IMMUNE NETWORK

Original Article

Check for updates

Simultaneous Epigenetic and Gene Expression Profiling at Single Cell Resolution Uncovers Stem-Like Treg Subsets Induced With Oligonucleotide Expansion in Humans

*Correspondence to
Jong-Il Kim
Genomic Medicine Institute, Medical Research Center, Seoul National University, 103 Daehak-ro, Jongno-gu, Seoul 03080, Korea.
Email: jongil@snu.ac.kr

Hyun Je Kim
Department of Biomedical Sciences, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 03080, Korea.
Email: hjkim0518@gmail.com

†Hyo Jeong Nam and Jeong-Ryeol Gong contributed equally to this work.

ORCID iDs
Hyo Jeong Nam
https://orcid.org/0000-0003-4987-0747
Jeong-Ryeol Gong
https://orcid.org/0000-0001-7057-4489
Yong-Hee Kim
https://orcid.org/0000-0001-6282-2980

Hyo Jeong Nam [1,2,†], **Jeong-Ryeol Gong** [3,†], **Yong-Hee Kim** [1,4], **Thuy Nguyen-Phuong** [1,2], **Nari Byun** [1,2,5], **Jeong Heon Yoon** [5], **Yong Chan Kim** [5], **Hyunwoo Chung** [1], **Brian Hyohyoung Lee** [1,2], **Haeyoon Kwon** [6], **Woochan Lee** [7], **Sung-Jun Kang** [1,2], **Kyunghyuk Park** [7], **Bukyoung Cha** [7], **Jong-Il Kim** [1,7,*], **Hyun Je Kim** [1,2,7,8,*]

[1]Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul 03080, Korea
[2]Department of Microbiology and Immunology, Seoul National University College of Medicine, Seoul 03080, Korea
[3]Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea
[4]Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA
[5]Teraimmune, Inc., Gaithersburg, MD 20878, USA
[6]Department of Medicine, Seoul National University College of Medicine, Seoul 03080, Korea
[7]Genomic Medicine Institute, Medical Research Center, Seoul National University, Seoul 03080, Korea
[8]Department of Dermatology, Seoul National University Hospital, Seoul 03080, Korea

ABSTRACT

Tregs play a central role in maintaining immune tolerance. Recent progress in the clinical application of Tregs underscores their potential for cell therapy. Nevertheless, a notable hurdle remains in producing functional Tregs *in vitro*. There is also a lack of detailed studies evaluating the function of Tregs during their *ex vivo* expansion process. Our prior investigation showed that the *ex vivo* expansion with oligonucleotides produces FoxP3[high]Helios[high] subsets. To investigate how oligonucleotides in culture media influence on gene expression and epigenetic states at single cell resolution, we sorted Tregs from healthy individuals and profiled *in vitro* oligonucleotide-expanded and non-expanded Tregs. We discovered a subset of Tregs, specifically enriched in expanded Tregs (seTregs), through oligonucleotide-induced expansion. seTregs showed an enhancement in both stem-like characteristics and functional attributes. Through analysis of histone modification data and gene regulatory networks, we elucidated *IKZF2* (Helios) as a pivotal transcription factor in generating these cell subsets. We believe these findings offer insights into evaluating functional regulation of *in vitro* expanded Tregs aimed at manufacturing Treg-based cell therapies.

Keywords: Regulatory T cell; Epigenetics; Transcriptomics

Thuy Nguyen-Phuong (iD)
https://orcid.org/0000-0002-8782-8163
Nari Byun (iD)
https://orcid.org/0000-0002-0508-6910
Hyunwoo Chung (iD)
https://orcid.org/0000-0001-9103-0560
Brian Hyohyoung Lee (iD)
https://orcid.org/0000-0002-1291-3208
Haeyoon Kwon (iD)
https://orcid.org/0009-0001-3422-0007
Woochan Lee (iD)
https://orcid.org/0000-0002-1092-7331
Sung-Jun Kang (iD)
https://orcid.org/0000-0001-8126-4072
Kyunghyuk Park (iD)
https://orcid.org/0000-0002-3141-5087
Bukyoung Cha (iD)
https://orcid.org/0000-0002-2859-0685
Jong-Il Kim (iD)
https://orcid.org/0000-0002-7240-3744
Hyun Je Kim (iD)
https://orcid.org/0000-0003-4467-0949

Abbreviations
ATAC, assay for transposase-accessible chromatin; ChIP-seq, chromatin immunoprecipitation sequencing; DEG, differentially expressed gene; DORC, Domain of Regulatory Chromatin; ETS1, erythroblast transformation-specific proto-oncogene 1; GO, gene ontology; GSEA, gene set enrichment analysis; NES, normalized enrichment score; LSI, latent semantic indexing; RUNX1, runt-related transcription factor 1; seTregs, specifically enriched in expanded Tregs; snATAC-seq, single-nucleus assay for transposase-accessible chromatin sequencing; snRNA-seq, single-nucleus RNA sequencing; TF, transcription factor; UMAP, Uniform Manifold Approximation and Projection.

INTRODUCTION

The human immune system employs complex mechanisms to defend against foreign entities such as protozoa, fungi, bacteria, and viruses. However, an overactive immune response to non-self Ags may damage tissue. To mitigate this, the immune system has evolved regulatory mechanisms, including Tregs (1). Following its discovery by Shimon Sakaguchi, three separate research teams independently identified FOXP3 as the key transcription factor (TF) directing the functional program of Tregs (2-4). Therefore, enhancing the expression of FOXP3 and other TFs, or modifying epigenetic regulations that control their expression, may offer viable strategies to improve the functionality of Tregs (5).

Numerous studies have emphasized the effectiveness of *ex vivo* expanded Treg-infusion therapy in addressing diseases such as graft-versus-host disease (6). Phase I clinical trials have additionally shown the potential of Treg therapy in alleviating autoimmune disorders, including type 1 diabetes (NCT02772679, NCT01210664), systemic lupus erythematosus (NCT02428309), inflammatory bowel disease (NCT03185000), Alzheimer's disease (NCT03865017), and pemphigus (NCT03239470).

A significant barrier to the clinical implementation of Treg therapies is the challenge of obtaining sufficient numbers of functionally stable Tregs via *ex vivo* expansion. Prior work, however, as reported by Kim et al. (7), has shown that the use of DNA oligonucleotides in Treg cultures promotes stable activation of these cells following isolation from human donors. Yet, to date, such cells have not been scrutinized at the single cell level.

In our current investigation, Tregs sourced from various healthy donors were treated with DNA oligonucleotides, cryopreserved, and subsequently subjected to multiome analyses to concurrently assess the transcriptome and epigenome within individual cells. Given that heterogeneity of Tregs was largely governed by epigenetic modifications influencing gene expression, discerning these epigenetic alterations alongside gene expression was suitable for a comprehensive annotation and comparative analysis between different Treg populations. Our study sought to uncover how DNA oligonucleotides trigger the activation of Tregs, thereby enhancing our knowledge of the factors that govern Treg function. This thorough analysis led to the identification of a subset of Tregs, specifically enriched in expanded Tregs (seTregs), which were rarely found in non-expanded Tregs. The seTreg group displayed both stemness and improved functionality, suggesting a potent Treg phenotype.

In conclusion, the discovery and characterization of seTregs opens new avenues for the development of Treg-based therapies, potentially offering more effective and targeted approaches for treating autoimmune and inflammatory diseases. The enhanced functionality and stability of seTregs, particularly through the upregulation of key TFs like IKZF2, suggest their significant therapeutic potential. Future research may focus on translating these findings into clinical applications, anticipating that seTregs could lead to breakthroughs in the treatment of conditions that are currently challenging to manage. The integration of seTregs into clinical practice holds promise for improving patient outcomes and expanding the therapeutic arsenal against autoimmune and inflammatory disorders.

MATERIALS AND METHODS

Enrichment, sorting, and expansion of Treg from human PBMCs

PBMCs from humans were acquired from CGT Global, previously known as StemExpress (CGT Global, Folsom, CA, USA), and used under the approval of the Institutional Review Board: C-2401-150-1506. Cells were enriched using human CD4 MicroBeads (Miltenyi Biotec Inc., San Jose, CA, USA) following the manufacturer's instructions. For cell sorting, we stained the enriched cells using anti-hCD4-FITC, anti-hCD25-PE-Cy7, and anti-hCD127-APC (BioLegend, San Diego, CA, USA). Cell sorting for both Treg (CD4$^+$CD25^{high}CD127^{low}) and conventional T cells (CD4$^+$CD25^{low}CD127$^+$) was conducted on the BD FACSMelody™ cell sorter (BD Biosciences, San Jose, CA, USA) (**Supplementary Fig. 1**).

Isolated cells were cultured in T cell media (RPMI [Cytiva, Marlborough, MA, USA] with 10% FBS [Sigma-Aldrich, St. Louis, MO, USA] and 300 IU/ml of human recombinant IL-2 [R&D systems, Minneapolis, MN, USA]) with the addition of TransAct (#200-076-204; Miltenyi Biotec Inc.) directly on day 0, and oligonucleotides (2 μM, 5'-AATCGTAACCGTCGTATCGGCGAT-3') for 2 wk. Rapamycin (100 nM, Sigma-Aldrich) was introduced from day 3 to day 6 for 72 h. Post 2 wk culture, cells were transitioned to oligonucleotide-free media until frozen.

For the second stimulation of cultured Tregs, autologous CD4 T cells treated with mitomycin C (10 μg/ml; Sigma-Aldrich) served as feeder cells at a 1:10 ratio, supplemented with 50 ng/ml of anti-h-CD3 Ab (clone OKT3) from Miltenyi Biotec Inc. A second stimulation was deemed unnecessary if Treg cells exhibited robust growth between days 6–8.

Nuclei isolation for single-nuclei library generation

Thawed frozen Tregs and conventional T cells were washed with chilled PBS+0.04% BSA, followed by counting with a Countess 3 FL (Invitrogen, Waltham, MA, USA). Nuclei were isolated from 0.5 to 1 million cells, spun at 300rcf for 5 min at 4°C, and resuspended in 100 μl Lysis Buffer as per guidelines provided by 10X Genomics (10X Genomics, Pleasanton, CA, USA). After a 3-min incubation on ice, cells were suspended in 1ml chilled wash buffer, washed thrice, and resuspended in chilled diluted nuclei buffer. The nuclei concentration was determined using Countess 3 FL (Invitrogen).

Single-nuclei multiome assay for transposase-accessible chromatin (ATAC) + gene expression sequencing

GEMs and libraries were constructed following the Chromium Next GEM Single-Cell Multiome ATAC + Gene Expression User Guide (10X Genomics). DNA transposition was performed by adding transposition mix to isolated nuclei. GEMs were generated using transposed nuclei, Gel Beads, Master Mix, and Partitioning Oil on a Chip J (10X Genomics), incubated in a thermal cycler, and then purified using SPRI beads for library construction. Libraries were sequenced using Illumina NovaSeq (Illumina, Inc., San Diego, CA, USA), targeting >50,000 reads per nucleus for both GEX and ATAC sequencing.

Single-nuclei gene expression and ATAC data processing and quality control

Mapping reads to the hg38 reference genome and barcode filtering were conducted using Cell Ranger (10x Genomics). Expression matrices were built using Seurat v4.9.9 with cells expressing more than 500 genes and less than 15% mitochondrial genes passing quality filters (**Supplementary Fig. 2**). The ATAC data object was created using Signac v1.12.0 and

Seurat v4.9.9. (8,9). Peaks were aligned to the hg38 reference, and Seurat objects were created based on peaks. Only cells meeting specific nucleosome signal and transcription start site enrichment criteria were analyzed further. Integration of gene expression and ATAC data within the same cells was facilitated by comparing cell barcodes.

Data normalization, batch effect correction by integration, and data visualization

RNA data were normalized using LogNormalize. This process involves dividing the expression measurement of each feature by the total expression in each cell, then multiplying by a scale factor of 10,000, and finally performing a natural logarithm transformation in Seurat v4.9.9. We implemented principle component analysis to correct for batch effects between RNA datasets before the Harmony integration (10). Using latent semantic indexing (latent semantic indexing [LSI]: term frequency-inverse document frequency followed by singular value decomposition) in Signac v1.12.0, the ATAC data has been normalized to adjust for variations in cellular sequencing depth and across peaks to give rare peaks higher values. Batch effect correction was conducted with Harmony and dimensionality reduction was calculated using iterative LSI. Twenty nearest neighbors for each cell were calculated based on the weighted combination of ATAC and RNA using FindMultiModalNeighbors in Seurat v4.9.9 with each integrated data.

Identification and analysis of differentially expressed genes (DEGs)

The FindAllMarkers from Seurat was used to find distinct DEGs inside each cluster, and non-parametric Wilcoxon rank sum tests were configured to assess the significance of each individual DEG. The downstream analysis employed the DEGs considered significant, with an adjusted p-value of less than 0.05. Then, significant genes were selected as input to perform Gene Ontology (GO) analysis through Clusterprofiler (11). A heatmap or dot plot was generated using the normalized gene expression levels for each type of cell.

Identification cluster by gene set enrichment analysis (GSEA)

Log2 fold change in gene expression comparisons between groups was used to rank all 36,601 genes for the GSEA. fGSEA v1.24.0 was used to produce an enrichment plot and identify significantly enriched gene sets based on gene ranks (12). Using msigdbr v7.5.1, appropriate gene sets were obtained from MSigDB v2023.2.Hs. Specifically, MSigDB was specifically used for obtaining reference gene sets from the H (hallmark), C2 (curated), and C7 (immunologic signature) categories. The p-value was used to estimate the significance of the enrichment scores. When a gene set's p-value was less than 0.05, it was deemed substantially enriched.

chromVAR motif analysis

TF accessibility deviations were calculated using chromVAR v1.20.2, comparing the actual accessibility to the average of all the cells or samples (13). The analysis utilized the human genome reference hg38 and TF binding motif references from the JASPAR2020 Core Vertebrates database as its references (14).

Inference of super-enhancer regions and computation of super-enhancer associated gene module scores

We predict enhancer regions using H3K27ac chromatin immunoprecipitation sequencing (ChIP-seq) data (GSE43119) and rank enhancers based on the peak intensity of H3K27ac. Furthermore, super enhancers were defined as regions up to the point of intersection with a

line having a slope of 1, starting from the highest-ranked enhancer. Subsequently, the nearest genes to the determined super enhancers were identified using 'ClosestFeature' function in the Signac package. And module scoring of gene expression levels was performed based on gene modules, allowing for comparison across clusters.

Inference and analysis of TF-regulated genes across clusters

We retrieved the regions where each TF binds from GSE43119. To infer genes regulated by TFs, we used the 'ClosestFeature' function in the signac package to identify the nearest genes to those regions. Using those genes, GO analysis was conducted to identify pathways inferred to be regulated by TFs in the expression. Furthermore, module scoring was performed based on the expression of these genes to compare across each cluster.

Inference of cis-regulatory interactions and definition of a TF-gene regulatory network

FigR (15) was run as described at https://buenrostrolab.github.io/FigR/. Using the 'runGenePeakcorr' function from the FigR package to assess peak-gene association linkage and classify significant correlations with p-value<0.05. The 'dorcJPlot' function is utilized to identify Domain of Regulatory Chromatin (DORC) genes with peak-gene associations, with a cutoff of 5 applied. Subsequently, we assess TF-gene associations by determining if motif sequences are present in peaks associated with genes based on TF binding sequence motifs, and infer regulatory networks.

RESULTS

Single-nucleus RNA sequencing (snRNA-seq) and single-nucleus ATAC sequencing (snATAC-seq) uncover distinct subsets within sorted and expanded human Tregs

To assess transcription and chromatin accessibility information at a single cell resolution, snRNA-seq and snATAC-seq were performed using sorted CD4$^+$CD25^{high}CD127^{low} cells and *ex vivo* expanded Tregs with rapamycin, IL-2, and oligonucleotides on the 10X Genomics Chromium platform (**Fig. 1A**). The purity and heterogeneity of non-expanded Tregs and expanded Tregs was evaluated (**Supplementary Fig. 3**). Cell clusters were determined using the weighted nearest neighbor clustering algorithm in Seurat (8), which is designed for integrating RNA and ATAC data. These were visualized through Uniform Manifold Approximation and Projection (UMAP) analysis. Ultimately, nine clusters within both non-expanded Tregs and expanded Tregs were identified (**Fig. 1B**).

We validated the identity of these clusters based on the expression of canonical marker genes (**Fig. 1C**). Naive Treg cells were characterized by high levels of CCR7 and TCF7 expression. Employing the top 10 signature genes and signature TFs for each cluster, those with reduced expression of CCR7 and TCF7 were categorized as FOXP3^{high} Tregs, SGMS1 Tregs, Helios SGMS1 Tregs, Helios Tregs, TGFB1 Tregs, and JUN Tregs. To further confirm and characterize each cluster, GSEA was conducted, uncovering unique attributes such as effector characteristics in FOXP3^{high} Tregs and the augmented mTOR signaling pathway in SGMS1 Tregs. TNF downstream pathway was elevated in Helios Tregs, TGFB1 Tregs showed significant activation of the TGF-β pathway compared to other clusters, and the AP1 pathway was predominant in JUN Tregs, FOXP3^{high} Tregs (5) and Helios Treg (16) (**Fig. 1C**, **Supplementary Fig. 4**, **Supplementary Table 1**).
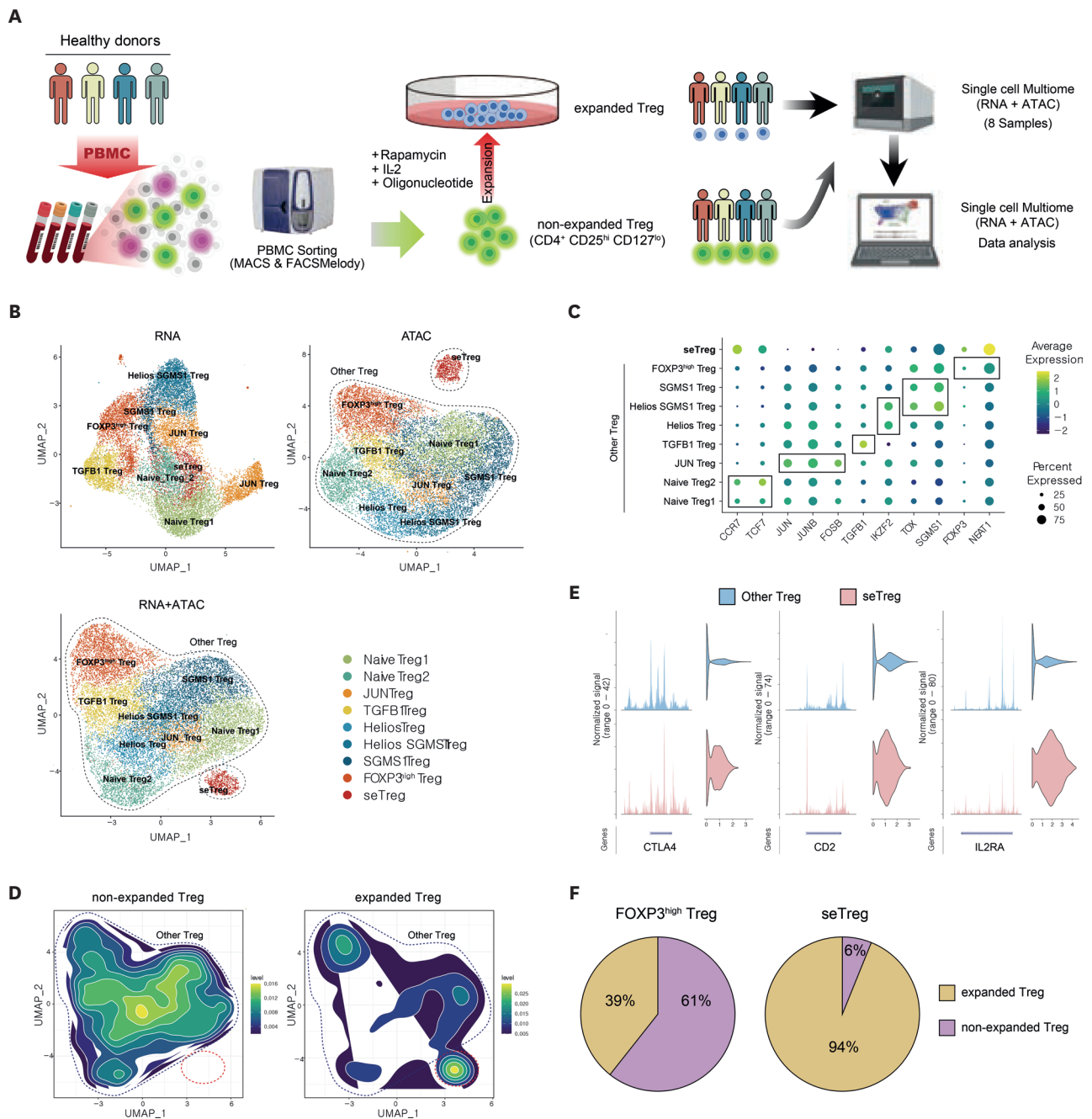
Figure 1. snRNA-seq and snATAC-seq reveals distinctive subsets among human sorted Treg and expanded Treg. (A) Schematic representation of profiling methods and cell types. (B) UMAP visualization of single cells derived from combining non-expanded Tregs and expanded Tregs population, showing UMAP coordinates defined by RNA, ATAC, and RNA+ATAC. The points, distinguished by cluster-specific colors, are determined through clustering based on combined RNA and ATAC data. Cell types are then assigned to these clusters based on the expression of marker genes. (C) Dot plots showing representative genes that were differently expressed in each cluster. Dot size shows the percentage of cells expressing each gene, and dot color indicates the average expression level. FC denotes fold change. (D) Contour plot illustrating the point density on the UMAP coordinates derived from combining RNA and ATAC data separately for non-expanded Tregs and expanded Tregs. The color gradient indicates the degree of density, with lighter shades representing higher density. (E) Comparison of gene expression differences in CTLA4, CD2, IL2RA among other Tregs and seTregs using violin plots, along with coverage plots depicting chromatin accessibility at transcription start site locations. (F) Pie charts to illustrate the distribution of non-expanded Treg and expanded Tregs proportions within FOXP3high Tregs and seTregs.

Furthermore, we identified a distinct cluster specific to *ex vivo* expanded Tregs, referred to as seTregs (**Fig. 1D**, **Supplementary Fig. 5**). To investigate the distribution of Treg clusters between sorted and expanded Tregs, we analyzed their density distribution by UMAP embedding space. Within the expanded Treg population, we observed a significant increase of seTregs and FOXP3high Tregs. Reasonably, naive Tregs were present in minimal numbers. Other Treg clusters, with the exception of seTregs, FOXP3high Tregs, and naive Tregs, were notably scarce or virtually undetectable. This pattern suggests a skewed differentiation of specific Treg subsets, potentially indicative of the influence of the expansion conditions on Treg differentiation and subset representation (**Fig. 1D**).

In addition, we grouped all non-seTregs together ("Other Tregs") to compare gene expression and chromatin accessibility to seTregs. The gene expression and chromatin accessibility peaks associated with Treg activation, including CTLA4, CD2, and IL2RA, were significantly increased among seTregs compared to other Tregs. CTLA4 is involved in Treg suppression through its higher affinity to CD28, while IL2RA contributes via IL2 consumption. CD2 signaling induces the expression of FOXP3, a master regulator of Treg function. These indicate the potential for seTregs to possess high functionality (**Fig. 1E**) ([17]-[19]). By analyzing the ratio of non-expanded Tregs to expanded Tregs in FOXP3high Tregs and seTreg clusters, we discovered that 94% of the seTreg cluster resided within the expanded Treg population (**Fig. 1F**). Therefore, seTregs emerge as a distinct cluster with potentially higher functional properties resulting from *ex vivo* expansion of Tregs by the addition of the oligonucleotides.

The seTregs exhibit superior functionality and stemness features relative to other Tregs

To further elucidate the characteristics of the identified seTregs, we compared the top 10 DEGs and TFs between clusters (**Fig. 2A and B**). The seTregs showed elevated expression of the DEGs associated to naive Tregs, such as BACH2 and LEF1 and also to FOXP3high Tregs including CTLA4, MAF and TOX2. To explore relationships between Treg subsets, we generated Spearman correlation graphs using all genes and accessible chromatin regions. This analysis indicated that seTregs share a closer correlation with naive Tregs and FOXP3high Tregs compared to other Treg subsets (**Fig. 2C**). Further, when comparing gene expression among all Treg subsets using naive Treg1 and FOXP3high Tregs signatures, the seTreg subset was uniquely characterized by elevated signatures of both naive Treg1 and FOXP3high Tregs, distinguishing them from other Treg subsets (**Fig. 2D**). GSEA further confirmed these findings showing seTregs also revealed a high normalized enrichment score (NES) for both *in vitro* induced Treg signature and the Treg signature in breast cancer tumor microenvironment (**Fig. 2E**, **Supplementary Table 1**), with significant NES and p-value noted for signatures associated with naive CD4 T cells (**Fig. 2F**, **Supplementary Table 1**). Most Tregs found within tumor tissue (TIL-Tregs) exhibited an activated and effector-memory phenotype ([20]). Focusing on these seTregs correlated subsets, namely naive Treg1, 2, and FOXP3high Tregs, we noted that canonical genes typical of naive Tregs including, TCF7 and SATB1 were predominantly expressed in naive Treg1, 2, and seTregs, while the genes indicative of activated Tregs including CCR4 and NEAT1 were more pronounced in seTregs than in FOXP3high Tregs. The effector Treg-associated genes such as GBP2, ITGA4, TPR, and LIMS1 were highly expressed in seTregs, while the tumor-infiltrating Treg-associated genes such as CTLA4, ICOS, and IL12RB2 were also most highly expressed in FOXP3high Tregs and seTregs (**Fig. 2G-J**) ([5]). These findings suggest that seTregs combine the stemness of naive Tregs with the functional attributes of activated Tregs.

Figure 2. The seTreg shows a correlation with both naive Treg and FOXP3^high Treg. (A) Heatmap displaying the top 10 marker genes (by fold change) for each subset, excluding ribosomal and mitochondrial genes. (B) Heatmap depicting the top 10 transcription factors for each cluster across the entire dataset. (C) Correlograms visualizing the correlation of single-cell gene expression profiles between clusters (D) Scatter plot using the average expression levels of signature genes for FOXP3^high Treg and naive Treg 1 for each cluster. (E, F) GSEA plots obtained by using the indicated gene sets. p-values from GSEA are indicated. (G-J) Dot plots illustrating the expression patterns of genes associated with naive Treg, effector Treg, activated Treg, and tumor infiltrating Treg in each cluster, including seTreg, FOXP3 Treg, naive Treg 1, and naive Treg.

The seTregs exhibit high accessibility to the super-enhancer regions of activated Tregs and increased activity of Treg-associated TFs

As seTregs have been shown to share features with the activated signatures of Tregs, notably FOXP3[high] Tregs, we decided to further explore whether seTregs display epigenetic characteristics of activated Tregs. To do this, we identified 558 super-enhancer regions using public H3K27ac ChIP-Seq data from activated Tregs (**Fig. 3A**) (21). To evaluate the accessibility of these super-enhancer regions in seTregs, we performed addmodulescore analysis using for super-enhancer-associated genes. Compared across the nine Treg subtypes, seTregs showed the highest score which indicates greater accessibility in the super-enhancer regions, resulting in increased gene expression compared to other Treg subtypes (**Fig. 3B**).
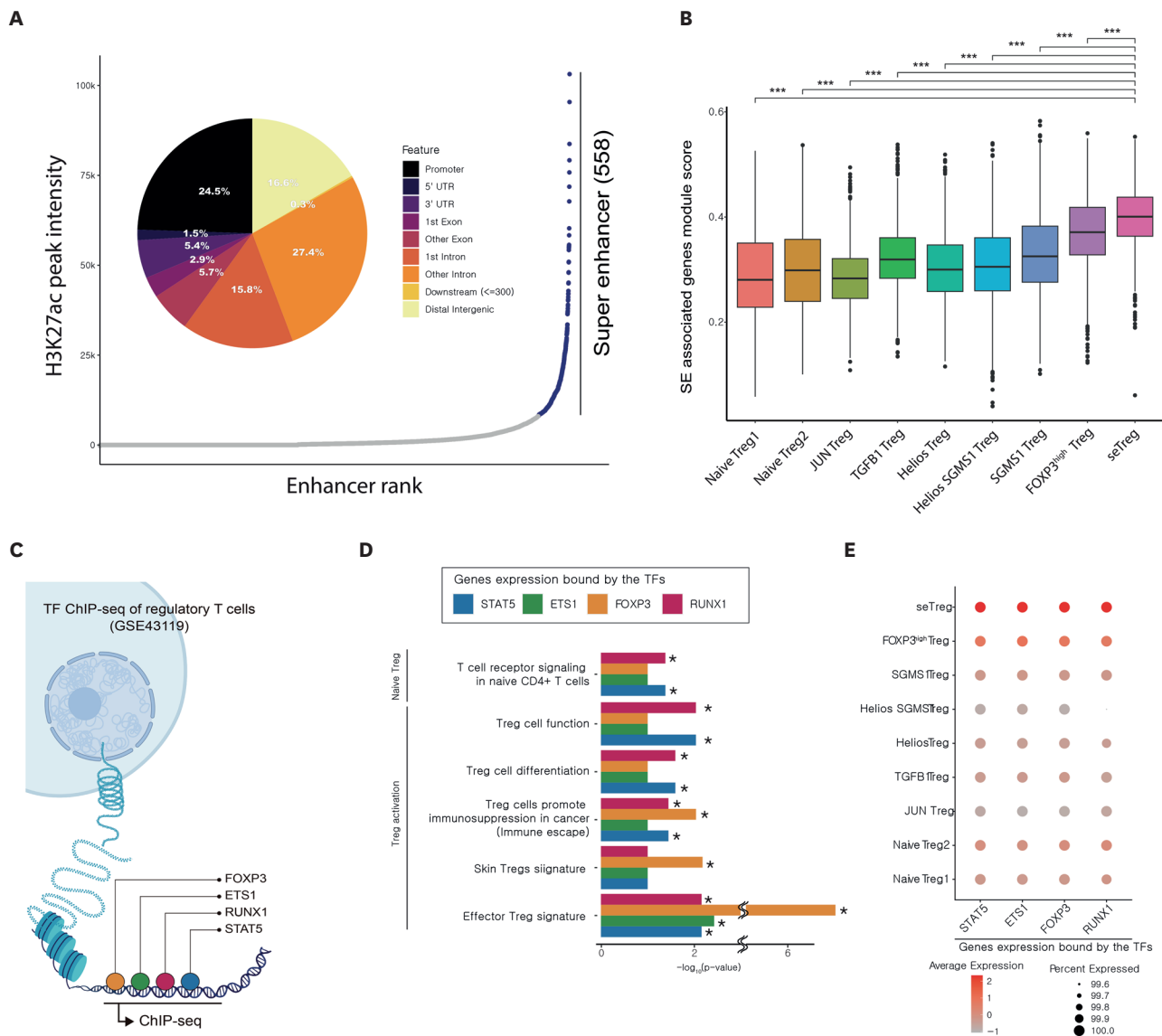

Figure 3. seTregs are more activated compared to other Tregs. (A) Super enhancers in activated Tregs were identified as described in the method section. (B) Box plot illustrating the difference in super-enhancer scores among Treg subsets within activated Tregs. (C) Scheme for profiling the ChIP-seq of Treg-associated transcription factors, namely STAT5B, RUNX1, ETS1, and FOXP3. (D) Enriched GO biological process terms for genes regulated by STAT5B, RUNX1, ETS1, and FOXP3. (E) Dot plot representing the scoring of gene expression levels regulated by binding of STAT5B, RUNX1, ETS1, and FOXP3.

Additionally, we aimed to identify the binding sites of TFs known to be essential for Treg cell function, including FOXP3, STAT5, erythroblast transformation-specific proto-oncogene 1 (ETS1), and runt-related transcription factor 1 (RUNX1), and to assess their accessibility in our dataset (22-25). Using TF ChIP-seq data from Schmidl et al.'s (23) study, we identified the DNA regions where each TF binds among Treg subsets (**Fig. 3C**). Next, we conducted GO analysis on genes predicted to be expressed upon each TF's binding and examined the gene sets regulated by each TF. The STAT5, ETS1, FOXP3, and RUNX1 were shown to regulate genes involved in Treg activation, such as those promoting suppression in cancer and contributing to the effector Treg signature. STAT5 and RUNX1 also displayed roles in TCR signaling in naive CD4 T cells (**Fig. 3D**).

We determined that gene expression driven by the four TFs, deemed crucial for Treg function, was highest in seTregs (**Fig. 3E**). Thus, seTregs demonstrate the epigenetic traits of highly activated Tregs. Moreover, due to the high binding activity of each TF, it is anticipated that seTregs will possess enhanced functionality, driven by the active participation of these TFs.

The seTregs exhibit simultaneous stemness and functionality through IKZF2 regulation

As other regulators may also play a crucial role in conferring characteristics of activation and stemness in seTregs, we identified DORC, which are cis-regulatory elements with a large number of linked peaks associated with genes (**Fig. 4A**). These DORC elements could be essential for determining cell identities (15,26). Through this approach, we ranked all genes on Tregs by the number of correlated peaks and identified 26 genes as DORCs, each characterized by five or more cis-regulatory elements showing a correlation with gene expression levels (**Fig. 4B**) (26). When comparing the gene expression levels of DORCs between FOXP3[high] Tregs and seTregs, *IKZF2* was the most significantly expressed gene in seTregs (**Fig. 4C**). Consequently, we concentrated on IKZF2 and undertook a DORC network analysis, which included downstream transcription factors regulated by IKZF2, specifically MAF, BHLHE40, ZNF287, and THRA. This analysis identified numerous genes presumed to be regulated by IKZF2, and we clarified the network among these genes (**Fig. 4D**). GO analysis on the genes within this network revealed pathways related to suppressive function, such as IL-10 signaling and the induced Treg pathway, alongside the stemness pathway of naive Tregs (**Fig. 4E**). Thus, we propose IKZF2 as a key factor in augmenting both the stem-like phenotype and potentially enhanced functionality of seTregs.

DISCUSSION

In this study, the use of oligonucleotides during the culture of isolated Treg cells led to the identification of a distinct cluster within the expanded Treg population, characterized by stem-like features and potentially enhanced functionality compared to other subsets. This discovery suggests a potential method for generating robust Tregs that retain their suppressive function, offering promising implications for Treg application in clinical trials.

The categorization of *ex vivo* expanded Tregs along with non-expanded natural Tregs into diverse subtypes, has enriched our understanding of the Treg population's heterogeneity. Interrogating these subsets by integrating gene expression with chromatin accessibility has allowed for a more detailed classification of Treg cellular subsets, highlighting the complexity within Treg populations.
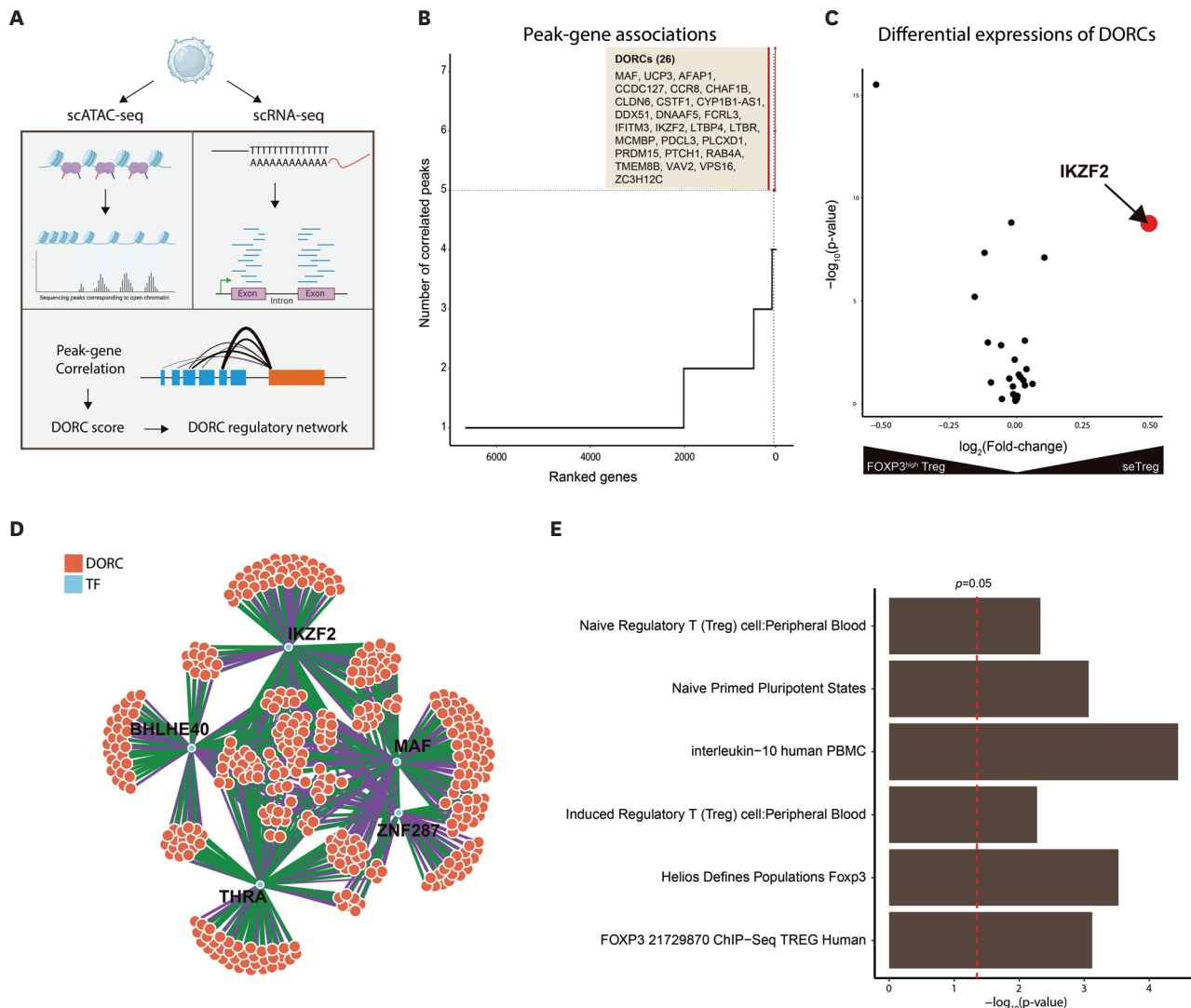
Figure 4. DORC of seTreg has been examined. (A) Scheme for analysis of cis-regulatory elements and expression of genes. (B) The quantity of each gene's significantly correlated peaks (p<0.01). (C) Volcano plot illustrating the comparison of DORC gene expression levels between FOXP3^{high} Treg and seTreg. (D) A network view depicting the DORC associated with each transcription factor. (E) Enriched GO terms derived from DORC genes associated with IKZF2 shown in (D).

A comparative analysis between expanded and non-expanded Tregs identified a unique subset (seTreg) highly enriched in the expanded Treg group, displaying traits of both naive and activated phenotypes. Krishna et al. (27) also showed a similar phenomenon in CD8 T cells, where a subset displaying both naive and activated characteristics was identified. This subset interestingly also displayed a resistance to exhaustion (27). Notably, this subset displayed epigenetic phenotypes that underscore the importance of these Treg subsets in achieving enhanced functionality without the exhaustion typically seen in activated CD8 T cells.

Furthermore, our integrated analysis, including H3K27ac ChIP-seq data from prior Treg studies, identified that epigenetic characteristics of seTregs are similar to those in activated Tregs. Remarkably, seTregs exhibited elevated expression and transcriptional activity of essential TFs involved in Treg function such as FOXP3, STAT5B, RUNX1, and ETS1, compared to FOXP3^{high} Tregs. It might mean that the seTregs are more activated than activated Tregs in peripheral circulation. Therefore, we found that Tregs expanded with oligonucleotides might demonstrate enhanced functionality over natural Tregs using the inferred TF activity analysis.

Through gene regulatory network reconstruction, we identified IKZF2 as a crucial TF of the seTreg subset. This finding is consistent with our earlier research showing an increase in Helios, a protein implicated in stemness and functionality, following expansion with oligonucleotides (7). Multiple prior studies have been conducted in both murine models (28,29) and human cells (30,31), which have elucidated the role of Helios in regulating the stability of FOXP3 expression and the suppressive function of Tregs. Significantly, it was observed that reducing Helios levels in human Tregs led to a decline in Foxp3 expression and compromised their ability to exert immunosuppressive effects (30). Our findings further support this notion, as we observe that Helios governs the functionality of seTregs by orchestrating downstream transcription factors and genes.

We sought to explore the receptor for the oligomer. Kim et al. (7) provided evidence suggesting that the oligonucleotide is internalized by Tregs in granule-like organelles within the cytoplasm via receptor-mediated endocytosis.

In conclusion, our findings reveal the potential for generating highly functional Tregs with pronounced stemness through expansion with oligonucleotides. The prospective use of these functionally activated Tregs in clinical therapy for autoimmune conditions and chronic inflammation is apparent. Additionally, our insights into IKZF2 modulation present a viable strategy for future clinical trials to regulate Treg functionality, offering deeper insights into the intrinsic mechanisms governing Tregs.

ACKNOWLEDGEMENTS

SUPPLEMENTARY MATERIALS

Supplementary Table 1
The results of GSEA analysis for each cluster

Supplementary Figure 1
Flow cytometry gating strategy for sorting Tregs from human PBMCs.

Supplementary Figure 2
Quality control in multimodal single-cell assay.

Supplementary Figure 3
Flow cytometry graph demonstrating the purity and heterogeneity of non-expanded Tregs and expanded Tregs.

Supplementary Figure 4
GSEA plots obtained by using the indicated gene sets for each Treg subsets.

Supplementary Figure 5

A stacked bar graph showing the proportions of each Treg subset for expanded Treg and sorted Treg.

REFERENCES

1. Sakaguchi S, Sakaguchi N, Asano M, Itoh M, Toda M. Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *J Immunol* 1995;155:1151-1164. **PUBMED** | **CROSSREF**

2. Fontenot JD, Gavin MA, Rudensky AY. Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. *Nat Immunol* 2003;4:330-336. **PUBMED** | **CROSSREF**

3. Hori S, Nomura T, Sakaguchi S. Control of regulatory T cell development by the transcription factor Foxp3. *Science* 2003;299:1057-1061. **PUBMED** | **CROSSREF**

4. Khattri R, Cox T, Yasayko SA, Ramsdell F. An essential role for Scurfin in CD4+CD25+ T regulatory cells. *Nat Immunol* 2003;4:337-342. **PUBMED** | **CROSSREF**

5. Luo Y, Xu C, Wang B, Niu Q, Su X, Bai Y, Zhu S, Zhao C, Sun Y, Wang J, et al. Single-cell transcriptomic analysis reveals disparate effector differentiation pathways in human T$_{reg}$ compartment. *Nat Commun* 2021;12:3913. **PUBMED** | **CROSSREF**

6. Ferreira LM, Muller YD, Bluestone JA, Tang Q. Next-generation regulatory T cell therapy. *Nat Rev Drug Discov* 2019;18:749-769. **PUBMED** | **CROSSREF**

7. Kim YC, Bhairavabhotla R, Yoon J, Golding A, Thornton AM, Tran DQ, Shevach EM. Oligodeoxynucleotides stabilize Helios-expressing Foxp3+ human T regulatory cells during in vitro expansion. *Blood* 2012;119:2810-2818. **PUBMED** | **CROSSREF**

8. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. Integrated analysis of multimodal single-cell data. *Cell* 2021;184:3573-3587.e29. **PUBMED** | **CROSSREF**

9. Stuart T, Srivastava A, Madad S, Lareau CA, Satija R. Single-cell chromatin state analysis with Signac. *Nat Methods* 2021;18:1333-1341. **PUBMED** | **CROSSREF**

10. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh PR, Raychaudhuri S. Fast, sensitive and accurate integration of single-cell data with harmony. *Nat Methods* 2019;16:1289-1296. **PUBMED** | **CROSSREF**

11. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012;16:284-287. **PUBMED** | **CROSSREF**

12. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. Preprint at https://www.biorxiv.org/content/10.1101/060012v3 (2016).

13. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods* 2017;14:975-978. **PUBMED** | **CROSSREF**

14. Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, Modi BP, Correard S, Gheorghe M, Baranašić D, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2020;48:D87-D92. **PUBMED**

15. Kartha VK, Duarte FM, Hu Y, Ma S, Chew JG, Lareau CA, et al. Functional inference of gene regulation using single-cell multi-omics. *Cell Genom* 2022;2:100166. **PUBMED** | **CROSSREF**

16. Thornton AM, Lu J, Korty PE, Kim YC, Martens C, Sun PD, Shevach EM. Helios⁺ and Helios⁻ Treg subpopulations are phenotypically and functionally distinct and express dissimilar TCR repertoires. *Eur J Immunol* 2019;49:398-412. **PUBMED** | **CROSSREF**

17. Chinen T, Kannan AK, Levine AG, Fan X, Klein U, Zheng Y, Gasteiger G, Feng Y, Fontenot JD, Rudensky AY. An essential role for the IL-2 receptor in T$_{reg}$ cell function. *Nat Immunol* 2016;17:1322-1333. **PUBMED** | **CROSSREF**

18. Kashiwakura Y, Sakurai D, Kanno Y, Hashiguchi M, Kobayashi A, Kurosu A, Tokudome S, Kobata T, Kojima H. CD2-mediated regulation of peripheral CD4(+) CD25(+) regulatory T-cell apoptosis accompanied by down-regulation of Bim. *Immunology* 2013;139:48-60. **PUBMED** | **CROSSREF**

19. Takahashi T, Tagami T, Yamazaki S, Uede T, Shimizu J, Sakaguchi N, Mak TW, Sakaguchi S. Immunologic self-tolerance maintained by CD25(+)CD4(+) regulatory T cells constitutively expressing cytotoxic T lymphocyte-associated antigen 4. *J Exp Med* 2000;192:303-310. **PUBMED** | **CROSSREF**

20. Devi-Marulkar P, Fastenackels S, Karapentiantz P, Goc J, Germain C, Kaplon H, Knockaert S, Olive D, Panouillot M, Validire P, et al. Regulatory T cells infiltrate the tumor-induced tertiary lymphoïd structures and are associated with poor clinical outcome in NSCLC. *Commun Biol* 2022;5:1416. **PUBMED** | **CROSSREF**

21. Mijnheer G, Lutter L, Mokry M, van der Wal M, Scholman R, Fleskens V, Pandit A, Tao W, Wekking M, Vervoort S, et al. Conserved human effector Treg cell transcriptomic and epigenetic signature in arthritic joint inflammation. *Nat Commun* 2021;12:2710. **PUBMED** | **CROSSREF**

22. Ohkura N, Yasumizu Y, Kitagawa Y, Tanaka A, Nakamura Y, Motooka D, Nakamura S, Okada Y, Sakaguchi S. Regulatory T cell-specific epigenomic region variants are a key determinant of susceptibility to common autoimmune diseases. *Immunity* 2020;52:1119-1132.e4. **PUBMED** | **CROSSREF**

23. Schmidl C, Hansmann L, Lassmann T, Balwierz PJ, Kawaji H, Itoh M, Kawai J, Nagao-Sato S, Suzuki H, Andreesen R, et al. The enhancer and promoter landscape of human regulatory and conventional T-cell subpopulations. *Blood* 2014;123:e68-e78. **PUBMED** | **CROSSREF**

24. Ono M, Yaguchi H, Ohkura N, Kitabayashi I, Nagamura Y, Nomura T, Miyachi Y, Tsukada T, Sakaguchi S. Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature* 2007;446:685-689. **PUBMED** | **CROSSREF**

25. Polansky JK, Schreiber L, Thelemann C, Ludwig L, Krüger M, Baumgrass R, Cording S, Floess S, Hamann A, Huehn J. Methylation matters: binding of Ets-1 to the demethylated Foxp3 gene contributes to the stabilization of Foxp3 expression in regulatory T cells. *J Mol Med (Berl)* 2010;88:1029-1040. **PUBMED** | **CROSSREF**

26. Ma S, Zhang B, LaFave LM, Earl AS, Chiang Z, Hu Y, Ding J, Brack A, Kartha VK, Tay T, et al. Chromatin potential identified by shared single-cell profiling of RNA and chromatin. *Cell* 2020;183:1103-1116.e20. **PUBMED** | **CROSSREF**

27. Krishna S, Lowery FJ, Copeland AR, Bahadiroglu E, Mukherjee R, Jia L, Anibal JT, Sachs A, Adebola SO, Gurusamy D, et al. Stem-like CD8 T cells mediate response of adoptive cell immunotherapy against human cancer. *Science* 2020;370:1328-1334. **PUBMED** | **CROSSREF**

28. Kim HJ, Barnitz RA, Kreslavsky T, Brown FD, Moffett H, Lemieux ME, Kaygusuz Y, Meissner T, Holderried TA, Chan S, et al. Stable inhibitory activity of regulatory T cells requires the transcription factor Helios. *Science* 2015;350:334-339. **PUBMED** | **CROSSREF**

29. Sugita K, Hanakawa S, Honda T, Kondoh G, Miyachi Y, Kabashima K, Nomura T. Generation of Helios reporter mice and an evaluation of the suppressive capacity of Helios(+) regulatory T cells in vitro. *Exp Dermatol* 2015;24:554-556. **PUBMED** | **CROSSREF**

30. Getnet D, Grosso JF, Goldberg MV, Harris TJ, Yen HR, Bruno TC, Durham NM, Hipkiss EL, Pyle KJ, Wada S. A role for the transcription factor Helios in human CD4(+)CD25(+) regulatory T cells. *Mol Immunol* 2010;47:1595-1600. **PUBMED** | **CROSSREF**

31. Elkord E, Abd Al Samid M, Chaudhary B. Helios, and not FoxP3, is the marker of activated Tregs expressing GARP/LAP. *Oncotarget* 2015;6:20026-20036. **PUBMED** | **CROSSREF**