PNAS Nexus, 2026, **5**, pgag386
https://doi.org/10.1093/pnasnexus/pgag386
Published: 10 November 2026
Research report

Human naive regulatory CD4 T-cell identity: Transcriptional and chromatin determinants

Alexandre A. S. F. Raposo[1,2], Pedro Rosmaninho[1,2], Beatriz Moleirinho[1,2], Afonso R. M. Almeida[1,2], and Ana E. Sousa[1,2]*

[1]GIMM—Gulbenkian Institute for Molecular Medicine, 1649-028 Lisbon, Portugal
[2]Faculdade de Medicina, Universidade de Lisboa, 1649-028 Lisbon, Portugal
*To whom correspondence should be addressed: Emails: ana.e.sousa@gimm.pt, asousa@medicina.ulisboa.pt
Edited By Hidde Ploegh

Abstract

Human naive regulatory CD4 T cells (Tregs) are a thymus-derived population essential for peripheral tolerance and tissue homeostasis, yet the mechanisms that preserve their steady-state identity remain poorly understood. Naive Tregs co-express two superimposed programs: Treg lineage and naive T-cell identity. We resolved this entanglement into orthogonal programs by confronting transcription and chromatin accessibility profiles across thymic, naive, and memory Treg and conventional CD4 T-cell populations. The core Treg program, comprising 293 genes common to all three compartments, revealed a conserved regulatory architecture with AP1/ETS members organizing gene-regulatory modules. The lineage-independent naive program identified 175 genes peaking specifically in naive cells, with overrepresentation analysis supporting translational preparedness as a naive hallmark alongside cell trafficking, and POU2F1/OCT1 emerging as key regulator. KLF/SP factors provided quantitative fine-tuning to both programs. These findings establish a two-program model of human naive Treg identity with implications for immune senescence and Treg-based cell therapies.

Keywords human regulatory T cells, naive CD4 T cells, naive regulatory CD4 T cells, gene expression regulation, RNA-seq/ATAC-seq

Significance Statement

Regulatory T cells suppress excessive or misdirected immune responses, protecting the body from autoimmune disease. They emerge from the thymus in a naive state before differentiating in the periphery into different populations with distinct modes of suppression of responses not only against self but also pathogens and cancer, limiting immunopathology and modulating immunotherapies. As the thymus shrinks with age, this naive pool becomes increasingly important to sustain T-cell diversity throughout human life. Combining transcriptome with epigenome profiling, this study identifies the gene regulatory program that maintains human naive regulatory T cells, distinguishing it from the one defining their general suppressive identity. Our findings may inform mechanisms of immune aging and new strategies for treating autoimmunity or improving transplant tolerance.

Introduction

Human CD4 T cells committed to the regulatory lineage during their development in the thymus are essential to maintain self-tolerance, control inflammation, and ensure tissue homeostasis (1–3). They egress the thymus as naive regulatory CD4 T cells (nTreg), defined by the expression of the lineage hallmark transcription factor (TF) FOXP3 and a naive phenotype best defined by CD45RA^{+}CD45ROnegCCR7^{+} (3–8). nTregs continuously circulate between blood and lymphoid tissues, ensuring the replenishment of the antigen-experienced Treg compartment, while simultaneously maintaining the breadth of the regulatory T-cell receptor (TCR) repertoire required to sustain tolerance across the full diversity of self and environmental antigens encountered over a lifetime (3–11). Most circulating Tregs in childhood and nearly half of those in healthy adults display a naive phenotype (6–8). Although their frequency decreases with age-associated thymic involution, the decline appears to be less pronounced than that of naive conventional CD4 T cells (nTconv) (6, 9). Their preferential use as starting material for

Received: September 3, 2026. **Accepted:** November 2, 2026
© The Author(s) 2026. Published by Oxford University Press on behalf of National Academy of Sciences.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

adoptive Treg-cell therapies reflects another facet of nTreg: they yield stably suppressive populations upon ex vivo expansion, whereas memory Tregs, the antigen-experienced fraction, are more plastic and prone to acquiring effector properties, likely due to the inclusion of peripherally induced Tregs (5). Yet, despite the nTreg's biological and translational importance, the molecular mechanisms that sustain nTreg identity at steady state remain poorly characterized, with most studies conducted in activated or memory Tregs (12–14). Moreover, while murine models have provided foundational insights into T-cell biology, a well-defined naive counterpart equivalent to the human CD45RA$^+$ naive pool is difficult to define, rendering murine models an imperfect surrogate for the study of nTreg identity (2, 4, 9, 10).

Characterizing the molecular basis of nTreg identity presents a challenge that is at once biological and analytical. Human nTregs are maintained through tonic TCR signaling from self-peptide–major histocompability complex (MHC) interactions and interleukin 7 (IL7)-dependent survival signals, despite their low levels of expression of the IL7 receptor, the same homeostatic signals that sustain their conventional counterparts (9, 15, 16). Superimposed on this naive identity is the Treg lineage program established during thymic development and epigenetically consolidated as illustrated by complete demethylation of the conserved noncoding sequence 2 (CNS2) within the *FOXP3* locus (17), elevated BCL2 expression (4, 9), and transforming growth factor beta–dependent signaling (18). nTregs, therefore, co-express transcriptional programs to govern both the Treg and naive identities. Profiling studies to date, including comprehensive multiomic atlases (19, 20), population-scale RNA-seq resources (21), and single-cell transcriptomic analyses of the human Treg compartment (22, 23), have not resolved these two component programs. Single-cell approaches are designed to resolve discrete populations and reveal Treg heterogeneity, rather than to untangle coexisting transcriptional programs common to each population (22, 23). Unmasking these programs requires analytical frameworks that model genome-wide expression variation across relevant cellular states simultaneously, most effectively when applied to sorted populations defined by well-established phenotypic markers.

The naive transcriptional program was largely defined as a quiescent, lymphoid-tissue-homing state. The TFs KLF2, FOXP1, FOXO1, BACH2, TCF7, and LEF1 have been shown to promote quiescence and to be enriched in human naive CD4 T cells, whereas effector TFs, such as TBX21 (T-BET), IRF4, and PRDM1 (BLIMP1), are reciprocally repressed (24, 25). Notwithstanding the recognized role of FOXP3 in Treg lineage specification (26), the human thymic Treg transcriptional program is organized into discrete gene-regulatory modules (GRMs), defined by the differential binding of TF families, notably AP1, ETS, and KLF/SP members, as we uncovered by integrating transcriptomic and chromatin accessibility data (27).

Here, by modeling genome-wide gene expression across thymic, naive, and memory differentiation stages within a unified linear framework, we uncoupled the transcriptional program defining human Treg and naive T-cell identities, enabling each to be characterized independently of the other. We then analyzed the regions of open chromatin (ROC) and the binding landscape of TFs at the promoters of signature genes across compartments and lineages, to uncover the regulatory architecture governing each program. Despite the increasingly recognized heterogeneity of thymically committed human Tregs and of peripheral Treg functions (28, 29), this approach revealed the cis-regulatory logic and hierarchy of TFs that sustain nTreg identity at steady state, which are relevant for immune competence throughout life and can be harnessed to develop Treg-based therapies with broad clinical application.

Results

The Treg signature shared between human nTregs and their thymic and memory counterparts

We started by investigating the transcriptional features defining naive CD4 T cells with Treg identity. The expression signature for human regulatory CD4 T cells was identified across their differentiation, from thymus to naive, and then memory compartments (Fig. 1A). Thymic samples were obtained during pediatric corrective cardiac surgery upon exclusion of immunodeficiency or syndromic diseases. Due to the impossibility of obtaining the required number of lymphocytes from paired pediatric thymus and blood samples, peripheral CD4 T cells were isolated from young blood donors, all females at birth to minimize the sex-associated variance. We quantified RNA-seq data obtained from purified CD4 single-positive (CD4SP) thymocytes from three children and paired naive (nTregs) and memory (mTregs) regulatory subsets from three young adults (Fig. S1A), and modeled genome-wide differential expression across the compartments according to a generalized linear model (see Materials and methods), in contrast with the respective Tconv counterparts (Figs. 1A and S1B). We selected as differentially expressed genes (DEGs) defining the Treg signature ($\log_2$FC $\geq \pm 1$, false-discovery rate [FDR] < 0.05) 293 genes for which the difference of expression between Tregs and Tconvs was consistently up-regulated (209 DEGs up, red) or down-regulated (84 DEGs down, blue) in all compartments (Figs. 1B and S2 and Data file S1).

The functional annotation of the Treg signature using gene set enrichment analysis (GSEA) captured both the core regulatory programs, such as IL2/STAT5 signaling and mTORC1-dependent metabolic control, in parallel with the suppression of pathways antagonistic to FOXP3 function, most notably WNT/beta-catenin signaling (Fig. 1C and D and Data file S1).

Overall, we defined the transcriptional identity of human Tregs established during thymic development, and preserved across the naive and memory states, providing a stable foundation of the human Treg lineage.

The regulation of the Treg signature in human nTregs

We then asked whether the regulatory mechanisms underlying the Treg signature were also conserved across thymic, naive, and memory compartments, specifically addressing the modulations for the naive state. We analyzed chromatin accessibility at the promoters of Treg signature genes and found that most ROC were preserved across all three compartments: 84% in thymic Tregs (tTreg),

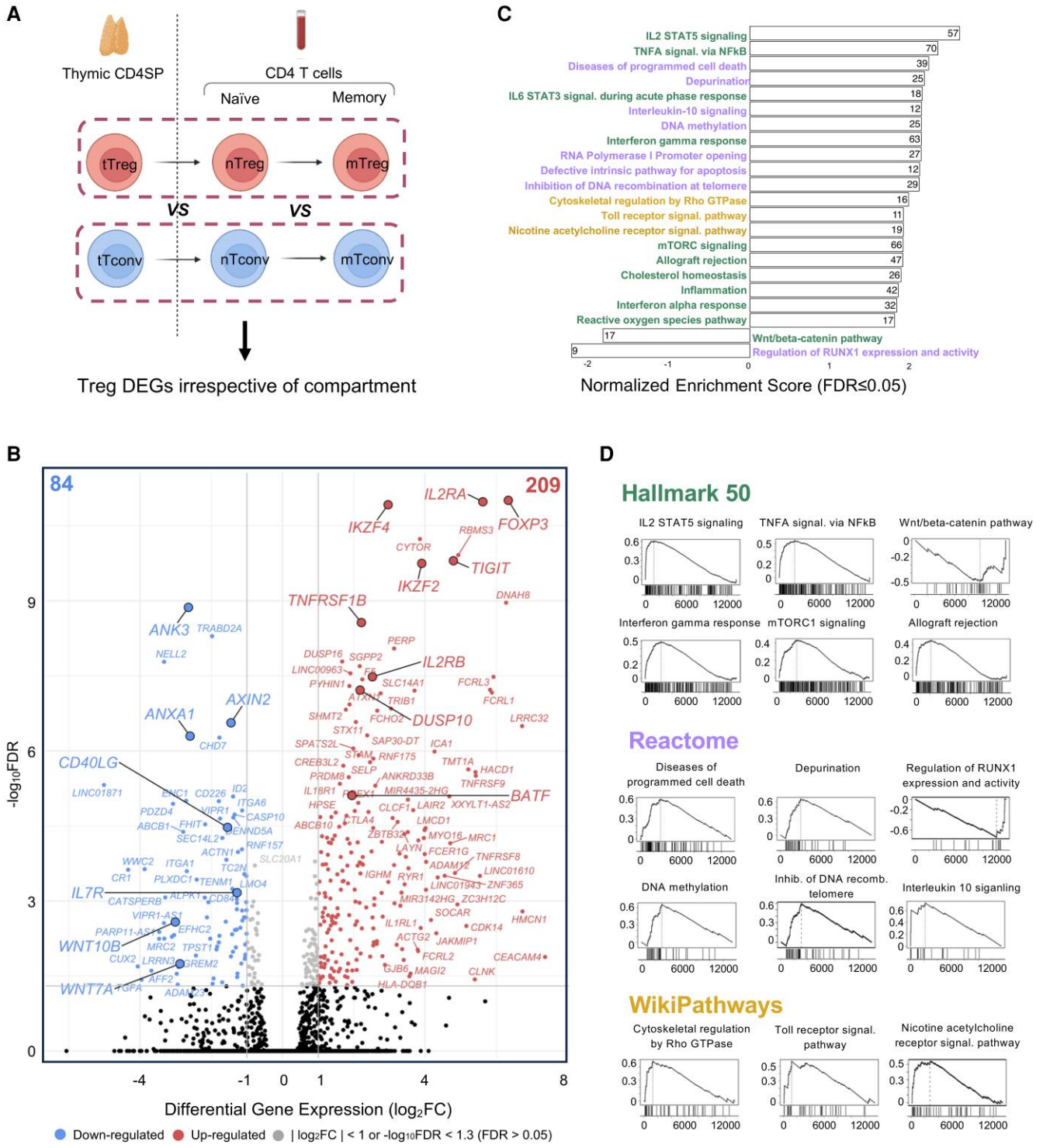

Figure 1 Regulatory CD4 T-cell signature shared between the naive and the thymic and memory compartments. A) Outline of the generalized linear model strategy to obtain the human regulatory (Treg) differential gene expression versus conventional (Tconv) CD4 T cells irrespective of compartment (tReg, thymic Treg; nTreg, naive Treg; mTreg, memory Treg; tTconv, thymic Tconv; nTconv, naive Tconv; mTconv, memory Tconv). B) Volcano plot for DEGs ($log_2FC \geq \pm 1$, FDR < 0.05) between Treg and Tconv: 209 DEGs up (red) and 84 DEGs down (blue) in Treg relative to Tconv; non-DEGs in gray and black. C) GSEA of the Treg signature, obtained by comparison against Hallmark 50 (green), WikiPathways (yellow), and Reactome signature (purple) collections. D) Normalized enrichment scores graphs for GSEA signatures highlighted in (C).

95% in nTregs, and 90% in mTregs, with only a minority showing compartment-restricted accessibility (Fig. 2A and Data file S2). We exploited the sequencing depth of ATAC-seq to perform digital genomic footprinting, which identifies TFs and quantifies respective occupancy through local depletion of Tn5 transposase insertions at binding sites, estimating the strength of interaction between TF with respective cis-regulatory sites (Data file S2). Analysis of TF-binding sites (TFBSs) within the shared ROCs revealed that most of the sites were likewise common to all compartments: 90% in tTregs, 94% in nTregs, and 92% in mTregs (Fig. 2B

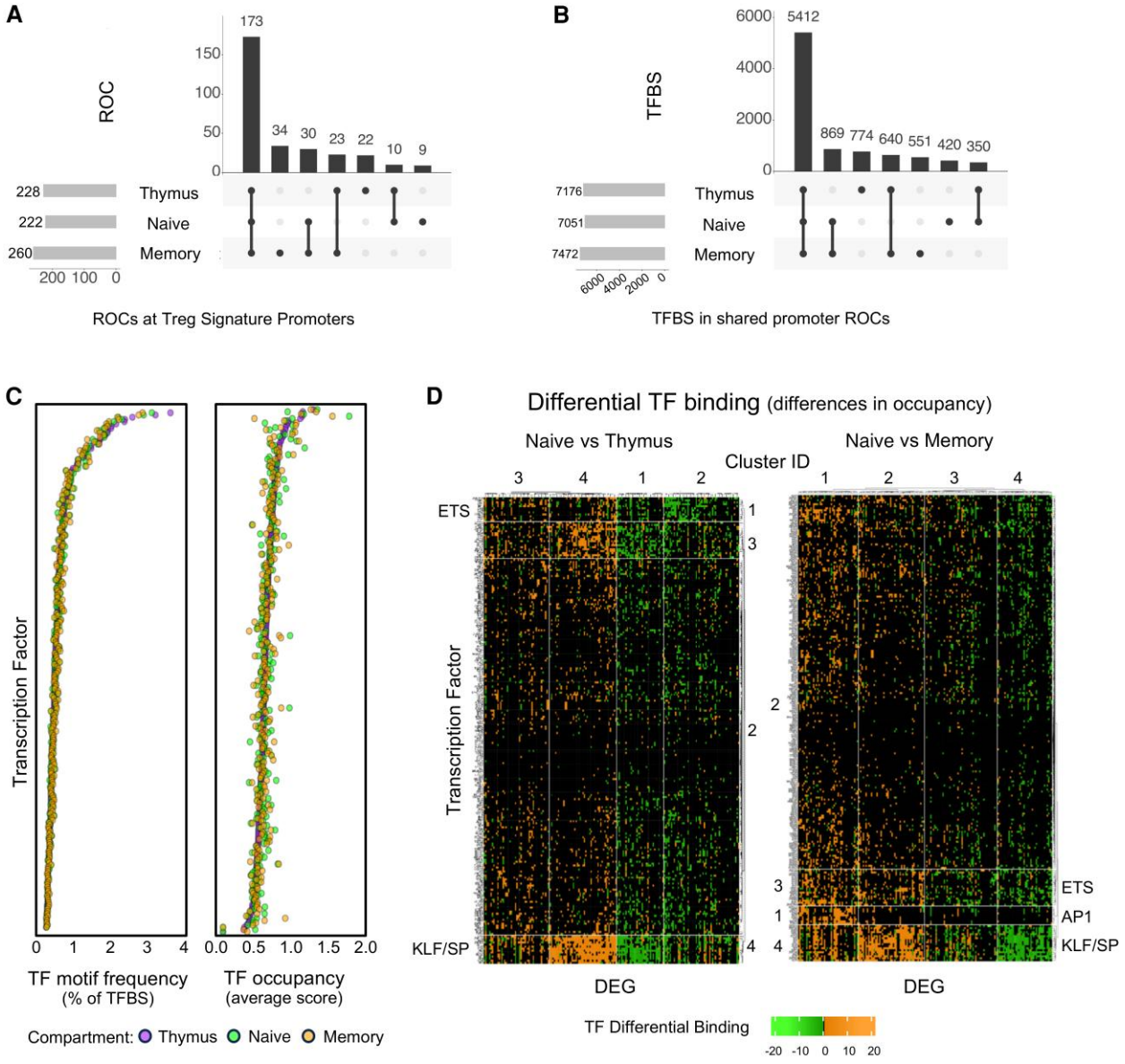

Figure 2 Regulation of the regulatory CD4 T-cell signature in human naive cells. A) ROC falling on the promoters of Treg signature genes in thymic Treg (tTreg), naive Treg (nTreg), and memory Treg (mTreg), with UpSet plot displaying intersections between the indicated sets. Horizontal bars (left) show the total membership of each set. Vertical bars (top) show the size of each intersection, ordered by decreasing size. Filled circles in the matrix below indicate which sets participate in each intersection; sets connected by a vertical line contribute jointly, whereas an isolated filled circle denotes a unique set membership. B) UpSet plot of TFBSs found in shared ROCs at Treg signature gene promoters across all three Treg subsets (tTreg, nTreg, and mTreg), displaying intersections between the indicated sets as in A. C) Left: Frequency of TFBS per TF (left) and average occupancy of TFBS per TF (right) at Treg signature gene promoters; purple, tTreg; green, nTreg; yellow, mTreg. D) Heat maps of the differential TF binding to the Treg signature DEGs promoters between nTregs and tTregs (left) and between nTregs and mTregs (right); labeled rows (TF) indicate clusters with higher TF binding in nTreg (orange). Left panel: ETS family TFs (TF-cluster 1/DEG-cluster 3); KLF/SP family TFs (TF-cluster 4/DEG-cluster 4). Right panel: AP1 family TFs (TF-cluster 1/DEG-cluster 1); ETS family TFs (TF-cluster 3/DEG-cluster 1); KLF/SP family TFs (TF-cluster 4/DEG-cluster 2). E) Scatter plots of TFBS frequency versus average occupancy per TF at Treg signature gene promoters in tTregs (left), nTreg (center), and mTreg (right) promoters; colors indicate relevant families of TFs.

and Data file S3). We also calculated the fraction of binding sites for each TF relative to all the TFBS at Treg gene promoters (TF motif frequency) and the average occupancy score of each TF (Fig. 2C and Data file S4). Despite the overall conservation of TFBS and similar TF motif frequencies, we observed quantitative differences in TF occupancy between compartments for the same TFs (Fig. 2C and Data file S4).

To systematically identify these dynamic regulatory patterns, we quantified differential binding for all TF–Treg signature gene interactions across developmental transitions, comparing naive with thymus or memory Tregs (Figs. 2D and S3A and Data file S4). The differential binding profiles for TF–target gene interactions allowed us to identify combinations of TF and DEG clusters, revealing discrete GRMs (27). Notably, we identified a GRM characterized by higher occupancy in naive than memory Tregs of many AP1 family members of TFs (BATF, FOS, FOSL2, JUNB, and JUND) and others (BACH1, BACH2, MAFK, and NFE2L2), with likely functional implications (Fig. 2D, TF cluster 1/DEG cluster 1, Data file S4). The AP1 complex target genes, including *BATF* itself and *WEE1*, *DUSP10*, and *CTLA4* pointing to concurrent maintenance of cell cycle restraint and suppressive competence (24, 26), while down-regulated targets include *GIMAP5*, a naive T-cell survival factor in humans (30, 31), and the effector-associated *CXCR3* (20, 24). KLF/SP TFs, mostly included in the TF clusters numbered as 4 in both transitions, exhibited compartment-specific differential binding across four distinct GRMs, with TF-DEG clusters characterized by high and low binding in nTreg of KLF/SP family to different genes in both "Naive versus Thymus" and "Naive versus Memory" (Fig. 2D and Data file S4). This regulatory function underscores the versatile role of KLF/SP family TFs in preserving the Treg phenotype, operating through quantitatively modulated binding to distinct target gene subsets during differentiation (27).

To characterize the TF engagement underlying these dynamic regulatory patterns, we plotted binding frequency against average occupancy scores (Fig. 2E). The KLF/SP family showed medium-to-low occupancy across medium-to-high binding frequencies in both Tregs and Tconvs (Figs. 2E and S3B and Data file S4). KLF/SP factors thus constitute the most frequently bound TF family at Treg signature gene promoters across thymic, naive, and memory compartments (Fig. 2E, top frequency TFs: ZNF148, KLF5, MAZ, ZNF263, SP2, SP3, SP1, KLF4, KLF2, and KLF16), and, therefore, their regulatory activity likely operates through compartment-specific adjustments in binding strength. Conversely, NFY pioneer factors displayed low binding frequency coupled with high occupancy at Treg signature gene promoters in both Tregs and Tconvs (Figs. 2E and S3B). The trimeric NFY complex (NFYA/NFYB/NFYC) has a generalized role in opening chromatin and recruiting other trans-acting factors across diverse developmental contexts, binding promoters initially and preceding histone posttranslational modification deposition (32). Therefore, NFY displays pioneer activity at promoters irrespective of cell type (32), suggesting a housekeeping function rather than a Treg-specific regulatory contribution, in accordance with the similar Tconv pattern (Fig. S3B and Data file S4). ETV3, in contrast, displays high occupancy at a restricted subset of promoters in Tregs and not in Tconvs (Figs. 2E and S3B and Data file S4). Among the ETV3 target genes, there were key genes such as *IL2RA*, *IL2RB*, and *NOSIP* in the three compartments and *IL7R*, *ANXA1*, and *GIMAP5* in both naive and memory Tregs (Data file S4), genes implicated in cell survival in the periphery (9, 16, 24, 30, 31, 33). This is particularly noteworthy given that ETV3 functions in the immune system have been mostly studied in dendritic cells, and our data support a Treg-intrinsic regulatory role that has not been previously described (34).

Collectively, we showed that the Treg transcriptional signature was epigenetically consolidated during thymic development and maintained with remarkable fidelity in the periphery. Our findings suggest that a fine-tuning dynamic modulation of TF occupancy likely defines the functional differentiation stages, prominently through the versatile activity of KLF/SP factors across distinct target gene subsets, without compromising the integrity of the core Treg program.

The signature shared between human nTregs and their conventional counterparts

Secondly, we examined the transcriptional features defining the naive identity, by further exploiting the linear model analysis strategy to compare expression profiles across thymic, naive, and memory compartments, ensuring that only transcriptional variation common to both regulatory and conventional T cells was attributed to the naive subset. We could thus select for gene expression changes specifically associated with the naive T-cell identity and independent of the regulatory or conventional lineage commitment. The naive CD4 T-cell signature was thus defined by those genes whose expression changed significantly both after thymic egression to the naive compartment and then upon differentiation into memory cells (Fig. 3A). Differential expression analysis ($\log2FC \geq \pm1$, $FDR < 0.05$) identified 1,947 genes up-regulated and 2,047 genes down-regulated in naive versus thymic T cells, alongside 821 genes up-regulated, and 940 genes down-regulated in naive versus memory T cells (Fig. 3B). The intersection of these two transitions yielded 1,167 DEGs constituting a lineage-independent naive T-cell signature (Fig. 3B and Data file S5). Then, we categorized the 1,167 naive signature genes according to their expression kinetics across compartments (Fig. 3C). This analysis revealed four distinct temporal patterns: 215 "NaiveDOWN" genes with lower expression in naive cells compared to both thymic and memory compartments; 411 "Rising" genes with progressively increasing expression from thymus through naive to memory; 366 "Falling" genes with progressively decreasing expression across the same differentiation trajectory; and 175 "NaiveUP" genes which expression peaks in the naive compartment (Fig. 3C and Data file S5). Expression heat maps confirmed these kinetic patterns were conserved across both Treg and Tconv lineages (Fig. 3C, right), demonstrating that naive T cells can be defined by a lineage-independent transcriptional program with distinct temporal dynamics during T-cell development and differentiation. The four kinetic patterns reveal that the naive compartment represents an actively enforced transcriptional identity.

Genes peaking specifically in the naive compartment (NaiveUP, Data file S5) were dominated by ribosomal protein subunits alongside the canonical naive CD4 marker *MAL* (35) and quiescence-associated regulator *TOB1* (36), suggesting that translational preparedness and active suppression of proliferation are defining features of naive T-cell identity (37, 38).

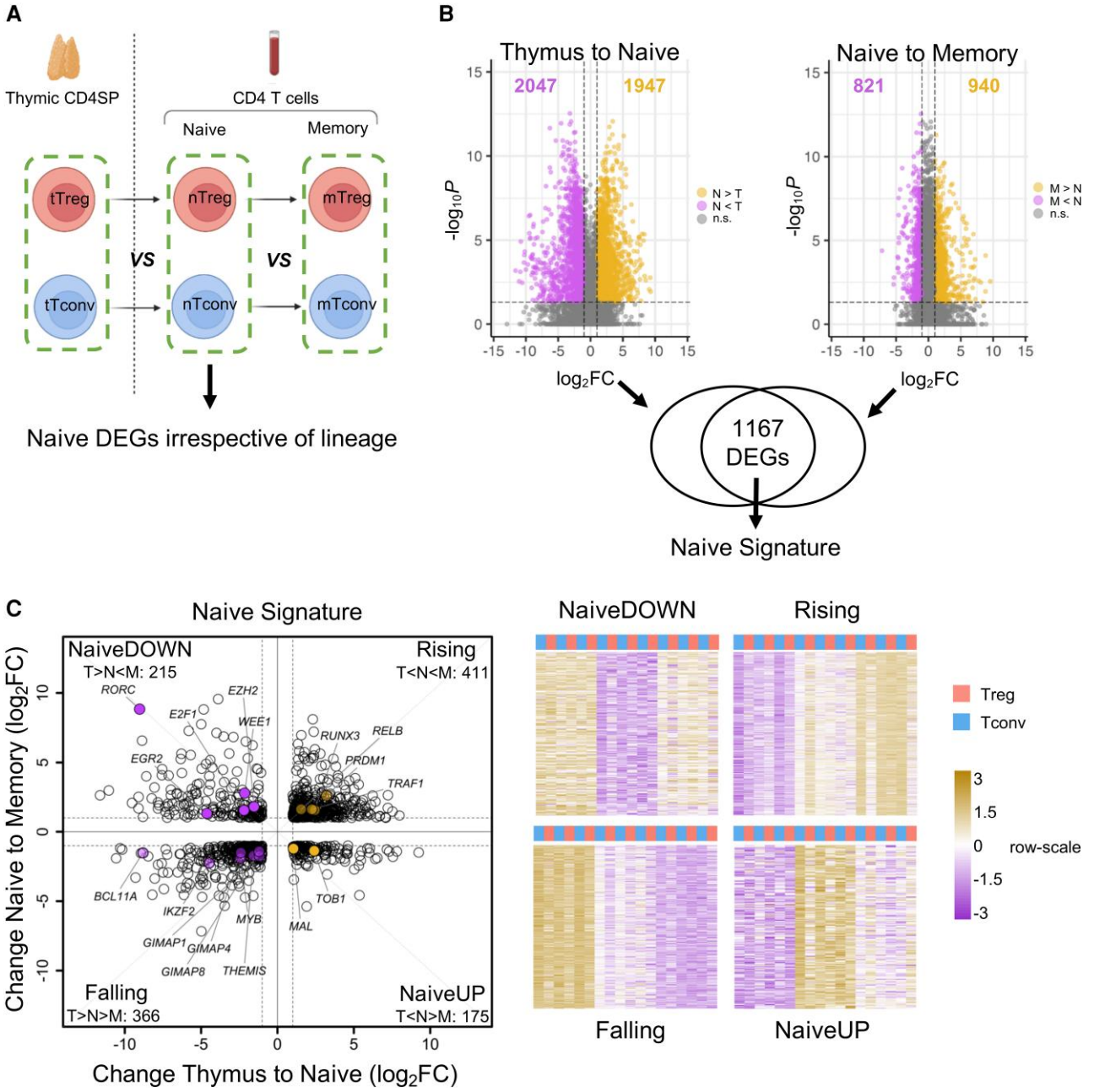

Figure 3 Naive CD4 T-cell signature shared between regulatory and conventional human CD4 T cells. A) Outline of the generalized linear model strategy to identify naive-specific DEGs irrespective of regulatory (Treg) or conventional (Tconv) lineage commitment. B) Volcano plots of DEG ($\log_2FC \geq \pm 1$, FDR < 0.05) across consecutive compartment transitions; non-DEGS in black. Left, naive versus thymus: 1,947 DEGs up-regulated in naive ("N > T," gold) and 2,047 DEGs down-regulated in naive ("N < T," purple). Right, memory versus naive: 821 DEGs up-regulated in naive ("M < N," purple) and 940 DEGs down-regulated in naive ("M > N," gold). The 1,167 DEGs common to both transitions constitute the naive signature of expression. C) Left, scatter plot of fold-change naive versus thymus (*x*-axis) against fold-change memory versus naive (*y*-axis) for the 1,167 naive signature genes, resolving four expression kinetic classes: "NaiveUP" (175 genes, higher in naive than in thymus and memory), "NaiveDOWN" (215 genes, lower in naive than in thymus and memory), "Rising" (411 genes, progressively higher from thymus to memory), and "Falling" (366 genes, progressively lower from thymus to memory). Right, heat maps of expression (CPM) for these same genes (rows) corresponding to each of the four patterns, with replicate samples grouped by compartment, in alternating lineages (columns: three replicates per cell type; red, Treg; blue, Tconv).

Naive T cells have been shown to possess a translational machinery that is assembled but largely idle: they harbor large numbers of preformed ribosomes, accumulate repressed mRNAs encoding metabolic enzymes, and are poised for immediate translational engagement upon TCR stimulation without requiring de novo ribosome biogenesis and additional layers of post-transcriptional regulation (37, 38). Our data further support a readiness transcriptional program of naive cells (37, 38) and suggest that it is shared by human nTregs.

Conversely, NaiveDOWN genes, specifically depleted in the naive compartment, include not only chemokine receptors associated with thymic development and activated memory states (*CCR4* and *CCR9*) but also key effector TFs (*EGR2*, *EGR3*, *IRF4*, *MAF*, and *RORC*) and proliferation regulators (*MKI67*, *EZH2*, *E2F1*, and *WEE1*),

underscoring naive state preservation through active transcriptional suppression of both activation and proliferative programs (Data file S5).

The Rising and Falling gene sets capture the longitudinal trajectory flanking this naive state: Falling genes, including thymic selection and survival factors (*THEMIS*, *GIMAP1*, *GIMAP4*, and *GIMAP8*) alongside TFs with prominent thymic roles (*MYB*, *BCL11A*, and *IKZF2*), reflect the progressive shedding of thymic identity following egression (Data file S5); while Rising genes, enriched for effector and memory commitment factors (*PRDM1*, *TBX21*, and *RUNX3*) and costimulatory mediators (*TRAF1* and *RELB*), may indicate a differentiation potential accumulating incrementally throughout peripheral T-cell life (Data file S5).

Together, these kinetic patterns define the naive compartment as a distinct, actively maintained transcriptional state. We found limited overlap between the Treg and naive signatures (Fig. S4A and Data file S6), demonstrating the uncoupling power of the linear model, which resolved naive identity from lineage commitment integrity and revealed a transcriptional program largely shared between regulatory and conventional naive CD4 T cells. Nevertheless, there were a few intersections at specific nodes (Fig. S4A and B and Data file S6). The dominant overlap category, NaiveDOWN/TregUP, includes genes associated with activated/effector Tregs, such as *TIGIT*, *TNFRSF4* (*OX40*/CD134), *IL2RB*, *IL18R1*, and *IRF4*, indicating that Treg identity in the naive compartment is in part achieved by actively overcoming naive-imposed transcriptional repression of key regulatory genes (Fig. S4A and B and Data file S6). Strikingly, *WEE1*, a cell cycle kinase that enforces G2/M checkpoint arrest, also falls in this category, suggesting that nTregs maintain a distinct cell-cycle regulatory state relative to nTconvs despite their shared quiescent phenotype, as previously suggested (9, 10). Conversely, the NaiveUP/TregDOWN category includes the negative regulator of NO synthase *NOSIP* (Fig. S4A and B and Data file S6), and a complementary pattern emerged from the NaiveRising/TregUP category, enriched for activation and inflammatory signaling components, including *TRAF1*, *TBX21*, *MAP3K8*, *MAP3K5*, *TBK1*, *ICAM1*, and *IL15RA* (Fig. S4A and B and Data file S6), further suggesting that nTconvs and nTregs shared an activation-responsive gene set before antigen encounter, consistent with a poised state for rapid effector deployment upon stimulation (37). Notably, the NaiveFalling/TregUP category contains *IKZF2* (Helios) and *LAYN* (Fig. S4A and B and Data file S6), two canonical Treg identity markers whose selective retention despite their overall declining trajectory possibly reflects active maintenance of thymic Treg identity in the periphery.

The functional annotation (overrepresentation analysis, ORA) on the complete 1,167-gene naive signature and revealed three distinct gene sets: one predominantly associated with "signaling receptor activity" and G protein-coupled receptor (GPCR) terms, another enriched for "nonsense-mediated decay (NMD)" functions and "cell adhesion mediator activity," while "cell junction" contained genes present in both groups (Fig. S5 and Data file S7). Among the contributing DEGs to the multiple terms related to GPCR signaling (Fig. S5A and B and Data files S5 and S7), there were transmembrane signaling molecules (e.g. *GNG8* and *GNGT2*); the chemokine receptors *CCR9* and *CCR4*, both repressed in naive T cells (NaiveDOWN) and playing roles in thymocytes and memory cell traffic; *GPR18* which was specifically up-regulated in the naive compartment (NaiveUP) and encodes an atypical lipid-sensing GPCR; and the Rising DEGs *ANXA1*, encoding the phospholipase A2 inhibitor Annexin A1, and *CXCR5.* The "NMD independent of the EJC" term includes (Fig. S5C, Data files S5 and S7), in addition to the ribosomal genes mostly NaiveUP DEGs, *ETF1* and *NCBP1* (39, 40)*: ETF1*, the translation termination factor that recognizes premature termination codons (PTCs) to trigger NMD, was less expressed in naive clusters relative to memory, consistent with its Rising classification, whereas *NCBP1*, a component of the cap-binding complex required for the pioneer round of translation that licences NMD, was specifically depleted in naive cells relative to both thymic and memory compartments (NaiveDOWN).

Collectively, our findings suggested that naive T-cell identity, irrespective of lineage relies on a coordinated adhesion program regulating lymphocyte recirculation and distinctive posttranscriptional regulatory mechanisms, which may ensure naive T-cell quiescence while preserving their poised state for rapid activation by enabling tight posttranscriptional control concomitant with accessible chromatin and ongoing transcription.

The regulation of the naive signature in human nTregs

Genes with expression patterns specifically enriched (NaiveUP) or depleted (NaiveDOWN) in the naive compartment, rather than those following progressive trends, are most likely to define the unique molecular features of the naive state independent of lineage commitment. Therefore, we focused on these gene groups for subsequent analysis of chromatin accessibility and TF-binding patterns that specify naive cells within human regulatory CD4 T cells.

To determine whether the expression of naive signature genes in nTreg and nTconv is dependent on lineage-specific chromatin accessibility, we quantified ROC at promoters of NaiveUP and NaiveDOWN genes in both nTreg and nTconv (Fig. 4A and Data file S8). The majority of ROC associated with both NaiveUP (4,831/6,214, 78%) and NaiveDOWN gene promoters (5,102/6,847, 75%) are present in the two lineages. Notably, even those ROC specific to one lineage fall predominantly within promoters harboring shared regulatory regions (688 nTreg-specific and 695 nTconv-specific for NaiveUP promoters; 812 nTreg-specific and 933 nTconv-specific for NaiveDOWN promoters). This pattern holds for both NaiveUP and NaiveDOWN genes, indicating that the chromatin accessibility landscape at naive signature gene promoters was mostly common to regulatory and conventional lineages.

We next asked whether the maintenance of a common naive signature could contain significant lineage-specific TF binding despite shared chromatin accessibility. Using digital genomic footprinting as described above, we quantified TFBSs within ROC at naive signature gene promoters that were shared between nTreg and nTconv (Fig. 4B, Data files S8 and S9). The vast majority of TFBS detected in these shared ROC were also common to both lineages in the naive compartment (4,567/5,369, 85%, for NaiveUP; 5,944/7,040, 84%, for NaiveDOWN), suggesting that a similar landscape of TF binding to the genome governed the naive signature in both regulatory and conventional lineages.

Differences in TF activity might nonetheless exist between these lineages. We first interrogated the differential expression

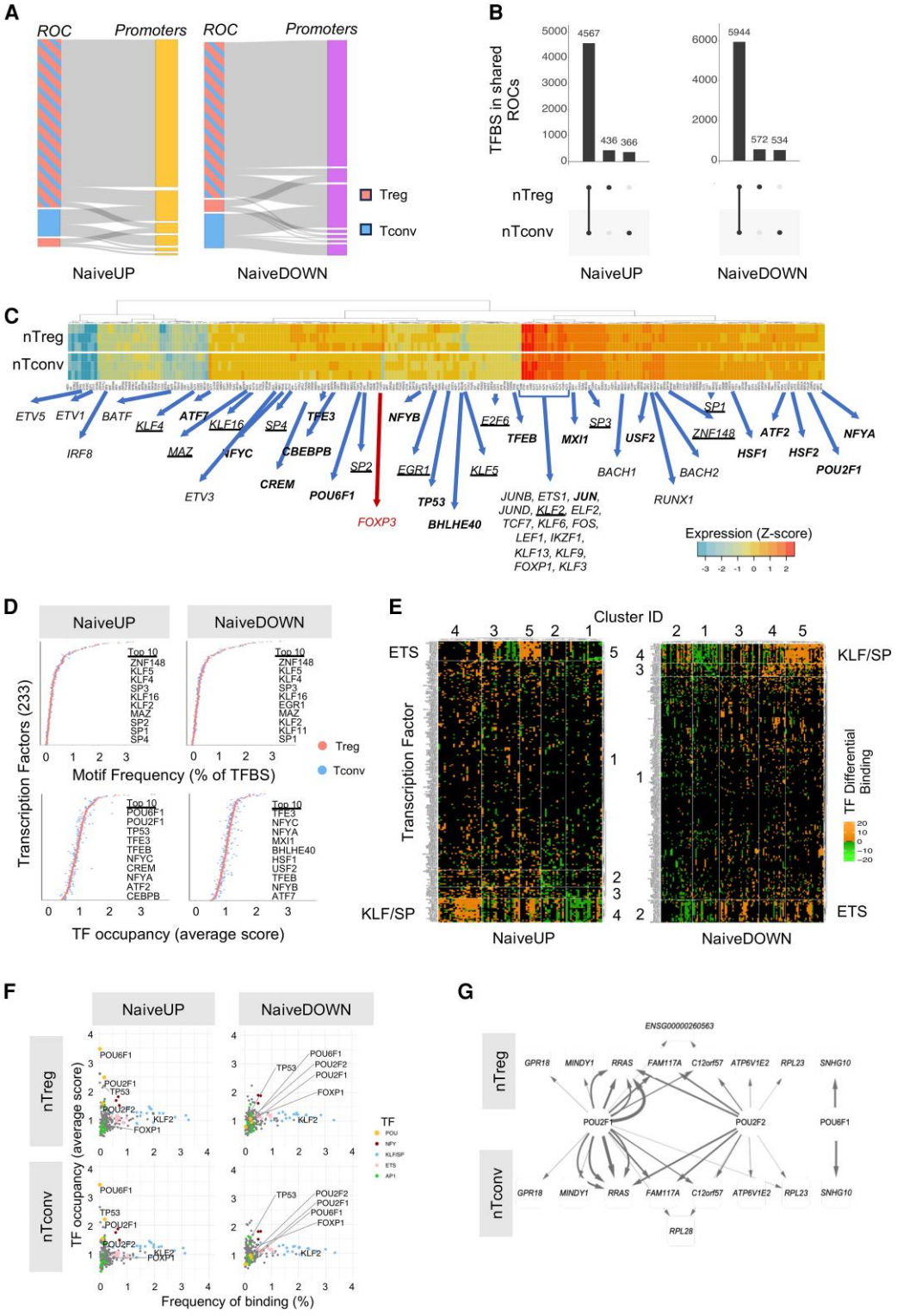

Figure 4 Regulation of the naive signature in human regulatory naive CD4 T cells. A) ROC falling on promoters of NaiveUP (yellow) and NaiveDOWN DEGs (purple): only in regulatory naive T cells (nTreg, red); only in conventional naive T cells (nTconv, blue); shared ROCs (red and blue). B) Analysis of TFBSs found in shared ROCs at the promoters of genes uniquely up-regulated (NaiveUP, left) and down-regulated (NaiveDOWN, right) in the naive compartment according to their presence in nTreg and nTconv. C) Expression levels of the analyzed 233 TFs identified in detected binding sites in nTreg and nTconv, highlighting relevant TFs already described and according to the pattern of binding to the naive signature genes: underlined, more frequently bound; bold, high binding scores; red arrow indicating FOXP3, down-regulated in nTconvs. D) Frequency of binding sites per TF in the promoters of genes uniquely up-regulated (NaiveUP, left) and down-regulated (NaiveDOWN, right) and TF average occupancy per binding site in NaiveUP and NaiveDOWN gene promoters in nTreg (red) and nTconv (blue); top 10 TF in nTreg are listed in each graph. E) Heat maps of differential TF binding between nTreg and nTconv for NaiveUP DEGs (left) and for NaiveDOWN DEGs (right); columns, DEGs; rows, TFs. Clusters of high TF binding in nTregs (orange) correspond to the KLF and ETS family TFs in both heat maps. F) Scatter plots of TFBS frequency versus average occupancy per TF at NaiveUP (left) and NaiveDOWN (right) gene promoters in nTreg (top) and nTconv (bottom); colors indicate relevant families of TFs. G) POU family TF binding targets among NaiveUP DEGs in nTreg (top) and nTconv (bottom); arrow thickness indicates relative occupancy score.

among all 233 TFs identified in TFBS detected for shared ROC, comparing their expression levels in both nTreg and nTconv (Fig. 4C and Data file S9). As expected, some TFs, such as *FOXP3*, were differentially expressed between Tregs and Tconvs (Treg signature, Data file S1). Remarkably, except for the few factors that are themselves part of the Treg signature, expression levels were similar between the two lineages (Fig. 4C and Data file S9). Thus, the conserved naive signature appears to be maintained by comparable TF repertoires operating on shared chromatin landscapes, independent of lineage commitment.

We then explored these TF activity data further by quantifying their motif frequency and their average occupancy score at the promoters of NaiveUP and NaiveDOWN genes in both nTreg and nTconv (Fig. 4D). The TF-DEG clusters were characterized by high and low binding in nTregs of KLF/SP family to both gene sets (Fig. 4E and Data file S9). KLF/SP factors showed higher occupancy at 23 NaiveUP target genes in nTregs relative to nTconvs, spanning the naive CD4 marker *MAL*, ribosomal proteins (*RPS5* and *RPL23*), metabolic regulators (*NPC2*, *ATP6V1G1*, *LIPT2*, and *NT5DC1*), signaling molecules (*PLCL1* and *DYRK2*), developmental factors (*NOG* and *CDNF*), and several lncRNAs (Data file S9). KLF/SP binding extended equally to 18 NaiveDOWN genes, including cell cycle regulators (*MCM6*, *GMNN*, and *WEE1*), signaling components (*GAB2*, *STAT5A*, and *EGR2*), and metabolic enzymes (*MAT2A* and *ATP1B3*), supporting that KLF/SP recruitment contributed to both the activation and repression arms of the naive program (Data file S9). Moreover, the analysis of TF frequency versus occupancy (Fig. 4F) also showed high frequency of binding and low occupancy of KLF/SP TFs for naive DEGs, regardless of lineage or differential expression pattern. Importantly, a group of TFs, including TP53, an important cell cycle controller and apoptosis regulator, and the POU factors POU2F1, POU2F2, and POU6F1, featured high occupancy at gene promoters of NaiveUP DEGs, a regulatory mechanism seemingly shared between nTreg and nTconv (Fig. 4F and Data file S9). Conversely, their occupancy at NaiveDOWN promoters falls within the range of other TFs (Fig. 4F and Data file S9).

POU6F1 regulates cell proliferation, and POU2F1/OCT1 plays an established role in maintaining naive CD4 T cells through association with the repressive NuRD complex, and in modulating activation responses (41–43). We therefore examined genes targeted by POU factors in both lineages (Fig. 4G). Shared targets include *GPR18*, a GPCR identified in the ORA analysis (Fig. S5A and B); the ribosomal protein subunit *RPL23*, included in the NMD biological pathway in the ORA analysis (Fig. S5A and C) also regulated by KLF/SP; *MINDY1*, a deubiquitinating enzyme; *ATP6V1E2*, involved in cellular homeostasis; *RRAS*, a regulator of MAPK and PI3K signaling; the long noncoding RNA *SNHG10*, implicated in transcription regulation; and *FAM117A* and *C12orf57*, both potentially involved in gene expression. *RPL28*, encoding another ribosomal protein subunit, is targeted exclusively in nTconv, while a single target (ENSG00000260563) was specific to nTreg.

Our findings revealed that naive identity is encoded by a common regulatory network, rather than lineage-specific rewiring, with KLF/SP factors providing quantitative fine-tuning through differential occupancy, possibly ensuring Treg function, while POU/OCT factors constitute a high-occupancy layer selectively driving NaiveUP gene activation, coordinating, across both lineages, a program linking ribosomal protein expression likely ensuring naive cell preparedness.

Discussion

This study established human nTreg identity governance by two largely independent transcriptional programs, whose regulatory architectures were both conserved and layered. The core Treg expression signature, established in the thymus, was maintained across peripheral compartments through a common chromatin landscape, with NFY providing constitutive promoter scaffolding, ETV3 emerging as a key repressive gene regulator, and KLF/SP factors serving as quantitative fine-tuners of occupancy repressing and activating distinct target gene subsets. Superimposed on this was a lineage-independent naive program that was not a transitional default but an actively enforced state, defined above all by translational preparedness, with elevated ribosomal protein gene levels and high POU2F1 (OCT1) occupancy at ribosomal gene promoters, poising naive T cells for rapid activation while maintaining quiescence. KLF/SP-mediated differential occupancy and AP1-driven naive-specific modules emerged as the main calibrators of gene expression at the nodes where regulatory identity diverges from conventional naive identity. Thus, we defined a layered architecture in which nTreg identity was actively maintained at steady state, with each component representing a potential point of dysregulation in autoimmunity or senescence. The translational implications of this architecture extend from the erosion of the naive pool in thymic loss and aging to the optimization of adoptive Treg therapies.

Our approach allowed us to define a 1,167-gene naive signature, lineage-independent and organized into four kinetic categories whose coexistence establishes the naive compartment as an actively maintained state rather than a transitional default between the thymus and memory populations. The observed overrepresentation of EJC-independent NMD machinery genes is concentrated among those up-regulated specifically in naive cells, where it is exclusively constituted by ribosomal subunits: *ETF1*, a molecular sensor of PTCs (39), was low in the naive compartment and increased with differentiation; and *NCBP1* (*CBP80*) was specifically low in the naive compartment, meaning that the pioneer round of mRNA translation mediating NMD is attenuated (40). Thus, rather than indicating active surveillance, this NMD enrichment is better interpreted in the converse sense: naive T cells are deficient in NMD through the coordinated down-regulation of mRNA licensing and stop codon recognition (40). In the absence of commensurate surveillance machinery, elevated ribosome abundance supports a model in which naive T cells maintain assembled ribosomes in a poised state, available for immediate deployment upon activation rather than engaged in ongoing posttranscriptional quality control (37). This is consistent with the observation that naive T cells harbor preformed ribosomes and repressed mRNAs primed for translation (38). In mice, ribosome biogenesis is itself rate-limiting for CD8 T-cell activation, underscoring the functional significance of this preassembled translational capacity (44), and recent findings have linked high ribosomal protein expression to longevity (45).

Our data also revealed POU factors, particularly POU2F1 (OCT1), as high-occupancy regulators selectively driving gene overexpression in the naive compartments in both lineages. Pou2f1 has been shown to maintain T-cell quiescence through association with the repressive NuRD complex, which sustains chromatin compaction and gene methylation at target loci; upon TCR stimulation, Pou2f1 dissociates from NuRD and recruits the OCA-B

coactivator and histone demethylase Jmjd1a, transitioning target genes from a poised to an active state (41, 42). Pou2f1 additionally maintains inter-chromosomal associations between cytokine gene loci in resting T cells, acting with CTCF and cohesin to enforce quiescence, positioning it as a chromatin organizer that poises T cells for controlled activation (43). The high POU2F1 occupancy observed here at ribosomal protein gene promoters provides the first evidence in human naive T cells, suggesting that this regulatory logic extends to translational preparedness: by maintaining accessible chromatin at ribosomal protein loci in resting cells, POU2F1 may sustain the preassembled translational capacity that enables rapid protein synthesis upon activation, while simultaneously maintaining the NMD machinery in a poised but not fully engaged state. That this architecture is preserved in nTregs specifically, a compartment whose rapid suppressive deployment is essential for peripheral tolerance, positions POU2F1-mediated translational poising as a previously unrecognized regulatory substrate with potential relevance to immune homeostasis and adoptive Treg therapy. Notably, naive CD4 T cells from supercentenarians have been shown to harbor high ribosomal gene expression and low inflammatory cytokine production (45), raising the possibility that POU2F1-mediated regulation of ribosomal protein genes constitutes a molecular signature of healthy immune aging.

As expected, the naive signature also captured the canonical lymphoid trafficking program, including GPCR-mediated homing signals and cell adhesion mediators, which is consistent with the established role of these pathways in nTreg recirculation between blood and secondary lymphoid organs (4, 6–8).

The intersection of the naive and regulatory programs revealed regulatory nodes at which Treg lineage commitment was superimposed on naive identity, mostly related to effector and cell cycling genes. This lineage-specific inflection appeared to be operating primarily through differential KLF/SP occupancy at distinct subsets of overexpressed and repressed target genes, with KLF/SP factors serving as quantitative controllers, calibrating gene expression within shared chromatin architecture rather than imposing lineage-specific regulatory rewiring.

The study was naturally constrained by the impossibility of obtaining paired thymic and peripheral blood samples from the same individuals. However, the experimental design, i.e. comparing across well-defined human compartments and lineages, was specifically structured to capture the relevant population-level biology. Accordingly, the study did not explore the heterogeneity of thymically committed Tregs and the multiple functional differentiation pathways of the peripheral memory Tregs, which would be better addressed with a single-cell approach. Several questions that arose directly from these findings require experimental validation, as illustrated below. The preferential ETV3 occupancy observed in Tregs raises the possibility of a cell-intrinsic repressive role analogous to its tolerogenic function in dendritic cells (34), which warrants direct functional interrogation in purified human Tregs. Whether GPR18-mediated lipid sensing directly instructs nTreg-cell positioning within secondary lymphoid organs, as suggested for GPR183/EBI2 (46), and how lineage-specific modulation of *ANXA1* expression is coordinated with differentiation, represent open questions with direct relevance to the maintenance of peripheral tolerance. Also, the role of posttranscriptional regulation mediated by NMD in the active maintenance of naive T-cell identity should be addressed through splicing analysis combined with NMD inhibition in primary cells (39, 40). Moreover, the convergence of POU2F1 occupancy at ribosomal protein loci raises the possibility that progressive loss of POU2F1-mediated translational preparedness constitutes a molecular mechanism of immune senescence, a causal contribution that remains to be directly tested. Finally, while this study focused on promoter-proximal chromatin accessibility, the enhancer landscape, including distal regulatory elements and their lineage-specific activity, was not systematically examined (27, 47, 48), leaving open the extent to which enhancer-driven regulation contributes to compartment and lineage-specific gene expression within these programs.

Overall, by resolving the molecular architecture that sustained human nTreg identity at steady state, we defined a set of regulatory interactions, spanning chromatin organization, TF occupancy, and posttranscriptional control, that constitute candidate targets for the selective expansion, stabilization, or functional reinforcement of this population in the clinical contexts where its erosion is most consequential: autoimmunity, immune senescence, and adoptive Treg therapy.

Materials and methods

Human samples

Thymic samples were obtained during pediatric reconstructive cardiac surgery, using tissue that would be otherwise discarded. Children with evidence of immunodeficiency, syndromic diseases and in the first month of life were excluded. Consecutive thymic samples were used (RNA-seq: female, 1.5 months; female, 12 months; male, 26.5 months. ATAC-seq: male, 1 month; female, 5.8 months; male, 9.3 months). Since it was not possible to obtain enough blood from thymus donor children to perform the required cell sorting, paired naive and memory peripheral blood samples were obtained from six young adult blood donors (three for RNA-seq and three for ATAC-seq), all female at birth and aged between 22 and 33 years old. All participants or their legal representatives provided written informed consent. The study was approved by the ethical boards of the Lisbon Academic Medical Center and of the Hospital de Santa Cruz, Unidade Local de Saude Lisboa Ocidental, Carnaxide, Portugal.

Cell sorting and flow cytometry analysis

Thymocytes isolated by Ficoll–Hypaque (GE Healthcare) from cell suspensions obtained by thymic tissue manual dispersion were sort-purified to obtain mature CD4SP regulatory (Tregs) and conventional (Tconvs) thymocytes (purities above 95%), based on the surface expression of CD4, CD8, CD27, CD25, and CD127 using an FACS Aria III (BD Biosciences), as previously described (27, 29), and illustrated in Fig. S1A. CD3 was intentionally not stained to avoid possible signaling, and the sorting strategy was validated in parallel through the analysis of CD3 and intracellular FOXP3 staining in a Fortessa flow cytometer (BD Biosciences), as we previously showed (27, 29). Peripheral mononuclear cells were isolated from buffy-coats by Ficoll–Hypaque (GE Healthcare), enriched for CD4 T cells by magnetic isolation using the untouched human CD4 T-cell enrichment kit (EasySep, StemCell Technologies), surface

stained for TCRαβ, CD4, CCR7, CD45RO, CD25, and CD127, and sort-purified into naive and memory Tregs and their Tconv counterparts, using an FACS Aria III, as we described (29). Analysis was performed using FlowJo v10 software (BD Biosciences).

RNA-seq

RNA was extracted from cell pellets of 600,000 sorted thymic, naive, and memory Tregs and Tconvs from three different thymuses and three different buffy coats, using the AllPrep DNA/RNA kit (QIAGEN) and following the manufacturer's instructions. A total of 18 libraries (i.e. three replicates per thymic, naive, and memory Tregs and Tconvs) were built selecting for polyadenylated RNA and then sequenced at both ends by high-throughput parallel sequencing (RNA-seq) in an Illumina Hiseq4000 sequencer (BGI Tech Solutions, Hong Kong, China). Raw sequencing was processed and analyzed with SAMtools (https://www.htslib.org/), and sequence quality assessed with FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The resulting ca. 200 million paired-end reads per biological replicate (PE100) were uniquely mapped and annotated to the human genome (hg38) with TopHat2 (https://github.com/DaehwanKimLab/tophat2), transcript expression quantified with R package HTSeq (https://htseq.readthedocs.io/en/latest/; count per million, CPM), with exclusion of genes with <1 CPM in more than two libraries (Data file S10). These libraries were used in previous studies (27, 29). Data accession: BioStudies E-MTAB-11211 for thymic data and E-MTAB-13930 for peripheral CD4 T-cell data.

Differential expression analysis

Libraries were scaled by trimmed mean of *M*-values (TMM) normalization and corrected for heterogeneity of samples specific to contrast matrix with weighted scaling based on voom followed by the quantification of differential expression with R package limma/edgeR (https://bioconductor.org/packages/release/bioc/html/limma.html). To extract single-outcome gene signatures from multiparametric predictors, we fitted all the data to a general linear model with lmFit and defined an experimental design with levels "compartment" and "celltype," and an interaction term. For the Treg signature, we considered the contrast between Tregs and Tconvs quantified by its average across compartments (after regression); for the naive signature, the contrast results in consecutive differential expressions between thymic and naive compartments, and then between naive and memory compartments, which are consistent among Tregs and Tconvs. Conversion between annotations was made with R biomaRt. FDR corresponds to adjusted *P*-value with Holm correction. The cut-off for expression of 2-fold change warrants the selection for differences with potential biological relevance and was overall on-pair with the chosen FDR < 0.05.

Functional annotation of expression signatures

ORA of the naive-expression signature made with gProfiler (https://elixir-europe.org/services/database/g:profiler), g:SCS multiple testing adjustment. GSEA of the Treg-expression signature made with GSEA module from WebGestalt, Bonferroni–Hochberg multiple testing adjustment (FDR). Threashold for both, $P_{adj} < 0.05$. The interactomes were generated by CytoScape 3.10.2 (https://cytoscape.org/) from GeneMania (https://genemania.org/) output, selecting the attributes "Physical interaction" and "Pathway," and spatially arranged by the yFiles Organic Edge Router layout, or Hierachical layout. We performed GSEA against Reactome (https://reactome.org/), WikiPathways (https://www.wikipathways.org/), and Hallmark 50 collection (https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp) from the Human MSigDB (Fig. 1C and D and Data file S1). This analysis revealed enrichment for the core Treg signaling pathway "IL-2/STAT5 signaling," other pathways consistent with Treg function ("IL-10 signaling" and "TNFA signaling via NFkB"), metabolic regulation ("mTORC signaling"), immune activation ("Allograft rejection," "Toll receptor signaling," and "interferon gamma response"), cytoskeletal regulation ("by Rho GTPase"), and epigenetic regulation ("DNA methylation"), alongside negative enrichment for WNT/beta-catenin signaling and RUNX1 regulation.

ATAC-seq processing and analysis

ATAC-seq data were generated from additional three thymuses and three buffy coats, yielding a minimum of ∼200 million PE50 reads per sample, in a total of 18 libraries (i.e. three replicates per thymic, naive, and memory Tregs and Tconvs), with sequencing quality assessed using FastQC, and processed as we previously described (27). Briefly, reads were aligned to the human reference genome (hg38) using Bowtie2. After processing, each library was used by the peak-caller MACS2 (https://github.com/macs3-project/MACS/wiki/Install-macs2) to detect the respective ROC. A unified set of genomic coordinates was created by merging peaks from all samples to create the reference landscape for the entirety of cell types considered, labeling each ROC with a unique Peak_ID. The original ATAC signal of each cell type was then mapped to the ROCs of the reference landscape using BAMscale (https://github.com/NLM-DIR/BAMscale) to obtain count libraries consistent within the full dataset. These were normalized to sequencing depth in CPM, excluded of peaks with <1 CPM in more than two libraries, scaled by TMM normalization and corrected for heterogeneity of samples specific to contrast matrix with weighted scaling based on voom. Promoter ROCs were defined as those intersecting human promoter regions from the FANTOM5 project (https://www.nature.com/collections/jcxddjndxy) and accessed through the Eukaryotic Promoter Database (https://epd.expasy.org/epd/human/human_database.php?db=human) using BEDtools (https://github.com/arq5x/bedtools2). Data accession: BioStudies E-MTAB-11220 for thymic samples and E-MTAB-16827 for peripheral CD4 T-cell samples.

TF-binding quantification

Characterization of TF activity followed the protocol that we previously described (27). Briefly, we used a digital genomic footprinting algorithm, TOBIAS 0.12.6 (https://github.com/loosolab/TOBIAS), to quantify occupancy and identify TFBSs.

TOBIAS-based footprinting is direct measure of occupancy distinct from motif enrichment: a bound TF physically protects DNA from tn5 tagmentation, creating a localized signal depression—the footprint—flanked by the elevated accessibility of the surrounding open chromatin. The TOBIAS score quantifies that depression relative to background, and bound versus unbound status is called per motif instance: TFBS were classified as "bound" for footprint scores with $P < 0.01$. TFBS consensus motifs were identified by matching to positional weight matrixes in the JASPAR Core database (2024 update, selecting 639 motif profiles matching "Homo Sapiens species" + "Latest Version"). Differential binding/occupancy between cell types and the definition of TF-DEG clusters was performed as before (27). The interactomes were generated by CytoScape 3.10.2, selecting the attributes "binding score," and spatially arranged by the Hierachical layout.

Statistical analyses and visualization

All quantifications and statistical significance were calculated with R (R Core Team [2023]. R: A Language and Environment for Statistical Computing), unless indicated otherwise. All charts and graphs created with packages from R/Bioconductor. Packages: heat maps were created with "ComplexHeatmap" or "Heatmap.2," volcano plots with "enhancedVolcano," upset plots with "UpSetR," Sankey plots with "networkD3," and other charts or graphs with the "ggplot2" package. Analysis design diagrams adapted from BioRender.

Acknowledgments

The authors thank Miguel Abecasis, MD, and Rui Anjos, MD, from Hospital de Santa Cruz/Unidade Local de Saúde Lisboa Ocidental (ULSLO), Instituto Português de Saude e Transplantação (IPST), and all parents and children for the access to thymic tissue and all donors of blood samples; Ana Godinho-Santos, Ana Serra-Caetano, André M.C. Gomes, Diana Santos, Helena Nunes-Cabaço, Madalena Marques, and Yumie Tokunaga, for the assistance in sample processing; Susana Paço and Miguel Ângelo-Dias for technical assistance; the GIMM flow cytometry facility for the isolation of cell populations.

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Competing Interest

The authors declare that they have no competing interests.

Funding

This work was supported by grants PAC-PRECISE-LISBOA-01-0145-FEDER-016394 and LISBOA-01-0145-FEDER-007391, co-funded by European Regional Development Fund (ERFD) through POR Lisboa 2020—Programa Operacional Regional de Lisboa PORTUGAL 2020 and Fundação para a Ciência e a Tecnologia (FCT); by national funding through FCT, under the project PTDC/MED-IMU/0938/2020, DOI: 10.54499/PTDC/MED-IMU/0938/2020 to A.E.S. P.R. was co-funded by FCT and ERDF. A.A.S.F.R. was funded by FCT (CEECIND/01474/2017).

Author Contributions

Alexandre A.S.F. Raposo (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing—original draft, Writing—review & editing), Pedro Rosmaninho (Formal analysis, Investigation, Writing—review & editing), Beatriz Moleirinho (Investigation, Writing—review & editing), Afonso R. M. Almeida (Investigation, Writing—original draft, Writing—review & editing), and Ana E. Sousa (Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Visualization, Writing—original draft, Writing—review & editing)

Data Availability

This study did not generate new unique reagents. The datasets supporting the conclusions of this article are available in the following manner: De-identified human RNA-seq and ATAC-seq have been deposited at BioStudies/ArrayExpress repository and are publicly available as of the date of publication. Accession numbers: https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-11211; https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-13930; https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-11220; https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-16827.

References

1. Sakaguchi S, *et al.* 2020. Regulatory T cells and human disease. *Annu Rev Immunol*. 38:541–566.
2. Sakaguchi S, Yamaguchi T, Nomura T, Ono M. 2008. Regulatory T cells and immune tolerance. *Cell*. 133:775–787.
3. Caramalho Í, Nunes-Cabaço H, Foxall RB, Sousa AE. 2015. Regulatory T-cell development in the human thymus. *Front Immunol*. 6:395.
4. Miyara M, *et al.* 2009. Functional delineation and differentiation dynamics of human CD4+ T cells expressing the FoxP3 transcription factor. *Immunity*. 30:899–911.
5. Hoffmann P, *et al.* 2006. Only the CD45RA+ subpopulation of CD4+CD25high T cells gives rise to homogeneous regulatory T-cell lines upon in vitro expansion. *Blood*. 108:4260–4267.
6. Seddiki N, *et al.* 2006. Persistence of naive CD45RA+ regulatory T cells in adult life. *Blood*. 107:2830–2838.
7. Silva SL, Sousa AE. 2016. Establishment and maintenance of the human naïve CD4+ T-cell compartment. *Front Pediatr*. 4:119.
8. Valmori D, Merlo A, Souleimanian NE, Hesdorffer CS, Ayyoub M. 2005. A peripheral circulating compartment of natural naive $CD4^+$ Tregs. *J Clin Invest*. 115:1953–1962.
9. Silva SL, *et al.* 2016. Human naïve regulatory T-cells feature high steady-state turnover and are maintained by IL-7. *Oncotarget*. 7:12163–12175.

10. den Braber I, *et al.* 2012. Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. *Immunity*. 36:288–297.
11. Lythe G, Callard RE, Hoare RL, Molina-París C. 2016. How many TCR clonotypes does a body maintain? *J Theor Biol*. 389:214–224.
12. Ferraro A, *et al.* 2014. Interindividual variation in human T regulatory cells. *Proc Natl Acad Sci U S A*. 111:E1111–E1120.
13. Miragaia RJ, *et al.* 2019. Single-cell transcriptomics of regulatory T cells reveals trajectories of tissue adaptation. *Immunity*. 50:493–504.e7.
14. Cuadrado E, *et al.* 2018. Proteomic analyses of human regulatory T cells reveal adaptations in signaling pathways that protect cellular identity. *Immunity*. 48:1046–1059.e6.
15. Sprent J, Surh CD. 2011. Normal T cell homeostasis: the conversion of naive cells into memory-phenotype cells. *Nat Immunol*. 12:478–484.
16. Azevedo RI, *et al.* 2009. IL-7 sustains CD31 expression in human naive CD4+ T cells and preferentially expands the CD31+ subset in a PI3K-dependent manner. *Blood*. 113: 2999–3007.
17. Ohkura N, Sakaguchi S. 2020. Transcriptional and epigenetic basis of Treg cell development and function: its genetic anomalies or variations in autoimmune diseases. *Cell Res*. 30:465–474.
18. Moreau JM, Velegraki M, Bolyard C, Rosenblum MD, Li Z. 2022. Transforming growth factor–β1 in regulatory T cell biology. *Sci Immunol*. 7:eabi4613.
19. Giles JR, *et al.* 2022. Human epigenetic and transcriptional T cell differentiation atlas for identifying functional T cell-specific enhancers. *Immunity*. 55:557–574.e7.
20. Graessel A, *et al.* 2015. A combined omics approach to generate the surface atlas of human naive CD4+ T cells during early T-cell receptor activation. *Mol Cell Proteomics*. 14:2085–2102.
21. Beumer N, *et al.* 2025. DNA hypomethylation traits define human regulatory T cells in cutaneous tissue and identify their blood recirculating counterparts. *Nat Immunol*. 26: 1315–1328.
22. Luo Y, *et al.* 2021. Single-cell transcriptomic analysis reveals disparate effector differentiation pathways in human Treg compartment. *Nat Commun*. 12:3913.
23. Zemmour D, *et al.* 2018. Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nat Immunol*. 19:291–301.
24. Choi JO, Seo Y, Hwang SS. 2025. Guardians of silence: transcriptional networks in T cell quiescence. *Exp Mol Med*. 57: 1663–1672.
25. Garaud S, *et al.* 2017. FOXP1 is a regulator of quiescence in healthy human CD4+ T cells and is constitutively repressed in T cells from patients with lymphoproliferative disorders. *Eur J Immunol*. 47:168–179.
26. Rudensky AY. 2011. Regulatory T cells and Foxp3. *Immunol Rev*. 241:260–268.
27. Raposo AASF, *et al.* 2024. Decoding mutational hotspots in human disease through the gene modules governing thymic regulatory T cells. *Front Immunol*. 15:1458581.
28. Moleirinho B, *et al.* 2025. A backbone-based flow cytometry approach to decipher regulatory T cell trajectories in the human thymus. *Front Immunol*. 16:1553535.
29. Raposo AASF, *et al.* 2025. The distinctive signature of regulatory CD4 T cells committed in the human thymus. *Front Immunol*. 16:1553554.
30. Patterson AR, *et al.* 2018. Gimap5-dependent inactivation of GSK3β is required for CD4+ T cell homeostasis and prevention of immune pathology. *Nat Commun*. 9:430.
31. Park AY, *et al.* 2024. GIMAP5 deficiency reveals a mammalian ceramide-driven longevity assurance pathway. *Nat Immunol*. 25:282–293.
32. Dolfini D, Imbriano C, Mantovani R. 2025. The role(s) of NF-Y in development and differentiation. *Cell Death Differ*. 32: 195–206.
33. Pekalski ML, *et al.* 2013. Postthymic expansion in human CD4 naive T cells defined by expression of functional high-affinity IL-2 receptors. *J Immunol*. 190:2554–2566.
34. Adams NM, *et al.* 2026. Transcription factor Etv3 controls the tolerogenic function of dendritic cells. *Science*. 391:eads1246.
35. Leitner J, *et al.* 2021. Differentiation and activation of human CD4 T cells is associated with a gradual loss of myelin and lymphocyte protein. *Eur J Immunol*. 51:848–863.
36. Tzachanis D, *et al.* 2001. Tob is a negative regulator of activation that is expressed in anergic and quiescent T cells. *Nat Immunol*. 2:1174–1182.
37. Ricciardi S, *et al.* 2018. The translational machinery of human CD4+ T cells is poised for activation and controls the switch from quiescence to metabolic remodeling. *Cell Metab*. 28: 895–906.e5.
38. Wolf T, *et al.* 2020. Dynamics in protein translation sustaining T cell preparedness. *Nat Immunol*. 21:927–937.
39. Kolakada D, *et al.* 2024. A system of reporters for comparative investigation of EJC-independent and EJC-enhanced nonsense-mediated mRNA decay. *Nucleic Acids Res*. 52:e34.
40. Kurosaki T, Popp MW, Maquat LE. 2019. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol*. 20:406–420.
41. Shakya A, *et al.* 2015. Oct1 and OCA-B are selectively required for CD4 memory T cell function. *J Exp Med*. 212:2115–2131.
42. Shakya A, Kang J, Chumley J, Williams MA, Tantin D. 2011. Oct1 is a switchable, bipotential stabilizer of repressed and inducible transcriptional states. *J Biol Chem*. 286:450–459.
43. Hwang SS, Kim LK, Lee GR, Flavell RA. 2016. Role of OCT-1 and partner proteins in T cell differentiation. *Biochim Biophys Acta*. 1859:825–831.
44. Tan TCJ, *et al.* 2017. Suboptimal T-cell receptor signaling compromises protein translation, ribosome biogenesis, and proliferation of mouse CD8 T cells. *Proc Natl Acad Sci U S A*. 114:E6117–E6126.
45. Zhu H, *et al.* 2023. Human PBMC scRNA-seq–based aging clocks reveal ribosome to inflammation balance as a single-cell aging hallmark and super longevity. *Sci Adv*. 9:eabq7599.
46. Baptista AP, *et al.* 2019. The chemoattractant receptor Ebi2 drives intranodal naive CD4+ T cell peripheralization to promote effective adaptive immunity. *Immunity*. 50: 1188–1201.e6.
47. Vahedi G, *et al.* 2015. Super-enhancers delineate disease-associated regulatory nodes in T cells. *Nature*. 520: 558–562.
48. Kundaje A, *et al.* 2015. Integrative analysis of 111 reference human epigenomes. *Nature*. 518:317–330.