PNAS RESEARCH ARTICLE | IMMUNOLOGY AND INFLAMMATION

An integrated transcription factor framework for Treg identity and diversity

Kaitavjeet Chowdhary[a], Juliette Léon[a,b], Diane Mathis[a], and Christophe Benoist[a,1]

Contributed by Christophe Benoist; received June 5, 2024; accepted July 12, 2024; reviewed by Tadatsugu Taniguchi and Ye Zheng

Vertebrate cell identity depends on the combined activity of scores of transcription factors (TF). While TFs have often been studied in isolation, a systematic perspective on their integration has been missing. Focusing on FoxP3+ regulatory T cells (Tregs), key guardians of immune tolerance, we combined single-cell chromatin accessibility, machine learning, and high-density genetic variation, to resolve a validated framework of diverse Treg chromatin programs, each shaped by multi-TF inputs. This framework identified previously unrecognized Treg controllers (*Smarcc1*) and illuminated the mechanism of action of FoxP3, which amplified a pre-existing Treg identity, diversely activating or repressing distinct programs, dependent on different regulatory partners. Treg subpopulations in the colon relied variably on FoxP3, Helios+ Tregs being completely dependent, but RORγ+ Tregs largely independent. These differences were rooted in intrinsic biases decoded by the integrated framework. Moving beyond master regulators, this work unravels how overlapping TF activities coalesce into Treg identity and diversity.

autoimmunity | immunoregulation | gene expression

Significance

The genetic regulatory network of the crucial Treg lineage is controlled by a number of transcriptional regulators, foremost FoxP3. How these various inputs are coordinated is poorly understood. Combining machine learning and high-density genetic variation, with several modes of genetic engineering for validation, we arrive at a coherent multifactor model that accounts for the role of FoxP3, its activating and repressive effects, and how these are integrated in different Treg phenotypes.

Cell identities are defined by characteristic gene expression programs controlled by specific transcription factors (TFs), which bind and regulate target *cis*-regulatory elements (CRE) (1). TFs act combinatorially, by binding to adjacent regulatory elements, assembling into complexes, or organizing into transcriptional networks. Some TFs ("master TFs") initiate cell-type-specific programs (1, 2). Others are required for responses to extracellular cues, such as from cell cross talk or environmental perturbations, leading to further differentiation or cell-state adaptations (3). TF action is context-dependent and can flip between activating or repressive at different CREs, influenced by local cofactors, ligand-mediated allosteric changes, or posttranslational modifications. Countless studies have tackled how individual TFs, or small combinations thereof, partake in determining cell-type identity or diversifying cell states. However, beyond focused one-TF forays, an integrated perspective of how the hundreds of TFs expressed by any one cell combine to form core cell identity and enable functional plasticity has been missing.

Regulatory T cells (Treg), a subset of CD4+ T lymphocytes, are among the most closely studied cell-types (4–7). They are dominant controllers of immunologic and organismal homeostasis, with diverse effector functions distributed across varied phenotypic poles (8). For example, distinct Treg phenotypes preferentially restrain Type-1 or -2 inflammation (9, 10), and unique Treg populations in nonlymphoid tissues facilitate tissue regeneration (11), enforce tolerance to commensal microbes (12), or control extraimmunologic consequences of inflammation (13). Treg specialization is undergirded by characteristic molecular programs (9, 10, 14–19). This specialization is often considered in terms of a "one TF-one state" model (4, 20), in which the expression of single context-specific TFs (T-bet, PPARγ, RORγ, cMAF, BATF), along with FoxP3, mediates differentiation of each Treg subpopulation (9, 18, 21–25), although more combinatorial models have also been considered (26, 27). However, how the many TFs expressed in Tregs are systematically organized to determine Treg identity and diversity remains uncharted. Furthermore, although FoxP3 expression defines Treg identity, its mechanism of action has not been resolved. It is neither necessary nor sufficient to establish Treg identity, Treg-specific transcriptomes only partially require FoxP3 (28–33), and its molecular mechanism of action is debated. An integrative view of how FoxP3 affects TF control across diverse Treg states has so far been elusive.

Here, we integrated several orthogonal strategies to systematically connect TFs to their target Treg programs (Fig. 1*A*), focusing on the role of TFs in modulating chromatin accessibility to study the most proximal effects of TF action without confounders from transcriptional bursting or stability. Collectively, these results offer a holistic and clarifying perspective on how one cell-type coordinates combinations of TFs and CREs to achieve both its identity and phenotypic diversification.

Author affiliations: [a]Department of Immunology, Harvard Medical School, Boston, MA 02115; and [b]INSERM UMR 1163, Imagine Institute, University of Paris, Paris, France 75015

Author contributions: K.C., J.L., D.M., and C.B. designed research; K.C. and J.L. performed research; K.C. and J.L. analyzed data; C.B. supervised; and K.C., J.L., and C.B. wrote the paper.

Reviewers: T.T., The University of Tokyo; and Y.Z., Salk Institute for Biological Studies.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: cbdm@hms.harvard.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2411301121/-/DCSupplemental.

Published August 28, 2024.

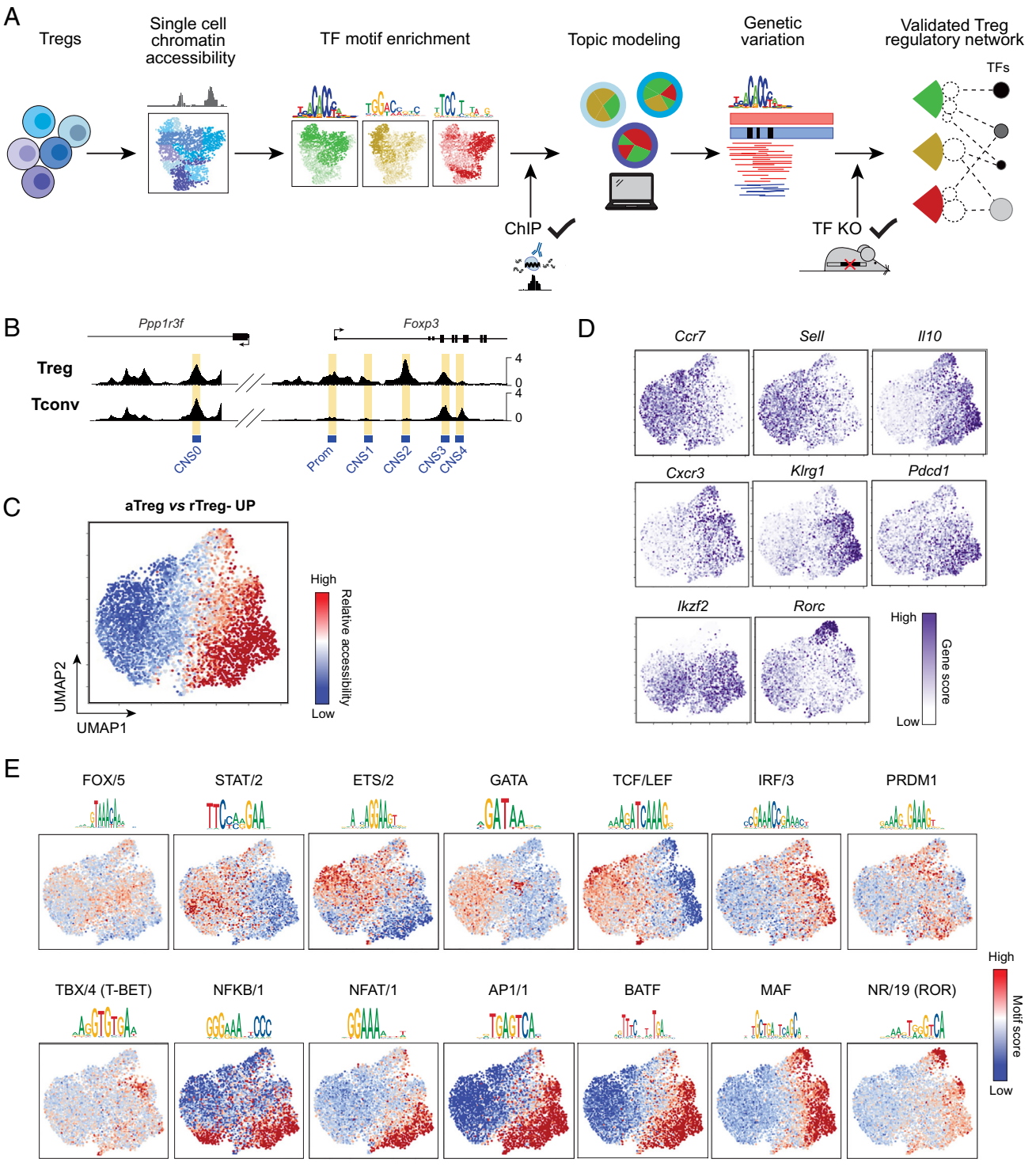

Fig. 1. Single-cell ATAC-seq reveals imbricated TF activities across diverse Treg cell states. (*A*) Experimental overview. Single-cell chromatin accessibility profiling was used to link TF activities to diverse Treg chromatin states. Continuous Treg cell states were first annotated using TF motif enrichment. Topic modeling was used to learn groups of covarying 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 framework was validated using TF binding (ChIP-seq, CUT&RUN) and knockout datasets. All generated datasets and associated metrics are described in Dataset S1. (*B*) Aggregated accessibility profiles of splenic Treg and Tconv single cells at the *Foxp3* locus from scATAC-seq data generated from a *Foxp3*$^{IRES\text{-}GFP}$ reporter mouse; highlights indicate conserved noncoding 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. 34) 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" (35) 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.

Results

Imbricated Combinations of TF Activities Underlie Treg Diversity. To connect the activity of TFs to the diversity of Treg states, we generated single-cell chromatin accessibility profiles (scATAC-seq) from splenic green fluorescent protein (GFP)+ Treg or T conventional (Tconv) cells from a *Foxp3*$^{\text{IRES-GFP}}$ reporter mouse (*SI Appendix*, Fig. S1 *A–E*). Tregs were clearly distinct from Tconv in a two-dimensional Uniform Manifold Approximation and Projection (UMAP) visualization (*SI Appendix*, Fig. S1*C*), and genome traces of aggregated reads recapitulated known patterns of *Foxp3* locus accessibility (Fig. 1*B* and *SI Appendix*, Fig. S1*F*). Within the Treg pool, cells separated broadly, as indicated by relative accessibility of open chromatin region (OCR) signatures that distinguish aTreg and rTreg populations (Fig. 1*C* and Dataset S2) (34). More granular annotation using "gene scores," chromatin-based proxies for gene expression (36), confirmed chromatin signatures (e.g., *Il10* vs. *Ccr7* and *Sell*) and further delineated aTreg populations marked by differential gene scores for *Cxcr3*, *Klrg1*, and *Pdcd1* (Fig. 1*D*), each markers of distinct, physiologically important Treg poles (9, 17, 18, 37). Rare populations were also represented, such as cells with high *Rorc* gene scores (24, 25), corresponding to a Treg subset that dominates in the colon, but also present at low levels in the spleen (24, 25) (Fig. 1*D*).

With this landscape of spleen Treg heterogeneity, we turned 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" (35) to reduce redundancy (Fig. 1*E* and *SI Appendix*, Fig. S2*A*). Activation-related motifs (AP-1, NF-AT, or NF-κB) had the greatest variability across single cells, cleanly partitioning the Treg pool (Fig. 1*E* and *SI Appendix*, Fig. S2*B*). 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. Accessibility of binding motifs of TFs classically ascribed to Th and Treg functional subtypes (GATA, TBX, BCL6) were broadly distributed, without the discrete and mutually exclusive patterns of activity that one would expect from the commonly evoked Treg sublineages (4, 20). This perspective also synthesized disparate results about individual TFs and their relevance to Treg physiology. For instance, although both cMaf and RORγ have been reported to be active within the same population (38), Treg-specific cMaf knockouts have broader phenotypes (39). 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. Similarly, while some studies have suggested BATF to be a main driver of tissue-Treg programs, its ablation in Tregs affected only a fraction of tissue-Treg-related OCRs (18, 40). Here, relative accessibility of the BATF motif did not stand out from several other motifs with similar patterns. 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 overlapping combinations of multiple factors constitute the diversity of Treg programs.

OCR Usage in Treg Single Cells. To understand how this phenotypic variance at the cell level arose from the activity of individual OCRs, we visualized the landscape of OCR usage across Treg single cells ("OCR UMAP"; *SI Appendix*, Fig. S3*A*). OCR activity formed graded continua, with most variability contributed by non-TSS (transcriptional start site), distal OCRs, consistent with prior observations (41–44) (*SI Appendix*, Fig. S3 *B* and *C*). Cell- and state-specific OCRs grouped together on the OCR UMAP visualization (*SI Appendix*, Fig. S3*D* and Dataset S2). Positioning TFs motifs onto the OCR UMAP demarcated distinct [STAT, NF-κB, and NR/19 (RORγ)] yet overlapping (NF-AT, AP-1, BATF) accessibility patterns in the OCR space (*SI Appendix*, Fig. S3*E*). Experimentally observed TF binding sites (from chromatin immunoprecipitation with sequencing (ChIP-seq) and Cleavage Under Targets and Release Using Nuclease (CUT&RUN)) (34, 45–49) (Dataset S6) also had variegated configurations (*SI Appendix*, Fig. S3*F*), with preferential but not exclusive binding to distinct groups of state-specific OCRs (e.g., Lef1, TCF-1 among rTreg-biased loci and Bach2, JunD among aTreg-biased loci). FoxP3-binding had a narrow footprint on the OCR landscape (*SI Appendix*, Fig. S3*F*). 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 CREs (50–52). This multiplicity confers evolutionary robustness but also enables the expression of one gene in different differentiated cell types (53–55). 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. 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) (56), also annotating OCRs within 15 kb of each TSS that did not meet correlation criteria (*SI Appendix*, Fig. S3*G* and Dataset S7). While some genes had linked OCRs with homogeneous activity patterns (e.g., *Rorc, Klrg1*), others had more varied distributions (e.g., *Ccr7*) or even multiple subpatterns (e.g., *Tigit, Pdcd1*, *Ctla4*), with disparate accessibility profiles among loci previously shown to control *Foxp3* expression (57). Thus, genes vary in how flexibly their associated OCRs are used across Treg states.

An Integrated Treg TF Framework. Individual TFs were thus unable to order Treg diversity but could one instead group OCRs with covarying accessibility into a stable framework of coregulated modules and then determine the TF combinations that drive them? We combined the power of computational machine learning and the causality that can be inferred from high-density genetic variation to establish such a framework. We employed a machine learning approach, probabilistic topic modeling, to categorize covarying OCRs into "topics" (58). Topic modeling is well suited to sparse, single-cell genomic data (59, 60), allowing for assignment of multiple programs per OCR, thus better reflecting the continuous structure of Treg OCR usage. We modified a previous method for topic modeling of scATACseq data (59) for robustness by using an ensemble approach to create a consensus set of 17 reproducible topics (*SI Appendix*, Computational Note 1; Dataset S3). Each topic captured diverse and subtle accessibility patterns across Treg single cells (*SI Appendix*, Fig. S4 *A* and *B*) with preferential accessibility in rTreg, aTreg, or Tconv cells (*SI Appendix*, Fig. S4*C*). Some topics, such as those (6 and 8) with an overrepresentation at TSS, were broadly accessible, while others captured highly specific gradations in accessibility. We computed TF motif enrichment within each topic, which proved differently distributed between distal enhancer or TSS OCRs (*SI Appendix*, Fig. S4*D* and Dataset S4). Distal enrichments highlighted state-specific TF connections, no topic being defined uniquely by enrichment of motifs of any one TF or family (Fig. 2*A*). While topics more accessible in aTregs were enriched in motifs for AP-1, NF-κB, and nuclear receptors, Tconv- and rTreg-preferential topics instead included TCF/LEF, ETS, and RUNX families (consistent with ref. 34; Fig. 2*A*). Topic modeling results were

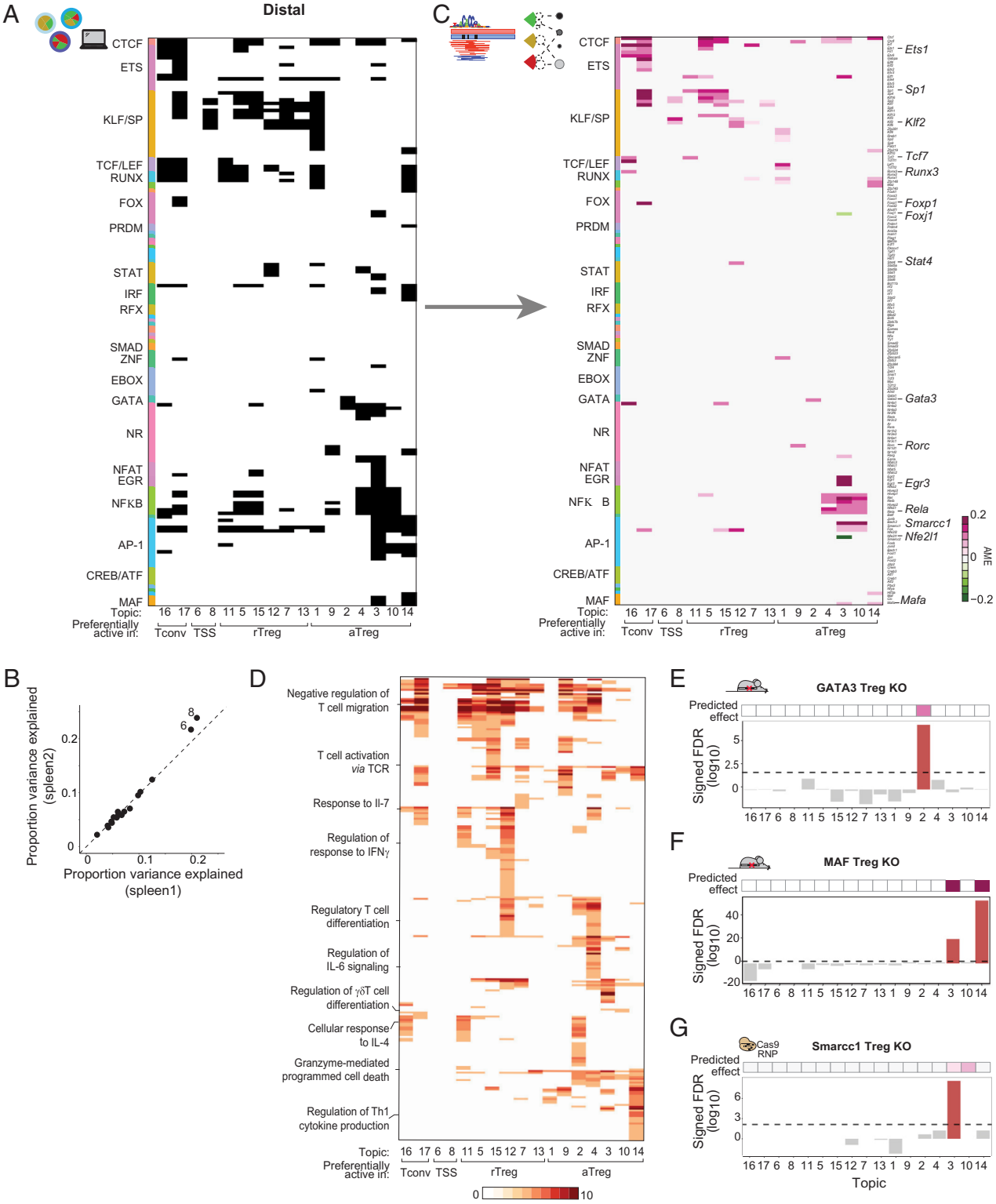

Fig. 2. Topic modeling and high-density genetic variation learn an integrated Treg TF framework. (*A*) Enrichment of TF motifs within distal OCRs from each topic; black indicates enrichment FDR < 1×10^{-10}. Motifs are organized by TF family (row annotation). (*B*) Proportion of variance in accessibility explained by each topic in two independently generated spleen Treg scATAC datasets. (*C*) Topic-specific AME (FDR < 0.10) of motifs in each Treg topic. Heatmap in same order as in *A* 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. (*D*) Gene Ontology gene sets significantly enriched among regulatory regions in each topic [using GREAT (61) analysis]. Heatmap indicates fold change of enrichment relative to background. Full table of enrichments and pathway names is provided in Dataset S5. (*E*) Enrichment [signed $\log_{10}$(FDR), permutation test] in each topic of Gata3-dependent OCRs (GATA-motif containing OCRs decreased in accessibility > twofold in *Foxp3-cre×Gata3*$^{fl/fl}$ Treg-specific Gata3 KO vs. *Foxp3-cre×Gata3*$^{+/+}$ WT scATAC-seq). Panel above indicates predicted topic effect based on topic TF framework in *C*. (*F*) Enrichment [signed $\log_{10}$(FDR), permutation test] in each topic of c-Maf-dependent OCRs [MAF-motif containing OCRs decreased in accessibility > twofold in *Foxp3-cre×Maf*$^{fl/fl}$ Treg-specific c-Maf KO vs. *Foxp3-cre×Maf*$^{+/+}$ WT (62)]. Panel above indicates predicted topic effect based on topic TF framework in *C*. (*G*) 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*$^{IRES\text{-}GFP}$ Tregs electroporated with CRISPR/Cas9 ribonucleoprotein complexes carrying *Smarcc1*-targeting vs. control gRNAs and transferred for 1 wk into Treg-depleted *Foxp3*^{DTR} hosts). Panel above indicates predicted topic effect based on topic TF framework in *C*.

highly reproducible, with excellent concordance between two independent spleen Treg scATAC-seq datasets generated months apart, in terms of variance explained (Fig. 2*B*) and of topic/TF framework accessibility (*SI Appendix*, Fig. S3*E*). They also overlapped well with published TF-binding datasets (ChIP-seq or CUT&RUN; *SI Appendix*, Fig. S3*F*).

We next exploited naturally occurring genetic variation to functionally identify causal relationships among the initial links from topic modeling. Wild-derived Cast/Eij (Cast) mice differ from reference C57BL/6 (B6) mice by approximately 20 million variants (63). B6xCast F1 offspring have been used to causally link sequence variation to chromatin features (34, 64–68). 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 allele-specific Treg scATAC profiles from B6xCast F1 mice (*SI Appendix*, Fig. S5*A*; n = 5,980 Tregs), modifying a published pipeline (69) to assign informative reads to their allele of origin (*SI Appendix*, Fig. S5*A*). To link changes in TF motifs with corresponding shifts in allele-specific chromatin accessibility ["allelic motif effect" (AME)], we adapted previous analytical frameworks (34, 67) to the single-cell, topic modeling context. For each topic, we identified the cells with active accessibility of topic OCRs containing each candidate motif and computed the F1 AME for motifs in each topic only in these relevant cells (*SI Appendix*, Computational Note 2). Because they measured the impact of variation in TF motif fit on accessibility at topic OCRs across the two chromosomes of each F1 cell, these topic-specific AMEs provided a causally determined refinement of the initial TF-topic framework (*SI Appendix*, Fig. S5*B*), yielding an intricate view of regulators of Treg chromatin programs, in breadth and context-specificity (Fig. 2*C* and Dataset S4). The vast majority of AMEs corresponded to positive contributions of TF motifs to chromatin accessibility, very few motifs (e.g., Foxj1 and Nfe2l1) having repressive effects. For ETS, KLF/SP, and TCF/LEF, their genetic effects overlapped with motif enrichments primarily in rTreg- and Tconv-biased topics. AMEs sharpened the scope of AP-1 and NF-κB activity, narrowing their multitopic enrichment to a cluster of aTreg-preferential topics (3, 4, 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.

We used GREAT (61) to connect OCRs to genes and identify functional pathways enriched in the regulatory regions defined by each topic (Fig. 2*D* and Dataset S5). Overall, each topic encompassed a limited set of biological functionalities, with segregation of some functions (e.g., regulation of Th1 or IL6 cytokines), although no particular biological process or molecular function was uniquely associated with any one of the topics. This shuffling between chromatin programs and cell function is not unexpected, as coregulated transcriptional modules routinely encompass several biochemical or cell biological functions, but they imply that these Topics can be used to mark specific Treg subsets or functionalities.

To validate our Treg regulatory framework, we analyzed chromatin of mice with Treg-specific deletion of several exemplar TFs. We generated scATAC-seq profiles of Tregs with conditional Gata3 ablation and control littermates [samples hashtagged in the same run for optimal comparability (70), hereafter "multiplexed"]. Gata3-dependent OCRs (Dataset S2) in these data were enriched (FDR < 0.01) only in Topic 2, the topic predicted to be under Gata3 control (Fig. 2*E*). MAF family-dependent OCRs (Dataset S2) in published ATAC-seq data from cMaf-deficient Tregs (62) were enriched only in Topics 3 and 14 (Fig. 2*F*), consistent with framework predictions. Finally, we tested the model's ability to nominate a previously unrecognized regulator of Treg biology, beyond the "usual suspects". Smarcc1 (BAF155) is a core subunit of all mammalian SWI/SNF chromatin remodeling complexes (71, 72), associated with T cell activation (68, 73) or regulation of *Foxp3* expression (74), but not previously implicated in control of Treg chromatin. Our framework predicted a specific effect of Smarcc1 in aTreg-biased topics 3 and 10, suggesting a role for Smarcc1 in control of aTreg-specific chromatin accessibility. We used CRISPR-Cas9 ribonucleoprotein complexes to engineer Tregs carrying *Smarcc1* deletions, parked the edited cells for 1 wk in vivo before bulk ATAC-seq. Smarcc1-dependent OCRs were enriched for Topic 3 OCRs, as predicted by our framework (Fig. 2*G*). Topic 10 was not represented, potentially reflecting redundancy between Smarcc1 and Smarcc2 paralogs (71, 75). Thus, experimental validations supported predicted TF links, identifying unappreciated regulatory connections.

Tissue-Specific Adaptations Map to the Treg Framework. We then asked whether this integrated framework would capture changes associated with Treg physiology, like tissue-residence. Tregs in nonlymphoid tissues adapt by elaborating specialized regulatory programs (7). The spleen contains precursors of this multistep adaptation (15, 16, 18, 19, 37, 76), and circulating tissue-Tregs, such that a framework derived from total splenic Tregs should contain modules linked to tissue adaptation. Indeed, OCRs associated with tissue-Tregs (15) were enriched in topics 3 and 14 in spleen Tregs (*SI Appendix*, Fig. S6*A*). We generated multiplexed scATAC-seq profiles of Tregs isolated from spleen and colonic lamina propria from the same *Foxp3*$^{\text{IRES-GFP}}$ reporter mice (*SI Appendix*, Fig. S6 *B* and *C*). Consistent with predictions, topics 3 and 14 explained more variance and had increased accessibility within colonic Treg cells (Fig. 3 *A* and *B*). Overlaying differential accessibility onto the framework (using the expanded, enrichment-based version for breadth) showed its granular reorganization and distinguished two groups of motifs (Fig. 3*C*). One group had monomorphic changes in accessibility (e.g., higher accessibility in colon Tregs of bZIP, MAF, and IRF motifs in topics 3, 10, and 14; lower accessibility in colon of ETS and KLF/SP members in rTreg-preferential topics). The second group demonstrated the power of partitioning changes by epigenomic program: TCF/LEF, RUNX, NF-Kb family motifs had variably positive or negative effects depending on their presence in OCRs from different topics (Fig. 3*C*; expanded network in *SI Appendix*, Fig. S6*D*). Some topics were even more specific: even though they were originally learned in spleen Tregs, topics 9 and 10 demarcated RORγ+ and Helios+ Tregs in the colon (Fig. 3*D* and Dataset S2). Thus, the framework captured tissue-Treg-specific specializations and identified which factors' activities were molded by cell state–specific inputs.

FoxP3 Control of Treg Identity. Our framework provided an opportunity to systematically interrogate FoxP3's mechanism of action, which remains poorly understood. Unlike traditional lineage-defining master regulators, FoxP3 does not act as a pioneer factor (44, 45) (consistent with this notion, FoxP3 itself does not appear in the framework) and is neither fully necessary nor sufficient to establish Treg identity. Treg-like cells ("Treg wannabes") can develop in its absence, and Treg-specific gene expression signatures consist of both FoxP3-independent and -dependent components (28–33). While FoxP3 interacts with a large array of other regulators (77–80), there is unsettled debate as to whether FoxP3 acts directly as an activator (29, 81–83), a repressor (77, 78, 84, 85), or both, depending on its interacting cofactors (80) or indirectly by tuning the expression of other

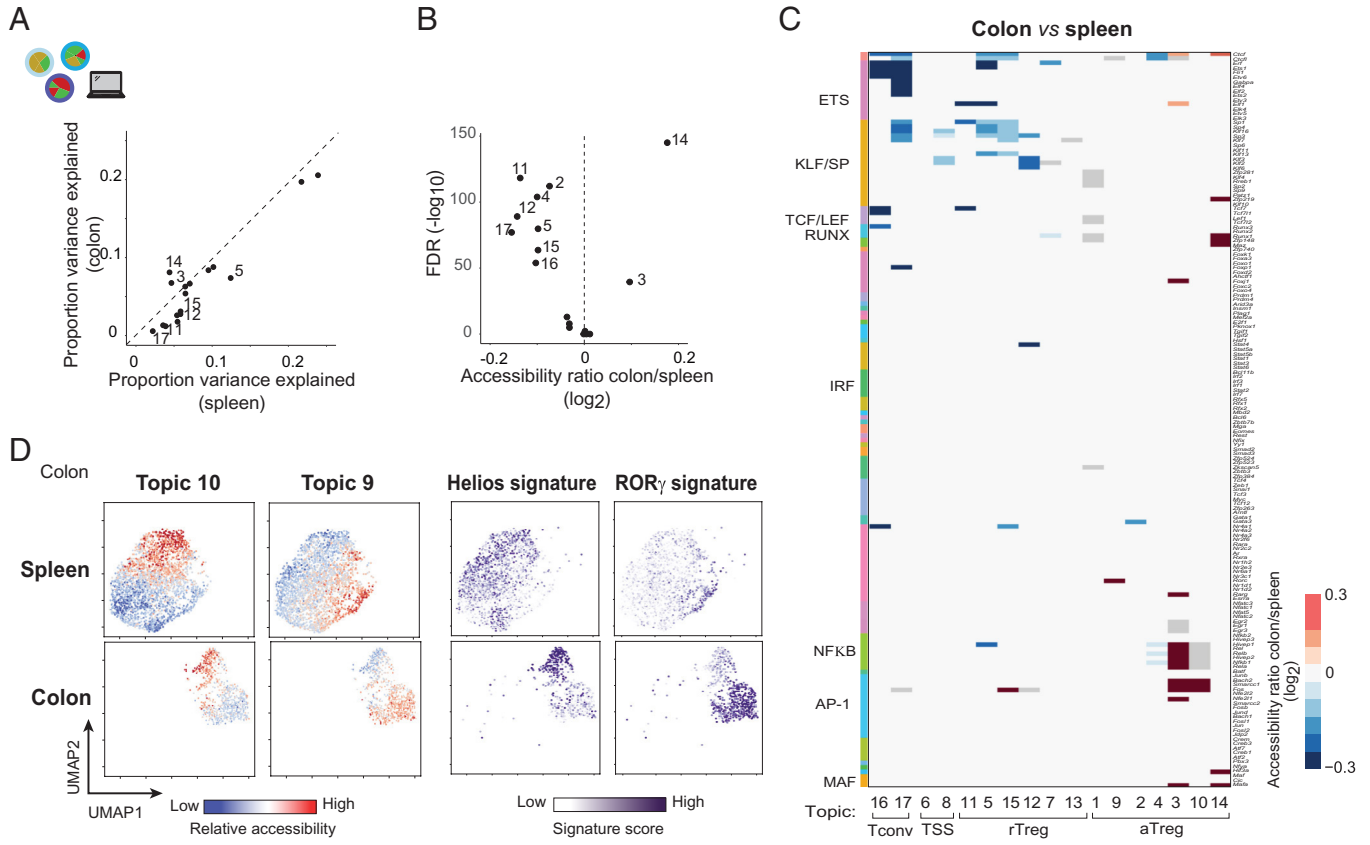

Fig. 3. Adaptations to tissues and cytokines mapped to the Treg TF framework. (*A*) Proportion of variance in accessibility explained by each topic in colon or spleen Treg scATAC data (from the same *Foxp3*$^{IRES\text{-}GFP}$ reporter mouse). (*B*) Differential accessibility per topic between aggregated colon and spleen Treg scATAC profiles. (*C*) Differential accessibility per motif in each topic (distal OCRs, FDR < 0.05) between colon and spleen Treg overlaid onto motif to topic connections from Fig. 2*C* (enrichment and F1 AME). Gray indicates significant topic-motif enrichment from Fig. 2*A* but no significant change in accessibility across colon vs. spleen comparisons. (*D*) 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.

TFs (34). 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 (Foxp3fs327-GFP/Foxp3-Thy1.1 mice, "KO" in Fig. 4*A*) (83). 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. 4*A*) are similarly constructed, with a functional FoxP3 encoded upstream of the GFP reporter.

We sorted GFP+ Treg and GFP- Tconv from both WT and KO heterozygotes for multiplexed scATAC-seq (Fig. 4*A* and *SI Appendix*, Fig. S7*A*). While Tconv from WT and KO mice mostly comingled on the UMAP visualization, as expected from the unperturbed environments, FoxP3-deficient Treg-like cells were shifted wholesale from the region occupied by WT Tregs (Fig. 4*B* and *SI Appendix*, Fig. S7 *B* and *C*). FoxP3-deficient Treg-like cells had diminished accessibility of aTreg-specific OCRs, most KO Tregs being in a resting-like chromatin state (Fig. 4*B*). This inability of FoxP3-deficient Tregs to progress to activated states was confirmed by flow cytometry (Fig. 4*C*). When overlaid onto the TF regulatory framework (Fig. 4*D*), FoxP3's positive and negative effects split cleanly into different chromatin programs, each driven by distinct TF ensembles (Fig. 4*D*). 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. 4*D* and *SI Appendix*, Fig. S7*D*). 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. Importantly, the framework systematically localized the specific genomic programs and partner cofactors that contribute to positive or negative FoxP3 function. This decomposition between FoxP3's positive and negative effects could be highlighted at exemplar loci (Fig. 4*E*).

If FoxP3 does not entirely define Treg identity (29–33, 87), what are the Treg wannabes that develop in its absence? We compared, in the Treg framework, KO Treg-like cells and Tconv. Strikingly, these FoxP3-independent effects (Fig. 4*F*) seemed a carbon-copy of the FoxP3-dependent effects observed when comparing WT and KO Tregs (Fig. 4*D*), as also evidenced by comparing the ratios of topic accessibility (*SI Appendix*, Fig. S7 *E* and *F*). Relative to Tconv cells, KO Treg-like cells induced Topic 4 [controlled by NF-κB, known to be important to Treg identity (47, 88) and repressed Tconv-preferential Topics 16 and 17. We attempted to relate these patterns to TF status in Treg progenitors in the thymus, before FoxP3 expression, by analyzing previously published Treg differentiation datasets (86). Interestingly, members of these same TF families displayed related expression changes (e.g., upregulation of NF-κB, AP-1 families, downregulation of ETS, KLF, TCF/LEF families—(Fig. 4*G*) in early Treg progenitors, stages in which pre-Treg cells are thought to receive strong

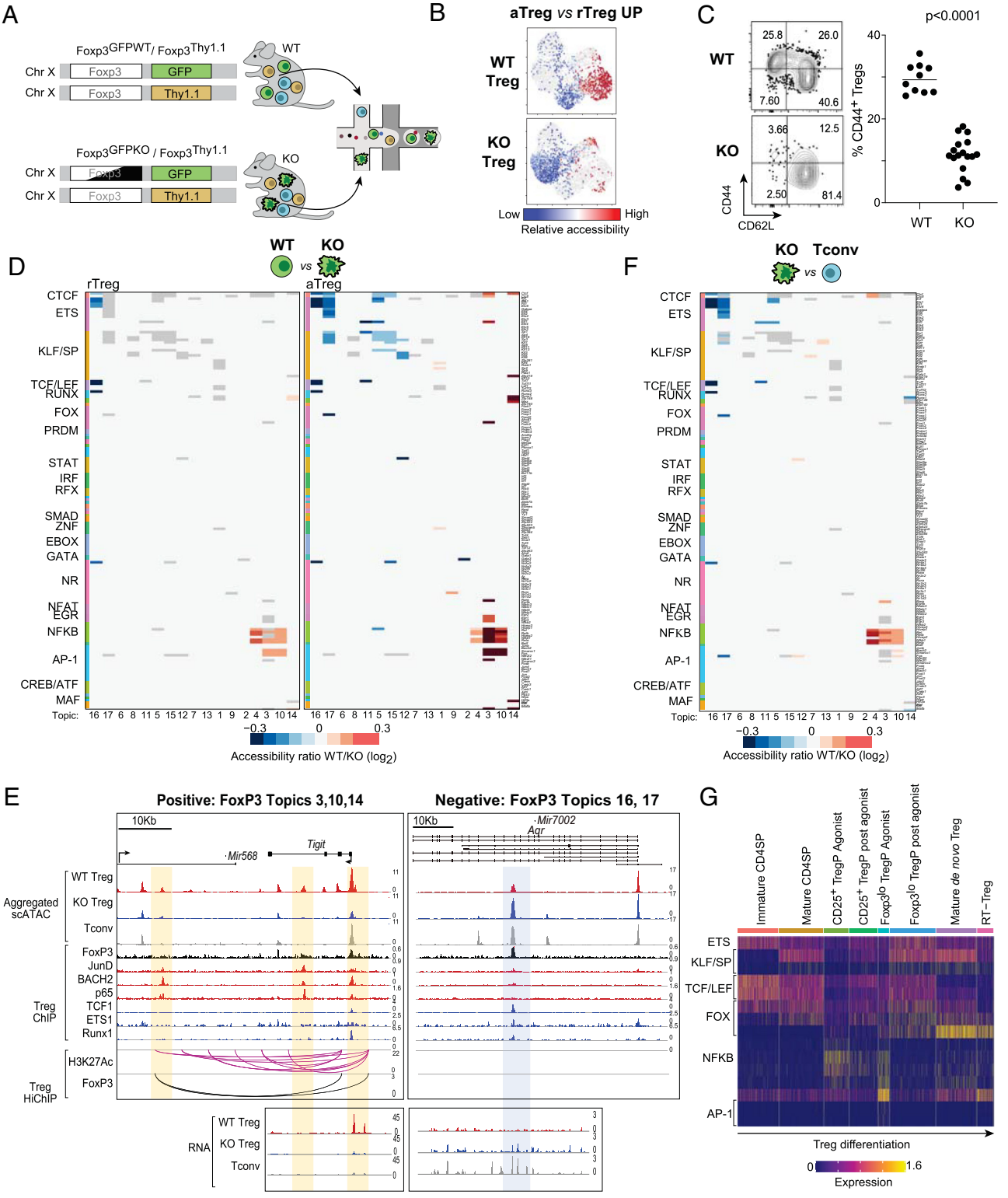

Fig. 4. FoxP3 effects and Treg identity in the Treg regulatory framework. (*A*) Experimental Scheme. FoxP3-deficient (KO) and -sufficient (WT) GFP+ Tregs were sorted along with GFP- Tconv from Foxp3^{fs327}-GFP/Foxp3-Thy1.1 or Foxp3^{wt}-GFP/Foxp3-Thy1.1 heterozygous female mice for multiplexed scATAC-seq. (*B*) 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 of Treg and Tconv from FoxP3 WT or KO heterozygous female mice. (*C*) Proportion of rTreg and aTreg populations in FoxP3 WT or KO populations by CD44 and CD62L flow cytometry with quantification across biological replicates. (*D*) 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 Fig. 2*C*. Gray indicates significant motif to topic connection from Fig. 2*C* but no significant change in accessibility across FoxP3 comparisons. (*E*) Exemplar loci demonstrating FoxP3 positive (*Left*) and negative (*Right*) effects on accessibility and dependence on distinct cofactors as identified by the topic-motif Treg framework. (*Top* tracks) Aggregated scATAC-seq reads from WT Treg (red), KO Treg (blue), or Tconv (gray) cells from heterozygous female mice. (*Middle*) Treg ChIP-Seq or CUT&RUN tracks of indicated TFs identified as contributing to FoxP3 positive (red) or negative (blue) action. (*Bottom*) Treg H3K27Ac or FoxP3 HiChIP. RNA expression from aggregated scRNA-seq at these loci is displayed below. Highlights indicate OCRs belonging to Topics 3, 10, 14 (orange, left, positive FoxP3 effect) or Topics 16, 17 (blue, right, negative FoxP3 effect). (*F*) Differential accessibility per motif in each topic (distal OCRs, FDR < 0.05) between KO rTreg and Tconv cells for motif to topic connections from Fig. 2*C*. Gray indicates significant motif to topic connection from Fig. 2*C* but no significant change in accessibility across differential comparisons. (*G*) Normalized expression of TFs in scRNA-seq of early thymic Treg differentiation (86). TregP: Treg Progenitor; RT-Treg: recirculating/long-term resident Treg.

TCR agonist signals, suggesting that the Treg wannabe program has an analog in early Treg progenitors). 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 maturation to aTreg, where FoxP3 protects Treg identity, suppresses Tconv programs and enables the induction of aTreg-specific chromatin programs, which underlie 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. 5*A*), 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. 5*B*), with a quasiabsence of Helios+ Tregs (Fig. 5 *B* and *C*), consistent with a drop in *Ikzf2* accessibility in the genomic data (*SI Appendix*, Fig. S8 *A* and *B*). These results confirm a report from van der Veeken et al. (89).

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 wk), the fraction of total GFP+ Tregs decreased strongly in the KO colon but not in WT (Fig. 5*D*) where the drop in RORγ+ Tregs was balanced by an increase of the Helios+ pool (Fig. 5*E*). The persisting Treg-like cells in KO mice remained RORγ$^+$, and no Helios+ Tregs emerged (Fig. 5*E*). 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. 5*F*), as previously observed (89). 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. 5*F*). 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 spleens and colons of WT and KO heterozygous females. In accordance with the scATACseq, KO Tregs were shifted in the UMAP relative to WT Tregs (Fig. 5*G* and *SI Appendix*, Fig. S8 *C* and *D*) and had high *Il17a* expression in the colon (Fig. 5*H*). 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. 5 *G* and *H* and *SI Appendix*, Fig. S8*D*). They retained the ability to express IL10 (Fig. 5*J*), maintained Treg signature scores closer to WT Treg than to Tconv (Fig. 5*I*), and matched WT Tregs in their accessibility pattern at the *Foxp3* locus (*SI Appendix*, Fig. S8*E*). Thus, KO Tregs did not become Th17 cells but maintained several facets of Treg identity while simply derepressing cytokine production. Notably, another population of *Gata3+* KO Treg-like cells down-regulated *Ikzf2* expression and expressed type 2 cytokine transcripts (*Il4, Il5, Il13*; *SI Appendix*, Fig. S8 *C* and *D*), indicating that the partial cytokine derepression matched the *Rorc+* or *Gata3+* tone of the cells. In summary, while not required for RORγ+ Treg differentiation, FoxP3 was required to repress Teff cytokines induced in the gut environment.

Why did colon Helios+ and RORγ+ Tregs show such divergent susceptibility to the loss of FoxP3? To identify programmatic traits that might explain this differential susceptibility, we compared chromatin programs of RORγ+ vs. Helios+ splenic aTreg cells (Fig. 5*K*). RORγ+ Tregs had higher accessibility of the RORγ-dependent topic 9 (as expected), but lower accessibility of NF-Kb and AP-1 motifs within topics 4, 3, and 10, which correspond precisely to those positively activated by FoxP3 (Fig. 4 *D* and *E*). Thus, the lower dependence on FoxP3 activity seems foretold by the topic activity of RORγ+ Treg chromatin. We reanalyzed in this framework scAT-ACseq from colonic Tregs with conditional knockouts in either *Ikzf2* or *Rorc* (39) (Fig. 5*L*). Helios was necessary for the activity of these NF-Kb/AP-1 dependent OCRs, while the deletion of RORγ had no effect. These results indicate that these FoxP3-dependent OCRs, enriched in NF-Kb and AP-1 motifs, further require Helios for their full activity (Fig. 5*L*), suggesting that Helios and FoxP3 may collaborate in a feed-forward loop that reinforces TCR-dependent aTreg programs. Thus, our Treg framework provides a basis for the differential dependence on FoxP3 across Treg subpopulations.

Discussion

This work thus provides an unbiased, genome-wide map connecting TFs to the Treg identity and to the full diversity of Treg epigenomic programs. Combinatorial control of gene expression is a classic notion (90), but the dominant paradigm in immunological gene regulation is often that phenotypic specialization is mediated by single master TFs (3, 4, 20, 91). Topic modeling quantitatively decomposed the combinatorial architecture, and natural genetic variation at single-cell resolution critically identified causal connections between TF-binding motifs and target programs. This framework systematically and precisely operationalized TF context-specificity, bringing out divergent activities of a given TF at different OCRs or topics. It also enabled finding of less characterized regulators, e.g., the validation of Smarcc1 as a controller of aTreg-specific chromatin, and several other TFs not previously implicated in Treg biology appear in our framework.

The Treg TF framework was portable across tissues and stimulation conditions. Though it was developed from splenic Tregs, the network's capture of tissue-specific effects supported the notion that tissue-Treg programs represent amplifications of patterns already present in Tregs from lymphoid tissues (15, 16, 18, 19, 37, 76). Decomposition of effects by topics importantly separated tissue-responsive TFs that acted monomorphically (e.g., bZIP, IRF, ETS, KLF/SP) from those whose activity was malleable depending on the epigenomic program (e.g., RUNX, TCF/LEF, NF-Kb). The latter may represent differences in TFs more dynamically responsive to environmental signals. The framework's reproducibility means that it could become a portable and reusable scaffold for interpretation of Treg chromatin studies elsewhere, moving beyond unconnected single studies. We have implemented a web application (https://cbdm.connect.hms.harvard.edu/Topic_Plotting/) that allows any investigator to upload single-cell chromatin data and receive its decomposition in terms of these topics.

While previous studies have debated whether FoxP3 acts as an activator or repressor, the present results resolve that it acts as both, consistent with the notion that FoxP3 is highly dependent on its cofactors (80). FoxP3's repressive and activating functions affected different chromatin topics with distinct motif contributions. The OCR membership of each of these topics provides genomic definition to the locations where FoxP3 might assemble into different molecular complexes spatially segregated in the nucleus with distinct cofactor composition and opposing functions (80). The results also offer an integrated model for the relationship between FoxP3 and Treg identity. FoxP3 amplified a program of Treg identity present in Treg wannabes, with a striking superimposition between the pre-existing program and what FoxP3 amplifies and which appears early in thymic Treg differentiation upon agonist signaling. **How** this concordance is achieved molecularly is an

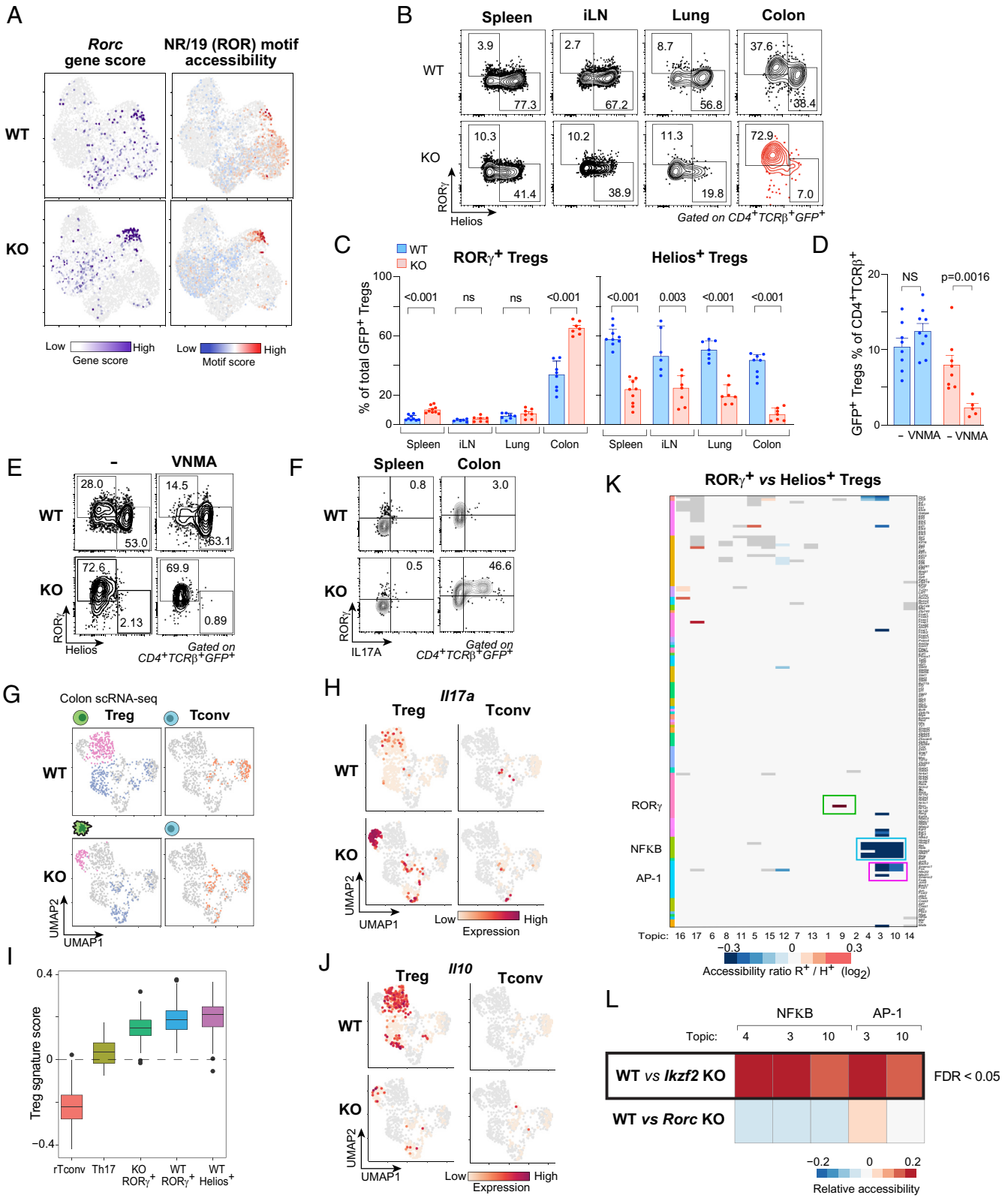

Fig. 5. FoxP3-independent RORγ+ Treg-like cells. (*A*) *Rorc* gene scores and NR/19 (contains RORγ) motif accessibility (chromVAR score) in WT and KO Treg populations in scATAC-seq data from Fig. 4. (*B*) Proportion of RORγ+ vs. Helios+ Tregs among WT or KO Tregs in heterozygous female mice across organs. (*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*. (*K*) Differential accessibility per motif in each topic (distal OCRs, FDR < 0.05) between WT RORγ+ and Helios+ aTreg cells for motif to topic connections from Fig. 2*C*. Gray indicates significant motif to topic connection from Fig. 2*C* but no significant change in accessibility across FoxP3 comparisons. (*L*) Differential accessibility per motif in each topic (aggregating motif family members contributing to each topic from Fig. 2*C*) in WT vs. *Ikzf2* conditional KO (*Top*) or WT vs. *Rorc* conditional KO colon Treg scATAC data (39). Comparisons in outlined boxes are significant at FDR < 0.05.

open question. One interpretation is that it replaces a similar Forkhead factor that would anchor the Tconv-wannabe distinction or that FoxP3 is uniquely attuned to cooperate with the TF ensemble that supports the Treg-like wannabe status, for which these data provide candidates.

The finding 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 (89). We found FoxP3-deficient RORγ+ Treg-like cells to be microbe-dependent, as are normal FoxP3-sufficient counterparts. Chromatin opening at the *Foxp3* locus resembled that of normal Tregs. They produced IL17A but only in the colon, and without diversion to a Th17-like transcriptome, suggesting that cytokine derepression 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 (92–94). However, extrinsic inputs cannot be the sole determinants of this program, as an intrinsic bias toward an increased RORγ/Helios ratio remained even in the context of antibiotic treatment. TF framework analysis suggested an origin for this bias in differential reliance on NF-kB and AP-1 dependent topics and a collaborative feed-forward loop between Helios and FoxP3. While RORγ+ Treg-like cells provided an alternative fate for FoxP3-deficient cells, they were unable to replace all of FoxP3's functions, such as repression of cytokine genes upon stimulation.

Overall, Treg cell identity and diversity are not defined by monomorphic TFs, into TF-controlled sublineages, but by a framework of programs, across which TFs have context-dependent activity. This work provides a first step toward integrating the global and granular architecture of such a framework.

Materials and Methods

Mice. *Foxp3*$^{\text{IRES-GFP}}$ mice (95) and *Foxp3*$^{fs327\text{-GFP}}$ x *Foxp3*$^{tm10.1(\text{Casp9, -Thy1})\text{Ayr}}$ (or Foxp3$^{\text{Thy1.1}}$) (83) heterozygous female mice were bred on the C57Bl/6J background. To generate B6/Cast F1 mice, Cast/EiJ males (Jackson Labs, strain #000928) were crossed with C57BL/6J females (Jackson Labs, strain #000664).

Processing and Sequencing of T Lymphocytes. Immunocytes from lymphoid organs and colon were prepared as previously described (96, 97). For scATAC-seq or scRNA-seq, Tregs (Cd19-, Cd11c-, Cd8-, Cd4+, TCRb+, GFP+, or CD25hi) and Tconv (Cd19-, Cd11c-, Cd8-, Cd4+, TCRb+, GFP-, or CD25lo) cells were processed per the 10× Chromium Next GEM Single-Cell ATAC manual, with hashtagging per condition using a modification of the ASAP-seq strategy (70), or the 10× Genomics Chromium Single-Cell 5′ v2 with Feature Barcoding kit. Bulk ATAC-seq libraries were prepared using the ImmGen ATAC-seq protocol (44).

Data Analysis. scATAC-seq data analysis was performed using Signac v1.4 (98), using a common set of open chromatin regions throughout the study. For topic modeling (see *SI Appendix*, Computational Note 1), we used an ensemble modeling approach building on cisTopic (59) to determine a set of reproducible topics and representative OCRs. For B6/Cast F1 data, we adapted the diploid pseudogenome alignment strategy implemented in the lapels/suspenders pipeline (69) for allele-specific mapping. For AME calculations, we built on recent work (34, 67), focusing comparisons to only relevant OCRs (with motif in corresponding topic) within relevant cells (with accessibility at these sites) to identify the causal contributions of motifs to each topic (see *SI Appendix*, Computational Note 2). Relative accessibility of motifs and OCR sets was calculated using chromVAR (99). Differential accessibility on the motif-topic framework was calculated using the $\log_2$ Fold Change in average accessibility of topic OCRs containing the motif of interest between conditions, with significance by *t* test with Benjamini–Hochberg correction.

Data, Materials, and Software Availability. Raw and processed data files have been deposited at the Gene Expression Omnibus [GEO: GSE216910 (100) and GSM5712663 (101)].

ACKNOWLEDGMENTS. 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 Bauer Core Facility at Harvard University for scATAC-seq; the Harvard Medical School 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; and D. Mallah for assistance with web tool development. This work was funded by grants from the NIH to D.M. and C.B. (AI150686, AI165697, AI125603). K.C. was supported by National Institute of General Medical Sciences (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.

1. T. I. Lee, R. A. Young, Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237–1251 (2013).
2. T. R. Sorrells, A. D. Johnson, Making sense of transcription networks. *Cell* **161**, 714–723 (2015).
3. S. Heinz, C. E. Romanoski, C. Benner, C. K. Glass, The selection and function of cell type-specific enhancers. *Nat. Rev. Mol. Cell Biol.* **16**, 144–154 (2015).
4. S. Z. Josefowicz, L. F. Lu, A. Y. Rudensky, Regulatory T cells: Mechanisms of differentiation and function. *Annu. Rev. Immunol.* **30**, 531–564 (2012).
5. P. Georgiev, L. M. Charbonnier, T. A. Chatila, Regulatory T cells: The many faces of Foxp3. *J. Clin. Immunol.* **39**, 623–640 (2019).
6. C. Benoist, D. Mathis, Treg cells, life history, and diversity. *Cold Spring Harb. Perspect. Biol.* **4**, a007021 (2012).
7. A. R. Muñoz-Rojas, D. Mathis, Tissue regulatory T cells: Regulatory chameleons. *Nat. Rev. Immunol.* **21**, 597–611 (2021).
8. D. J. Campbell, M. A. Koch, Phenotypical and functional specialization of FOXP3$^+$ regulatory T cells. *Nat. Rev. Immunol.* **11**, 119–130 (2011).
9. M. A. Koch *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. Y. Zheng *et al.*, Regulatory T-cell suppressor program co-opts transcription factor IRF4 to control T(H)2 responses. *Nature* **458**, 351–356 (2009).
11. D. Burzyn *et al.*, A special population of regulatory T cells potentiates muscle repair. *Cell* **155**, 1282–1295 (2013).
12. S. K. Lathrop *et al.*, Peripheral education of the immune system by colonic commensal microbiota. *Nature* **478**, 250–254 (2011).
13. M. Feuerer *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).
14. M. Feuerer *et al.*, Genomic definition of multiple ex vivo regulatory T cell subphenotypes. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5919–5924 (2010).
15. J. R. Dispirito *et al.*, Molecular diversification of regulatory T cells in nonlymphoid tissues. *Sci. Immunol.* **3**, eaat5861 (2018).
16. C. Li *et al.*, TCR transgenic mice reveal stepwise, multi-site acquisition of the distinctive fat-Treg phenotype. *Cell* **174**, 285–299 (2018).
17. D. Zemmour *et al.*, Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nat. Immunol.* **19**, 291–301 (2018).
18. M. Delacher *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).
19. M. Delacher *et al.*, Single-cell chromatin accessibility landscape identifies tissue repair program in human regulatory T cells. *Immunity* **54**, 702–720 (2021).
20. J. L. Trujillo-Ochoa, M. Kazemian, B. Afzali, The role of transcription factors in shaping regulatory T cell identity. *Nat. Rev. Immunol.* **23**, 842–856 (2023).
21. Y. Chung *et al.*, Follicular regulatory T cells expressing Foxp3 and Bcl-6 suppress germinal center reactions. *Nat. Med.* **17**, 983–988 (2011).
22. M. A. Linterman *et al.*, Foxp3$^+$ follicular regulatory T cells control the germinal center response. *Nat. Med.* **17**, 975–982 (2011).
23. Y. Wang, M. A. Su, Y. Y. Wan, An essential role of the transcription factor GATA-3 for the function of regulatory T cells. *Immunity* **35**, 337–348 (2011).
24. E. Sefik *et al.*, Individual intestinal symbionts induce a distinct population of RORγ$^+$ regulatory T cells. *Science* **349**, 993–997 (2015).
25. C. Ohnmacht *et al.*, The microbiota regulates type 2 immunity through RORγ$^+$ T cells. *Science* **349**, 989–993 (2015).
26. W. Fu *et al.*, A multiply redundant genetic switch "locks in" the transcriptional signature of regulatory T cells. *Nat. Immunol.* **13**, 972–980 (2012).
27. K. Schumann *et al.*, Functional CRISPR dissection of gene networks controlling human regulatory T cell identity. *Nat. Immunol.* **21**, 1456–1466 (2020).
28. M. A. Gavin *et al.*, Foxp3-dependent programme of regulatory T-cell differentiation. *Nature* **445**, 771–775 (2007).
29. J. A. Hill *et al.*, Foxp3 transcription-factor-dependent and -independent regulation of the regulatory T cell transcriptional signature. *Immunity* **27**, 786–800 (2007).
30. W. Lin *et al.*, Regulatory T cell development in the absence of functional Foxp3. *Nat. Immunol.* **8**, 359–368 (2007).
31. N. Ohkura *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).
32. L. M. Charbonnier *et al.*, Functional reprogramming of regulatory T cells in the absence of Foxp3. *Nat. Immunol.* **20**, 1208–1219 (2019).
33. D. Zemmour *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).

34. J. van der Veeken *et al.*, The transcription factor Foxp3 shapes regulatory T cell identity by tuning the activity of trans-acting intermediaries. *Immunity* **53**, 971–984 (2020).
35. J. Vierstra *et al.*, Global reference mapping of human transcription factor footprints. *Nature* **583**, 729–736 (2020).
36. J. M. Granja *et al.*, ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.* **53**, 403–411 (2021).
37. C. Li *et al.*, PPARγ marks splenic precursors of multiple nonlymphoid-tissue Treg compartments. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2025197118 (2021).
38. M. Xu *et al.*, c-MAF-dependent regulatory T cells mediate immunological tolerance to a gut pathobiont. *Nature* **554**, 373–377 (2018).
39. D. Ramanan *et al.*, Homeostatic, repertoire and transcriptional relationships between colon T regulatory cell subsets. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2311566120 (2023).
40. N. Hayatsu *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).
41. N. D. Heintzman *et al.*, Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
42. Y. Shen *et al.*, A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).
43. M. R. Corces *et al.*, Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat. Genet.* **48**, 1193–1203 (2016).
44. H. Yoshida *et al.*, The cis-regulatory atlas of the mouse immune system. *Cell* **176**, 897–912 (2019).
45. R. M. Samstein *et al.*, Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell* **151**, 153–166 (2012).
46. Y. Kitagawa *et al.*, Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment. *Nat. Immunol.* **18**, 173–183 (2017).
47. H. Oh *et al.*, An NF-κB transcription-factor-dependent lineage-specific transcriptional program promotes regulatory T cell identity and function. *Immunity* **47**, 450–465 (2017).
48. F. M. Grant *et al.*, BACH2 drives quiescence and maintenance of resting Treg cells to promote homeostasis and cancer immunosuppression. *J. Exp. Med.* **217**, e20190711 (2020).
49. T. Sidwell *et al.*, Attenuation of TCR-induced transcription by Bach2 controls regulatory T cell differentiation and homeostasis. *Nat. Commun.* **11**, 252 (2020).
50. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
51. M. Osterwalder *et al.*, Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239–243 (2018).
52. M. Gasperini *et al.*, A genome-wide framework for mapping gene regulation via cellular genetic screens. *Cell* **176**, 377–390 (2019).
53. F. Spitz, E. E. Furlong, Transcription factors: From enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012).
54. M. Gasperini, J. M. Tome, J. Shendure, Towards a comprehensive catalogue of validated and target-linked human enhancers. *Nat. Rev. Genet.* **21**, 292–310 (2020).
55. A. Panigrahi, B. W. O'Malley, Mechanisms of enhancer action: The known and the unknown. *Genome Biol.* **22**, 108 (2021).
56. V. K. Kartha *et al.*, Functional inference of gene regulation using single-cell multi-omics. *Cell Genom.* **2**, 100166 (2022).
57. K. Wang, W. Fu, Transcriptional regulation of Treg homeostasis and functional specification. *Cell Mol. Life Sci.* **77**, 4269–4287 (2020).
58. D. M. Blei, A. Y. Ng, M. I. Jordan, Latent Dirichlet allocation. *J. Mach. Learn. Res.* **3**, 993–1022 (2003).
59. C. Bravo Gonzalez-Blas *et al.*, cisTopic: cis-regulatory topic modeling on single-cell ATAC-seq data. *Nat. Methods* **16**, 397–400 (2019).
60. P. Bielecki *et al.*, Skin-resident innate lymphoid cells converge on a pathogenic effector state. *Nature* **592**, 128–132 (2021).
61. C. Y. McLean *et al.*, GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
62. C. Neumann *et al.*, c-Maf-dependent Treg cell control of intestinal TH17 cells and IgA establishes host-microbiota homeostasis. *Nat. Immunol.* **20**, 471–481 (2019).
63. T. M. Keane *et al.*, Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289–294 (2011).
64. S. Heinz *et al.*, Effect of natural genetic variation on enhancer selection and function. *Nature* **503**, 487–492 (2013).
65. T. Vierbuchen *et al.*, AP-1 transcription factors and the BAF complex mediate signal-dependent enhancer selection. *Mol. Cell* **68**, 1067–1082 (2017).
66. V. M. Link *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).
67. J. van der Veeken *et al.*, Natural genetic variation reveals key features of epigenetic and transcriptional memory in virus-specific CD8 T cells. *Immunity* **50**, 1202–1217 (2019).
68. Y. Zhong *et al.*, Hierarchical regulation of the resting and activated T cell epigenome by major transcription factor families. *Nat. Immunol.* **23**, 122–134 (2022).
69. S. Huang *et al.*, A novel multi-alignment pipeline for high-throughput sequencing data. *Database (Oxford)* **2014**, bau057 (2014).
70. E. P. Mimitou *et al.*, Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat. Biotechnol.* **39**, 1246–1258 (2021).
71. N. Mashtalir *et al.*, Modular organization and assembly of SWI/SNF family chromatin remodeling complexes. *Cell* **175**, 1272–1288 (2018).
72. R. C. Centore *et al.*, Mammalian SWI/SNF chromatin remodeling complexes: Emerging mechanisms and therapeutic strategies. *Trends Genet.* **36**, 936–950 (2020).
73. S. M. Jeong *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).
74. C. S. Loo *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).
75. J. E. Otto *et al.*, Structural and functional properties of mSWI/SNF chromatin remodeling complexes revealed through single-cell perturbation screens. *Mol. Cell* **83**, 1350–1367 (2023).
76. R. J. Miragaia *et al.*, Single-cell transcriptomics of regulatory T cells reveals trajectories of tissue adaptation. *Immunity* **50**, 493–504 (2019).
77. Y. Wu *et al.*, FOXP3 controls regulatory T cell function through cooperation with NFAT. *Cell* **126**, 375–387 (2006).
78. M. Ono *et al.*, Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature* **446**, 685–689 (2007).
79. D. Rudra *et al.*, Transcription factor Foxp3 and its protein partners form a complex regulatory network. *Nat. Immunol.* **13**, 1010–1019 (2012).
80. H. K. Kwon, H. M. Chen, D. Mathis, C. Benoist, Different molecular complexes that mediate transcriptional induction and repression by FoxP3. *Nat. Immunol.* **18**, 1238–1248 (2017).
81. C. Chen *et al.*, Transcriptional regulation by Foxp3 is associated with direct promoter occupancy and modulation of histone acetylation. *J. Biol. Chem.* **281**, 36828–36834 (2006).
82. B. Li *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).
83. R. N. Ramirez *et al.*, FoxP3 associates with enhancer-promoter loops to regulate Treg-specific gene expression. *Sci. Immunol.* **7**, eabj9836 (2022).
84. L. A. Schubert *et al.*, Scurfin (FOXP3) acts as a repressor of transcription and regulates T cell activation. *J. Biol. Chem.* **276**, 37672–37679 (2001).
85. A. Arvey *et al.*, Inflammation-induced repression of chromatin bound by the transcription factor Foxp3 in regulatory T cells. *Nat. Immunol.* **15**, 580–587 (2014).
86. D. L. Owen, R. S. La Rue, S. A. Munro, M. A. Farrar, Tracking regulatory T cell development in the thymus using single-cell RNA sequencing/TCR sequencing. *J. Immunol.* **209**, 1300–1313 (2022).
87. N. Sugimoto *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).
88. D. L. Owen *et al.*, Thymic regulatory T cells arise via two distinct developmental programs. *Nat. Immunol.* **20**, 195–205 (2019).
89. J. van der Veeken *et al.*, Genetic tracing reveals transcription factor Foxp3-dependent and Foxp3-independent functionality of peripherally induced Treg cells. *Immunity* **55**, 1173–1184 (2022).
90. P. J. Mitchell, R. Tjian, Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371–378 (1989).
91. W. Sungnak, C. Wang, V. K. Kuchroo, Multilayer regulation of CD4 T cell subset differentiation in the era of single cell genomics. *Adv. Immunol.* **141**, 1–31 (2019).
92. S. Keerthivasan *et al.*, β-Catenin promotes colitis and colon cancer through imprinting of proinflammatory properties in T cells. *Sci. Transl. Med.* **6**, 225ra28 (2014).
93. A. Osman *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).
94. J. Quandt *et al.*, Wnt-β-catenin activation epigenetically reprograms Treg cells in inflammatory bowel disease and dysplastic progression. *Nat. Immunol.* **22**, 471–484 (2021).
95. E. Bettelli *et al.*, Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. *Nature* **441**, 235–238 (2006).
96. J. Leon *et al.*, Mutations from patients with IPEX ported to mice reveal different patterns of FoxP3 and Treg dysfunction. *Cell Rep.* **42**, 113018 (2023).
97. D. Ramanan *et al.*, An immunologic mode of multigenerational transmission governs a gut Treg setpoint. *Cell* **181**, 1276–1290 (2020).
98. T. Stuart *et al.*, Single-cell chromatin state analysis with Signac. *Nat. Methods* **18**, 1333–1341 (2021).
99. A. N. Schep, B. Wu, J. D. Buenrostro, W. J. Greenleaf, chromVAR: Inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017).
100. K. Chowdhary, D. Mathis, C. Benoist, Single-cell epigenomics and high-density genetic variation resolve the architecture of the Treg regulatory network [scATAC-seq]. Gene Expression Omnibus. https://www.ncbi.xyz/geo/query/acc.cgi?acc=GSE216910. Deposited 31 October 2022.
101. scATAC.Treg.Tconv. Gene Expression Omnibus. https://www.ncbi.xyz/geo/query/acc.cgi?acc=GSM5712663. Deposited 2 December 2021.