nature immunology

ARTICLES

https://doi.org/10.1038/s41590-021-00980-8

Check for updates

Essential role of a ThPOK autoregulatory loop in the maintenance of mature CD4+ T cell identity and function

Jayati Basu[1], Bernardo S. Reis[2], Suraj Peri[3], Jikun Zha[1], Xiang Hua[1], Lu Ge[1], Kyle Ferchen[4], Emmanuelle Nicolas[1], Philip Czyzewicz[1], Kathy Q. Cai[5], Yinfei Tan[6], Juan I. Fuxman Bass[7], Albertha J. M. Walhout[7], H. Leighton Grimes[4], Sergei I. Grivennikov[8,9], Daniel Mucida[2] and Dietmar J. Kappes[1]✉

The transcription factor ThPOK (encoded by the *Zbtb7b* gene) controls homeostasis and differentiation of mature helper T cells, while opposing their differentiation to CD4+ intraepithelial lymphocytes (IELs) in the intestinal mucosa. Thus CD4 IEL differentiation requires *ThPOK* transcriptional repression via reactivation of the *ThPOK* transcriptional silencer element (Sil^{ThPOK}). In the present study, we describe a new autoregulatory loop whereby ThPOK binds to the Sil^{ThPOK} to maintain its own long-term expression in CD4 T cells. Disruption of this loop in vivo prevents persistent ThPOK expression, leads to genome-wide changes in chromatin accessibility and derepresses the colonic regulatory T (T_{reg}) cell gene expression signature. This promotes selective differentiation of naive CD4 T cells into GITR^{lo}PD-1^{lo}CD25^{lo} (Triple^{lo}) T_{reg} cells and conversion to CD4+ IELs in the gut, thereby providing dominant protection from colitis. Hence, the ThPOK autoregulatory loop represents a key mechanism to physiologically control ThPOK expression and T cell differentiation in the gut, with potential therapeutic relevance.

CD4 and CD8 T cell subsets develop in the thymus and play key roles in orchestrating cell-mediated immune responses. Major histocompatibility complex (MHC) specificity of the T cell receptor (TCR) expressed by double-positive (DP) CD4+CD8+ thymocytes drives CD4 versus CD8 lineage commitment. The transcription factor ThPOK acts as a 'master regulator' of CD4 commitment, which promotes the helper T cell lineage gene expression program and represses the T cell cytotoxic program[1–4]. ThPOK expression in thymocytes is regulated primarily at the transcriptional level via several stage- and lineage-specific *cis* elements, most notably the ~400-bp silencer (Sil^{ThPOK})[5,6], which has been functionally conserved since the divergence of marsupial and placental mammals 165 million years ago[7].

In peripheral CD4 T cells, persistent expression of ThPOK is required to maintain the helper T cell lineage gene expression program[4,8] and support effective CD4+ T cell memory[9]. Conversely, downmodulation of ThPOK expression drives conversion of mature CD4 T cells into class II-restricted cytotoxic T lymphocytes[10] or CD4+CD8αα IELs[11], and this conversion requires Sil^{ThPOK} activity[10]. The ability of Sil^{ThPOK} to toggle between on and off states in peripheral CD4 T cells has major consequences for T cell function and differentiation. For example, in an adoptive T cell transfer model of colitis, Sil^{ThPOK} actively downmodulates ThPOK expression by naive CD4 T cells arriving at the site of inflammation or infection, thereby controlling their effector fate[12]. Given the critical roles of ThPOK in CD4 T cell function and differentiation, understanding the molecular basis for its control by Sil^{ThPOK} is both fundamentally important and crucial for designing new therapeutic strategies against autoimmunity and infectious diseases.

In the present study, we identify an 'anti-silencer' element within Sil^{ThPOK}, which is essential for persistent ThPOK expression in mature CD4 T cells and consequently for maintenance of mature CD4 T cell identity and function. This element acts as a ThPOK-binding site, and drives a positive autoregulatory loop by coordinating enhancer–promoter interactions. We further demonstrate that the ThPOK-mediated, positive autoregulatory loop controls tissue-specific T_{reg} cell differentiation and T_{reg} cell conversion to CD4+ IELs. Finally, combining ThPOK chromatin immunoprecipitation sequencing (ChIP–seq), assay for transposase-accessible chromatin using sequencing (ATAC-seq) and RNA-sequencing (RNA-seq) data, we explore how the genome-wide regulatory landscape of naive CD4 T cells is fundamentally dependent on the ThPOK-mediated autoregulatory loop.

Results

Identification of an anti-silencer element. Sil^{ThPOK} activation is crucial for conversion of naive CD4 T cells into CD4+ IELs in the gut[11]. To understand how Sil^{ThPOK} toggles between alternate on/off states, we generated a series of knockout mouse lines carrying various deletions within Sil^{ThPOK} (Fig. 1a–c). Two mutants (ThPOK$^{\Delta OC74}$,

[1]Blood Cell Development and Cancer, Fox Chase Cancer Center, Philadelphia, PA, USA. [2]Laboratory of Mucosal Immunology, The Rockefeller University, New York, NY, USA. [3]Biostatistics and Bioinformatics, Fox Chase Cancer Center, Philadelphia, PA, USA. [4]Division of Immunobiology and Center for Systems Immunology, Cincinnati Children's Hospital 10 Medical Center, Cincinnati, OH, USA. [5]Cancer Signaling and Epigenetics, Fox Chase Cancer Center, Philadelphia, PA, USA. [6]Cancer Biology, Fox Chase Cancer Center, Philadelphia, PA, USA. [7]Program in Systems Biology, Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA, USA. [8]Cancer Prevention and Control, Fox Chase Cancer Center, Philadelphia, PA, USA. [9]Cedars-Sinai Medical Center, Departments of Medicine and Biomedical Sciences, Samuel Oschin Comprehensive Cancer Institute, Los Angeles, CA, USA. ✉e-mail: dietmar.kappes@fccc.edu

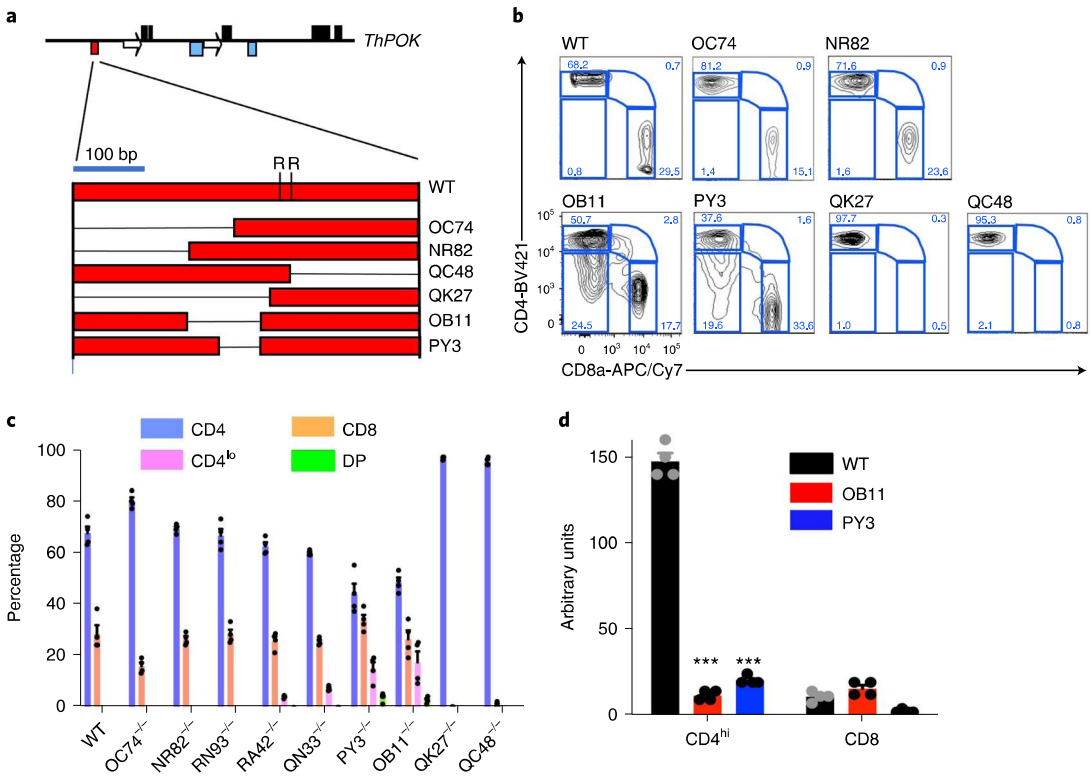


Fig. 1 | Genetic mapping of an anti-silencer element in mature CD4 T cells. a, Organization of mouse *ThPOK* gene (top) and diagram of silencer deletion mutants (bottom). Black boxes, blue boxes, arrows and red boxes indicate exons, enhancers, silencers and promoters, respectively. Deletions within the silencer are indicated by thin black lines. 'R' indicates positions of conserved Runx-binding sites. **b**, Flow cytometric analysis of CD4 and CD8a expression in gated TCRβ$^+$ peripheral blood lymphocytes (PBLs) of WT mice and homozygous mutant lines, as designated in **a**. **c**, Plots indicating percentage of SP CD4, CD8, CD4$^+$CD8$^+$ (DP) and CD4$^{lo/-}$ cells within gated TCRβ$^+$ PBLs of indicated strains ($n = 4$ for all groups). **d**, RNA collected from freshly isolated cells before probing for ThPOK expression by qPCR. Data represent four technical replicates, each derived from pooled RNA of three animals. Data are presented as mean ± s.e.m. A P value < 0.05 was considered to be significant. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. The experiment was repeated three times with similar results.

ThPOK$^{\Delta NR82}$) did not change peripheral T cell composition compared with wild-type (WT) controls, suggesting that Sil^{ThPOK} activity was unaffected, whereas two others showed complete absence of peripheral CD8 T cells (ThPOK$^{\Delta QC48}$, ThPOK$^{\Delta QK27}$), suggesting that Sil^{ThPOK} was constitutively turned off during thymic development, thereby preventing CD8 development (Fig. 1b,c). Finally, mutants ThPOK$^{\Delta PY3}$ and ThPOK$^{\Delta OB11}$ displayed profound alterations in mature T cell phenotype (Fig. 1b,c): (1) peripheral CD4 T cells defined by high CD4 levels (CD4^{hi}) were greatly diminished relative to CD8 T cells; (2) two unusual T cell (TCRβ$^+$) populations appeared, which, respectively, expressed both CD4 and CD8 (DP), or displayed low CD4 levels while lacking CD8 (CD4^{lo} cells).

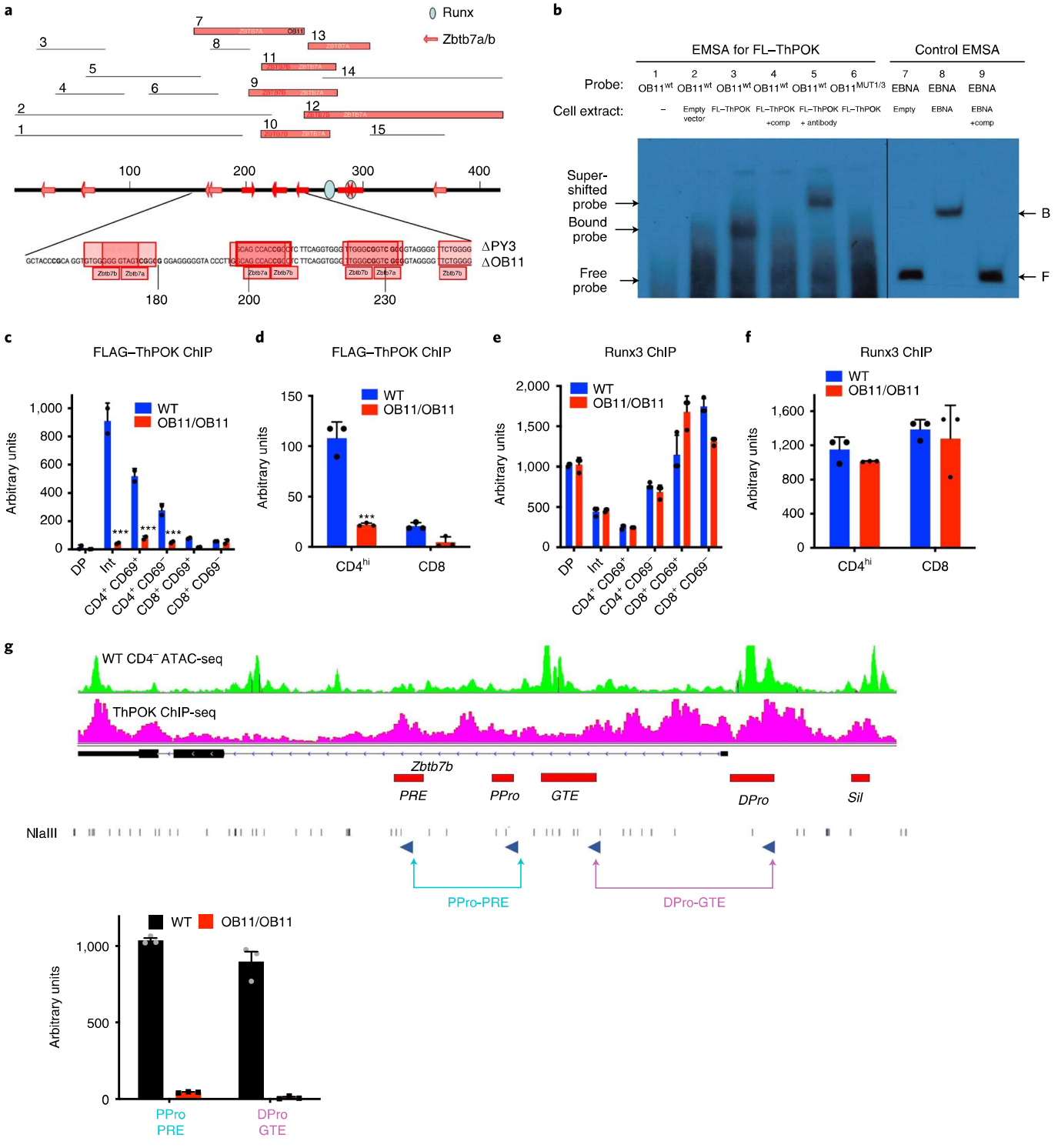
Quantitative reverse transcription PCR (RT–qPCR) analysis showed that ThPOK messenger RNA is almost absent in all

Fig. 2 | Regulation of anti-silencer by ThPOK binding. a, Numbered bars (top) representing 15 different mouse *ThPOK* silencer fragments that were tested as bait in Y1H analyses against ThPOK as prey (Zbtb7b). Fragments that bind ThPOK, and/or the related LRF (Zbtb7a) factor, are indicated by thick red bars, whereas those that do not show binding are marked as thin black lines. The thick black line (middle) represents the full-length silencer, with positions of consensus Zbtb7b/a-binding sites (as predicted using the JASPAR algorithm) marked by red arrows and conserved Runx-binding sites by blue ovals. The sequence of regions deleted in OB11 and PY3 mutants, with positions of predicted Zbtb7b/a-binding sites, are marked in red. **b**, EMSA analysis, using the 100-bp OB11 region (lanes 1–6) or 60-bp EBNA (Epstein–Barr nuclear antigen)-binding site control probe (lanes 7–9). Biotinylated probes were incubated with cell extracts from NIH 3T3 cells transfected with empty vector, FL-ThPOK or EBNA expression constructs, as indicated. In some lanes, unlabeled competitor (comp) DNA (lanes 4 and 9), anti-FL-ThPOK antibody (lane 5) or mutant OB11 probe, in which the consensus ThPOK-binding sites are destroyed (lane 6), were added. The experiment was repeated three times with similar results. **c**–**f**, ChIP analysis with antibodies against indicated FL-ThPOK (**c**,**d**) or Runx3 (**e**,**f**) for indicated sorted thymocyte (**c**,**e**) or LN T cell (**d**,**f**) subsets (Int, CD4$^+$CD8^{lo} subset). Red and blue bars indicate WT and ThPOK$^{OB11/OB11}$ cells, respectively. Data represent two (**c**) or three (**d**–**f**) technical replicates, each derived from pooled sorted cells of three animals. Data are presented as mean ± s.d. The experiment was repeated three times with similar results. **g**, A 3C assay performed with sorted peripheral CD4 T cells from WT or ThPOK$^{OB11/OB11}$ mice. Primer positions relative to ThPOK enhancers and promoters are marked (blue arrows, top panel). Quantitative PCR was performed to reveal interactions between indicated elements (lower panel). DPro, distal promoter; GTE, general T lymphoid element; PPro, proximal promoter; PRE, proximal regulatory element; Sil, silencer. Data were obtained from three technical replicates, each derived from pooled sorted samples of three animals, and presented as mean ± s.e.m. The experiment was repeated three times with similar results. A P value < 0.05 was considered to be significant. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

peripheral T cell subsets from ThPOK$^{\Delta OB11/\Delta OB11}$ and ThPOK$^{\Delta PY3/\Delta PY3}$ mice, indicating reactivation of Sil^{ThPOK} (Fig. 1d). In contrast, thymocytes from both lines showed normal stage-specific expression of ThPOK, except for a modest reduction at the most mature (CD69^{-}TCRβ^{+}) single-positive (SP) CD4 stage (Extended Data Fig. 1b), and a normal ratio of SP CD4:CD8 thymocytes (Extended Data Fig. 1a). Furthermore, Sil$^{\Delta OB11/\Delta OB11}$ β_2-microglobulin knockout ($\beta_2m^{-/-}$) mice generated only SP CD4 thymocytes, indicating normal lineage commitment of class II-restricted thymocytes (Extended Data Fig. 1c).

In summary, these results indicate that the OB11 region of Sil^{ThPOK} is dispensable for Sil^{ThPOK} activity in thymocytes, but essential to prevent inappropriate Sil^{ThPOK} activation in mature peripheral CD4 T cells.

Y1H analysis of ThPOK binding to the Sil^{ThPOK} element. The 100-bp OB11 deletion region contains four bioinformatically predicted ThPOK-binding sites, three of which overlap with the 50-bp PY3 deletion region (Fig. 2a). To directly assess ThPOK binding, we performed yeast-1-hybrid (Y1H) analysis[13–15] using 15 overlapping



40- to 100-bp fragments, spanning the whole Sil^{ThPOK} (Fig. 2a). It is noteworthy that ThPOK-binding sites identified by Y1H analysis map to a minimal 100-bp region (between fragments 5 and 14), coinciding with the OB11 region (Fig. 2a). Electrophoretic mobility shift assay (EMSA) confirmed that a 100-bp probe spanning the OB11 region binds ThPOK (Fig. 2b, lanes 3 and 5). Mutation of two predicted ThPOK-binding sites within the OB11 region (sites 1 and 3) abolished ThPOK binding by Y1H (Fig. 2a), and abrogated ThPOK binding in EMSA (Fig. 2b, lane 6). Mutation of site 3, but not site 1, partially impaired ThPOK binding in EMSA, suggesting greater importance of site 3 relative to site 1, but also partial functional redundancy between this and other sites (Extended Data Fig. 2a). To evaluate the contribution of individual binding sites in vivo, we generated three additional mutant lines with smaller deletions or mutations within the OB11 region ($ThPOK^{\Delta RA42}$, $ThPOK^{\Delta RN93}$ and $ThPOK^{\Delta QN33}$) (Extended Data Fig. 2b,c,d). Mutants $ThPOK^{\Delta RA42}$ (sites 3 and 4 ablated) and $ThPOK^{\Delta QN33}$ (sites 1, 3 and 4 ablated) show a slight shift to the $CD4^{lo}$ phenotype, indicating that site 2 is individually important for ThPOK expression, but partially redundant with other sites. Finally, $ThPOK^{\Delta RN93/\Delta RN93}$ mice, which mutate only site 1, exhibit no phenotype, indicating that site 1 is functionally redundant. Sites 2 and 3 are perfectly conserved between humans and mice, consistent with the important functional roles of these two sites (Extended Data Fig. 2d).

To confirm ThPOK binding to the $Sil^{\Delta OB11}$ silencer in vivo, we crossed $ThPOK^{OB11}$ mice with FLAG–ThPOK mice, encoding FLAG-tagged ThPOK, to generate compound heterozygous FLAG–ThPOK/$ThPOK^{\Delta OB11}$ mice and assessed ThPOK binding to the $Sil^{\Delta OB11}$ and Sil^{WT} elements by anti-FLAG ChIP using allele-specific primers. This demonstrated that ThPOK binds specifically to the Sil^{WT} but not the $Sil^{\Delta OB11}$ allele in thymocytes (Fig. 2c) and peripheral T cells (Fig. 2d). Of note, binding of Runx3 to Runx consensus motifs adjacent to the OB11 region was unaffected in the $Sil^{\Delta OB11}$ allele (Fig. 2e,f). Several enhancers and promoters have been identified at the ThPOK locus[5,6,16] (Fig. 2g). We used the 3C approach to test whether ThPOK binding to the OB11 region affects promoter and enhancer interactions in mature naive CD4 T cells. This revealed clear interaction between *ThPOK* enhancers and promoters in WT CD4 T cells under homeostatic conditions, which were abolished in $ThPOK^{\Delta OB11/\Delta OB11}$ T cells (Fig. 2h). We conclude that the 100-bp OB11 region is exclusively responsible for ThPOK binding to Sil^{ThPOK} and regulates enhancer–promoter interaction at the *ThPOK* locus.

ThPOK binding to Sil^{ThPOK} drives positive feedback loop. To assess the effect of the OB11 deletion on ThPOK expression at the single-cell level, we generated a green fluorescent protein (GFP) reporter allele under control of the $Sil^{\Delta OB11}$ element ($Sil^{\Delta OB11.GFP}$; Fig. 3a). Sorted CD4 thymocytes from heterozygous $ThPOK^{\Delta OB11.GFP/+}$ mice show high GFP expression, whereas mature CD4 T cells mostly lack GFP mRNA and protein expression, indicating selective block of reporter expression in peripheral CD4 T cells (Fig. 3b–d). Importantly, GFP transcription from the $Sil^{\Delta OB11.GFP}$ allele is not rescued by ThPOK expressed from the WT ThPOK allele in $Sil^{\Delta OB11.GFP/+}$ mice, confirming that the $Sil^{\Delta OB11.GFP}$ allele is insensitive to positive regulation by ThPOK (Fig. 3c).

Next, we crossed the $ThPOK^{\Delta OB11}$ strain to a second reporter line in which GFP is knocked into the WT ThPOK locus ($ThPOK^{GFP}$ mice), to create compound heterozygous $ThPOK^{GFP/\Delta OB11}$ mice (Fig. 3a–c). In the $ThPOK^{GFP}$ allele, GFP acts as a 'gene trap', precluding expression of ThPOK, so that the $ThPOK^{\Delta OB11}$ allele provides the only potential source of ThPOK in heterozygous $ThPOK^{GFP/\Delta OB11}$ mice. We observe considerably reduced GFP levels in $ThPOK^{GFP/\Delta OB11}$compared with $ThPOK^{GFP/+}$ mice (Fig. 3c,d), implying that continued ThPOK expression is necessary to prevent reactivation of the Sil^{ThPOK} in peripheral CD4 T cells. We also crossed $ThPOK^{\Delta OB11}$ mice to a third GFP reporter line in which Sil^{ThPOK} is deleted ($ThPOK^{\Delta Sil.GFP}$ mice; Fig. 3a). Importantly, GFP expression is fully restored in CD4 T cells of compound heterozygous $ThPOK^{\Delta Sil.GFP/\Delta OB11}$ mice compared with $ThPOK^{GFP/\Delta OB11}$ mice, showing that reduced reporter expression by $ThPOK^{GFP/\Delta OB11}$ CD4 T cells is silencer dependent (Fig. 3c, third column).

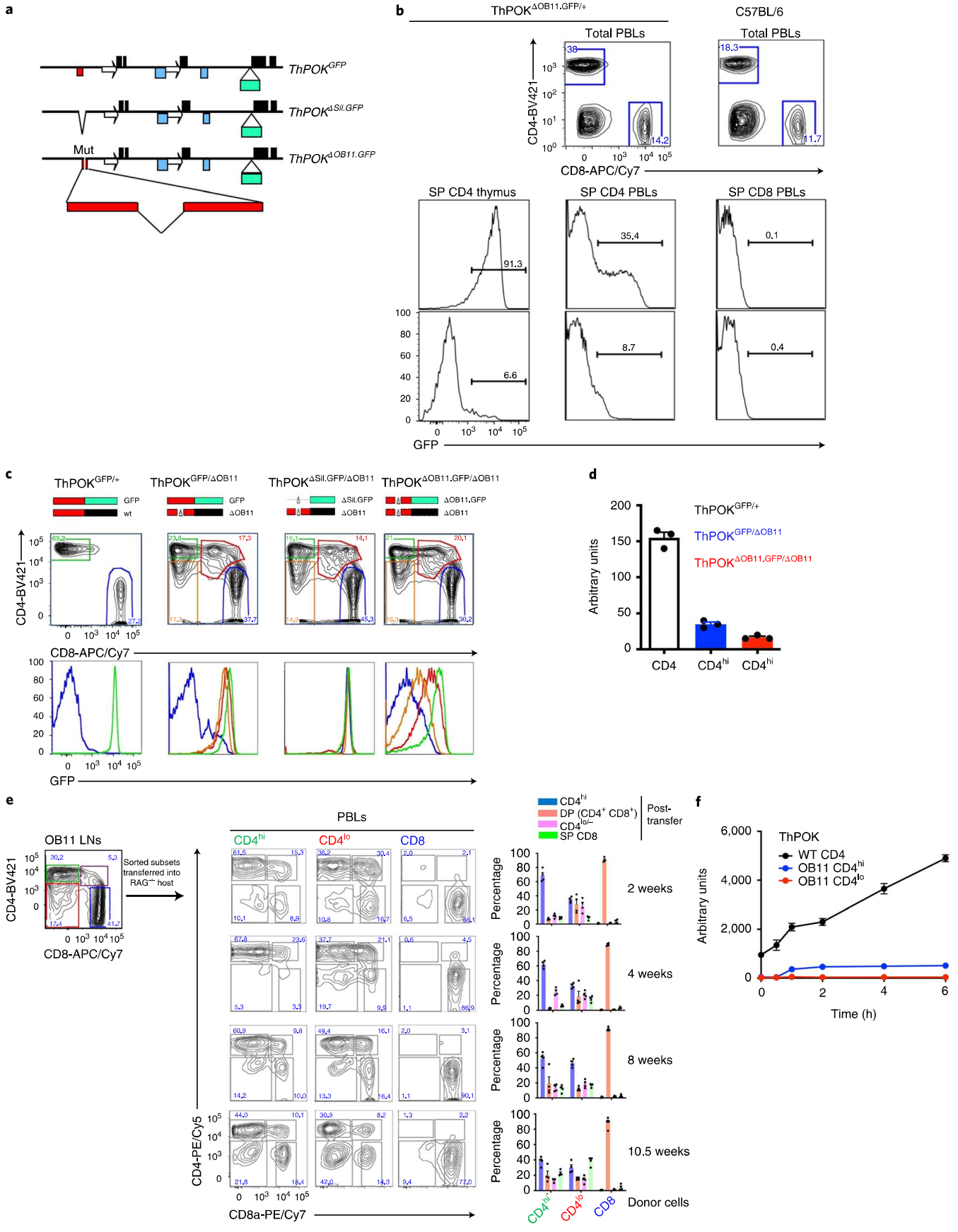
We noted that 35% of mature CD4 T cells from $ThPOK^{\Delta OB11.GFP/+}$ mice express GFP (Fig. 3b), although they show minimal GFP mRNA expression (at the population level). Such GFP^+ cells may represent recent thymic immigrants that retain long-lived GFP produced in the thymus, as in RAG–GFP mice[17], or T cells that have transiently reactivated the ThPOK locus post-thymically. To test the latter possibility, we adoptively transferred sorted GFP^- T cell subsets from $ThPOK^{\Delta OB11/\Delta OB11.GFP}$ and $ThPOK^{GFP/+}$ mice into $RAG^{-/-}$ hosts (Fig. 3e; data not shown). Of transferred $CD4^{hi}$ T cells, 20% expressed GFP 4 weeks after adoptive transfer (Extended Data Fig. 3b), consistent with post-thymic reactivation of the reporter allele. To assess whether this reactivation might depend on TCR-mediated induction, we treated CD4 T cells from $Sil^{\Delta OB11/\Delta OB11}$ versus WT mice with anti-TCRβ. TCR stimulation did not markedly increase ThPOK mRNA expression by $CD4^{hi}$ or $CD4^{lo}$ T cells from $ThPOK^{\Delta OB11/\Delta OB11}$ mice, implying that the OB11 allele is not subject to TCR-mediated reactivation (Fig. 3f). The surface expression of the TCR was unaffected (Extended Data Fig. 3c).

Cell transfer experiments showed further that $CD4^{hi}$ cells and $CD4^{lo}$ cells, from $ThPOK^{\Delta OB11/\Delta OB11.GFP}$ mice, are phenotypically plastic, that is, give rise to $CD4^{hi}$, $CD4^{lo}$, DP and CD8 subsets on adoptive transfer (Fig. 3e). In contrast, transfer of CD8 T cells from $ThPOK^{\Delta OB11/\Delta OB11.GFP}$ mice gave rise almost exclusively to CD8 T cells, indicating that SP CD8 T cells are phenotypically stable/terminal (Fig. 3e).

In summary, these results indicate that the ThPOK autoregulatory loop is essential for sustaining ThPOK expression and CD4 T cell lineage stability.

Fig. 3 | Loss of anti-silencer destabilizes CD4 T cell phenotype. a, Schematics of $ThPOK^{GFP}$, $ThPOK^{\Delta Sil.GFP}$ and $ThPOK^{\Delta OB11.GFP}$ reporter alleles. **b**, Flow cytometric analysis of CD4 and CD8a expression in total PBLs of WT mice and $ThPOK^{\Delta OB11.GFP/+}$ mice (top), and GFP expression of gated SP CD4 and CD8 PBLs (bottom rows). Ten animals of each genotype were analyzed in three separate experiments. **c**, Flow cytometric analysis of CD4 and CD8a expression in gated $TCR\beta^+$ PBLs (top), or GFP expression of gated $CD4^{hi}$ (green), $CD4^{lo/-}$ (orange), $CD4^+CD8^+$ (red) or CD8 PBLs (blue), of indicated mice. The bottom row illustrates *ThPOK* allele combinations. The red box represents silencer (deletions indicated by thin black lines); green and black boxes indicate GFP or ThPOK exons. A total of nine animals of each genotype were analyzed in three separate experiments. **d**, The qPCR analysis of ThPOK mRNA expression in sorted CD4 T cells from indicated mice. Data are obtained from three technical replicates, containing pooled mRNA from three animals, and are presented as mean ± s.e.m. The experiment was repeated three times. **e**, Reconstitution of $RAG^{-/-}$ hosts with $CD4^{hi}$, $CD4^{lo}$ and SP CD8 cells from $ThPOK^{\Delta OB11.GFP/\Delta OB11}$ donor mice. The left panel shows sort gates for donor cell isolation; the next three columns display CD4 and CD8 expression by PBLs from reconstituted hosts at indicated times after transfer of $CD4^{hi}$, $CD4^{lo}$ or CD8 T cells. The plots on the right indicate the percentage of indicated T cell subsets within $TCR\beta^+$ PBLs of mice reconstituted with the indicated $ThPOK^{\Delta OB11/\Delta OB11}$ donor cells. Donor cells were sorted from pooled LN samples of three $ThPOK^{\Delta OB11.GFP/\Delta OB11}$ mice, and four host animals were reconstituted with each sorted cell type. The experiment was performed three times. The bar graphs on the right present data from one experiment and are presented as mean ± s.e.m. **f**, Sorted T cell populations from the indicated mice (obtained from pooled LNs of four mice each) were stimulated with anti-TCRβ. The qPCR analysis of ThPOK mRNA was carried out at the indicated time points. Each data point represents three technical replicates, each derived from pooled sorted samples of three animals, and are presented as mean ± s.e.m.

Altered gene expression profile in ThPOK$^{\Delta OB11/\Delta OB11}$ T cells. To assess how the anti-silencer regulates ThPOK-dependent gene expression, we carried out RNA-seq analysis of sorted naive T cell subsets from ThPOK$^{\Delta OB11/\Delta OB11}$ and WT mice. This revealed that 1,724, 1,862 and 1,570 genes were differentially expressed (at least twofold) between WT CD4 T cells and CD4^{hi}, DP or CD4^{lo}



ThPOK$^{\Delta OB11/\Delta OB11}$ subsets, respectively (Fig. 4a,b and Extended Data Fig. 4b). Altogether 2,770 distinct genes were differentially expressed in one or more ThPOK$^{\Delta OB11/\Delta OB11}$ subsets compared with WT CD4 T cells. Of these 2,770 differentially expressed genes (DEGs), 789 (28%) show similar expression trend (up or down) for all ThPOK$^{\Delta OB11/\Delta OB11}$ T cell subsets, whereas the remainder show divergent expression between subsets, implying that phenotypically distinct ThPOK$^{\Delta OB11/\Delta OB11}$ T cell subsets are also distinct at the level of gene expression (Fig. 4a–c).

We noted that 760 (27%) of the 2,770 DEGs are CD8 like, consistent with derepression of a CD8 T cell transcriptional program (Fig. 4c). On the other hand, 1,042 DEGs are not CD8 like (false discovery rate (FDR) = 5%), but rather emulate a gene expression pattern characteristic of the early T cell progenitor stage (ETP)-to-double-negative stage 2 (DN2) transition in early thymopoiesis (including *Id2*, *Kit*, *Sca1*, *Pgk1*, *Meis1*, *Cd82* and *Sox13*) (Fig. 4g and Extended Data Fig. 4a), or of granulocyte differentiation (including *Mpo*, *Tyrobp*, *Fcerg1* and *Lyz*, a marker of DN thymocytes destined to differentiate to the granulocytic lineage[18]).

Next, to identify direct ThPOK targets among DEGs, we intersected our RNA-seq data with α-ThPOK ChIP–seq data from WT CD4 T cells[19] (Gene Expression Omnibus (GEO) accession no. GSE116506). Given the known functional antagonism between ThPOK and Runx3, we further intersected our RNA-seq data with α-Runx3 ChIP–seq data from WT CD8 T cells[20] (GEO accession no. GSE124912). This revealed 4,574 sites (associated with 3,200 genes) that are DUAL targets of ThPOK and Runx3 (that is, bind within 400 bp of each other) (Fig. 4d,f and Extended Data Fig. 4c,d). Among 789 DEGs common to all ThPOK$^{\Delta OB11/\Delta OB11}$ T cell subsets, 212 (24%) were associated with such dual target sites (Fig. 4a(bottom panel),d,e), including several critical transcriptional regulators of CD8 differentiation (for example, Eomes, RORα and T-bet) (Supplementary Table 1a,b). Among these 212 genes, 76 are downmodulated in all ThPOK$^{\Delta OB11/\Delta OB11}$ subsets, and of these 15 (20%) are implicated in regulation of TCR signaling (marked by pink boxes in Supplementary Table 1), especially dampening of TCR signaling (including *Camk2d*, *Kidins220*, *Pag1* and *Prkd2*) (Supplementary Table 1a).

Finally, genome-wide intersection of ThPOK ChIP–seq data indicated enrichment for Ets-domain family transcription factor and CTCF-binding sites in close proximity to ThPOK-binding sites (Fig. 4h). CTCF sites were predominantly associated with ThPOK sites not linked with Runx3 sites.

Altered chromatin accessibility in ThPOK$^{\Delta OB11/\Delta OB11}$ T cells. Our RNA-seq and T cell transfer experiments (Fig. 3e) indicate that ThPOK$^{\Delta OB11/\Delta OB11}$ T cell subsets exhibit phenotypic and transcriptional metastability. Metastable states in progenitors have been shown to be associated with epigenomic plasticity[21,22], metastable chromatin accessibility and multilineage transcriptional signatures[23,24]. To evaluate chromatin organization, we carried out ATAC-seq on naive ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} and WT CD4 T cells. This revealed strikingly altered chromatin accessibility, such that regions normally open in WT CD4 T cells are diminished in height in a genome-wide manner in ThPOK$^{\Delta OB11/\Delta OB11}$ T cells (Fig. 5a and Extended Data Fig. 5a). Nevertheless, we identified 2,767 differentially accessible chromatin regions (DACRs) that were selectively more open in ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} T cells. Intersection of ATAC-seq and ThPOK ChIP–seq data revealed that 566 of these selectively open DACRs are direct ThPOK target sites (Fig. 5b). These 566 regions showed enrichment of Runx-binding motifs compared with regions selectively open in WT CD4 T cells, suggesting a major role of ThPOK in countering Runx-dependent gene activation in CD4 T cells (Fig. 5c and Extended Data Fig. 5b). In contrast, regions selectively open in WT CD4 T cells showed enrichment of chromatin territory-defining regulatory factors Ctcf and Ctcfl (Boris), suggesting a possible role of ThPOK in collaborating with these factors to maintain accessible chromatin territory (Fig. 5c). DACRs that bind ThPOK map to 533 independent genes, of which 296 are more accessible in ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} T cells (Fig. 5b). Of these genes, 190 (80%) are relatively upmodulated in ThPOK$^{\Delta OB11/\Delta OB11}$ versus WT CD4 T cells, and associated DACRs tend to fall within introns rather than promoters (Fig. 5d).

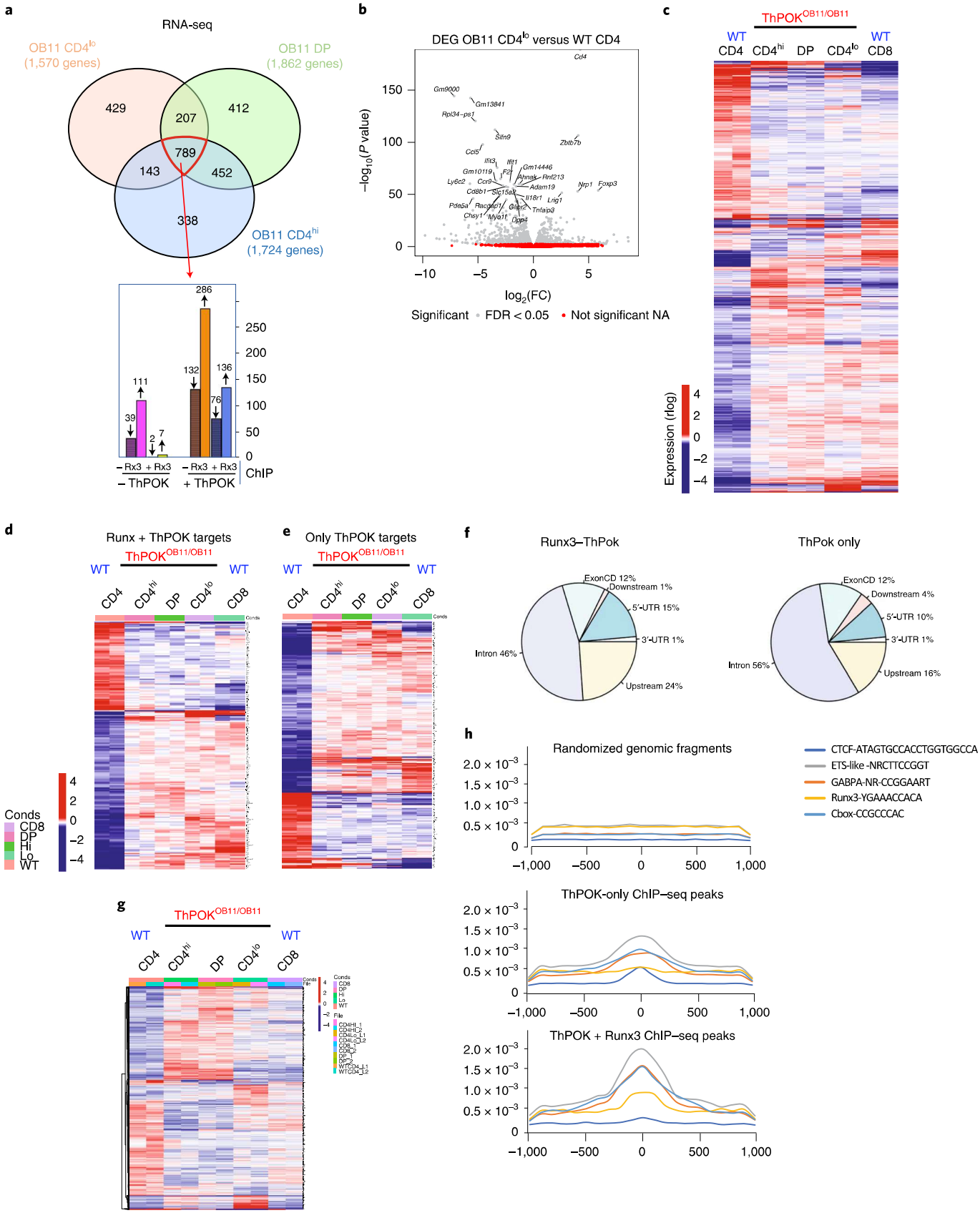
T cell 'identity'-determining genes are frequently associated with superenhancers (SEs)[25,26]. To explore whether ThPOK might modulate SE accessibility, we intersected ThPOK-binding sites that are differentially accessible in ThPOK$^{\Delta OB11/\Delta OB11}$ versus WT CD4 T cells with a public SE dataset (dbSUPER) (https://asntech.org/dbsuper)[27]. Strikingly, 67 (11%) and 98 (14%) of selectively open DACRs in ThPOK$^{\Delta OB11/\Delta OB11}$ T cells and WT CD4 T cells, respectively, mapped to helper T cell (Th)-specific SEs (Fig. 5e), but not to SEs associated with thymic development or early B cell development (Extended Data Fig. 5c). These Th SEs are closely linked to genes encoding Th master regulator T-bet, as well as critical cytokines (*Ifng*, *IL10*), chemokines (*Ccl4*) and cytokine receptors (*IL12rb*, *Tnfrsf4*), *Runx1/3*, *Gzmb*), TCR signal pathways (*Egr1*, *Nfatc2*, *Grb2*) and epigenetic regulators (*Dnmt3a*). Importantly, most of these genes are disregulated in ThPOK$^{\Delta OB11/\Delta OB11}$ versus WT CD4 T cells, supporting a key role for ThPOK in regulation of Th SE opening as well as their activity (Fig. 5f).

Finally, as many DEGs from ThPOK$^{\Delta OB11/\Delta OB11}$ T cells match the ETP gene expression pattern (Extended Data Fig. 4a), we compared chromatin accessibility between ThPOK$^{\Delta OB11/\Delta OB11}$ T cells and WT thymocyte developmental stages (Immgen). Indeed, regions that are less accessible in ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} T cells are also less accessible in ETPs (DN1–DP) and vice versa (Fig. 5g).

Fig. 4 | Ablation of anti-silencer leads to deregulation of the CD4 T cell gene expression program. a, Venn diagram illustrating the intersection between gene subsets that are differentially expressed between the indicated T cell subsets. The number of DEGs for each subset is shown in brackets below the subset name. The bar graph at the bottom indicates the number of 789 commonly misregulated DEGs (upward and downward arrows indicate up- and downregulated genes, respectively) that are direct targets of ThPOK, Runx3, or both ThPOK and Runx3, as indicated. **b**, Volcano plot illustrating gene expression differences between OB11 CD4^{lo} and WT CD4 T cells. Gray dots represent genes differentially expressed (adjusted $P < 0.05$) between samples. Genes with the largest negative or positive standardized mean difference are marked (Wald's two-sided test). DEGs with $P_{adj} < 0.05$ are considered significant. FC, fold-change. **c**, Heatmap displaying hierarchical clustering of DEGs for indicated T cell subsets. Analysis is restricted to the union of all 2,770 genes differing in expression between any OB11 T cell subset and WT CD4 T cells (FDR of 5%). Red and blue indicate increased or decreased levels, respectively. **d**,**e**, Heatmaps displaying hierarchical clustering of DEGs that are commonly misregulated between any OB11 T cell subset and WT CD4 T cells (789 genes) and are targets of either both ThPOK and Runx3 (212 genes) (**d**) or ThPOK alone (318 genes) (**e**) (FDR = 5%). Red and blue indicate increased or decreased gene expression levels, respectively. **f**, Pie charts illustrating distribution of ThPOK-binding sites within DEGs that are direct targets of ThPOK (right) or ThPOK and Runx3 (left). **g**, Heatmap displaying hierarchical clustering of DEGs for indicated T cell subsets. Analysis is restricted to 1,042 DEGs with expression that differs between any ThPOK$^{\Delta OB11/\Delta OB11}$ subset and WT CD4 T cells, but not between WT CD4 and WT CD8 T cells, that is, genes that are not CD8 like (FDR = 5%). Red and blue indicate increased or decreased gene expression levels, respectively. **h**, Co-distribution of ThPOK ChIP–seq peaks with other transcription factor–binding sites for ThPOK peaks associated with Runx motifs (bottom), ThPOK peaks not associated with Runx motifs (middle) or random genomic fragments not bound by ThPOK (top).

In summary, ThPOK$^{\Delta OB11/\Delta OB11}$ T cells display profound changes in gene expression and chromatin accessibility, including SE accessibility, resulting in a metastable state characterized by transdifferentiation to the CD8 lineage and de-differentiation to the ETP stage.

Altered T cell function in ThPOK$^{\Delta OB11/\Delta OB11}$ T cells. As the ThPOK autoregulatory loop plays an important role in regulation of genes involved in TCR signal strength, we compared proliferation and cytokine production of sorted WT and ThPOK$^{\Delta OB11/\Delta OB11}$ T cells in



response to weak (anti-TCRβ) or strong (anti-CD3/CD28) TCR signals. Strikingly, T cells from ThPOK$^{\Delta OB11/\Delta OB11}$ mice exhibit stronger proliferative response and more sustained ERK phosphorylation compared with WT CD4 T cells in response to weak stimulation (Extended Data Fig. 6a,b), supporting the ThPOK autoregulatory loop limiting TCR signal responsiveness. Furthermore, we uncovered promiscuous deregulation of multiple cytokine genes (*Ifng*, *IL9* and *IL17*) and associated transcription factors (*Tbx21*, *Rorc*, *Ahr*, *Irf4*) in ThPOK$^{\Delta OB11/\Delta OB11}$ T cells. It is interesting that we identified ThPOK ChIP–seq peaks within the *Ifng* and *Tbx21* SEs (Fig. 5f), the *Ahr* promoter, as well as within genes encoding *Rorc*, the interleukin (IL)-9-inducing factor *Irf4*, and to the CNS25 regulatory region of the *IL9* gene (Extended Data Fig. 6e), suggesting that the ThPOK autoregulatory feedback loop is required to maintain 'naiveness' in CD4 T cells (Extended Data Fig. 6c,d). Finally, activated ThPOK$^{\Delta OB11/\Delta OB11}$ T cells display strong skewing toward a T_H1/inflammatory response relative to WT CD4 T cells, consistent with the known role of TCR signal strength in controlling Th differentiation[28] (Extended Data Fig. 6d).

Anti-colitogenic activity of ThPOK$^{\Delta OB11/\Delta OB11}$ T cells. Conditional T_{reg} cell-specific targeting of ThPOK causes substantial reduction in mature T_{reg} cells but no overt autoimmune or inflammatory phenotype[11,29]. Similarly, ThPOK$^{\Delta OB11/\Delta OB11}$ mice show no evidence of autoimmunity at least till age 15 months. To evaluate nT_{reg} (natural T_{reg}) cell development and differentiation, we introduced a Foxp3–red fluorescent protein (RFP) reporter allele on to the ThPOK$^{\Delta OB11/\Delta OB1}$ background. We observed a notable decrease in peripheral T_{reg} cells in ThPOK$^{\Delta OB11/\Delta OB11}$ mice. It is interesting that the anti-colitogenic GITR^{lo}PD-1^{lo} CD25^{lo} (Triple^{lo}) subset[30] was markedly increased relative to the GITR^{hi}PD-1^{hi}CD25^{hi} (Triple^{hi}) (where GITR is glucocorticoid-Induced TNFR-related protein and PD-1 is programmed cell death protein 1) subset in peripheral ThPOK$^{\Delta OB11/\Delta OB11}$ T_{reg} cells (Fig. 6a). Of note, in WT mice Triple^{lo} T_{reg} cells express a much lower level of ThPOK than Triple^{hi} T_{reg} cells, suggesting that relative ThPOK expression may contribute to alternate differentiation of these subsets under physiological circumstances (Fig. 6b). Next. we tested the effect of autoregulatory loop ablation on iT_{reg} (induced T_{reg}) cell generation. Although both WT and ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{hi} T cells cultured under T_{reg} cell-polarizing conditions were able to induce Foxp3, Foxp3 levels were lower for ThPOK$^{\Delta OB11/\Delta OB11}$ T cells. Strikingly, ThPOK$^{\Delta OB11/\Delta OB11}$ T cells gave rise to large numbers of Triple^{lo} T_{reg} cells, which also expressed lower levels of FR4 and CD73 compared with WT T cells (Fig. 6c). Hence, ThPOK promotes the Triple^{hi} T_{reg} cell fate at the expense of the Triple^{lo} T_{reg} cell fate.

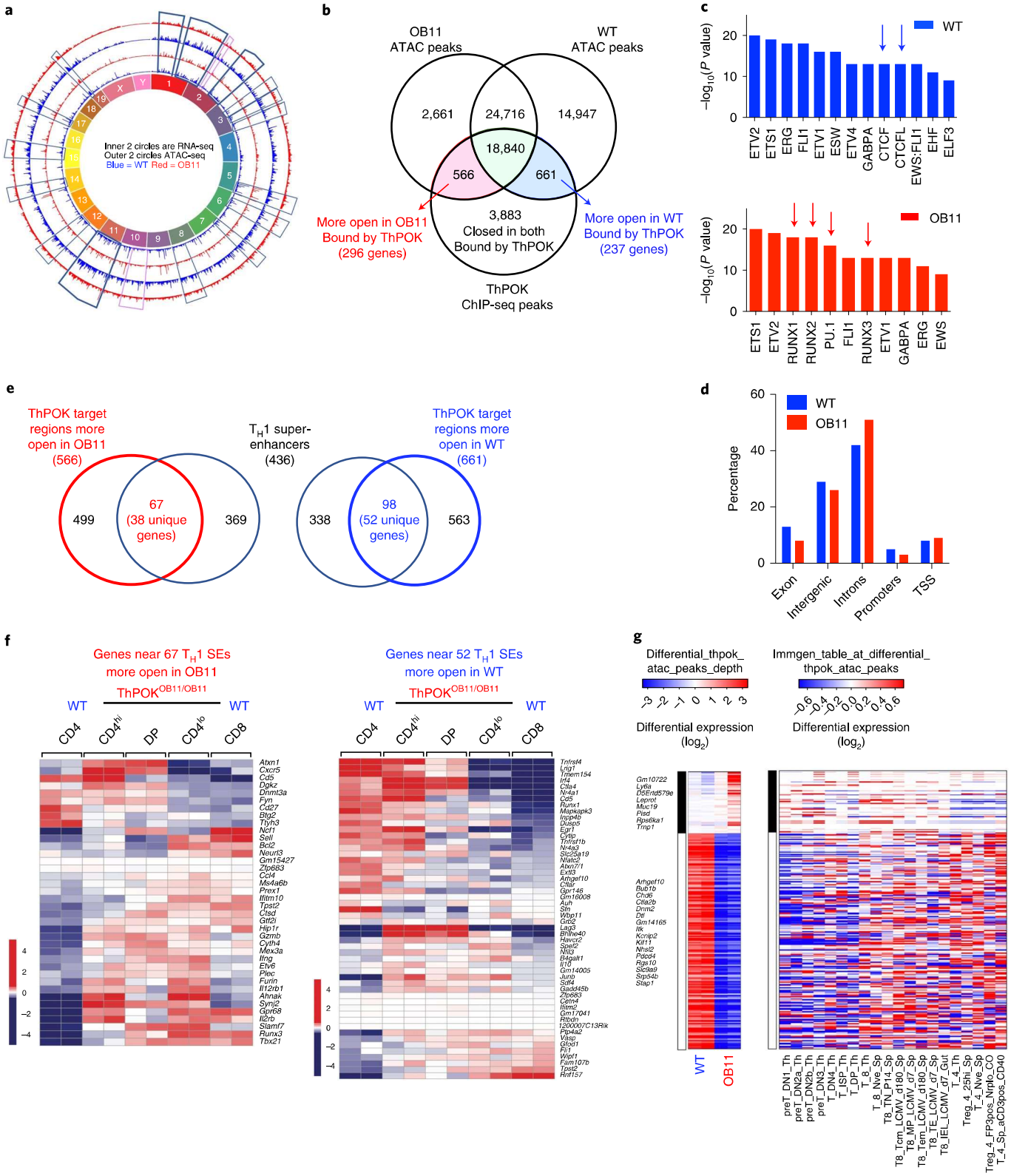
Intestinal T_{reg} cells cannot cross the lamina propria, but can interconvert to Foxp3^{-}CD8αα^{+}CD4^{+} IELs (CD4^{+} IELs) by downmodulating ThPOK in a microbiota-dependent fashion and these accumulate in the intestinal epithelium[11]. As ablation of the ThPOK autoregulatory loop leads to an increase in Triple^{lo} T_{reg} cells, and Triple^{lo} T_{reg} cells have been implicated in protection from colitis[30], we hypothesized that the ThPOK autoregulatory loop might play an important role in regulating T_{reg} cell identity and CD4^{+} IEL formation from T_{reg} cells. Consistent with this notion, the proportion of T_{reg} cells among colonic lymphoplasmacytic lymphomas (cLPLs) was substantially decreased (twofold) in ThPOK$^{\Delta OB11/\Delta OB11}$ versus WT mice, and the representation of Triple^{lo} cells was increased with concomitant decrease in Triple^{hi} cells (Fig. 6d). Concomitantly, CD4^{+} IEL precursors (CD4^{+}CD8α^{+}CD8β^{lo}) were strongly increased in the lamina propria of ThPOK$^{\Delta OB11/\Delta OB11}$ mice compared with WT mice (Fig. 6e). Hence, although CD4^{+} IELs are generated only in the intestinal mucosa in WT mice, in the absence of the ThPOK autoregulatory loop, they are already generated before traveling to the intestinal epithelium.

Recent single-cell transcriptomic analysis defined organ-specific gene expression signatures for T_{reg} cells transiting from the lymph node (LN) to barrier tissues including the colon[31]. Comparison of gene expression signatures of ThPOK$^{\Delta OB11/\Delta OB11}$ naive CD4^{hi} and DP T cells with the reported gene signature of WT colonic T_{reg} cells showed upmodulation of ~70% of the colon-specific T_{reg} cell signature genes, consistent with rewiring of the CD4 T cell gene expression program to favor their conversion into colonic T_{reg} cells (Fig. 7a).

Finally, we analyzed the functional consequences of the altered T_{reg} cell compartment in ThPOK$^{\Delta OB11/\Delta OB11}$ mice for trinitrobenzene sulfonic acid (TNBS)-induced colitis[32]. Although ThPOK$^{\Delta OB11/\Delta OB11}$ and WT mice showed similar initial pathology, ThPOK$^{\Delta OB11/\Delta OB11}$ mice exhibited improved recovery and longevity (Fig. 7b,c and Extended Data Fig. 7e). To directly assess the anti-colitogenic role of ThPOK$^{\Delta OB11/\Delta OB11}$ CD4 T cells, we next employed a T cell transfer model of colitis[33]. Although naive WT CD4 T cells readily induced colitis, both ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{hi} and CD4^{lo} subsets did not, although transferred T cells were present in circulation. Finally, we conducted co-transfers of naive WT (CD45.1^{+}) and naive ThPOK$^{\Delta OB11/\Delta OB11}$ (CD45.2^{+}) T cells into the same hosts (Fig. 7f–h). Colonic inflammation was substantially suppressed by the presence of ThPOK$^{\Delta OB11/\Delta OB11}$ T cells, reflecting the active anti-colitogenic activity of these cells (Fig. 7f–h). In contrast to WT T cells, ThPOK$^{\Delta OB11/\Delta OB11}$ T cells readily converted to CD4^{+}CD8αα T cells and CD8αβ^{+} T cells (Fig. 7k). Although the T_{reg} cell frequency was the same among both WT and ThPOK$^{\Delta OB11/\Delta OB11}$ co-transferred T cells in both intestinal and colonic lamina propria, ThPOK$^{\Delta OB11/\Delta OB11}$ T_{reg} cells in colonic lamina propria of co-transferred mice were remarkably enriched for Triple^{lo} T_{reg} cells compared with T_{reg} cells from mice reconstituted with WT T cells only(Fig. 7j). WT (CD45.1^{+}) T cells from mice co-transferred with ThPOK$^{\Delta OB11/\Delta OB11}$ T cells were also enriched for Triple^{lo} T_{reg} cells (Fig. 7j), suggesting a *trans* effect of ThPOK$^{\Delta OB11/\Delta OB11}$ cells. To demonstrate the T cell-autonomous basis of this phenotype, we performed a similar co-transfer experiment using naive CD4 T cells from Ox40–Cre ThPOK$^{Fl/Fl}$ mice, in which ThPOK is selectively ablated only in mature peripheral T cells. Consistent with our other results, OX40ΔThPOK CD4 T cells strongly suppress colitis induction by naive WT CD4 T cells even at a 5:1 ratio (Fig. 7l–o), accompanied by a striking enhancement of CD4^{+}CD8aa^{+} IEL generation.

Fig. 5 | Loss of anti-silencer function causes genome-wide change in chromatin accessibility. a, Circular plot of mouse chromosomes. The outer two rings represent ATAC-seq peaks. The inner two rings represent RNA-seq peaks. Blue and red rings represent WT CD4 and OB11 CD4^{lo} T cells, respectively. **b**, Venn diagram indicating intersection between ATAC-seq peaks that are selectively open in WT CD4 and OB11 CD4^{lo} T cells and α-ThPOK ChIP–seq peaks, as indicated. **c**, Relative enrichment of transcription factor–binding sites associated with open chromatin region in WT CD4 T cells (top) or ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} T cells (bottom) (Wilcoxon's rank sum test with *P* value adjusted using the Benjamini–Hochberg method). **d**, Relative distribution of DACRs in ThPOK$^{\Delta OB11/\Delta OB1}$ CD4^{lo} (red) versus WT CD4 T cells (blue) with respect to gene organization. TSS, transcriptional start site. **e**, Venn diagram indicating intersection between DACRs that are selectively open in WT CD4 and ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} T cells, and helper T cell–associated SEs, as indicated. **f**, Heatmaps showing relative expression of genes associated with 67 and 98 SEs that are selectively open in ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} versus WT CD4 T cell subsets, respectively (from **e**). **g**, Heatmap showing top 250 regions with highest differential accessibility between WT CD4 and ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} T cells, filtered for genes that are also differentially expressed between these cell types (but not differentially expressed between WT CD4 and WT CD8 T cells) (left panel). Comparison of chromatin accessibility of these 250 regions with a T cell developmental accessibility panorama (Immgen) (right panel).

Transforming growth factor (TGF)-β and retinoic acid are required for downmodulation of ThPOK and gain of CD8α expression by intestinal CD4+ T cells[10], suggesting that TGF-β-dependent binding of SMAD to Sil^{ThPOK} may interrupt the ThPOK autoregulatory loop. Indeed, our bioinformatic and Y1H analysis, as well as public Smad4 ChIP–seq data (GEO accession no. GSM2706519), indicate the presence of SMAD sites within Sil^{ThPOK}. To establish functional relevance of the TGF-β–Smad axis for Sil^{ThPOK} function, we performed anti-Smad ChIP assay on CD4 cells from FLAG–ThPOK homozygous mice activated with anti-CD3/CD28 in the



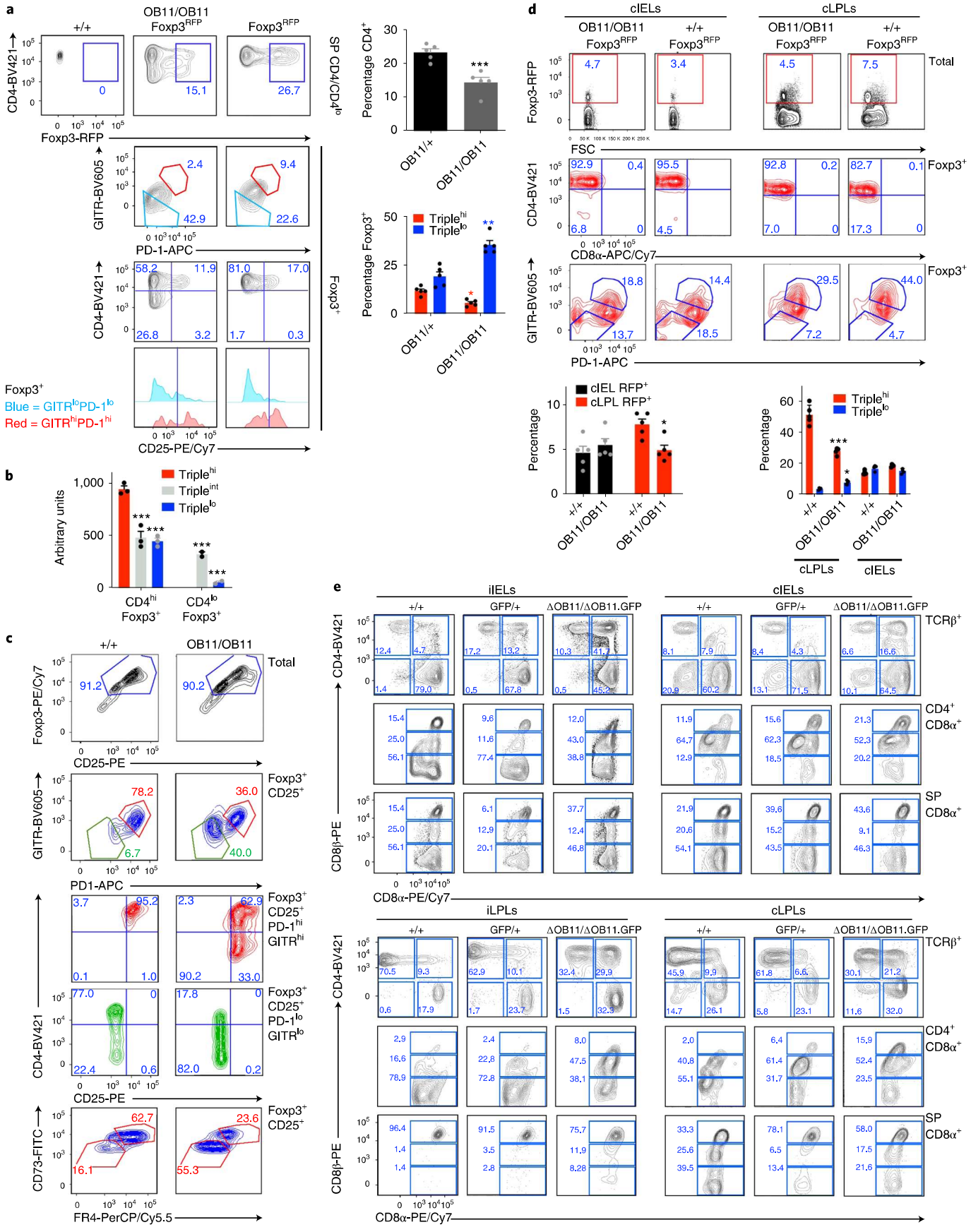

a
OB11/OB11
Foxp3RFP
+/+
Foxp3RFP
CD4-BV421
Foxp3-RFP
GITR-BV605
PD-1-APC
CD4-BV421
CD25-PE/Cy7
Foxp3+
Blue = GITRloPD-1lo
Red = GITRhiPD-1hi
SP CD4+/CD4lo
Percentage CD4+
OB11/+
OB11/OB11
Percentage Foxp3+
Triplehi
Triplelo
b
Arbitrary units
Triplehi
Triplelnt
Triplelo
CD4hi Foxp3+
CD4lo Foxp3+
c
+/+
OB11/OB11
Total
Foxp3-PE/Cy7
CD25-PE
GITR-BV605
PD1-APC
Foxp3+ CD25+
Foxp3+ CD25+ PD-1hi GITRhi
Foxp3+ CD25+ PD-1lo GITRlo
CD4-BV421
CD73-FITC
FR4-PerCP/Cy5.5
d
cIELs
cLPLs
Total
Foxp3-RFP
FSC
Foxp3+
CD8α-APC/Cy7
GITR-BV605
PD-1-APC
cIEL RFP+
cLPL RFP+
Percentage
cLPLs
cIELs
e
iIELs
cIELs
iLPLs
cLPLs
+/+
GFP/+
ΔOB11/ΔOB11.GFP
TCRβ+
CD4+ CD8α+
SP CD8α+
CD4-BV421
CD8β-PE
CD8α-PE/Cy7

Fig. 6 | Ablation of anti-silencer promotes differentiation of Triple^{lo} T_{reg} cell subset. a, Flow cytometric analysis of Foxp3 versus CD4 (top row), PD-1 versus GITR (second row), CD4 versus CD25 (third row), or CD25 expression in gated GITR^{lo}PD-1^{lo} or GITR^{hi}PD-1^{hi} LN (mesenteric) T cells. Bar graphs on the right represent the percentage of T_{reg} cells among the total CD4 lymphocytes (top), or the percentage of Triple^{lo} and Triple^{hi} cells among Foxp3$^+$CD4$^+$ T cells of ThPOK$^{\Delta OB11/+}$ or ThPOK$^{\Delta OB11/\Delta OB11}$ mice ($n=5$ for all groups). **b**, RNA was collected from WT CD4$^+$Foxp3$^+$ T cells subsetted into Triple^{hi}, Triple^{int} and Triple^{lo} subsets, as indicated, and assayed for ThPOK mRNA expression by qPCR. Data were obtained from three technical replicates, containing pooled mRNA from three animals. The experiment was repeated three times. **c**, WT or ThPOK$^{\Delta OB11/\Delta OB11}$ CD4 LN T cells were sorted and pooled from three animals per genotype and subjected to in vitro T_{reg} cell polarization. Representative flow cytometric analysis is shown for Foxp3 versus CD25 (top row), PD-1 versus GITR (second row), CD4 versus CD25 (third and fourth rows), and CD73 versus FR4 expression (total or gated subsets, as indicated). The experiment was repeated three times. **d**, Flow cytometric analysis of Foxp3 versus forward scatter (FSC) (top row), CD4 versus CD8α (second row), and PD-1 versus GITR (bottom row) by total or gated Foxp3$^+$ cLPL or cIEL T cells, as indicated. The bar graphs at the bottom represent the percentage of T_{reg} cells among the total population (left), or the percentage of Triple^{lo} and Triple^{hi} cells among gated T_{reg} cells, of ThPOK$^{+/+}$ or ThPOK$^{\Delta OB11/\Delta OB11}$ mice, as indicated ($n=5$ animals for all groups, examined over two independent experiments). Data were analyzed using one-way ANOVA with Bonferroni's correction (lower panels). Asterisks represent the *P* value. *$P=0.0136$ (left panel); *$P=0.0109$, ***$P=0.0001$ (right panel). **e**, Flow cytometric analysis of CD4 versus CD8α (top row), or CD8α versus CD8β (second and third rows) expression, for either total TCRβ$^+$ cells or indicated gated subset of freshly isolated intestinal (i)IEL, cIEL, iLPL or cLPL populations, as indicated. The data represent eight animals per genotype, examined over three independent experiments. The bar graphs in **a**, **b** and **d** are presented as mean ± s.e.m. A *P* value <0.05 was considered to be significant. *$P<0.05$, **$P<0.01$, ***$P<0.001$.

presence or absence of TGF-β, revealing strong TGF-β-dependent Smad4 binding to Sil^{ThPOK} accompanied by substantial loss of ThPOK binding (Fig. 7o).

Altogether, our data suggest that under physiological conditions the ThPOK autoregulatory loop inhibits untimely rewiring of circulating naive CD4 T cells toward highly suppressive barrier-specific T_{reg} cells.

Discussion

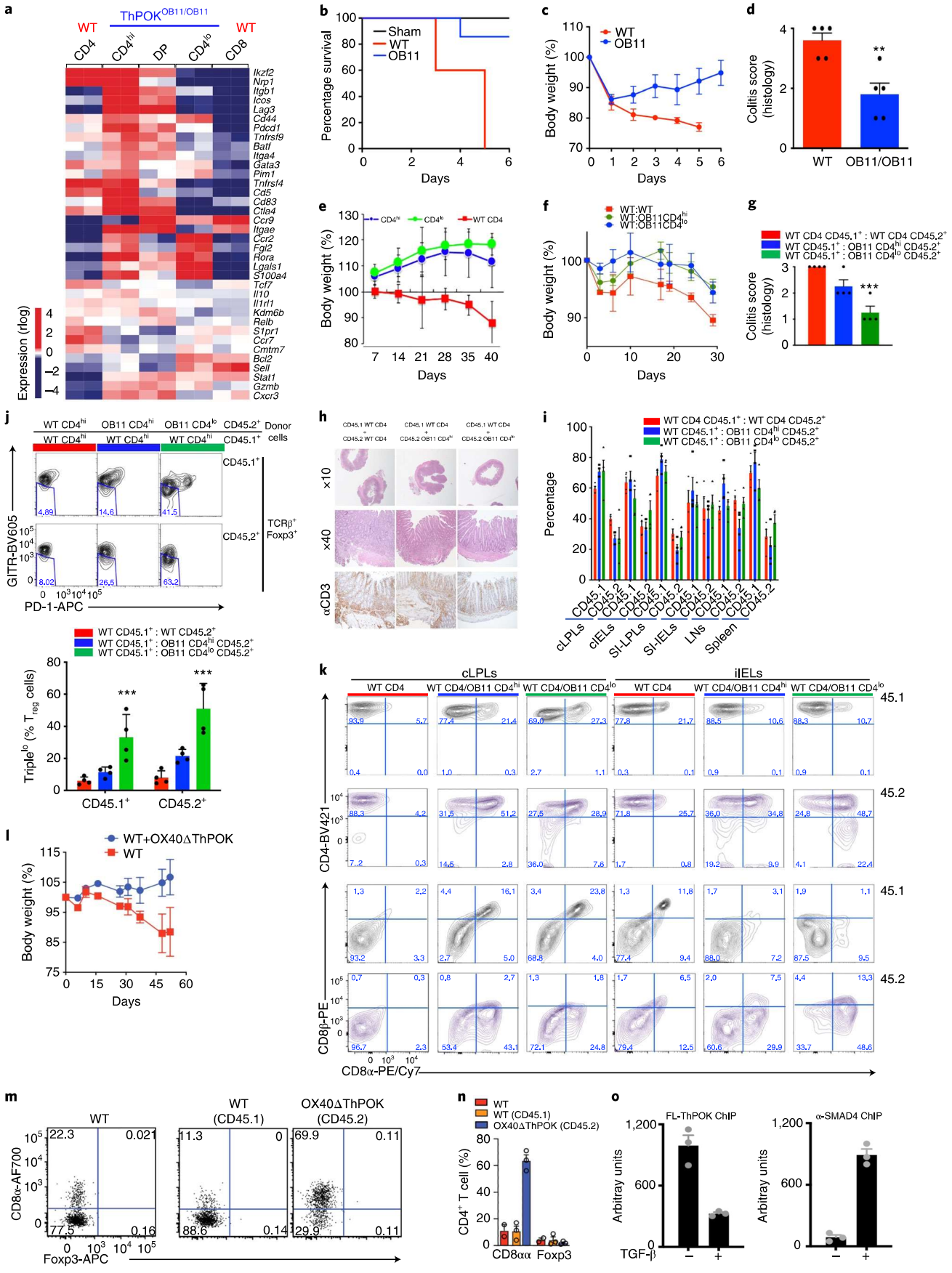
In the present study, we identify a new regulatory switch, which is an anti-silencer located within the Sil^{ThPOK} element, necessary for maintenance of helper T cell (Th cell) identity. The anti-silencer mediates recruitment of ThPOK, thereby driving a positive feedback loop that stabilizes *ThPOK* transcription. Disruption of this loop leads to drastic change in genome-wide chromosomal accessibility, perturbed gene expression and adoption of a metastable state in CD4 T cells. Using in vitro T_{reg} cell culture and in vivo colitis models, we further demonstrate a crucial role of this switch for tissue-specific T_{reg} cell differentiation and maintenance of T_{reg} cell integrity in the intestinal microenvironment to prevent premature CD4 IEL generation.

The metastable state induced by anti-silencer ablation is characterized by three distinct and mutually interconvertible cellular phenotypes, CD4^{hi}, CD4^{lo} and DP, each with divergent gene expression profiles. Permanent escape from the metastable condition is possible by conversion to the CD8 state, because CD4^{hi} and CD4^{lo} cells can give rise to CD8 T cells, but not vice versa. It will be interesting to determine in future whether interconversion between subsets and escape to the CD8 fate are determined stochastically or by external stimuli.

Stochastic cell fate decisions often involve bistable switching of developmental genes between periods of transcriptional bursts and transcriptional inactivity[34–37]. Given that a substantial fraction of adoptively transferred ThPOK$^{\Delta OB11/\Delta OB11}$ T cells express GFP even many weeks after transfer, despite a lack of steady-state ThPOK mRNA, we propose that ablation of the anti-silencer element restricts *ThPOK* transcription to short transcriptional bursts. This could be directly tested in future using live cell RNA-imaging approaches.

Mechanistically, ThPOK acts to maintain lineage stability by: (1) controlling SE accessibility and activity; (2) regulating chromatin accessibility by directly targeting lamina-associated genes (for example, *Lmnb1*, *Lmna*, *Lbr*) and mediators of transcriptional silencing (for example, *Satb1*, *Rreb1*, *Hdac1/7/10*, *Tle1*, *Bcor*, *Etv3*, *Dnmt3a*, *Kdm1*), consistent with drastic reduction in accessible chromatin in ThPOK$^{\Delta OB11/\Delta OB11}$ T cells; (3) functionally interacting with Ctcf in limiting spread of closed chromatin, consistent with genome-wide enrichment of Ctcf motifs near ThPOK-binding sites; and (4) masking Runx- and PU.1-binding sites, thereby preventing multilineage gene derepression. Furthermore, the chromatin accessibility pattern of ThPOK$^{\Delta OB11/\Delta OB11}$ T cells strikingly resembles that of the ETP, suggesting that the ThPOK autoregulatory loop plays a key role in preventing de-dedifferentiation to the ETP. Indeed, we observe derepression of ETP–DN2 thymocyte (for example, *Kit*, *Meis1*, *Pgk1*, *Flt3*, *Id2*)[18] and granulocyte-biased DN thymocyte gene programs (for example, *Fcgr1*, *MPO*, *Lyz2*)[38,39].

Fig. 7 | Anti-silencer-deficient CD4 T cells display anti-colitogenic activity. a, Heatmap of colonic T_{reg} cell signature gene expression for indicated ThPOK$^{\Delta OB11/\Delta OB11}$ and WT T cell subsets. Red and blue indicate increased or decreased expression, respectively. **b–d**, Survival plot (**b**), weight plot (**c**) and colitis severity score (**d**) after TNBS treatment of indicated mouse strains ($n=10$ for all groups). Histopathology was performed on five animals per genotype at the end of the experiment. Asterisks represent the *P* value. **$P=0.0023$. **e**, Weight plot of host RAG$^{-/-}$ mice after transfer of sorted naive ThPOK$^{\Delta OB11/\Delta OB11}$ or WT T cell subsets, as indicated. Donor cells were sorted from pooled LN samples of three CD45.1$^+$ WT or CD45.2$^+$ ThPOK$^{\Delta OB11.GFP/\Delta OB11}$ mice ($n=5$ host animals reconstituted with each cell type). The experiment was performed three times. **f–j**, Similar analysis to that in **e**, except all animals received a co-transfer of WT CD4 T cells ($n=4$ host animals reconstituted with each cell combination): weight plot (**f**), colitis severity score at day 30 (**g**), histopathological analysis of colon at day 30 (**h**), proportions of CD45.1$^+$ and CD45.2$^+$ cells for indicated cell subset at day 30 (**i**), and flow cytometric analysis of PD-1 versus GITR expression for indicated subsets at day 30 (**j**). SI, small intestine. The experiment was repeated three times. The bar graph indicates the proportions of Triple^{lo} cells among Foxp3$^+$ T_{reg} cells ($n=4$ animals). **k**, Flow cytometric analysis of CD4, CD8α and CD8β expression for TCRβ$^+$ cells from different gut locations. The colitis score (**g**) was analyzed by one-way ANOVA with Bonferroni's correction. The asterisks represent the *P* value. ***$P=0.0004$. **l–n**, Adoptive transfer of sorted OX40ΔThPOK or WT CD4 T cells into RAG$^{-/-}$ hosts: weight plot (**l**), flow cytometric analysis of CD8α versus Foxp3 expression (**m**), and proportions of CD8αα and Foxp3$^+$ cells among CD4$^+$ IELs (**n**). For **l–n**, $n=3$ (WT$^+$Ox40 co-transfer) or 2 (WT only), and the experiment was repeated twice. **o**, CD4 T cells were sorted from pooled LN samples of three FL-ThPOK knockin mice, followed by anti-CD3/CD28 stimulation, in the presence/absence of 2 ng ml^{-1} TGF-β for 48 h. SMAD4 and ThPOK binding to the silencer were determined by ChIP assay using anti-SMAD4 antibody and anti-FLAG antibody ($n=3$ technical replicates). The experiment was repeated three times. The data in **c–g**, **i**, **j**, **l**, **n** and **o** represent the mean ± s.e.m. A *P* value <0.05 was considered to be significant. *$P<0.05$, **$P<0.01$, ***$P<0.001$.


a
ThPOK OB11/OB11
WT
CD4
CD4hi
DP
CD4lo
CD8
WT
Ikzf2
Nrp1
Itgb1
Icos
Lag3
Cd44
Pdcd1
Tnfrsf9
Batf
Itga4
Gata3
Pim1
Tnfrsf4
Cd5
Cd83
Ctla4
Ccr9
Itgae
Ccr2
Fgl2
Rora
Lgals1
S100a4
Tcf7
Il10
Il1r1
Kdm6b
Relb
S1pr1
Ccr7
Cmtm7
Bcl2
Sell
Stat1
Gzmb
Cxcr3
Expression (rlog)
b
Sham
WT
OB11
Percentage survival
Days
c
WT
OB11
Body weight (%)
Days
d
Colitis score (histology)
WT
OB11/OB11
e
CD4hi
CD4lo
WT CD4
Body weight (%)
Days
f
WT:WT
WT:OB11CD4hi
WT:OB11CD4lo
Body weight (%)
Days
g
WT CD4 CD45.1+ : WT CD4 CD45.2+
WT CD45.1+ : OB11 CD4hi CD45.2+
WT CD45.1+ : OB11 CD4lo CD45.2+
Colitis score (histology)
j
WT CD4hi
OB11 CD4hi
OB11 CD4lo
CD45.2+ Donor cells
WT CD4hi
WT CD4hi
WT CD4hi
CD45.1+
CD45.1+
CD45.2+
TCRβ+ Foxp3+
GITR-BV605
PD-1-APC
WT CD45.1+ : WT CD45.2+
WT CD45.1+ : OB11 CD4hi CD45.2+
WT CD45.1+ : OB11 CD4lo CD45.2+
Triple lo (% Treg cells)
CD45.1+
CD45.2+
h
CD45.1 WT CD4 + CD45.2 WT CD4
CD45.1 WT CD4 + CD45.2 OB11 CD4hi
CD45.1 WT CD4 + CD45.2 OB11 CD4lo
x10
x40
αCD3
i
WT CD4 CD45.1+ : WT CD4 CD45.2+
WT CD45.1+ : OB11 CD4hi CD45.2+
WT CD45.1+ : OB11 CD4lo CD45.2+
Percentage
CD45.1
CD45.2
cLPLs
cIELs
SI-LPLs
SI-IELs
LNs
Spleen
k
cLPLs
iILEs
WT CD4
WT CD4/OB11 CD4hi
WT CD4/OB11 CD4lo
WT CD4
WT CD4/OB11 CD4hi
WT CD4/OB11 CD4lo
45.1
45.2
45.1
45.2
CD4-BV421
CD8β-PE
CD8α-PE/Cy7
l
WT+OX40ΔThPOK
WT
Body weight (%)
Days
m
WT
WT (CD45.1)
OX40ΔThPOK (CD45.2)
CD8α-AF700
Foxp3-APC
22.3
0.021
77.5
0.16
11.3
0
88.6
0.14
69.9
0.11
29.9
0.11
n
WT
WT (CD45.1)
OX40ΔThPOK (CD45.2)
CD4+ T cell (%)
CD8αα
Foxp3
o
FL-ThPOK ChIP
α-SMAD4 ChIP
Arbitray units
TGF-β
–
+

The ThPOK-mediated autoregulatory loop is further required to maintain the quiescent state of naive CD4 T cells, as evidenced by derepression of effector cytokine expression, and dispensability of co-stimulation for activation and proliferation of naive ThPOK$^{\Delta OB11/\Delta OB11}$ T cells. TCR 'tickling' by self-MHC[40] and active TGF-β signaling[41] have been implicated in maintaining the quiescent state in naive T cells. Our RNA-seq data suggest that the absence of the ThPOK autoregulatory loop leads to downmodulation of *TGFβR* as well as genes involved in restraining TCR signaling, indicating an important role for the ThPOK-mediated autoregulatory loop in preventing T cell hyperactivation and autoimmunity. ThPOK$^{\Delta OB11/\Delta OB11}$ T cell subsets also exhibit upmodulation of genes related to cytotoxic CD4 T cells arising in influenza-infected lungs and tumor sites (for example, *Anxa1*, *Id2*, *Lat2*, *Nkg7* and *Eomes*)[42–47], suggesting a key role of the ThPOK autoregulatory loop in limiting cytotoxic CD4 T cell generation in infection and cancer.

On arrival in the gut, T_{reg} cells convert to CD4 IELs by downmodulating ThPOK in a microbiota-dependent fashion[11], which is a crucial step for maintenance of gut immune homeostasis. In the present study, we show that that abrogation of the ThPOK autoregulatory loop promotes selective differentiation of naive CD4 T cells into anti-colitogenic Triple^{lo} T_{reg} cells, and their conversion to CD4$^+$ IELs in the gut. Recent findings from our lab suggest that TCR signaling plays a major role in T_{reg} cell-to-IEL conversion[48]. Future experiments will address how modulation of ThPOK expression affects TCR signaling to ensure selective T_{reg} cell repertoire-specific conversion to IELs.

We conclude that perturbation of the ThPOK autoregulatory feedback loop is a physiological mechanism for the post-thymic generation of T_{reg} cells and CD8aa IELs in the colon, and potentially other mucosal barriers. This controls a delicate balance between protective immunity and absence of tissue-destructive autoimmune responses. ThPOK misregulation via interruption of this autoregulatory loop may be responsible for other pathological conditions, for example, development of autoaggressive pathogenic T cells in a model of type 1 diabetes mellitus characterized by an unusual diabetogenic CD4^{lo} T cell subset[49], and the appearance of DP CD4$^+$/CD8$^+$ T cells in HIV-infected individuals[50].

Online content

Any methods, additional references, Nature Research reporting summaries, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-021-00980-8.

Received: 21 February 2020; Accepted: 21 June 2021;
Published online: 26 July 2021

References

1. Dave, V. P., Allman, D., Keefe, R., Hardy, R. R. & Kappes, D. J. HD mice: a novel mouse mutant with a specific defect in the generation of CD4$^+$ T cells. *Proc. Natl Acad. Sci. USA* **95**, 8187–8192 (1998).
2. He, X. et al. The zinc finger transcription factor Th-POK regulates CD4 versus CD8 T-cell lineage commitment. *Nature* **433**, 826–833 (2005).
3. Keefe, R., Dave, V., Allman, D., Wiest, D. & Kappes, D. J. Regulation of lineage commitment distinct from positive selection. *Science* **286**, 1149–1153 (1999).
4. Wang, L. et al. The zinc finger transcription factor Zbtb7b represses CD8-lineage gene expression in peripheral CD4$^+$ T cells. *Immunity* **29**, 876–887 (2008).
5. He, X. et al. CD4-CD8 lineage commitment is regulated by a silencer element at the ThPOK transcription-factor locus. *Immunity* **28**, 346–358 (2008).
6. Setoguchi, R. et al. Repression of the transcription factor Th-POK by Runx complexes in cytotoxic T cell development. *Science* **319**, 822–825 (2008).
7. Mookerjee-Basu, J. et al. Functional conservation of a developmental switch in mammals since the Jurassic age. *Mol. Biol. Evol.* **36**, 39–53 (2019).
8. Vacchio, M. S. et al. A ThPOK-LRF transcriptional node maintains the integrity and effector potential of post-thymic CD4$^+$ T cells. *Nat. Immunol.* **15**, 947–956 (2014).
9. Ciucci, T. et al. The emergence and functional fitness of memory CD4$^+$ T cells require the transcription factor Thpok. *Immunity* **50**, 91–105.e4 (2019).
10. Mucida, D. et al. Transcriptional reprogramming of mature CD4$^+$ helper T cells generates distinct MHC class II–restricted cytotoxic T lymphocytes. *Nat. Immunol.* **14**, 281–289 (2013).
11. Sujino, T. et al. Tissue adaptation of regulatory and intraepithelial CD4$^+$ T cells controls gut inflammation. *Science* **352**, 1581–1586 (2016).
12. Reis, B. S., Rogoz, A., Costa-Pinto, F. A., Taniuchi, I. & Mucida, D. Mutual expression of the transcription factors Runx3 and ThPOK regulates intestinal CD4$^+$ T cell immunity. *Nat. Immunol.* **14**, 271–280 (2013).
13. Fuxman Bass, J. I. et al. Human gene-centered transcription factor networks for enhancers and disease variants. *Cell* **161**, 661–673 (2015).
14. Reece-Hoyes, J. S. et al. Yeast one-hybrid assays for gene-centered human gene regulatory network mapping. *Nat. Methods* **8**, 1050–1052 (2011).
15. Reece-Hoyes, J. S. et al. Enhanced yeast one-hybrid assays for high-throughput gene-centered regulatory network mapping. *Nat. Methods* **8**, 1059–1064 (2011).
16. Muroi, S. et al. Cascading suppression of transcriptional silencers by ThPOK seals helper T cell fate. *Nat. Immunol.* **9**, 1113–1121 (2008).
17. Boursalian, T. E., Golob, J., Soper, D. M., Cooper, C. J. & Fink, P. J. Continued maturation of thymic emigrants in the periphery. *Nat. Immunol.* **5**, 418–425 (2004).
18. Zhou, W. et al. Single-cell analysis reveals regulatory gene expression dynamics leading to lineage commitment in early T cell development. *Cell Syst.* **9**, 321–337.e9 (2019).
19. Vacchio, M. S. et al. A Thpok-directed transcriptional circuitry promotes Bcl6 and Maf expression to orchestrate T follicular helper differentiation. *Immunity* **51**, 465–478.e6 (2019).
20. Istaces, N. et al. EOMES interacts with RUNX3 and BRG1 to promote innate memory cell formation through epigenetic reprogramming. *Nat. Commun.* **10**, 3306 (2019).
21. Zhang, J. A., Mortazavi, A., Williams, B. A., Wold, B. J. & Rothenberg, E. V. Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. *Cell* **149**, 467–482 (2012).
22. Zheng, S., Papalexi, E., Butler, A., Stephenson, W. & Satija, R. Molecular transitions in early progenitors during human cord blood hematopoiesis. *Mol. Syst. Biol.* **14**, e8041 (2018).
23. Diao, H. & Pipkin, M. E. Stability and flexibility in chromatin structure and transcription underlies memory CD8 T-cell differentiation. *F1000Res.* **8**, 1278 (2019).
24. Chen, H. et al. Single-cell trajectories reconstruction, exploration and mapping of omics data with STREAM. *Nat. Commun.* **10**, 1903 (2019).
25. Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **53**, 307–319 (2013).
26. Vahedi, G. et al. Super-enhancers delineate disease-associated regulatory nodes in T cells. *Nature* **520**, 558–562 (2015).
27. Khan, A. & Zhang, X. dbSUPER: a database of super-enhancers in mouse and human genome. *Nucleic Acids Res.* **44**, D164–D171 (2016).
28. Snook, J. P., Kim, C. & Williams, M. A. TCR signal strength controls the differentiation of CD4$^+$ effector and memory T cells. *Sci. Immunol.* **3**, eaas9103 (2018).
29. Carpenter, A. C. et al. Control of regulatory T cell differentiation by the transcription factors Thpok and LRF. *J. Immunol.* **199**, 1716–1728 (2017).
30. Wyss, L. et al. Affinity for self antigen selects T_{reg} cells with distinct functional properties. *Nat. Immunol.* **17**, 1093–1101 (2016).
31. Miragaia, R. J. et al. Single-cell transcriptomics of regulatory T cells reveals trajectories of tissue adaptation. *Immunity* **50**, 493–504.e7 (2019).
32. Neurath, M., Fuss, I. & Strober, W. TNBS-colitis. *Int. Rev. Immunol.* **19**, 51–62 (2000).
33. Powrie, F. et al. Inhibition of Th1 responses prevents inflammatory bowel disease in *scid* mice reconstituted with CD45RB^{hi} CD4$^+$ T cells. *Immunity* **1**, 553–562 (1994).
34. Larsson, A. J. M. et al. Genomic encoding of transcriptional burst kinetics. *Nature* **565**, 251–254 (2019).
35. Suter, D. M. et al. Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**, 472–474 (2011).
36. Chubb, J. R., Trcek, T., Shenoy, S. M. & Singer, R. H. Transcriptional pulsing of a developmental gene. *Curr. Biol.* **16**, 1018–1025 (2006).
37. Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309 (2006).
38. De Obaldia, M. E., Bell, J. J. & Bhandoola, A. Early T-cell progenitors are the major granulocyte precursors in the adult mouse thymus. *Blood* **121**, 64–71 (2013).
39. Rothenberg, E. V., Hosokawa, H. & Ungerback, J. Mechanisms of action of hematopoietic transcription factor PU.1 in initiation of T-cell development. *Front. Immunol.* **10**, 228 (2019).
40. Takada, K. & Jameson, S. C. Naive T cell homeostasis: from awareness of space to a sense of place. *Nat. Rev. Immunol.* **9**, 823–832 (2009).

41. Tu, E. et al. T cell receptor-regulated TGF-β type I receptor expression determines T cell quiescence and activation. *Immunity* **48**, 745–759.e6 (2018).
42. Marshall, N. B. et al. NKG2C/E marks the unique cytotoxic CD4 T cell subset, ThCTL, generated by influenza infection. *J. Immunol.* **198**, 1142–1155 (2017).
43. Pearce, E. L. et al. Control of effector CD8$^+$ T cell function by the transcription factor eomesodermin. *Science* **302**, 1041–1043 (2003).
44. Curran, M. A. et al. Systemic 4-1BB activation induces a novel T cell phenotype driven by high expression of eomesodermin. *J. Exp. Med.* **210**, 743–755 (2013).
45. Hirschhorn-Cymerman, D. et al. Induction of tumoricidal function in CD4$^+$ T cells is associated with concomitant memory and terminally differentiated phenotype. *J. Exp. Med.* **209**, 2113–2126 (2012).
46. Qui, H. Z. et al. CD134 plus CD137 dual costimulation induces eomesodermin in CD4 T cells to program cytotoxic Th1 differentiation. *J. Immunol.* **187**, 3555–3564 (2011).
47. Wagner, D. H. Jr. et al. Expression of CD40 identifies a unique pathogenic T cell population in type 1 diabetes. *Proc. Natl Acad. Sci. USA* **99**, 3782–3787 (2002).
48. Bilate, A. N. et al. T cell receptor is required for differentiation, but not maintenance, of intestinal CD4$^+$ intraepithelial lymphocytes. *Immunity* **53**, 1001–1014.e20 (2020).
49. Waid, D. M., Vaitaitis, G. M. & Wagner, D. H. Jr. Peripheral CD4^{lo}CD40$^+$ auto-aggressive T cell expansion during insulin-dependent diabetes mellitus. *Eur. J. Immunol.* **34**, 1488–1497 (2004).
50. Flamand, L. et al. Activation of CD8$^+$ T lymphocytes through the T cell receptor turns on CD4 gene expression: implications for HIV pathogenesis. *Proc. Natl Acad. Sci. USA* **95**, 3111–3116 (1998).

Methods

Mice. All experimentation involving animals was approved by the Institutional Animal Care and Use Committee of Fox Chase Cancer Center, or the Rockefeller University. RAG1$^{-/-}$, β_2m$^{-/-}$, Foxp3$^{IRES\text{-}mRFP}$, B6.SJL-Ptprc^{a}Pepc^{b}, Zbtb7b^{Fl} mice lines were procured from Jackson Laboratory. All other mice lines described in the present study have been generated by the FCCC Transgenic Facility on the C57BL/6 strain of *Mus musculus*. Animals used in all experiments were aged 6–12 weeks, and males and females were used in equal proportions (no difference was noted between males and females in any experiment). Animal care was in accordance with the National Institutes of Health (NIH) guidelines. Mice were maintained on a 12 h light:dark cycle, at 24 °C and 50% humidity.

ZFN-mediated gene targeting in mouse embryos. Site-specific mutagenesis was carried out[7]. Briefly, a pair of zinc finger nuclease (ZFN) RNAs that recognize a specific target site was designed and generated by Millipore-Sigma (Genome Editing division). The ZFN target sequence near Sil^{ThPOK} is ACCGCTACCCTAACCcataaCTGGAAGGGGTTTAG (capital letters denote nucleotides actually bound by right and left ZFN proteins). The ZFN upstream of the first coding exon of ThPOK, where the GFP reporter complementary DNA was inserted, is CTGAACCGCAGTTCCCTTgtgcatGACATGAAAGGTGGTTTGG.

The mRNAs encoding the two site-specific ZFNs (50 ng µl^{-1}) were introduced together with double-stranded DNA-targeting constructs bearing the desired mutations/deletions into one-cell mouse oocytes by pronuclear injection, and injected oocytes were transferred to a pseudopregnant surrogate mother. Targeting constructs contained 1.5- and 0.8-kb arms of homology on either side of the desired mutations/deletion. Positive founder pups were identified based on a reduced size of PCR product using primers F1 (5′-ATCCCTACGAAGAAGCCTCT-3′) and R1 (5′-AGGCTTTCCATGTCAGGGTC-3′), and mated to C57BL/6 mice to generate stable heritable knockin lines.

Antibodies. All fluorescently labeled antibodies used were obtained from commercial sources (eBioscience, BioLegend or BD, as indicated) and were validated by the supplier: Thy1-FITC (clone 30-H12; BioLegend, catalog no. 105306, lot no. B224687), Thy1-APC (30-H12; BioLegend, catalog no. 140312, lot no. B274408), TCRβ-APC (clone H57-597; BioLegend, catalog no. 109212, lot no. B273860), TCRβ-PE (clone H57-597; eBioscience, catalog no. 12-5961-83, lot no. E01951-1631), TCRβ-FITC (clone H57-597; BD, catalog no. 553171, lot no. 01934), CD4-BV421 (clone RM4-5; BioLegend, catalog no. 100544, lot no. B293278), CD4-PE/Cy5 (clone RM4-5; BioLegend, catalog no. 100514, lot no. B231410), CD25-PE (clone PC61; BD, catalog no. 553866, lot no. 5047982), CD25-PE/Cy7 (clone PC61; BD, catalog no. 552880, lot no. 36592), CD8a-APC/Cy7 (clone 53-6.7; BioLegend, catalog no. 100714, lot no. B276265), CD8a-FITC (clone 53-6.7; BioLegend, catalog no. 100706, lot no. B168591), CD8a-PE/Cy7 (clone 53-6.7; BioLegend, catalog no. 100722), CD8a-AF700 (clone 53-6.7; BioLegend, catalog no. 100730), CD8b-PE (clone 53-5.8; BD, catalog no. 553041, lot no. 3193966), CD69-PE/Cy7 (clone H1.2F3; eBioscience, catalog no. 25-0691-82, lot no. E07583-1635), CD62L-APC (clone MEL-14; BD, catalog no. 553152, lot no. 7075853), CD45.1-BV605 (clone A20; BioLegend, catalog no. 110738), CD45.2-Alexa Fluor-647 (clone 104; BioLegend, catalog no. 109818, lot no. B181101), CD73-FITC (clone TY/11.8; BioLegend, catalog no. 127219), FR4-PerCP/Cy5.5 (clone 12A5; BioLegend, catalog no. 125018), Foxp3-PE/Cy7 (clone FJK-16S; eBioscience, catalog no. 25-5773-82, lot no. E07638-1634), Foxp3-APC (clone FJK-16S; eBioscience, catalog no. 17-5773-82), CD44-BV605 (clone IM7; BioLegend, catalog no. 103047, lot no. B288308), GITR-Biotin (clone DTA-1; BioLegend, catalog no. 126305, lot no. B254051), PD1-APC (clone 29F.1A12; BioLegend, catalog no. 135210), streptavidin-BV605 (BioLegend, catalog no. 405229, lot no. B242305). Anti-SMAD4 was procured from Abcam (clone EP618Y, catalog no. ab215968). Unlabeled anti-CD28 (clone 37.51; BioLegend, catalog no. 102131), TCRβ (clone H57-597; BioLegend, catalog no. 102214, lot no. B218495), CD3e (clone 145-2C11; BioLegend, catalog no. 100340, lot no. B260927), IL-4 (clone 11B11; catalog no. 504115, lot no. B213842) and anti-interferon (IFN)-γ (clone XMG1.2; catalog no. 505827, lot no. B218186) were from BioLegend. Anti-SMAD4 was procured from Abcam.

EMSA. Nuclear extracts were prepared from human embryonic kidney 293T cells (obtained from American Type Culture Collection, catalog no. CRL-1573) transfected with Flag-tagged murine ThPOK constructs cloned into the pcDNA3 expression vector. Negative controls included nuclear extracts from cells transfected with vector alone. ThPOK expression was verified by immunoblot analysis (data not shown) and used as a ThPOK protein source for binding assay. DNA-binding probes were generated by annealing synthetic double-stranded oligonucleotides corresponding to the OB11 region and end-labeling with polynucleotide kinase and digoxigenin-11-ddUTP using an EMSA Kit (Sigma-Aldrich). The anti-Flag antibody (Sigma-Aldrich) was used for 'supershifting' ThPOK protein–DNA complexes.

Quantitative RT–PCR. The qPCR was carried out according to the probe-based method and analyzed by the comparative C_t method (compared with Rps6).

Th cell polarization. Spleen and LN cells were flow cytometrically sorted for CD44^{-}CD69^{-}CD62l^{hi} naive CD4^{+} T cells. Cells were activated with 5 µg ml^{-1} of plate-bound anti-CD3/CD28 antibodies (BioLegend) with IL-2 (25 IU ml^{-1}). For Th0 conditions, anti-IL-4 (11B11; 20 µg ml^{-1}) and anti-IFN-γ (20 µg ml^{-1}) were added; for T_H1 conditions, anti-IL-4 (11B11; 20 µg ml^{-1}) and IL-12 (10 ng ml^{-1}) were added; for T_H2 conditions, anti-IFN-γ (20 µg ml^{-1}) and IL-4 (20 ng ml^{-1}) were added; for T_H9 conditions, anti-IFN-γ (20 µg ml^{-1}), IL-4 (20 ng ml^{-1}) and TGF-β (2 ng ml^{-1}) were added; for T_H17 conditions, IL-6 (10 ng ml^{-1}), TGF-β (2 ng ml^{-1}), anti-IFN-γ (20 µg ml^{-1}) and IL-4 (20 ng ml^{-1}) were added. For T_{reg} cell polarization, TGF-β (10 ng ml^{-1}), IL-2 (100 U ml^{-1}), anti-IFN-γ (20 µg ml^{-1}) and IL-4 (20 ng ml^{-1}) were added. Mouse IL-12 and IL-4 were from PeproTech, TGF-β and IL-6 were from R&D Systems, human IL-2 was from Roche, and neutralizing antibodies to mouse IFN-γ (XMG1.2) and IL-4 (11B11) were from BD Bioscience. Cells were cultured for 5 d in RPMI medium 1640 containing 10 mM 4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid, pH 7.0, 10% (v:v) fetal bovine serum, 2 mM L-glutamine, antibiotics (complete medium) and 50 µM 2-mercaptoethanol.

Proliferation assays. Flow cytometrically sorted T cell subsets were incubated with 5- (and 6-)carboxyfluorescein diacetate succinimidyl ester (CFSE, Dojindo, 5 µM) in phosphate-buffered saline (37 °C, 15 min). CFSE-labeled cells were cultured in plate-bound anti-TCRβ or anti-CD3/CD28 antibodies for 1–4 d before flow cytometric analysis. Generation analysis was performed using FlowJo software.

Intracellular staining. Cells were fixed in 100 µl of Cytofix/Cytoperm solution for 30 min at 4 °C, washed twice in perm/wash solution, pelleted by centrifugation and resuspended in 100 µl of perm/wash solution with or without (fluorescence minus one control) fluorochrome-conjugated antibody at room temperature, using the BD Fixation/Permeabilization Solution Kit (catalog no. 554714), according to the manufacturer's instructions.

ChIP. Thymocyte or naive (SP CD62l^{hi}CD44^{lo}) peripheral T cell subsets (1×10^6) were purified by flow cytometry from compound heterozygous ThPOK$^{FH/\Delta OB11}$ mice. Chromatin crosslinking and immunoprecipitation were performed using the iDeal ChIP–seq kit for Transcription Factors (Diagenode) according to the manufacturer's protocol. The anti-Flag and anti Runx3 (Abcam, catalog no. 11905) antibody was used for immunoprecipitation and purified DNA sequences were analyzed by qPCR using the WT allele (forward GGCGCGCAGTTATAAATAG, reverse CCCCTACCGCGACCGCCCAA) and OB11 allele-specific (forward CAGT TATAAATAGAGGCTT, reverse CTGCCTCCGCTTCCCTCGAA) primers.

ATAC-seq. Flow cytometrically sorted naive CD4 cells (5×10^4) from WT or CD4^{lo} cells from OB11 mice were pelleted at 500*g* for 5 min at 4 °C, and whole-cell pellets resuspended in 50 µl of ATAC-RSB buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Igepal CA-630), 0.1% Tween-20 (Sigma-Aldrich) and 0.01% digitonin, and kept on ice for 3 min. Then, 1 ml of cold ATAC-RSB + 0.1% Tween-20 was added, and samples centrifuged at 500*g* for 10 min at 4 °C (fixed-angle rotor) to obtain a nuclear pellet. The transposase reaction of open chromatin was achieved by resuspending free nuclei in tagmentation mix (22.5 µl of tagment DNA buffer, 2.5 µl of tagment DNA enzyme, 25 µl of H_2O; Illumina, catalog no. FC-121-1030) and incubating at 37 °C for 30 min. Purification of DNA was performed with a Diapure kit (Diagenode), according to the manufacturer's protocol. Barcoding and amplification were performed using Nextera Index Kit (Illumina, catalog no. FC-121-1011)[51]. Amplified ATAC-seq libraries were purified using Gene-Read Size Selection Kit (QIAGEN, catalog no. 180514) according to the manufacturer's protocol. The quality and quantity of the final ATAC-seq libraries were assessed with the High Sensitivity DNA kit (Agilent, catalog no. 5067-4626) run on an Agilent 2100 Bioanalyzer. ATAC-seq libraries were sequenced using Illumina 125-bp paired-end sequencing on a HiSeq2500 platform, generating between 38 and 43 million reads per condition per biological replicate.

After quality control by FastQC (https://www.bioinformatics.babraham.ac.uk/projects/:fastqc), ATAC-seq reads after were aligned to mouse genome (mm10) using Bowtie2 (ref. [52]). Samples were filtered for regions blacklisted by the ENCODE project[53] and de-duplicated using Picard tools (http://broadinstitute.github.io/picard). Alignment coordinates were converted to BED format using BEDTools v.2.17 (ref. [54]) and peak calling was performed using MACS2 (ref. [55]) with default parameters. To visualize the relative occupancy of ThPOK along with the ATAC-seq peaks and the relative relationship to genes, we binned the chromosomes into consecutive 10-kb regions and plotted the read distribution of ATAC-seq and ThPOK ChIP–seq data.

ChIP–seq data analysis. We obtained ThPOK ChIP–seq data from Vacchio et al.[19] via the GEO (accession no. GSE116506). ChIP–seq reads were aligned to mouse genome (mm10) using the BWA aligner[56]. Peak calling was performed using MACS2 (ref. [55]). Similarly, CD8 T cell Runx3 ChIP–seq peaks were obtained from Istaces et al.[20] through the GEO (accession no. GSM3559330). Peak annotations and motif enrichment analyses were performed using HOMER tools[57].

Bulk RNA-seq. For each population, 10^5 cells were sorted into RLT lysis buffer (QIAGEN) containing 1% 2-mercaptoethanol and total RNA purified using the

RNA Microprep kit (Zymo Research). All resulting RNA was used as an input for cDNA synthesis using the SMART-Seq v.4 kit (Takara Bio) and ten cycles of PCR amplification. Next, 1 ng of cDNA was converted to a sequencing library using the NexteraXT DNA Library Prep Kit and NexteraXT indexing primers (Illumina) with ten additional cycles of PCR. Final libraries were pooled at equimolar ratios and sequenced on a HiSeq2500 using 100-bp paired-end sequencing. Quality-controlled reads after FastQC were aligned to mouse genome (mm10) using Tophat2 (PMID: 23618408). For counting reads from the resulting BAM files for ATAC-seq, ChIP–seq and RNA-seq, HTSeq[58] was used with default parameters. To identify DEGs, we applied the Deseq2 algorithm[59]. Genes with an FDR < 0.05 and absolute $\log_2$(fold-change) > 1 and < −1 were considered to be significant. Enrichment analyses were done using the DAVID bioinformatics resource[60].

Colitis models. Adoptive T cell transfer model: colitis was induced after transfer of 5×10^5 sorted naive T cells into Rag1$^{-/-}$ mice[61]. Recipient mice were monitored regularly for signs of disease, including weight loss, hunched appearance, piloerection of the coat and diarrhea, and analyzed at various times after the initial transfer or when they reached 90% of their initial weight.

TNBS colitis model: colitis was induced by intrarectal transfer of TNBS[62]. Recipient mice were monitored regularly for signs of disease, including weight loss, hunched appearance, piloerection of the coat and diarrhea, and analyzed at various times after the initial transfer or when they reached 80% of their initial weight.

3C assay. Quantitative analysis of chromosome conformation capture assays has been performed[63] using 4-bp cutter NlaIII, and ThPOK BAC plasmid was used as the positive control.

Statistics and reproducibility. No statistical method was used to determine sample size. Instead, sample sizes were rationalized by weighing sufficient replication (to determine the extent of biological variation) with a reduction of total animals used. Data were excluded only for technical reasons, such as low cell viability. With regard to replication, all RNA- and ATAC-seq analyses were performed on two to three independent samples, and all in vivo analyses were performed on a total of three to six animals per genotype (across at least three separate experiments). All attempts at replication were successful. Randomization was not used; assignment to experimental groups was based on genotype. To exclude physiological and environmental covariates, mice of different genotypes were derived from the same litters as control mice (as much as possible), or co-housed before analysis. The investigators were blinded to allocation for histopathological analysis of colitis induction samples. Other blinding was not possible because genotyping was necessary for all other experiments. Statistical analysis for nonsequencing data was performed using GraphPad Prism software. Data were analyzed by applying unpaired, two-tailed Student's *t*-tests, and one-way analysis of variance (ANOVA) with Bonferroni's correction. A P value < 0.05 was considered to be significant. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. All statistical analyses performed for sequencing data are mentioned elsewhere in Methods.

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

Data availability

All newly generated sequencing data for this study can be accessed at the GEO under accession code GSE168772. All other data that support the findings of this study are available from the corresponding author upon request.

References

51. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21.29.1–21.29.9 (2015).
52. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
53. Hoffman, M. M. et al. Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res.* **41**, 827–841 (2013).
54. Quinlan, A. R. BEDTools: the Swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinform.* **47**, 11.12.1–11.12.34 (2014).
55. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).
56. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
57. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
58. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
59. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
60. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
61. Mucida, D. et al. Reciprocal T_H17 and regulatory T cell differentiation mediated by retinoic acid. *Science* **317**, 256–260 (2007).
62. Antoniou, E. et al. The TNBS-induced colitis animal model: an overview. *Ann. Med. Surg.* **11**, 9–15 (2016).
63. Hagège, H. et al. Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat. Protoc.* **2**, 1722 (2007).

Acknowledgements

This work was supported by NIH grants (nos. R01 AI068907 and R01 GM107179 (to D.J.K.), R35 GM122502 (to J.F.B. and A.J.M.W.), R01CA227629 and R01CA218133 (to S.G.) and P30 CA006927 (FCCC Comprehensive Cancer Center Core Grant)). We thank the following core facilities of the Fox Chase Cancer Center for their assistance: Flow Cytometry, Cell Culture, DNA Sequencing, Genomic and Laboratory Animal. We thank D.L. Wiest and G. Rall for critical comments and suggestions on the manuscript.

Author contributions

J.B. and D.J.K. conceived the project. J.B., B.S.R., S.P., X.H., A.J.M.W., H.L.G., S.G., D.M. and D.J.K. carried out the methodology. J.B., B.S.R., S.P., J.Z., X.H., L.G., K.F., E.N., P.C., K.Q.C., Y.T., J.I.F.B. and D.J.K. performed the investigations. J.B. and D.J.K. wrote the original draft of the manuscript. J.B., A.J.M.W., S.G., D.M. and D.J.K. reviewed and edited the manuscript. D.J.K. acquired the funding. J.B. and D.J.K. were overall supervisors.

Competing interests

The authors declare no competing interests.

Additional information

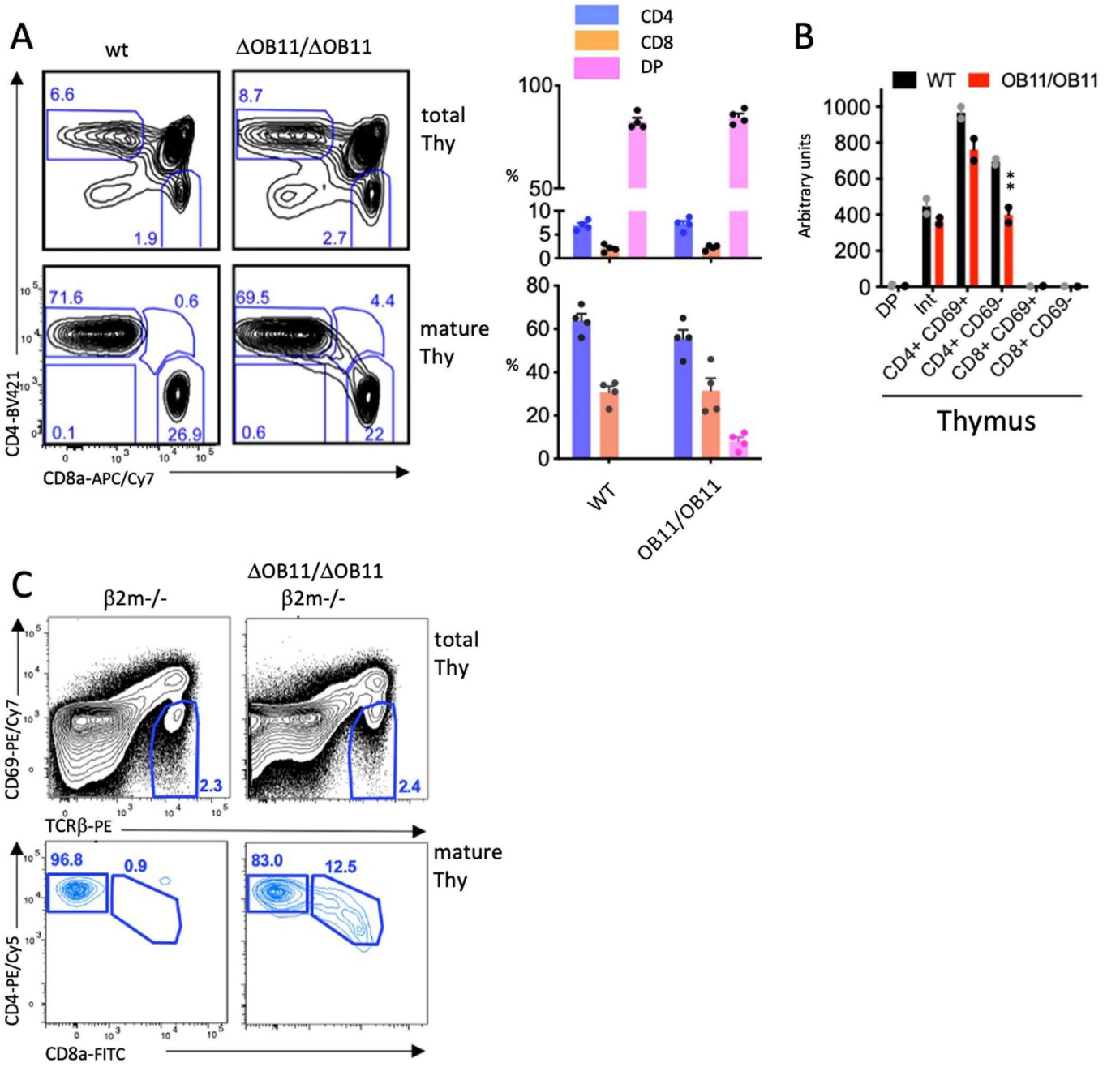
Extended data Extended data are available for this paper at https://doi.org/10.1038/s41590-021-00980-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-021-00980-8.

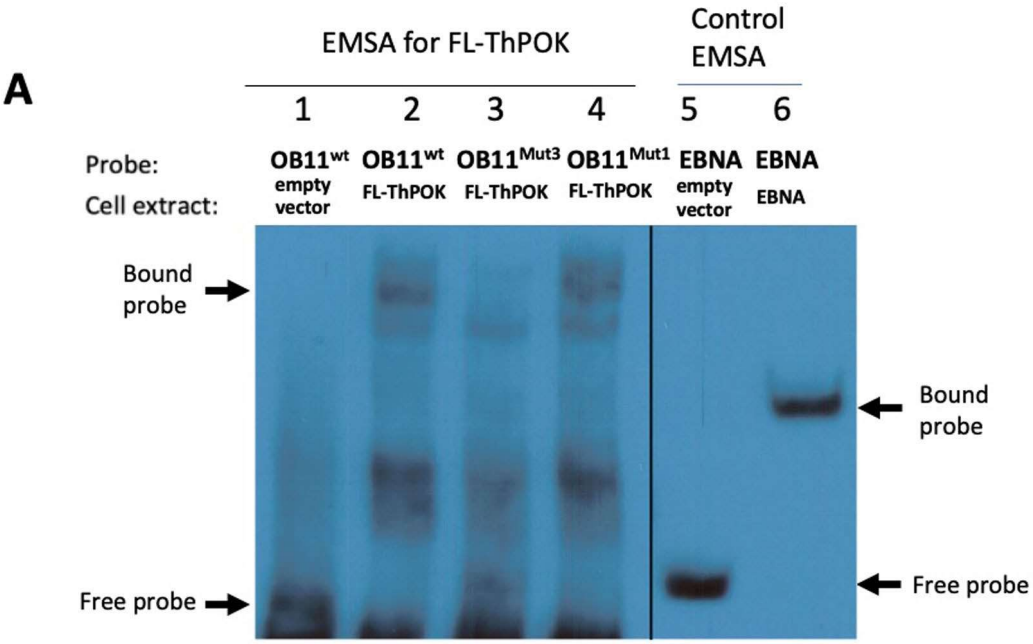
Correspondence and requests for materials should be addressed to D.J.K.

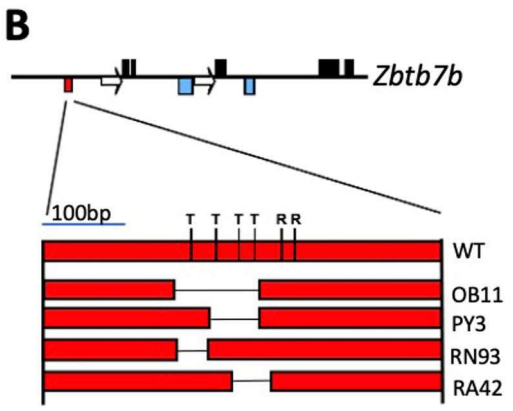
Peer review information *Nature Immunology* thanks Jonathan Kaye and Harinder Singh for their contribution to the peer review of this work. Laurie Dempsey was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

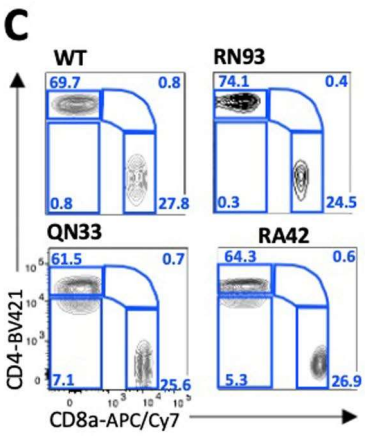
Reprints and permissions information is available at www.nature.com/reprints.

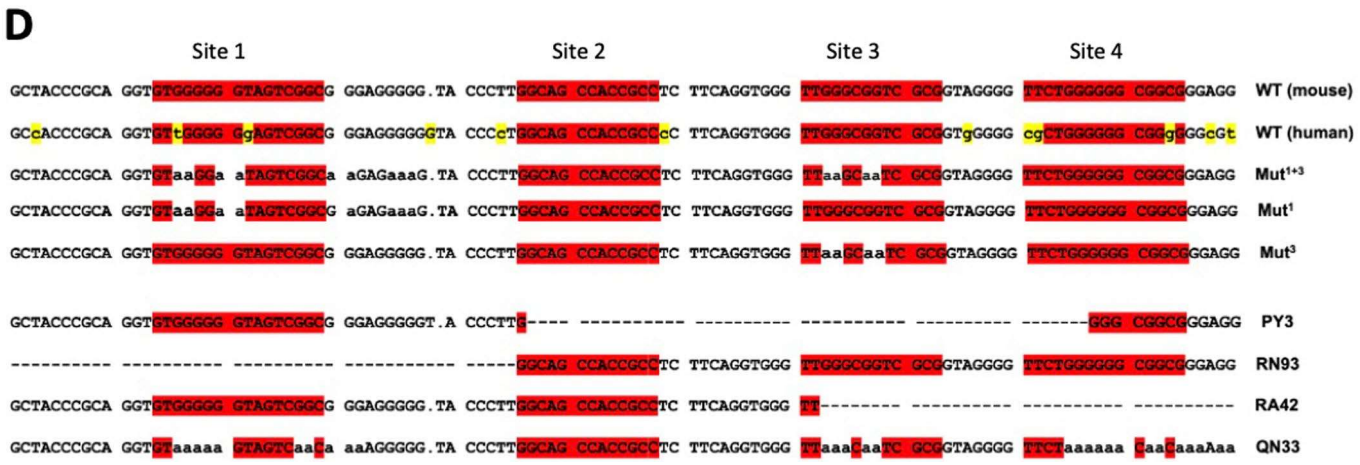


Extended Data Fig. 1 | Ablation of anti-silencer does not substantially disrupt development of class II-restricted thymocytes. a, Flow cytometric analysis of CD4, and CD8a expression in total thymocytes (top row), or gated mature (TCRβ + CD69- CD24- CD62L +) thymocytes of WT mice and homozygous ThPOK$^{\Delta OB11/\Delta OB11}$ mice, as indicated. A total of 4 animals of each genotype were analyzed over 2 independent experiments. Plots at right indicate % of SP CD4, CD8, and CD4 + 8 + (DP) thymocytes among total thymocytes (top panel), or gated mature thymocytes (as defined in panel a) (n = 4, for each strain). **b**, RNA was collected from freshly sorted thymocyte subsets, as indicated, before probing for ThPOK expression by qPCR. Data represent 2 technical replicates, each derived from pooled RNA of 3 animals. Data are presented as mean values +/- SEM. Experiment was repeated 3 times. **c**) Flow cytometric analysis of TCRβ and CD69 expression in total thymocytes (top row), or gated mature thymocytes of β2m-/- ThPOK +/+ and β2m-/- ThPOK$^{\Delta OB11/\Delta OB11}$ mice, as indicated. Data are representative of 6 animals of each genotype that were analyzed in 3 separate experiments.

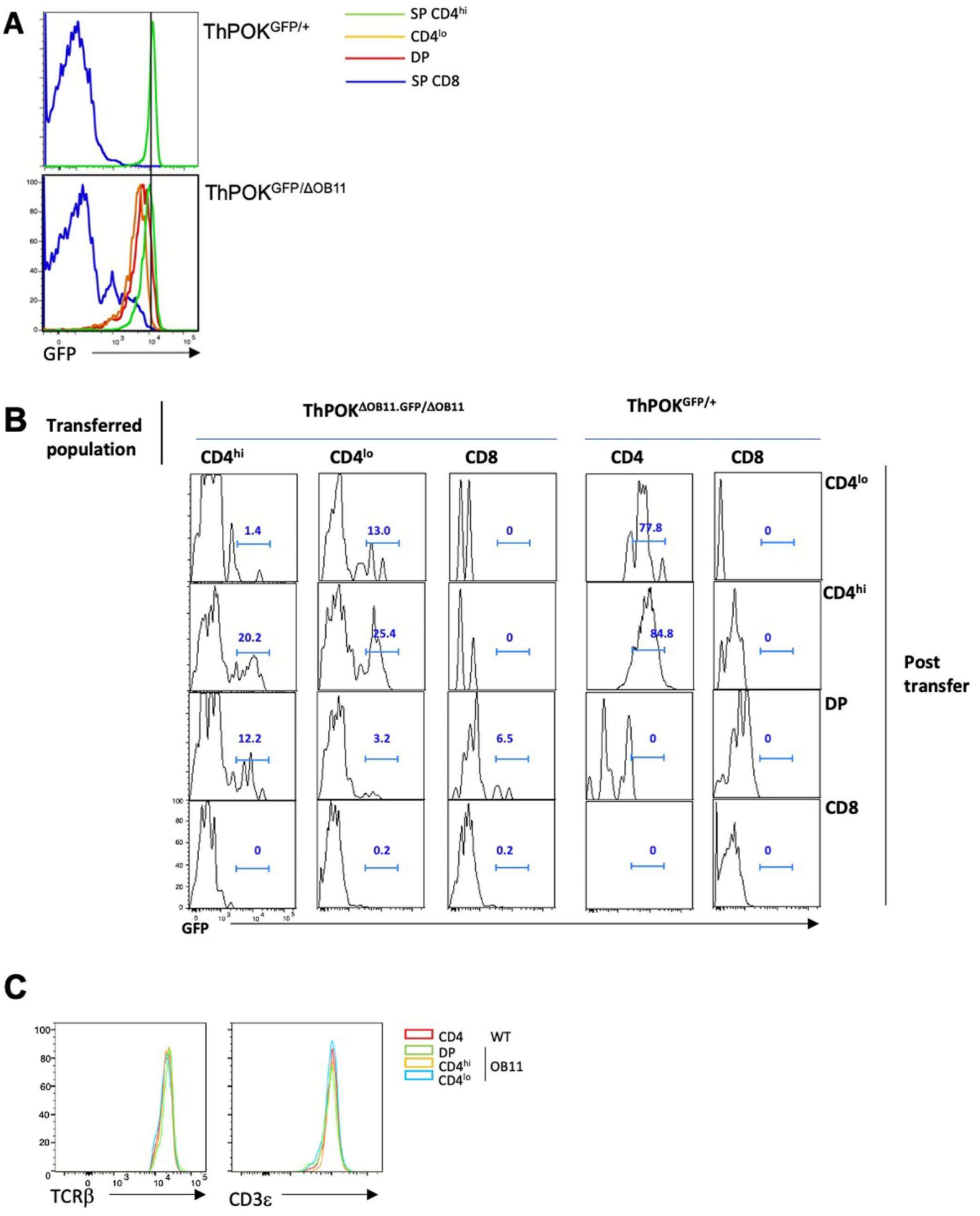




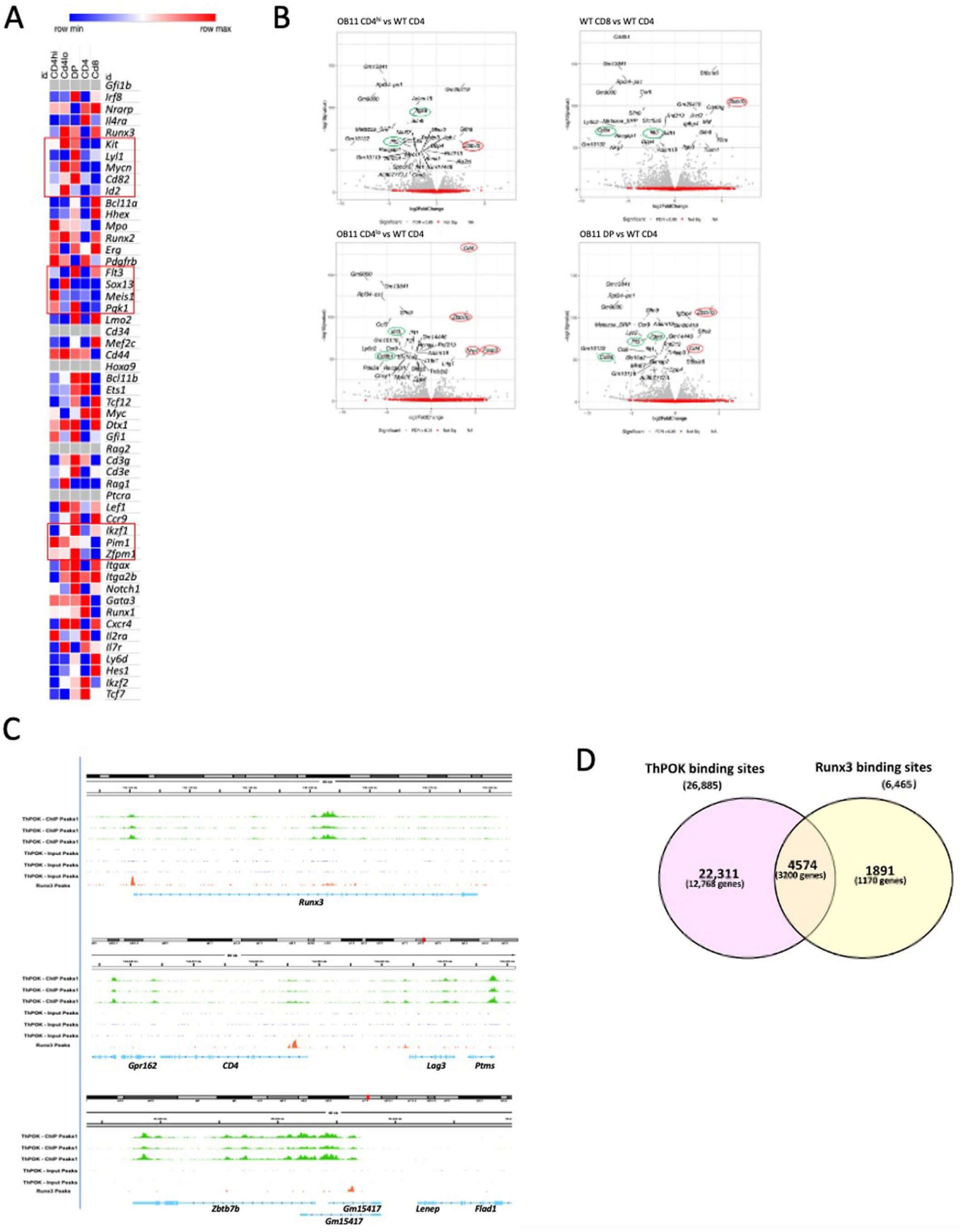




Extended Data Fig. 2 | Functional analysis of ThPOK binding sites in anti-silencer. a, EMSA analysis, using the 100 bp OB11 region (lanes 1-4) or EBNA control DNA (lanes 5,6) as probes. Biotinylated probes were incubated with cell extracts from NIH 3T3 cells transfected with empty vector, Fl-ThPOK or EBNA expression constructs, as indicated. In some lanes mutant OB11 probes in which the consensus ThPOK binding sites 1 (lane 4) or 3 (lane 3) are disrupted were added. Experiment was repeated 3 times with similar results. **b**, Organization of mouse *ThPOK* gene (top), and diagram of silencer deletion mutants (bottom). Black boxes, blue boxes, arrows, and red boxes indicate exons, enhancers, silencers and promoters, respectively. Deletions within the silencer are indicated by thin black lines. "R" indicates positions of conserved Runx binding sites. **c**, Flow cytometric analysis of CD4, and CD8a expression in gated TCRβ + PBLs of WT mice and homozygous mutant lines, as designated in panel b. A total of 9 animals of each genotype were analyzed in 3 separate experiments. **d**, Sequence of WT murine OB11 region (top row) aligned with corresponding human sequence (second row; mismatches between human and mouse are highlighted in yellow), as well as sequences of indicated mutants used in EMSA or for generation of mutant alleles.

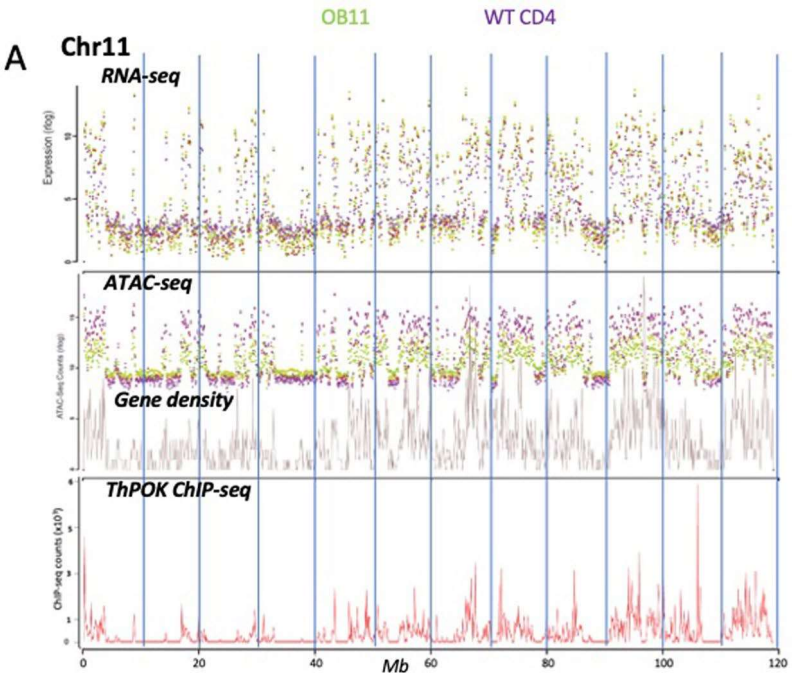


Extended Data Fig. 3 | Ablation of anti-silencer leads to transient transcription at ThPOK locus. a, Flow cytometric analysis of GFP expression in gated CD4^{hi} (green), CD4$^{lo/-}$ (orange), DP (CD4+CD8+) (red), or CD8 PBLs (middle) of ThPOK$^{GFP/+}$ or ThPOK$^{GFP/\Delta OB11}$ mice. Black line indicates MFI of GFP expression by ThPOK$^{GFP/+}$ CD4 T cells, for comparison (same mice as in Fig. 3c). **b**, Flow cytometric analysis of GFP expression by gated CD4^{hi}, CD4$^{lo/-}$, CD4+CD8+(DP), and CD8 PBLs from RAG-/- hosts reconstituted with indicated sorted T cell subsets from ThPOK$^{GFP/+}$ or ThPOK$^{\Delta OB11.GFP/\Delta OB11}$ mice, as indicated (same mice as in Fig. 3e). **c**) Surface TCRβ and CD3ε expression of indicated Sil$^{\Delta OB11/\Delta OB11}$ or WT T cell subset after activation.



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Genes misregulated upon anti-silencer ablation are often direct targets of ThPOK and Runx3 binding. a, Heat map displaying relative expression of genes associated with ETP-DN2 transition for indicated ThPOK$^{\Delta OB11/\Delta OB11}$ and WT T cell subsets. Red indicates increased gene expression levels; blue indicates decreased levels. **b**, Volcano plots illustrating gene expression differences between individual sorted ThPOK$^{\Delta OB11/\Delta OB11}$ subsets and WT CD4 T cells (Wald 2-sided test). The grey dots represent genes differentially expressed (adjusted $P < 0.05$) between samples. Genes with the largest negative or positive standardized mean difference are marked. Note diminished ThPOK (Zbt7b) expression in all samples. **c**, ThPOK-ChIP and Runx3-Chip profiles and Refseq gene positions across chromosome regions containing Runx3, CD4 and ThPOK (Zbtb7b) genes, as marked. Input is already subtracted for Runx3 peaks. **d**, Venn diagram indicating intersection between ThPOK and Runx3 ChIP-seq peaks (from WT CD4 and WT CD8 cells, respectively).



B

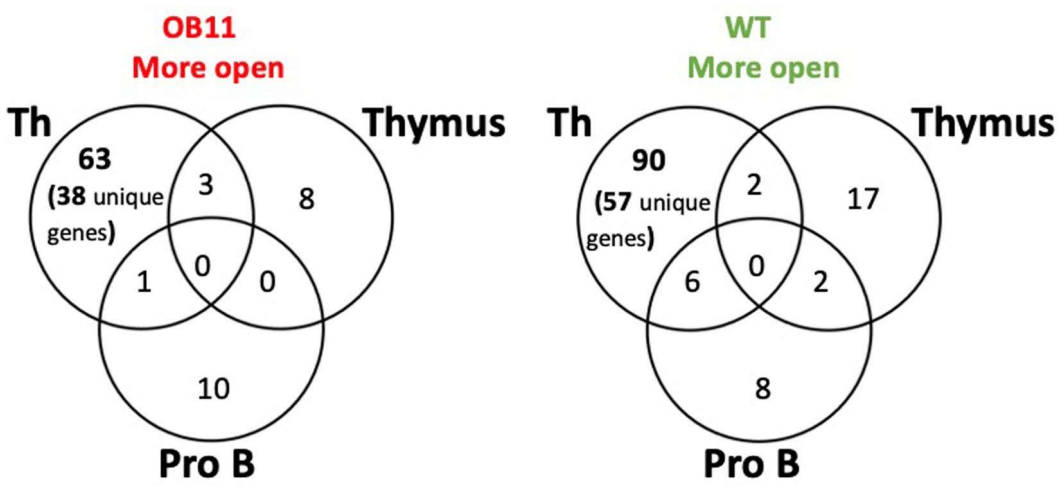
OB11 CD4lo

	Name	P-value
1	ETS1-ChIP-seq; Jurkat (GSE17954)/Homer	1e-20
2	Etv2 (ETS)-ChIP-seq; ES (GSE59402)/Homer	1e-19
3	Runx1-ChIP-seq; HPC7 (GSE22178)/Homer	1e-18
4	Runx2-ChIP-seq; PCa (GSE33889)/Homer	1e-18
5	Runx1-ChIP-seq; Jurkat (GSE17954)/Homer	1e-16
6	PU.1(ETS)-ChIP-seq; ThioMac (GSE21512)/Homer	1e-16
7	Fli1(ETS)-ChIP-seq; CD8 (GSE20898)/Homer	1e-13
8	Runx-ChIP-seq; CD4 (Barski et al)/Homer	1e-13
9	Etv1 (ETS)-ChIPseq; CD4 (GSE22441)/Homer	1e-13
10	Runx1-ChIPseq; Jurkat (GSE29180)/Homer	1e-13
11	GABPA (ETS)-ChIPseq; Jurkat (GSE17954)/Homer	1e-13
12	ETV4 (ETS)-ChIPseq; HepG2(ENCODE)/Homer	1e-13
13	Ets1-ChIPseq; CD4 (Barski et al)/Homer	1e-12
14	ERG (ETS)-ChIPseq; VCaP (GSE14097)/Homer	1e-11
15	EWS:ERG (ETS)-ChIPseq; CADO (SRA014231)/Homer	1e-9

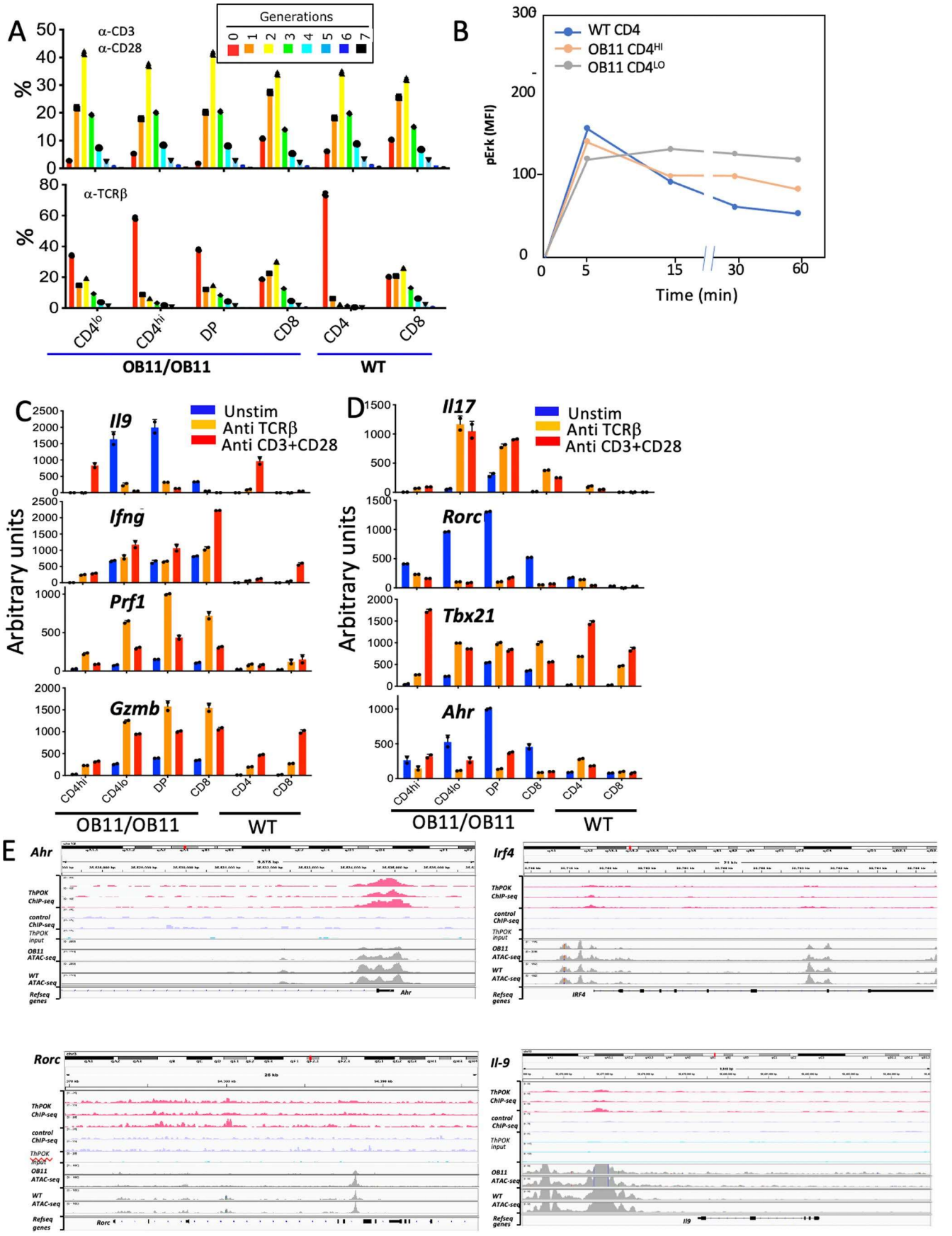
WT

	Name	P-value
1	Etv2 (ETS)-ChIP-seq; ES (GSE59402)/Homer	1e-20
2	ETS1-ChIP-seq; ES (GSE17954)/Homer	1e-19
3	ERG (ETS)-ChIP-seq; VCaP (GSE14097)/Homer	1e-18
4	Fli1 (ETS)-ChIP-seq; CD8 (GSE20898)/Homer	1e-18
5	ETV1 (ETS)-ChIP-seq; GIST48 (GSE22441)/Homer	1e-16
6	ESW:ERG(ETS)-ChIP-seq; CADO (SRA014231)/Homer	1e-16
7	ETV4 (ETS)-ChIP-seq; HepG2 (ENCODE)/Homer	1e-13
8	GABPA (ETS)-ChIP-seq; Jurkat (GSE17954)/Homer	1e-13
9	ETS1-ChIPseq; CD4 (Barski et al)/Homer	1e-13
10	CTCF-ChIPseq; CD4 (Barski et al)/Homer	1e-13
11	CTCFL-ChIPseq; K562 (GSE32465)/Homer	1e-13
12	EWS:Fli1(ETS)-ChIPseq; SK-N-MC[SRA014231)/Homer	1e-13
13	Ets1-ChIPseq; CD4 (Barski et al)/Homer	1e-12
14	EHF (ETS)-ChIPseq; LoVo (GSE49402)/Homer	1e-11
15	ELF3 (ETS)-ChIPseq; PDAC (GSE64557)/Homer	1e-9

C

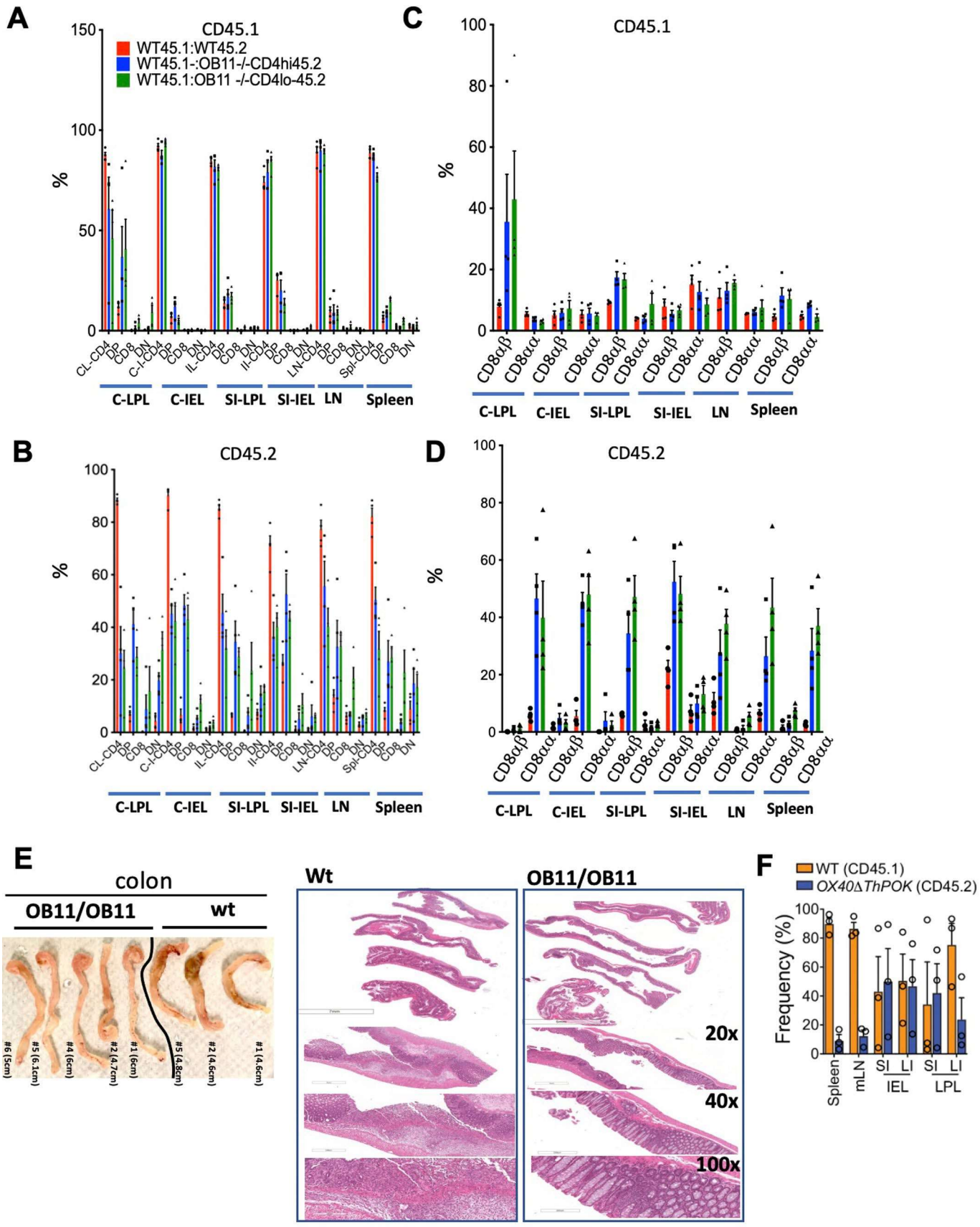


Extended Data Fig. 5 | Altered chromatin accessibility in ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} T cells. a, Comparison of distribution of RNA-seq reads (top row), ATAC-seq reads (second row, top), genes (second row, bottom), and ThPOK-ChIP seq reads per 100 kb bins across one chromosome (Chr11). **b**, TF consensus motif enrichment in DACRs selectively open in ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} T cells or WT CD4 T cells, as indicated (Wilcoxon Rank Sum Test with p value adjusted by Benjamini Hochberg method). **c**, Venn diagram illustrating intersection between super-enhancers from Th cells, thymocytes and proB cells, as indicated.



Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Effect Anti-silencer element ablation on CD4 T cell proliferation and cytokine gene expression. a, Bar graph denoting the percentage of cells assigned to each division cycle based on CFSE dilution following stimulation with anti-TCRβ (bottom) or anti-CD3ε/CD28 (top) for 48 h. **b**, Time course of phospho-ERK expression in response to anti- TCRβ stimulation of indicated WT or OB11 T cell subsets. **c**, **d**, Bar graphs illustrating deregulated expression of indicated mRNAs according to qPCR. Sorted OB11 or WT T cell subsets were cultured with anti-TCRβ or anti-CD3ε/CD28 stimulation, or in the absence of stimulation, as indicated. N = 2 (technical replicates, each derived from pooled RNA of 4 animals for each genotype). Experiment was repeated twice. Data are presented as mean values +/- SEM. **e**, ThPOK-ChIP-seq (for WT CD4 T cells, biological triplicates; pink tracks), and ATAC-seq (for ThPOK$^{\Delta OB11/\Delta OB11}$ CD4^{lo} T cells or WT CD4 T cells, as indicated, biological duplicates; grey tracks) profiles for gene loci.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Colitis induction in WT and ThPOK$^{\Delta OB11/\Delta OB11}$ mice. a–d, Bar graphs depicting the percentage of CD45.1^{+} (**a**, **c**) or CD45.2^{+} (**b**, **d**) cells within each indicated intestinal or splenic population. Data is derived from cotransfer experiment depicted in Fig. 7f. Donor cells were sorted from pooled LN samples of 3 CD45.1+ WT or CD45.2+ ThPOK$^{\Delta OB11.GFP/\Delta OB11}$ mice. N = 4 (# of host animals reconstituted with each indicated cell combination). The experiment was repeated 3 times. **e**, Image of whole colons, or stained cross-sections of the indicated mouse strains after colitis induction with TNBS. Experiment was repeated 2 times with similar results. **f**, Frequency of indicated cell type of OX40ΔThPOK or WT origin at indicated organ site following co-transfer of sorted OX40ΔThPOK or WT CD4 T cells (same experiment as in Fig. 7l–n). N = 3 (WT + Ox40 co transfer), or 2 (WT only), and the experiment was repeated 2 times. Data in panels a-d and f are presented as mean values +/- SEM.

nature research

Corresponding author(s): Dietmar Kappes

Last updated by author(s): 2021-5-26

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	For ATAC-seq, libraries were sequenced using Illumina 125-bp paired-ends sequencing on a HiSeq2500 platform, generating between 38 and 43 million reads per condition per biological replicate. For RNA-seq, final cDNA libraries were pooled at equimolar ratios and sequenced on a HiSeq2500 using 100-bp paired-end sequencing.
Data analysis	For analysis of ATAC-seq data, quality control was performed by FastQC version 0.11.5 (URL:https://www.bioinformatics.babraham.ac.uk/projects/:fastqc/), and reads were aligned to mouse genome (mm10) using BoWTie2 version 2.3.4.3 (PMID: 22388286). Samples were filtered for regions blacklisted by the ENCODE project and deduplicated using Picard tools, version 2.1.1 (http://broadinstitute.github.io/picard). Alignment coordinates were converted to BED format using BEDTools v.2.17 and peak calling was performed using MACS2 with default parameters. For RNA-seq, quality controlled reads after FastQC were aligned to mouse genome (mm10) using Tophat2, version 2.1.1. For counting reads from the resulting BAM files for ATAC-Seq, and RNA-Seq, HTSeq was used with default parameters. To identify differentially expressed genes (DEG), we applied Deseq2 algorithm. Genes with an FDR < 0.05 and absolute log2(FC) > 1 and < -1 were considered to be significant. Enrichment analyses were done using DAVID bioinformatics resource. HOMER version used was 4.10.4. Flow cytometric data was analyzed using FlowJo software (versions 9.3.3, 10.1 or 10.2, FlowJo, Ashland, OR, USA).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

GEO accession number GSE168772 for RNA-Seq and ATAC-Seq raw and processed data/metadata.
No data restrictions

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. Instead, sample sizes were rationalized by weighing sufficient replication (to determine the extent of biological variation) with reduction of total animals used.
Data exclusions	Data was excluded only for technical reasons, such as low cell viability.
Replication	Results were confirmed by analysis of individual biological replicates, and all attempts at replication were successful.
Randomization	Sample allocation was not random. Instead, biological controls were included in all experiments.
Blinding	The investigators were blinded to allocation for histopathological analysis of colitis induction samples. Other blinding was not possible since genotyping was necessary for all other experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

All fluorescently labeled antibodies used were obtained from commercial sources (eBioscience, Biolegend, or BD, as indicated) and were validated by the supplier: Thy1-FITC (clone 30-H12; BioLeg Cat # 105306; Lot B224687), Thy1-APC (30-H12; BioLeg Cat # 140312; Lot # B274408), TCRβ-APC (clone H57-597; BioLeg Cat # 109212; Lot # B273860), TCRβ-PE (clone H57-597; eBio Cat # 12-5961-83; Lot # E01951-1631), TCRβ-FITC (clone H57-597; BD Cat # 553171; Lot # 01934), CD4-BV421 (clone RM4-5; BioLeg Cat # 100544; Lot # B293278), CD4-PE/Cy5 (clone RM4-5; BioLeg Cat # 100514; Lot # B231410), CD25-PE (clone PC61; BD Cat # 553866; Lot # 5047982), CD25-PE/Cy7 (clone PC61; BD Cat # 552880; Lot # 36592), CD8a-APC/Cy7 (clone 53-6.7; BioLeg Cat # 100714; Lot # B276265), CD8a-FITC (clone 53-6.7; BioLeg Cat # 100706; Lot # B168591), CD8a-PE/Cy7 (clone 53-6.7; BioLeg Cat # 100722), CD8a-AF700 (clone 53-6.7; BioLeg Cat #100730), CD8b-PE (clone 53-5.8; BD Cat # 553041; Lot # 3193966), CD69-PE/Cy7 (clone H1.2F3; eBio Cat # 25-0691-82; Lot # E07583-1635), CD62L-APC (clone MEL-14; BD Cat # 553152; Lot # 7075853), CD45.1-BV605 (clone A20; BioLeg Cat # 110738), CD45.2-AlexaFluor647 (clone 104; BioLeg Cat # 109818; Lot # B181101), CD73-FITC (clone TY/11.8; BioLeg Cat # 127219), FR4-PerCP/Cy5.5 (clone 12A5; BioLeg Cat # 125018), Foxp3-PE/Cy7 (clone FJK-16S; eBio Cat # 25-5773-82; Lot #

E07638-1634), Foxp3-APC (clone FJK-16S; eBio Cat #17-5773-82), CD44-BV605 (clone IM7; BioLeg Cat # 103047; Lot # B288308), GITR-Biotin (clone DTA-1; BioLeg Cat # 126305; Lot # B254051), PD1-APC (clone 29F.1A12; BioLeg Cat # 135210), Streptavidin-BV605 (BioLeg Cat # 405229; Lot # B242305). Anti-SMAD4 (clone EP618Y; Cat # ab215968) was procured from Abcam. Unlabelled anti-CD28 (clone 37.51; BioLeg Cat # 102131), TCRb (clone H57-597; BioLeg Cat # 102214; Lot # B218495), CD3e (clone 145-2C11; BioLeg Cat # 100340; Lot # B260927), IL-4 (clone 11B11; Cat # 504115; Lot # B213842) and anti–IFN-γ (clone XMG1.2; Cat # 505827; Lot # B218186) were from BioLegend (San Diego, CA). Anti SMAD4 procured from ABCAM.

Validation

All antibodies were validated by the source company.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Male and female mice were maintained on a C57BL/6 background and were analyzed between 5 and 12 weeks of age.
Wild animals	The study did not involve samples collected from wild animals.
Field-collected samples	The study did not involve samples collected from the field
Ethics oversight	All experimentation involving animals was approved by Institutional Animal Care and Use Committee (IACUC), of Fox Chase Cancer Center, or Rockefeller University.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Mice were euthanized using carbon dioxide followed by cervical dislocation. Peripheral blood was collected with heparin-coated capillaries. Mesenteric lymph nodes, thymus and spleen were harvested immediately after euthanasia and stored in cold medium (2% FBS, RPMI) under sterile conditions. Single-cell sell suspensions were obtained by crushing organs through a 40 µm cell strainer (Becton, Dickinson and Company). Prior to analytical flow cytometry, blood, spleen and lymph node samples were layered over Lympholyte-M (Cedarlane), centrifuged at 1,200rpm for 30', and interface collected. For isolation of IEL and LPLs, intestines were dissected longitudinally in cold HBSS containing 5% FBS and 10 mM EDTA, cut into small pieces (about 5 mm), shaken at 37 'C for 20 min and passed through a metal sieve. This procedure was repeated on remaining tissue pieces for 2-3 times until supernatants appeared clear, and filtered through the metal sieve. Accumulated cells were pelleted, resuspended in 40% percoll, layered over 80% percoll, and centrifuged at at 2200 rpm for 25 minutes at room temperature (with slow acceleration and 0 deceleration). IELs were collected from the interphase between 40 and 80% percoll. For LPL preparation residual tissues were further minced and digested in 5% FBS in HBSS containing 1 mg/ml collagenase type VIII (Sigma) for 25-30 min with shaking at 37 0C. After that, cells were filtered through consecutive 100 µm and 40 µm cell strainers. Cells were pelleted and run over Percoll gradient as above.
Instrument	Flow cytometry analyses were conducted on a FACS LSRII or FACS LSRFortessa (Becton, Dickinson, and Company). Cell sorting was performed on a FACSAria II (Becton, Dickenson, and Company).
Software	FACS data were collected using FACS Diva version 7.0 or 9.0, and data was analyzed using FlowJo software (versions 9.3.3, 10.1 or 10.2, FlowJo, Ashland, OR, USA).
Cell population abundance	No post-sort analysis was performed.
Gating strategy	For all analyses, PI+ or 7AAD+ cells were excluded, then debris was excluded using a FSC-A vs SSC-A gate, then doublets were excluded using FSC-W vs FSC-H or SSC-W vs SSC-H gates for all downstream gating. Fluorescent minus-one controls were used in some circumstances to assist in discriminating between positive and negative signal, while other gating was performed according to previously published strategies.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.