaient comparables à celle des souris totalement dépourvues de FoxP3 (KO), bien que l'apparition des symptômes était retardée. Ce phénotype était cohérent avec d'autres mutations du FKHD rapportées dans la littérature. En revanche, les mutations dans les autres domaines - C168Y, K199del, R309Q, F324L, R337Q - n'ont pas montré de phénotype clair à l'état stable. Les souris mâles et femelles provenant de ses lignées mutantes étaient dénuées de toute inflammation et leurs Tregs globalement semblables aux Tregs. Cependant, l'induction d'inflammation tissulaire (colite à DSS ou dermatite induite par MC903) ainsi que le croisement avec des allèles NOD ont révélé des maladies tissulaires spécifiques : dermatite, colite, diabète. De plus, analysant les Tregs de la rate chez la femelle (sans le potentiel facteur de confusion de l'inflammation), ces mutations avaient des impacts subtils mais distincts sur la signature ARN dépendante de FoxP3 ainsi que sur l'accessibilité de la chromatine. Ceci suggère que ces mutations pourraient perturber l'interaction de FoxP3 avec certains cofacteurs spécifiques. Malheureusement, aucune interaction claire n'a pu être identifiée au cours de cette analyse, possiblement biaisée par le fait que cette analyse a été réalisée sur des Tregs de la rate à l'état stable. L'impact mutationnel est très probablement dévoilé à l'état dynamique (post-trigger) et au niveau tissulaire. Enfin, en collaboration avec le laboratoire du Dr. Sun Hur, nous avons pu mettre en évidence un effet spécifique de la mutation R337Q, en comparaison à l'absence complète de FoxP3. R337Q impacte immédiatement la structure de FoxP3 et celui-ci ne peut

former qu'une dimérisation de type « swapped », et perdant l'une des formes fonctionnelles de FoxP3, à savoir le « head-to-head » dimère. Ce travail a également été publié dans *Immunity*.

Enfin, dans le dernier article, en utilisant un modèle de souris hétérozygotes FoxP3 KO, nous avons analysé l'impact global de FoxP3 sur la chromatine et sur le réseau de facteurs de transcription définissant les Tregs. Ce réseau transcriptionnel a été initialement établi par une analyse de l'accessibilité de la chromatine de Tregs de souris WT en cellule unique (scATACseq). Par le biais de modélisations mathématiques, cette analyse a révélé l'implication et l'imbrication de nombreux facteurs de transcription dans le contrôle épigénétique des Tregs. Dans un second temps, j'ai généré un modèle de souris FoxP3 KO associée à un GFP reporter. Ceci permet le marquage *in vivo* des Tregs wannabe, qui ont un locus FoxP3 accessible. En réalisant la même analyse sur ses Tregs Foxp3 KO wannabe, en particulier chez les femelles héterozygotes dépourvues d'inflammation pour étudier l'effet intrinsèque de FoxP3, nous avons pu montrer que FoxP3 influence ce réseau à travers des rôles à la fois répressifs et activateurs. Nous avons aussi identifié une sous-population de Tregs, les Tregs RORγ+, qui semble fonctionner indépendamment de FoxP3, et qui prédomine dans le colon. Cette population FoxP3-indépendante a ensuite été caractérisée sur le plan cytométrique et sur le plan transcriptomique par single cell RNAseq (scRNAseq) dans la rate et le colon. Les résultats suggèrent que ces cellules ont un profil à la frontière entre les lymphocytes T régulateurs et les lymphocytes Th17 mais qu'elles se rapprochent plus du profil d'expression génique des T régulateurs. Leur potentiel rôle fonctionnel n'a pas pu être étudié mais ils semblent dépendent du microbiote.

Discussion :
La discussion aborde en détail les principales découvertes mises en évidence par cette thèse. Parmi celles-ci, nous pouvons énumérer : (1) L'identification de l'origine de l'effet extrinsèque qui est communément partagé entre les cellules T régulatrices et les cellules T conventionnelles chez les patients atteints du syndrome IPEX ; (2) L'importance prédominante du domaine Forkhead dans la stabilité des Tregs. (3) L'interaction de la protéine FoxP3 avec ses cofacteurs via ses autres domaines à l'origine d'une tolérance spécifique d'organe ; (4) L'analyse du double rôle de FoxP3, qui se manifeste non seulement comme stabilisateur de l'identité des Tregs (rôle répresseur), mais aussi comme un élément essentiel dans leur fonction suppressive (rôle activateur) ; (5) La suggestion d'un potentiel rôle fonctionnel pour ce que nous avons nommé les "Tregs wannabe" chez les mères hétérozygotes.

Conclusion :
En conclusion, il apparaît que la physiopathologie de chaque mutation faux-sens est intrinsèquement liée à la position de cette mutation. Ceci est clairement démontré par le fait qu'une mutation affectant directement le FKHD peut, à elle seule, déclencher la maladie. Cependant, pour les mutations qui ne touchent pas directement le FKHD et qui perturbent probablement les interactions FoxP3-cofacteurs, il est nécessaire d'avoir une combinaison de facteurs à la fois génétiques et environnementaux pour que leurs effets se manifestent cliniquement. Cette interprétation donne une lumière nouvelle à l'hétérogénéité clinique souvent observée dans le syndrome IPEX. D'un point de vue moléculaire, le FKHD semble jouer un rôle central dans la stabilité générale du phénotype des Tregs et, par conséquent, dans leur capacité globale à supprimer. Tandis que d'autres parties de la molécule semblent être impliquées dans des fonctions plus ciblées, qui varient probablement en fonction du tissu en question et des cofacteurs associés.

APPENDIX E, *LISTE DES ÉLÉMENTS SOUS DROITS*

Université Paris Cité

LISTE DES ÉLÉMENTS SOUS DROITS

Liste de **tous les éléments retirés** de la version complète de la thèse

faute d'en détenir les droits

Illustrations, figures, images...

Légende de l'image	N° de l'image	Page(s) dans la thèse
Figure 13. Regulatory T (Treg) cell heterogeneity and suppression of distinct classes of the immune response, from (Josefowicz, Lu and Rudensky, 2012).	Figure 2	15
Figure 14. Structure of the Forkhead DNA binding domain from (Dai *et al.*, 2021)	Figure 3	19
Figure 15: Typical and unusual clinical features in IPEX, from (Consonni, Ciullini Mannurita and Gambineri, 2021).	Figure 6	27
Figure 7: Schematic representation of all reported FOXP3 mutations in IPEX patients from Park et al.	Figure 7	30
Figure 11. The FoxP3 interactome, figure from (Georgiev, Charbonnier and Chatila, 2019)	Figure 11	40