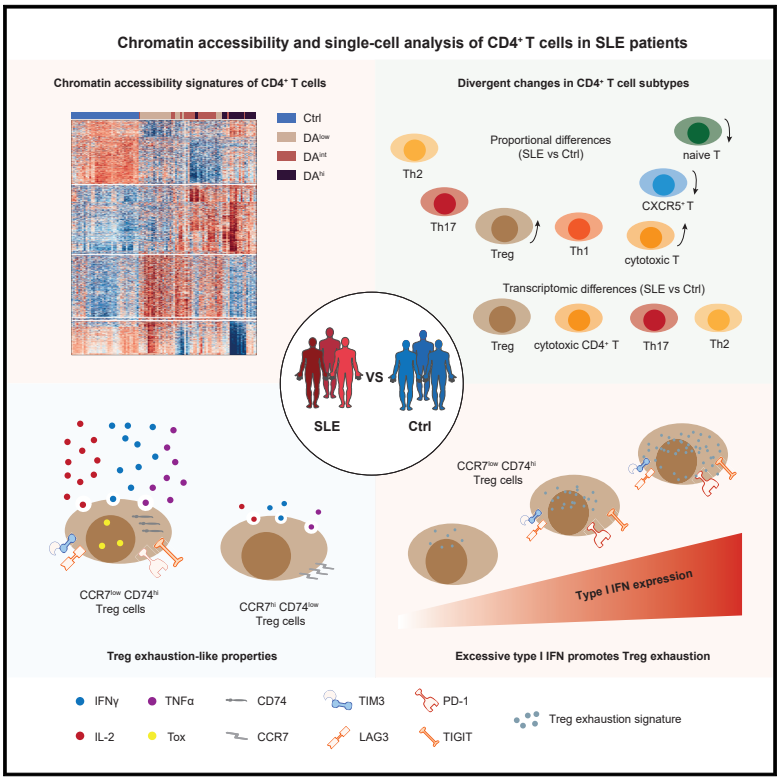
Cell Reports

Article

Single-cell transcriptome profiling and chromatin accessibility reveal an exhausted regulatory CD4+ T cell subset in systemic lupus erythematosus

Graphical abstract



Highlights

- Chromatin accessibility landscape of CD4$^+$ T cells in SLE patients
- Single-cell atlas of CD4$^+$ T cells from severe-stage SLE patients
- Exhaustion-like properties in CCR7^{low}CD74^{hi} Treg cells from SLE patients
- Type I interferon promotes Treg exhaustion-like properties in SLE patients

Authors

Chuang Guo, Qian Liu, Dandan Zong, ..., Quan Wu, Xiaomei Li, Kun Qu

Correspondence

qukun@ustc.edu.cn

In brief

Guo et al. present chromatin accessibility landscapes and single-cell transcriptome profiling of CD4$^+$ T cells from a total of 72 SLE patients and 30 healthy control individuals to identify type I interferon-induced functional exhaustion-like properties in Treg cells.

Guo et al., 2022, Cell Reports *41*, 111606
November 8, 2022
https://doi.org/10.1016/j.celrep.2022.111606

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Single-cell transcriptome profiling and chromatin accessibility reveal an exhausted regulatory CD4+ T cell subset in systemic lupus erythematosus

Chuang Guo,[1,5] Qian Liu,[1,5] Dandan Zong,[1,5] Wen Zhang,[1,2] Zuqi Zuo,[1] Qiaoni Yu,[1] Qing Sha,[1] Lin Zhu,[1] Xuyuan Gao,[1] Jingwen Fang,[1,3] Jinhui Tao,[1] Quan Wu,[4] Xiaomei Li,[1] and Kun Qu[1,2,6,*]

[1]Department of Rheumatology and Immunology, the First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui 230021, China

[2]Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, Anhui 230088, China

[3]HanGene Biotech, Xiaoshan Innovation Polis, Hangzhou, Zhejiang 31200, China

[4]Central Laboratory of Medical Research Center, the First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui 230021, China

[5]These authors contributed equally

[6]Lead contact

*Correspondence: qukun@ustc.edu.cn

https://doi.org/10.1016/j.celrep.2022.111606

SUMMARY

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease, and CD4$^+$ T cells are known to promote SLE development. Here, we explore heterogeneities in the CD4$^+$ T cell regulome and their associations with SLE pathogenesis by performing assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) and single-cell transcriptome sequencing (single-cell RNA sequencing [scRNA-seq]) of peripheral CD4$^+$ T cells from 72 SLE patients and 30 healthy controls. Chromatin accessibility signatures of CD4$^+$ T cells are correlated with disease severity. Further, we generate 34,176 single-cell transcriptomes of healthy and SLE CD4$^+$ T cells and reveal transcriptional dysfunction of regulatory T (Treg) cells, identifying two Treg subpopulations, among which the CCR7^{low}CD74^{hi} Treg subgroup features type I interferon-induced functional exhaustion in SLE patients. These transcriptome-level findings for SLE Tregs are mirrored in trends from the ATAC-seq data. Our study establishes a rich empirical foundation for understanding SLE and uncovers previously unknown contributions of Treg with exhaustion-like properties to SLE pathogenesis.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a complex systemic autoimmune disease caused by perturbations in self-tolerance, leading to the activation of autoreactive B cell and T cell immune responses in multiple tissues (Tsokos, 2020). Compelling evidence has shown that epigenetic modifications (e.g., gene regulation that involves chromatin modifications or chromatin accessibility) of T cells are involved in SLE pathogenesis (Tsokos et al., 2016). Specifically, increased acetylation of histone H3/H4 was identified in SLE CD4$^+$ T cells (Hu et al., 2008), and decreased methylation of costimulatory molecules (e.g., *CD40LG*) and interferon (IFN) signature genes (e.g., *IFI44L*) (Hedrich et al., 2017) in CD4$^+$ T cells was found to correlate with the extent of inflammation and tissue damage in SLE patients (Coit et al., 2016). Therefore, understanding the impacts of global chromatin accessibility in CD4$^+$ T cells and linking differential accessibility to SLE disease activity can deepen our understanding of SLE pathogenesis.

Recent studies have employed single-cell transcriptome sequencing (single-cell RNA sequencing [scRNA-seq]) technology to delineate the landscapes of immune cells in the peripheral blood and kidneys of SLE patients (Arazi et al., 2019; Nehar-Belaid et al., 2020). These foundational studies have provided an empirical basis for dissecting the cellular heterogeneity of patients' immune systems. However, the current lack of data for specific cell populations—for example, CD4$^+$ T cells—has thus far limited our capacity to make data-driven inferences about the specific impacts of distinct immune cell types on this autoimmune disease.

Regulatory CD4$^+$ T (Treg) cells are a subset of CD4$^+$ T cells known to function in maintaining immune tolerance and are essential for immune system homeostasis (Sakaguchi et al., 2020). Multiple studies on patients with autoimmune diseases have observed attributes of defective Treg cells (Dominguez-Villar and Hafler, 2018; Wing et al., 2019), such as increased proportions of Th1-like Treg cells, which produce the inflammatory cytokine IFN-γ (Sumida et al., 2018). Interleukin (IL)-17-producing Foxp3-expressing Treg cells were also reported by a previous study (Komatsu et al., 2014). However, there is increasing awareness that some trends observed in early studies may be misleading, owing to the introduction of potentially confounding factors related to the identification of Treg

cells (based on the expression of surface markers) (Dominguez-Villar and Hafler, 2018). Single-cell profiling technologies provide an alternative method for identifying and investigating Treg cells under pathological conditions.

Type I IFNs were reported to be blood transcriptional signatures in SLE patients (Banchereau et al., 2016). Several studies have suggested that excess levels of type I IFNs may induce Treg dysregulation, with obvious trends to differing extents in SLE patients with different disease severities (Crow et al., 2019); in contrast, other studies have reported positive effects of IFN mediation on Treg expansion (or function) under inflammatory conditions (Lee et al., 2012). It is thus clear that the impact(s) of type I IFNs on Treg suppression remain quite controversial in autoimmune disease studies.

Here, we utilized the assay for transposase-accessible chromatin sequencing (ATAC-seq) and scRNA-seq to analyze peripheral $CD4^+$ T cells from SLE patients and healthy controls and generated high-resolution landscapes of the epigenome and single-cell transcriptome of $CD4^+$ T cells *in vivo*. We found major epigenetic regulatory patterns corresponding to SLE disease activity based on in-depth analyses of our ATAC-seq data. Additionally, single-cell transcriptome analysis enabled the identification of $CD4^+$ T cell subtypes in SLE patients that were distinct from those of healthy controls and enabled an exploration of their potential contributions to SLE pathogenesis. Ultimately, we discovered that $CCR7^{low}CD74^{hi}$ Treg cells feature type I IFN-associated functional exhaustion in SLE patients as well as consistently dysregulated patterns at the transcriptome and chromatin accessibility. We also detected apparently similar Treg cells in $CD4^+$ T cell datasets from patients with ulcerative colitis and multiple sclerosis.

RESULTS

Chromatin accessibility landscapes of $CD4^+$ T cells from healthy controls and SLE patients

We generated 102 high-resolution ATAC-seq profiles of primary $CD4^+$ T cells and conducted a genome-wide analysis to map the locations and profile the accessibilities of diverse regulatory elements (Figure 1A). The examined cells were freshly isolated from peripheral blood mononuclear cells (PBMCs) using fluorescence-activated cell sorting (FACS) (Figures S1A and S1B). Our dataset included a total of 65 samples from 63 SLE patients (58 of which had clinical information) and 37 samples from 25 healthy controls (Table S1). For each SLE patient, detailed medication history and comorbidities at the time of blood draw were recorded and summarized (Table S1). Specifically, the patient cohort comprised 40 individuals with new-onset disease (24 of whom were drug naive). According to the disease activity index (DAI) determined using the SLEDAI-2K method (Yee et al., 2007), 38 SLE patients were at a severe stage, with a DAI > 11 (24 of whom were freshly diagnosed). The average DAI was ~15 in this patient cohort for bulk ATAC-seq of primary $CD4^+$ T cells (Table S1).

Each ATAC-seq library was sequenced to obtain an average of more than 20 million paired-end reads (Table S1). With this dataset, we identified a total of 103,317 high-quality peaks for DNA accessibility in healthy and SLE $CD4^+$ T cells using the previously reported tool ATAC-pipe (Zuo et al., 2019); these high-quality peaks had strong signal-to-noise ratios (Figures S1C and S1D) and reproducibility (Figures S1E and S1F). Confirming the plausibility of our analysis, our ATAC-seq analysis successfully detected focal enrichments for open chromatin around the *CD3D* and *CD4* loci (Figures S1G and S1H) but not for *CD8A* (Figure S1I). In addition, the number of accessible peaks approached saturation (~100,000) when the number of samples from SLE patients reached 60, indicating that the sample size of our cohort was sufficiently large to identify most of the DNA accessible sites in $CD4^+$ T cells from patients (Figure 1B). Thus, our ATAC-seq analysis of primary $CD4^+$ T cells yielded a large-scale and reliable dataset of genome-wide chromatin accessibility profiles for both healthy individuals and those with SLE.

Chromatin accessibility signatures of SLE $CD4^+$ T cells

We next explored the chromatin accessibility signatures related to SLE disease activity. We first performed a principal-component analysis (PCA) of all the ATAC-seq samples; the healthy control samples were closely clustered together, whereas the SLE patient samples were more diffusely distributed (Figure 1C, average Euclidean distance at the PCA space: controls = 0.09 versus patients = 0.17, $p < 10^{-88}$), indicating a higher extent of heterogeneity for the chromatin accessibility of primary $CD4^+$ T cells from SLE patients (Figures S1J and S1K). We observed no clear pattern that can distinguish between drug-treated SLE patients and drug-naive SLE patients (Figure S1L). This substantial heterogeneity among the SLE patient samples apparently masked obvious differences between the SLE patients and healthy controls per se, as, when we tried to identify uniform chromatin accessibility differences in SLE versus healthy control samples, we found no lupus- or autoimmune disease-related ontologies among the 3,563 sites with differential DNA accessibility (3.4% of total peaks) ($|\log_2$ fold change$| > 1$, false discovery rate [FDR] <0.05, $p < 0.05$; Figures S1M and S1N).

We therefore adopted an alternative approach to help identify informative chromatin accessibility signatures related to SLE disease activity. We extracted the most variable chromatin-accessible sites among all patients (coefficient of variation >0.5) and performed a PCA of the patient samples based on the intensity values for these peaks (STAR methods). The patient samples were clustered into three distinct groups (Figures 1D and 1E). Because these three groupings were supported by the detection of strong correlations with their respective patient disease activity (DA) scores (Figure 1F), we defined these patient groups as DA^{low}, DA^{int}, and DA^{hi}.

We also assessed potential impacts of a set of comorbidities in healthy controls and in DA^{low}, DA^{int}, and DA^{hi} patients (Figure S2A). We found that the DA^{hi} SLE patients showed severe disease activity with high proportions for most of the comorbidities, such as hematuria, leukopenia, pericarditis, and pyuria (Figure S2A). The DA^{int} SLE patients showed moderate disease activity, and the DA^{low} SLE patients had mild disease activity (Figure 1F). These results demonstrated that the three groupings classified based on the divergence of the chromatin accessibility of peripheral $CD4^+$ T cells were closely related to the clinical severity of SLE.

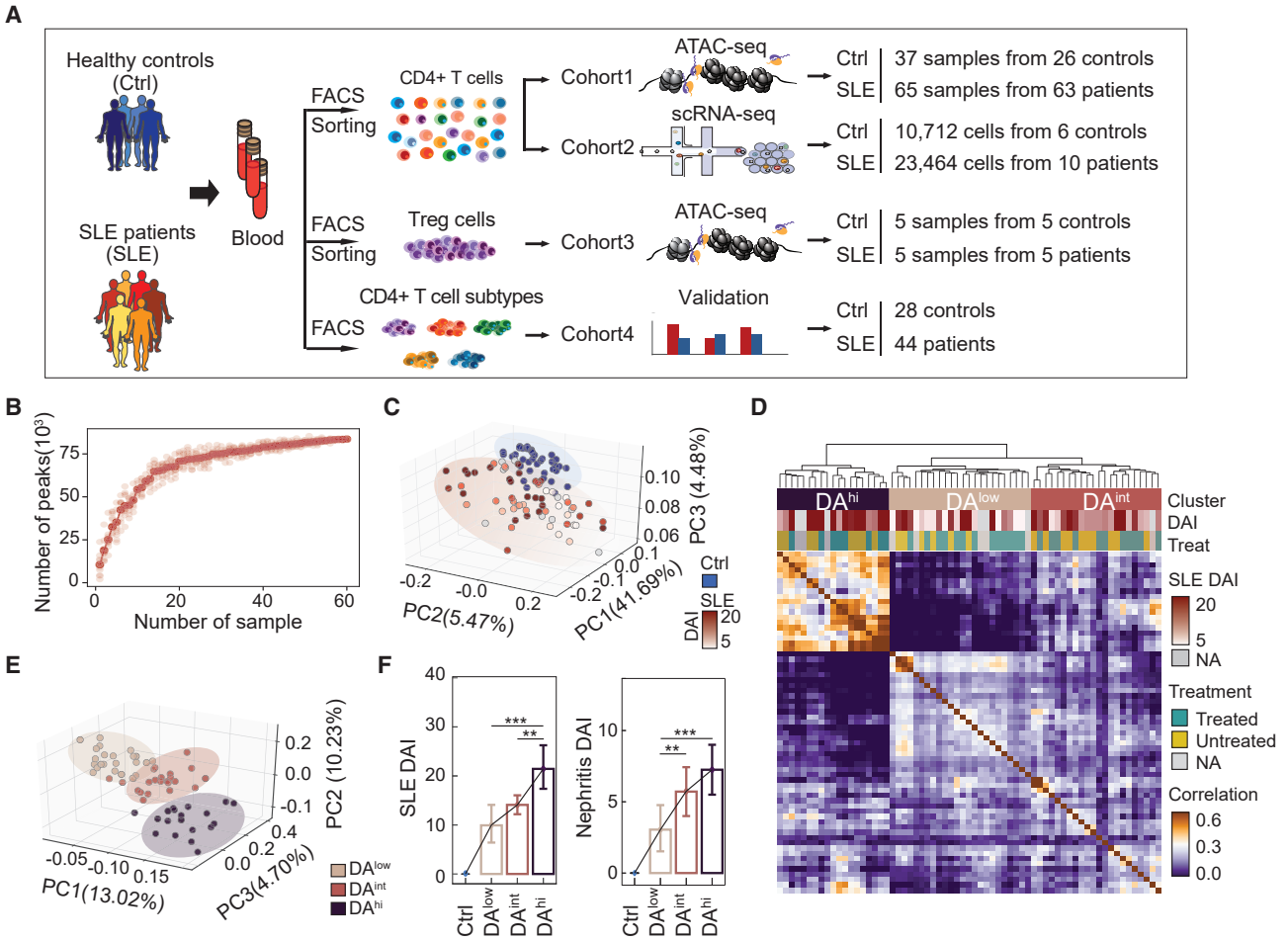


Figure 1. Landscape of the heterogeneity of DNA accessibility in SLE peripheral CD4⁺ T cells

(A) Flowchart depicting the overall design of the study. Blood draws were obtained from a total of 72 SLE patients (SLE) and 30 healthy controls. $CD4^+$ T cells were purified by FACS and used for assay for transposase-accessible chromatin sequencing (ATAC-seq) and single-cell RNA-seq (scRNA-seq) library construction. The ATAC-seq dataset comprised 37 samples from 26 healthy controls and 65 samples from 63 SLE patients. The scRNA-seq dataset comprised a total of 34,177 cells, with 10,711 cells from six healthy controls and 23,466 cells from 10 SLE patients.

(B) Saturation curve of the number of accessible peaks along with the number of samples. Each peak number represents the average of 10 random samplings in the indicated number of samples. The light red dots represent the 10 random sampling results.

(C) PCA of the top variant (coefficient of variation >0.2) chromatin-accessible elements for all ATAC-seq samples from healthy controls and SLE patients. Each dot represents a sample, and samples are color coded by their clinical status: healthy samples are marked in blue, and SLE samples are colored in red, with shading according to their SLE disease activity index (DAI).

(D) Heatmap showing the unsupervised hierarchical clustering of all SLE samples. Each square was color coded by the Pearson correlation between row and column sample according the peak intensity of the top variant chromatin-accessible elements. Color bars in the right panel indicate the range of SLE DAI values (top), the treatment information (middle), and the range of Pearson correlation coefficient values (bottom). Top: first row, samples were categorized into three groups, DA^{low}, DA^{int}, DA^{hi}; second row, the SLE DAI of each SLE patient; third row, the treatment information.

(E) PCA of the top variant (coefficient of variation >0.45) chromatin-accessible elements for all ATAC-seq samples from SLE patients. Each dot represents an SLE sample, and samples are color coded by their clinical groups.

(F) Bar plots showing the SLE DAI (left) and the nephritis DAI (right) in healthy controls and in DA^{low}, DA^{int}, and DA^{hi} SLE patients. $^{**}p < 0.01$, $^{***}p < 0.001$, Student's t test.

Next, we performed a pairwise differential analysis of each patient group and the healthy controls (STAR methods), thereby identifying a total of 11,774 DNA elements across the genome with significant differential chromatin accessibility (Figure 2A). Unsupervised hierarchical clustering grouped these differentially accessible peaks into five distinct clusters of regulatory elements (Figure 2A). Clusters I–III comprised chromatin regions that were more accessible in primary $CD4^+$ T cells from healthy controls, DA^{hi} and DA^{int} SLE patients, and DA^{int} and DA^{low} SLE patients, respectively. Cluster IV consisted of a small number of peaks that could distinguish the SLE patients from the healthy controls. In cluster V, the $CD4^+$ T cells from DA^{hi} SLE patients contained fewer accessible peaks than those from the other patient groups.

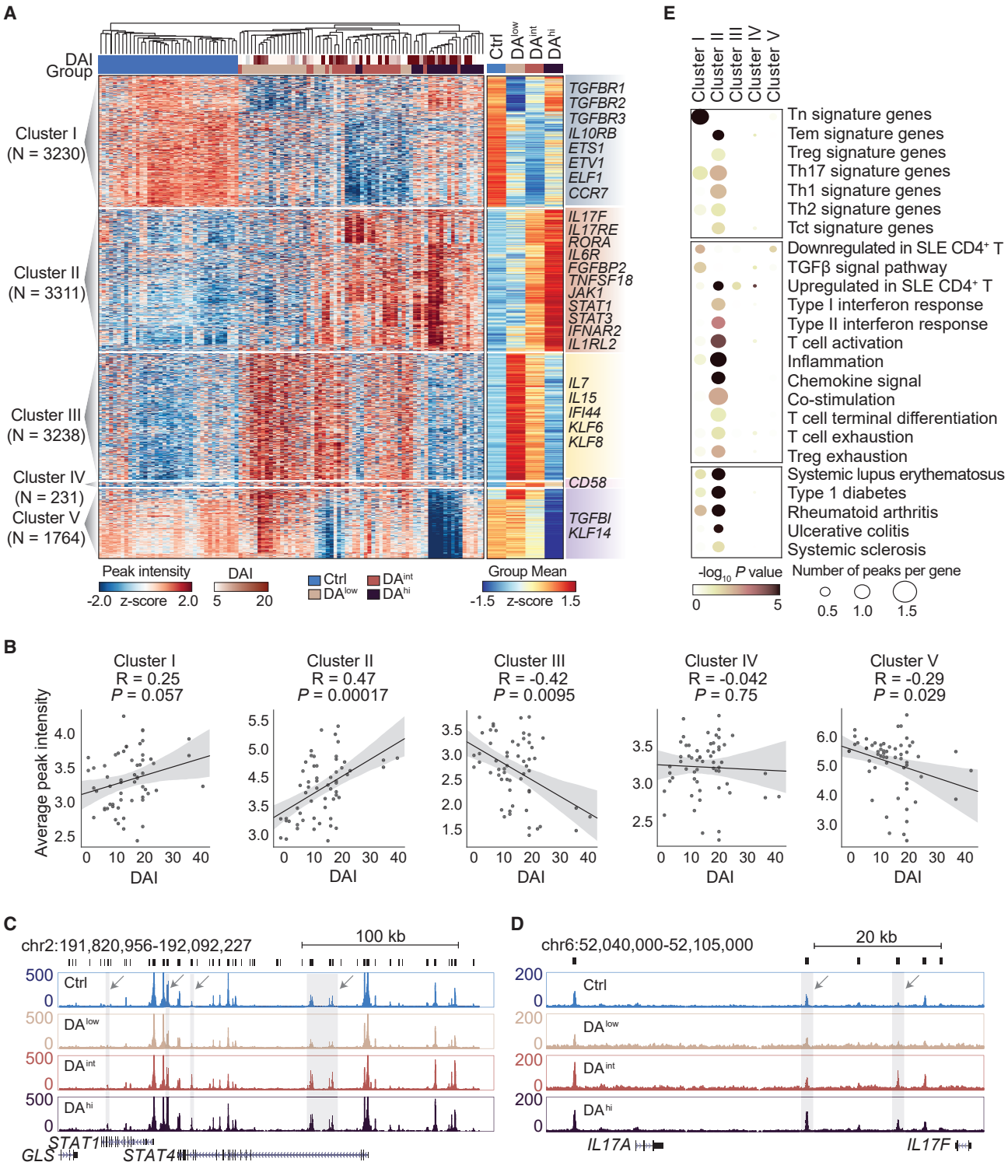


Figure 2. The epigenetic signatures of different groups of SLE patients

(A) Heatmap of differentially accessible peaks in CD4$^+$ T cells based on pairwise comparisons between healthy controls and DA^{low}, DA^{int}, and DA^{hi} SLE patients (false discovery rate [FDR] <0.05, $\log_2$ fold change >1.2). Each column is a sample, and each row is a differentially accessible peak. Samples and elements were organized by 2D unsupervised hierarchical clustering. The differentially accessible peaks were classified into five clusters (I–V). The color scale of the heatmap shows the relative ATAC-seq peak intensity. The right panel heatmap shows the average peak intensity and representative genes around the peaks (<2 kb) in normal controls and in DA^{low}, DA^{int}, and DA^{hi} SLE patients. Top: first row, samples are colored by their SLE DAIs; second row, samples are categorized into four clinical groups: normal controls and DA^{low}, DA^{int}, and DA^{hi} SLE patients.

(legend continued on next page)

To explore the relationship(s) between the distinct chromatin accessibilities of primary CD4$^+$ T cells and the clinical statuses of the patients, we performed a correlation analysis of the average intensity value of each peak cluster versus the DA of the SLE patients. We observed that the accessible peaks from cluster I showed a trend toward a positive correlation with the SLEDAI score ($p = 0.057$; Figure 2B). Notably, we found that the chromatin accessibility intensity of peaks in cluster II, but no other clusters, was significantly positively correlated with SLE DA ($p = 0.00017$, $R = 0.47$; Figure 2B). In contrast, the chromatin accessibility intensity of peaks in cluster III was significantly negatively correlated with SLE DA ($p = 0.0095$, $R = -0.42$; Figure 2B).

We next investigated the potential biological functions of cluster II peaks; as we were still aiming to identify chromatin accessibility signatures related to SLE, we used genomic regions enrichment of annotations tool (GREAT) (McLean et al., 2010) for functional annotation and conducted an enrichment analysis of the peak-associated genes from our analysis against SLE signature gene sets reported from previous SLE studies (Azizi et al., 2018; Hutcheson et al., 2008; Lowther et al., 2016; Smillie et al., 2019; Takeshita et al., 2015) (Table S2; STAR methods). Many genes known to regulate SLE tissue injury (Goropevsek et al., 2017), including *STAT1*, *STAT4*, and *IFNAR*—together with signature genes of Th17 cells, which are known as major contributors to SLE (Goropevsek et al., 2017), such as *IL17F*, *RORA*, and *IL6R*—were associated with cluster II peaks (Figure 2A). We also found that the cluster II peaks associated with the aforementioned genes had elevated chromatin accessibilities in DA^{hi} severe-stage SLE patients compared with DA^{int} and DA^{low} patients (Figures 2C, 2D, and S2B–S2D).

Furthermore, gene set enrichment and disease ontology analyses (STAR methods) showed that genes associated with cluster II peaks had strong enrichment for predicted functions related to the "inflammation" ($p < 10^{-5}$) and "IFN response" ($p < 10^{-2}$) terms. There was also enrichment for genes known to contribute to autoimmune diseases, including SLE ($p < 10^{-5}$), rheumatoid arthritis ($p < 10^{-5}$), and type 1 diabetes ($p < 10^{-5}$) (Figures 2E and S2E). Interestingly, genes involved in T cell exhaustion and Treg cell exhaustion (e.g., *PDCD1* and *TIGIT*) (Lowther et al., 2016; Yang et al., 2017) were highly enriched for cluster II peaks ($p < 10^{-2}$), which were more accessible in DA^{hi} and DA^{int} SLE patients (Figures 2E and S2F). These results suggested that the cluster II peaks represent an SLE disease chromatin signature associated with abnormal inflammation and immune responses in primary CD4$^+$ T cells. Increased accessibility of cluster II peaks may induce the dysfunction of CD4$^+$ T cell subtypes, such as Th17 cells and Treg cells, and promote the aggravation of SLE.

Single-cell atlas of CD4$^+$ T cells from severe-stage SLE patients

CD4$^+$ T cells are heterogeneous and comprise multiple subtypes, including naive, memory, effector, and Treg populations. Effector CD4$^+$ T cells can be further divided into cytokine-polarized T helper (Th) 1, Th2, and Th17 cells (Sallusto, 2016). Previous studies on blood from SLE patients have reported a decreased proportion of naive T (Tn) cells and an increased proportion of effector T (Teff) cells (Suarez-Fueyo et al., 2016). The data for altered Treg proportions in blood from SLE patients are inconclusive, and this idea remains controversial (Scheinecker et al., 2020). Chromatin accessibility analysis of our bulk ATAC-seq data indicated that particular CD4$^+$ T cell subtypes may be involved in SLE pathogenesis (Figures 2A and 2B). We therefore used the 10X platform to perform scRNA-seq analysis of the primary CD4$^+$ T cells from six healthy control individuals and 10 severe-stage SLE patients (average DAI of ~22), a different cohort from that in the bulk ATAC-seq analysis (Figure 1A; Table S1).

After quality control (QC) filtering (Figures S3A–S3D), we obtained a total of 34,176 high-quality single cells, of which 23,464 were from patients and 10,712 were from healthy controls. We then applied Seurat (Stuart et al., 2019) to integrate the cells from SLE patients and healthy controls and identified 16 clusters of CD4$^+$ T cell subtypes (Figures 3A and 3B). We also used Harmony (Korsunsky et al., 2019) to assess the accuracy and robustness of our single-cell clustering results and found strong correlations of the identified CD4$^+$ T cell clusters and gene expression patterns between the two integration methods (Figures S3E–S3I). These results support that we obtained a reliable and high-quality atlas of single-cell transcriptomes of CD4$^+$ T cells from severe-stage SLE patients and healthy controls.

Next, we analyzed known marker genes for particular CD4$^+$ T cell subtypes (STAR methods) and evaluated their expression profiles within each of the 16 cell clusters to identify the cell types in each cluster (Figure 3C). We identified seven major cell subtypes, namely CCR7$^+$ Tn, CXCR5$^+$ T, TBX21$^+$ Th1, GATA3$^+$ Th2, RORC$^+$ Th17, FOXP3$^+$ Treg, and GNLY$^+$ Tct cells (Figure 3D). Although the other clusters each fitted clearly within one of the seven major cell subtypes, the cells of cluster 15 appeared to be in a mixed state among the known CD4$^+$ T cell subtypes and were defined as "Tx", while the cells of cluster 16 exhibited a small number of genes and were named "undefined" (Figure S3J). Further, we obtained signature gene sets for multiple known CD4$^+$ T cell subtypes from previously reported studies (Azizi et al., 2018; Hutcheson et al., 2008; Lowther et al., 2016; Smillie et al., 2019; Takeshita et al., 2015) and found remarkable enrichment in the corresponding cell subtypes we

(B) Scatterplot representing the linear regressions of the SLE DAI and the average peak intensity of peaks from clusters I–V. Each dot is an SLE patient, and the upper and lower limits of the light gray shading represent the 95% confidence intervals. The two-tailed t-statistic p values and coefficients (R) of Pearson's correlation are shown at the bottom right.

(C and D) Normalized ATAC-seq profiles at the loci *STAT1* (C) and *IL17A* (D).

(E) Functional annotations of peaks from clusters I–V. Shown are the enrichment p values (dot color) and average peak numbers per gene (dot size) of predefined sets in peaks from clusters I–V. The gene sets were classified into three categories: cell subtype-specific gene signatures, SLE-associated functional gene sets, and autoimmune disease risk gene sets (Tn, naive CD4$^+$ T cells; Tem, effector memory CD4$^+$ T cells; Treg, regulatory CD4$^+$ T cells; Tct, cytotoxic CD4$^+$ T cells).

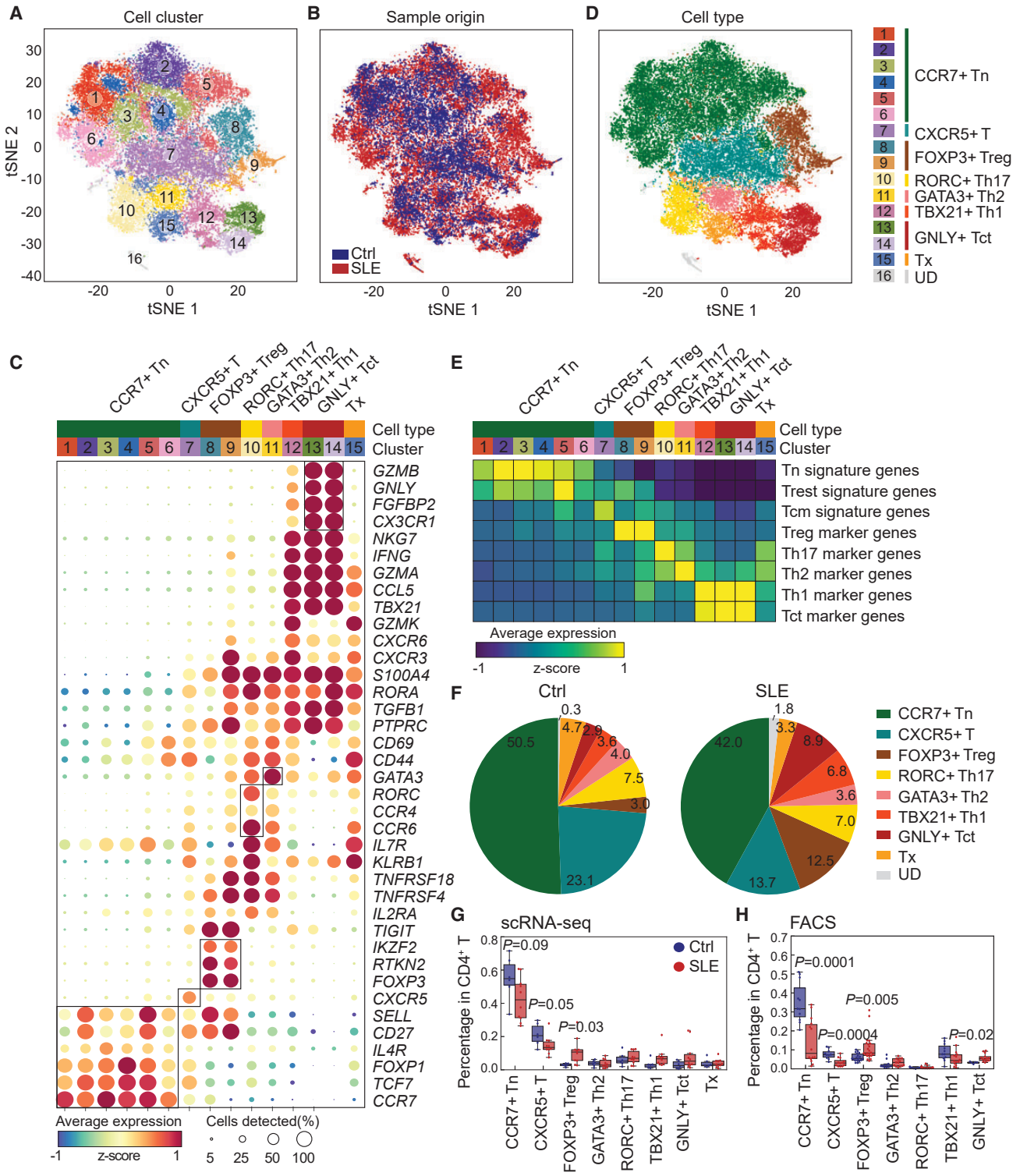


Figure 3. Single-cell atlas of peripheral CD4⁺ T cells from normal controls and SLE patients

(A and B) Two-dimensional t-distributed stochastic neighbor embedding (t-SNE) of the 34,098 CD4⁺ T cells profiled herein, with each cell color coded by the cell cluster (A) and sample type of origin (B).

(C) Heatmap showing the expression of selected marker genes for CD4⁺ T subtypes in the 15 cell clusters. The dot size represents the percentage of cells expressing the selected marker genes (right). The dot color represents the mean expression levels of the selected marker genes.

(D) t-SNE of the 34,098 profiled CD4⁺ T cells, with each cell color coded by subtype.

(legend continued on next page)

classified through our scRNA-seq analysis (Figure 3E), confirming our cell subtype identification.

We then explored the distributions of the seven major cell subtypes in samples from SLE patients and healthy controls. Multiple studies on PBMCs from SLE patients have reported a decrease in the proportion of naive $CD4^+$ T cells and an increase in effector $CD4^+$ T cells compared with cells from healthy controls (Suarez-Fueyo et al., 2016). We also observed decreases in the proportions of $CCR7^+$ Tn and $CXCR5^+$ T cells in SLE patients and noted increases in the proportions of multiple Teff cell subtypes, including Treg cells and cytotoxic $CD4^+$ T cells (Tct), by comparing the proportion of each cell type directly (Figure 3F).

We next applied Demuxlet (Kang et al., 2018) and Souporcell (Heaton et al., 2020) to deconvolute the cell sources and identified the sources of 15,301 (45%) and 31,752 (93%) sequenced single cells, respectively (STAR methods). We then calculated and compared the proportions of each major cell subtype in SLE patients with those in healthy controls and found that Treg cells were significantly increased in SLE patients ($p < 0.05$) (Figures 3G and S3K). Pursuing this, we used flow cytometry to measure the distributions of the seven major cell subtypes in $CD4^+$ T cells from the PBMCs of an independent cohort of 27 severe-stage SLE patients and 15 healthy controls (Figure S3L; Table S1) and again noted a remarkable and significant decrease in the proportion of $CCR7^+$ Tn cells in the SLE patients (2.44-fold, $p = 0.0001$) as well as significant increases in the proportions of Treg cells (1.82-fold, $p = 0.005$) and Tct cells (1.76-fold, $p = 0.02$) (Figure 3H).

Divergent changes in the single-cell transcriptomes of $CD4^+$ T cell subtypes from severe-stage SLE patients

To explore the functional divergence of each $CD4^+$ T cell subtype in the disease state, we performed pairwise comparisons of the $CD4^+$ T cell subtypes from SLE patients and the corresponding cell subtypes from healthy controls, which identified a total of 1,331 differentially expressed genes (DEGs) (Figures 4A and 4B; Table S3; STAR methods). Among these DEGs, 197 (28.67%) and 183 (28.42%) genes were up- and downregulated, respectively, in more than five cell subtypes (defined as "shared DEGs"), while 259 (37.70%) and 277 (43.01%) genes were up- and downregulated, respectively, in fewer than three cell subtypes ("specific DEGs") (Figures 4A, 4B, and S4A–S4C). When accounting for the number of specific DEGs for each of the seven major cell subtypes, there were clearly more specific DEGs that were either up- or downregulated in Treg, Th2, Th17, and Tct cells than in the other cell subtypes (i.e., naive $CD4^+$ T cells, $CXCR5^+$ T cells, Th1 cells) (Figures 4C, S4D, and S4E). Both Euclidean distance and Spearman correlation analyses indicated that the Treg cells showed the largest differences in gene expression between SLE patients and healthy controls (Figure 4D). These results indicate that Treg, Tct, Th17, and Th2 cells dominated the transcriptomic changes in $CD4^+$ T cells from severe-stage SLE patients.

Given that nuclear factor κB (NF-κB) signaling pathway genes are known to promote the inflammatory development and progression of SLE (Brightbill et al., 2018), we consistently observed strong expression of these genes, including *NFKB1*, *NFKB2*, and *REL*, in Th17 cells from severe-stage SLE patients (Figure S5A). In addition, strong expression of the effector molecule CCL20 was reported to mediate the Th17-driven inflammatory process and to be associated with SLE pathogenesis (Koga et al., 2016), and this gene was expressed at significantly higher levels in SLE Th17 cells than in healthy Th17 cells (Figure S5B; $p = 1.40e^{-110}$, Mann-Whitney U test). SLE Tct cells showed obviously increased levels of a series of cytotoxic molecules, including *GZMH*, *NKG7*, and *GZMB*, supporting previous claims about their potential involvement in exacerbating autoimmune diseases (Takeuchi and Saito, 2017) (Figure S5C).

Interestingly, our single-cell analysis revealed that quite a few T cell exhaustion-related genes, such as *TIGIT*, *PDCD1*, and *LAG3* (Wherry and Kurachi, 2015), were significantly upregulated in SLE Treg cells (Figures 4A, 4E, and S5D). We purified Treg cells from an additional cohort of healthy controls and severe-stage SLE patients (Table S1) and then performed qPCR and flow cytometry analysis of these genes, and found that each gene was expressed at a significantly higher level in the patients (Figures 4F and S5E; $p < 0.05$, Student's t test). We also found that the signature score of Treg exhaustion-like properties was higher in SLE Tregs than in normal Tregs (Figure 4G; $p = 2.01 \times 10^{-29}$). Chromatin accessibility analysis of the isolated Treg cells using ATAC-seq (Table S1) also indicated that these loci were more accessible in SLE patients than in healthy controls (Figures 4H–4J and S5F).

To further explore this apparent dysfunction of SLE Treg cells, we isolated Treg cells from another cohort of four healthy controls and four severe-stage SLE patients using flow cytometry (Table S1) and examined their abilities to inhibit Teff cells by a carboxyfluorescein succinimidyl ester (CFSE)-based T cell suppression assay. We found that Treg cells isolated from SLE patients showed lower suppressive activity than healthy controls (Figures 4K and S5G). Together, these results indicate that SLE Treg cells may undergo functional exhaustion and thus fail to suppress the over-reactive immune system. Moreover, Treg with exhaustion-like properties was also observed in the

(E) Heatmap showing the expression of the cell-type-specific signature genes in the 15 cell clusters.
(F) Pie chart showing the proportions (i.e., percentage of total sequenced $CD4^+$ T cell complement) of each $CD4^+$ T cell subtype in normal controls and SLE patients.
(G) Boxplot showing the percentages of $CD4^+$ T cell subsets in normal controls and SLE patients in our scRNA-seq data. Each dot is a normal control/SLE patient. The cells from each individual were identified by the published tool Demuxlet (Student's t test, the upper, center, and lower lines indicate the 75% quantile + 1.5 × interquartile range (IQR), 50% quantile, and 25% quantile − 1.5 × IQR, respectively).
(H) Flow cytometry analysis of a larger cohort of normal controls and SLE patients (cohort independent of those used in the scRNA-seq analysis), showing the proportions of $CCR7^+$ T cells, $CXCR5^+$ T cells, Treg cells, Th2 cells, Th17 cells, Th1 cells, and Tct cells among the gated $CD4^+$ T cells. (Student's t test, the upper, center, and lower lines indicate the 75% quantile + 1.5 × interquartile range (IQR), 50% quantile, and 25% quantile − 1.5 × IQR, respectively).

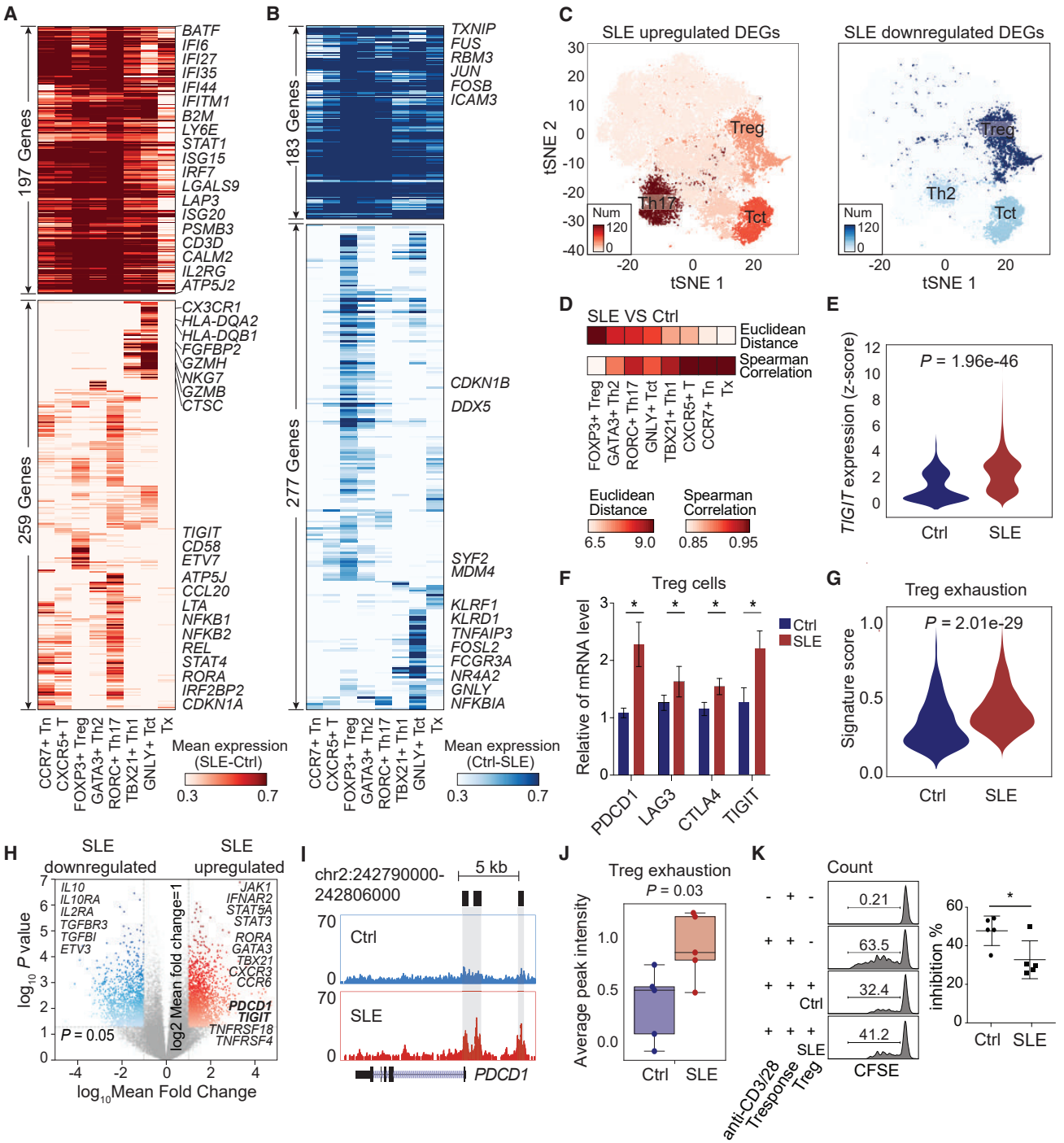


Figure 4. Compositional and functional differences between normal and SLE CD4⁺ T subtypes

(A and B) Heatmap showing the normalized expression of DEGs between normal controls and SLE patients in each of the CD4⁺ T cell subtypes. The upregulated and downregulated DEGs in SLE are shown in (A) and (B), respectively. The upper panels in (A) and (B) show the shared DEGs across all of the CD4⁺ T cell subtypes. The lower panels in (A) and (B) show the specific DEGs in each of the CD4⁺ T cell subtypes.

(C) t-SNE showing the numbers of upregulated DEGs (left) and downregulated DEGs (right) specific to SLE in all CD4⁺ T cell subtypes between normal controls and SLE patients.

(D) Heatmap showing the results of Euclidean distance (upper) and Spearman correlation (lower) analyses of gene expression between healthy controls and SLE patients in each cell subtype.

(E) Violin plot showing the expression of TIGIT in Tregs from healthy controls and SLE patients (Mann-Whitney U test).

(legend continued on next page)

inflamed intestinal mucosa of patients with ulcerative colitis (Smillie et al., 2019) (Figures S5H–S5K).

CCR7^{low}CD74^{hi} Treg cells from SLE patients are functionally exhausted

A previous study identified three phenotypically and functionally distinct human Treg cell subpopulations based on the expression of CD45RA and FOXP3: CD45RA^{+}FOXP3^{low} naive Treg (nTreg) cells, CD45RA^{-}FOXP3^{hi} effector Treg (eTreg) cells, and CD45RA^{-}FOXP3^{low} fraction (Fr. III) cells (Miyara et al., 2009). In our single-cell transcriptomes of CD4^{+} T cells from SLE patients and healthy controls, we identified two subpopulations of Treg cells, which we defined as Treg1 and Treg2, from a total of 3,239 cells (Figure 5A). We first obtained the cell subtype-specific genes of the nTreg, eTreg, and Fr. III cells from published microarray datasets (Cuadrado et al., 2018) (Figure S6A; STAR methods). Correlation analysis of the expression of these specific genes between Treg1 and Treg2 cells with nTreg, eTreg, and Fr. III cells (Figure S6B) indicated that Treg1 cells typically exhibited marker expression profiles similar to that of nTreg cells, whereas the Treg2 cell profiles appeared more similar to eTreg and Fr. III cells (Figure S6B).

We then utilized Monocle3 (Cao et al., 2019) to perform a pseudotime analysis of the differentiation potential of Treg cell subsets from healthy controls and SLE patients to investigate the specific trajectories of Treg1 and Treg2 cells. We found that Treg2 trajectory-enriched cells strongly expressed genes related to T cell exhaustion and Treg cell exhaustion (Figure 5B), while Treg1 trajectory-enriched cells exhibited higher levels of lineage-specific genes for Th1, Th2, Th17, and Tct cells (Figures S6C, S6D, and Table S4).

Among the 1,064 significant DEGs between Treg1 and Treg2 cells (Table S4), we identified two putative surface marker genes, *CCR7* and *CD74*, that could be used to distinguish the two Treg subpopulations (Figures 5C–5E). Flow cytometry analysis indicated that the proportions of both Treg1 (CCR7^{hi} CD74^{low}) and Treg2 (CCR7^{low}CD74^{hi}) cells were significantly higher in severe SLE patients than in healthy controls (Figure 5F; Student's t test, *p <0.05). There were 830 and 818 DEGs between healthy controls and SLE patients for Treg1 and Treg2 cells, respectively (Figures 6A and 6B; Table S4). We noted that the SLE Treg2 cells highly expressed signature genes of T cell exhaustion, such as *PDCD1* and *LAG3* (Figures 6C and S6E–S6G) (Lowther et al., 2016). We validated the expression trends for these genes in Treg cells from healthy controls and SLE patients by qPCR and flow cytometry (Figures 6D and S6H). These results suggested that CCR7^{low}CD74^{hi} Treg2 cells, which are an effective Treg subpopulation in SLE patients, may undergo functional exhaustion that weakens their immunosuppressive function.

We then examined peripheral blood samples from SLE patients using flow cytometry to detect the secretion of IFN-γ, IL-2, and tumor necrosis factor alpha (TNFα) in CCR7^{hi}CD74^{low} Treg1 and CCR7^{low}CD74^{hi} Treg2 cells upon phorbol 12-myristate 13-acetate (PMA)/ionomycin stimulation. In contrast to the characteristics of CD8^{+} T cell exhaustion, which exhibits an insufficient ability in producing IFN-γ, IL-2, and TNFα, we found significantly higher proportions of IFN-γ ($p < 0.05$, Student's t test), IL-2 ($p < 0.01$, Student's t test), and TNFα (Figure 6E; $p < 0.01$, Student's t test) production by Treg2 cells compared with Treg1 cells in SLE patients. As previous studies have shown that these cytokine-producing Treg cells are phenotypically Th1-like Tregs, which contribute to the observed Treg dysfunction in autoimmune diseases (Dominguez-Villar et al., 2011; Korn et al., 2007; McClymont et al., 2011), this result provides further protein-level support for the functional exhaustion-like properties of the CCR7^{low}CD74^{hi} Treg2 cells in SLE patients.

We also assessed peripheral blood samples from SLE patients using flow cytometry to detect the expression of exhaustion-associated transcription factors (Tcf1 and Tox) in CCR7^{hi}CD74^{low} Treg1 and CCR7^{low}CD74^{hi} Treg2 cells upon *in vitro* PMA/ionomycin stimulation. We observed that there were no significant differences in the expression of Tcf1 between Treg1 and Treg2 cells (Figure 6F). Notably, we found that Treg2 cells displayed significantly higher expression levels of Tox than Treg1 cells (Figure 6F; $p < 0.05$, Student's t test). These data are consistent with the findings from previous reports showing Tox as a promoting factor for T cell exhaustion (Kim et al., 2020; Scott et al., 2019) and provide additional support for the conclusion that CCR7^{low}CD74^{hi} Treg2 cells are Tregs with exhaustion-like properties.

Excessive type I IFN production promotes Treg cell exhaustion in SLE patients

A previous study showed that chronic stimulation by persistent antigens or inflammatory signals is sufficient to drive the exhaustion and functional disability of T cells (Rome et al., 2020). Therefore, we sought to investigate whether persistent stimulation of inflammatory signals can induce Treg cell exhaustion. We first defined a Treg cell exhaustion signature score (TES) as the average expression of genes involved in Treg cell exhaustion and then ranked the genes whose expression was significantly correlated with the TES score across all Treg cells (Figure 7A; Pearson's correlation >0.4, p <0.05). We performed functional annotation analyses of these genes by Metascape (Zhou et al., 2019) and gene set enrichment analysis (GSEA) (Mootha et al., 2003) according to the correlation coefficients between the

(F) qPCR showing the expression of *TIGIT*, *CTLA4*, *PDCD1*, and *LAG3* in Tregs from healthy controls (n = 8) and SLE patients (n = 8). *$p < 0.05$, Student's t test.
(G) Violin plot showing the signature score of Treg exhaustion-like properties in Tregs from healthy controls and SLE patients (Mann-Whitney U test).
(H) Volcano plot showing the significantly differential (Student's t test $p < 0.05$, $|\log_2$ fold change$| > 1$) peaks between healthy controls and SLE patients; genes whose promoters were nearest to these peaks are shown.
(I) Normalized ATAC-seq profiles in Treg cells from normal controls (blue) and SLE patients (red) at the loci of *PDCD1*.
(J) Boxplots showing the chromatin accessibility signature scores of Treg exhaustion-like properties in Treg ATAC-seq samples from healthy controls and SLE patients (Student's t test; the upper, center, and lower lines indicate the 75% quantile + 1.5 × interquartile range (IQR), 50% quantile, and 25% quantile − 1.5 × IQR, respectively).
(K) The ratios of the suppressive function of Treg cells from healthy controls and SLE patients.

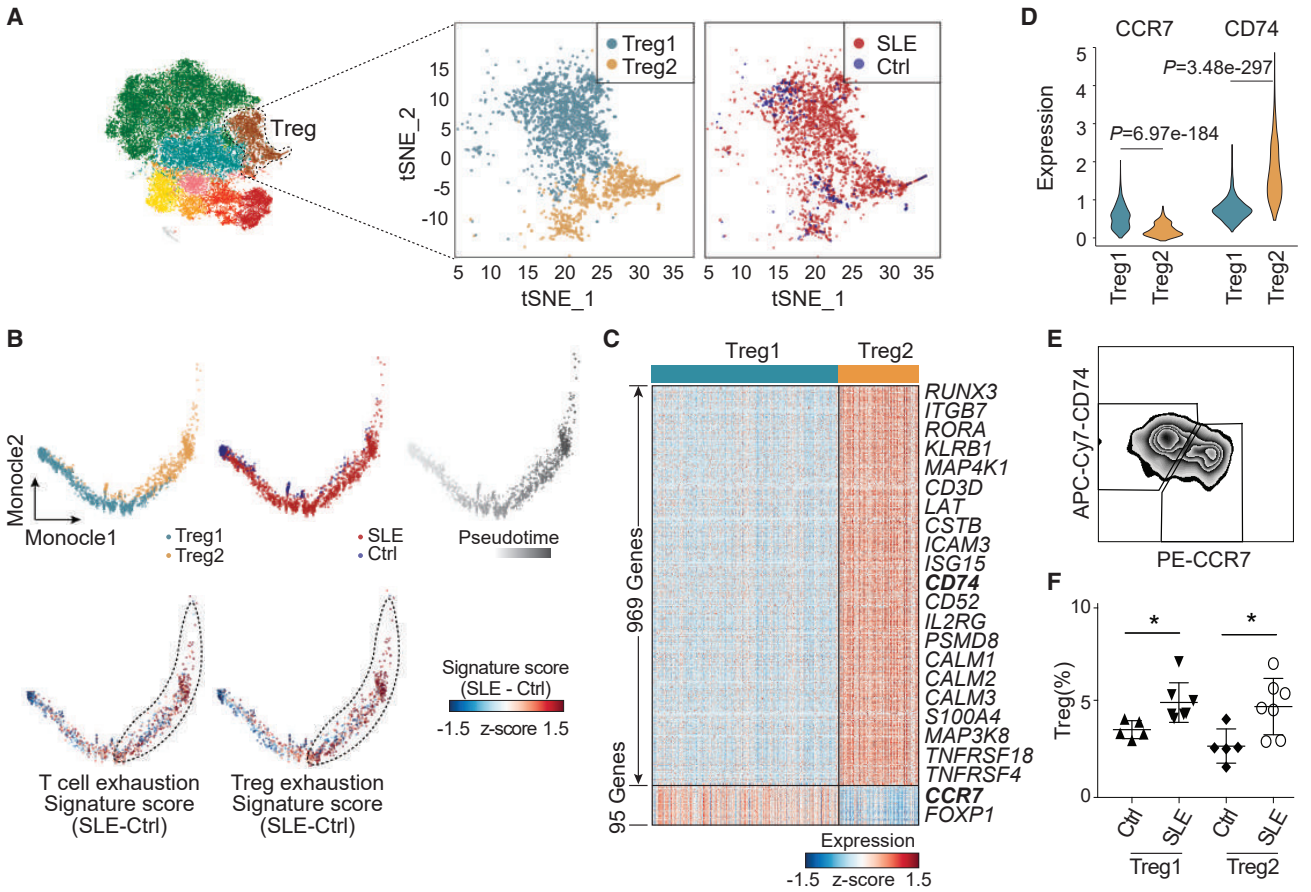


Figure 5. Treg cell subsets and their functional divergence in SLE patients

(A) t-SNE of 3,239 Treg cells profiled in this study, with each cell color coded by the cell cluster and sample type of origin.

(B) Trajectory of Treg cells determined using the Monocle 2 algorithm. The cell clusters, sample types of origin, pseudotimes, and differences in the signature scores of T cell exhaustion and Treg exhaustion-like properties between SLE patients (SLE) and healthy controls (Ctrl) are colored.

(C) Heatmap showing the normalized expression of the DEGs (Mann-Whitney U test $p < 0.01$, FDR < 0.05, expression difference >0.4) between Treg cells from healthy controls and SLE patients.

(D) Violin plot showing the normalized expression of *CCR7* and *CD74* in Treg1 and Treg2 cells from healthy controls and SLE patients.

(E) Representative contour plot showing *CCR7* and *CD74* expression in gated $CD4^+CD25^+CD127^-$ Treg cells isolated from SLE patients. Data are representative of at least five independent experiments.

(F) The percentages of $CCR7^{hi}CD74^{low}$ Treg1 cells and $CCR7^{low}CD74^{hi}$ Treg2 cells in $CD4^+$ T cells from normal controls and SLE patients (*$p < 0.05$, Student's t test).

expression of genes and TES (STAR methods). The genomic features of IFN signaling (Figure 7B) and response to type I IFNs (Figure 7C) were significantly enriched. The average expression levels of genes associated with Treg cell exhaustion and genes involved in the type I IFN signaling pathway were also found to be highly correlated along the Treg cell pseudotime trajectory in SLE patients (Figures 7D–7F and S7A). In addition, the chromatin accessibility level of Treg cells between the exhaustion signature genes and type I IFN genes showed a close correlation (Figure S7B). We also found that the expression of IFN-induced genes was higher in SLE Tregs than in Th1, Th2, Th17, and Tct cells, suggesting that Treg cells from SLE patients are more sensitive to IFN stimulation than Th1, Th2, Th17, and Tct cells (Figures S7C–S7F). We therefore hypothesize that persistent stimulation of type I IFN signaling may promote the Treg cell exhaustion observed among the primary $CD4^+$ T cells of SLE patients.

Pursuing this, we purified primary Treg cells from healthy controls and performed transcriptome sequencing analysis of Treg cells that were stimulated with different concentrations of type I IFN via multiple protocols. As expected, we noticed that the expression of genes involved in the type I IFN signaling pathway was successfully induced (Figure 7G). We also found that persistent type I IFN stimulation was sufficient to induce the expression of Treg exhaustion-associated genes and genes upregulated in SLE Treg2 cells (Figure 7G; Table S4; STAR methods). The average expression of Treg exhaustion-associated genes in each sample was also found to be significantly positively correlated with that of genes involved in the type I IFN signaling pathway ($p = 0.047$, $R = 0.5$; Figures 7H and S7G), suggesting that type I IFN signaling

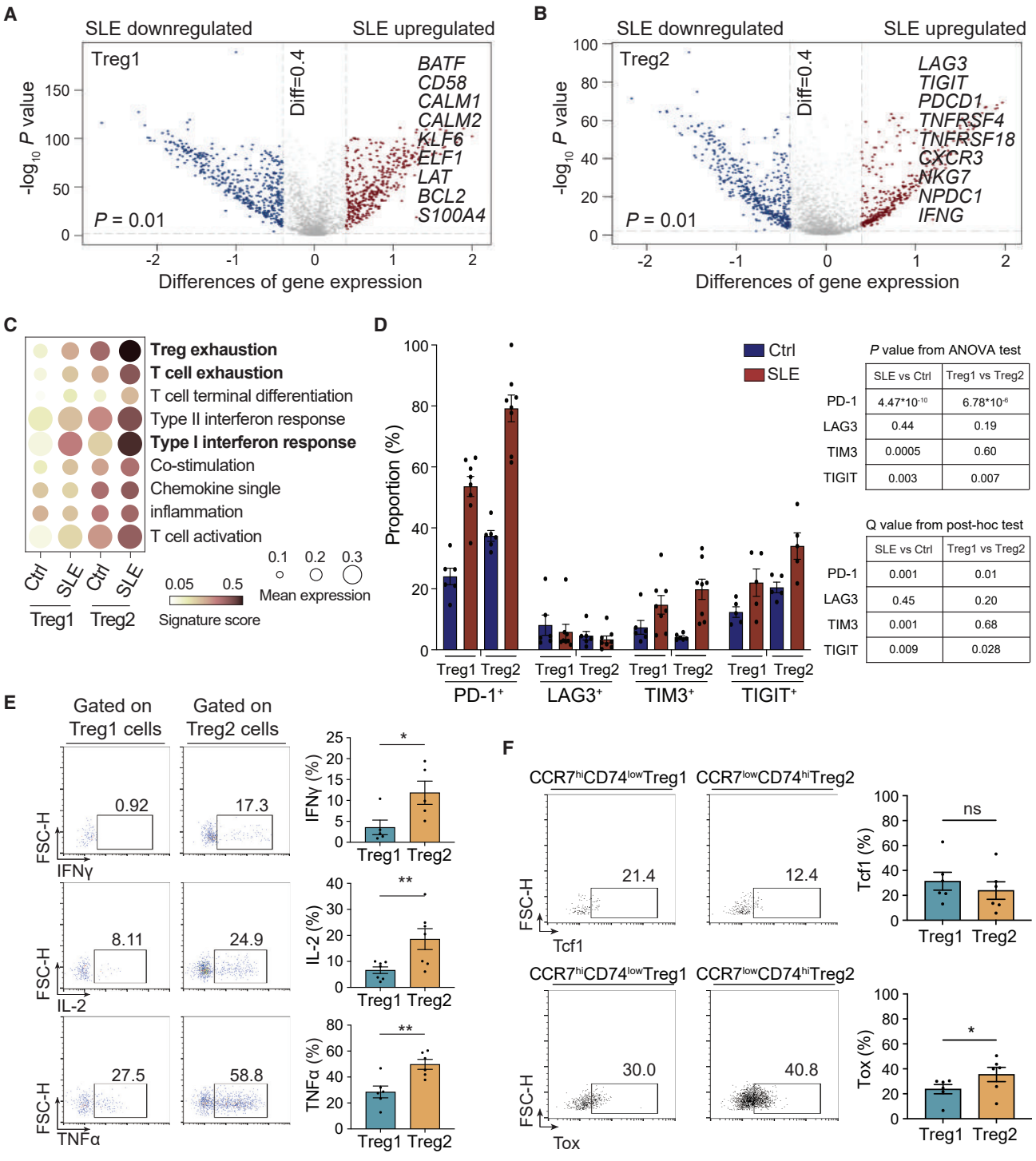


Figure 6. Exhaustion-like properties in CCR7^{low}CD74^{hi} Treg2 cells

(A and B) Volcano plot showing the DEGs (Mann-Whitney U test $p < 0.01$, FDR < 0.05, expression difference >0.4) between healthy controls in Treg1 (A) and Treg2 cells (B).

(C) Dot plot showing the enrichment of selected gene sets in Treg1 cells and Treg2 cells from healthy controls and SLE patients. The dot color represents the enrichment score, and the dot size represents the average gene expression of selected gene sets.

(D) Bar plots displaying the frequencies of PD-1, LAG3, TIM3, and TIGIT in CCR7^{hi}CD74^{low} Treg1 cells and CCR7^{low}CD74^{hi} Treg2 cells from SLE patients and healthy control individuals. ANOVA followed by *post hoc* multiple comparisons test.

(legend continued on next page)

is responsible for the Treg exhaustion-like properties in SLE patients. In addition, a CFSE-based T cell suppression assay indicated that IFNα-treated Treg cells decreased their capacity to inhibit the proliferation and activation of Teff cells (Figure 7I), further supporting that persistent exposure to activated IFN signaling can induce Treg cell exhaustion. In summary, we found that the Treg cells from SLE patients exhibited exhaustion characteristics that were related to IFN signaling and verified *in vitro* that chronic IFN stimulation could induce Treg dysfunction.

DISCUSSION

SLE is a complex systemic disease characterized by a wide spectrum of clinical manifestations and has substantial interindividual heterogeneity (Lisnevskaia et al., 2014). $CD4^+$ T cells from the peripheral blood are an essential contributor and indicator of immune overactivation (Caielli et al., 2018), and the chromatin accessibility of this cell type is highly informative for their identity, activity state, and regulatory programs (Qu et al., 2017). In the present study, we used ATAC-seq to survey the landscape of active regulatory DNA in a single cell type ($CD4^+$ T) sorted from peripheral blood samples of a large cohort of SLE patients. Further analysis of data from this large cohort identified three groups of patients with divergent transcriptional regulatory patterns, and these same groups were clustered similarly based on their DAI. Notably, the SLE disease signature evident in chromatin accessibility patterns was most obvious in severe-stage patients; perhaps patterns in patients with mild or moderate DA can be identified in studies with larger patient cohorts. Our analysis detected a correlation between the epigenome of $CD4^+$ T cells and SLE clinical states and provided a rich empirical foundation from the perspective of epigenetic regulation for understanding SLE heterogeneity.

Previous studies have reported an expansion of PD-1^+ $CXCR5^+CD4^+$ T (Tfh) cells and PD-$1^+CXCR5^-CD4^+$ T (Tph) cells in SLE patients (He et al., 2013; Kim et al., 2018; Makiyama et al., 2019). However, researchers have paid minimal attention to changes in the proportion of PD-$1^-CXCR5^+CD4^+$ T cells in disease states. The flow cytometry analysis of blood samples from lupus patients showed that the patients exhibited higher proportions of Tfh and Tph and a lower proportion of PD-1^- $CXCR5^+$ $CD4^+$ T cells than healthy control individuals (Lin et al., 2019). In line with these results are data from peripheral blood of patients with Sjögren's syndrome showing diminished CXCR5 expression in T cell subsets (Aqrawi et al., 2018). Therefore, the finding of a significant decrease in the fraction of $CXCR5^+$ T cells in SLE patients in our study may be associated with an expansion of Tph cells. In addition, CXCR5 is also expressed on 20%–25% of peripheral blood human central memory $CD4^+$ T cells, and these cells are a heterogeneous pool consisting of functionally distinct Th1-, Th2-, and Th17-like subsets (Chevalier et al., 2011). The reduction of $CXCR5^+$ T cells in SLE patients may also be attributed to the activation of these central memory $CD4^+$ T cells that transformed into effector T cells, supported by the results that SLE patients exhibited higher proportions of Th1, Th2, Th17, and Tct cells compared with healthy control individuals (Figures 3G and 3H). Nevertheless, biological function of the $CXCR5^+$ cells, especially the PD-$1^-CXCR5^+CD4^+$ T cells, remains to be further investigated.

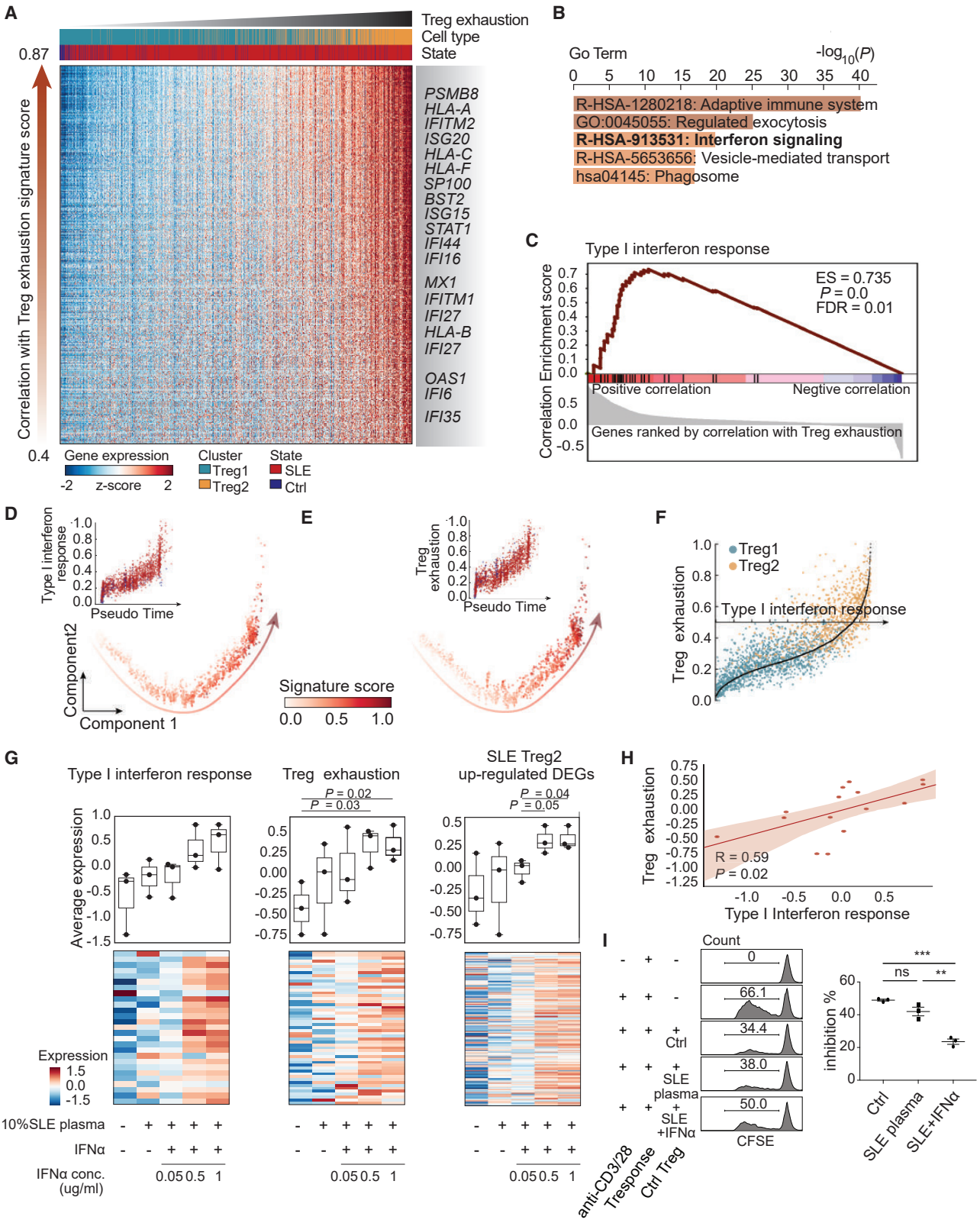
Since Treg cells function in maintaining immune homeostasis (Dominguez-Villar and Hafler, 2018; Sakaguchi et al., 2020; Wing et al., 2019), it seems counterintuitive that Treg numbers were increased in patients with an autoimmune disease. We explored the heterogeneity and dysfunction of SLE Treg cells in SLE patients and healthy controls, identified two distinct subsets of Treg cells, and confirmed their existence by flow cytometry. In healthy controls, these two subsets have polarized cell activities, separately displaying characteristics of naive and effector Treg cells. While both subsets of Treg cells were significantly increased in SLE patients, it was interesting to note that both SLE Treg1 and Treg2 cells in SLE patients adopted an effector-like state, with SLE Treg2 cells tending to display markers of functional exhaustion. ATAC-seq data also supported that this SLE Treg2 exhaustion phenotype results from epigenetic regulation. In addition, when we expanded the scope of our analysis to other autoimmune diseases by examining scRNA-seq? data for Tregs from patients with ulcerative colitis, we again observed upregulated Treg exhaustion-associated genes (Figures S5J–S5M), indicating that these Treg2-like exhausted cells may be features of multiple autoimmune diseases.

High levels of PD-1 expression on human Treg cells were described as a dysfunctional Treg exhaustion phenotype (Lowther et al., 2016). These exhausted Tregs exhibited high secretion of IFN-γ and were identified in healthy individuals and were enriched in tumor infiltrates (Lowther et al., 2016). Blockade of PD-1 activity reinvigorated the regulatory ability of dysfunctional Treg cells (Yang et al., 2017). In the present study, we identified $CCR7^{low}CD74^{hi}$ Treg cells with exhaustion-like properties (e.g., high expression of *PD-1*, *TIM3*, *TIGIT*, *IFN-γ*, *IL-2*, *TNFα*, *Tox*) that enriched in severe-stage SLE patients, as measured by scRNA-seq and supported by protein-level examination results. Treg exhaustion phenotype was also accompanied by several properties, such as an increase in pFoxo1 (Ser319), shortened telomere length, and decreased telomere-specific demethylated region (TSDR) (Lowther et al., 2016). However, the limited number of $CCR7^{low}CD74^{hi}$ Treg cells from human peripheral blood hindered the further in-depth investigation of the phenotypic and functional exhaustion of this Treg cell subpopulation.

Several studies have characterized the impact of type I IFNs on the function of Treg cells in a variety of contexts, leading to contrasting results (Piconese et al., 2015). It has been shown

(E) Expression of IFNγ, IL-2, and TNFα in gated Treg1 cells and Treg2 cells from SLE patients. Flow cytometry analysis of IFN-γ, IL-2, and TNFα expression in Treg1 cells and Treg2 cells from SLE patients with PMA and ionomycin stimulation *in vitro* (left). Bar plots displaying the frequencies of IFN-γ^+, IL-2^+, and TNFα^+ in Treg1 cells and Treg2 cells from SLE patients with PMA and ionomycin stimulation *in vitro* (right). *$p < 0.05$, **$p < 0.01$, Student's t test.

(F) Expression of Tcf1 and Tox in gated Treg1 cells and Treg2 cells from SLE patients. Flow cytometry analysis of Tcf1 and Tox expression in Treg1 cells and Treg2 cells from SLE patients with PMA and ionomycin stimulation *in vitro* (left). Bar plots displaying the frequencies of Tcf1$^+$ and Tox$^+$ in Treg1 cells and Treg2 cells from SLE patients with PMA and ionomycin stimulation *in vitro* (right). *$p < 0.05$, Student's t test.



(legend on next page)

that type I IFN signaling regulates the dynamic balance between human-activated regulatory and effector T cells and attenuates Treg cell function in viral infection and in the tumor microenvironment (Bacher et al., 2013; Gangaplara et al., 2018; Golding et al., 2010; Hashimoto et al., 2014; Srivastava et al., 2014). In line with these studies, we found that type I IFN signaling is responsible for the Treg exhaustion-like properties and Treg dysfunction in SLE patients, as shown by the fact that type I IFN stimulation was sufficient to induce the expression of Treg exhaustion-associated genes and genes upregulated in SLE Treg2 cells (Figure 7G). IFN-mediated positive effects on Treg function have also been reported under stress conditions (Lee et al., 2012; Metidji et al., 2015). This is likely because chronic or sustained exposure to type I IFN signaling may give rise to an opposite effect on Treg homeostasis and functions (Piconese et al., 2015).

The use of glucocorticoids or other immunosuppressive agents is still the mainstay of SLE management (Kiriakidou, 2013). However, these agents have substantial adverse effects and do not typically confer adequate therapeutic efficacy (Lisnevskaia et al., 2014). A recent clinical trial of anifrolumab, a human monoclonal antibody that targets the type I IFN receptor, showed a significant therapeutic effect in treating SLE (Morand et al., 2020). Our results also support type I IFN as an effective drug target for the treatment of severe SLE patients and provide a potential treatment strategy for this disorder, suppression of Treg with exhaustion-like properties. Overall, we delineated the chromatin accessibility and single-cell transcriptome atlases of $CD4^+$ T cells from SLE patients and healthy controls and found that chronic IFN signaling pathway induction may induce peripheral Treg cell exhaustion in SLE patients. Our study provides a rich source of data offering an epigenetic and transcriptomic view of $CD4^+$ T cell-related functions in terms of SLE clinical manifestations and immunopathogenesis, supporting a deeper understanding of this autoimmune disease.

Limitations of the study

Although there have been a few papers supporting an exhausted phenotype in human Treg cells (Lowther et al., 2016; Yang et al., 2017), the concept of Treg exhaustion has not yet been widely recognized. A major limitation of this study is the lack of protein-level intervention experiments, such as knockout or CRISPR of exhaustion-related genes such as PD-1, IFNAR1, and LKB1 in primary cells, and test the cellular functions thereafter. Such experiments are technically not implemented due to the limited number of $CCR7^{low}CD74^{hi}$ Treg cells and the fragile state of human primary cells. Thereby, further studies are necessary to uncover the molecular features, particularly the exhaustion-like properties, of the human $CCR7^{low}CD74^{hi}$ Treg cells.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human subjects
- METHOD DETAILS
  - Cell isolation
  - Intracellular staining
  - Single-cell RNA-seq library preparation and sequencing
  - ATAC-seq library preparation and sequencing
  - Assessment of the inhibitory function of Treg cells
  - *In vitro* stimulation of Treg with IFN-α
  - Bulk RNA-seq library preparation and sequencing
  - Real-time quantitative polymerase chain reaction
  - ATAC-seq primary data processing and peak calling
  - Normalization of ATAC-seq profiles
  - Saturation curve of the sample and peak numbers
  - Correlation analysis of all SLE patients

Figure 7. Chronic type I IFN stimulation promotes Treg cell exhaustion in SLE patients

(A) Heatmap showing the normalized expression of genes whose expression was correlated with the signature score of Treg exhaustion-like properties in all Treg cells. Each row depicts a gene, and each column depicts a cell. Genes were ranked by the Pearson's correlation of their expression and signature score of Treg exhaustion-like properties. Cells are ranked by their signature score of Treg exhaustion-like properties. Top: first row, each cell is colored by the cell subtype; second row, each cell is categorized by the sample type of origin.

(B) Functional annotation of the genes shown in (A) by the online tool Metascape.

(C) Genes were ranked by their gene expression correlation with the signature score of Treg exhaustion-like properties among all Treg cells. Gene set enrichment analysis (GSEA) showed the enrichment of genes in response to type I IFN among the above preranked genes.

(D and E) Trajectory of Treg cells using the Monocle 2 algorithm. The signature scores of the type I IFN response (D) and Treg exhaustion-like properties (E) are colored. The attached panel at the top left corner shows the scatterplot of pseudotime and the signature scores of the type I IFN response (D) and Treg exhaustion-like properties (E).

(F) Scatterplot showing the signature score of Treg exhaustion-like properties. Each dot is a Treg cell, and Treg cells are ranked by the IFN response signature score along the x axis.

(G) Tregs from healthy controls were stimulated with different concentrations of IFNα and SLE plasma. The boxplots in the upper panel and heatmap in the lower panel show the average expression of genes involved in the type I IFN response, Treg exhaustion-like properties, and SLE Treg2 upregulated DEGs (IFN-induced genes were excluded from this gene list). Each dot in the boxplot is an RNA sequencing sample, and each row in the heatmap is a gene (Student's t test; the upper, center, and lower lines indicate the 75% quantile + 1.5 × interquartile range [IQR], 50% quantile, and 25% quantile −1.5 × IQR, respectively).

(H) Scatterplot representing the linear regressions between the signature scores of the type I IFN response and Treg exhaustion-like properties. The two-tailed t-statistic p value and Pearson's correlation coefficient (*R*) are shown at the bottom.

(I) The ratio of the suppressive function of Treg cells from healthy controls (Ctrl) with persistent stimulation of 10% SLE plasma (SLE plasma) and 0.5 μg/mL IFNα plus 10% SLE plasma (SLE + IFNα) for 7 days.**$p < 0.01$, ***$p < 0.001$, ns not significant, Student's t test.

- Summary of the clinical information for the three SLE patient groups
- Acquisition of microarray and RNA-seq datasets
- Acquisition of gene sets
- Screening the differential peaks between three SLE patient groups and healthy controls
- Functional annotation of peak clusters
- scRNA-seq primary data processing and gene expression imputation
- Cell clustering
- Determination of the sample identity for each cell type
- Identifying the statistically significant differences in cell proportions
- Scoring gene signatures based on gene expression and chromatin accessibility
- Screening the differentially expressed genes between healthy controls and SLE patients for each $CD4^+$ T cell subtype
- Construction of the functional change profiles in all cell subsets
- Estimating the differences in gene expression patterns between Treg cells from healthy controls and SLE patients
- Screening of differentially expressed genes between two Treg clusters (Treg1 and Treg2)
- Identifying the SLE Treg2 up-regulated DEGs
- Identifying the signature genes for three published Treg groups
- Screening specific genes for Th1, Th2, Th17, and Tct cells
- Identifying the Treg exhaustion-associated functional gene set

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.111606.

ACKNOWLEDGMENTS

Funding: this work was supported by the National Key R&D Program of China (2020YFA0112200 to K.Q.), the National Natural Science Foundation of China grants (T2125012, 91940306, 31970858, and 31771428 to K.Q.; 32270978 to C.G.), CAS Project for Young Scientists in Basic Research (YSBR-005 to K.Q.), Anhui Province Science and Technology Key Program (202003a07020021 to K.Q.), and the Fundamental Research Funds for the Central Universities (YD2070002019, WK9110000141, and WK2070000158 to K.Q.). We thank the USTC supercomputing center and the School of Life Science Bioinformatics Center for providing computing resources for this project.

AUTHOR CONTRIBUTIONS

K.Q. conceived and supervised the project. C.G. and D.Z. performed the experiments and conducted all the sample preparation for next-generation sequencing with help from Q.S., L.Z., Y.L., Y.S., and X.G. Q.L. performed the data analysis with help from W.Z., Z.Z., Q.Y., J.F., P.D., and P.C. W.Z. also contributed to the revision of the manuscript. X.L., Q.W., and J.T. provided blood samples from SLE patients and healthy control individuals. K.Q., C.G., and Q.L. wrote the manuscript with the help of all other authors.

DECLARATION OF INTERESTS

J.F. is the chief executive officer of HanGene Biotech. The other authors declare no competing interests.

Received: March 24, 2022
Revised: August 2, 2022
Accepted: October 14, 2022
Published: November 8, 2022

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Aqrawi, L.A., Ivanchenko, M., Björk, A., Ramírez Sepúlveda, J.I., Imgenberg-Kreuz, J., Kvarnström, M., Haselmayer, P., Jensen, J.L., Nordmark, G., Chemin, K., et al. (2018). Diminished CXCR5 expression in peripheral blood of patients with Sjogren's syndrome may relate to both genotype and salivary gland homing. Clin. Exp. Immunol. *192*, 259–270.

Arazi, A., Rao, D.A., Berthier, C.C., Davidson, A., Liu, Y., Hoover, P.J., Chicoine, A., Eisenhaure, T.M., Jonsson, A.H., Li, S., et al. (2019). The immune cell landscape in kidneys of patients with lupus nephritis. Nat. Immunol. *20*, 902–914.

Azizi, E., Carr, A.J., Plitas, G., Cornish, A.E., Konopacki, C., Prabhakaran, S., Nainys, J., Wu, K., Kiseliovas, V., Setty, M., et al. (2018). Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. Cell *174*, 1293–1308.e36.

Bacher, N., Raker, V., Hofmann, C., Graulich, E., Schwenk, M., Baumgrass, R., Bopp, T., Zechner, U., Merten, L., Becker, C., et al. (2013). Interferon-alpha suppresses cAMP to disarm human regulatory T cells. Cancer Res. *73*, 5647–5656.

Banchereau, R., Hong, S., Cantarel, B., Baldwin, N., Baisch, J., Edens, M., Cepika, A.M., Acs, P., Turner, J., Anguiano, E., et al. (2016). Personalized immunomonitoring uncovers molecular networks that stratify lupus patients. Cell *165*, 1548–1550.

Boland, B.S., He, Z., Tsai, M.S., Olvera, J.G., Omilusik, K.D., Duong, H.G., Kim, E.S., Limary, A.E., Jin, W., Milner, J.J., et al. (2020). Heterogeneity and clonal relationships of adaptive immune cells in ulcerative colitis revealed by single-cell analyses. Sci. Immunol. *5*, eabb4432.

Brightbill, H.D., Suto, E., Blaquiere, N., Ramamoorthi, N., Sujatha-Bhaskar, S., Gogol, E.B., Castanedo, G.M., Jackson, B.T., Kwon, Y.C., Haller, S., et al. (2018). NF-kappaB inducing kinase is a therapeutic target for systemic lupus erythematosus. Nat. Commun. *9*, 179.

Caielli, S., Veiga, D.T., Balasubramanian, P., Athale, S., Domic, B., Murat, E., Banchereau, R., Xu, Z., Chandra, M., Chung, C.H., et al. (2018). A CD4(+) T cell population expanded in lupus blood provides B cell help through interleukin-10 and succinate. Nat. Med. *25*, 75–81.

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502.

Chevalier, N., Jarrossay, D., Ho, E., Avery, D.T., Ma, C.S., Yu, D., Sallusto, F., Tangye, S.G., and Mackay, C.R. (2011). CXCR5 expressing human central memory CD4 T cells and their relevance for humoral immune responses. J. Immunol. *186*, 5556–5568.

Coit, P., Dozmorov, M.G., Merrill, J.T., McCune, W.J., Maksimowicz-McKinnon, K., Wren, J.D., and Sawalha, A.H. (2016). Epigenetic reprogramming in naive CD4+T cells favoring T cell activation and non-Th1 effector T cell immune response as an early event in lupus flares. Arthritis Rheumatol. *68*, 2200–2209.

Crow, M.K., Olferiev, M., and Kirou, K.A. (2019). Type I interferons in autoimmune disease. Annu. Rev. Pathol. *14*, 369–393.

Cuadrado, E., van den Biggelaar, M., de Kivit, S., Chen, Y.Y., Slot, M., Doubal, I., Meijer, A., van Lier, R.A.W., Borst, J., and Amsen, D. (2018). Proteomic analyses of human regulatory T cells reveal adaptations in signaling pathways that protect cellular identity. Immunity *48*, 1046–1059.e6.

Dominguez-Villar, M., Baecher-Allan, C.M., and Hafler, D.A. (2011). Identification of T helper type 1-like, Foxp3+ regulatory T cells in human autoimmune disease. Nat. Med. *17*, 673–675.

Dominguez-Villar, M., and Hafler, D.A. (2018). Regulatory T cells in autoimmune disease. Nat. Immunol. *19*, 665–673.

Gangaplara, A., Martens, C., Dahlstrom, E., Metidji, A., Gokhale, A.S., Glass, D.D., Lopez-Ocasio, M., Baur, R., Kanakabandi, K., Porcella, S.F., et al. (2018). Type I interferon signaling attenuates regulatory T cell function in viral infection and in the tumor microenvironment. PLoS Pathog. *14*, e1006985.

Gladman, D.D., Ibañez, D., and Urowitz, M.B. (2002). Systemic lupus erythematosus disease activity index 2000. J. Rheumatol. *29*, 288–291.

Golding, A., Rosen, A., Petri, M., Akhter, E., and Andrade, F. (2010). Interferon-alpha regulates the dynamic balance between human activated regulatory and effector T cells: implications for antiviral and autoimmune responses. Immunology *131*, 107–117.

Goropevsek, A., Holcar, M., and Avčin, T. (2017). The role of STAT signaling pathways in the pathogenesis of systemic lupus erythematosus. Clin. Rev. Allergy Immunol. *52*, 164–181.

Hashimoto, H., Ueda, R., Narumi, K., Heike, Y., Yoshida, T., and Aoki, K. (2014). Type I IFN gene delivery suppresses regulatory T cells within tumors. Cancer Gene Ther. *21*, 532–541.

He, J., Tsai, L.M., Leong, Y.A., Hu, X., Ma, C.S., Chevalier, N., Sun, X., Vandenberg, K., Rockman, S., Ding, Y., et al. (2013). Circulating precursor CCR7(lo) PD-1(hi) CXCR5(+) CD4(+) T cells indicate Tfh cell activity and promote antibody responses upon antigen reexposure. Immunity *39*, 770–781.

Heaton, H., Talman, A.M., Knights, A., Imaz, M., Gaffney, D.J., Durbin, R., Hemberg, M., and Lawniczak, M.K.N. (2020). Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. Nat. Methods *17*, 615–620.

Hedrich, C.M., Mäbert, K., Rauen, T., and Tsokos, G.C. (2017). DNA methylation in systemic lupus erythematosus. Epigenomics *9*, 505–525.

Hu, N., Qiu, X., Luo, Y., Yuan, J., Li, Y., Lei, W., Zhang, G., Zhou, Y., Su, Y., and Lu, Q. (2008). Abnormal histone modification patterns in lupus CD4+T cells. J. Rheumatol. *35*, 804–810.

Huang, M., Wang, J., Torre, E., Dueck, H., Shaffer, S., Bonasio, R., Murray, J.I., Raj, A., Li, M., and Zhang, N.R. (2018). SAVER: gene expression recovery for single-cell RNA sequencing. Nat. Methods *15*, 539–542.

Hutcheson, J., Scatizzi, J.C., Siddiqui, A.M., Haines, G.K., Wu, T., Li, Q.Z., Davis, L.S., Mohan, C., and Perlman, H. (2008). Combined deficiency of proapoptotic regulators Bim and Fas results in the early onset of systemic autoimmunity. Immunity *28*, 206–217.

Kang, H.M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., McCarthy, E., Wan, E., Wong, S., Byrnes, L., Lanata, C.M., et al. (2018). Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. Nat. Biotechnol. *36*, 89–94.

Kim, K., Park, S., Park, S.Y., Kim, G., Park, S.M., Cho, J.W., Kim, D.H., Park, Y.M., Koh, Y.W., Kim, H.R., et al. (2020). Single-cell transcriptome analysis reveals TOX as a promoting factor for T cell exhaustion and a predictor for anti-PD-1 responses in human cancer. Genome Med. *12*, 22.

Kim, S.J., Lee, K., and Diamond, B. (2018). Follicular helper T cells in systemic lupus erythematosus. Front. Immunol. *9*, 1793.

Kiriakidou, M. (2013). Systemic lupus erythematosus. Ann. Intern. Med. *159*. ITC4-Itc416.

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. *22*, 568–576.

Koga, T., Otomo, K., Mizui, M., Yoshida, N., Umeda, M., Ichinose, K., Kawakami, A., and Tsokos, G.C. (2016). Calcium/calmodulin-dependent kinase IV facilitates the recruitment of interleukin-17-producing cells to target organs through the CCR6/CCL20 Axis in Th17 cell-driven inflammatory diseases. Arthritis Rheumatol. *68*, 1981–1988.

Komatsu, N., Okamoto, K., Sawa, S., Nakashima, T., Oh-hora, M., Kodama, T., Tanaka, S., Bluestone, J.A., and Takayanagi, H. (2014). Pathogenic conversion of Foxp3+ T cells into TH17 cells in autoimmune arthritis. Nat. Med. *20*, 62–68.

Korn, T., Reddy, J., Gao, W., Bettelli, E., Awasthi, A., Petersen, T.R., Bäckström, B.T., Sobel, R.A., Wucherpfennig, K.W., Strom, T.B., et al. (2007). Myelin-specific regulatory T cells accumulate in the CNS but fail to control autoimmune inflammation. Nat. Med. *13*, 423–431.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lee, S.E., Li, X., Kim, J.C.K., Lee, J., González-Navajas, J.M., Hong, S.H., Park, I.K., Rhee, J.H., and Raz, E. (2012). Type I interferons maintain Foxp3 expression and T-regulatory cell functions under inflammatory conditions in mice. Gastroenterology *143*, 145–154.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. *1*, 417–425.

Lin, J., Yu, Y., Ma, J., Ren, C., and Chen, W. (2019). PD-1+CXCR5-CD4+T cells are correlated with the severity of systemic lupus erythematosus. Rheumatology *58*, 2188–2192.

Lisnevskaia, L., Murphy, G., and Isenberg, D. (2014). Systemic lupus erythematosus. Lancet *384*, 1878–1888.

Lowther, D.E., Goods, B.A., Lucca, L.E., Lerner, B.A., Raddassi, K., van Dijk, D., Hernandez, A.L., Duan, X., Gunel, M., Coric, V., et al. (2016). PD-1 marks dysfunctional regulatory T cells in malignant gliomas. JCI Insight *1*, e85935.

Luckheeram, R.V., Zhou, R., Verma, A.D., and Xia, B. (2012). CD4(+)T cells: differentiation and functions. Clin. Dev. Immunol. *2012*, 925135.

Makiyama, A., Chiba, A., Noto, D., Murayama, G., Yamaji, K., Tamura, N., and Miyake, S. (2019). Expanded circulating peripheral helper T cells in systemic lupus erythematosus: association with disease activity and B cell differentiation. Rheumatology *58*, 1861–1869.

McClymont, S.A., Putnam, A.L., Lee, M.R., Esensten, J.H., Liu, W., Hulme, M.A., Hoffmüller, U., Baron, U., Olek, S., Bluestone, J.A., et al. (2011). Plasticity of human regulatory T cells in healthy subjects and patients with type 1 diabetes. J. Immunol. *186*, 3918–3926.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Metidji, A., Rieder, S.A., Glass, D.D., Cremer, I., Punkosdy, G.A., and Shevach, E.M. (2015). IFN-alpha/beta receptor signaling promotes regulatory T cell development and function under stress conditions. J. Immunol. *194*, 4265–4276.

Miyara, M., Yoshioka, Y., Kitoh, A., Shima, T., Wing, K., Niwa, A., Parizot, C., Taflin, C., Heike, T., Valeyre, D., et al. (2009). Functional delineation and differentiation dynamics of human CD4+ T cells expressing the FoxP3 transcription factor. Immunity *30*, 899–911.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. *34*, 267–273.

Morand, E.F., Furie, R., Tanaka, Y., Bruce, I.N., Askanase, A.D., Richez, C., Bae, S.C., Brohawn, P.Z., Pineda, L., Berglind, A., et al. (2020). Trial of anifrolumab in active systemic lupus erythematosus. N. Engl. J. Med. *382*, 211–221.

Morita, R., Schmitt, N., Bentebibel, S.E., Ranganathan, R., Bourdery, L., Zurawski, G., Foucat, E., Dullaers, M., Oh, S., Sabzghabaei, N., et al. (2011). Human blood CXCR5(+)CD4(+) T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion (vol 34, pg 108, 2011). Immunity *34*, 135.

Nehar-Belaid, D., Hong, S., Marches, R., Chen, G., Bolisetty, M., Baisch, J., Walters, L., Punaro, M., Rossi, R.J., Chung, C.H., et al. (2020). Mapping

systemic lupus erythematosus heterogeneity at the single-cell level. Nat. Immunol. *21*, 1094–1106.

Patil, V.S., Madrigal, A., Schmiedel, B.J., Clarke, J., O'Rourke, P., de Silva, A.D., Harris, E., Peters, B., Seumois, G., Weiskopf, D., et al. (2018). Precursors of human CD4+ cytotoxic T lymphocytes identified by single-cell transcriptome analysis. Sci. Immunol. *3*, eaan8664.

Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. *9*, 171–181.

Piconese, S., Pacella, I., Timperi, E., and Barnaba, V. (2015). Divergent effects of type-I interferons on regulatory T cells. Cytokine Growth Factor Rev. *26*, 133–141.

Qu, K., Zaba, L.C., Satpathy, A.T., Giresi, P.G., Li, R., Jin, Y., Armstrong, R., Jin, C., Schmitt, N., Rahbar, Z., et al. (2017). Chromatin accessibility landscape of cutaneous T cell lymphoma and dynamic response to HDAC inhibitors. Cancer Cell *32*, 27–41.e4.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rome, K.S., Stein, S.J., Kurachi, M., Petrovic, J., Schwartz, G.W., Mack, E.A., Uljon, S., Wu, W.W., DeHart, A.G., McClory, S.E., et al. (2020). Trib1 regulates T cell differentiation during chronic infection by restraining the effector program. J. Exp. Med. *217*, e20190888.

Sakaguchi, S., Mikami, N., Wing, J.B., Tanaka, A., Ichiyama, K., and Ohkura, N. (2020). Regulatory T cells and human disease. Annu. Rev. Immunol. *38*, 541–566.

Sallusto, F. (2016). Heterogeneity of human CD4+T cells against microbes. Annu. Rev. Immunol. *34*, 317–334.

Schafflick, D., Xu, C.A., Hartlehnert, M., Cole, M., Schulte-Mecklenbeck, A., Lautwein, T., Wolbert, J., Heming, M., Meuth, S.G., Kuhlmann, T., et al. (2020). Integrated single cell analysis of blood and cerebrospinal fluid leukocytes in multiple sclerosis. Nat. Commun. *11*, 247.

Scheinecker, C., Göschl, L., and Bonelli, M. (2020). Treg cells in health and autoimmune diseases: new insights from single cell analysis. J. Autoimmun. *110*, 102376.

Scott, A.C., Dündar, F., Zumbo, P., Chandran, S.S., Klebanoff, C.A., Shakiba, M., Trivedi, P., Menocal, L., Appleby, H., Camara, S., et al. (2019). TOX is a critical regulator of tumour-specific T cell differentiation. Nature *571*, 270–274.

Sharma, S., Jin, Z., Rosenzweig, E., Rao, S., Ko, K., and Niewold, T.B. (2015). Widely divergent transcriptional patterns between SLE patients of different ancestral backgrounds in sorted immune cell populations. J. Autoimmun. *60*, 51–58.

Smillie, C.S., Biton, M., Ordovas-Montanes, J., Sullivan, K.M., Burgin, G., Graham, D.B., Herbst, R.H., Rogel, N., Slyper, M., Waldman, J., et al. (2019). Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. Cell *178*, 714–730.e22.

Srivastava, S., Koch, M.A., Pepper, M., and Campbell, D.J. (2014). Type I interferons directly inhibit regulatory T cells to allow optimal antiviral T cell responses during acute LCMV infection. J. Exp. Med. *211*, 961–974.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21.

Suarez-Fueyo, A., Bradley, S.J., and Tsokos, G.C. (2016). T cells in systemic lupus erythematosus. Curr. Opin. Immunol. *43*, 32–38.

Sumida, T., Lincoln, M.R., Ukeje, C.M., Rodriguez, D.M., Akazawa, H., Noda, T., Naito, A.T., Komuro, I., Dominguez-Villar, M., and Hafler, D.A. (2018). Activated beta-catenin in Foxp3(+) regulatory T cells links inflammatory environments to autoimmunity. Nat. Immunol. *19*, 1391–1402.

Takeshita, M., Suzuki, K., Kassai, Y., Takiguchi, M., Nakayama, Y., Otomo, Y., Morita, R., Miyazaki, T., Yoshimura, A., and Takeuchi, T. (2015). Polarization diversity of human CD4(+) stem cell memory T cells. Clin. Immunol. *159*, 107–117.

Takeuchi, A., and Saito, T. (2017). CD4 CTL, a cytotoxic subset of CD4(+) T cells, their differentiation and function. Front. Immunol. *8*, 194.

Tsokos, G.C. (2020). Autoimmunity and organ damage in systemic lupus erythematosus. Nat. Immunol. *21*, 605–614.

Tsokos, G.C., Lo, M.S., Costa Reis, P., and Sullivan, K.E. (2016). New insights into the immunopathogenesis of systemic lupus erythematosus. Nat. Rev. Rheumatol. *12*, 716–730.

Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., et al. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. *42*, D1001–D1006.

Wherry, E.J., and Kurachi, M. (2015). Molecular and cellular insights into T cell exhaustion. Nat. Rev. Immunol. *15*, 486–499.

Wilk, A.J., Rustagi, A., Zhao, N.Q., Roque, J., Martínez-Colón, G.J., McKechnie, J.L., Ivison, G.T., Ranganath, T., Vergara, R., Hollis, T., et al. (2020). A single-cell atlas of the peripheral immune response in patients with severe COVID-19. Nat. Med. *26*, 1070–1076.

Wing, J.B., Tanaka, A., and Sakaguchi, S. (2019). Human FOXP3(+) regulatory T cell heterogeneity and function in autoimmunity and cancer. Immunity *50*, 302–316.

Yang, K., Blanco, D.B., Neale, G., Vogel, P., Avila, J., Clish, C.B., Wu, C., Shrestha, S., Rankin, S., Long, L., et al. (2017). Homeostatic control of metabolic and functional fitness of Treg cells by LKB1 signalling. Nature *548*, 602–606.

Yee, C.S., Farewell, V., Isenberg, D.A., Rahman, A., Teh, L.S., Griffiths, B., Bruce, I.N., Ahmad, Y., Prabu, A., Akil, M., et al. (2007). British Isles Lupus Assessment Group 2004 index is valid for assessment of disease activity in systemic lupus erythematosus. Arthritis Rheum. *56*, 4113–4119.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523.

Zuo, Z., Jin, Y., Zhang, W., Lu, Y., Li, B., and Qu, K. (2019). ATAC-pipe: general analysis of genome-wide chromatin accessibility. Brief. Bioinform. *20*, 1934–1943.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
APC anti-human CD196 (CCR6) Antibody	Biolegend	cat#353416; RRID:AB_10945155
APC anti-human CD197 (CCR7) Antibody	Biolegend	cat#353214; RRID:AB_10917387
APC anti-human CD25 Antibody	eBioscience	cat#17-0259-42; RRID:AB_1582219
APC/Cyanine7 anti-human CD45 Antibody	Biolegend	cat#304014; RRID:AB_314402
Brilliant Violet 421™ Anti-human CD127 Antibody	Biolegend	cat#562437; RRID:AB_11153481
Brilliant Violet 421™ anti-human CD197 (CCR7) Antibody	Biolegend	cat#353207 ; RRID:AB_10915137
FITC anti-human CD4 Antibody	Biolegend	cat#300538 ; RRID:AB_2562052
PE anti-human CD127 Antibody	eBioscience	cat#12-1278-41; RRID:AB_10853334
PE anti-human CD185 (CXCR5) Antibody	Biolegend	cat#356904; RRID:AB_2561813
PE anti-human CD25 Antibody	BD Bioscience	cat#560989; RRID:AB_10563905
PE anti-human FOXP3 Antibody	Biolegend	cat#320208; RRID:AB_492982
PE/Cyanine7 anti-human CD183 (CXCR3) Antibody	Biolegend	cat#353720; RRID:AB_11219383
PE/Cyanine7 anti-human CD74 Antibody	Biolegend	cat#357609; RRID:AB_2721663
PE/Cyanine7 Anti-Human CD45RA Antibody	Biolegend	cat#304125; RRID:AB_10709440
PerCP/Cyanine5.5 anti-human CD3 Antibody	Biolegend	cat#344808; RRID:AB_10640736
Chemicals, peptides, and recombinant proteins		
penicillin/streptomycin	Gibco	cat#15140122
Glutamine	Gibco	cat#25030081
sodium pyruvate	Thermo Fisher Scientific	cat#11360070
nonessential amino acids	Solarbio	cat#N1250
human recombinant IFN-α	Abcam	cat#ab48750
Critical commercial assays		
10X Chromium Single Cell 3′ Kit	10X Genomics	cat#120237
CD4+CD25+ Regulatory T Cell Isolation Kit	Miltenyi Biotec	cat#130091301
CellTrace™ CFSE Cell Proliferation Kit	Thermo Fisher Scientific	cat#C34554
Dynabeads™ Human T-Activator CD3/CD28	Invitrogen	cat#11161D
Roswell Park Memorial Institute (RPMI) 1640 medium	HyClone	cat#SH30809.01
Maxima H Minus Reverse Transcriptase	Thermo Fisher Scientific	cat#EP0751
SYBR Green PCR Master Mix	Applied Biosystems	cat#4344463
Tn5 transposome	Vazyme Biotech	cat#TD501
Nonidet P40 Substitute	Roche	cat#11332473001
MinElute PCR Purification Kit	Qiagen	cat#28006
Foxp3/Transcription Factor Staining Buffer	eBioscience	cat#00-5523-00
SPRIselect beads	Beckman Coulter	cat#B23318
foetal calf serum	Gibco	cat#16170078
Deposited data		
MSigDB pathways	Liberzon et al., 2015	http://software.broadinstitute.org/gsea/msigdb/index.jsp
Microarray data of blood CD4$^+$ T from SLE patients and controls	N/A	Gene Expression Omnibus (GEO) (GSE4588)
Microarray data of healthy and SLE blood CD4$^+$ T	Sharma et al., 2015	Gene Expression Omnibus (GEO) (GSE55447)
Microarray data of healthy and SLE blood CD4$^+$ T	Hutcheson et al., 2008	Gene Expression Omnibus (GEO) (GSE10325)

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Microarray data of CD4$^+$ T cell subsets	Takeshita et al., 2015	Gene Expression Omnibus (GEO) (GSE61697)
RNA-seq data of three Treg groups (nTreg, eTreg, Fr. III)	Cuadrado et al., 2018	Gene Expression Omnibus (GEO) (GSE90600)
Single cell RNA-seq data of colon mucosa of UC patients and controls	Smillie et al., 2019	Single Cell Portal: SCP259
Single-cell RNA-seq data of PBMCs from UC patients and controls	Boland et al., 2020	Gene Expression Omnibus (GEO) (GSE125527)
Single-cell RNA-seq data of PBMCs from healthy controls	Wilk et al., 2020	Gene Expression Omnibus (GEO) (GSE150728)
Single-cell RNA-seq data of PBMCs from SLE patients	Schafflick et al., 2020	Gene Expression Omnibus (GEO) (GSE137029)
Single-cell RNA-seq data of PBMCs from multiple sclerosis patients	Schafflick et al., 2020	Gene Expression Omnibus (GEO) (GSE138266)
ATAC-seq and scRNA-seq of CD4$^+$ T cells from SLE patients and healthy controls	This paper	Genome Sequence Archive (GSA) for Humans (HRA000676)
Bulk RNA-seq of Treg cells stimulated with/without IFN-α	This paper	Genome Sequence Archive (GSA) for Humans (HRA000676)
Software and algorithms		
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
DESeq	Anders and Huber, 2010	http://bioconductor.org/packages/release/bioc/html/DESeq.html
BedTools	Quinlan and Hall, 2010	http://bedtools.readthedocs.io/en/latest/
ATAC-pipe	Zuo et al., 2019	https://github.com/QuKunLab/ATAC-pipe
MetaScape	Zhou et al., 2019	http://metascape.org/gp/index.html#/main/step1
GREAT	McLean et al., 2010	http://bejerano.stanford.edu/great/public/html/
GSEA	Mootha et al., 2003	https://www.gsea-msigdb.org/gsea/index.jsp
Demuxlet	Kang et al., 2018	https://github.com/statgen/demuxlet
Souporcell	Heaton et al., 2020	https://github.com/wheaton5/souporcell
CellRanger V2.0	10X Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger
Seurat V3.0 (R package)	Stuart et al., 2019	https://satijalab.org/seurat/
SAVER (R package)	Huang et al., 2018	https://github.com/mohuangx/SAVER
Code in this study	This paper	https://github.com/QuKunLab/SLE; https://zenodo.org/record/7118733. (https://doi.org/10.5281/zenodo.7118733)

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by lead contact, Kun Qu (qukun@ustc.edu.cn).

Materials availability

All non-commercial reagents used in this paper are available from the lead contact upon request.

Data and code availability

The processed data reported in this paper and raw data are available for download from the Genome Sequence Archive (GSA) for Humans at GSA: https://bigd.big.ac.cn/gsa-human/browse/HRA000676. The analysis codes supporting the current study are available at Github: https://github.com/QuKunLab/SLE. The code used in this study has also been publicly deposited at Zenodo: https://zenodo.org/record/7118733. (https://doi.org/10.5281/zenodo.7118733). Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects

Peripheral blood samples were collected from SLE patients and healthy controls at the First Affiliated Hospital of University of Science and Technology of China. Informed consent was obtained from the patients. Study procedures were followed in accordance with protocols approved by the ethics committee of the University of Science and Technology of China. Detailed clinical information for the patients is described in Table S1. The 24 clinical and laboratory parameters for SLE patients were collected according to the internationally certified method SLEDAI-2K (Gladman et al., 2002).

The ATAC-seq data of $CD4^+$ T cells were obtained from a cohort of 56 female and 7 male SLE patients, with a mean age of 35.48; the scRNA-seq data of $CD4^+$ T cells were obtained from a cohort of 10 female SLE patients, with a mean age of 31.33; the FACS experiment of $CD4^+$ T cells was obtained from a cohort of 27 female SLE patients, with a mean age of 39.5; the ATAC-seq data of Treg cells were obtained from a cohort of 5 female SLE patients, with a mean age of 41; the quantitive PCR experiment of Treg cells was obtained from a cohort of 8 female SLE patients, with a mean age of 37.3; the assay of the suppressive function for Treg cells was obtained from a cohort of 4 female SLE patients, with a mean age of 28.75. Detailed information is also available in Table S1.

METHOD DETAILS

Cell isolation

Peripheral blood was drawn in a green-top blood collection tube. PBMCs were then prepared by Ficoll-Paque density gradient centrifugation and stained with fluorochrome-labelled anti-human monoclonal antibodies (Biolegend). Bulk $CD4^+$ T cells were sorted with CD45 (clone HI30), CD4 (clone RPA-T4), and CD3 (SK7) using a SH800S flow cytometer (SONY). For $CD4^+$ T helper cell subtypes, cells were identified as previously described (Morita et al., 2011). Briefly, naïve cells were identified as $CD4^+$ $CCR7^+$ $CD45RA^+$, $CXCR5^+$ T cells as $CD4^+$ $CD25^-$ $CD45RA^-$ $CXCR5^+$, Th1 cells as $CD4^+$ $CD25^-$ $CD45RA^-CXCR3^+$ $CCR6^-$, Th2 cells as $CD4^+$ $CD25^-$ $CD45RA^-$ $CXCR3^-$ $CCR6^-$, Th17 cells as $CD4^+$ $CD25^-$ $CD45RA^-$ $CXCR3^-$ $CCR6^+$, Treg cells as $CD4^+$ $FOXP3^+$, and Tct cells as $CD4^+$ $CD25^-$ $CD45RA^+$ $CCR7^-$. The post-sort purities were confirmed to be >95% prior to ATAC-seq. For single-cell RNA-seq, $CD4^+$ T cells were sorted and cryopreserved according to the 10X Genomics official recommendation.

Intracellular staining

PBMCs were prepared and stained using the indicated human mAbs. Homologous IgGs served as the negative control. FACS surface marker staining was performed according to the Biolegend antibody instructions. For intracellular staining, cells were blocked, stained with Foxp3 (clone 259D) or CD74 (clone pin.1) and then washed with Foxp3/Transcription Factor Staining Buffer (eBioscience, cat no. 00-5523-00) according to the manufacturer's instructions.

Single-cell RNA-seq library preparation and sequencing

Single-cell suspensions were prepared as described in the 10X Genomics protocol. Briefly, we sorted $CD4^+$ T cells from 10 SLE patients and 6 healthy controls. The cells reached a final viability of 85%. We then resuspended the cells at a concentration of 700 cells/μL and mixed the same sample types immediately according to the 10X Genomics Chromium single-cell protocol for the v2 reagent kit (10X Genomics). Cell suspensions were loaded onto a chromium single-cell chip along with reverse transcription (RT) master mix and 3′ gel beads. After generation of single-cell gel beads in emulsion (GEMs). RT was performed using a C1000 Touch™ Thermal Cycler (Bio-Rad) using the manufacturer's standard parameters. cDNA was amplified and purified with SPRIselect beads (Beckman Coulter). Single-cell 3′ libraries were then constructed following fragmentation, end repair, polyA tailing, adaptor ligation and size selection. Single-cell sequencing libraries were generated with one sample index for each sample and sequenced on the Illumina HiSeq X-Ten platform.

ATAC-seq library preparation and sequencing

ATAC-seq of $CD4^+$ T cells was performed as previously described with minor modifications. Briefly, $CD4^+$ T cells were sorted using a SH800S sorter (SONY). Samples were lysed in cold lysis buffer (10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$ and 0.1% NP-40

(Roche)) for 3 min on ice to prepare the nuclei. Immediately after cell lysis, nuclei were centrifuged at 500 g for 5 min, and the supernatant was discarded carefully. Nuclei extracts were then incubated with Tn5 transposome (Vazyme Biotech, cat no. TD501) at 37°C for 30 min. After DNA purification with the MinElute Kit (Qiagen), PCR was performed to amplify the library for 12–15 cycles based on the quantitative data regarding the optimum number of PCR cycles. The PCR conditions were as follows: 98°C for 30 sec and cycling at 98°C for 10 sec, 63°C for 30 sec and 72°C for 1 min. After PCR amplification, the sample libraries were purified and sequenced on the Illumina HiSeq X-Ten platform with the 150-bp paired-end configuration.

Assessment of the inhibitory function of Treg cells

For functional analysis, $CD4^+CD25^+$ T cells and $CD4^+CD25^-$ T cells were purified from PBMCs using a $CD4^+CD25^+$ Regulatory T Cell Isolation Kit (Miltenyi Biotec, cat no. 130091301). Briefly, a non-$CD4^+$ cocktail and anti-biotin beads were used for the isolation of $CD4^+$ T cells. After detachment, the cells were washed, and $CD4^+CD25^+$ Treg cells were positively selected using CD25 microbeads. The cells were reanalysed after sorting and routinely showed a purity greater than 95%. The negative fraction of $CD4^+CD25^-$ T cells from healthy controls served as effector cells and were stained with a CellTrace™ CFSE Cell Proliferation Kit (Thermo Fisher Scientific, cat no. C34554) at 1 μm for 15 min. Then, $CD4^+CD25^-$ T cells (2*10∧4) were incubated with Dynabeads™ Human T-Activator CD3/CD28 (Invitrogen, cat no. 11161D) at a 20:1 ratio in Roswell Park Memorial Institute (RPMI) 1640 medium (HyClone, cat no. SH30809.01) supplemented with 10% foetal calf serum (Gibco, cat no. 16170078), 100 U/mL penicillin/streptomycin (Gibco, cat no. 15140122), 2 mM glutamine (Gibco, cat no. 25030081), sodium pyruvate (Thermo Fisher Scientific, cat no. 11360070) and nonessential amino acids (Solarbio, cat no. N1250). Purified $CD4^+CD25^+$ Treg cells from healthy controls or SLE patients were added to the culture at a 1:0 or 1:1 ratio. After 4 days, the proliferation of $CD4^+CD25^-$ T cells was determined by assessing CFSE dilution by flow cytometry. The data are expressed as the percent inhibition of cell proliferation according to the following formula: inhibition% = 100-(cell proliferation ratio at 1:1/cell proliferation ratio at 1:0).

In vitro stimulation of Treg with IFN-α

PBMCs from healthy controls were isolated by Ficoll-Paque gradient centrifugation and then stimulated with human recombinant IFN-α(0 μg/mL, 0.25 μg/mL, 0.5 μg/mL, 1 μg/mL) (Abcam cat no. ab48750). The cells were cultured for 7 days in RPMI 1640 medium supplemented with 10% heat-inactivated human serum or 10% SLE serum. After 7 days, cells were harvested and isolated by flow cytometry using the same sorting strategy as that used for $CD4^+CD25^+CD127^{low}CCR7^{low}$ Treg2 cells. Finally, RNA-seq library construction and sequencing were performed using the same strategies described above.

Bulk RNA-seq library preparation and sequencing

Treg ($DAPI^-CD45^+CD3^+CD4^+CD25^+CD127^{low}$) and $CCR7^-$ Treg2 ($DAPI^-CD45^+CD3^+CD4^+CD25^+CD127^{low}CCR7^{low}$) cells were sorted by flow cytometry. For Treg and $CCR7^-$ Treg2 cells, up to 1000 cells were collected directly in a 0.2 mL PCR tube (KIRGEN, Cat No. KG2511), and the RNA-seq library was constructed using the Smart-seq2 method (Picelli et al., 2014) and sequenced on the Illumina NovaSeq 6000 system; at least 20 million paired reads were generated per sample.

Real-time quantitative polymerase chain reaction

The selected candidate genes were validated by qPCR. Briefly, cDNA was synthesized with Maxima H Minus Reverse Transcriptase (Thermo Fisher Scientific, cat no. EP0751) in accordance with the manufacturer's instructions. Two-step PCR was performed using SYBR Green PCR Master Mix (Applied Biosystems, cat no. 4344463) in accordance with the manufacturer's instructions on a LightCycler96 fluorescence sequence detection system (Roche). Gene expression was quantified relative to that of the housekeeping gene *GAPDH* and normalized to the control by the standard $2^{-\Delta\Delta CT}$ calculation. The primer sequences were as follows: *GAPDH*, 5′- GGAGCGAGATCCCTCCAAAAT-3′ and 5′- GGCTGTTGTCATACTTCTCATGG-3′; *PDCD1*, 5′-ACGAGGGACAATAGGA GCCA-3′ and 5′-GGCATACTCCGTCTGCTCAG-3′; *LAG3*, 5′-GCCTCCGACTGGGTCATTTT-3′ and 5′- CTTTCCGCTAAGTGG TGATGG-3′; *CTLA4*, 5′-CATGATGGGGAATGAGTTGACC-3′ and 5′- TCAGTCCTTGGATAGTGAGGTTC-3′; *TIGIT*, 5′-TGGTCGCG TTGACTAGAAAGA-3′ and 5′- GGGCTCCATTCCTCCTGTC-3′.

ATAC-seq primary data processing and peak calling

ATAC-seq raw data were processed using the published ATAC-seq pipeline ATAC-pipe (Zuo et al., 2019). Sequencing reads were mapped using the "–MappingQC" module with option "-c 50" in ATAC-pipe. Adapter sequences were trimmed, and reads were mapped to hg19 using Bowtie2 (Langmead and Salzberg, 2012). PCR duplicates were removed as described. Mapped reads were then shifted +4/-5 bp depending on the read strand such that the first base of each mapped read represented the Tn5 cleavage position. All mapped reads were then extended to 50 bp centred by the cleavage position. Reads mapped to repeated regions and chromosome M were removed. We used the "–PeakCalling" module in ATAC-pipe with the options "–p1 3 –q1 5 –f1 1 -w 50" to call the peaks. The peaks were then filtered, and enriched regions were identified as those with a posterior probability >0.99. Samples from the same cell type classified under the same clinical condition (healthy, SLE) were grouped for peak calling, and peaks for all categories were then merged together to generate a peak list. The number of raw read counts mapped to each peak in each sample was quantified by this module in ATAC-pipe. We then obtained an N×M data matrix where N indicated the number of merged peaks, M indicated the number of samples, and the matrix value $D_{i,j}$ represented the raw read counts falling in peak i (i = 1 to N) of sample

j (j = 1 to M). After manually removing the peaks mapping to chromosome Y, this data matrix was then normalized by the ''normalize.quantiles'' function of the ''preprocessCore'' package in R, and the normalized matrix was then then $\log_2$ transformed. To avoid the effect of the difference between males and females, we excluded the peaks that were significantly differential in male and female samples (N = 212, screened by all male samples vs all female samples, $|\log_2$ fold change$|>1.2$, FDR<0.05). This final data matrix was used for downstream analysis.

Normalization of ATAC-seq profiles

The UCSC Genome Browser provides aligned annotation tracks and can also display ATAC-seq signal peaks. To avoid the impacts of variable sequencing depths and signal-to-noise ratios among different samples, we normalized the ''bedGraph'' files before uploading. To do this, we used the R package DESeq (Anders and Huber, 2010) to calculate the size factor for every sample via the sample×peak raw count matrix (output file of ATAC-pipe). These size factors can be used to suitably measure the depth of data at chromatin-accessible regions in all samples. To obtain the standardized data, the raw count in the ''bedGraph'' file for each sample was divided by the size factor. Normalized ''bedGraph'' files were then converted to the UCSC Genome Browser input format ''bigWig'' using the UCSC tool ''bedGraphToBigWig''.

Saturation curve of the sample and peak numbers

Given the heterogeneity of $CD4^+$ T cell subsets, complexity of SLE and variety between individuals, a sufficient number of samples is necessary to ensure that all chromatin-accessible regions of SLE $CD4^+$ T cells are detectable. To construct the saturation curve of the numbers of SLE samples used and peaks detected, we first randomly selected a certain number of SLE samples. Next, we obtained the peak list by the same peak calling method described above, and the number of peaks called from this random selection was obtained. This random selection was performed on 1 to 63 samples 10 times. The x-axis of the saturation curve depicts the number of samples used, while the y-axis depicts the number of peaks called from 10 random selections.

Correlation analysis of all SLE patients

To obtain a representative correlation landscape of all samples from SLE patients, we first filtered variant peaks among all SLE samples based on a coefficient of variation (COV) greater than a certain threshold, and multiple variant peak lists were obtained from a COV threshold ranging from 0.2 to 0.7 (step by 0.02). Then, these sample×variant peak count matrixes were used to calculate the sample×sample Pearson's correlation matrixes. The average of these correlation matrixes was used for the final data presentation. After unsupervised hierarchical clustering (Seaborn clustermap, with parameter metric = 'Euclidean', method = 'complete'), SLE patients were divided into three distinct groups.

Summary of the clinical information for the three SLE patient groups

The SLE DAI was calculated for each SLE patient with detailed clinical information according to the internationally certified method SLEDAI-2K (Gladman et al., 2002). Nephritis is a serious and common comorbidity in SLE patients. To quantify the severity of nephritis for each patient, we defined the sum of the four nephritis-related indicators—haematuria, proteinuria, pyuria, and tubular urine—in SLEDAI-2K as the nephritis DAI. We compared the SLE and nephritis DAIs among the three patient groups and summarized the ratio of patients with each comorbidity for each patient group.

Acquisition of microarray and RNA-seq datasets

We downloaded the following bulk datasets for comparison to our single-cell data: (1) the microarray gene expression data of three $CD4^+$ T cell subsets in each developmental stage, Tn/Tem/Tcm, downloaded from the Gene Expression Omnibus: GSE61697 (Takeshita et al., 2015); (2) the microarray gene expression data of healthy and SLE blood $CD4^+$ T cells, downloaded from the Gene Expression Ominibus: GSE4588, GSE10325 and GSE55447 (Hutcheson et al., 2008; Sharma et al., 2015); (3) the bulk RNA-seq of three Treg subtypes, nTreg, effect Treg, and Fr. III Treg, downloaded from the Gene Expression Omnibus: GSE90600 (Cuadrado et al., 2018); (4) the single-cell RNA-seq expression data of colon mucosa immune cells from ulcerative colitis patients and controls from the Single Cell Portal: SCP259 (Smillie et al., 2019); (5) the single-cell RNA-seq expression data of PBMCs from ulcerative colitis patients and controls from the Gene Expression Omnibus: GSE125527 (Boland et al., 2020); (6) the single-cell RNA-seq expression data of PBMCs from healthy controls from the Gene Expression Omnibus: GSE150728 (Wilk et al., 2020); (7) the single-cell RNA-seq expression data of PBMCs from SLE patients from the Gene Expression Omnibus: GSE137029; and (8) the single-cell RNA-seq expression data of PBMCs from multiple sclerosis patients from the Gene Expression Omnibus: GSE138266 (Schafflick et al., 2020).

Acquisition of gene sets

All gene sets used in this study and their citations are listed in Table S2. Marker gene lists of Th1 cells (CXCR3, TBX21, IFNG, GZMB, TNF), Th2 cells (GATA3, IL4, IL5, IL13), Th17 cells (CCR6, RORC, IL17A, IL17F), Treg cells (FOXP3, RTKN2, IKZF2, IL10, TGFB2, TGFB3, CTLA4, IL2RA), Tct cells (NKG7, GZMH, FGFBP2) and resting T cells (CCR7, SELL, CD27) were obtained from published studies (Luckheeram et al., 2012; Patil et al., 2018). Signature genes of naïve T (Tn), effector memory T (Tem), and central memory T (Tcm) cells were obtained by differential expression analysis of Tn/Tem/Tcm cells with all other cell types (Tn features: Tn vs Tem and Tcm; Tem features: Tem vs Tn and Tcm; Tcm features: Tn and Tem; $|\log_2$ fold change$|>1$, FDR<0.05) (Cuadrado et al., 2018).

However, we obtained 0 Tcm cell signature genes using this method, potentially because Tcm cells are intermediate between Tn and Tem cells. Risk genes of multiple autoimmune diseases, such as SLE, rheumatoid arthritis, type 1 diabetes inflammatory bowel disease, and ulcerative colitis, were obtained from their reported genes in the GWAS catalogue (v 1.0.2) (Welter et al., 2014). SLE $CD4^+$ T cell up-/downregulated gene sets were screened from the published dataset GSE10325 (Hutcheson et al., 2008) (p value<0.01, $\log_2$ fold change>1). The gene sets of biological processes in MSigDB (Liberzon et al., 2015) were also used in our analysis. In addition, we excluded the gene sets that were irrelevant to the immune functions of $CD4^+$ T cells (such as those related to the nervous system, embryonic development, reproduction and cellular dynamics) and the gene sets with more than 1500 or less than 10 genes. The final MSigDB gene sets used in our downstream analysis are listed in Table S2.

Screening the differential peaks between three SLE patient groups and healthy controls

Differential analysis was performed on $CD4^+$ T cell ATAC-seq samples from each SLE patient group with all samples from the healthy controls. Peaks with $|\log_2$ fold change$|>1.2$ and FDR<0.05 were defined as differential peaks. After unsupervised hierarchical clustering of the sample×peak count matrix, 11775 differential peaks were divided into Clusters I-V.

Functional annotation of peak clusters

We first obtained the regulatory regions of all genes using the ''basal plus extension'' model in GREAT (McLean et al., 2010) (proximal: 5 kb upstream, 1 kb downstream; plus distal: up to 500 kb). Then, we overlaid all peaks with these regulatory regions using ''intersectBed'' in bedtools (Quinlan and Hall, 2010) and constructed a gene×peak regulatory matrix in which each element was a 0 (peak not in the regulatory region of a gene) or 1 (peak in the regulatory region of a gene). The regulatory elements of a gene set were defined as peaks located in the regulatory regions of all genes in this gene set. For each gene set, we identified all its regulatory elements and then summarized the number of these regulatory elements in each peak cluster and the average number of regulatory elements per gene. We calculated the statistical significance of the enrichment of certain functional gene sets in a certain peak cluster compared with all peaks (background) by Fisher's exact test.

scRNA-seq primary data processing and gene expression imputation

Cell Ranger v2.0 was used to demultiplex the FASTQ reads, align them to the hg19 human transcriptome, and extract their ''cell'' and ''UMI'' barcodes. The output of this pipeline is a digital gene expression (DGE) matrix, which records the number of UMIs for each gene that are associated with each cell barcode. Next, we created the Seurat objects for three DGE matrixes of three batches using Seurat3.0, and merged them into a big Seurat object. Cells with fewer than 400 genes, more than 4000 genes and a percentage of mitochondrial genes greater than 5% were removed, and genes expressed in fewer than 10 cells were removed. Gene expression was normalized by ''LogNormalize'' in Seurat3.0. To further recover the expression of genes, we input a normalized expression matrix into the published imputation tool SAVER (Huang et al., 2018) with the default settings. The output of SAVER was the recovered gene expression matrix, which was used for downstream analysis.

Cell clustering

We first created the Seurat objects for all three batches of 10X single-cell RNA-seq data and identified the top 1000 variable features (genes in the Y chromosome were eliminated) individually for each object with default parameters. Next, we identified anchors for these three objects and integrated them with the top 15 canonical correlation analysis (CCA) dimensions. Then, we performed PCA on the integrated object and found cell clusters with the top 25 PCA dimensions (resolution set to 1).

Determination of the sample identity for each cell type

To capture interindividual variability in this population genetics study, we sequenced a large number of cells each from 4–6 individuals. Therefore, the sample (healthy controls and SLE patients) from which each cell type was derived was unknown. The published tool Demuxlet can be used to determine the sample identity of each single cell based on a comparison of single nucleotide polymorphisms (SNPs) between individuals and cells (Kang et al., 2018). To obtain the input file for Demuxlet (VCF file containing the SNPs of all samples), we first aligned the ATAC-seq sequencing reads to the reference genome hg19 using the ''–MappingQC'' module with option ''-c 100'' in ATAC-pipe. Then, we integrated these output bam files into a single mpileup file using ''samtools mpileup''. Next, SNPs were called by VarScan (Koboldt et al., 2012) with the options ''–min-coverage 5 –p-value 0.01 –output-vcf 1''. The output VarScan file was the VCF file that was used as the input file for Demuxlet, and Demuxlet was then run with default settings. Only the cells in the ''.best'' file (a Demuxlet output file) marked by ''SNG'' (singlet) could be determined. In total, we identified 47% of all cells. To further confirm the sample identities of more cells, we applied another tool, Souporcell (Heaton et al., 2020), to classify all $CD4^+$ T cells by genotype. A total of 93% of the $CD4^+$ T cells were divided into 10 clusters, and cells in the same cluster derived from the same sample. We identified the samples in each cell cluster by overlapping the results of Demuxlet and Souporcell.

Identifying the statistically significant differences in cell proportions

To identify changes in the proportion of each cell subset in blood from healthy controls and SLE patients, we used two statistical tests that each captured distinct but complementary types of information: (1) Student's t-test and (2) Fisher's exact test. The proportion of each $CD4^+$ T cell subtype was calculated for each sample according to the results of Demuxlet and Souporcell. We compared the

proportions of $CD4^+$ T cell subtypes between healthy controls and SLE patients (Student's t-test). However, a t-test was used to examine each cell subset independently; for cell proportions, an increase in the percentage of one cell subset would necessitate decreases in the percentages of other cell subsets. We therefore performed Fisher's exact test on the numbers of cells from each subset that were isolated from the blood of healthy individuals and patients with SLE, and the odds ratios and *p* values were used to measure the enrichment of each cell subset in each clinical state. These two methods reflected almost the same statistically significant differences.

Scoring gene signatures based on gene expression and chromatin accessibility

We defined the signature score to compare the degrees of enrichment of a gene set among cells or samples. For gene expression data, we first calculated the sums of the normalized expression across all genes in a signature gene set and then normalized them to values from 0 to 1 among cells/samples $(normalized\ L_i = (L_i - min(L))/(max(L) - min(L)))$, which was deemed the gene expression-signature score. For chromatin accessibility data (ATAC-seq), we first overlaid all ATAC-seq peaks with the regulatory regions (defined in "functional annotation of peak clusters") of all genes by "intersectBed" in bedtools and constructed connections between peaks and genes. We then calculated the sums of the normalized peak intensities across all peaks that were connected to genes in a signature gene set and then normalized them to values from 0 to 1 among samples. This was deemed the chromatin accessibility signature score.

Screening the differentially expressed genes between healthy controls and SLE patients for each $CD4^+$ T cell subtype

We first obtained the log-normalized gene expression matrix (after SAVER imputation) through the method described above, removed genes with a maximum expression value of less than 1 across all cells, and then converted the expression to a standardized z-score across cells. For each $CD4^+$ T cell subtype, the DEGs were screened by three thresholds: (1) *p* value <0.01 (Mann-Whitney U test), FDR <0.05 (Bonferroni correction), (2) average z score of cells from healthy controls (AZ_n) >0.1 or average z score of cells from SLE patients (AZ_p) >0.1, and (3) difference between AZ_n and AZ_p >0.4. The DEGs of each cell subset are listed in Table S3. We summarized the number of upregulated and downregulated DEGs for each cell subset and showed them on a t-SNE scatter plot. The DEGs between Treg cells from healthy controls and SLE patients were also screened in this manner and are listed in Table S4.

Construction of the functional change profiles in all cell subsets

To explore the functional differences in each cell subtype between healthy controls and SLE patients, we first calculated the difference in average expression (DAV) between healthy and SLE cells $(DAV = AZ_p - AZ_n)$ for all cell subtypes and obtained a differential expression matrix (DEM) in which each row was a gene and each column was a cell subtype. To construct a complete profile of functional changes, we used all MSigDB gene sets and created a 0/1 matrix in which each column was a functional gene set, each row was a gene, and each value was 0 (row gene was not in the column gene set) or 1 (row gene existed in the column gene set). The DEM and 0/1 matrix were used as the "input file" and "gene sets file" of Genomica, respectively, and "Module Map" (parameters: exclude gene sets with less than 3 genes or more than 1000 genes; expression level $\geq$0.3 denoted upregulation, expression level $\leq$ -0.3 denoted downregulation; value was displayed as the negative log *p* value of gene hit enrichment) was then run to obtain the enrichment of each gene set in each cell subset. The top enriched gene sets are shown in Table S3.

Estimating the differences in gene expression patterns between Treg cells from healthy controls and SLE patients

We first obtained the log-normalized gene expression matrix (after SAVER imputation) and calculated the average gene expression (AGE) for each $CD4^+$ T cell type from healthy controls and SLE patients. For each $CD4^+$ T cell subtype, genes with AGEs greater than 0.2 $(AGE_{cells\ from\ healthy\ controls} > 0.2\ or\ AGE_{cells\ from\ SLE\ patients} > 0.2)$ were used to calculate the Euclidean distance and Spearman correlation between healthy controls and SLE patients.

Screening of differentially expressed genes between two Treg clusters (Treg1 and Treg2)

We first obtained the log-normalized gene expression matrix (after SAVER imputation) through the method described above, removed genes with maximum expression values of less than 1 across all Treg cells, and then converted expression to a standardized z-score across all Treg cells. The DEGs between Treg1 and Treg2 cells were screened by a *p* value <0.01 (Mann-Whitney U test), FDR <0.05 (Bonferroni correction), and the difference in the average z score between Treg1 and Treg2 cells was more than 0.4 $(|AZ_{Treg1\ cells} - AZ_{Treg2\ cells}| > 0.4)$. The DEGs between Treg1 and Treg2 cells are listed in Table S4.

Identifying the SLE Treg2 up-regulated DEGs

We identified the SLE Treg2 up-regulated DEGs by the same thresholds described in "Screening of differentially expressed genes between two Treg clusters (Treg1 and Treg2).", and the interferon response genes were removed in the analysis of Figure 6A, the SLE Treg2 up-regulated DEGs were listed in Table S4.

Identifying the signature genes for three published Treg groups

Previous studies defined three fractions of $FOXP3^+CD4^+$ T cells: naïve Tregs (nTregs), effector Tregs (eTregs) and nonsuppression Tregs (Fr. III). We used the bulk RNA-seq data to identify their signature genes (Cuadrado et al., 2018). We performed differential

expression analysis (log_2 fold change>1, p value<0.05 (t-test) and FDR <0.5 (Bonferroni correction)) for each Treg group versus all other Treg groups (eTreg vs nTreg and Fr. III; nTreg vs eTreg and Fr. III; Fr. III vs eTreg and nTreg). The signature genes of these three Treg groups are listed in Table S2. The relative expression of all signature genes in eTreg, nTreg and Fr. III cells are shown in the heatmap.

Screening specific genes for Th1, Th2, Th17, and Tct cells

We used the single-cell gene expression data of healthy controls to identify the specific genes of Th1, Th2, Th17 and Tct cells. Differential expression analysis (log_2 fold change >0.2 and Mann-Whitney U test p value <0.01) was performed for each subtype of Teff cells versus all other subtypes of Teff cells. The specific genes of each Teff subtype are listed in Table S4.

Identifying the Treg exhaustion-associated functional gene set

We used two different methods that each captured distinct types of information to identify Treg exhaustion-associated functions: (1) Metascape gene list annotation and (2) GSEA. We described each of these below. We first calculated the Pearson's correlation between normalized gene expression and the signature score of Treg exhaustion-like properties for all detected genes across all Treg cells and then identified the genes whose expression was significantly correlated with the signature score of Treg exhaustion-like properties (correlation >0.4). We uploaded this gene list to the Metascape main page and clicked ''express analysis''. The Metascape's report page showed the functional annotations of Treg exhaustion-associated genes. We showed the top enriched functions via a bar plot. To further verify the correlation between Treg exhaustion-like properties and the type I IFN response, we performed GSEAPreranked analysis of Treg-detected genes (ranked by their Pearson correlation with the signature score of Treg exhaustion-like properties) and a type I IFN response gene list.

QUANTIFICATION AND STATISTICAL ANALYSIS

The analysis, software, and quantification methodology that are specific to ATAC-seq, RNA-seq, and scRNA-seq experiments are included under the relevant subsections of the method details section. Information regarding replicate numbers is provided in figure legends. If error bars are used in figures, information about what error bars represent is also provided in the figure legend. If the degree of significance is provided in the figure legend, further details regarding the statistical test used are provided in the relevant subsections of the method details that are specific to the analysis being performed.

Flow cytometry data were analysed using FlowJo V.X.0.7 software (Tree Star). Statistical analyses and approximations were performed with GraphPad Prism 7 software (GraphPad).

Statistical significance was analyzed with two-tailed Student's-tests, Mann-Whitney U test, or two-way ANOVA. A p-value of less than 0.05 was considered statistically significant, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$. Data are presented as the upper, centre, and lower lines indicate the 75% quantile +1.5 * interquartile range (IQR), 50% quantile, and 25% quantile −1.5 * IQR, respectively which are indicated in the figure legends.

Cell Reports, Volume *41*

Supplemental information

Single-cell transcriptome profiling and chromatin accessibility reveal an exhausted regulatory CD4+ T cell subset in systemic lupus erythematosus

Chuang Guo, Qian Liu, Dandan Zong, Wen Zhang, Zuqi Zuo, Qiaoni Yu, Qing Sha, Lin Zhu, Xuyuan Gao, Jingwen Fang, Jinhui Tao, Quan Wu, Xiaomei Li, and Kun Qu

Supplemental Figures

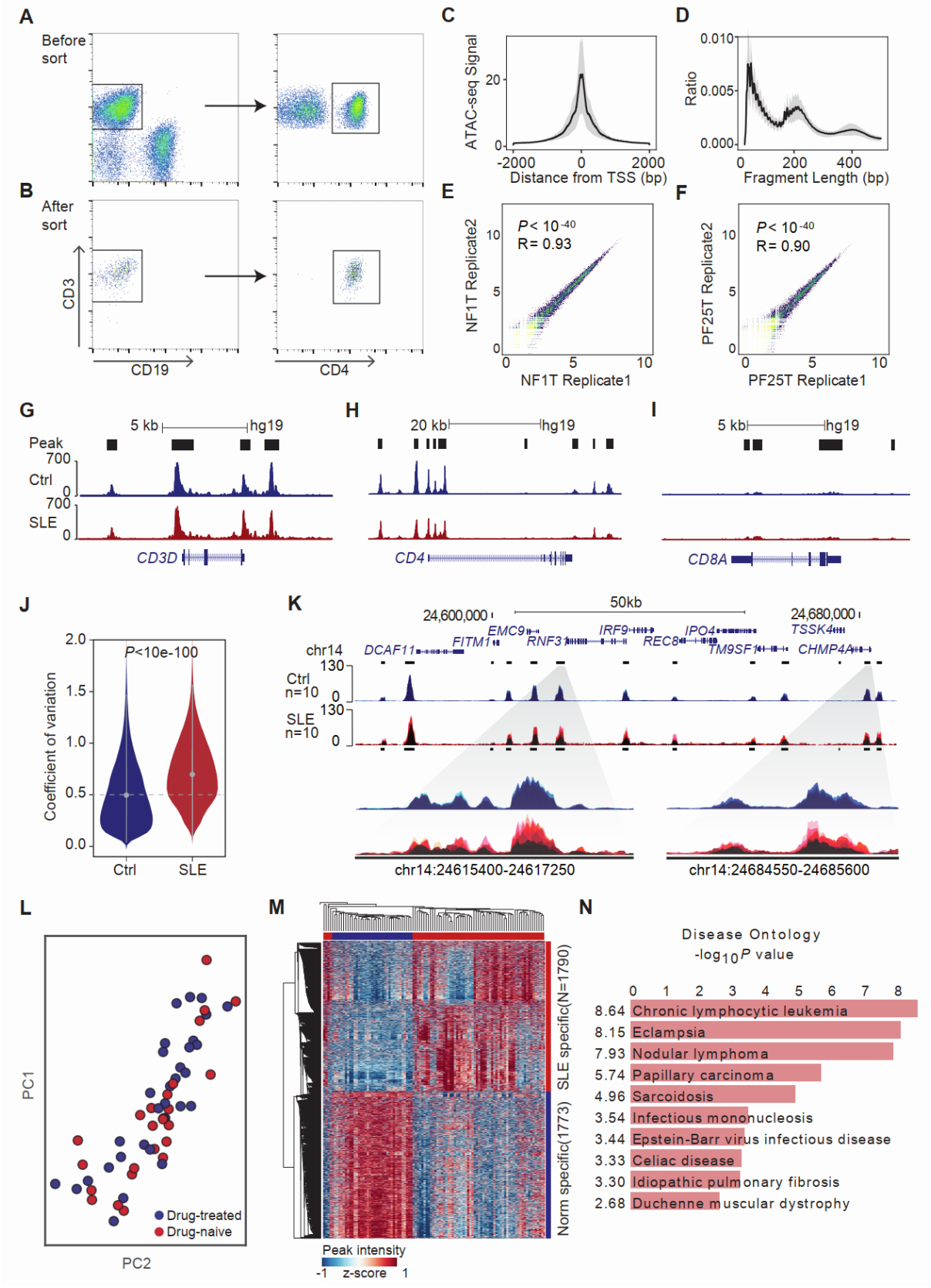


Figure S1. Heterogeneity of the chromatin accessibility of SLE CD4$^+$ T cells. Related to Figure 1.

(A-B) Pseudocolor dot plots showing the gating strategy of peripheral CD4$^+$ T cells (A) and a representative example for the detection of the purity of sorted CD4$^+$ T cells (B).

(C-D) Plots of insert size distributions (C) and transcription start site (TSS) enrichment scores (D) of all ATAC-seq samples from healthy controls and SLE patients. Light shades indicate the 95% confidence interval.

(E-F) Scatter plots of ATAC-seq signals for all merged peaks from a pair of biological replicates from healthy controls (E) and SLE patients (F).

(G-I) Normalized ATAC-seq signal profiles around *CD3D* (G), *CD4* (H), *CD8A* (I) in merged samples from healthy controls (upper, n = 25) and SLE patients (lower, n = 63).

(J) Violin plot showing the coefficient of variation of all peaks of samples from healthy controls (blue) and SLE patients (red).

(K) The UCSC genome browser plot showing the ATAC-seq signal intensities of samples from 10 random-selected healthy controls (blue) and from 10 random-selected SLE patients (red). The bottom rows showed the detailed accessible chromatin at two selected regulatory regions.

(L) Principal component analysis of the top variant (coefficient of variation > 0.2) chromatin accessible regions for ATAC-seq samples from 32 drug-treated SLE patients (blue) and 24 drug-naive SLE patients (red). Each dot represents an SLE sample, and samples are color-coded by their clinical groups.

(M) Heatmap showing differentially accessible peaks between CD4$^+$ T cells from healthy controls and those from SLE patients.

(N) Bar plot showing the disease ontology of highly accessible peaks in SLE patients compared with healthy controls.

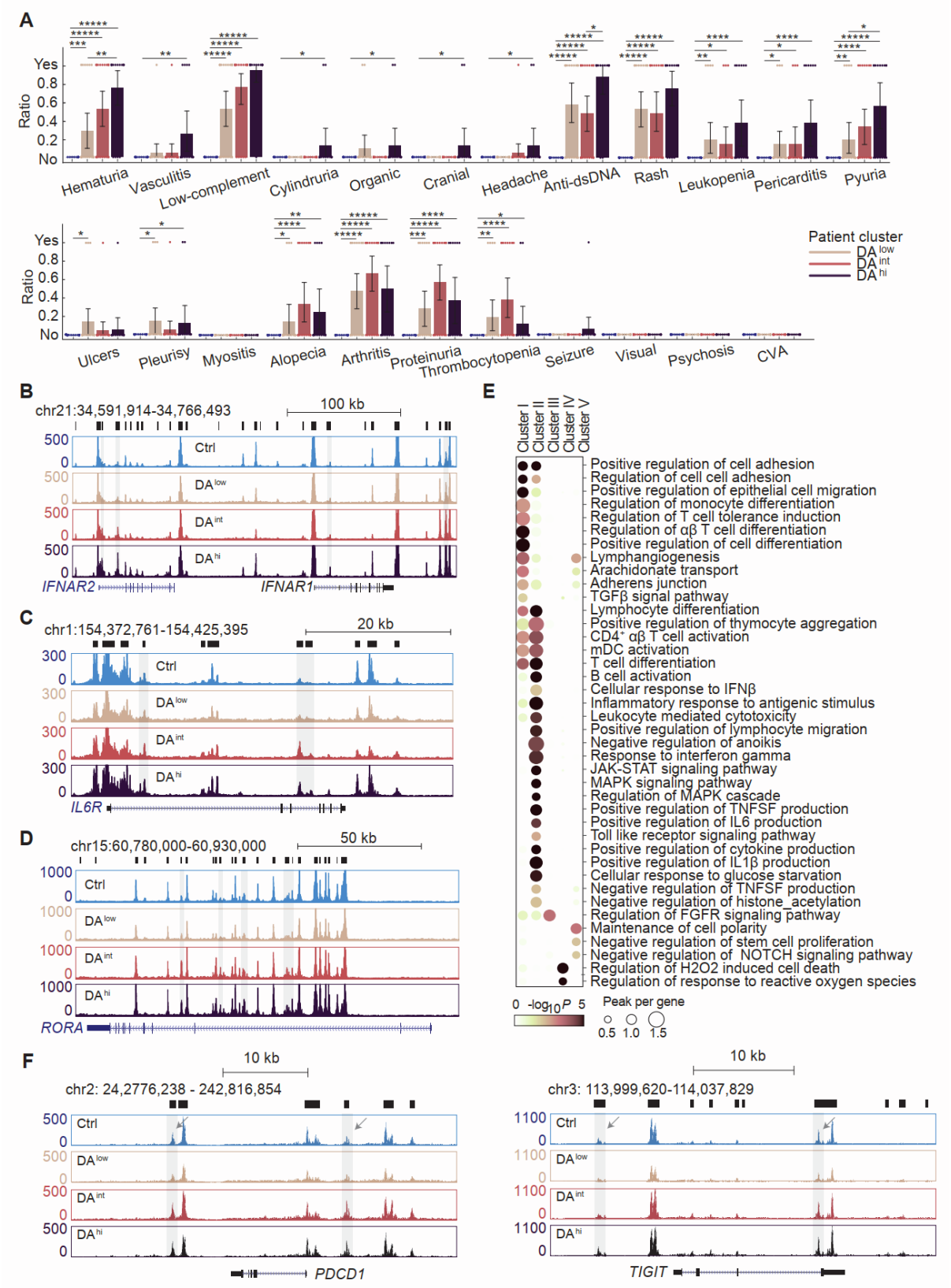

A
Ratio
Yes
No
Hematuria
Vasculitis
Low-complement
Cylindruria
Organic
Cranial
Headache
Anti-dsDNA
Rash
Leukopenia
Pericarditis
Pyuria
Ulcers
Pleurisy
Myositis
Alopecia
Arthritis
Proteinuria
Thrombocytopenia
Seizure
Visual
Psychosis
CVA
Patient cluster
DA low
DA int
DA hi
B
chr21:34,591,914-34,766,493
100 kb
Ctrl
IFNAR2
IFNAR1
C
chr1:154,372,761-154,425,395
20 kb
IL6R
D
chr15:60,780,000-60,930,000
50 kb
RORA
E
Cluster I
Cluster II
Cluster III
Cluster IV
Cluster V
Positive regulation of cell adhesion
Regulation of cell cell adhesion
Positive regulation of epithelial cell migration
Regulation of monocyte differentiation
Regulation of T cell tolerance induction
Regulation of αβ T cell differentiation
Positive regulation of cell differentiation
Lymphangiogenesis
Arachidonate transport
Adherens junction
TGFβ signal pathway
Lymphocyte differentiation
Positive regulation of thymocyte aggregation
CD4+ αβ T cell activation
mDC activation
T cell differentiation
B cell activation
Cellular response to IFNβ
Inflammatory response to antigenic stimulus
Leukocyte mediated cytotoxicity
Positive regulation of lymphocyte migration
Negative regulation of anoikis
Response to interferon gamma
JAK-STAT signaling pathway
MAPK signaling pathway
Regulation of MAPK cascade
Positive regulation of TNFSF production
Positive regulation of IL6 production
Toll like receptor signaling pathway
Positive regulation of cytokine production
Positive regulation of IL1β production
Cellular response to glucose starvation
Negative regulation of TNFSF production
Negative regulation of histone_acetylation
Regulation of FGFR signaling pathway
Maintenance of cell polarity
Negative regulation of stem cell proliferation
Negative regulation of NOTCH signaling pathway
Regulation of H2O2 induced cell death
Regulation of response to reactive oxygen species
0 -log10P 5
Peak per gene
0.5 1.0 1.5
F
10 kb
chr2: 24,2776,238 - 242,816,854
PDCD1
10 kb
chr3: 113,999,620-114,037,829
TIGIT

Figure S2. Functional annotation of peaks of Cluster I-V in DA^{low}, DA^{int}, and DA^{hi} SLE patient. Related to Figure 2.

(A) Bar plots showing the proportion of samples with different comorbidities in healthy controls and in DA^{low}, DA^{int}, and DA^{hi} SLE patients.

(B-D) Normalized ATAC-seq profiles at the loci *IFNAR1* (B), *IL6R* (C), and *RORA* (D).

(E) Functional annotations of peaks in Cluster I-V. Shown are the enrichment *P* value (dot color) and the average peak number per gene (dot size) of MSigDB gene sets in Cluster I-V peaks.

(F) Normalized ATAC-seq profiles at the loci *PDCD1* and *TIGIT.*

A

Batch name	Number of samples that pooled together in each library	Captured cells	Median genes per cell	Mean reads per cell	Cells passed QC
Batch1	6 Controls	10715	1148	58569	10712
Batch2	6 SLE patients	11922	1237	52984	11922
Batch3	4 SLE patients	11816	837	36966	11542

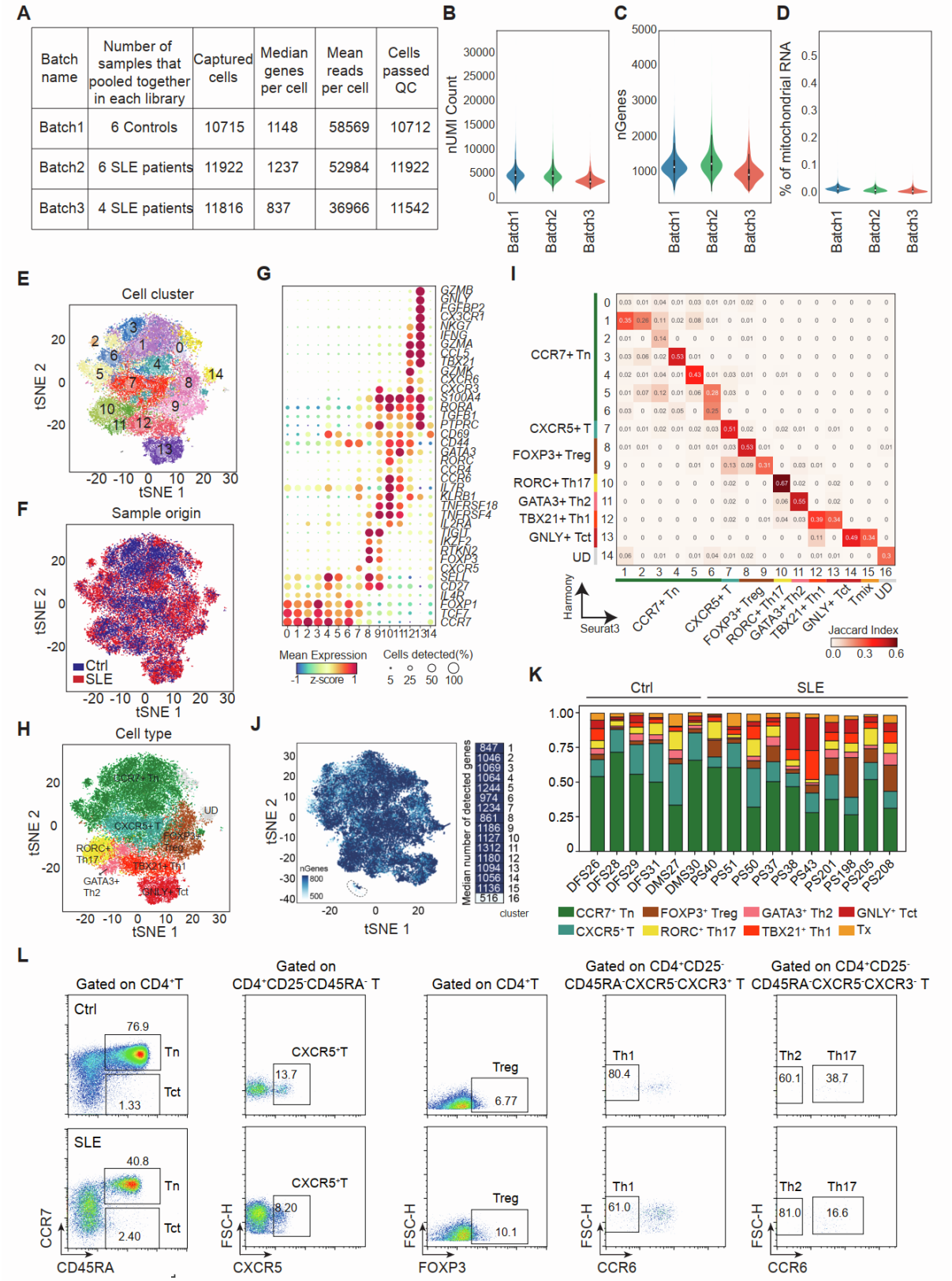


Figure S3. Quality control and integrative analysis of the single-cell dataset for CD4$^+$ T cells from healthy controls and SLE patients. Related to Figure 3.

(A) Summary of number of samples that pooled together in each library, captured cells, median genes per cell, median reads per cell per cell, and the number of cells that passed quality control (QC) in distinct batches of single-cell data from healthy controls or SLE patients.

(B-D) Box plots showing the UMI number (B), gene number (C), and percentage of mitochondrial RNA (D) in distinct batches of the single-cell data from healthy controls and SLE patients.

(E-F) t-SNE of single-cell transcriptomes of 34,177 CD4$^+$ T cells integrated by Harmony, with each cell color-coded by cell clusters (E) and sample type of origin (F).

(G) Heatmap showing the expression of selected marker genes for CD4$^+$ T subtypes in the 15 cell clusters. The dot size represents the percentage of cells expressing the selected marker genes (right panel). The dot color represents the mean expression levels of selected marker genes.

(H) tSNE of the 34,177 CD4$^+$ T cells integrated by Harmony, with each cell color-coded by cell subtypes.

(I) Jaccard similarities between the cell clusters with the integration processed by Seurat3.0 and with the integration processed by Harmony.

(J) t-SNE of single-cell transcriptomes of CD4$^+$ T cells colored by the mean number of detected genes. The right panel shows the mean number of detected genes in each cell cluster.

(K) Bar plot showing the proportion of CCR7$^+$ Tn, CXCR5$^+$ T, Treg, Th17, Th2, Th1, Tct, and Tx cells for each healthy control individual and each SLE patient.

(L) The representative pseudocolor dot plots showing the proportions of CCR7$^+$ T cells, CXCR5$^+$ T cells, FOXP3$^+$ Treg cells, Th1 cells, Th2 cells, Th17 cells, and Tct cells among the gated CD4$^+$ T cells in SLE patients and in healthy control individuals.

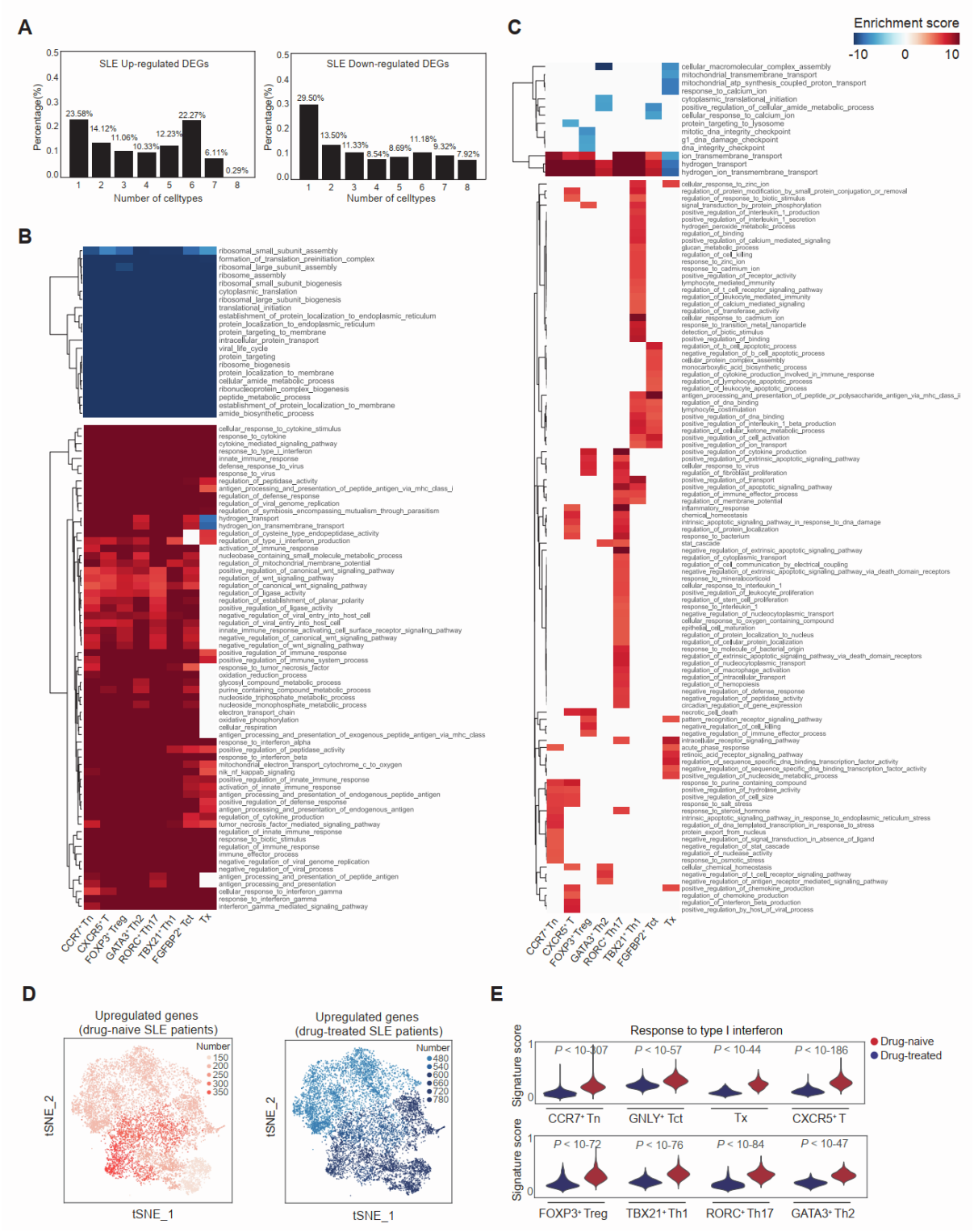

A
SLE Up-regulated DEGs
SLE Down-regulated DEGs
Percentage(%)
Number of celltypes
23.58%
14.12%
11.06%
10.33%
12.23%
22.27%
6.11%
0.29%
29.50%
13.50%
11.33%
8.54%
8.69%
11.18%
9.32%
7.92%
B
C
Enrichment score
-10
0
10
CCR7+ Tn
CXCR5+ T
FOXP3+ Treg
GATA3+ Th2
RORC+ Th17
TBX21+ Th1
FGFBP2+ Tcl
Tx
D
Upregulated genes
(drug-naive SLE patients)
Upregulated genes
(drug-treated SLE patients)
Number
tSNE_1
tSNE_2
E
Response to type I interferon
Signature score
Drug-naive
Drug-treated
P < 10-107
P < 10-57
P < 10-44
P < 10-186
P < 10-72
P < 10-76
P < 10-84
P < 10-47
CCR7+ Tn
GNLY+ Tcl
Tx
CXCR5+ T
FOXP3+ Treg
TBX21+ Th1
RORC+ Th17
GATA3+ Th2

Figure S4. The functional divergence of CD4$^+$ T cell subtypes in SLE patients. Related to Figure 4.

(A) Bar plot showing the percentages of upregulated DEGs (left) and downregulated DEGs (right) shared by different number of CD4$^+$ T cell subtypes between healthy controls and SLE patients.

(B) Heatmap showing the functional annotations of shared DEGs (n > 5) in CD4$^+$ T cell subtypes. Upper panel shows the biological functions enriched in CD4$^+$ T cell subtypes from healthy controls (blue). Lower panel shows the biological functions enriched in CD4$^+$ T cell subtypes from SLE patients (red).

(C) Heatmap showing the functional annotations of specific DEGs (n < 3) in CD4$^+$ T cell subtypes. Upper panel shows the biological functions enriched in CD4$^+$ T cell subtypes from healthy controls (blue). Lower panel shows the biological functions enriched in CD4$^+$ T cell subtypes from SLE patients (red).

(D) tSNE showing the numbers of upregulated DEGs (left) and downregulated DEGs (right) specific to SLE in all CD4$^+$ T cell subtypes between drug-treated SLE patients and drug-naive SLE patients.

(E) Violin plot showing the signature score of response to type I interferon in CD4$^+$ T cell subtypes from drug-treated SLE patients and drug-naive SLE patients (Mann-Whitney U test).

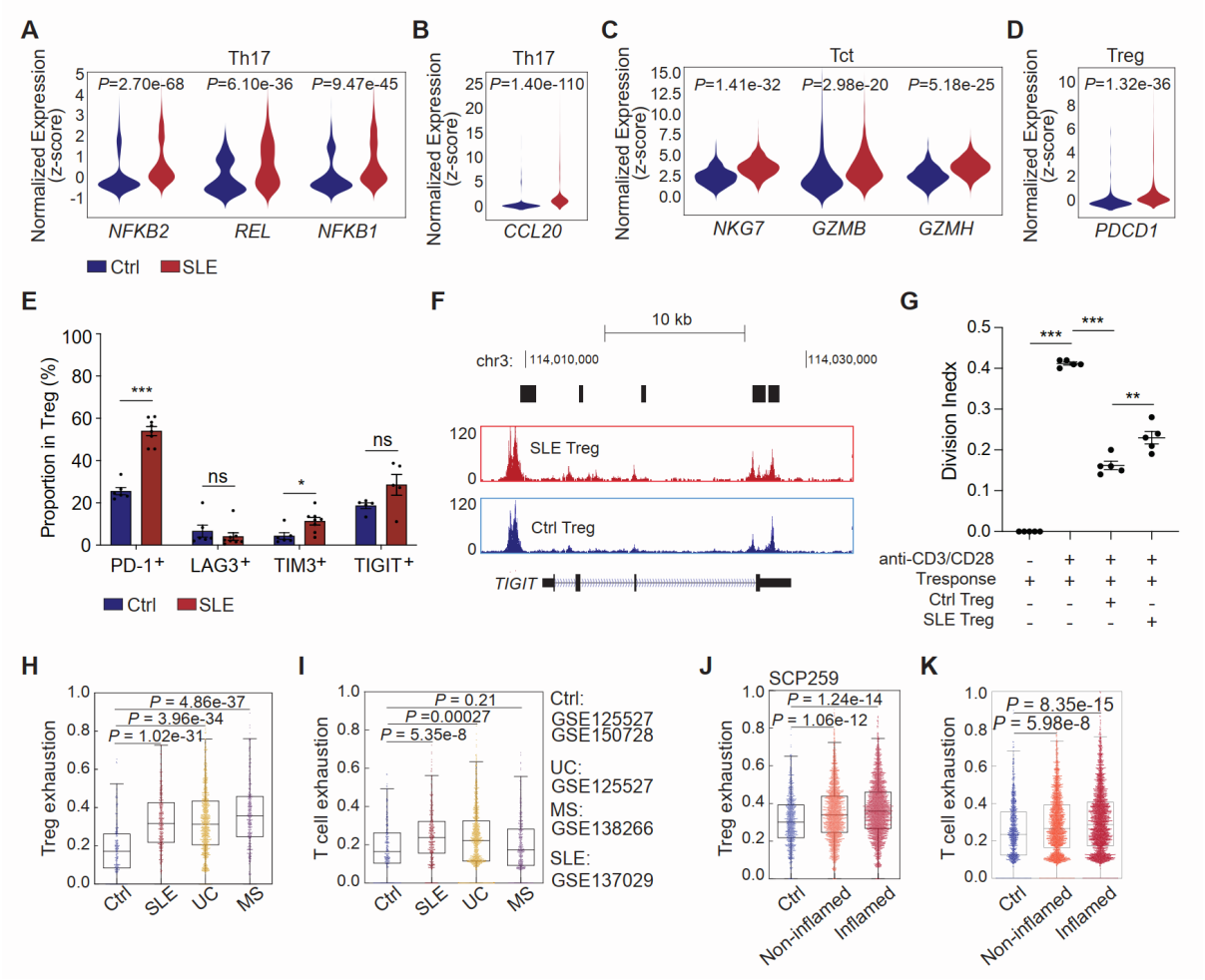

A
Th17
Normalized Expression (z-score)
P=2.70e-68
P=6.10e-36
P=9.47e-45
NFKB2
REL
NFKB1
Ctrl
SLE
B
Th17
P=1.40e-110
CCL20
C
Tct
P=1.41e-32
P=2.98e-20
P=5.18e-25
NKG7
GZMB
GZMH
D
Treg
P=1.32e-36
PDCD1
E
Proportion in Treg (%)
PD-1+
LAG3+
TIM3+
TIGIT+
ns
*
***
F
10 kb
chr3:
114,010,000
114,030,000
SLE Treg
Ctrl Treg
TIGIT
G
Division Inedx
anti-CD3/CD28
Tresponse
Ctrl Treg
SLE Treg
H
Treg exhaustion
P = 4.86e-37
P = 3.96e-34
P = 1.02e-31
Ctrl
SLE
UC
MS
I
T cell exhaustion
P = 0.21
P =0.00027
P = 5.35e-8
Ctrl:
GSE125527
GSE150728
UC:
GSE125527
MS:
GSE138266
SLE:
GSE137029
J
SCP259
Treg exhaustion
P = 1.24e-14
P = 1.06e-12
Non-inflamed
Inflamed
K
T cell exhaustion
P = 8.35e-15
P = 5.98e-8

Figure S5. DEGs between healthy controls and SLE patients in $CD4^+$ T subtypes and exhaustion signature score in several autoimmune diseases. Related to Figure 4.

(A-B) Violin plot showing the normalized expression of *NFKB2*, *REL*, *NFKB1*, (A) and *CCL20* (B) in Th17 cells from healthy controls and SLE patients, Mann-Whitney U test.

(C) Violin plot showing the normalized expression of *NKG7*, *GZMB*, and *GZMH* in Tct cells from healthy controls and SLE patients, Mann-Whitney U test.

(D) Violin plot showing the normalized expression of *PDCD1* in Treg cells from healthy controls and SLE patients, Mann-Whitney U test.

(E) Bar plots displaying the frequencies of PD-1, LAG3, TIM3, and TIGIT in $FOXP3^+$ Treg cells from SLE patients and healthy control individuals. * $P < 0.05$; *** $P < 0.001$; Student's t-test.

(F) Normalized ATAC-seq profiles of Treg cells from SLE patients and healthy controls at *TIGIT* locus.

(G) The ratios of the division index in effector $CD4^+CD25^-$ T cells from healthy controls and SLE patients.

(H-I) Treg exhaustion (H) and T cell exhaustion (I) signature score in peripheral blood Treg cells from healthy controls and from patients with SLE, Ulcerative Colitis (UC), Multiple Sclerosis (MS). Mann-Whitney U test.

(J-K) Treg exhaustion (J) and T cell exhaustion (K) signature score in Treg cells from intestinal mucosa from healthy controls and from patients with non-inflamed and inflamed biopsies of Ulcerative Colitis. Mann-Whitney U test.

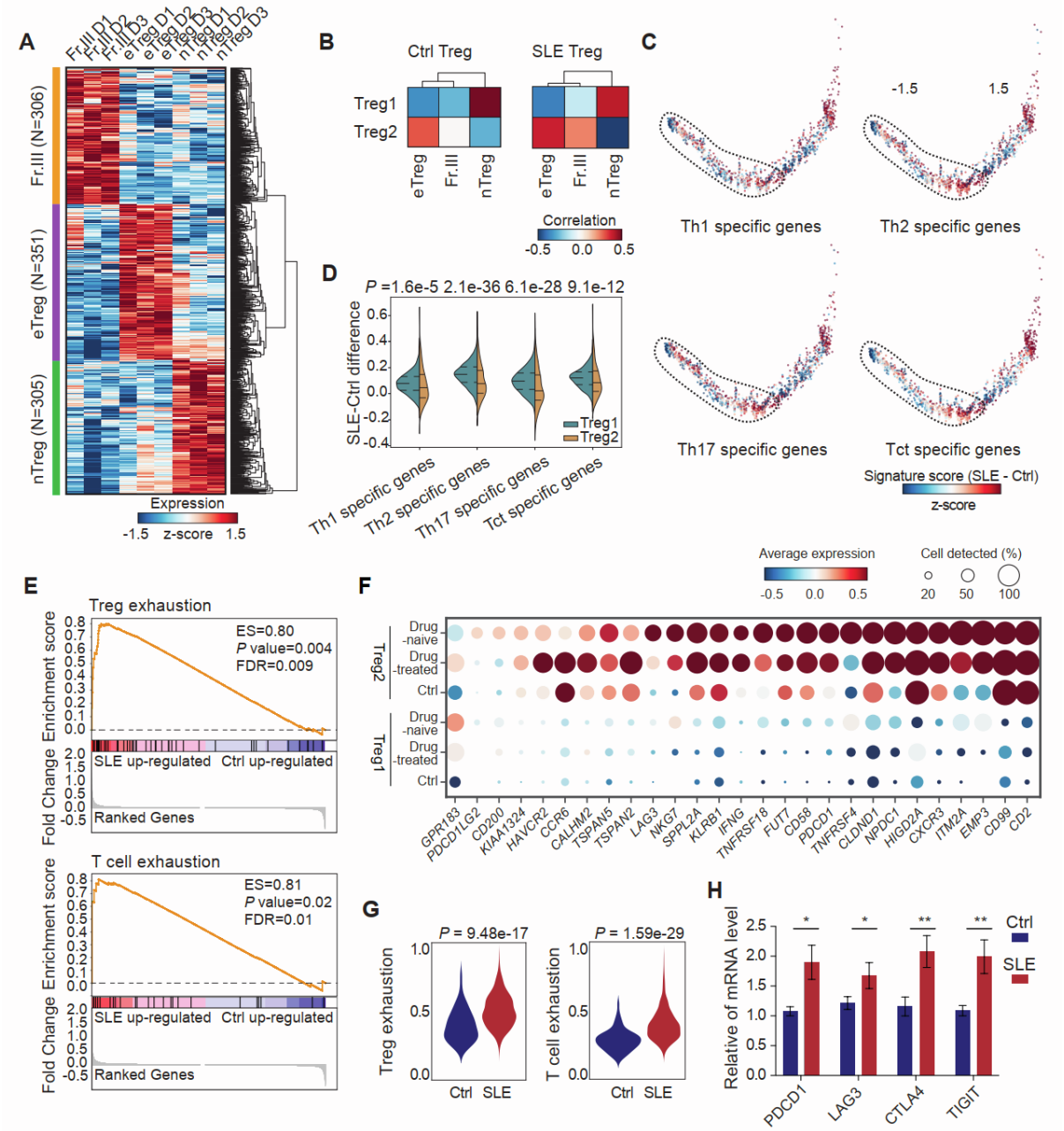

A
Fr.III D1
Fr.III D2
Fr.III D3
eTreg D1
eTreg D2
eTreg D3
nTreg D1
nTreg D2
nTreg D3
Fr.III (N=306)
eTreg (N=351)
nTreg (N=305)
Expression
-1.5 z-score 1.5
B
Ctrl Treg
SLE Treg
Treg1
Treg2
eTreg
Fr.III
nTreg
Correlation
-0.5 0.0 0.5
C
-1.5
1.5
Th1 specific genes
Th2 specific genes
Th17 specific genes
Tct specific genes
Signature score (SLE - Ctrl)
z-score
D
P =1.6e-5 2.1e-36 6.1e-28 9.1e-12
SLE-Ctrl difference
Treg1
Treg2
Th1 specific genes
Th2 specific genes
Th17 specific genes
Tct specific genes
E
Treg exhaustion
ES=0.80
P value=0.004
FDR=0.009
Fold Change Enrichment score
SLE up-regulated
Ctrl up-regulated
Ranked Genes
T cell exhaustion
ES=0.81
P value=0.02
FDR=0.01
F
Average expression
-0.5 0.0 0.5
Cell detected (%)
20 50 100
Treg2
Treg1
Drug -naive
Drug -treated
Ctrl
GPR183
PDCD1LG2
CD200
KIAA1324
HAVCR2
CCR6
CALHM2
TSPAN5
TSPAN2
LAG3
NKG7
SPPL2A
KLRB1
IFNG
TNFRSF18
FUT7
CD58
PDCD1
TNFRSF4
CLDND1
NPDC1
HIGD2A
CXCR3
ITM2A
EMP3
CD99
CD2
G
P = 9.48e-17
P = 1.59e-29
Treg exhaustion
T cell exhaustion
Ctrl SLE
H
Relative of mRNA level
Ctrl
SLE
PDCD1
LAG3
CTLA4
TIGIT

Figure S6. Treg cell exhaustion in SLE patients. Related to Figures 5 and 6.

(A) Heatmap showing the DEGs between nTreg, eTreg and Fr. III Treg cells.

(B) Heatmap showing the pearson correlation of the expression of DEGs (shown in A) between nTreg, eTreg and Fr. III Treg cells with Treg1 and Treg2 cells from healthy controls (left) and those from SLE patients (right).

(C) Trajectory of Treg cells determined using the Monocle 2 algorithm. The differences in the signature scores of Th1, Th2, Th17, and Tct specific genes between SLE patients and healthy controls (SLE - Ctrl) were coloured.

(D) Violin plot showing the difference of the signature scores of Th1, Th2, Th17, and Tct specific genes (SLE - Ctrl) between Treg1 and Treg2 cells. Mann-Whitney U test.

(E) Gene set enrichment analysis (GSEA) showing the enrichment of Treg exhaustion and T cell expression in pre-ranked genes. Genes were ranked by the differences of the expression in Treg2 cells between SLE and healthy controls.

(F) Scatter plot showing the expression of T cell exhaustion signature genes in Treg1 and Treg2 cells from healthy controls, drug-treated SLE patients, and drug-naive SLE patients. The dot color represents the mean gene expression level. The dot size represents the proportion of cells expressing the gene.

(G) Violin plot showing the signature score of Treg exhaustion-like properties (left panel) and T cell exhaustion (right panel) in Tregs from healthy controls and SLE patients (Student's t-test).

(H) qPCR showing the expression of *TIGIT, CTLA4, PDCD1* and *LAG3* in Treg2 cells from healthy controls and SLE patients.

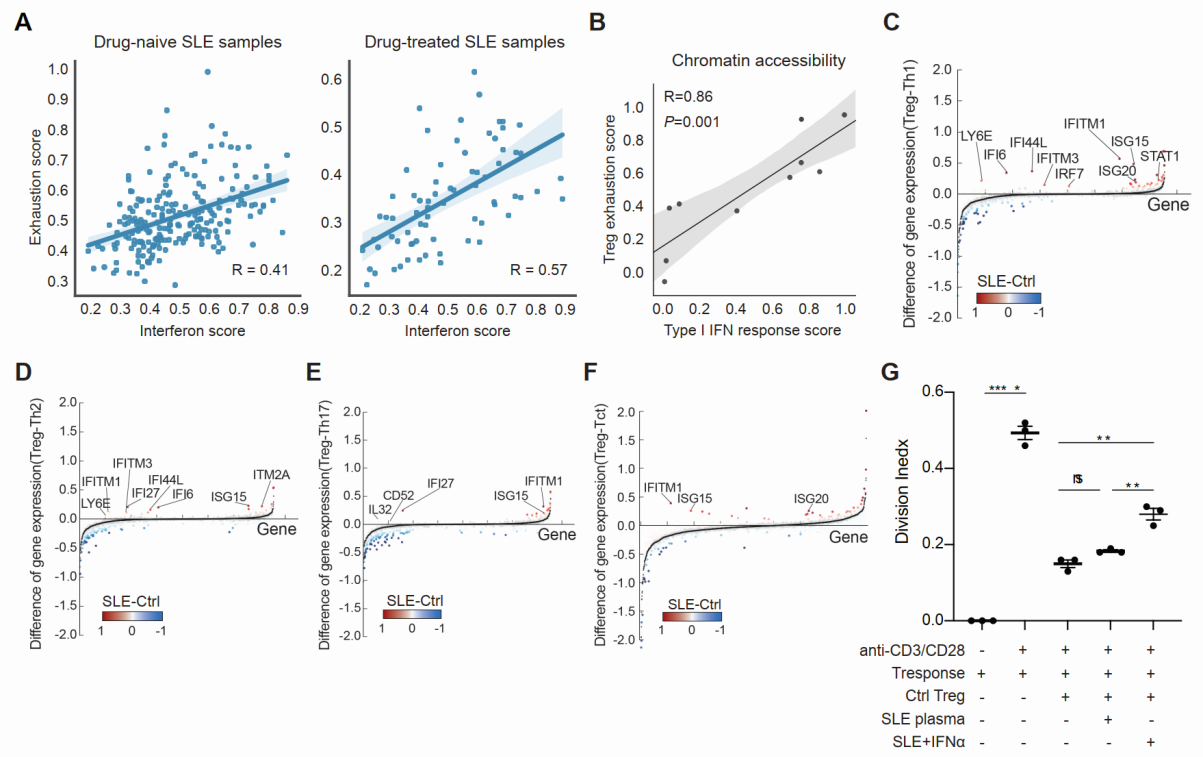


Figure S7. Type I interferon signalling is associated with Treg cell exhaustion-like properties in SLE patients. Related to Figure 7.

(A) Scatter plot representing the linear regressions between the signature scores of the type I IFN response and Treg exhaustion in Treg cells from drug-treated SLE patients and drug-naive SLE patients. The two-tailed t-statistic P value and Pearson's correlation coefficient (R) are shown at the bottom.

(B) Scatter plot representing the linear regressions of chromatin accessibility signature score of Treg exhaustion and those of type I interferon response. The *P* value and correlation R were estimated by the Pearson correlation analysis. Two-tailed t-statistic *P* value and coefficient (R) of Pearson's correlation were shown in top left.

(C-F) Difference of gene expression between Treg and Th1 (C), Th2 (D), Th17 (E), and Tct (F) in Treg cells from healthy controls (black dot) and SLE Treg cells (colored by the difference between SLE and Ctrl). Each dot is a gene, ranked by difference between cells from SLE patients and healthy controls.

(G) The ratios of the division index in effector $CD4^{+}CD25^{-}$ T cells from healthy controls (Ctrl) with persistent stimulation of 10% SLE plasma (SLE plasma) and 0.5 μg/mL IFNα plus 10% SLE plasma (SLE+IFNα) for 7 days.