ARTICLE OPEN

Single-cell identifies and validates human circulating Treg subtype/state Treg^{fci} in non-small cell lung cancer

Xuanqi Liu[1,2], Yifei Liu[3], Wanxin Duan[2], Jianming Weng[4], Hao Dai[5], Fangming Liu[2], Lingyan Wang[2], Mengjia Qian[2], Bijun Zhu[2], Shuyao Wu[6], Youqiong Ye[6], Muhan Yuan[2], Shaojun Hong[4], Mingjie Wang[7], Yunfeng Cheng[2]✉, Tianxiang Chen[8,9]✉, Duojiao Wu[2]✉, Yuanlin Song[1,10]✉ and Xiangdong Wang[1,2]✉

Regulatory T cells (Treg) play immunosuppressive roles in lung cancer and are key biomarkers in cancer immunology. Using scRNA-seq and stereo-seq, we aimed to identify a novel circulating Treg subtype in non-small cell lung cancer (NSCLC). We developed a regional overlap-expression rate (rOER) system with cell identity marker gene panels (ciMGPs) to quantify T cell subtype specificity across pre/post-operative blood, 31 diseases, and in 17 organs. A Treg subtype (Treg^{fci}, FOXP3^{+}CTLA4^{+}IL2RA^{+}) was identified and validated for disease-, organ-, and time-specific characteristics across spatializations, temporalizations, diseases, and prognoses. Treg^{fci} was prominently enriched in pre-operative blood and NSCLC tissue but markedly decreased or was absent post-operatively. Spatial transcriptome and multiplex immunostaining revealed that Treg^{fci} localized primarily to the interface between NSCLC and normal tissues, as well as in specific micro-niches. ETS Proto-Oncogene 1 (*ETS1*) was upregulated in Treg^{fci}, confirmed by knockdown and knockout experiments. *ETS1* facilitated Treg^{fci} migration from circulation into tumor tissues via cancer cell-derived chemoattraction. Interactions between *ETS1* and Treg^{fci} ciMGPs, and between intracellular organelles, were found to modulate metabolism and mitochondrial function. These findings highlight the importance of Treg^{fci} in understanding the molecular mechanism by which Tregs maintain the balance between systemic and local immune ecosystems. ETS1's regulation of Treg^{fci} transcriptomic and metabolic profiles were confirmed in human PBMCs treated with an *ETS1* inhibitor and in lung tissues from mice$^{ETS1\text{-}cKO}$ mice. Furthermore, Treg^{fci} and *ETS1* represent a distinct molecular signature of T cells and potential therapeutic targets for clinical intervention in lung cancer.

Signal Transduction and Targeted Therapy (2026)11:211 ; https://doi.org/10.1038/s41392-026-02677-6

INTRODUCTION

Circulating T cells are critical for maintaining systemic immune homeostasis, orchestrating inter-organ immune coordination, and modulating local microenvironments.[1] The discovery of T cell subtypes/states (Ts/s) and identities have expanded rapidly with the advent of single-cell RNA sequencing (scRNA-seq) and the development of cell identity marker gene panels (ciMGPs). Spatial transcriptomics (Stereo-seq) demonstrates that the distribution and dynamics of Ts/s within tumor microenvironments (TME) are correlated with cancer malignancy, tumor invasion, formation of anti-tumor immune niches, and therapeutic responsiveness.[2–5] Circulating T cells play a critical role in stabilizing systemic immune function, mimicking spatialized microenvironmental defense, and reflecting inter-organ communication.[2,3,6] scRNA-seq can refine the abundance, diversity, and characteristics of activated or exacerbated T cells and receptors, as biomarkers for the severity and progression of immune-related adverse events in melanoma patients treated with immune checkpoint blockade agents.[7,8] The delineation of Ts/s defined by scRNA-seq provides novel insights into the T cell atlas, disease-specific T cell patterns, and interactions between Ts/s, which are expected to be translated into clinical applications as part of the clinical biochemistry of hematology.[9–11] However, the clinical translation of scRNA-seq-defined Ts/s remains challenging due to cost burden, technical complexity, regulations, equipment certifications, standardization, repeatability, data interpretation, bioinformatic security, variability in physician expertise, and legal frameworks. Among these, a major hurdle lies in the precise selection of cell Ts/s, which is complicated by the multifaceted nature of T cell development, localization, regulation, functional plasticity, and metabolic heterogeneity.[12]

The accuracy and specificity of Ts/s are highly dependent on the reproducibility and overlap of the selected marker gene number and function, the dynamic transitions from naïve to active Ts/s,

[1]Department of Pulmonary and Critical Care Medicine, Zhongshan Hospital, Fudan University Shanghai Medical College, Shanghai, China; [2]Shanghai Institute of Clinical Bioinformatics, Zhongshan Hospital Institute of Clinical Biomedicine, Shanghai, China; [3]Clinical Center for Molecular Diagnosis and Therapy, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, China; [4]Department of Pathology, Zhangzhou Hospital of Fujian Medical University, Zhangzhou, China; [5]State Key Laboratory of Cell Biology, Shanghai Key Laboratory of Molecular Andrology, Shanghai Institute of Biochemistry and Cell Biology, CAS Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai, China; [6]Shanghai Institute of Immunology, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [7]Department of Gastroenterology, Ruijin Hospital, Shanghai Jiao Tong University, School of Medicine, Shanghai, China; [8]Department of Thoracic Surgery, Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [9]Shanghai Key Laboratory of Thoracic Tumor Biotherapy, Shanghai, China and [10]Shanghai Key Laboratory of Lung Inflammation and Injury, Shanghai Respiratory Research Institute, Department of Pulmonary Medicine, Zhongshan Hospital, Fudan University, Shanghai, China
Correspondence: Yunfeng Cheng (cheng.yunfeng@zs-hospital.sh.cn) or Tianxiang Chen (c.tianxiang@gmail.com) or Duojiao Wu (wu.duojiao@zs-hospital.sh.cn) or Yuanlin Song (song.yuanlin@zs-hospital.sh.cn) or Xiangdong Wang (xdwang@fuccb.com)
These authors contributed equally: Xuanqi Liu, Yifei Liu, Wanxin Duan, Jianming Weng, Hao Dai.

Received: 12 May 2025 Revised: 4 February 2026 Accepted: 8 March 2026
Published online: 03 June 2026

and disease-specific phenomes. The evidence-based evaluation systems should represent the disease specificity and clinical utility of ciMGPs for diagnostic purposes. In order to facilitate clinical translation, a robust framework for cell identity assessment is urgently required to quantify and classify the biological and clinical specificities of ciMGPs. Analysis of confidence intervals-expression quantitative trait loci demonstrated that the definition of cell types and identities in peripheral blood mononuclear cells (PBMCs) is essential for elucidating the immune functions, cellular heterogeneity, and disease pathogenesis.[13] Nevertheless, it remains unclear precisely how circulating Ts/s profiles correspond to clinical data in patients with lung cancer. Taking into consideration that Ts/s in peripheral blood are accessible in the clinical setting, there is a pressing need for precise and efficient strategies to delineate characteristics specific to disease nature, severity, stage and response of Ts/s, and to define the gene expression overlaps and changes of ciMGPs for each Ts/s. Expression profiles of Ts/s-specific ciMGPs reflect systemic immune cell development, compositional patterns, transcriptional programs, and genetic variants associated with immune dysregulation or cancer progression.[13–15] The prior evidence suggests that the specificity and diversity of Ts/s reflect the spatialization of tumor microenvironmental inflammation, balance between T cell activation and exhaustion, and patient prognosis.[2,3,16,17] With the rapid advancement of single-cell measurements, numerous immune cell subtypes were identified for deeper insights into the molecular mechanisms and disease pathogenesis.[18]

The clinical specificity and utility of T cell ciMGPs require multidimensional validation to delineate the physiological range and pathological alterations of circulating Ts/s. The primary aim of the present study was to establish a new evaluation system for the validation of novel subtypes of T regulatory cells (Treg) in the circulation of patients with non-small cell lung cancer (NSCLC). This was achieved by systematically assessing the diversity and specificity of Ts/s in healthy individuals, the correspondence between circulating and tissue-resident T cell profiles in lung cancer, dynamic changes prior to and following surgical intervention, spatial distribution across various organs, and disease-specific signatures across various lung diseases and extrapulmonary malignancies. A regional overlap-expression rate (rOER) metric, a screening and evaluation system, was developed to quantify Ts/s ciMGPs expression specificity. Ts/s ciMGPs were categorized through bibliometric and nomenclature-based screening from our own and public databases. NSCLC-specific Ts/s ciMGPs were selected and validated based on the expression dynamics, disease specificity, altered patterns, and prognostic values by profiling single-cell transcriptomes of peripheral T cells from paired pre- and post-operative blood samples, and matched NSCLC and para-NSCLC tissues. The specificity of these ciMGPs was confirmed by comparing the expression patterns with Ts/s profiles of pulmonary diseases to validate the NSCLC specificity, with Ts/s profiles of other extrapulmonary diseases to confirm the disease specificity, and with other organs/tissues to define the tissue specificity.

One of the Treg subtypes, Treg^{fci} with *CD4*, *CD45*, *FOXP3*, *CTLA4*, and *IL2RA* (5g-Treg^{fci}) and Treg^{fci} with *FOXP3*, *CTLA4*, and *IL2RA* (3g-Treg^{fci}) were compared, and the latter was used as the representative panel description, as clarified in Section 10 of the Supplementary Materials. This panel was used in the current study to enhance classification accuracy and robustness, particularly to mitigate gene dropout effects in scRNA-seq and Stereo-seq data. The disease-, organ-, and time-specific characteristics of Treg^{fci} were validated across various spatializations, temporalizations, diseases, and prognoses. Moreover, the present study characterized the biological functions of Treg^{fci} under healthy and disease conditions, with or without stimulations, and in co-culture with NSCLC cells or human normal airway epithelial cells, following individual or combined downregulation of Treg^{fci} ciMGPs. During the evaluation, *ETS1* (ETS Proto-Oncogene 1, Transcription Factor) emerged as one of the key Treg^{fci}-specific regulators and exhibited significant alterations in transcription-factor networks. Functional perturbations of *ETS1* introduced through knockin, knockdown, or knockout approaches in cells and animal models demonstrated its critical role in regulating Treg^{fci} ciMGPs and immune functions during lung cancer. The roles of *ETS1* in reprogramming of transcriptomic and metabolic profiles were confirmed in Treg^{fci} from human PBMCs treated with *ETS1* inhibitor or from mouse$^{ETS1\text{-}wt}$ or mice$^{ETS1\text{-}cKO}$ blood and lung tissues.

RESULTS

Selection of rOER-based Ts/s ciMGPs

To initiate the selection of T cell ciMGPs, Ts/s ciMGPs were collected from our own data, 11 self-established scRNA-seq databases, and publicly available scRNA-seq repositories (Fig. 1a). About 300 candidates of ciMGPs were shortlisted from an initial screening and selection for T cell annotations in the subsequent scRNA-seq analyses. A refined list of 170–262 ciMGPs was retained for downstream analysis, after the removal of duplicate entries or synonymous gene sets across species and tissues. The corresponding T-cell subsets were consistently renamed based on standardized CD markers or characteristic functional genes to ensure uniform nomenclature, as detailed in Supplementary Table 1. T cell ciMGPs in human blood samples and tissues were subjected to specificity evaluation by assessing the cutoff values, number of overlaps, rOER filtering, and ciMGP sorting (Fig. 1a).

To define the specificity of each ciMGP in the annotation of scRNA-seq data, the overlapping regions of the target ciMGP mRNA expression were quantified and visualized using boxplots. The target panel box between the corresponding Q3 and Q1 was defined as the repeatable range of non-target ciMGPs, as explained in Fig. 1b. The cutoff value was set at 10% of the total number of ciMGPs identified per sample by scRNA-seq and adjusted according to the number of detected subtypes, as described in Section 6 of the Supplementary Materials. The threshold of the selection was defined by re-calculating and establishing more than 70% sensitivity, less than 25% false positive rates, and more than 60% reproducibility and consistency in the regional line of Number30%/Number60% and the ternary classification-ss ciMGP (Supplementary Table 13). Based on the proportion of ciMGPs overlapping within the repeatable range, ss-ciMGPs were defined as those with $Number_{>30\%}$ <cutoff value (Fig. 1c), sa-ciMGPs as those between $Number_{>30\%}$ >cutoff value and $Number_{>60\%}$ < cutoff value (Fig. 1d), and sr-ciMGPs as those with $Number_{>60\%}$ >cutoff value (Fig. 1e). The trade-off values between sensitivity and specificity were detailed in Supplementary Table 17. The intermediate thresholds between 10–20% provided the optimal balance between sensitivity, specificity, and reproducibility, as detailed in Section 7 of the Supplementary Materials and listed in Supplementary Table 17.

To expound the presence of selected ciMGPs in normal PBMCs, 89,982 PBMCs were clustered from a cohort of healthy individuals ($n = 31$), of which 123 Ts/s were detectable. Clinical reference ranges for those Ts/s mainly included the mean, median, standard deviation, and standard error values (Table 1 and Supplementary Table 2.1), to mimic the report format of clinical biochemistry or hematology, although the number of clinical samples was limited and the accuracy requires further validation. Among the identified ciMGPs, 29 (39%) were defined as ss-ciMGPs, 8 (11%) as sa-ciMGPs, and 37 (50%) as sr-ciMGPs in healthy individuals (Supplementary Table 2.2). To explore the distribution and repeatability, the number (Supplementary Table 3) and rOER values (Supplementary Table 4) of ss-ciMGPs (Supplementary Tables 3.1 and 4.1), sa-ciMGPs (Supplementary Tables 3.2 and 4.2), and sr-ciMGPs (Supplementary Tables 3.3 and 4.3) were quantified across 10% rOER intervals. The ss-ciMGPs, sa-ciMGPs, and sr-ciMGPs exhibited distinct rOER distributions. The largest overlapping region of the

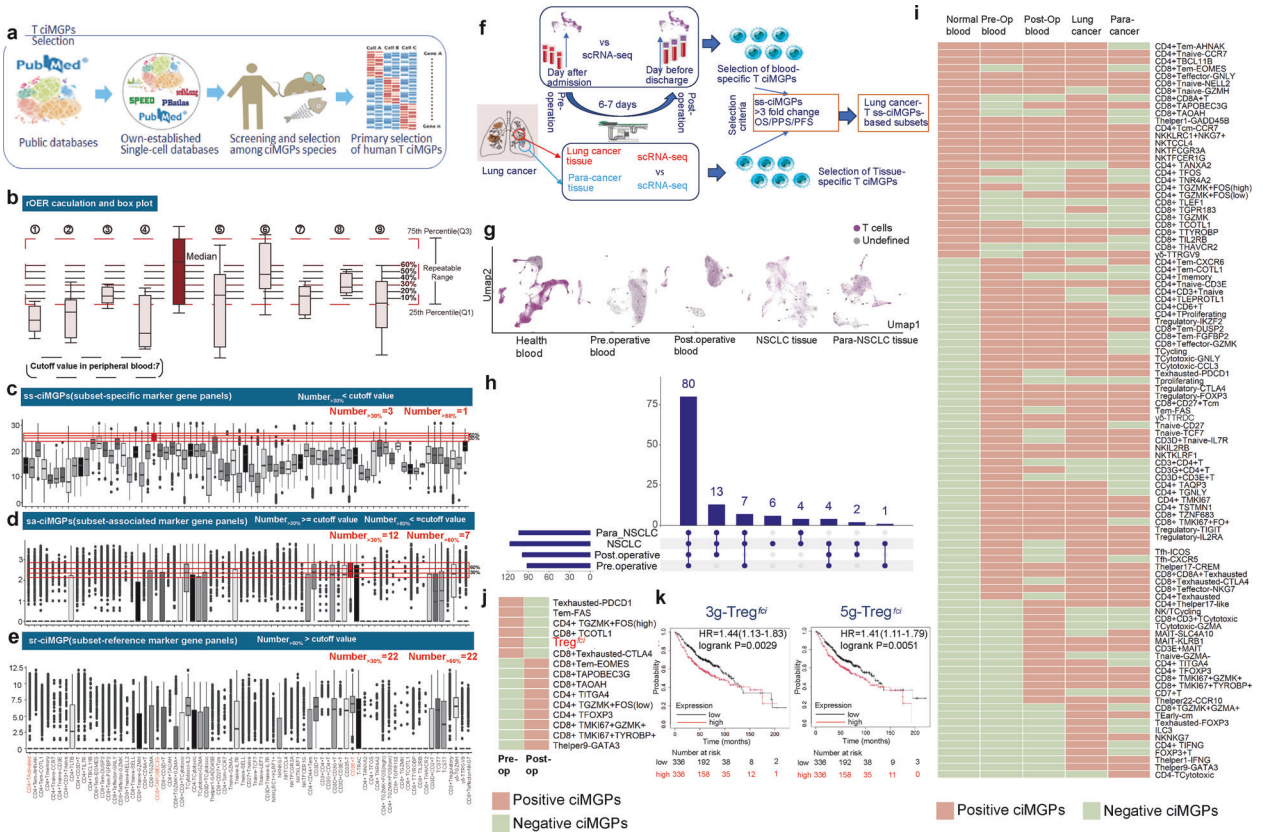


Fig. 1 Workflow and classification of T ciMGPs across immune compartments and clinical conditions. **a** A total of 300 candidate marker gene panels associated with distinct T cell subsets and functional states across species (human, mouse, and zebrafish) from various sources and selections. **b** Strategy for evaluating the specificity of ciMGPs using regional overlap expression rate (rOER). Columns 1–9, which represent the mean expression level of each ciMGP across the scRNA-seq datasets under evaluation that were used for the repeatability assessment. The red boxplot summarizes the distribution (Q1–median–Q3) of rOER values for the target ciMGPs within the focal T-cell subtype. The corresponding black boxplots, which represent non-target ciMGPs used as internal controls for specificity. **c** Subset-specific ciMGPs (ss-ciMGPs) as ciMGPs in fewer than the threshold number of subsets with rOER >30%. **d** Subset-associated ciMGPs (sa-ciMGPs) as ciMGPs with intermediate rOER values (30–60%) in a greater-than-threshold number of subsets. **e** Subset-reference ciMGPs (sr-ciMGPs) were characterized as ciMGPs expressed in more than the threshold number of subsets with rOER >60%, representing broadly expressed ciMGPs with limited discriminatory power. **f** Schematic overview of the study design for validation of circulating T ciMGPs. **g** UMAP visualization of T cell subset distribution based on rOER-defined subset specificity. **h** Upset plot illustrating intersections of identified T ciMGPs across NSCLC patient blood and lung tissue samples. **i** Heatmap displaying the presence (orange) or absence (green) of ss-ciMGPs in pre- or post-operative blood, NSCLC, or para-NSCLC lung tissues. **j** Heatmap showing appearance and variation of rOER-defined T cell ciMGPs in pre- or post-operative blood. **k** The impact of the low (black line) and high expression (red line) of 5g- or 3g-Treg^{fci} ciMGP signatures on the long-term overall survival rate of 672 patients with lung cancer over 200 months by Kaplan–Meier survival analysis

ss-ciMGPs was <10% in the $_{<30\%}$rOER, where the overlaps of $^{CD8+}$T cells, $T_{helper1\text{-}GADD45B}$, NK cells, and γδ-$T_{TRGV9}$ were almost zero.

The clinical characteristics of the NSCLC patients selected for scRNA-seq exhibited less obvious respiratory symptoms or abnormal physical findings, although sputum production, pain, or sleep disturbance occurred, as detailed in Supplementary Tables 18 and 19. Among the 12 patients, 4 or 2 had the smoking or family history of lung cancer. One had pleural effusion indicative of metastasis. Abnormalities in clinical biochemistries included elevated C-reactive protein (3/12), liver dysfunction (4/12), electrolyte disturbances (4/12), renal function changes (4/12), and elevated tumor markers (CEA, Cyfra21-1, 2/12). The cumulative abnormality score ranged from 13 to 37, indicating substantial inter-patient variability. Across the multicenter cohort of 124 NSCLC patients, the majority were recruited with the widest age range (22–89 years), as detailed in Section 17 of the Supplementary Materials.

Identification of NSCLC-specific T cell ciMGPs

To identify NSCLC-specific T cell ciMGPs, this study recruited patients aged between 47–78 years at the initial clinical visit, with disease stages ranging from I to III (Supplementary Table 5). scRNA-seq was performed on 122,747 T cells from peripheral blood samples prior to and following surgical resection, and from tumor and adjacent para-tumor lung tissues of patients with NSCLC (Fig. 1f). Approximately 9915 or 11,675 T cells were analyzed from pre- or post-operative blood samples, whereas 11,742 or 10,268 T cells were profiled from NSCLC or para-NSCLC tissues, respectively (Fig. 1g). The prevalence of T cell ciMGPs after the primary screening were validated and 93 or 99 T cell ciMGPs were detected in pre- or post-operative blood samples, while 117 or 104 T cell ciMGPs in NSCLC or para-NSCLC tissues, respectively (Fig. 1h). Of those, 80 ciMGPs appeared in all samples, 13 in post-operative blood, NSCLC, and para-NSCLC samples, and 4 in pre- and post-operative blood, and NSCLC samples, as depicted in Fig. 1h. The number (Supplementary Table 6) and rOER values (Supplementary Table 7) of ciMGPs were categorized as ss-ciMGPs (Supplementary Tables 6.1 and 7.1), sa-ciMGPs (Supplementary Tables 6.2 and 7.2), and sr-ciMGPs (Supplementary Tables 6.3 and 7.3). The ss-/sa-/sr-ciMGPs were 63/ 9/21 or 65/15/19 in pre- or post- operative blood, 78/19/ 20 in NSCLC tissue, and 58/ 21/25 in para-NSCLC tissue. The distribution of ciMGP categories

Table 1. The mean values of subset-specific cell identity marker gene panels (ss-ciMGPs) in blood before and after the operation in patients with lung cancer

Subsets names	Reference values (cells/μL)	Before operation	After operation
$^{CD4+}T_{em\text{-}AHNAK}$	0–334	12	14
$^{CD4+}T_{em\text{-}CXCR6}$	0–0	6↑	2↑
$^{CD4+}T_{em\text{-}COTL1}$	51–331	1↓	3↓
$^{CD4+}T_{memory}$	0–616	27	53
$^{CD4+}T_{naive\text{-}CCR7}$	0–295	5	19
$^{CD4+}T_{naive\text{-}CD3E}$	0–342	5	19
$^{CD4+CD3+}T_{naive}$	0–352	4	10
$^{CD4+}T_{LEPROTL1}$	0–0	1↑	2↑
$^{CD4+CD6+}T$	0–0	0	0
$^{CD4+}T_{BCL11B}$	0–139	1	2
$^{CD4+}T_{Proliferating}$	0–0	2↑	2↑
$^{CD4+}T_{helper17\text{-}like}$	0–0	0	1↑
$T_{regulatory\text{-}IKZF2}$	0–0	1↑	1↑
$^{CD8+}T_{em\text{-}EOMES}$	0–20	0	0
$^{CD8+}T_{em\text{-}DUSP2}$	29–291	4↓	3↓
$^{CD8+}T_{em\text{-}FGFBP2}$	51–331	1↓	3↓
$^{CD8+}T_{effector\text{-}GNLY}$	0–312	8	9
$^{CD8+}T_{effector\text{-}GZMK}$	52–336	14↓	18↓
$^{CD8+}T_{naive\text{-}NELL2}$	0–104	1	2
$^{CD8+}T_{naive\text{-}GZMH}$	0–292	6	5
$^{CD8+CD8A+}T$	22–264	0↓	0↓
$^{CD8+}T_{APOBEC3G}$	0–57	0	0
$^{CD8+}T_{AOAH}$	0–26	0	0
$NK/T_{Cycling}$	0–0	2↑	1↑
$T_{Cycling}$	0–0	0	1↑
$^{CD8+CD3+}T_{Cytotoxic}$	22–380	8↓	8
$T_{Cytotoxic\text{-}GNLY}$	0–0	1↑	0
$T_{Cytotoxic\text{-}CCL3}$	0–0	1↑	0
$T_{Cytotoxic\text{-}GZMA}$	20–432	15↓	13↓
$T_{exhausted\text{-}PDCD1}$	0–0	0	0
$T_{helper1\text{-}GADD45B}$	0–234	13	31
$T_{proliferating}$	0–0	7↑	8↑
$T_{regulatory\text{-}CTLA4}$	0–0	1↑	0
$T_{regulatory\text{-}FOXP3}$	0–0	1↑	4↑
$^{CD8+CD27+}T_{cm}$	0–268	1	1
$^{CD4+}T_{cm\text{-}CCR7}$	0–136	0	1
$T_{em\text{-}FAS}$	0–0	0	0
$\gamma\delta\text{-}T_{TRDC}$	0–0	24↑	23↑
$MAIT_{\text{-}SLC4A10}$	0–0	0	1↑
$MAIT_{\text{-}KLRB1}$	0–0	0	2↑
$^{CD3E+}MAIT$	0–0	0	2↑
$T_{naive\text{-}GZMA\text{-}}$	0–1100	9	23
$T_{naive\text{-}CD27}$	0–1402	9	17
$T_{naive\text{-}TCF7}$	0–760	10	35
$^{CD3D+}T_{naive\text{-}IL7R}$	0–1292	13	33
$NK_{KLRC1+NKG7+}$	12–82	4↓	2↓
NK_{IL2RB}	0–0	2↑	3↑
NKT_{CCL4}	0–264	4	2
NKT_{FCGR3A}	2–160	13	11
NKT_{KLRF1}	38–558	30↓	24↓
NKT_{FCER1G}	0–94	9	7
$^{CD3+CD4+}T$	0–23	0	0
$^{CD3G+CD4+}T$	0–24	0	1
$^{CD3D+CD3E+}T$	0–23	0	0
$^{CD4+}T_{AQP3}$	0–0	0	2↑
$^{CD4+}T_{ITGA4}$	0–0	0	1↑
$^{CD4+}T_{ANXA2}$	0–48	0	0
$^{CD4+}T_{FOS}$	0–56	0	0
$^{CD4+}T_{NR4A2}$	0–100	0	0
$^{CD4+}T_{GZMK+FOS(high)}$	0–240	1	0
$^{CD4+}T_{GZMK+FOS(low)}$	2–34	0↓	1↓
$^{CD4+}T_{GNLY}$	0–0	1↑	1↑
$^{CD4+}T_{FOXP3}$	0–0	0	0
$^{CD4+}T_{MKI67}$	0–0	1↑	0
$^{CD4+}T_{STMN1}$	0–0	1↑	0
$^{CD8+}T_{LEF1}$	0–9	0	0
$^{CD8+}T_{GPR183}$	0–53	0	0
$^{CD8+}T_{GZMK}$	0–1	0	0
$^{CD8+}T_{COTL1}$	0–7	0	0
$^{CD8+}T_{ZNF683}$	0–0	0	0
$^{CD8+}T_{TYROBP}$	0–325	6	3
$^{CD8+}T_{IL2RB}$	0–99	3	1
$^{CD8+}T_{MKI67+GZMK+}$	0–0	0	0
$^{CD8+}T_{MKI67+FO+}$	0–0	1↑	1↑
$^{CD8+}T_{MKI67+TYROBP+}$	0–0	0	0
$^{CD8+}T_{HAVCR2}$	0–3	0	0
^{CD7+}T	114–1416	22↓	37↓
$T_{regulatory\text{-}TIGIT}$	0–0	0	0
$T_{regulatory\text{-}IL2RA}$	0–0	1↑	4↑
$T_{regulatory\text{-}CTLA4+FOXP3+IL2RA+}$	0–0	0	0
$T_{fh\text{-}ICOS}$	0–0	0	0
$T_{fh\text{-}CXCR5}$	0–0	0	0
$T_{helper17\text{-}CREM}$	0–0	0	0
$T_{helper22\text{-}CCR10}$	0–0	0	1↑
$\gamma\delta\text{-}T_{TRGV9}$	27–169	0↓	3↓
$^{CD8+CD8A+}T_{exhausted}$	0–0	1↑	1↑
$^{CD8+}T_{exhausted\text{-}CTLA4}$	0–0	1↑	0
$^{CD8+}T_{effector\text{-}NKG7}$	0–66	0	0
$^{CD4+}T_{exhausted}$	0–0	6↑	0

varied among samples, as detailed in Supplementary Table 8. Among ss-ciMGPs, 13 uniquely appeared in NSCLC tissues, including $^{CD4+CD6+}T$, $^{CD4+}T_{AQP3}$, $^{CD4+}T_{IFNG}$, $^{CD4+}T_{GNLY}$, $^{CD8+}T_{GPR183}$, $^{CD8+}T_{IL2RB}$, $^{FOXP3+}T$, $T_{fh\text{-}CXCR5}$, ILC3, $^{CD8+}T_{em\text{-}EOMES}$, $^{CD8+CD8A+}T$, $^{CD8+}T_{AOAH}$, and $T_{exhausted\text{-}FOXP3}$ (Supplementary Table 9.1), 1 ss-ciMGP was detected exclusively in pre- operative blood but not in lung cancer tissues, 6 were found only in post- operative blood but not in para-cancer tissues, and 18 or 11 ss-ciMGPs were specific to NSCLC or para-cancer tissues, with none detected in post- operative blood (Supplementary Tables 9.2 and 9.3). Some ciMGPs transitioned from ss-ciMGPs to sa- or sr-ciMGPs after surgery (e.g., $^{CD4+}T_{memory}$, $T_{proliferating}$, $T_{naive\text{-}TCF7}$, $^{CD3D+}T_{naive\text{-}IL7R}$, $^{CD3+CD4+}T$, $^{CD3D+CD3E+}T$, $^{CD4+}T_{FOS}$), while others from sa-/sr-ciMGPs

into ss-ciMGPs (e.g., $T_{naive\text{-}GZMA}$-, $^{CD7+}$T, NK/$T_{Cycling}$, $T_{Cytotoxic\text{-}GZMA}$, and $^{CD8+CD3+}T_{Cytotoxic}$) (Supplementary Table 9.4).

Validation of NSCLC-specific T ciMGPs

To assess the disease specificity of T cell ciMGPs, we compared the presence and frequency of ciMGPs across different anatomical locations and time points between patients with NSCLC and heathy individuals. Compared to healthy ss-ciMGPs, approximately 41%, 38%, 21%, or 45% of Ts/s were absent in pre-operative blood, post-operative blood, NSCLC tissues, or para-NSCLC tissues, respectively. An increased number of ss-ciMGPs were observed, amounting to 159%, 162%, 190%, and 145% of the baseline values observed in healthy controls, respectively, across sample types (Fig. 1i).

To evaluate the NSCLC specificity and rOER-based variability of T cell ciMGPs among lung diseases, the presence and frequency of T cell ss-, sa-, or sr-ciMGPs in lung tissues of NSCLC were compared with LUAD and para-LUAD, COPD, IPF, and SSC. Of 602,237 cells collected, about 600,853 were single T cells from peripheral blood. The numbers of ss-, sa- or sr-ciMGPs were 63, 19 or 22 in the healthy, 59, 23 or 26 in LUAD, 58, 20 or 26 in para-LUAD, 58, 20 or 20 in COPD, 65, 14 or 23 in IPF, and 48, 21 or 29 in SSC samples, respectively (Supplementary Table 10). $^{CD4+}T_{activated}$, $^{CD4-}T_{cytotoxic}$, $^{CD4+CD44+}T_{em}$ and $T_{helper22\text{-}CCR10}$ were uniquely in healthy lung, $T_{em\text{-}FAS}$, $^{FOXP3+}$T and $T_{fh\text{-}CXCR5}$ in LUAD, $T_{Cytotoxic\text{-}GZMA}$ and $^{CD7+}$T in COPD, or $^{CD8+}T_{GZMK+GZMA+}$ in IPF (Supplementary Fig. 1a). The variability of sa- or sr-ciMGPs expression across lung diseases was shown in Supplementary Fig. 1b, c. The similarity rates of ss-ciMGPs to healthy lung tissue were 78%, 84%, 79%, 86%, or 65% in LUAD, para-LUAD, COPD, IPF, or SSC, respectively. Approximately 10–17 ss-ciMGPs (approximately 15%) were newly identified in NSCLC lungs. The differences of NSCLC-specific ss-ciMGPs were noticed from LUAD (33%), COPD (35%), IPF (26%), or SSC (44%; Supplementary Fig. 1d). The mean cell numbers of ss-ciMGP in NSCLC, LUAD, para-cancer, COPD, IPF, and SSC were compared with those in healthy lung tissues (Table 2). In comparison to chronic lung diseases, $^{CD4+CD6+}$T, $T_{early\text{-}cm}$, $^{γδ-}T_{TRDC}$, NKIL2RB, and $^{CD4+}T_{AQP3}$, of NSCLC-specific ciMGPs were exclusively detected in NSCLC patients.

To evaluate organ/tissue-associated specificity, we systematically assessed the distribution of ss-, sa- or sr-ciMGPs in 336,456 single T cells of total 370,218 single mononuclear cells among 17 organs/tissues (e.g., lung, choroid, small intestine, cervix, skin, tooth, bone marrow, eyes, breast, liver, mouth, pancreas, nose, nasal mucosa or cortex, and gastric and renal tissues), as detailed in Supplementary Table 11. Comparative analysis revealed that the ss-ciMGPs profiles in extrapulmonary organs/tissues differed from those in the lung by 38–100% (Supplementary Fig. 1e) and in NSCLC tissues by 45–100% (Supplementary Fig. 1f). To further evaluate NSCLC specificity, this study analyzed the distribution of ss-, sa-, or sr-ciMGPs across 31 extrapulmonary diseases, as detailed in Section 8 of the Supplementary Materials, and compared the ss-MGPs profiles in extrapulmonary diseases with NSCLC. The difference of ss-ciMGPs between NSCLC and extrapulmonary diseases ranged from 26–100%, which likely reflected the variations in disease nature, severity, and anatomical location (Supplementary Fig. 1g).

To select NSCLC-specific Ts/s, we defined the selection criteria of candidate Ts/s, which should be the ss-ciMGPs, exhibit differential expression greater than three fold changes, as compared to corresponding controls, and demonstrate significant association with patient survival outcomes (Supplementary Fig. 2a). The selection of above three fold changes was evaluated by comparing T cell ciMGP expression between NSCLC and para-NSCLC tissues and between pre- and post-operative blood samples (Supplementary Table 16). Of those, 6 or 9 T-cell subsets were detectable in pre- or post-operative blood (Fig. 1j), and some in NSCLC tissues (Supplementary Fig. 2b). Approximately 7 or 54 T cell subsets changed more than three folds between pre- and post-operative blood (Supplementary Fig. 2c) or between NSCLC and para-NSCLC tissues (Supplementary Fig. 2d), respectively. A majority of subsets were upregulated in pre-operative blood or NSCLC tissue and declined following tumor resection or in para-NSCLC tissues. Conversely, the number of $T_{Cycling}$ were lower in pre-surgery and became higher after surgery, while $^{CD4+}T_{cm\text{-}CCR7}$, $T_{Cytotoxic\text{-}GNLY}$, $^{CD8+}T_{naive\text{-}SELL}$, and $T_{exhausted\text{-}PDCD1}$ were notably lower in para-NSCLC tissues. Functional comparisons of Ts/s demonstrated that T cells exhibited greater activation and accumulation in NSCLC tissues compared to para-NSCLC tissues. Significant changes in Ts/s between pre- and post-operative blood or between NSCLC and para-NSCLC tissues were presented in Supplementary Fig. 2e.

Of Ts/s, 26 were significantly associated with PFS, OS, and PPS rates in NSCLC (Supplementary Table 12). The OS of lung cancer patients demonstrated significant differences between low and high profiles of Ts/s (Supplementary Fig. 2f). Among those, the number and expression of $Treg^{fci}$ were significantly higher in NSCLC samples and in pre-operative blood, while was absent in post-operative blood (Fig. 1j). The NSCLC patients with high 5g- or 3g-$Treg^{fci}$ had markedly poorer OS compared to those with low expression (Fig. 1k, $p = 0.002$ or 0.005, respectively), which was also seen in LUAD patients (Supplementary Fig. 11a; $p < 0.01$). To evaluate the potential impacts of 5g- or 3g-$Treg^{fci}$ on long-term survival rate of patients with multiple cancers, as detailed in the Section 18 of Supplementary Materials, we re-evaluated and found that patients with the high expression of 5g- or 3g-$Treg^{fci}$ ciMGPs had a poor survival rate and there was no difference between 5g- and 3g-$Treg^{fci}$ (Supplementary Fig. 11b–f). $Treg^{fci}$ was selected as one of the NSCLC-specific Ts/s on basis of the high specificity of its ciMGPs, its unique appearance or greater than three-fold of ciMGPs gene expressions than controls, and its strong correlation with survival rates.

Validation of selected target $Treg^{fci}$

To further validate the specificity of human $Treg^{fci}$, we recruited 45 healthy volunteers aged between 20–50, performed scRNA-seq on PBMCs, and compared those from public datasets (Supplementary Fig. 3a). Among the identified Ts/s, 123 (63.1%) were detected in PBMCs from our cohort, of which 72 Ts/s were uniquely found in our cohort (Fig. 2a). About 71% of Ts/s in our cohort were ss-ciMGPs and larger than that in public databases (37%; Fig. 2b), whereas the functional compositions of Ts/s exhibited comparable patterns between our cohort and public databases (Fig. 2c). Within the Ts/s, NK and NKT cells accounted for 45% of the top 20 most significantly altered Ts/s (Supplementary Fig. 3b). UMAP visualization revealed strong associations between the T cell composition in blood and the tumor burden status (Fig. 2d), with substantial overlaps and unique subtype distributions in NSCLC tissues, adjacent non-tumor tissues, pre- and post-operative blood, and blood from healthy subjects (Fig. 2e). 3g-$Treg^{fci}$ was one of the few subtypes that appeared or disappeared in circulation before or after tumor resection (Fig. 2f), and was significantly enriched in NSCLC tissues (Fig. 2g). Upon comparing among Treg subsets, $Treg^{fci}$ had more than three-fold expression of ciMGPs and obvious differences in the subset abundance including 5g-$Treg^{fci}$ between NSCLC and para-NSCLC tissues (Supplementary Fig. 8a). The variations and consistencies of spatial distributions between 5g- and 3g-$Treg^{fci}$ were compared by evaluating the appearance and location, respectively (Supplementary Fig. 8b, c). To assess the potential influence of age, we measured and found that the appearance frequency of 5g- or 3g-$Treg^{fci}$ was similar between groups with ages <50 or ≥50 from public datasets (Supplementary Fig. 10a).

To confirm the detectability of $Treg^{fci}$ in blood, flow cytometric staining was performed for classic Treg with CD45+CD3+CD4+CD25+CD127− and $Treg^{fci}$ markers in samples

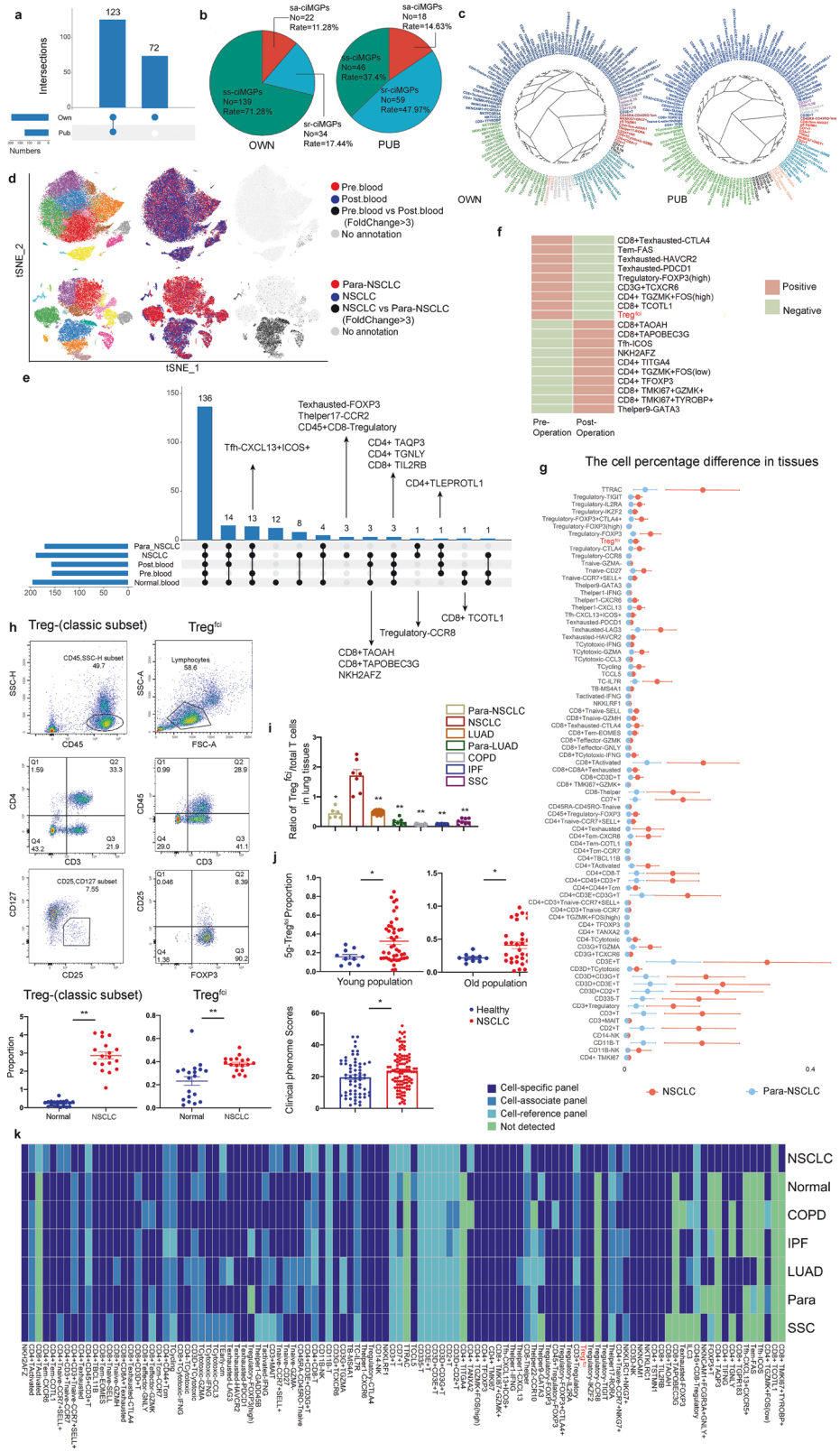
Table 2. The mean values (cells/μL) of subset-specific cell identity marker gene panels (ss-ciMGPs) in lung tissues in healthy or patients with non-small cell lung cancer (NSCLC), para-NSCLC, lung adenocarcinoma (LUAD), para-cancer tissue (Para), chronic obstructive pulmonary diseases (COPD), idiopathic pulmonary fibrosis (IPF), or systemic sclerosis-associated interstitial lung disease (SSC)

	Normal	NSCLC	Para-NSCLC	LUAD	Para-LUAD	COPD	IPF	SSC
$^{CD4+}T_{Activated}$	0–19	165↑	28↑	48↑	43↑	8	6	11
$^{CD4+}T_{em\text{-}AHNAK}$	14–130	209↑	146↑	40	53	82	110	83
$^{CD4+}T_{em\text{-}CXCR6}$	0–47	184↑	20	61↑	117↑	31	15	39
$^{CD4+}T_{em\text{-}COTL1}$	0–42	16	3	105↑	29	8	9	14
$^{CD4+}T_{naive\text{-}CCR7}$	0–2	7↑	2	7↑	2	1	1	1
$^{CD4+}T_{naive\text{-}CD3E}$	0–3	9↑	4↑	12↑	5↑	1	1	1
$^{CD4+CD3+}T_{naive}$	0–3	11↑	1	12↑	3	1	0	1
$^{CD4+CD6+}T$	0–0	2	0	0	0	0	0	0
$^{CD4+}T_{BCL11B}$	0–39	8	2	2	4	42↑	22	1
$^{CD4+}T_{Proliferating}$	0–47	92↑	74↑	74↑	19	19	64↑	62↑
$^{CD4+}T_{helper17\text{-}like}$	0–8	5	3	12↑	4	1	1	2
$T_{regulatory\text{-}IKZF2}$	0–3	37↑	9↑	6↑	3	2	3	0
$^{CD8+}T_{em\text{-}EOMES}$	0–1	1	0	1	2↑	0	0	0
$^{CD8+}T_{em\text{-}DUSP2}$	0–52	21	7	103↑	29	2	5	13
$^{CD8+}T_{em\text{-}FGFBP2}$	0–42	16	3	105↑	29	8	9	14
$^{CD8+}T_{effector\text{-}GNLY}$	0–78	7	8	16	28	12	8	11
$^{CD8+}T_{effector\text{-}GZMK}$	0–90	81	28	240↑	92↑	4	12	24
$^{CD8+}T_{naive\text{-}NELL2}$	0–1	1	0	2↑	1	1	0	1
$^{CD8+}T_{naive\text{-}SELL}$	0–19	53	14	54↑	17	13	17	16
$^{CD8+}T_{naive\text{-}GZMH}$	0–78	54	7	129↑	65	24	16	15
$^{CD8+CD8A+}T$	0–0	2↑	0	0	0	0	0	0
$^{CD8+}T_{APOBEC3G}$	0–0	2↑	0	0	0	0	0	0
$^{CD8+}T_{AOAH}$	0–15	0	0	1	1	24↑	16↑	0
$^{CD8+}T_{GZMK+GZMA+}$	0–135	101	40	341↑	134	3	11	20
$T_{Early\text{-}cm}$	0–0	2↑	1↑	0	0	0	0	0
$T_{Cycling}$	0–40	64↑	55↑	64↑	21	14	54↑	59↑
$^{CD8+CD3+}T_{Cytotoxic}$	0–56	67↑	13	128↑	56	11	10	10
$T_{Cytotoxic\text{-}GNLY}$	0–78	16	1	66	23	3	9	19
$^{CD4-}T_{Cytotoxic}$	0–71	68	14	142↑	64	19	15	15
$T_{Cytotoxic\text{-}CCL3}$	0–81	28	2	85↑	26	4	10	19
$T_{Cytotoxic\text{-}GZMA}$	0–121	104	22	211↑	102	33	24	26
$T_{exhausted\text{-}FOXP3}$	0–0	3↑	0	1↑	1↑	0	0	0
$T_{exhausted\text{-}PDCD1}$	0–0	26↑	1↑	12↑	1↑	0	0	0
$T_{helper1\text{-}GADD45B}$	0–344	32	123	178	135	76	134	88
$T_{proliferating}$	0–77	142↑	104↑	128↑	36	36	100↑	91↑
$T_{regulatory\text{-}CTLA4}$	0–5	88↑	14↑	27↑	12↑	3	5	3
$T_{regulatory\text{-}FOXP3}$	0–8	63↑	13↑	31↑	13↑	3	6	3
$^{CD8+CD27+}T_{cm}$	0–12	20↑	6	69↑	14↑	1	2	4
$^{CD4+}T_{cm\text{-}CCR7}$	0–5	2	0	16↑	5	2	3	5
$T_{em\text{-}FAS}$	0–0	1↑	0	0	0	0	0	0
γδ-T_{TRDC}	0–0	28↑	11↑	0	0	0	0	0
ILC3	0–0	0	0	1↑	0	0	0	0
$MAIT_{\text{-}SLC4A10}$	0–6	21↑	5	2	1	4	3	0
$MAIT_{\text{-}KLRB1}$	0–14	8	7	5	2	13	7	1
$^{CD3E+}$MAIT	0–16	33↑	10	5	2	12	7	1
$T_{naive\text{-}GZMA\text{-}}$	0–18	16	3	74↑	18	3	5	12
$T_{naive\text{-}SELL}$	0–42	41	10	141↑	38	13	20	27
$T_{naive\text{-}CD27}$	0–52	120↑	26	288↑	61↑	6	12	23
$T_{naive\text{-}TCF7}$	0–4	7↑	2	14↑	4	1	2	3
$T_{naive\text{-}LEF1}$	0–42	41	10	141↑	38	13	20	27
$^{CD3D+}T_{naive\text{-}IL7R}$	0–58	22	6	176↑	53	11	19	39

Table 2. continued

	Normal	NSCLC	Para-NSCLC	LUAD	Para-LUAD	COPD	IPF	SSC
$NK_{KLRC1+NKG7+}$	0–57	8	2	14	17	14	13	11
NK_{IL2RB}	0–0	6↑	3↑	0	0	0	0	0
NK_{NKG7}	0–553	82	61	125	195	220	138	80
NKT_{CCL4}	0–139	4	1	24	39	15	16	14
NKT_{FCGR3A}	0–36	19	10	6	11	22	14	4
NKT_{KLRF1}	0–165	22	13	34	82	28	17	9
NKT_{FCER1G}	0–35	7	6	15	21	6	5	8
$^{CD4+CD44+}T_{em}$	0–73	71	14	145↑	66	20	15	16
$^{CD4+}T_{AQP3}$	0–0	1↑	0	0	0	0	0	0
$^{CD4+}T_{ITGA4}$	0–0	0	0	0	0	0	0	0
$^{CD4+}T_{ANXA2}$	0–1	0	0	2↑	0	0	0	1
$^{CD4+}T_{FOS}$	0–4	1	1	5↑	3	0	0	2
$^{CD4+}T_{NR4A2}$	0–14	6	2	24↑	4	1	2	8
$^{CD4+}T_{GZMK+FOS(high)}$	0–26	2	0	31↑	13	0	2	7
$^{CD4+}T_{GZMK+FOS(low)}$	0–1	1	0	2↑	0	0	0	0
$^{CD4+}T_{IFNG}$	0–2	0	0	2	1	0	0	1
$^{CD4+}T_{GNLY}$	0–1	0	0	0	0	0	0	0
$^{CD4+}T_{FOXP3}$	0–0	3↑	1↑	1↑	1↑	0	0	0
$^{CD4+}T_{MKI67}$	0–2	22↑	4↑	6↑	3↑	1	1	0
$^{CD4+}T_{STMN1}$	0–7	42↑	14↑	30↑	10↑	2	2	1
$^{CD8+}T_{GPR183}$	0–1	0	0	2↑	0	0	0	0
$^{CD8+}T_{ZNF683}$	0–1	1	0	0	0	0	0	0
$^{CD8+}T_{TNF}$	0–0	0	0	0	0	0	0	0
$^{CD8+}T_{TYROBP}$	0–64	1	1	9	11	3	1	2
$^{CD8+}T_{IL2RB}$	0–13	1	0	1	1	0	0	0
$^{CD8+}T_{MKI67+GZMK+}$	0–1	4↑	2↑	10↑	1	0	0	0
$^{CD8+}T_{MKI67+FO+}$	0–45	25	52↑	55↑	22	14	67↑	65↑
$^{CD8+}T_{MKI67+TYROBP+}$	0–0	0	0	0	0	0	0	0
$^{CD8+}T_{HAVCR2}$	0–0	0	0	1↑	0	0	0	0
^{CD7+}T	0–242	447↑	47	698↑	342↑	81	84	121
$^{FOXP3+}T$	0–0	0	0	0	0	0	0	0
$T_{regulatory\text{-}TIGIT}$	0–1	64↑	10↑	9↑	5↑	0	2↑	1
$T_{regulatory\text{-}IL2RA}$	0–8	63↑	13↑	31↑	13↑	3	6	3
$T_{regulatory\text{-}CTLA4+FOXP3+IL2RA+}$	0–4	46↑	11↑	16↑	10↑	2	4	2
$T_{fh\text{-}ICOS}$	0–0	44↑	3↑	6↑	1↑	0	0	0
$T_{fh\text{-}CXCR5}$	0–0	2↑	0	1↑	0	0	0	0
$T_{helper1\text{-}IFNG}$	0–46	7	1	0	1	22	9	1
$T_{helper9\text{-}GATA3}$	0–0	4↑	0	0	0	0	0	0
$T_{helper17\text{-}CREM}$	3–97	5	2↓	48	15	81	62	44
$T_{helper22\text{-}CCR10}$	0–0	0	0	1↑	0	0	0	0
$γδ\text{-}T_{TRGV9}$	0–0	4↑	2↑	0	0	0	0	0
$^{CD8+CD8A+}T_{exhausted}$	0–19	55↑	10	108↑	15	3	4	5
$^{CD8+}T_{exhausted\text{-}CTLA4}$	0–2	85↑	5↑	48↑	8↑	1	2	2
$^{CD8+}T_{effector\text{-}NKG7}$	0–10	6	1	48↑	15↑	0	0	2
$^{CD4+}T_{exhausted}$	0–3	198↑	5↑	165↑	3	2		0

from healthy volunteers ($N = 19$), 5g-Treg^{fci} in healthy volunteers ($N = 19$), 3g-Treg^{fci} patients with NSCLC ($N = 18$), and 5g-Treg^{fci} in NSCLC patients ($N = 102$). The Treg and Treg^{fci} populations in blood were significantly increased in NSCLC patients, as compared to the healthy controls (Fig. 2h; $p < 0.01$). To assessed the disease and organ/tissue specificity of Treg^{fci}, 64 selected Ts/s, including Treg^{fci}, were evaluated in tissues/organs and pulmonary or extra-pulmonary diseases. To define the NSCLC specificity, we compared the number of Treg^{fci} among lung diseases and found that the ratio of Treg^{fci}/total T cells in NSCLC tissues was significantly higher than in para-NSCLC tissues, or in LUAD, para-LUAD, COPD, IPF, and SSC (Fig. 2i). To assess the age impact on Treg^{fci} abundance, we isolated and measured the number of 5g-Treg^{fci} from peripheral blood of an additional 76 NSCLC patients using



flow cytometry, including 43 young and 33 old patients, and found the number of 5g-Treg^{fci} was significantly higher in young and old NSCLC patients than in young and old healthy controls (Fig. 2j, $p < 0.05$, respectively). There was no difference between young and old health controls (Supplementary Fig. 10g). An additional study on immunohistochemistry staining of Treg^{fci} marker and target proteins was performed in young and old NSCLC and para-NSCLC tissues. The scores of Treg and Treg aggregates (Supplementary Fig. 10b, c) and the expression intensity of *ETS1* protein (Supplementary Fig. 10d, e) were higher

Fig. 2 The profiling of T cell ciMGPs across healthy individuals, NSCLC patients, lung diseases, and extrapulmonary conditions. **a** Upset plot showing the intersections of identified T ciMGPs in PBMCs. Each bar represents the number of elements in each intersection, while the matrix above indicates the presence (black) or absence (gray) of ciMGPs in each set. Intersections are ordered by frequency, with total set sizes displayed at the top. **b** Pie charts categorizing T ciMGPs identified in PBMCs, as determined using the rOER tool. These ciMGPs are categorized into ss-ciMGPs, sa-ciMGPs, and sr-ciMGPs. **c** Phylogenetic tree visualizing the clustering of 195 and 123 T cell ciMGPs using the rOER tool. **d** t-SNE plots displaying integrated scRNA-seq data from NSCLC patient blood and lung tissues. Clusters are annotated by rOER-defined subsets, >3-fold change between pre- and post-operative blood or between NSCLC and para-NSCLC tissues. **e** Upset plot illustrating identified T ciMGPs intersections across blood and lung tissue samples from NSCLC patients and the healthy volunteers. Each bar represents the number of elements in a specific intersection, with the corresponding intersection matrix above indicating presence (dark blue) or absence (gray) of elements in each set. **f** Heatmap showing the presence (orange) or absence (green) of T ciMGPs in pre- or post-operative blood of NSCLC patients. Each row and column represent ciMGPs and sample types. **g** Dot plot depicting the proportion of T ciMGPs with >3-fold changes between NSCLC and para-NSCLC tissues. **h** Flow cytometric analysis of classical Treg stained with CD4, CD3, CD45, CD25 and CD127 antibodies and $Treg^{fci}$ subset stained with FOXP3, CTLA4 and IL2RA antibodies (3g-$Treg^{fci}$) in peripheral blood obtained from healthy controls ($n = 19$) and NSCLC patients ($n = 18$). Bar plots show frequencies (mean ± SEM). **i** Bar plot showing the proportion of $Treg^{fci}$ in lung tissues of the normal, LUAD, adjacent para-LUAD, chronic obstructive pulmonary disease (COPD), idiopathic pulmonary fibrosis (IPF), and systemic sclerosis-associated interstitial lung disease (SSC-ILD) as determined from the scRNA-seq data. *stands for the p values < 0.05 vs. NSCLC patients. **j** Scatter plot showing the proportion of 5g-$Treg^{fci}$ in the peripheral blood of healthy volunteers ($N = 20$) and NSCLC patients ($N = 76$), stratified by age. *stands for the p values < 0.05 vs. health volunteers. The clinical phenome scores in NSCLC patients ($N = 124$), stratified by the age. *stands for the p values < 0.05 vs. the young group. **k** Heatmap comparing the distribution and variability of 108 selected T ciMGPs across NSCLC tissues with normal lung, LUAD, para-LUAD, COPD, IPF, and SSC

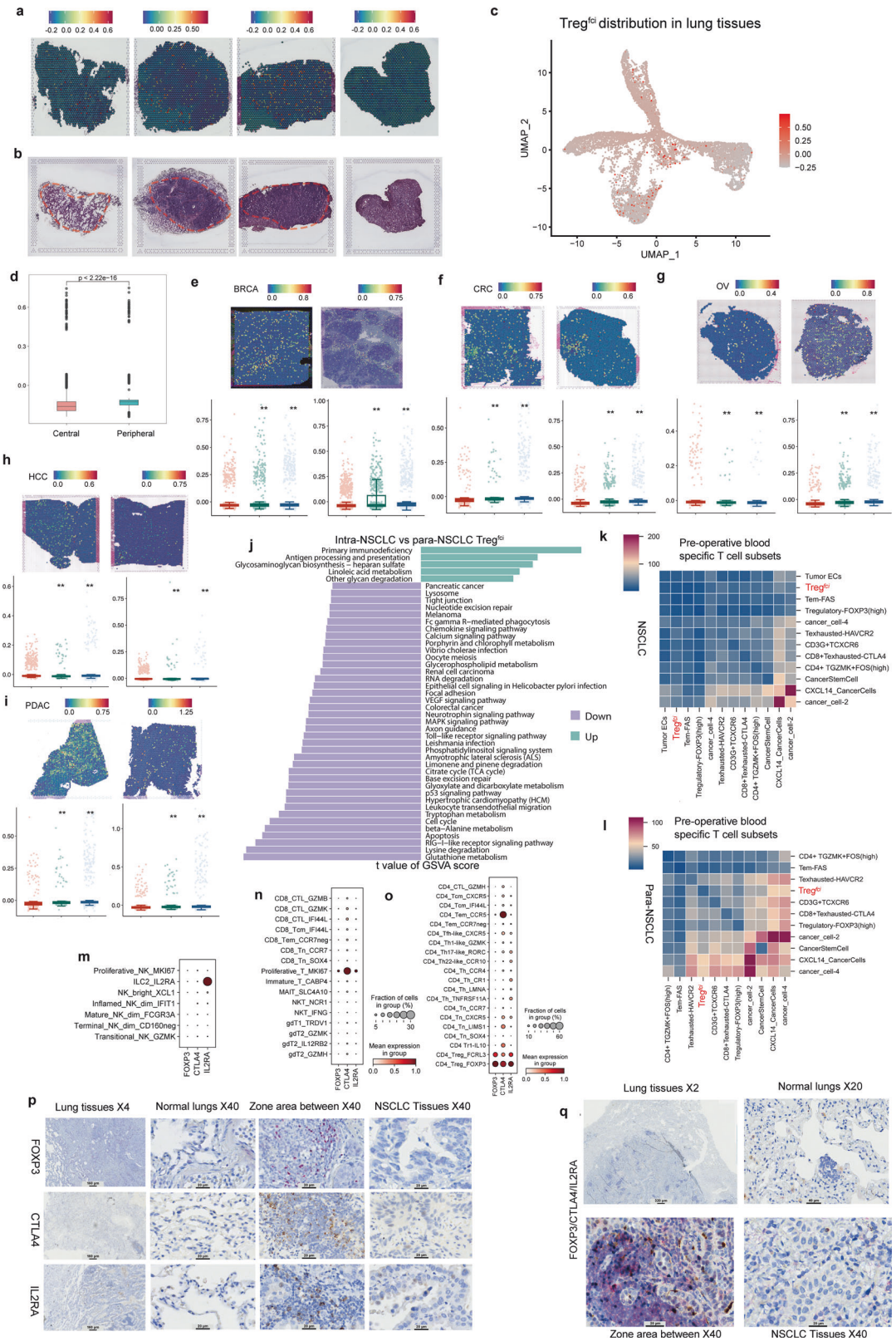
in intra- and peri-tumoral tissues of the old patients, compared with the young patients. The patterns of T cell subsets with ciMGP in NSCLC were compared with other lung diseases, as shown in Fig. 2k. Among 17 organs/tissues, $Treg^{fci}$ was primarily observed in breast, cervical, skin, small intestine and mouth tissue (Supplementary Fig. 9a). Of the 31 extrapulmonary diseases analyzed, $Treg^{fci}$ was highly represented in clear cell renal carcinoma (No. 16), and colorectal cancer (No. 17), as depicted in Supplementary Fig. 9b.

To elucidate the spatial distribution and biological characteristics of $Treg^{fci}$, stereo-seq measurements (Fig. 3a) and HE staining (Fig. 3b) were performed on matched lung tissue sections obtained from healthy donors and NSCLC patients. The $Treg^{fci}$, ILC3 (Supplementary Fig. 3c), and CD8 $T_{cytotoxic\text{-}IFNG}$(Supplementary Fig. 3d) were diffusely distributed across lung cancer and para-cancer tissues. $Treg^{fci}$ were significantly enriched in NSCLC tissues and exhibited markedly distinct spatial distribution patterns between tumor cores and adjacent areas, as depicted in the representative cases (Fig. 3a, panels 1–3). The substantial differences between the central and peripheral NSCLC tumor zones were noted (Fig. 3c). The $Treg^{fci}$ specificity of spatial distribution and appearance was evidenced by the higher expression of $Treg^{fci}$ ciMGPs in the center of NSCLC, compared with other cancers (Supplementary Fig. 12a).

The abundance of $Treg^{fci}$ was significantly lower in the central regions of tumors than in the peripheral areas (Fig. 3d; $p < 0.01$). The expression of $Treg^{fci}$ ciMGPs were enriched in the para-malignant zones but markedly declined in the malignant regions, thereby suggesting potential immune evasion mechanisms in advanced tumor states. To explore the unique patterns of $Treg^{fci}$ across various cancer types, the localization of $Treg^{fci}$ was evaluated in BRCA, CRC, HCC, OV, and PDAC. The density of $Treg^{fci}$ in the interface zone between the tumor and para-tumor interface was significantly higher than within the tumor tissue itself in 2 out of 6 BRCA patients, even higher than in the para-tumor regions (Fig. 3e; $p < 0.01$; Supplementary Fig. 3e). Similar distribution patterns were observed in CRC (Fig. 3f) and OV (Fig. 3g). In contrast, the interface zones in HCC (Fig. 3h) and PDAC (Fig. 3i) had significantly higher $Treg^{fci}$ density, as compared to para-cancer tissues ($p < 0.01$). The transcriptomic comparisons between the $Treg^{fci}$ in NSCLC and para-NSCLC revealed the upregulation of multiple signaling pathways, including primary immunodeficiency, linoleic acid metabolism, antigen processing and presentation, and glycan biosynthesis-heparan sulfate (Fig. 3j). To characterize $Treg^{fci}$-mediated cell-cell communication of pre-operative blood-specific Ts/s, those Ts/s ciMGPs present in NSCLC and para-NSCLC tissues were assessed. $Treg^{fci}$ interactions with Tem-FAS, $Treg^{fci}$ itself in NSCLC, and $T_{regulatory\text{-}\ FOXP3(high)}$ cells represented the weakest interactions (Fig. 3k). Pre-operative blood-specific T cell subsets exhibited considerably upregulated interactions with cancer ciMGPs detected in the para-NSCLC region, particularly between cancer-CXCL14 and cancer stem cells, as compared to the communication with the rest of the cell subsets (Fig. 3l). $T_{fh\text{-}ICOS}$ and $^{CD8}+T_{MKI67+GZMK+}$ showed stronger interactions with cancer ciMGPs in para-NSCLC region (Supplementary Fig. 3f), as compared to post-operative blood specific T cell subsets in NSCLC tissues.

The $Treg^{fci}$-mediated interactions with Tem-FAS and $Treg^{fci}$ itself were weak in para-NSCLC tissues (Fig. 3l). The expression of $Treg^{fci}$ ciMGPs was also validated using the recently published Chinese Immune Multi-Omics Atlas (CIMA), which includes sex-, age-, and genetics-related immune profiles from 428 adults and compares over 10 million immune cells.[19,20] According to the CIMA annotation, *IL2RA* was highly expressed in an ILC2 subset (Fig. 3m), *CTLA4* in a MKI67+CD8+T cell subset (Fig. 3n), and *IFI44L* in a $CD4^+$T cell subset. $Treg^{fci}$ ciMGPs were prominently expressed in two annotated Treg subsets: CD4_Treg_FCRL3 and CD4_Treg_-FOXP3 (Fig. 3o). The pre-switched memory B cell_JAM and DC_CLEC9A subsets also exhibited high expression of *IL2RA* and *CTLA4*, respectively (Supplementary Fig. 3g). To precisely map the spatial localization of $Treg^{fci}$, we conducted immunostaining with specific antibodies targeting individual ciMGP proteins and a multiplex panel of all three $Treg^{fci}$ markers. FOXP3-positive cells were predominantly localized within the TME (Fig. 3p), confirming the spatial enrichment of $Treg^{fci}$ in tumor tissues. The multiplex staining demonstrated that $FOXP3^+CTLA4^+IL2RA^+$-positive cells were predominantly enriched within the tumor tissue and adjacent areas (Fig. 3q).

To define the functional and regulatory heterogeneities of $Treg^{fci}$, we compared the transcriptomic profiles of $Treg^{fci}$ with pre- and post-operative T cells. The top 10 upregulated and downregulated genes were identified in $Treg^{fci}$ by comparing with other T cells in pre-operative blood (Fig. 4b), along with differences between $Treg^{fci}$ in pre-operative blood and total T cells in post-operative blood (Fig. 4a). Among DEGs, *ETS1* was significantly upregulated in pre-operative $Treg^{fci}$, as compared to pre- or post-operative T cells. *ETS1*, *RGS1*, and *TRBC2* consistently elevated in $Treg^{fci}$ of pre-operative blood. The inter-subset variations of pre-operative blood-specific Ts/s were investigated by comparing the quantitative frequencies (Fig. 4c) and visualizing the DEG profiles of 20 pre- versus post-operative blood-specific Ts/s in heatmaps, separately (Fig. 4d). *ETS1* emerged as the predominantly upregulated gene in $Treg^{fci}$. The ETS1 expression values were measured across all spots of spatial transcriptome



samples (Fig. 4e) and within the Treg^{fci} specifically (Fig. 4f), by comparing the central and peripheral tumor regions of NSCLC.

To understand the roles of ETS1 down-regulated Treg^{fci} in the human system, we established the special humanized animal model with Treg$^{ETS1\text{-}KD}$ or Treg$^{ETS1\text{-}NC}$ and found that the rate of human CD45 positive cells was more than 25% in humanized mice 2 weeks after human PBMC-engraft and those animals were subjected to tumor cell inoculation (Supplementary Fig. 12b). From day 3 to day 15, we observed that the increased proportion of human CD45 positive and the decreased mouse CD45 positive

Fig. 3 Spatial and functional profiling of Treg^{fci} subset in NSCLC and other solid tumors using transcriptomic and histological analyses. **a** Spatial transcriptomic visualization of Gene Set Variation Analysis (GSVA) scores for the Treg^{fci} gene signature across tissue sections from 3 NSCLC patients. Color-mapped GSVA scores illustrate the spatial heterogeneity of Treg^{fci} subset localization in tumor and para-tumor regions. **b** Representative Hematoxylin and eosin (H&E)-stained images and morphological architectures of NSCLC and non-NSCLC regions and the distribution of Treg^{fci}. **c** UMAP visualization of spatial transcriptomic spot distribution in NSCLC tissues. Each spot represents a defined spatial location and colored by transcriptional similarity. **d** Box plot comparing Treg^{fci}-associated gene signature scores between central and peripheral margin regions of NSCLC tissues. **e**–**i** Box plots based on SpatialTME database (http://www.spatialtme.yelab.site/) showing Treg^{fci} signature scores across tumor core (Mal); tumor-normal interface (Body); and adjacent normal tissue (nMal) in breast cancer (BRCA), colorectal cancer (CRC), hepatocellular carcinoma (HCC), ovarian cancer (OV), pancreatic ductal adenocarcinoma (PDAC). **j** Histogram of GSVA based pathway enrichment using differentially expressed genes (DEGs) of Treg^{fci} subsets between NSCLC and para-NSCLC tissues. Enriched 44 pathways were identified to characterize the functional divergence of Treg^{fci} subsets by comparing tumor versus non-tumor conditions. **k**, **l** Heatmaps of intercellular communication strength between 9 pre-operative blood–specific T ciMGPs and cancer ciMGPs in NSCLC and para-NSCLC region, using CellPhoneDB ligand–receptor interaction analysis. **m**–**o** Bubble plots illustrating the mRNA expression of Treg^{fci} ciMGPs across 43 annotated T lymphocyte subsets from 428 healthy individuals (CIMA cohort), including 20 CD4$^+$ T cell subsets, 7 CD8$^+$ T cell subsets, 9 natural killer (NK) and natural killer T (NKT) cell subsets, and 4 γδ T cell subsets. Each bubble represents a distinct subset-specific expression level, with bubble size proportional to the percentage of cells via the expression of the gene of interest and color intensity indicating the average expression level. **p** Individual immunohistochemical staining of CTLA4, FOXP3, and IL2RA (CD25) in normal and NSCLC tissues (×4 and ×40 magnifications). **q** Multiplex immunohistochemical co-expression staining of CTLA4, FOXP3 and IL2RA(CD25) in NSCLC tissues (×2, ×4, ×20, and ×40 magnifications). * and ** stand for *p* values less than 0.05 and 0.01, respectively, as compared with the corresponding control groups

cells to the live cells (Supplementary Fig. 12c, d). To better reflect human immunobiology and avoid the potential influence of allogeneic responses, we developed the autologously humanized organoid model with PBMCs and lung cancer cells from a patient and expanded the organoids by biweekly passaging at 1:2 split ratios (Supplementary Fig. 12e). The cell morphology within organoids appeared the origin of lung cancer cells.

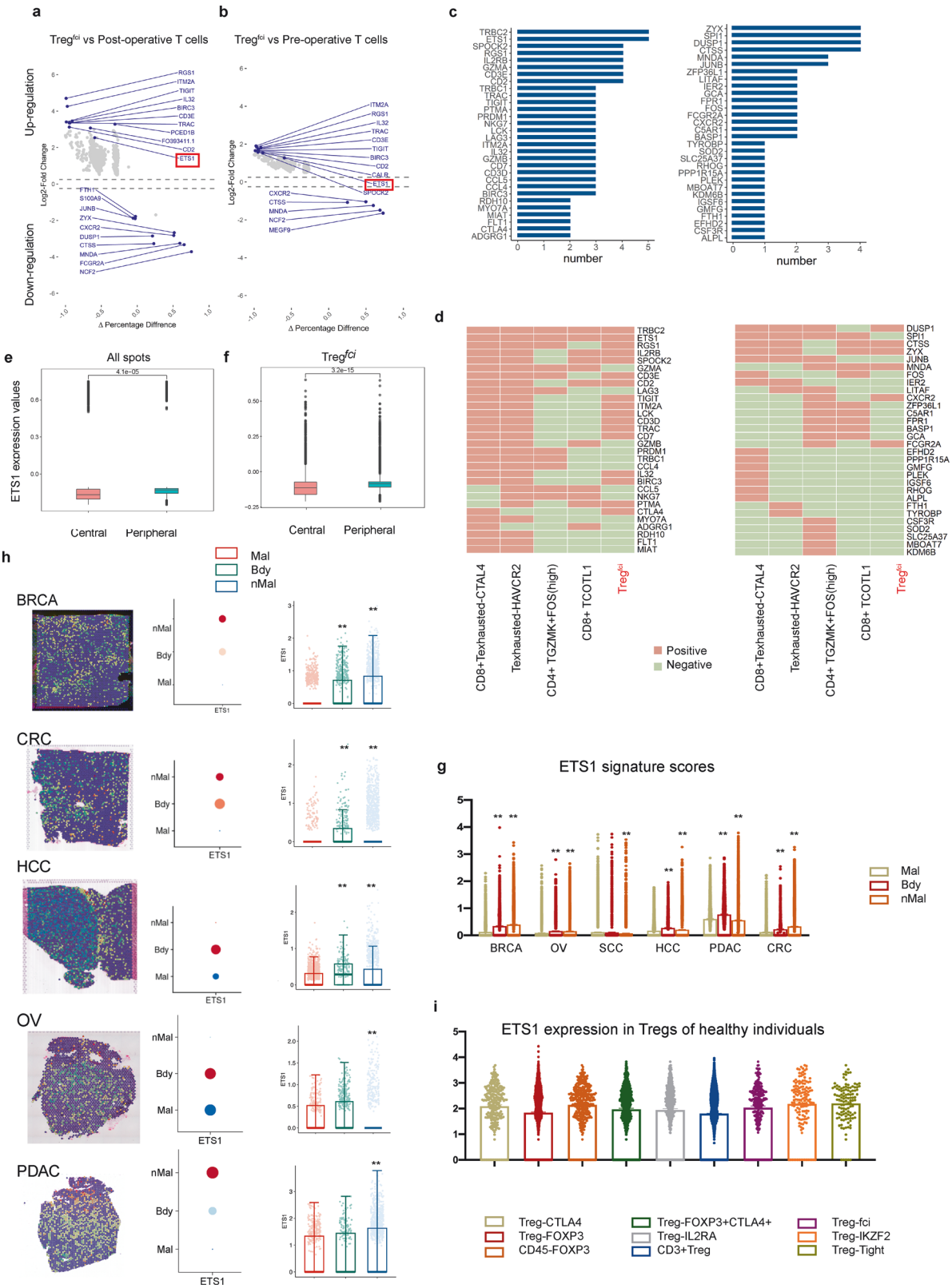
ETS1 expression significantly elevated in the peripheral tumor areas, as compared with the central areas ($p < 0.01$). Spatial analysis revealed that *ETS1* was primarily enriched in the interface zone between NSCLC and adjacent para-tumor tissues, across 6 extra-pulmonary cancers (Fig. 4g), as detailed in Fig. 4h. The *ETS1* expression in tumor/para-tumor interface and para-tumor regions increased in other cancer types, including BRCA, CRC, HCC, OV, and PDAC. In HCC tissues, *ETS1* significantly overexpressed in the interface zone, as compared to HCC and para-HCC tissues. The present study illustrated the *ETS1* expression across 16 defined Treg subsets under various physiological and pathological conditions, including healthy donors, NSCLC, and other lung diseases. In healthy peripheral blood, the *ETS1* expression was relatively higher in Treg^{CTLA4}, $_{CD45+}$Treg^{FOX3}, Treg^{IKZF2}, or Treg^{Tight}, but lower in $_{CD3+}$Treg and Treg^{FOXP3} (Fig. 4i).

Validation of Treg^{fci} biological functions

To assess the roles of each Treg^{fci} ciMGP in T cell function and interdependencies, we selectively downregulated the expression of *FOXP3*, *CTLA4*, and *IL2RA* (respectively, T$^{FOXP3\text{-}KD}$, T$^{CTLA4\text{-}KD}$, T$^{IL2RA\text{-}KD}$), or simultaneously downregulated all the ciMGPs present in T cells (T$^{fci\text{-}KD}$) using a lentiviral shRNA vector. Cellular biofunctions and cell-cell interactions were evaluated, as depicted in Fig. 5a. Our results demonstrated that the viability of T$^{FOXP3\text{-}KD}$, T$^{CTLA4\text{-}KD}$, and T$^{fci\text{-}KD}$ significantly reduced, as compared to the negative control T^{NC}. T$^{CTLA4\text{-}KD}$ (Fig. 5b) and T$^{IL2RA\text{-}KD}$ (Supplementary Fig. 4a) treated with PHA exhibited a loss of proliferative capacity by day 3 of culture. The expression levels of three genes markedly decreased in cells with either individual or combined knockdowns, especially *FOXP3* (Fig. 5c) in T$^{FOXP3\text{-}KD}$ and T$^{fci\text{-}KD}$, and *IL2RA* (Fig. 5d) and *CTLA4* (Fig. 5e) in T$^{CTLA4\text{-}KD}$ and T$^{fci\text{-}KD}$, respectively, compared to the T^{NC} control group, 48 h after pre-treatment with the vehicle ($p < 0.05$). PHA pre-treatment increased *IL2RA* expression in all T cell populations, as compared to the T^{NC}. PHA treat increased in *FOXP3* expression and decreased the expression of *IL2RA* and *CTLA4* levels in T$^{fci\text{-}KD}$.

To evaluate the regulatory roles of Treg^{fci} ciMGPs in the toxic effects of T cells, T$^{fci\text{-}NC}$ or T$^{fci\text{-}KD}$ were co-cultured with HBE or A549, with or without PHA stimulation. The killing capacities of T cells were explored across varying concentrations and time points. Residual target cell viability and morphology were visualized using crystal violet staining (Supplementary Fig. 4b). We selected the 1:10 ratio of A549 and T cells following PHA stimulation (Fig. 5f) and evaluated the impact of downregulating Treg^{fci} ciMGPs on T cell-killing capacities. Amongst those knockdown cells, T$^{fci\text{-}NC}$ exhibited the greatest sensitivity to PHA, independent of co-culture with A549 (Fig. 5f) or HBE (Fig. 5g), because there was a significant difference of viabilities between T$^{fci\text{-}NC}$ and T$^{fci\text{-}KD}$ treated with either vehicle or PHA. The viability levels of T^{fci} were significantly higher or lower when co-cultured with the 1:1 or 1:10 concentration of HBE, respectively (Fig. 5j, $p < 0.01$ vs. group without the co-culture). T$^{fci\text{-}KD}$ exacerbated the toxic effects of T$^{fci\text{-}NC}$ treated with PHA, as demonstrated by the comparison of T$^{fci\text{-}KD}$ and T$^{fci\text{-}NC}$ cell treated with vehicle, and by the reduced PHA sensitivity observed when comparing the extent of viability decline between vehicle- and PHA-treated T$^{fci\text{-}KD}$ and T$^{fci\text{-}NC}$ (Fig. 5g). However, the knockdown of individual *FOXP3* or *CTLA4* did not result in significant differences in toxicity with regard to treatment with either vehicle or PHA (Supplementary Fig. 4c). In addition, we also measured the effects of T$^{fci\text{-}NC}$ on survival of 8 cancer cell lines and found lower survival levels in HBE after vehicle stimulation and co-culture with T$^{fci\text{-}KD}$ and in A549, H460, H446, and SPC-A1 cells after PHA stimulation and co-culture with T$^{fci\text{-}KD}$, compared with T$^{fci\text{-}NC}$ with vehicle or PHA (Supplementary Fig. 13a). The survival of LTEP treated with vehicle and T$^{fci\text{-}KD}$ was significantly higher than that of T$^{fci\text{-}NC}$ with vehicle or T$^{fci\text{-}KD}$ with PHA. While the dynamical survivals of T$^{fci\text{-}KD}$ with vehicle or PHA were higher than those of T$^{fci\text{-}NC}$ (Supplementary Fig. 13b).

To investigate the regulatory roles of Treg^{fci} ciMGPs in modulating pro-inflammatory cytokines and Treg-associated effector mediators during T cell-epithelial interactions, HBE or A549 were seeded in the bottom chambers and T$^{fci\text{-}NC}$ or T$^{fci\text{-}KD}$ in the upper inserts. The levels of cytokine mRNA expression were measured 24 h after the co-culturing of cells, accompanied by treatment with the vehicle or PHA. Compared to T$^{fci\text{-}NC}$, levels of IL-6, IL-8, IL-1B, IL-4, and TNF-α mRNA expressions were significantly upregulated in PHA-treated T$^{fci\text{-}KD}$, whereas IL-10 expression was higher in T$^{fci\text{-}NC}$ (Fig. 5h). Protein levels of the Treg-associated effector cytokine IL-10 were markedly reduced in T$^{fci\text{-}KD}$, rather than IL-4 protein levels (Supplementary Fig. 6a). Those altered patterns were further modulated by co-culturing the T cells with HBE or A549. IL-6 levels were significantly elevated in vehicle-treated T$^{fci\text{-}KD}$, whereas IL-1β increased in T$^{fci\text{-}KD}$ treated with PHA. TNF-α and IL-10 were downregulated in T$^{fci\text{-}KD}$ with PHA and vehicle. Expression of TGF-β, TLR4, IL-27, and IL-35 varied under diverse conditions (Supplementary Fig. 4d). T$^{fci\text{-}NC}$ alone enhanced



the expression of IL-6, IL-8, and TNF-α in co-cultured A549 (Fig. 5i), and induced IL-8 expression in HBE (Fig. 5j), unlike $T^{fci\text{-}NC}$ treated with PHA or $T^{fci\text{-}KD}$. $T^{fci\text{-}KD}$ increased IL-6 expression and reduced TNF-α levels in HBE, which were reversed by PHA treatment. The expression of IL-6, IL-8, TNF-α and IL-1β in A549 or HBE treated with vehicle or PHA were also observed (Supplementary Fig. 14e). We also measured levels of TNFα, IL-1β, IL-6, and IL-8 proteins in the supernatants and found that TNFα levels were significantly higher in HBE co-cultured with $T^{fci\text{-}NC}$ treated with vehicle, as compared with $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ treated with PHA (Supplementary

Fig. 4 *ETS1*-associated transcriptomic landscape and spatial enrichment of Treg^{fci} subsets in NSCLC and diverse immunopathological settings. **a**, **b** Volcano plot showing DEGs identified from scRNA-seq data. **a** About 1363 DEGs with adjusted $p < 0.05$ and $\log_2FC > 1$ between pre-operative Treg^{fci} subsets and post-operative T cells. **b** About 995 DEGs between pre-operative Treg^{fci} subsets and non-Treg^{fci} T cell populations of NSCLC patients. **c** Bar plots displaying top 30 DEGs between pre-operative blood-specific T cell subsets and post-operative blood T cells of NSCLC patients. **d** Heatmap illustrating the expression profiles of the top 30 up-regulated/down-regulated DEGs between pre-operative blood-specific T cell subsets and post-operative blood T cells of NSCLC patients. Box plots comparing spatial distribution of *ETS1* (**e**) and *ETS1*+Treg^{fci} (**f**) signature scores between central and peripheral tumor regions in NSCLC, based on spatial transcriptomics. **g** Box plot showing *ETS1* exprssion across spatial spots in BRCA, CRC, HCC, OV, PDAC, SCC, analyzed by SpatialTME. **h** Spatial distribution of *ETS1* s in BRCA, CRC, HCC, OV, PDAC, and SCC was analyzed using SpatialTME. Representative spatial transcriptomic heatmaps visualize the regional expression of *ETS1* within tissue sections. Representative heatmaps illustrate spatial expression heterogeneity, where red indicates regions of high expression and blue denotes low expression. **i** Box plots showing ETS1 expression, the identification of 9 out of 16 Treg ciMGPs in healthy donors. * and ** stand for p values less than 0.05 and 0.01, respectively, as compared with the corresponding control groups

Fig. 14a, $p < 0.05$ or 0.01, respectively). Levels of IL-1b or IL-8 proteins changed in A549 and HBE co-cultured with T cells (Supplementary Fig. 14b or d), but not IL-6 (Supplementary Fig. 14c).

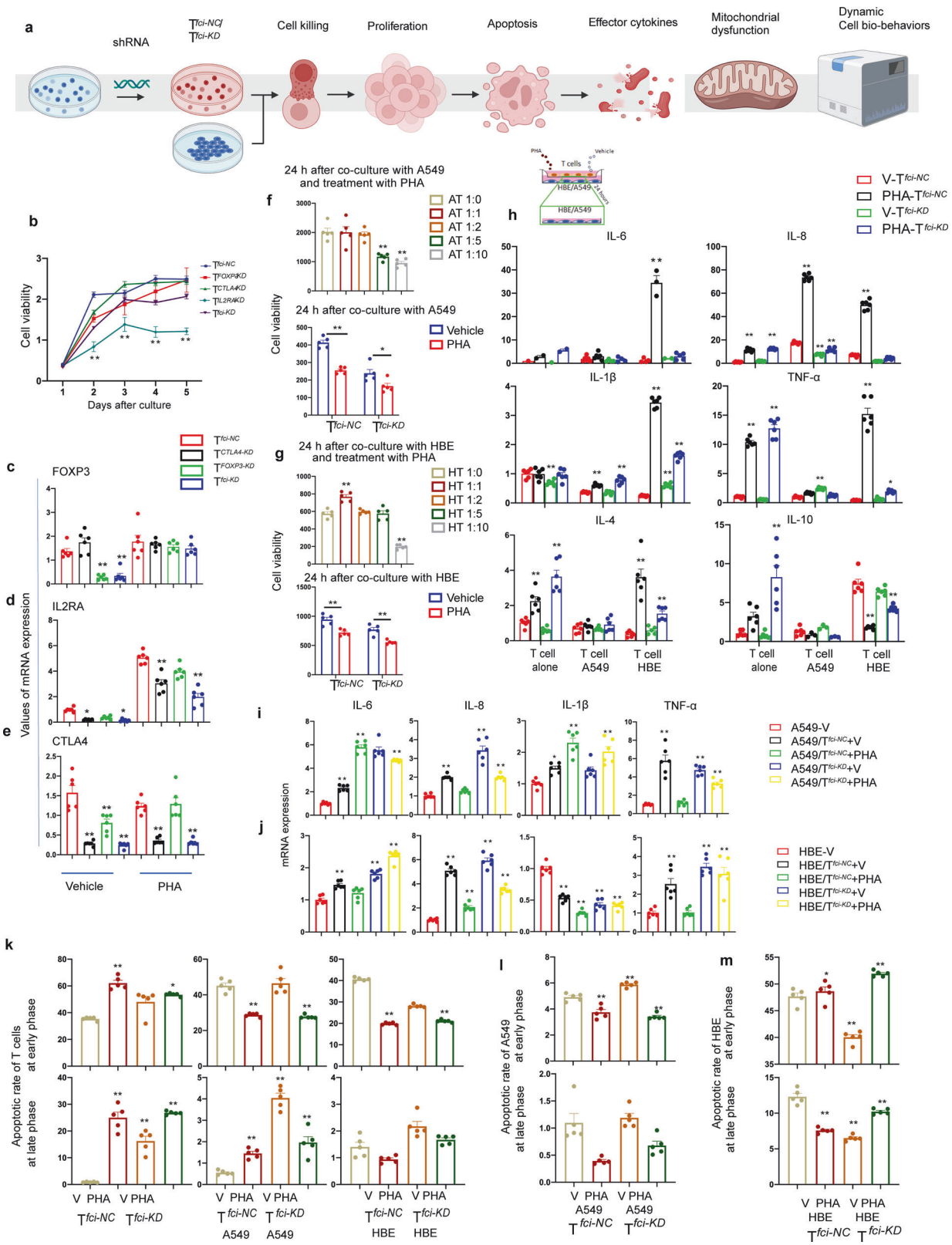
Apoptosis assays revealed that PHA stimulation significantly increased early and late apoptotic rates in T$^{fci\text{-}NC}$, as compared to those treated with vehicle. T$^{fci\text{-}KD}$ increased the baseline apoptotic rate and attenuated PHA-induced sensitivity of T$^{fci\text{-}KD}$ (Fig. 5k). Co-culturing with A549 or HBE reduced the T cell apoptosis, especially late-stage apoptosis, which declined nearly 10-fold, as compared to T cell alone (Supplementary Fig. 4e). PHA-induced apoptosis in A549/HBE co-cultured with T$^{fci\text{-}NC}$ or T$^{fci\text{-}KD}$ was minimal, except for the apoptotic rate in T$^{fci\text{-}NC}$ treated with PHA (Supplementary Fig. 4f, g). A549 reduced apoptosis rates when co-cultured with either T$^{fci\text{-}NC}$ or T$^{fci\text{-}KD}$ stimulated by PHA (Fig. 5l and Supplementary Fig. 4f). PHA-treated T$^{fci\text{-}KD}$ increased the early apoptotic rates of co-cultured HBE (Fig. 5m and Supplementary Fig. 4g). The Bax/Bcl-2 axis was shifted toward a pro-apoptotic state, indicating enhanced activation of apoptotic pathways determined by WB assays (Supplementary Fig. 6b).

To evaluate the influence of Treg^{fci} ciMGPs on T cell metabolism and cellular behavior, we assessed the mitochondrial function and bioenergetics following ciMGP knockdown and found that T$^{fci\text{-}KD}$ exhibited lower levels of ROS production, whereas co-culturing with HBE elevated ROS levels in T$^{fci\text{-}KD}$ and T$^{fci\text{-}NC}$ subjected to PHA stimulation (Supplementary Fig. 5a). The PHA treatment significantly increased intracellular calcium concentrations in T$^{fci\text{-}KD}$ with or without the co-culture with A549/HBE, whereas no comparable effect was observed in T$^{fci\text{-}NC}$ (Fig. 6a). The mitochondrial mass per cell increased in T$^{fci\text{-}NC}$ with PHA, whereas decreased in T$^{fci\text{-}KD}$ with vehicle or PHA (Fig. 6b). The OCR was comparable across HBE and A549 (Fig. 6c, d). Upon FCCP injection, T$^{fci\text{-}KD}$ with PHA exhibited the highest respiratory capacity, whereas other groups displayed relatively lower OCR levels (Fig. 6c). During the ROT/AA phase, the T$^{fci\text{-}KD}$ with PHA showed the greatest spare respiratory capacity during co-culture (Fig. 6d). The OCR-based assessments of mitochondrial function, including basal and maximal respiration, non-mitochondrial respiration, ATP turnover, and proton leak, offered additional insight into the bioenergetic reprogramming induced by ciMGP knockdown, as detailed in Supplementary Fig. 5b. T$^{fci\text{-}KD}$ had a marked reduction in maximal respiration and increased basal respiration and ATP turnover, whereas PHA activation significantly enhanced the maximal respiration capacity, basal respiration, and proton leak in T$^{fci\text{-}NC}$ (Supplementary Fig. 5b). To uncover the effects of Treg^{fci} ciMGPs in inter-organelle interactions, we assessed the morphology and ultrastructural contact sites between mitochondria and the endoplasmic reticulum (ER) and found that T$^{fci\text{-}KD}$ possessed more mitochondria-ER contacts than T$^{fci\text{-}NC}$, where PHA stimulation further enhanced these contacts in T$^{fci\text{-}NC}$ and T$^{fci\text{-}KD}$ (Fig. 6e and Table 3).

To elucidate the potential mechanisms by which Treg^{fci} ciMGPs contribute to alterations in molecular phenotypes and intra-/inter-cellular communications, we performed bulk RNA sequencing and compared the transcriptomic profiling between T$^{fci\text{-}NC}$ and T$^{fci\text{-}KD}$ treated with vehicle or PHA (Fig. 6f), between co-cultures with or without A549 (Fig. 6g), between cells with or without PHA (Fig. 6h), and between interactions of T$^{fci\text{-}NC}$ or T$^{fci\text{-}KD}$ with HBE (Fig. 6i). Treg$^{fci\text{-}KD}$ primarily altered gene expression signatures associated with key intracellular signaling pathways, including infection response, cellular senescence, the spliceosome, and RNA transport processes, as detailed in Supplementary Fig. 6c–e. Directional migration was sparsely detected in T$^{fci\text{-}NC}$, owing to the random motility. T$^{fci\text{-}KD}$ distinctly impaired the proliferative capacity, consistent with the previous CCK8 assay findings (Fig. 6j). Upon co-culturing with A549, PHA stimulation and T$^{fci\text{-}KD}$ reduced the proliferation of T$^{fci\text{-}NC}$ (Fig. 6k) and T$^{fci\text{-}KD}$ selectively affected cell phenotypes. The expression of the proliferation-associated protein CD39 decreased in T$^{fci\text{-}KD}$ compared with T$^{fci\text{-}NC}$, indicating a reduction in proliferative capacity. The activated T$^{fci\text{-}NC}$ significantly upregulated the activation marker CD69, rather than T$^{fci\text{-}KD}$, suggesting that the downregulation of T^{fci} ciMGPs impairs the activation of those cells (Supplementary Fig. 6c). The MSD increased in PHA-treated T$^{fci\text{-}NC}$ and peaked at 48 h. T$^{fci\text{-}KD}$ had higher MSD values than T$^{fci\text{-}NC}$, which was indicative of decreased kinetic stability. T$^{fci\text{-}KD}$ with PHA had the fluctuating speeds of migratory velocity, as depicted in Supplementary Fig. 5g.

To validate Treg^{fci} function in humanized immune systems, we established a human-Treg immune system in mice, implanted human lung cancer cells, and detected more than 25% human CD45-positive cells in the blood (Supplementary Fig. 12b–d). There was no difference in the in vivo tumor growth between animals treated with vehicle or TK216. To evaluate potential effects of autologous Treg^{fci} and lung cancer cells and minimize potential allogeneic responses, we established organoids with human NSCLC (Supplementary Fig. 12e) and treated them with autologous PBMCs pre-incubated with TK216, T$^{ETS1\text{-}KO}$, and T cells. The cell viability in the organoid with PBMCs pre-incubated with TK216 or T$^{ETS1\text{-}KO}$ was significantly lower than in those without treatments (Supplementary Fig. 12f). To define potential allogeneic responses, human primary NSCLC tissue cells were co-cultured with vehicle, T$^{fci\text{-}NC}$, T$^{fci\text{-}KD}$, and autologous CD4$^+$T cells from peripheral blood, after isolated and incubated for 24 h. Our results demonstrated that NSCLC cells were more independent and dispersed when co-cultured with vehicle, T$^{fci\text{-}NC}$, or autologous CD4$^+$T cells, while become more agglomerated and clustered when with T$^{fci\text{-}KD}$ (Supplementary Fig. 14f).

To investigate the roles of T^{fci}-lung cancer cell interactions in metabolism reprogramming, we measured the expression of metabolism-associated genes 24 h after mono- or co-culture between T$^{fci\text{-}NC}$ or T$^{fci\text{-}KD}$ with A549 or HBE and found a large number of metabolism-genes were up- (Table 4) or down-regulated (Supplementary Table 5) after pairwise comparisons. Of those, the aldo-keto reductase family 1 member B (FC = 52), annexin A1 (FC = 71) and 2 (FC = 41), caveolin 1 (FC = 20), and fatty acid 2-hydroxylase (FC = 17) significantly up-regulated in



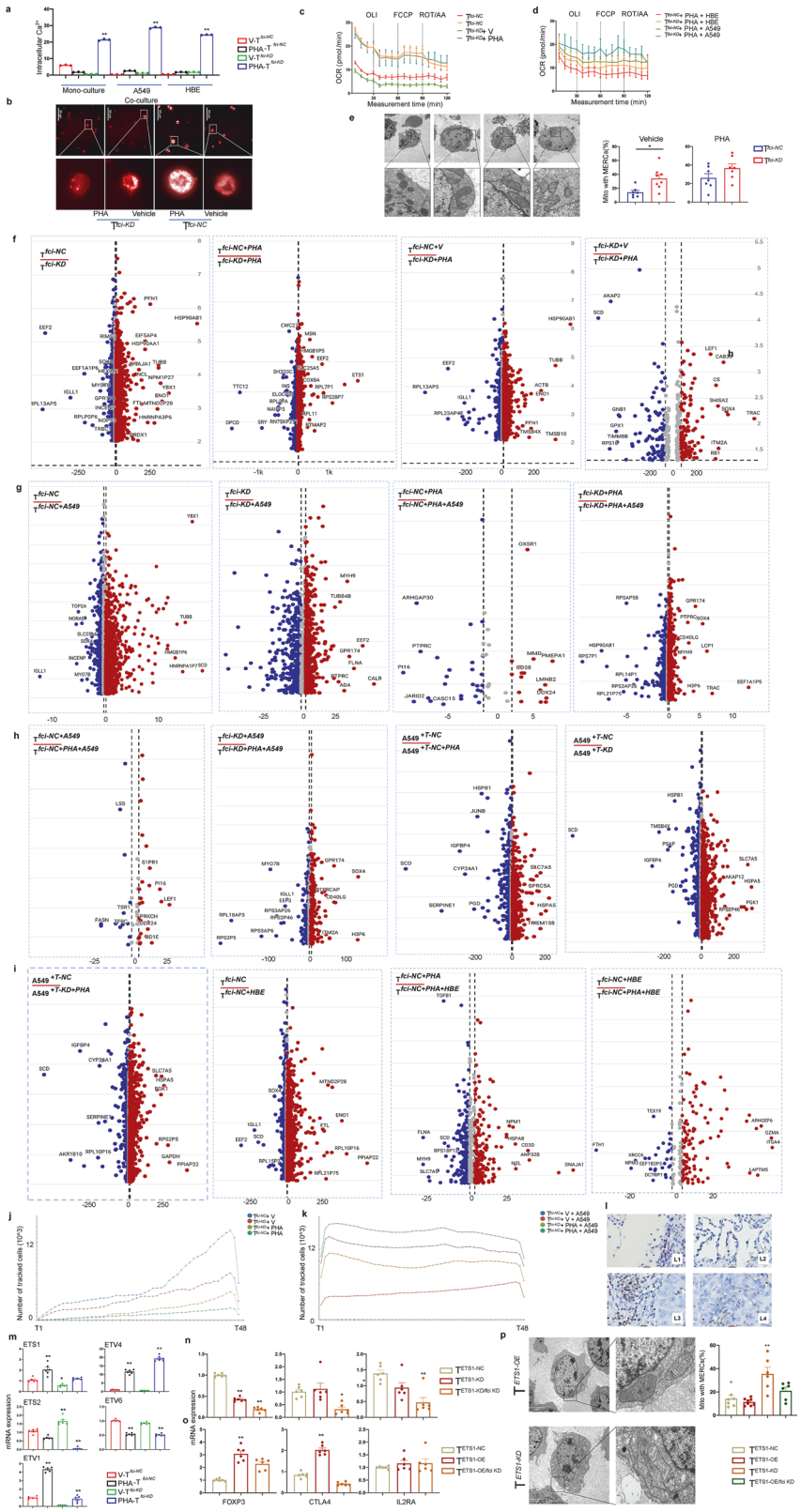
A549 co-cultured with $T^{fci\text{-}KD}$ or $T^{fci\text{-}KD}$ with PHA, compared to those with $T^{fci\text{-}NC}$ or $T^{fci\text{-}NC}$ with PHA. The $T^{fci\text{-}KD}$ metabolism influenced by A549 included SET binding factor 1 (FC = 11), protein kinase C theta (FC = 9.9), ubiquitin protein ligase E3 component N-recognin (FC = 7.4), O-linked N-acetylglucosamine transferase (FC = 6.5), and phosphoinositide 5-phosphatase (FC = 4). The altered molecular interactions and categories of metabolic processes in $T^{fci\text{-}KD}$-affected A549 (Supplementary Fig. 15a, b) were different from those in A549-affected $T^{fci\text{-}KD}$ (Supplementary Fig. 15d, e).

Fig. 5 Functional characterization of $T^{fci\text{-}KD}$ and epithelial co-culture response. **a** Workflow diagram showing generation of $T^{fci\text{-}KD}$ through lentiviral-mediated simultaneous knockdown of key ciMGPs, co-cultured with either normal human bronchial epithelial (HBE) cells or A549 cells. **b** Cell viability of T cells with knockdown of FOXP3 ($T^{FOXP3\text{-}KD}$), IL2RA ($^{TIL2RA\text{-}KD}$), CTLA4 ($T^{CTLA4\text{-}KD}$), or all three ciMGPs ($T^{fci\text{-}KD}$), and control ($T^{fci\text{-}NC}$). Quantitative real-time PCR validation of FOXP3 (**c**), CTLA4 (**d**), and IL2RA (**e**) genes in $T^{FOXP3\text{-}KD}$, $T^{CTLA4\text{-}KD}$, $T^{fci\text{-}KD}$, and $T^{fci\text{-}NC}$ treated with vehicle or PHA for 24 h. **f** Cell viability (upper) of A549 co-cultured with $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ 24 h after being treated with vehicle or PHA at 1:1, 1:2, 1:5, and 1:10 effector-to-target (E:T) ratios. Cell viability (lower) of A549 at a fixed E:T ratio of 1:10 24 h after being treated with vehicle or PHA. **g** Cell viability (upper) of HBE co-cultured with $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ 24 h after being treated with vehicle or PHA at 1:1, 1:2, 1:5, and 1:10 effector-to-target (E:T) ratios. Cell viability (lower) of HBE at a fixed E:T ratio of 1:10 24 h after being treated with vehicle or PHA. **h** $T^{fci\text{-}NC}$ and $T^{fci\text{-}KD}$ with vehicle or PHA (24 h) in the upper chambers co-cultured with A549 or HBE in the lower chambers for 24 h. T cells were harvested for qPCR analysis of inflammatory cytokines IL-6, IL-8, IL-1β, TNFα and Treg-associated cytokines IL-4 and IL-10. **i** Quantitative PCR analysis of IL-6, IL-8, IL-1β, and TNF-α, gene expression in A549 after co-cultured with $T^{fci\text{-}NC}$ and $T^{fci\text{-}KD}$, with or without PHA or vehicle. **j** Quantitative PCR analysis of IL-6, IL-8, IL-1β, and TNF-α gene expression in HBE after co-cultured with $T^{fci\text{-}NC}$ and $T^{fci\text{-}KD}$ treated with vehicle or PHA. **k** Flow cytometric analysis of apoptosis in $T^{fci\text{-}NC}$, PHA activated $T^{fci\text{-}NC}$, $T^{fci\text{-}KD}$, and PHA activated $T^{fci\text{-}KD}$ after co-cultured with A549/HBE. **l** Flow cytometric quantification of apoptosis in A549 after being co-cultured with $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ pretreated with vehicle or PHA was measured by Annexin V/PI staining. Representative dot plots and quantitative scatter plots are shown. **m** Apoptosis of HBE was assessed after co-culture with $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ treated with vehicle or PHA. Early and late apoptotic cell percentages were quantified by flow cytometry and presented as scatter plots. * and ** stand for *p* values less than 0.05 and 0.01, respectively, as compared with the corresponding control groups

To validate the interaction between $Treg^{fci}$ ciMGPs and *ETS1*, we determine *ETS1* protein location in the NSCLC tissues and found that *ETS1* predominantly expressed in lymphocytes, with nuclear localization observed at the interface between cancer cells and adjacent normal tissues (Fig. 6l). *ETS1*-positive cells primarily distributed along the boundary zones between NSCLC and para-NSCLC tissues (Fig. 6l panel 1, X40). Only a few ETS1-positive lymphocytes were observed in normal lung tissue (Fig. 6l panel 2), whereas higher densities appeared around NSCLC tissues (Fig. 6l panel 3) and occasionally infiltrated the centers and interstitial spaces in cancer cell niches (Fig. 6l panel 4). The expression of 5 *ETS1* family homologs was quantified in T cells. PHA stimulation markedly upregulated *ETS1*, *ETV1*, and *ETV4* in $T^{fci\text{-}NC}$, whereas *ETV6* and *ETS2* in $T^{fci\text{-}NC}$ and $T^{fci\text{-}KD}$, as compared to vehicle treat (Fig. 6m). The main signaling pathways between each two groups were presented in Supplementary Fig. 5e, f. Immunofluorescence staining revealed a reduced level of the ETS1 protein in $T^{fci\text{-}KD}$ (Supplementary Fig. 5h). *ETS1* and ETV4 levels were significantly lower in $T^{fci\text{-}KD}$, and became higher following PHA treatment, with ETV6 expression surpassing that of PHA-treated $T^{fci\text{-}NC}$. To illustrate the effects of *ETS1* on $Treg^{fci}$ ciMGPs, the *ETS1* levels were manipulated via overexpression and knockdown strategies in T^{fci}. It was found that *ETS1* knockdown reduced the expression of *FOXP3* and *IL2RA*, along with a further reduction in the expression of all 3 target genes when combined with $T^{fci\text{-}KD}$ (Fig. 6n). The *ETS1* overexpression enhanced the expression of *FOXP3* and *CTLA4*, whereas it slightly decreased the *IL2RA* expression (Fig. 6o). *ETS1* silencing significantly increased the number of mitochondria-ER contact sites (Fig. 6p) and inter-organelle communication. We generated a conditional knockout mouse model in which *Ets1* was specifically deleted in *Foxp3*-expressing cells ($Ets1^{flox/floxFOXP3\text{-}Cre}$), and established a murine lung cancer model through subcutaneous inoculation of Lewis lung carcinoma (LLC) cells (Supplementary Fig. 6d). To assess the effect of *Ets1* deficiency on the $Treg^{fci}$ compartment, peripheral white blood cells (WBCs), spleens, and lung tissues were collected from tumor-bearing $Ets1^{flox/floxFOXP3\text{-}Cre}$ mice and their $Ets1^{flox/flox}$ controls ($n = 6$ per group) and subjected to multiparametric flow cytometric analysis (Supplementary Fig. 6e). The frequency of $Treg^{fci}$ in WBCs and lung tissues showed no significant differences between groups ($p > 0.05$). In contrast, the proportion of $Treg^{fci}$ was significantly reduced in the spleens of $Ets1^{flox/floxFOXP3\text{-}Cre}$, indicating a tissue-specific dependency of $Treg^{fci}$ on *Ets1* expression.

To evaluate the regulatory roles of *ETS1* in transcriptomic profiles of $Treg^{fci}$, single-cell transcriptomics were measured after inhibiting *ETS1* protein function in human PBMCs using TK216 and observed about 0.3–0.4% $Treg^{fci}$ in PBMCs (Fig. 7a). About 98 genes significantly upregulated, while 109 down-regulated $Treg^{fci}$ with TK216 (Supplementary Fig. 16a) with increased changes of cell circle, IgA production, and metabolism pathways (Supplementary Fig. 16b), as compared with $Treg^{fci}$ with vehicle, of which top 20 were presented in Fig. 7b and 5 metabolism-associated genes changes in Fig. 7c. To evaluate the heterogeneity of *ETS1* roles between the circulation and lung tissues, an additional study was performed using mice with *Ets1*-cKO ($mice^{ETS1\text{-}cKO}$) or *Ets1*-wt ($mice^{ETS1\text{-}wt}$) and measured scRNA-seq of mouse blood WBCs and lung tissue on basis of the ciMGPs (Supplementary Fig. 16c and d or e and f). Mouse $Treg^{fci}$ in WBCs was 0.02–0.04% in the circulation (Fig. 7d) and 759 or 85 genes in blood significantly up- or down-regulated in $mice^{ETS1\text{-}cKO}$ with metabolic and signaling pathways (Supplementary Fig. 16e), as compared with $mice^{ETS1i\text{-}wt}$. The expression of the top 20 genes was shown in Fig. 7e and 19 metabolism-associated genes in Fig. 7f, with an increase in cardiac muscle contraction pathways and a decrease in the Jak-STAT signaling pathway (Supplementary Fig. 16d). The $Treg^{fci}$ percentage of mouse lung tissue cells was 0.26–0.28% (Fig. 7g) and 49 or 130 genes in lung tissues significantly up- or down-regulated in $mice^{ETS1\text{-}cKO}$. The panels of the top 20 genes in lung tissues of $mice^{ETS1\text{-}cKO}$ were different from those in peripheral blood (Fig. 7h) and the main changes of metabolism-associated genes were significantly lower in lung tissues of $mice^{ETS1\text{-}cKO}$ (Fig. 7i). The X-inactive specific transcript (XIST), immunoglobulin kappa constant (IGKC), or DEAD-Box Helicase 3 Y-Linked (DDX3Y) significantly up- (Fig. 7j) or down-regulated (Fig. 7k) in $Treg^{fci}$ of *ETS1*-inhibited human PBMCs or blood and lung tissues of $mice^{ETS1\text{-}cKO}$. The HLA-DRBS and PLCG2 significantly changed in human samples, while MT-ATP6 and MT-ND4 were only seen in mouse samples. To understand the potential specificity of those three factors among 17 Treg subsets, we annotated all Treg subsets on basis of the ciMGPs and analyzed the expression in human PBMCs. We found that the expression of IGKC mRNA was significantly higher in $Treg^{fci}$ and $T^{regulatory\text{-}FOXP3}$ in *ETS1*-inhibited human PBMCs (Fig. 7l), or blood (Fig. 7m) and lung tissues (Fig. 7n) of $mice^{ETS1\text{-}cKO}$. The expression of XIST or DDX3Y mRNA up- or down-regulated in most Treg subsets containing FOXP3, CTLA4, or IL2RA.

After the important roles of *Ets1*in the regulation of $Treg^{fci}$ transcriptomic and metabolism-associated genes were defined, we questioned whether the deletion of *Ets1* could influence the microenvironmental metabolism where $Treg^{fci}$ are present. To illustrate the links and roles of *Ets1*in metabolism, the circulating and lung tissue metabolomic profiles were measured in $mice^{ETS1\text{-}cKO}$ and $mice^{ETS1\text{-}wt}$. About 465 or 281 differential metabolites were identified in mouse lung tissue (Fig. 7o) and blood (Fig. 7t), of which 333 or 193 upregulated and 132 or 88 downregulated in the lung (Fig. 7p, q) or plasma (Fig. 7u, v). The contribution of variables and the influence and explanatory power of metabolic accumulation in lung tissues (Fig. 7p) and blood (Fig. 7u) were clearly distinguishing between $mice^{ETS1\text{-}cKO}$ and $mice^{ETS1\text{-}wt}$. The



heatmap of differential metabolites illustrates that alteration in metabolomic profiles in lung tissue exhibited the high consistency within the group (Fig. 7q), whereas those in plasma show greater variability (Fig. 7v). The *Z*-score values of the top 100 differential metabolites showed that the amplitude of differential metabolite changes in plasma (Fig. 7w) was higher than that in lung tissue (Fig. 7r) and differences in metabolites varied significantly between mice$^{ETS1\text{-}cKO}$ and mice$^{ETS1\text{-}wt}$. The differential metabolite-enriched pathways of lung tissue (Fig. 7s) and plasma (Fig. 7x) highly overlapped, including central carbon metabolism, TCA

Fig. 6 Transcriptomic and functional characterization of fci-modulated T cells reveal altered calcium. Signaling and mitochondrial dynamics under co-culture and activation conditions. **a** Flow cytometric analysis of intracellular calcium (Ca^{2+}) levels in $T^{fci\text{-}NC}$, PHA activated $T^{fci\text{-}NC}$, $T^{fci\text{-}KD}$, PHA activated $T^{fci\text{-}KD}$. Representative bar plot showing the intracellular calcium concentration using Fluo4 dye determined by flow cytometry. **b** Mitochondrial morphology in $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ treated with vehicle or PHA was visualized by high-content imaging. Scale bar = 50 μm. **c**, **d** Mitochondrial respiration of T cells measured using the Seahorse XF Analyzer. **c** Line graph showing the oxygen consumption rate (OCR) in $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ treated with vehicle or PHA. **d** OCR in $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ co-cultured with HBE or A549 treated with vehicle or PHA. **e** Transmission electron microscopic (TEM) analysis of mitochondria endoplasmic reticulum contact sites (MERCs) in $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ with vehicle or PHA. Scale bar = 5 μm. **f**–**i** Transcriptomic profiling of $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ with vehicle or PHA and co-cultured with HBE or A549 assessed by bulk RNA sequencing. Volcano plots illustrating the DEGs defined by a fold change >2 and $p < 0.05$ highlight upregulated and downregulated transcripts. **f** Transcriptomic changes between $T^{fci\text{-}NC}$ and $T^{fci\text{-}KD}$, PHA treated $T^{fci\text{-}NC}$ and PHA treated $T^{fci\text{-}KD}$, vehicle treated $T^{fci\text{-}NC}$ and PHA treated $T^{fci\text{-}NC}$, vehicle treated $T^{fci\text{-}KD}$ vs. PHA treated $T^{fci\text{-}KD}$. **g** Transcriptomic changes between $T^{fci\text{-}NC}$ and $T^{fci\text{-}NC}$ co-cultured with A549, $T^{fci\text{-}KD}$ and $T^{fci\text{-}KD}$ co-cultured with A549, PHA treated $T^{fci\text{-}NC}$ and PHA treated $T^{fci\text{-}NC}$ co-cultured with A549, PHA treated $T^{fci\text{-}KD}$ and PHA treated $T^{fci\text{-}KD}$ co-cultured with A549. **h** Transcriptomic changes between $T^{fci\text{-}NC}$ co-cultured with A549 and PHA treated $T^{fci\text{-}NC}$ co-cultured with A549, $T^{fci\text{-}KD}$ co-cultured with A549 and PHA treated $T^{fci\text{-}KD}$ co-cultured with A549, A549 co-cultured with $T^{fci\text{-}NC}$ and A549 co-cultured with PHA treated $T^{fci\text{-}NC}$, A549 co-cultured with $T^{fci\text{-}NC}$ and A549 co-cultured with $T^{fci\text{-}KD}$. **i** Transcriptomic changes between A549 co-cultured with $T^{fci\text{-}NC}$ and PHA treated A549 co-cultured with $T^{fci\text{-}KD}$, $T^{fci\text{-}NC}$ and $T^{fci\text{-}NC}$ co-cultured with HBE, PHA treated $T^{fci\text{-}NC}$ and PHA treated $T^{fci\text{-}NC}$ co-cultured with HBE, $T^{fci\text{-}NC}$ co-cultured with HBE and PHA treated $T^{fci\text{-}NC}$ co-cultured with HBE. **j** Quantitative analysis of proliferation of $T^{fci\text{-}NC}$, PHA activated $T^{fci\text{-}NC}$, $T^{fci\text{-}KD}$, PHA activated $T^{fci\text{-}KD}$ tracked every hour for 48 h using the high-content imaging system. **k** Quantitative analysis of proliferation of $T^{fci\text{-}NC}$, PHA activated $T^{fci\text{-}NC}$, $T^{fci\text{-}KD}$, PHA activated $T^{fci\text{-}KD}$ co-cultured with A549 tracked every hour for 48 housing. **l** Immunohistochemical staining (IHC) of human NSCLC tissue sections to assess expression levels of *ETS1*, FOXP3, CTLA4, and IL2RA. Representative images illustrate differential expression patterns within the tumor microenvironment. Scale bar = 40 μm. **m** Bar plots showing the mRNA expression of ETS family transcription factors (*ETS1*, ETS2, ETV1, ETV4, and ETV6) in $T^{fci\text{-}NC}$ and $T^{fci\text{-}KD}$, with or without vehicle or PHA. **n** Bar plots showing the mRNA expression of $Treg^{fci}$ ciMGPs (FOXP3, CTLA4, IL2RA) in $T^{fci\text{-}NC}$ and $T^{fci\text{-}KD}$ following lentiviral knockdown of *ETS1*. **o** Bar plot showing the mRNA expression of $Treg^{fci}$ ciMGPs in $T^{fci\text{-}NC}$ and $T^{fci\text{-}KD}$ following lentiviral overexpression of *ETS1*. **p** TEM was used to evaluate ultrastructural features of MERCs in *ETS1*-modulated T cells ($T^{ETS1\text{-}NC}$, $T^{ETS1\text{-}KD}$, $T^{ETS1\text{-}OE}$ and $T^{ETS1\text{-}OE/fci\text{-}KD}$). Representative images and quantitative measures of MERCs proportions are presented. Scale bar = 5 μm

cycle, or amino acid and cholesterol metabolism. The metabolic pathways and networks of significantly changed metabolites in mouse lung tissues and blood were shown in Supplementary Fig. 16g, h, respectively.

DISCUSSION

This study introduces a novel evaluation system based on the regional overlap-expression rate (rOER) of cell identity marker gene panels (ciMGPs) to address a critical limitation in single-cell technologies: the lack of quality control, precision, and repeatability in defining cell identities for clinical application. The application of single-cell measurements in clinical practice for monitoring disease progression and therapeutic effects has been primarily hindered by these challenges. For the first time, we present a conceptual framework to screen and validate the specificity and biological relevance of ciMGPs in human samples, defining a biologically acceptable ratio for each T cell subpopulation. By applying this rOER-based platform to identify disease-specific cell subtypes in non-small cell lung cancer (NSCLC), we selected and validated $Treg^{fci}$ as an independent Treg subset with clear biological and pathophysiological relevance. Although $Treg^{fci}$-representative genes (FOXP3, CTLA4, IL2RA) are also expressed in other Treg populations, the combinatorial integration of these markers, coupled with multi-dimensional validation across spatial, temporal, and disease contexts, establishes $Treg^{fci}$ as a distinct and clinically relevant subtype. Our findings underscore that ensuring ciMGP accuracy and specificity is pivotal for translating single-cell discoveries into diagnostic and therapeutic applications.

Our studies demonstrated the feasibility of performing exploratory, multi-dimensional scRNA-seq-based rOER analysis in NSCLC patients, establishing new criteria for disease-specific T cell subset selection, including >3-fold change, subset-specific marker panels, and prognostic relevance. $Treg^{fci}$ detection was highly dependent upon disease nature, location, and severity, with higher frequencies confirmed in peripheral blood and lung tissues through combinatorial marker protein labeling. Spatial distribution analysis revealed marked variation between tumor core, tumor-normal interface, and normal lung tissue. $Treg^{fci}$ exhibits unique biological features distinguishing it from the other 15 Treg subtypes defined herein and from previously published peripheral Treg subsets. Its key ciMGPs (CD45+CD4+FOXP3+CTLA4+IL2RA+) differ from those of CD4 Treg-FCRL3 and other subsets, highlighting the critical role of ciMGP composition in defining subtype identity, intercellular communication, and tumor microenvironment construction. The transcription factor FOXP3, along with IL2RA and CTLA-4, maintains immune self-tolerance and homeostasis, yet their random presence across Treg subgroups has historically made subtypes barely distinguishable. Notably, $Treg^{fci}$ demonstrated the highest enrichment of activation pathways related to inflammatory responses and cellular senescence, suggesting a potential mechanistic link between aging and NSCLC pathogenesis (Fig. 8).

Aging plays an important role in tumorigenesis and represents a risk factor for cancer development, with NSCLC closely associated with aging, inflammation, and high mortality due to progressive metabolic and T cell dysfunctions. We observed that $Treg^{fci}$ exhibits the highest number of activation pathways involved in inflammatory responses and cellular senescence, hypothesizing that $Treg^{fci}$ may mechanistically bridge aging and NSCLC. Our data demonstrate that intratumoral and peritumoral $Treg^{fci}$ aggregation was significantly lower in young NSCLC patients, while circulating $Treg^{fci}$ showed no age-related difference. This suggests that aging remodels immune compartments in a context-dependent fashion, with effects differing between blood and tissue niches. The functional significance of $Treg^{fci}$ is further underscored by its strategic spatial localization at the NSCLC-normal tissue interface and within micro-niches, where it appears to form a barrier zone regulating tumor invasion. This spatial positioning, confirmed by spatial transcriptomics and multiplex immunostaining, suggests that $Treg^{fci}$ functions as an intercellular coordinator, releasing cytokines, regulating cell cycles, and facilitating communication within the microenvironmental ecosystem. Such interface zones may act as "command posts" delivering intercellular messages or as "blindage" maintaining ecosystem homeostasis.

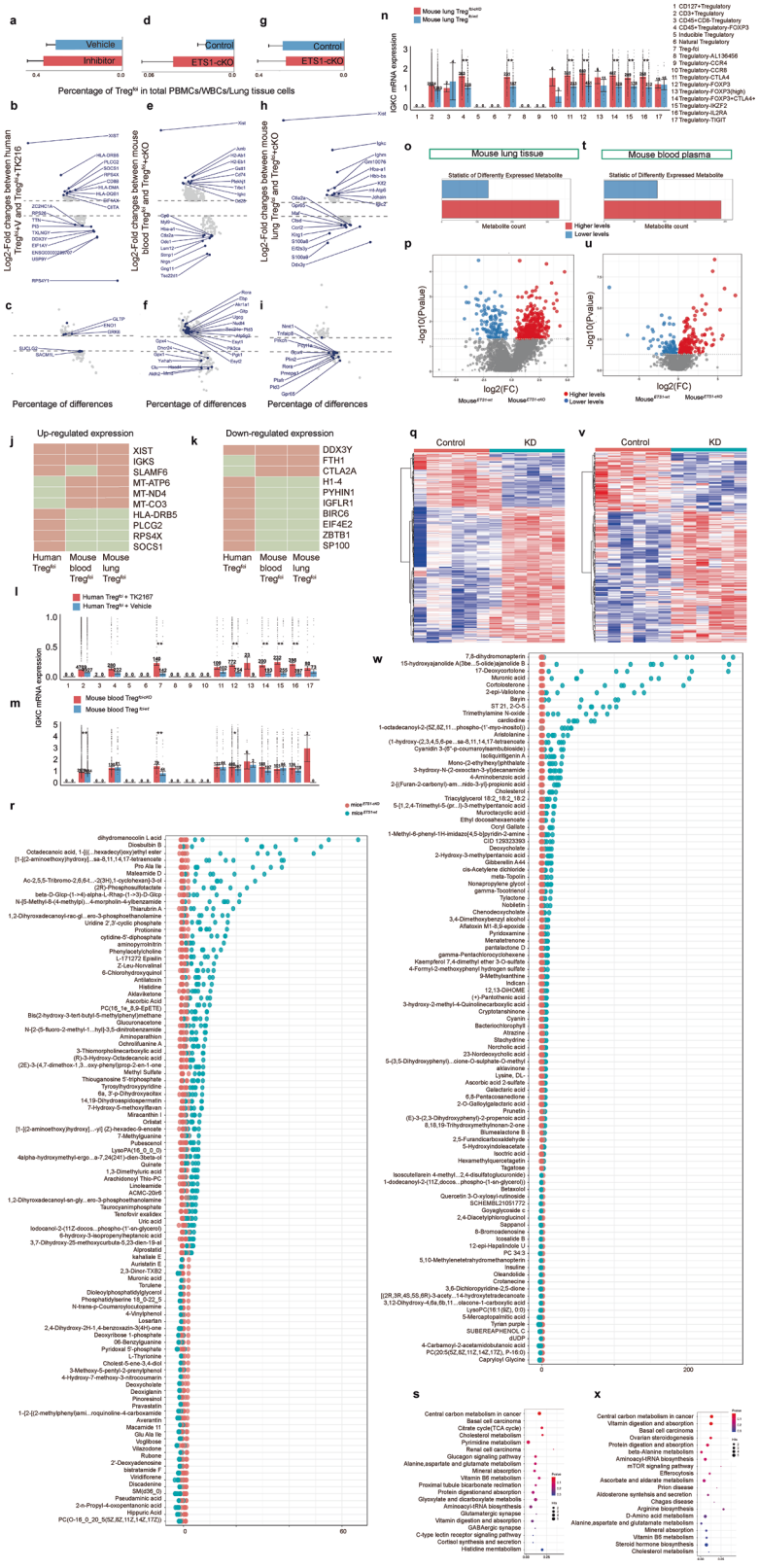
The combinatorial regulation of FOXP3, CTLA4, and IL2RA determines $Treg^{fci}$ biological functions, with synergistic interactions among these genes demonstrated through individual and combined knockdown experiments. Single-IL2RA knockdown caused high T cell death within 48 h, whereas combined $Treg^{fci\text{-}KD}$ cells survived, revealing remarkable self-regulation and compensation

Table 3. Top 20 of up-regulated lipid metabolism-associated genes ($n = 2000$) when compared between lung cancer cells A549 (A) co-cultured with T cells without sham $Treg^{fci}$ ciMGPs knockdown (T^{NC}) and A co-cultured with T cells with $Treg^{fci}$ ciMGPs knockdown (T^{KD}) ($A + T^{NC}$ vs. $A + T^{KD}$), between $A + T^{NC}$ and $A + T^{KD}$ treated with phytohemagglutinin (PHA) ($A + T^{NC}$ vs. $A + T^{KD} + PHA$), between $A + T^{NC}$ and $A + T^{NC}$ treated with PHA ($A + T^{NC}$ vs. $A + T^{NC} + PHA$), between T^{KD} and $A + T^{KD}$ co-cultured with A (T^{KD} vs. $T^{KD} + A$), between T^{KD} co-cultured with A and T^{KD} treated with PHA co-cultured with A ($T^{KD} + A$ vs. $T^{KD} + PHA + A$), between T^{KD} with PHA and T^{KD} with PHA co-cultured with A ($T^{KD} + PHA$ vs. $T^{KD} + PHA + A$), between T^{KD} with vehicle (V) and T^{KD} with PHA ($T^{KD} + V$ vs. $T^{KD} + PHA$), between T^{NC} and T^{KD} (T^{NC} vs. T^{KD}), between T^{NC} and T^{NC} with A (T^{NC} vs. $T^{NC} + A$), between T^{NC} and T^{NC} with human bronchial epithelial cells HBE (T^{NC} vs. $T^{NC} + HBE$), between T^{NC} with A and T^{NC} with PHA co-cultured with A ($T^{NC} + A$ vs. $T^{NC} + PHA + A$)), between T^{NC} with HBE and T^{NC} with PHA and HBE ($T^{NC} + HBE$ vs. $T^{NC} + PHA + HBE$), between T^{NC} with PHA and T^{KD} with PHA ($T^{NC} + PHA$ vs. $T^{KD} + PHA$), between T^{NC} with PHA and T^{NC} with PHA and A ($T^{NC} + PHA$ vs. $T^{NC} + PHA + A$), between T^{NC} with PHA and T^{NC} with PHA and HBE ($T^{NC} + PHA$ vs. $T^{NC} + PHA + HBE$), as well as between T^{NC} with V and T^{NC} with PHA ($T^{NC} + V$ vs. $T^{NC} + PHA$), respectively

$A + T^{NC}$ vs. $A + T^{KD}$	$A + T^{NC}$ vs. $A + T^{KD} + PHA$	$A + T^{NC}$ vs. $A + T^{NC} + PHA$	T^{KD} vs. $T^{KD} + A$	$T^{KD} + A$ vs. $T^{KD} + PHA + A$	$T^{KD} + PHA$ vs. $T^{KD} + PHA + A$	$T^{KD} + V$ vs. $T^{KD} + PHA$	T^{NC} vs. T^{KD}	T^{NC} vs. $T^{NC} + A$	T^{NC} vs. $T^{NC} + HBE$	$T^{NC} + A$ vs. $T^{NC} + PHA + A$	$T^{NC} + HBE$ vs. $T^{NC} + PHA + HBE$	$T^{NC} + PHA$ vs. $T^{KD} + PHA$	$T^{NC} + PHA$ vs. $T^{NC} + PHA + A$	$T^{NC} + PHA$ vs. $T^{NC} + PHA + HBE$	$T^{NC} + V$ vs. $T^{NC} + PHA$
PGK1	PGK1	PGK1	IDH2	STT3B	SERINC1	CS	ENO1	ENO1	ENO1	S1PR1	PRKCH	PSAP	PMEPA1	DNAJA1	ENO1
ENO1	ENO1	ENO1	ELOVL4	S1PR1	CPT1A	ECHS1	DNAJA1	HMGCS1	DNAJA1	PRKCH	S1PR1	ELOVL5	MMD	SUMO2	DNAJA1
LDHA	PFKP	LDHA	SBF1	PRKCH	S1PR1	GNAS	PGK1	SCD	PGK1	APOBR	CPT1A	PEBP1		CPT1A	PGK1
PFKP	PSAT1	PSAT1	PRKCQ	CPT1A	PPP1CA	BPNT1	PEBP1	ACAT2	PGAM1		ACADM	PSAT1		DBI	STT3B
PSAT1	AKR1B1	PFKP	CPT1A	RPIA	ESYT1	CPT1A	PGAM1	PGK1	PEBP1		GPR65	IDH2		SDHB	PGAM1
GPI	GPI	GPI	UBR5	ALOX5AP	PRKCH	S1PR1	THRAP3	DNAJA1	PSAT1		NMT2	THRAP3		PRKCE	PEBP1
ANXA1	ANXA1	PGAM1	ACO2	CS	ALOX5AP	NMT2	FH	IDI1	THRAP3		PRKCE	NFYA		PGM2	FH
PGAM1	PGAM1	PLIN2	ITPR3	ESYT1	SBF1	PRKCH	HSD17B10	ACLY	HSD17B10		HPGD	STARD7		VAPA	THRAP3
AKR1B1	DNAJA1	ALDOA	OGT	GPR65	PRKCQ	PRPS1	RPIA	SQLE	PRMT1		PLPP1	ARF3		ALG5	HSD17B10
ANXA2	PLIN2	HK1	S1PR1	PREX2	ALDH1A2	ESCO1	GPX1	FDPS	SUMO2		ANXA1	CSNK2B		SUCLA2	PRMT1
ALDOA	ALDOA	ALDOC	PRPS1	PIKFYVE	OGT	CSNK2A1	NDUFAB1	FASN	NDUFAB1		APOBR	SBF1		PLPP1	ACOT7
PLIN2	PTGES3	HILPDA	PLCG1	PLCH1	KAT6A	CPNE1	RAN	MSMO1	FH		ATP8A2	PMVK		RAB5A	PTGES3
GPX1	TXNRD1	ANGPTL4	PMEPA1	SOAT1	EP300	SLC25A29	SUMO2	DHCR24	RPIA		PLCH1	ATP6V0C		PLEKHA3	PRKCH
ANGPTL4	HK1	GBE1	AGPAT1	ATP8A2	UBR5	RDH14	PRMT1	PSAT1	ALDOA		SEC24B	STT3B			NDUFAB1
PYGL	ANGPTL4	SLC2A3	ACAA2	HPGD	CERK	MDH1	STARD7	PEBP1	PRPS2		DGKA	ACOT7			ARF3
CAV1	FA2H	GYS1	LPIN2	APOBR	PREX2	PIGX	ACOT7	LSS	ACOT7		INPP5B	SQLE			PPP1CA
PRMT1	ANXA2	HK2	FIG4	PRKAB1	SLC44A2	LPIN2	FABP5	HMGCR	GPI		PREX2	ARF1			SUMO2
FA2H	HILPDA	LPCAT1	PNPLA2	APOL6	GPR65		PRPS2	PGAM1	CSNK2B		PIK3CG	GPX1			GPI
PRPS1	SLC2A1	PRMT1	ACADS	CERS6	MAPKAPK2		PTGES3	STT3B	RAN			PIKFYVE			NMT2
PDIA3	CAV1	UGCG	KAT6A	H6PD	FIG4		NAA20	MVD	PTGES3			NUDT3			G6PD

Table 4. Top 20 of down-regulated lipid metabolism-associated genes ($n = 2000$) when compared between lung cancer cells A549 (A) co-cultured with T cells without sham $Treg^{fci}$ ciMGPs knockdown (T^{NC}) and A co-cultured with T cells with $Treg^{fci}$ ciMGPs knockdown (T^{KD}) ($A + T^{NC}$ vs. $A + T^{KD}$), between $A + T^{NC}$ and $A + T^{KD}$ treated with phytohemagglutinin (PHA) ($A + T^{NC}$ vs. $A + T^{KD} + PHA$), between $A + T^{NC}$ and $A + T^{NC}$ treated with PHA ($A + T^{NC}$ vs. $A + T^{NC} + PHA$), between T^{KD} and $A + T^{KD}$ co-cultured with A (T^{KD} vs. $T^{KD} + A$), between T^{KD} co-cultured with A and T^{KD} treated with PHA co-cultured with A ($T^{KD} + A$ vs. $T^{KD} + PHA + A$), between T^{KD} with PHA and T^{KD} with PHA co-cultured with A ($T^{KD} + PHA$ vs. $T^{KD} + PHA + A$), between T^{KD} with vehicle (V) and T^{KD} with PHA ($T^{KD} + V$ vs. $T^{KD} + PHA$), between T^{NC} and T^{KD} (T^{NC} vs. T^{KD}), between T^{NC} and T^{NC} with A (T^{NC} vs. $T^{NC} + A$), between T^{NC} and T^{NC} with human bronchial epithelial cells HBE (T^{NC} vs. $T^{NC} + HBE$), between T^{NC} with A and T^{NC} with PHA co-cultured with A ($T^{NC} + A$ vs. $T^{NC} + PHA + A$)), between T^{NC} with HBE and T^{NC} with PHA and HBE ($T^{NC} + HBE$ vs. $T^{NC} + PHA + HBE$), between T^{NC} with PHA and T^{KD} with PHA ($T^{NC} + PHA$ vs. $T^{KD} + PHA$), between T^{NC} with PHA and T^{NC} with PHA and A ($T^{NC} + PHA$ vs. $T^{NC} + PHA + A$), between T^{NC} with PHA and T^{NC} with PHA and HBE ($T^{NC} + PHA$ vs. $T^{NC} + PHA + HBE$), as well as between T^{NC} with V and T^{NC} with PHA ($T^{NC} + V$ vs. $T^{NC} + PHA$), respectively

$A + T^{NC}$ vs. $A + T^{KD}$	$A + T^{NC}$ vs. $A + T^{KD} + PHA$	$A + T^{NC}$ vs. $A + T^{NC} + PHA$	T^{KD} vs. $T^{KD} + A$	$T^{KD} + A$ vs. $T^{KD} + PHA + A$	$T^{KD} + PHA$ vs. $T^{KD} + PHA + A$	$T^{KD} + V$ vs. $T^{KD} + PHA$	T^{NC} vs. T^{KD}	T^{NC} vs. $T^{NC} + A$	T^{NC} vs. $T^{NC} + HBE$	$T^{NC} + A$ vs. $T^{NC} + PHA + A$	$T^{NC} + HBE$ vs. $T^{NC} + PHA + HBE$	$T^{NC} + PHA$ vs. $T^{KD} + PHA$	$T^{NC} + PHA$ vs. $T^{NC} + PHA + A$	$T^{NC} + PHA$ vs. $T^{NC} + PHA + HBE$	$T^{NC} + V$ vs. $T^{NC} + PHA$
FASN	PLAAT3	ABHD4	ACBD3	B4GALT2	LYPLA1	ELOVL1	CS	PAG1	FADS1	B4GALT2	VAPB	PRPS2	NEU3	LRP5	ELOVL4
MSMO1	PTGES	ACLY	PITPNB	CYB5R3	SDHB	CERS5	FADS1	MED18	SERINC1	VAPB	VLDLR	LIPF		PIK3CD	ITPR3
DHRS3	IDH1	CES1	ARF3	ATP8B3	NUDT3	GGPS1	SP1	PIK3AP1	PIKFYVE	LDLR	PMEPA1	LPGAT1		NCOR1	STARD4
SQLE	FADS2	ME1	LYPLA1	ABCG1	NAA20	KAT6A	GSK3B	MAPKAPK2	GSK3B	ACACA		FITM1		NMT2	SLC25A1
PLP2	DHRS3	ABCC1	DLAT	DGAT1	ARF3	RABGGTB	SERINC1	MMUT	LSS	LSS		PLA1A		NCOR2	ELOVL5
IDH1	FADS1	DHRS3	ZMPSTE24	VKORC1	STARD7	GLTP	PMEPA1	CPT1A	SREBF2	MSMO1		LTB4R2		RXRB	SPTSSA
FADS2	AKR1C1	FADS2	NAA20	PLCB3	SUMO2	DHCR24	PIKFYVE	PITPNM2	PLP2	FASN		HMGCS1		ABCC1	ACBD7
AKR1C1	PLP2	NPC2	ACAT2	ERLIN1	DLAT	ACAT2	SBF1	SERINC1	FASN	SCD		NUDT4		EP300	PITPNM2
FADS1	NPC2	FADS1	SQLE	SREBF1	MDH2	COQ3	CPT1A	PTEN	STT3B	NA		PGK1		ITPR2	SP1
PTGES	ABCC3	PTGES	NDUFAB1	CPNE7	LBR	PGM2	ACSF3	SPTSSA	MSMO1	NA		GM2A		UBR5	SREBF2
NPC2	AKR1C3	ABCC3	MSMO1	PITPNM2	RPIA	HK2	PITPNM2	ITPR3	SQLE	NA		PLCB2		PLCG1	GSK3B
ABCC3	G6PD	AKR1C1	NUS1	TKT	CSNK2B	NAA20	ITPR3	MED28	IDH2	NA		PIGP		MID1IP1	FADS1
B4GALT1	B4GALT1	B4GALT1	THRAP3	ALDH9A1	PI4K2B	PEBP1	SREBF2	ENO2	PTAR1	NA		RAN		CERK	PMEPA1
AKR1C2	PMEPA1	PMEPA1	SUMO2	ITPR3	PGAM1	KPNB1	HADHA	ACSF3	GM2A	NA		NR5A2		ITPR3	PLP2
PMEPA1	AKR1C2	NQO1	IDI1	MDH2	NUS1	SQLE	PLP2	ACBD7	FADS2	NA		BSCL2		SREBF2	FADS2
NQO1	NQO1	AKR1C2	RPIA	PRKCA	PRPS2	GPX1	PTAR1	PIKFYVE	ELOVL5	NA		ADH1C		FASN	GM2A
CYP24A1	PSAP	PSAP	PEBP1	PMEPA1	PGK1	SCD	GM2A	PRKCA	FDPS	NA		OGDH		ACLY	PTAR1
PSAP	CYP24A1	CYP24A1	PSAT1	PEBP1	PEBP1	NA	PRKCA	PLP2	PRKCA	NA		SCD		SBF1	PRKCA
SCD	SCD	SCD	SCD	ELOVL4	PSAT1	NA	PSAP	PTAR1	IDI1	NA		LDHA		SCD	INSIG1
ALDH1A1	ALDH1A1	ALDH1A1	DNAJA1	PSAT1	ENO1	NA	IDH2	GM2A	SCD	NA		ORMDL3		ARF1	SCD



capacity. PHA stimulation provoked increased expression of Treg ciMGPs and cytokines even with combinatorial gene silencing, indicating impressive compensatory mechanisms. Bulk RNA-seq analysis revealed enhanced intracellular RNA synthesis and movement in Treg$^{fci\text{-}KD}$, particularly spliceosome activity and RNA transport pathways. Metabolic profiling showed that Treg^{fci} ciMGPs strongly control and stabilize intracellular metabolism, with profiles varying according to stimulations, gene regulations, and cellular interactions.[29] Treg^{fci} -lung cancer interactions reprogrammed metabolic profiles bidirectionally: Treg^{fci} reprogrammed glycolysis,

Fig. 7 Single-cell transcriptomic and metabolomic profiles of human and mouse Treg^{fci} under conditions of either *ETS1* inhibition or deletion. The PBMCs from human healthy volunteers treated with vehicle (Vehicle) or ETS1 inhibitor (Inhibitor) for 24 h ($n = 11$/group) and mouse white blood cells and lung tissue cells of *Ets1*$^{flox/flox}$ (control) and Ets1$^{flox/floxFOXP3\text{-}Cre}$ (Ets1-cKO) conditional knockout mice were isolated and prepared for scRNA-seq. Treg^{fci} were annotated and selected on basis of Treg^{fci} ciMGPs and the differential expression of genes was analyzed. The percentages of Treg^{fci} were calculated by comparing with total human PBMCs (**a**), mouse blood WBCs (**d**), and mouse lung tissue cells (**g**). Representative changes in the top 20 genes or metabolism-associated genes from significantly different expression of genes were detected between human blood Treg^{fci} with *ETS1* inhibitor (Treg^{fci} + TK216) and Treg^{fci} with vehicle (Treg^{fci} + V) (**b** or **c**), between *Ets1* conditional knockout mice blood (mice$^{ETS1\text{-}cKO}$) and wild-type mice blood (mice$^{ETS1\text{-}wt}$) (**e** or **f**), and between Ets1 conditional knockout mice lung (mice$^{ETS1\text{-}cKO}$) and wild-type mice lung (mice$^{ETS1\text{-}wt}$) (**h** or **i**). The up- (**j**) or down-(**k**)regulated gene expression profiles were presented. The expression of IGKC mRNA among 17 Treg subsets was compared between human blood Treg^{fci}+TK216 and Treg^{fci} + V (**l**), between mice$^{ETS1\text{-}cKO}$ and mice$^{ETS1\text{-}wt}$ blood (**m**), and between mice$^{ETS1\text{-}cKO}$ and mice$^{ETS1\text{-}wt}$ lung (**n**). * and ** stand for the *p* values less than 0.05 and 0.01, respectively, when compared with mice$^{ETS1\text{-}wt}$. The number (**o**, **t**), the distribution volcano plots (**p**, **u**), the individual levels (**q**, **v**), *Z*-scores (**r**, **w**), and KEGG pathway enrichment analysis bubble plots (**s**, **x**) of significantly different metabolite levels in mouse lung tissues (**o**, **p**, **q**, **r**, **s**) or blood plasma (**t**, **u**, **u**, **w**, **x**) were compared between mice$^{ETS1\text{-}cKO}$ and mice$^{ETS1\text{-}wt}$

lipid raft organization, and nutrient acquisition in cancer cells through ANXA1/ANXA2 pathways, while cancer cells remodeled Treg^{fci} development, proliferation, and differentiation through PKC-θ-dominated metabolic pathways and NFκB signaling.[30,31]

ETS1 emerges as a central regulator of Treg^{fci} development and function, consistently upregulated in Treg^{fci} and confirmed through knockdown and knockout approaches. ETS1 facilitates Treg^{fci} migration from circulation into tumor tissues via chemoattraction by cancer cell-derived inflammatory mediators, and modulates intracellular metabolism and mitochondrial function through interactions with ciMGPs and organelles. ETS1 shapes Treg^{fci} lineage stability, metabolic programming, and immunosuppressive activity by enhancing FOXP3 transcription and regulating key metabolic genes including PGC1-α, HK2, and PKM2.[32–34] In conditional Treg ETS1-knockout mouse models, implanted NSCLC tumors exhibited slower growth and reduced Treg levels due to dysfunctional Treg-specific ETS1 transcriptional networks. ETS1 regulates transcriptomic and metabolomic profiles of Treg^{fci}, influencing XIST/DDX3Y balance and IGKC expression, which may mediate interactions with cancer cells. ETS1 also acts as a metabolic shift, upregulating glycolytic enzymes and transporters (GLUT1, ENO1, PEKP) and driving extensive metabolic reprogramming.[35,36]

Treg^{fci} holds significant clinical promise as a diagnostic biomarker and therapeutic target for NSCLC. Its high disease specificity, detectability in preoperative blood and tumor tissues, and marked decline following surgical resection underscore its dependence on tumor presence and utility for dynamic monitoring. Treg^{fci} exhibits high NSCLC-specificity and is readily detectable using flow cytometry and immunostaining, enhancing clinical feasibility (Fig. 8). However, challenges remain for clinical translation, including validation in large-scale cohorts, standardization of detection assays, and deeper mechanistic understanding of Treg^{fci} interactions.[37,38] The rapidly growing number of Treg subsets identified by scRNA-seq creates nomenclature confusion, with few studies validating subset functions and metabolic profiles. Future studies should include more matched control groups considering smoking, aging, and infection history. Despite implementing standardization strategies, limitations remain regarding individual phenome information, sample variations, and methodological differences. Nevertheless, this study provides an alternative approach for selecting and validating disease-specific cell subtypes, with Treg^{fci} and ETS1 representing key molecular signatures for understanding Treg-mediated immune regulation and developing novel therapeutic strategies in lung cancer.

MATERIALS AND METHODS

Human scRNA-seq and Stereo-seq

Study populations. A total of 45 healthy adult volunteers were recruited from the Health Examination Center of the Qingpu blood bank, between May 2023 and January 2024. All participants underwent physical examinations and routine laboratory testing to confirm the absence of acute or chronic illnesses. Inclusion criteria included: (1) age between 20 and 80 years; (2) no history of cardiovascular, respiratory, autoimmune, metabolic, or malignant diseases; (3) no current infections or recent vaccinations within the past 4 weeks; and (4) no use of immunosuppressive or anti-inflammatory medications. Exclusion criteria included: (1) any self-reported or clinically diagnosed disease; (2) pregnancy or lactation; or (3) failure to provide written informed consent. Peripheral blood samples (5–10 mL) were collected from each participant via venipuncture into EDTA-anticoagulated tubes and processed within 2 h for downstream single-cell or immunological assays.

Approximately 168 patients with NSCLC, including large cell lung cancer, lung adenocarcinoma (LUAD), and lung squamous cell carcinomas, were included in the present study from primary cohorts belonging to the Fudan University Zhongshan Hospital, the Shanghai Chest Hospital of Shanghai Jiaotong University, and the Fujian Medical University Zhangzhou Hospital, between October 2020 and July 2025. Clinical phenomic data were collected from patients and healthy volunteers and were scored according to the principles of the Digital Evaluation Scoring System (DESS), as reported previously.[39] The details of the DESS calculations and scores were described in Section 1 of the Supplementary Materials. Diagnoses of the disease were based on the 2015 WHO classification and further confirmed via histological examination.[40] The patients with NSCLC enrolled in the study were between 47 and 78 years old and were classified under IA–IIIB as per the TNM classification. The present study was approved by the Ethics Committee of Zhongshan Hospital, which is affiliated to Fudan University (Approval No. B2021-265). Written informed consent was obtained from all participants prior to the enrollment in the study. Of the total patient cohort, 168 were enrolled for analyses of clinical phenomic characters, 12 for circulating immune scRNA-seq, 4 for lung cancer tissue spatial transcriptomics, 76 for the isolation and flow cytometry measurements of Treg^{fci}, and 21 for immunohistological staining of Treg^{fci}.

Clinical phenomic characters were analyzed in patients from multiple clinical centers, of whom 48 samples from 12 NSCLC patients were collected for determining single-cell characteristics and spatial distributions, 76 samples from 124 (2 blood samples/patient) for validating the heterogeneity and subtype-specific features of Treg, and 21 samples of 32 lung cancer tissues for analyzing multiplex immunohistochemistry. All patients were enrolled between October 2020 and July 2025 into the primary cohorts, as detailed in Supplementary Table 19.

Blood and tissue preparations for scRNA-seq. Circulating blood cells were prepared and harvested from the healthy volunteers or NSCLC patients. Pre- and post-operative blood samples (5 mL each) were drawn into EDTA Vacutainer tubes (BD Biosciences, San Diego, CA, USA) within 3 h of hospital admission or in the morning prior to discharge (typically on post-operative day 5–7).

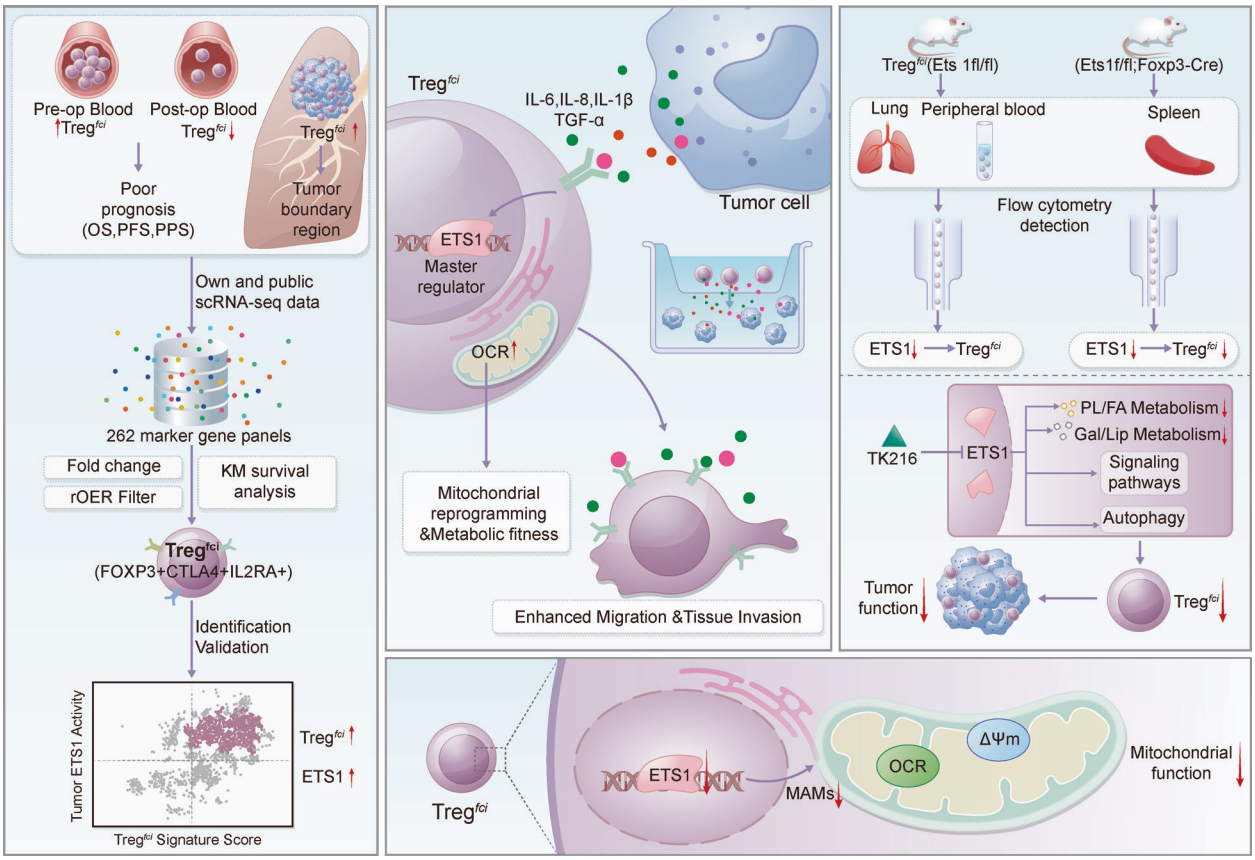


Fig. 8 A multi-layered strategy to dissect the role of regulatory T cells (Tregs) in non-small cell lung cancer (NSCLC) and the mechanistic contribution of ETS1. Data mining from public repositories enabled the systematic characterization of T cell subsets including Tregs, using marker-based profiling. The analyses of NSCLC-related datasets across six computational platforms revealed Treg enrichment, regional overexpression rates (rOER), and prognostic associations in diverse patient cohorts. Cross-tissue and cross-disease comparisons further established the landscape of Treg distribution in 6 pulmonary diseases, 17 tissues, and 31 disease contexts. Integrative approaches combining differentially expressed gene (DEG) screening, gene set variation analysis (GSVA), and cell communication analysis highlighted ETS1 as a candidate regulator of Treg biology. The relevance of Tregs across cancers was confirmed through flow cytometric phenotyping, while functional experiments demonstrated their impact on cytokine secretion, proliferation, and survival. Co-culture assays using NSCLC cells, organoids, and primary cells provided evidence that Tregs modulate tumor cell autophagy. ETS1 expression and activity within Tregs were further dissected using wild-type and knockout models, uncovering ETS1-dependent transcriptional programs. Gene interaction network analysis delineated ETS1-associated pathways, supporting its role as a central regulator of Treg function in NSCLC. This summarizes an integrated workflow spanning bioinformatics, cross-disease profiling, functional validation, and genetic perturbation, converging on ETS1 as a pivotal determinant of Treg-mediated immune regulation in NSCLC. Created with MedPeer (medpeer.cn)

Within 15 min of the collection, the blood samples were centrifuged at 400 × *g* for 15 min at room temperature. The interface layer (buffy coat) was carefully aspirated with a pipette and added to 10 mL of 1× Red Blood Cell Lysis Buffer (eBioscience, Catalog: 00-4333-57). The solution was centrifuged at 400 × *g* for 10 min, washed thrice, and re-collected for scRNA-seq.

The NSCLC and para-NSCLC tissue samples (>2 cm from tumor margin) were harvested during the resection of lung cancer tissues from the patients. The tissues were finely chopped and incubated in 1 mg/mL collagenase IV (C5138, Sigma, St. Louis, MO, USA) and 50 μg/mL DNase I (Sigma) for 45 min at 37 °C with agitation. The digested fragments were filtered through a 70 μm cell strainer (FAL352350, Corning, NY, USA) and washed with cold flow cytometry buffer (No. 554656, BD Biosciences). The cells were resuspended in Red Blood Cell Lysis Buffer (eBioscience) for 4 min at room temperature, accompanied by gentle intermittent agitation. The cell viability was assessed by trypan blue staining. Live single cells accounted for 90% of the whole population. An appropriate volume of single-cell suspension was calculated to approximately 10,000 live cells per sample.

Tissue sectioning and processing for Stereo-Seq. Tissue sampling, harvesting, and preparation for scRNA-seq and Stereo-seq were performed in accordance with previous reports.[41,42] Human samples, including NSCLC, para-NSCLC tissues (>2 cm from cancer edge) were collected intraoperatively in the operation rooms. Each sample was divided for pathology and spatial transcriptomics, after the macroscopic characters of the lung tissue samples were imaged and recorded immediately following the surgery. Tissues were immediately stored in pre-chilled physiological saline on ice, and then transferred to the laboratory on ice within 30 min. Upon arrival, the tissues were directly immersed within a solution of optimal cutting temperature compound in isopentane cooled with liquid nitrogen. Sample quality control and spatial sequencing were performed according to the reported protocol.[21] Total RNAs were extracted from tissues using the RNeasy Mini Kit (QIAGEN, Germany), and the RNA integrity number (RIN) was analyzed using the Agilent 2100 Bioanalyzer (Agilent Technologies, USA). The STOmics Gene Expression Set-S1 kit (BGI, 1000028493, Shenzhen, China) was utilized as per the standard protocol V1.1. The tissues were retrieved from −80 °C storage and equilibrated in a freezing

microtome for 30 min, and cryo-sectioned at 10 μm thickness in a cryostat (Leica, CM1950, Mikrosysteme Vertrieb GmbH, Wetzlar, Germany). The tissue sections were attached to the probe surface of the Stereo-seq chip, incubated at 37 °C for 3 min, and fixed in methanol at −20 °C for 30 min. After drying, the chips were permeabilized at 37 °C for 12 min using 0.1% PR Enzyme (BGI, Shenzhen, China) at pH 2.0 for RNA capturing. Tissue sections were stained with hematoxylin and eosin using a commercial kit (Solarbio, G1120, Beijing, China) for histological interpretation, and then imaged using a Motic microscope (Motic Co., Xiamen, China).

Spatially resolved transcriptomic analysis. Stereo-seq raw data were processed using the SAW pipeline to generate spot-level gene expression matrices and spatial coordinate files in the standard GEM format. These matrices were converted into a 10x Genomics-compatible feature-barcode matrix (MTX format) together with corresponding spatial position files using the Bin-to-MTX conversion script provided in the Stereo-seq analysis toolkit. Loupe Browser 6.0 was then used to visualize spatial gene expression patterns, perform graph-based clustering, implement dimensionality reduction (UMAP), and display transcriptional signatures across tissue regions. Differentially expressed genes (DEGs) across spatial clusters were examined within Loupe Browser. The graph-based clustering analysis was performed to identify clusters of spots within expression profiles. The spatial expression patterns of the curated set of 262 ciMGPs were re-analyzed and visualized using the same converted MTX-format Stereo-seq dataset. The scRNA-seq data generated from the NSCLC patient samples in this study have been deposited in the Genome Sequence Archive (GSA) at the National Genomics Data Center, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences. The datasets are publicly accessible under the accession number HRA011677 (https://ngdc.cncb.ac.cn/gsa-human/).

cDNA library construction and sequencing. The single-cell suspensions (from peripheral blood or dissociated tissues) were immediately processed for single-cell capture and library construction on the BD Rhapsody system (BD Biosciences) within 30 min. Single-cell suspensions were randomly distributed into approximately 200,000 microwells. Beads carrying unique molecular identifiers (UMI) and cell barcodes were loaded close to saturation. The polyadenylated RNA molecules were hybridized to the beads following exposure to cell lysis buffer. The beads were retrieved and reverse-transcribed. Each cDNA molecule incorporated a UMI and a cell-specific barcode, positioned at the 5'-end of the oligo-dT primer, enabling accurate quantification and cell-origin assignment. To generate whole transcriptome libraries, second-strand cDNA analysis, adapter ligation, and universal amplification were performed on the Rhapsody beads. The sequencing libraries were completed by amplifying the whole transcriptome and enriching the 3'-end of the transcripts, linked with the cell barcode and UMI using random priming PCR. Libraries were generated according to the BD Rhapsody single-cell whole transcriptome amplification (WTA) protocol. The microwell-based BD Rhapsody scRNA-seq platform was selected. Library concentration and size distribution were assessed using a Bioanalyzer 2100(Agilent), TapeStation 4200(Agilent) via a Qubit High Sensitivity DNA Assay (Thermo Fisher Scientific, Waltham, MA, USA).

Sequencing was performed on the NovaSeq 6000 System platform (Illumina, San Diego, CA, USA) with the S1 Reagent Kit v1.5 (200 cycles, 8 bp index read; Illumina), targeting a depth of approximately 50,000 reads per cell. This yielded at least 40 million paired reads per pool in a Paired-End-75 run. Raw scRNA-seq data were pre-processed using the BD Rhapsody Whole Transcriptome Assay Analysis Pipeline (v2.0). For each sample, gene-barcode matrices were generated by counting the UMIs, removing non-cell associated barcodes, and compiling barcoded cells and gene expression counts. This output was imported into the Seurat (v3.0.2) for quality control and further analysis of the scRNA-seq data. Cells with fewer than 200 or more than 6000 detected genes were excluded, and genes detected in fewer than three cells were filtered out.

Dimensionality reduction and clustering analysis of scRNA-seq data. Raw UMI counts were log-normalized prior to dimensionality reduction and clustering. The top 5000 most variable genes were identified using the "FindVariableFeatures" function in Seurat package (default parameters). Principal component analysis (PCA) was performed, and the first 50 principal components were selected for downstream analysis. Two-dimensional visualization was carried out using the t-distributed stochastic neighbor embedding (t-SNE) algorithm. Cell clustering was determined using the Leiden algorithm, which utilizes a shared nearest neighbor modularity optimization approach to identify transcriptionally distinct cell populations.

DEG analysis. DEGs among different Ts/s and groups, between pre- vs. post-operation blood, and between cancer vs. para-cancer lung tissue were analyzed using the "FindMarkers" function of the Seurat R package with the Student's *t* test. T cell subtype/cluster/state-specific and disease-related DEG datasets were established on the basis of the adjusted p value < 0.05, avg|Log2FC| > 2 after comparisons. The "up-regulated" or "down-regulated" DEGs in specific T cell subtypes/states were defined as the DEGs that increased or decreased significantly in comparison to the corresponding groups.

Functional annotation of gene set variation analysis (GSVA). To explore the potential biological functions and pathway activity of each T cell subtype/state phenotype, GSVA analysis was conducted by implementing the "gsva" function from the R.package GSVA (version 1.44.2) with method specification as "ssgsea". The gene sets, including the Hallmark and KEGG gene sets, were derived from the Molecular Signatures Database (MSigDB) and WikiPathways. GSVA enrichment scores were calculated for per T cell subtype/state across pathways. For visualization and comparative analysis, GSVA *Z*-scores for each pathway were calculated using the following formula: (score of a pathway in a given cell subset – average score of that pathway across all cell subsets) / standard deviation of that pathway across all subsets.

Development of rOER for validation of disease-specific Ts/s

Primary screening and standardization of T cell ciMGPs. To assess the accuracy and specificity of selected T cell ciMGPs in lung cancer, existing scRNA-seq papers and databases were examined and used to establish 11 scRNA-seq datasets from various species, tissues/organs and diseases,[22–28] as shown in Section 3 of the Supplementary Materials. The names used, ciMGPs, and corresponding sources were listed in Supplementary Table 1 of which 170–262 human ciMGPs were selected for further investigation in the concept study. To clarify the diversities and complexities of T cell ciMGPs and standardize the labeling in the current studies, we recoined the ciMGPs names of selected Ts/s according to the cluster of differentiation (CD), cell type (T), and specific marker gene names, as detailed in Supplementary Table 1. The gene name in the recoined subtype/state system was unique to each panel, or comprised a specific group of genes if the ciMGPs contained multiple genes. The ciMGPs were also discussed on the basis of functional Ts/s—including activated, regulatory, memory, effector memory, naïve, cytotoxic, exhausted, or helper clusters. The ciMGPs were labeled in the form of the CD molecules written as the left superscript, "T" as the cell type in the middle, and the italicized specific gene names of non-CD molecules as the right subscript, e.g., $^{CD4+}T_{naive\text{-}CCR7}$, $^{CD8+}T_{GZMK+GZMA+}$, or $^{CD45RA\text{-}CD45RO-}T_{em}$. The number of

ciMGP-based Ts/s was calculated in an automatic program using R Studio 4.2.1.

Calculation of ciMGP RNA expression. The level of ciMGP RNA expression was calculated as the sum of each panel gene expression. Data distribution was visualized using boxplots, followed by comparison of the characteristics of data obtained from various categories, to provide a rapid overview of the average value, data arrangement, and data distribution. The median values, interquartile ranges and whiskers, potential outliers, and skewness signs of the boxplots were interpreted as adjusted comparison criteria.[43] Mathematical expressions, along with the corresponding codes to integrate them in the R function, were mentioned in Section 4 of the Supplementary Materials. All boxplots were generated using the "boxplot" function in R package (R4.1.2), as detailed in Section 5 of the Supplementary Materials.

Based on the principle underlying Box and Whiskers plots,[43–45] the median and distribution of the dataset were described by the lower and upper hinges of ciMGP expression, including the first and third quartile (25th and 75th percentiles). The quartiles were composited of the median, upper (Q3), lower fourths (Q1), lower whisker (Q1 − 1.5 × the inter-quartile range), upper whisker (Q3 + 1.5 × the inter-quartile range), and outliers with individual points, as reported previously.[43,46,47] The data were divided into various sections in the boxplots, which contained approximately 25% of the data represented by points within the inner limit of the outliers. The outliers between the inner and outer limits were defined as mild outliers, whereas the outlier points outside the outer limit were considered as the extreme outliers.[47]

Calculation of rOER. rOER stands for a quantitative metric to evaluate the expression specificity and reproducibility of ciMGPs for T cell subtype. It was based on a regional statistical framework derived from the Box-and-Whisker's principle, with interval-based repeatable ranges, and was systematically optimized through multi-level tests of sensitivity, specificity, and reproducibility. Repeatable ranges stand for the interval of percentage values within which a cell ciMGP is consistently detected across samples. High rOER indicates a higher regional overlap-expression rate, meaning that the ciMGPs showed less subtype-specific expression across samples, lower specificity, and high cross-expression among subtypes. The rOER based on the principle of the Box and Whiskers plot was explained in detail in Section 6 of the Supplementary Materials, along with the rOER values < 10, 20, 30, 40, and 50%, and >60%. According to the concept test involving the scRNA-seq data of peripheral blood derived from healthy subjects ($n = 31$), 10% interval repeatable ranges were divided up to 100%. The number of ciMGPs at each interval exhibited a normal distribution, where the peak value appeared at approximately 30%. The interval >60% repeatable range contained few ciMGPs. The number of ciMGP-based Ts/s and the respective rOER values were analyzed via an automatic program. The optimal sensitivity, specificity, and reproducibility of ciMGP selection were evaluated using the regional line test, classification test (binary, ternary, and quintuple classification tests), and threshold test, during rOER establishment. Details of those test methods were presented in Section 7 of the Supplementary Materials and the workflow for multiple evaluations were depicted in Supplementary Fig. 7a, b. To define threshold choices, the sensitivity, false positive rate, reproducibility, and consistency of the obtained data were re-calculated and compared to those as follows: binary classification Number30%, binary classification-Number60%, ternary classification-ss ciMGP, and quintuple classification-level 1 (Section 7 of the Supplementary Materials and Supplementary Table 15). The formulae for the rOER calculations were presented below:

rOER = overlapped ciMGPs numbers in repeatable range/total ciMGPs numbers × 100% [e.g., rOER < 10% = the Q3 of the number of panels in the <10% repeatable range (number < 10%)/ the total cell type numbers × 100%, rOER > 60% = the Q3 of the number of panels in the >60% repeatable range (number > 60%)/ the total cell type numbers × 100%, and so on].

Determination of cell-cell interactions. Cell-cell interactions were assessed by evaluating the expression of immune-related ligands and the corresponding receptors. Potential ligand-receptor interactions between two cell types or subtypes were identified using CellphoneDB (version 4.0.0) and the receptor-ligand database (version 4.1.0; https://github.com/Teichlab/cellphonedb). CellphoneDB was run with the default parameters and returned 2923 interaction pairs of well-annotated ligands and receptors in total, including co-stimulators, chemokines, and cytokines. The possible interactions between the two cell types and subtypes were orchestrated based on the receptor-ligand pairs and average expression levels of the ligand in one Ts/s, along with the respective receptor in the other Ts/s. The rank of each ligand-receptor pair was calculated by dividing CellphoneDB total number of significant p values by the number of cluster-cluster comparisons.

Validation of selected Ts/s specificity

Validation of Ts/s categories based on ciMGPs and rOER. To assess the robustness of the concept study, the number of ciMGPs was refreshed up to 262 Ts/s to validate the specificity and repeatability of ciMGPs and rOER, and define the diversity of Ts/s changes. All 262 Ts/s based on ciMGPs were categorized into 3 types: subset-specific (ss-ciMGPs), subset-associated (sa-ciMGPs), and subset-reference (sr-ciMGPs). The prioritization strategies used to define the prioritization order were slightly modified in the validation study. When comparing two Ts/s with the same level of specificity, the subtype with the lesser $Number_{>30\%}$ or $Number_{>60\%}$ value was given higher priority. The total overlap number in the 100% repeatable range was compared as described in the concept study, and the Ts/s with $Number_{>30\%}$ or $Number_{>60\%}$ values less than the total overlapped number were given higher priority. To determine an appropriate fold change (FC) threshold for defining Ts/s-associated ciMGPs and ss-ciMGPs, the number of differentially expressed ciMGPs was systematically re-evaluated across a range of FC cutoffs between NSCLC and para-NSCLC and between pre- and post-operative peripheral blood samples. The ciMGP selection was examined at FC thresholds of >2, >3, >5, >10, >15, and >20, as detailed in Section 7 of the Supplementary Materials. An FC > 3 was selected as a biologically meaningful and statistically balanced threshold for downstream functional annotation and overlap analysis, as seen in Supplementary Table 16. The potential impacts of different threshold choices on the specificity and performance were evaluated by comparing the influences of various thresholds in the sensitivity, specificity, and reproducibility, as described in Section 7 of the Supplementary Materials.

Validation of Ts/s-associated survival rates. Survival curves of T cell subtypes were estimated by the Kaplan–Meier method, with statistical significance calculated by the log-rank test. To evaluate patient outcomes, overall survival (OS), post-progression survival (PPS), and progression-free survival (PFS) were generated using the online KMplot tool (http://kmplot.com/analysis/). The average expression level of the constituent genes of each ciMGP was calculated for each patient sample to represent the subtype abundance. For each T cell subtype, patients were stratified into high- and low- expression groups based on median ciMGPs levels, and Kaplan–Meier curves were plotted accordingly. Hazard ratios with 95% confidence intervals were calculated and the

corresponding log-rank *p* values were reported to assess the prognostic significance of each subtype.

Validation of disease- and organ/tissue-specificity. The specificity of lung cancer was evaluated by comparing with COPD as the representative of chronic local inflammation, idiopathic pulmonary fibrosis (IPF) as the representative of chronic tissue remodeling, and interstitial lung disease in systemic sclerosis (SSC) as the representative of chronic lung disease secondary to systemic immune disorder. The organ/tissue specificities of selected Ts/s were evaluated by comparing with 17 organs/tissues, including the nasal mucosa, nose, breast, gastric, tooth, lung, pancreatic, bone marrow, eye, renal, skin, small intestine, cortex, cervical, liver, mouth and choroid. The disease specificities of selected Ts/s were determined by comparing with 31 extrapulmonary diseases. Detailed information on validation of specificities was described in Section 8 of Supplementary Materials.

Validation of Treg^{fci} spatialization. In order to define the biological differences between 5g-Treg^{fci} and 3g-Treg^{fci}, the spatial consistency was evaluated in two sets of data: our in-house spatial transcriptomics dataset ($n = 4$, described above); and public datasets of spatial transcriptomes, including breast cancer (BRCA, $n = 6$), colorectal cancer (CRC, $n = 4$), hepatocellular carcinoma (HCC, $n = 3$), ovarian cancer (OV, $n = 8$), pancreatic ductal adenocarcinoma (PDAC, $n = 3$), and squamous cell carcinoma (SCC, $n = 4$), from the SpatialTME database website.[12] The spatial distribution patterns of 5g-Treg^{fci} or 3g-Treg^{fci} differed as depicted in Supplementary Fig. 8a–d. To determine the significance of Treg^{fci} localization, spatial gene expression data were integrated with TME annotations, as detailed in Section 9 of the Supplementary Materials.

All slides were normalized prior to downstream analysis to ensure data comparability across different spatial spots and tissue samples. To investigate the spatial distribution and relationship, *FOXP3*, *CTLA4*, *IL2RA*, and *ETS1* were selected and examined, and spatial gene expression patterns were visualized using the "SpatialFeaturePlot" function. To quantify expression differences across spatially annotated regions in the SpatialTME database (e.g., tumor core, invasive margin, stromal compartments), the "DotPlot" and "VlnPlot" functions from the Seurat package (version 3.0.2) were utilized. The Treg^{fci} spatialization was assessed by calculating a Treg^{fci} signature score using the "AddModuleScore" function in the Seurat package. The spatial distribution of the composite score was visualized using the same spatial plotting approach to maintain consistency. Statistical comparisons between spatial regions were performed using the Wilcoxon rank-sum test (two groups) or Kruskal–Wallis test (multiple groups) (version 0.6.0).[48]

Validation of Treg^{fci} functions

Construction of stable multi-gene knockdown. To evaluate the influence of each Treg^{fci} ciMGP or combination of genes on cell functions, mono-target and multi-target short hairpin RNA (shRNA) lentiviral delivery systems targeting *FOXP3*, *CTLA4*, and *IL2RA* were constructed. The shRNA sequences (selected from 3–5 candidates per gene) were validated and synthesized (GeneChem Co., Ltd., Shanghai, China). After the sequencing and function of each mono-gene for knockdown were screened and selected, the *FOXP3*, *CTLA4*, and *IL2RA* genes in T cells were simultaneously knocked-down by constructing a recombinant lentiviral vector that encoded multiple shRNA sequences, each of which specifically targeted a distinct gene of interest. The shRNA sequences used in the study were listed in Section 10 of Supplementary Materials. For combinatorial gene silencing, a recombinant lentiviral construct encoding three tandem U6 promoter-driven shRNA cassettes was designed to simultaneously target FOXP3, CTLA4, and IL2RA, according to a previous protocol.[49] The equal mixtures of three mono lentiviral vectors were produced for co-infection and validated for specificity and knockdown efficiency. The lentiviral particles were produced by transient co-transfection of HEK-293T cells with the shRNA plasmid(s) and the packaging plasmids pMD2.G and psPAX2, using Lipofectamine 2000 (Thermo Fisher Scientific). Viral supernatants were collected at 48 and 72 h following the transfection, filtered through a 0.45 μm filter (Thermo Fisher Scientific), and concentrated via ultracentrifugation at 25,000 rpm for 2 h at 4 °C. The viral titer was determined using a p24 antigen ELISA kit (QuickTiter™ Lentivirus Titer Kit, Cell Biolabs), and adjusted to $1–5 \times 10^8$ TU/mL prior to use. T cells were seeded at a density of 5×10^5 cells per 60 mm dish and transduced with the multi-target lentivirus or a mixture of the mono-target lentiviruses at a multiplicity of infection of 10–20, in the presence of 8 μg/mL Polybrene (GeneChem, Shanghai, China). After 24 h, the medium was replaced, and the cells were cultured for 48 h before being selected with puromycin (3 μg/mL, Beyotime, ST551) for 5–7 days. Transduction efficiency was initially assessed via fluorescent microscopy or flow cytometry, based on the co-expression of a green fluorescent protein (GFP) reporter gene encoded in the lentiviral backbone. Cell population with >80% GFP positivity were considered to be successfully transduced. The knockdown efficiency of each target gene was confirmed via qRT-PCR, wherein T cells transduced with a non-targeting shRNA lentivirus (shGFP) served as negative controls. The resulting stable cell lines with triple or mono gene knockdown were designated as Treg$^{fci\text{-}KD}$, Treg$^{foxp3\text{-}KD}$, Treg$^{il2ra\text{-}KD}$, Treg$^{ctla4\text{-}KD}$, and Treg$^{ets1\text{-}KD}$ or ETS1 up-regulated cell-lines as Treg$^{ets1\text{-}OE}$ for subsequent functional analyses.

Validation of Treg^{fci} ciMGPs and ETS1 protein expression. The protein expression and distribution of FOXP3, CTLA4, IL2RA, and *ETS1* in lung tissues of NSCLC patients were examined through immunohistochemistry. Lung cancer tissue samples were collected from young and old patients with NSCLC ($n = 21$; 11 females, 10 males; 8 young, 13 old), followed by formalin fixation and paraffin embedding in blocks. Sections of 4 μm thickness were cut from the blocks and mounted on glass slides for immunohistochemical staining. To enhance the exposure of epitopes masked during the fixation process, heat-induced antigen retrieval was performed. Sections were deparaffinized, rehydrated through a graded ethanol series, and subjected to heat-induced antigen retrieval in a pH 7.8 buffer for 10 min at 100 °C using a pressure cooker. To block nonspecific antibody binding and reduce background staining contamination, sections were incubated in a blocking solution containing 5% normal goat serum in phosphate-buffered saline (PBS) for 30 min at room temperature. Primary antibodies specific to FOXP3 (Cat#AB3199), IL2RA (Cat#AB2709), CTLA4 (Cat#AB3659) and *ETS1* (Cat#AB3669) were purchased from Xiamen Zhiwei Electronic and Pharmaceutical Co., Ltd. The sections were incubated with the primary antibodies overnight at 4 °C in a humidified chamber. After washing the sections with PBS thrice, a secondary antibody conjugated to horseradish peroxidase was applied. The sections were incubated with the secondary antibody for 60 min at room temperature. The bound antibodies were visualized using the EnVision Kit (Agilent DAKO, Santa Clara, USA) according to the manufacturer's instructions, to produce a brown chromogenic precipitate at the site of antibody binding. The nuclei were counterstained with hematoxylin. The stained sections were examined under a light microscope (Olympus BX51). The staining intensity and distribution of marker proteins were assessed in tumor cells and tumor-infiltrating lymphocytes (TILs).

The protein expression and distribution of *ETS1* in T cells were examined via cellular immunofluorescence staining. T cells were collected, washed twice with PBS, loaded onto poly-L-lysine-coated glass slides, and cytospun at $800 \times g$ for 5 min. The cells were immediately fixed in 4% paraformaldehyde (Sigma-Aldrich,

USA) at room temperature for 15 min, followed by three washes with PBS. The cell membranes were permeabilized with 0.1% Triton X-100 (Solarbio, China) in PBS for 10 min, followed by blocking with 5% bovine serum albumin (Sigma-Aldrich) in PBS for 1 h at room temperature. The cells were incubated with anti-*ETS1* antibody (Abclonal, China) and Cy3-conjugated antibody (Beyotime, China) in blocking buffer overnight at 4 °C in a humidified chamber. After three washes with PBS, the nuclei were counterstained with 1 μg/mL 4′,6-diamidino-2-phenylindole (DAPI) (Thermo Fisher Scientific) for 5 min in the dark. Slides were prepared with the cell and nuclei samples, followed by mounting with an anti-fade fluorescence mounting medium (Vectashield, Vector Laboratories) and sealing with coverslips. Fluorescence images were acquired using the Perkin-Elmer high-content automatic imaging system. Cy3 and DAPI signals were captured using appropriate excitation/emission filters. The quantitative fluorescence intensity analysis of *ETS1* localization was performed using the ImageJ software (NIH, USA). For each randomly selected field, the mean fluorescence intensity of *ETS1* within the cell cytoplasm and nucleus was measured. Each experiment was repeated at least 6 times to ensure reproducibility.

Immunoblotting assay. Cell samples were homogenized in RIPA lysis buffer (P0013B, Beyotime, Shanghai, China) and centrifuged at 14,000 × *g* for 30 min at 4 °C for Western blotting, as previously described.[50] Equal quantities of protein (30 μg) were loaded onto 7.5% − 12.5% SDS-PAGE gel, followed by electrophoresis and transfer to nitrocellulose membranes (0.45 μm for proteins >50 kDa, 0.2 μm for proteins ≤50 kDa). The membranes were blocked with 5% non-fat blocking grade milk (Bio-Rad, Hercules, CA, USA) and incubated overnight at 4 °C with the following primary antibodies: anti-caspase3 (1:1000, A0214, Abclonal, China), anti-cleaved caspase 3 (1:1000, 25128-1-AP, Proteintech, Wuhan, China), anti-Bax (1:1000, A19684, Abclonal, China), anti-Bcl2 (1:1000, A19693, Abclonal, China), anti-GAPDH (1:5000, A19056, Abclonal, China), anti-IL4 (1:1000, A23277, Abclonal, China), anti-IL10 (1:1000, A12255, Abclonal, China), anti-CD73 (1:1000, A25914, Abclonal, China), anti-CD29 (1:1000, A23497, Abclonal, China), anti-CD69 (1:1000, A26620, Abclonal, China), and anti-CD223 (1:1000, A2996, Abclonal, China). On the following day, the membranes were incubated with the appropriate secondary antibody (1:2000, AS097, Abclonal, Wuhan, China) at room temperature for 1 h. The immunoblots were visualized using an ECL Plus chemiluminescence reagent kit (P0018S, Beyotime, Shanghai, China). For quantitative analysis, the integrated density of each protein band was calculated using ImageJ software (ImageJ 1.5, NIH, USA). The resulting values were normalized to GAPDH as an internal control.

Construction of Treg-specific Ets1-deficient mice. To investigate the functional role of *Ets1* in Tregs, a Treg-specific *Ets1* conditional knockout mouse model was generated using the Cre-LoxP system. Guide RNAs (gRNAs) targeting the intronic regions flanking exons 7 and 8 were designed using CRISPR design tools to minimize off-target effects and synthesized. Cas9 mRNA (100 ng/μL) was co-injected with gRNAs and a donor plasmid containing 1–2 kb homologous arms on each side of the targeted region to facilitate homology-directed repair. Off-target sites with high similarity to the gRNA sequences were evaluated in silico, and gRNAs with minimal predicted off-target activity were selected. The gRNAs, Cas9 mRNA, and donor vector were microinjected into fertilized C57BL/6 mouse zygotes. The mouse embryos were transferred into the oviducts of pseudopregnant females to obtain founder (F0) mice. F0 offspring were screened via PCR and validated via Sanger sequencing to identify mice with the correct loxP-flanked alleles. Positive founders were bred with wild-type C57BL/6JCya mice to obtain heterozygous *Ets1*-floxed mice (*Ets1*fl/+), which were interbred to obtain homozygous *Ets1*fl/fl mice. To achieve Treg-specific deletion of *Ets1*, *Ets1*fl/fl mice were crossed with Foxp3YFP-Cre transgenic mice, wherein Cre recombinase was driven by the endogenous FOXP3 promoter. Breeding was performed by mating male Foxp3YFP-Cre$^{+/-}$; *Ets1*+/fl mice with female *Ets1*fl/fl mice, generating experimental mice (Foxp3YFP-Cre +/−; *Ets1*fl/fl, hereafter referred to as *Ets1*Treg-KO) and control littermates (Foxp3YFP-Cre−/−; *Ets1*fl/fl, referred to as *Ets1*fl/fl). All animals were maintained on a pure C57BL/6 genetic background to minimize strain-related immunological variability. Tail biopsies were collected for genotyping at 2–3 weeks of age using PCR with primers specific to the loxP-flanked region and the Foxp3$^{YFP\text{-}Cre}$ transgene. The PCR products were resolved on agarose gels and visualized under UV illumination.

Ets1$^{flox/flox}$ control and *Ets1*$^{flox/flox;FOXP3\text{-}Cre}$ conditional knockout mice (6–8 weeks old, female) were sourced from Cyagen Biosciences. To establish the lung cancer mouse model, 10^6 Lewis lung carcinoma cells (LLC cells, ATCC, CRL-1642) were suspended in 100 μL of sterile PBS and injected subcutaneously into the left axillary region of each mouse. Tumor development was monitored every 2–3 days using calipers to measure the tumor size. Mice were randomly assigned to different experimental groups ($n = 6$–8 animals/group). Tumor growth was monitored for approximately 30 days, or until tumors reached a maximum volume of 1500 mm^3, or if mice exhibited signs of distress or impaired mobility, at which point animals were euthanized according to institutional ethical guidelines. All animal procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of Zhongshan Hospital, Fudan University, and conducted in accordance with institutional and national ethical guidelines. Mice were housed under specific pathogen-free conditions with a 12-h light/dark cycle and provided with autoclaved food and water *ad libitum*.

Ets1 roles in metabolomics. To deeply define the potential influence of *Ets1* on metabolism, systemic (circulation) and local metabolism (lung) were measured in *Ets1*flox/flox control and *Ets1*flox/floxFOXP3-Cre conditional knockout mice ($n = 6$/group) using untargeted metabolomics. Plasma samples were obtained from peripheral blood of mice, and lung tissues were dissected, with half of the lobes reserved for metabolomics analysis. For untargeted metabolomics, plasma and lung tissue samples were kept on ice and vortexed thoroughly before extraction. Briefly, 300 μL of 80% methanol containing 2-chloro-phenylalanine (4 ppm) was added to each sample, vortexed for 1 min, and centrifuged at 12,000 rpm for 10 min at 4 °C. The supernatant was collected, filtered through a 0.22 μm membrane, and transferred into LC-MS vials for analysis. Chromatographic separation was performed on an ACQUITY UPLC® HSS T3 column (2.1 × 100 mm, 1.8 μm; Waters) using a Vanquish UHPLC system (Thermo Fisher Scientific, USA) with a flow rate of 0.3 mL/min and an injection volume of 5 μL. The mobile phase consisted of 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B) for positive mode, and 5 mM ammonium format in water (A) and acetonitrile (B) for negative mode. Mass spectrometry was carried out on an Orbitrap Exploris 120 mass spectrometer (Thermo Fisher Scientific, USA) equipped with an electrospray ionization (ESI) source operating in positive and negative ion modes. The MS parameters were set as follows: spray voltage 3.50 kV (positive) and −2.50 kV (negative), capillary temperature 325 °C, sheath gas 40 arb, auxiliary gas 10 arb, full MS resolution 60,000, scan range *m/z* 100–1000, and data-dependent MS/MS resolution 15,000 with normalized collision energy of 30%. Quality control samples were prepared by pooling equal aliquots of all test samples and were injected at regular intervals to monitor instrument stability.

Cell culture and reagents. Human T cell acute lymphoblastic leukemia cell lines (T^{fci}), human lung adenocarcinoma epithelial cell lines (A549), and human bronchial epithelial cell lines (HBE), human small cell lung carcinoma epithelial cell line (H446), human

lung large cell lung cancer cell line (H460), human lung adenocarcinoma (SPC-A1), human small cell lung carcinoma cell line (H1688), human lung large cell carcinoma (H661), and human lung adenocarcinoma cell line (LTEP), were obtained from ATCC. All cell lines were cultured in RPMI 1640 medium (Keygen Bio, KGM31800-500, Shanghai, China) supplemented with 10% fetal bovine serum (Hyclone, SH30084.03, Logan, UT, USA), 100 U/mL penicillin, and 100 μg/mL streptomycin at 37 °C in a humidified chamber containing 5% CO_2. For co-culture experiments, T^{fci} stably transfected with either control vector (V- $T^{fci\text{-}NC}$), PHA-stimulated control vector (PHA-$T^{fci\text{-}NC}$), $Treg^{fci}$ ciMGPs knockdown vector (V-$T^{fci\text{-}KD}$), or PHA-stimulated fci knockdown vector (PHA-$T^{fci\text{-}KD}$) were used ($n = 6$/group), as detailed in Section 11 of Supplementary Materials. The co-culture was performed in 24-well plates equipped with Transwell inserts (0.4 μm pore size, Corning). Adherent epithelial or lung cancer cells were seeded in the lower chamber, while the different $Treg^{fci}$ subtypes were seeded in the upper insert. The co-culture system was maintained for 24 h under standard culture conditions.

Quantitative reverse transcription PCR (qRT-PCR). Universal RNA Purification Kits (EZB-RN4, eBiosciences, Shanghai, China) were used to extract total RNA according to the manufacturer's protocol. RNA concentration and purity were determined by measuring A260/A280 ratios using a Nanodrop spectrophotometer (Thermo Fisher Scientific, USA), and RNA integrity was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, USA). A total of 500 ng of total RNA was reverse-transcribed into cDNA using the PrimeScript RT Master Mix (RR036A, Takara, Japan). The real-time PCR was performed using TB Green Premix Ex Taq (RR420A, Takara, Japan) with the primers listed in Supplementary Table 2 on the ABI 7000 PCR instrument (Applied Biosystems, USA) by two-stage program parameters. The expression level of the target gene was normalized using the GAPDH housekeeping gene and the expression level was calculated by the comparative method ($2^{-\Delta\Delta Ct}$). The primer design and DNA oligonucleotide sequences of GAPDH, FOXP3, CTLA4, IL2RA, IL-10, TGFβ, EBI3, TLR4, IL-27, IL-4, IL-6, IL-1β ETS1, ETS2, ETV1, ETV4, ETV6, TNFα, and IL-8 were detailed in the Section 12 of Supplementary Materials.

Measurements of cytokine proteins. Levels of TNF-α, IL-1β, IL-6, and IL-8 proteins in culture supernatants across different experiments were measured using ELISA. We repeated the measurement of supernatant levels of TNF-α, IL-1β, IL-6, and IL-8 proteins 24 h after $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ culture pre-treated with vehicle or PHA (V-$T^{fci\text{-}NC}$, PHA-$T^{fci\text{-}NC}$, V-$T^{fci\text{-}KD}$, and PHA-$T^{fci\text{-}KD}$), or co-culture of A549 or HBE with $T^{fci\text{-}NC}$ or $T^{fci\text{-}KD}$ pre-treated with vehicle or PHA ($n = 6$/group). Each was performed with technical duplicates, and analyzed using standard ELISA procedures. The detailed methodology was described in Section 13 of the Supplementary Materials.

Co-culture assay. T cells were co-cultured with either A549 or HBE at a 10:1 ratio for 24 h. For the indirect co-culture model, A549 or HBE were seeded on the lower chamber of a 24-well plate at a density of 1.5×10^5 per well and allowed to adhere overnight. T cells were added to the upper chamber at a density of 1.5×10^6 cells per insert and cultured until the cells were confluent. Co-cultures were maintained in RPMI 1640 medium supplemented with 10% FBS for 24 h.

Crystal violet staining. Viable cell numbers were determined via crystal violet staining. A549, HBE, H446, H460, SPC-A1, H1688, H661 and LTEP cells were plated in 6-well plates/24-well plates overnight. After treatment, the RPMI 1640 culture medium was removed, and the cells were washed with PBS, fixed with paraformaldehyde for 10 min at room temperature, and washed twice. Cells were stained with 0.5% crystal violet (Beyotime Institute of Biotechnology, Cat number: KGA317, Jiangsu, China) for 15 min at room temperature. After staining, cells were washed twice with distilled water to remove excess dye, air-dried, and then imaged under a microscope. Images of stained cells were captured using an inverted microscope (Olympus IX71) at ×20 magnification. The number of viable cells was quantified by counting stained cells in randomly selected fields using ImageJ software (NIH, USA). All experiments were repeated twice ($n = 6$/group).

Cell Counting Kit-8 (CCK-8) assay. Cell viability was assessed using CCK-8 (Keygen, Catalog KGA317, Jiangsu, China) according to the manufacturer's instructions. Briefly, 2000 cells/well were seeded into 96-well plates and subjected to transient transfection with plasmid vectors. At pre-determined time points, 10 μL of CCK-8 reagent was added to each well, and cells were incubated for 2 h at 37 °C. Absorbance at 450 nm was measured using a microplate reader (BioTek, USA). Each experiment was repeated thrice ($n = 6$/group).

Apoptosis assay. For analyzing apoptosis, single-cell suspensions were subjected to flow cytometric staining. The T cells were harvested and processed using the Apoptosis Detection Kit (Beyotime, Shanghai, China) and analyzed using a Beckman FACS flow cytometer.10. The percentage of apoptotic cells was determined via flow cytometry using the Alexa Fluor 488/annexin V/PI Cell Apoptosis Kit (Beyotime, Shanghai, China) according to the manufacturer's protocol.

Reactive oxygen species (ROS) assay. Levels of intracellular ROS were measured via flow cytometry using an ROS detection kit (S0033, Beyotime, Shanghai, China). Briefly, the cells were washed with PBS after treatment, incubated with 15 μM 2',7'-dichlorofluorescin diacetate (DCFH2-DA) for 30 min at 37 °C in the dark, and analyzed using a FACSCalibur flow cytometer (BD Biosciences). Intracellular ROS levels were expressed as the mean DCFDA fluorescence intensity of the cells.

Intracellular calcium assay. Intracellular levels of calcium signaling were measured using a Fluo-4 Calcium Assay Kit according to the manufacturer's protocol (S1061S, Beyotime, Shanghai, China). After treatment, medium from the culture plate was removed, and 100 μl of Fluo-4 AM dye loading solution was added to each well quickly. The cells were incubated at 37 °C for 30 min, followed by an additional 30 min at room temperature in the dark. Intracellular Ca^{2+} signals were measured using flow cytometry with excitation at 494 nm and emission at 516 nm. Potential fluorescence by organic transporters outside the cells was inhibited by adding probenecid, and the baseline signal was reduced.

Mitochondrial stain and high-content screening assay. Cellular mitochondria were stained with MitoTracker Deep Red (MTDR) (40743ES50, Yeasen, Shanghai, China) and the fluorescence was determined. To assess the colocalization of nuclei and mitochondria, cells were stained with 25 nM Mito Tracker Deep Red FM (MTDR; Yeasen, cat# 40743ES50) in pre-warmed culture medium for 30 min at 37 °C under 5% CO_2 to label functional mitochondria. After staining, cells were washed with PBS and fixed with 4% paraformaldehyde for 15 min at room temperature. Following fixation, nuclei were incubated with 0.5 μg/mL Hoechst 33342 (Beyotime, Shanghai, China) at room temperature for 10 min. The emission/excitation wavelengths of the mitochondria and Hoechst stain were 665/644 nm and 460/346 nm, respectively. The cells were screened using the Perkin-Elmer high-content automatic imaging system in confocal mode, and the results were analyzed in the Perkin-Elmer Harmony image analysis software (PerkinElmer Life Sciences).

Oxygen consumption rate (OCR) measurement. The OCR (in pmol/min) of cultured T cells was determined using a Seahorse XF96 Extracellular Flux Analyzer (Seahorse Bioscience) as per the manufacturer's instructions. T cells were plated at a density of 200,000 or 50,000 cells/well in 24 or 96-well Seahorse assay plates and incubated for 1 h in a 37 °C incubator without CO_2. The culture medium was replaced with XF assay medium containing unbuffered DMEM, 5.5 mM glucose, and 0.5 mM carnitine. Values of OCR were measured using the XF96 analyzer in XF base medium (pH 7.4) containing 1 μM mitochondrial respiration inhibitors oligomycin (OLI), 2 μM carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone (FCCP), and 1 μM Antimycin A (AA). The data were analyzed using the Seahorse XF Mito Stress Test Reporter Generator package.

Cell trajectory assays. T cells at 2×10^5 cells/mL/well were plated on 24-well plates for imaging using the Perkin-Elmer high-content automatic imaging system (PerkinElmer Life Sciences). The Perkin-Elmer Harmony image analysis software (PerkinElmer Life Sciences) was used to calculate the cell numbers, mean square displacement (MSD), and velocity of the T cells post-migration. The values calculated for the areas were inversely proportional to the amount of cell migration. Following a 48-h migration period, cell migration data were acquired through a high-content screening approach in the Operetta high-content screening system (Perkin Elmer). Data were presented as the mean value per group, along with the standard error of the mean (SEM).

Transmission electron microscopy. T cells were harvested, washed twice with PBS, and immediately fixed in 2.5% glutaraldehyde (Biossci, Wuhan, China) in 0.1 M phosphate buffer (pH 7.4) at 4 °C overnight. After three washes with phosphate buffer (10 min each), the samples were post-fixed with 1% osmium tetroxide for 2 h at room temperature. Cells were then dehydrated through a graded ethanol series with each step lasting 15 min (30%, 50%, 70%, 90%, and 100%), followed by two 15 min washes in 100% ethanol. The dehydrated samples were embedded in epoxy resin (SPI-PON 812), polymerized at 60 °C for 48 h, and sectioned into ultrathin slices (70–90 nm) using an ultramicrotome (Leica EM UC7, Leica Microsystems, Germany). The sections were mounted onto copper grids and stained with 2% uranyl acetate for 15 min, followed by 2.6% lead citrate for 10 min at room temperature. Transmission electron microscopy was performed using a Hitachi HT7800 electron microscope (Hitachi, Japan) at an accelerating voltage of 80 kV. Imaging and data acquisition were conducted by Biohao Biotechnology Co., Ltd. (Wuhan, China).

Human PBMC isolation assay. PBMCs were harvested from the healthy subjects and patients with NSCLC and isolated via the density gradient centrifugation using Ficoll (#1114546, Axis-Shield/Alere Technologies AS, Germany). Blood was carefully layered over Ficoll and centrifuged at $400 \times g$ for 25 min at room temperature without brake. The PBMC layer was collected, washed twice with PBS, and counted using a hemocytometer. The collection and analysis of the patient samples were performed in accordance with the Declaration of Helsinki and the protocols were approved by the Institutional Review Board of Zhongshan Hospital. Written informed consent for genomic analysis was obtained from the participant.

Flow cytometry assay. Surface staining for flow cytometry was performed as per standard flow cytometry protocols. Briefly, PBMCs (~1×10^6 cells 100 μL per group) were collected from patients with NSCLC or healthy subjects, and first stained with FVS510 Live/Dead dye (Cat#564406, BD Biosciences) for 20 min at room temperature in the dark. Surface markers were labeled using the following antibodies for the classical Treg identification ($CD3^+CD4^+CD25^+CD127^-CD45^+$): FITC-conjugated anti-CD3 (Cat#555339, clone UCHT1), APC-conjugated anti-CD4 (Cat#555349, clone RPA-T4), PE-conjugated anti-CD25 (Cat#555432, clone M-A251), BV421-conjugated anti-CD127 (Cat#562436, clone HIL-7R-M21), and APC-Cy7-conjugated anti-CD45 (Cat#557833, clone HI30), and for $Treg^{fci}$ identification ($CD3^+CD4^+FOXP3^+IL2RA^+CTLA4^+$): FITC-conjugated anti-CD3 (Cat#555339, clone UCHT1); APC-conjugated anti-CD4 (Cat#555349, clone RPA-T4); PE-conjugated anti-IL2RA/CD25 (Cat#555432, clone M-A251); PE-Cy7-conjugated anti-CTLA4/CD152 (Cat#567341, clone BNI3) and AF647-conjugated anti-FOXP3 (Cat#560045, clone PCH101, BioLegend). For intracellular staining, AF647-conjugated anti-FOXP3 antibody was applied for 30 min at 4 °C in the dark, after the cells were fixed and permeabilized using BD Cytofix/Cytoperm kit (BD Biosciences) or FOXP3/Transcription Factor Staining Buffer Set (Cat#00-5523-00, eBioscience) according to the manufacturer's instructions. Data were collected on an LSR III flow cytometer (BD Biosciences, California, USA) and analyzed using the FlowJo software v10.7.1 (FlowJo LLC). Gating strategies were verified using single-stain. All experiments were performed in three independent replicates.

Bulk RNA-seq assay. Total RNA was extracted using TRIzol reagent (Beyotime, R0016, Suzhou, China), and RNA quality and integrity were assessed using the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). The samples with an RIN > 7.0 were used for library construction. The poly(A)+ mRNA was enriched using oligo(dT)-conjugated magnetic beads (New England Biolabs), followed by fragmentation at 94 °C and synthesis of double-stranded cDNA using the NEBNext Ultra Directional RNA Library Prep Kit. After undergoing end repair and adapter ligation, PCR amplification was performed (typically 15 cycles). The resulting libraries were purified using AMPure XP beads (Beckman Coulter, USA). Library quality was evaluated via qRT-PCR, and libraries with an effective concentration >2 nM were pooled according to the target sequencing depth. Sequencing was performed using NovaSeq 6000 platform (paired-end) by Shanghai Biotechnology Corporation (Shanghai, China).

Treg effects on autologous NSCLC organoid. Autologous PBMC and cancer tissues were collected from four patients with NSCLC into tumor storage solution. Organoids were generated, passaged for 3–4 generations, and cryopreserved, as detailed in Section 14 of the Supplementary Materials. The recovered organoids were between the 4th and 6th generations, and the culture medium was changed every 3 days. PBMCs were cultured overnight at 37 °C and used for stimulating the autologous NSCLC organoids. After collection, cells were washed twice with staining buffer (DPBS supplemented with 1% FBS) and stained with Live/dead, Annexin V-FITC/PI, CD3, CD4, CD25, CD127, PD1, IFNγ. The cells were collected on BD Lyric flow cytometer and the analysis was performed using FlowJo software. Prior to co-culture with autologous T cells, tumor organoids were pre-stimulated with IFNγ to enhance antigen presentation. Plate-bound anti-CD28 and IL-2 were added to provide co-stimulation and support T cell proliferation, respectively. PBMC were stimulated weekly with autologous tumor organoids. Tumor recognition by CD8+ T cells was evaluated at baseline and after 2 weeks of co-culture, by staining for IFNγ and PD1.

ETS1 roles in human and mouse $Treg^{fci}$. To evaluate direct roles of *ETS1* in $Treg^{fci}$, human $Treg^{fci}$ were isolated from PBMCs of healthy volunteers ($n = 22$) and treated with vehicle (1% DMSO) or *ETS1* inhibitor (TK216, HY-122903, ShangHai Caerulum Pharma Discovery Co., Ltd., Shanghai, China) at 0.25 μM/ml for 24 h ($n = 11$/group). TK216, an orally active ETS inhibitor, inhibits *ETS* family protein function through the protein-protein interactions of EWS-FLI1.[70] Afterward, cells were collected and prepared for scRNA-seq as described above. Peripheral WBCs and lung tissues were

obtained from twelve mice (6 Ets1fl/fl control and 6 Ets1fl/fl; Foxp3YFP-Cre$^{+/-}$). WBCs were isolated following red blood cell lysis, while lung lobes were immediately dissociated into single-cell suspensions within 30–60 min of harvest for scRNA-seq. Tissues were transported on ice, minced on ice, and digested with 0.25% Trypsin (Thermo Fisher, Cat. no. 25200-072) and 10 μg/mL DNase I (Sigma, Cat. no. 11284932001) in PBS supplemented with 5% FBS (Thermo Fisher, Cat. no. SV30087.02) at 37 °C with gentle shaking (50 rpm) for 40 min. Cell suspensions were collected every 20 min to maximize yield and viability, filtered through a 40 μm nylon strainer (Corning, Cat. no. 352340), and subjected to red blood cell lysis (Thermo Fisher, Cat. no. 00-4333-57). The suspensions with >90% viability were used for library preparation and were loaded at 2×10^5 cells/mL onto the ABcellar® Single Cell Chip Set (ABclonal, RK30150). Gel Beads-in-emulsions (GEMs) were generated using the ABcellar® Single Cell Encapsulation Instrument (ABclonal, AI50001), containing the single cell and barcode-labeled gel bead for mRNA capture. The cDNAs were amplified and used for library construction according to the ABcellar® Single Cell 3' RNA-seq Lib Prep Kit protocol (ABclonal, RK20390). Library quality was evaluated with an Agilent High Sensitivity DNA Chip on a Bioanalyzer 2100 (Agilent Technologies) and quantified with the Qubit High Sensitivity DNA Assay Kit (Thermo Fisher Scientific). Sequencing was performed on an Illumina NovaSeq 6000 platform with 2×150 bp paired-end reads. Primary data processing was conducted with the ABcellar v1.2.5 pipeline using default parameters. FASTQ files generated from Illumina sequencing were aligned to the mouse reference genome (GRCm38) using the STAR algorithm. Unique molecular identifiers were counted to generate gene-barcode matrices. The resulting expression matrix was imported into Seurat (v3.0.2) in R for quality control, normalization, dimensionality reduction, clustering, and downstream analyses.

Statistical analysis

Data are presented as the means ± SEM. Group comparisons were performed using the Mann–Whitney *U* test, using the Benjamini–Hochberg false discovery rate (FDR) procedure in multiple groups across multiple conditions. The one-way analysis of variance (ANOVA) was applied for comparison among multi-groups, followed by post-hoc Tukey's test as needed. The hypergeometric test was used to evaluate statistical significance in the functional enrichment analysis of gene lists. To evaluate the potential influence of age variation on 5g- or 3g-Treg^{fci}, we re-evaluated our own and public datasets and categorized them into young (<65 years) and old (≥65 years) groups, as detailed in Section 15 of the Supplementary Materials. Raw sequencing data (fastq files) and gene-barcode matrices were processed and analyzed as detailed in Section 16 of the Supplementary Materials. The mRNA expression levels of each T ciMGP were extracted from the raw gene expression matrix, and the number of T cell subsets was summarized as SEM. Regional differences of Treg^{fci} scores were assessed using the Wilcoxon rank-sum test, and correlations between individual marker expression and composite Treg^{fci} scores were analyzed by Spearman's correlation. All statistical analyses and graphical visualizations were conducted using R v.4.1.2. The two-tailed *p* value less than 0.05 was considered statistically significant.

DATA AVAILABILITY

Patient data for non-small-cell lung cancer cases described in this study are publicly available in the Genome Sequence Archive (GSA) under accession HRA011677.

ACKNOWLEDGEMENTS

We especially appreciate Professor Dongshen Chen and his team for long-term scientific collaboration and database sharing, especially for those valuable and respectful considerations and suggestions, which help improve the quality of the manuscript. We also appreciate Professor Chuanyu Liu for human lymphocyte analyses, and BoHao Biotech for technical supports. Noncommunicable Chronic Diseases-National Science and Technology Major Project (2024ZD0529300). Fujian Province Medical and Health Senior Talent Team Introduction Project.

AUTHOR CONTRIBUTIONS

Xuanqi Liu contributed to study design, experimental performance, data analysis, and manuscript writing, Yifei Liu to data analysis and programing, Wanxin Duan to data analysis and manuscript writing, Jianming Weng to immunohistology and spatial analysis, Hao Dai to data analysis and computational programing, Fangmin Liu, Lingyan Wang, Mengjia Qian, Bijun Zhu, Shuyao Wu, Youqiong Ye, Shaojun Hong, and Wang Mingjie to experimental performance, Tianxiang Chen to study design and sampling, Yunfeng Chen to study design and in vitro study, Duojiao Wu to study design, validation of immune cell function, data analysis, and manuscript writing, Yuanlin Song to study design, clinical data analysis, and manuscript writing, and Xiangdong Wang to study design, data analysis, and manuscript writing. All authors have read and approved the article.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41392-026-02677-6.

Competing interests: The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Ali, M. et al. T cells targeted to TdT kill leukemic lymphoblasts while sparing normal lymphocytes. *Nat. Biotechnol.* **40**, 488–498 (2022).
2. Sun, Y. et al. Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. *Cell* **184**, 404–421 (2021).
3. Sun, Y. F. et al. Dissecting spatial heterogeneity and the immune-evasion mechanism of CTCs by single-cell RNA-seq in hepatocellular carcinoma. *Nat. Commun.* **12**, 4091 (2021).
4. Pai, J. A. et al. Lineage tracing reveals clonal progenitors and long-term persistence of tumor-specific T cells during immune checkpoint blockade. *Cancer Cell* **41**, 776–790 (2023).
5. Zhong, J. et al. Single-cell RNA sequencing reveals the molecular features of peripheral blood immune cells in children, adults and centenarians. *Front. Immunol.* **13**, 1081889 (2022).
6. Liu, Y. et al. Immune phenotypic linkage between colorectal cancer and liver metastasis. *Cancer Cell* **40**, 424–437 (2022).
7. Lozano, A. X. et al. T cell characteristics associated with toxicity to immune checkpoint blockade in patients with melanoma. *Nat. Med.* **28**, 353–362 (2022).
8. Andor, N. et al. Single-cell RNA-Seq of follicular lymphoma reveals malignant B-cell types and coexpression of T-cell immune checkpoints. *Blood* **133**, 1119–1129 (2019).
9. Fang, H. et al. Can single cell RNA sequencing reshape the clinical biochemistry of hematology: new clusters of circulating blood cells. *Clin. Transl. Med.* **11**, e671 (2021).
10. Villani, A. C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).
11. Papalexi, E. & Satija, R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.* **18**, 35–45 (2018).
12. Liu, X., Wang, D. C., Powell, C. A. & Wang, X. Challenges of clinical translation from single-cell sequencing to measures in clinical biochemistry of haematology: definition of immune cell identities. *Clin. Transl. Med.* **13**, e1401 (2023).
13. Yazar, S. et al. Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. *Science* **376**, eabf3041 (2022).
14. Perez, R. K. et al. Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. *Science* **376**, eabf1970 (2022).
15. Suo, C. et al. Mapping the developing human immune system across organs. *Science* **376**, eabo0510 (2022).
16. Zhang, L. et al. Clinical and translational values of spatial transcriptomics. *Signal Transduct. Target. Ther.* **7**, 111 (2022).
17. Chen, Y., Shen, J., Kasmani, M. Y., Topchyan, P. & Cui, W. Single-cell transcriptomics reveals core regulatory programs that determine the heterogeneity of circulating and tissue-resident memory CD8(+) T cells. *Cells* **10**, 2143 (2021).
18. Zhang, Q. et al. Landscape and dynamics of single immune cells in hepatocellular carcinoma. *Cell* **179**, 829–845 (2019).

19. Yin, J. et al. Single-cell genomics elucidates molecular variations and regulatory mechanisms in circulating immune cells. *bioRxiv* https://doi.org/10.1101/2025.01.26.634963 (Preprint) (2025).
20. Liu, Z. et al. A single-cell atlas reveals immune heterogeneity in anti-PD-1-treated non-small cell lung cancer. *Cell* **188**, 3081–3096 (2025).
21. Phillips, J. D. et al. The effect of lung resection for NSCLC on circulating immune cells: a pilot study. *Curr. Oncol.* **30**, 5116–5134 (2023).
22. Liu, X. et al. Evaluation of pulmonary single-cell identity specificity in scRNA-seq analysis. *Clin. Transl. Med.* **12**, e1132 (2022).
23. Liu, X., Powell, C. A. & Wang, X. Forward single-cell sequencing into clinical application: understanding of cancer microenvironment at single-cell solution. *Clin. Transl. Med.* **12**, e782 (2022).
24. AlMusawi, S., Ahmed, M. & Nateri, A. S. Understanding cell-cell communication and signaling in the colorectal cancer microenvironment. *Clin. Transl. Med.* **11**, e308 (2021).
25. Ortega, M. A. et al. Using single-cell multiple omics approaches to resolve tumor heterogeneity. *Clin. Transl. Med.* **6**, 46 (2017).
26. Ohkura, N. & Sakaguchi, S. Transcriptional and epigenetic basis of Treg cell development and function: its genetic anomalies or variations in autoimmune diseases. *Cell Res.* **30**, 465–474 (2020).
27. Chalepaki, A. M. et al. A multi-omics analysis of effector and resting treg cells in pan-cancer. *Comput. Biol. Med.* **189**, 110021 (2025).
28. Schneider, J. L. et al. The aging lung: physiology, disease, and immunity. *Cell* **184**, 1990–2019 (2021).
29. Golzari-Sorkheh, M. & Zuniga-Pflucker, J. C. Development and function of FOXP3+ regulators of immune responses. *Clin. Exp. Immunol.* **213**, 13–22 (2023).
30. Kim, G. R. & Choi, J. M. Current understanding of cytotoxic T lymphocyte antigen-4 (CTLA-4) signaling in T-cell biology and disease therapy. *Mol. Cells* **45**, 513–521 (2022).
31. Shouse, A. N., LaPorte, K. M. & Malek, T. R. Interleukin-2 signaling in the regulation of T cell biology in autoimmunity and cancer. *Immunity* **57**, 414–428 (2024).
32. Mouly, E. et al. The Ets-1 transcription factor controls the development and function of natural regulatory T cells. *J. Exp. Med.* **207**, 2113–2125 (2010).
33. Ouyang, W. & O'Garra, A. IL-10 family cytokines IL-10 and IL-22: from basic science to clinical translation. *Immunity* **50**, 871–891 (2019).
34. Zhou, Y. T. et al. Atg7 in CD4(+) T cells improves intestinal mucosal inflammation by regulating Ets1-mediated T cell differentiation. *Clin. Transl. Med.* **15**, e70462 (2025).
35. Nakamura, S. et al. Single-cell transcriptome analysis reveals functional changes in tumour-infiltrating B lymphocytes after chemotherapy in oesophageal squamous cell carcinoma. *Clin. Transl. Med.* **13**, e1181 (2023).
36. Wang, Z. et al. Landscape of gut mucosal immune cells showed gap of follicular or memory B cells into plasma cells in immunological non-responders. *Clin. Transl. Med.* **14**, e1699 (2024).
37. Liao, S. et al. Stereo-cell: spatial enhanced-resolution single-cell sequencing with high-density DNA nanoball-patterned arrays. *Science* **389**, eadr0475 (2025).
38. Wang, X., Duan, W., Liu, X. & Fan, J. An important step to translate single-cell measurement into clinical practice: stereoscopic cells. *Clin. Transl. Med.* **15**, e70304 (2025).
39. Benegiamo, G. et al. COX7A2L genetic variants determine cardiorespiratory fitness in mice and human. *Nat. Metab.* **4**, 1336–1351 (2022).
40. Travis, W. D. et al. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J. Thorac. Oncol.* **10**, 1243–1260 (2015).
41. Wei, X. et al. Single-cell Stereo-seq reveals induced progenitor cells involved in axolotl brain regeneration. *Science* **377**, eabp9444 (2022).
42. Liu, X. et al. Clinical challenges of tissue preparation for spatial transcriptome. *Clin. Transl. Med.* **12**, e669 (2022).
43. Buttarazzi, D., Pandolfo, G. & Porzio, G. C. A boxplot for circular data. *Biometrics* **74**, 1492–1501 (2018).
44. Royeen, C. B. The boxplot: a screening test for research data. *Am. J. Occup. Ther.* **40**, 569–571 (1986).
45. Simpson, R. J. Jr., Johnson, T. A. & Amara, I. A. The box-plot: an exploratory analysis graph for biomedical publications. *Am. Heart J.* **116**, 1663–1665 (1988).
46. Pagano, M. The histogram and boxplot: a picture is worth a thousand words. *Nutrition* **8**, 374–375 (1992).
47. Zhao, H. et al. Comprehensive assessment of harmful heavy metals in contaminated soil in order to score pollution level. *Sci. Rep.* **12**, 3552 (2022).
48. Jovic, D. et al. Single-cell RNA sequencing technologies and applications: a brief overview. *Clin. Transl. Med.* **12**, e694 (2022).
49. Weng, Y. et al. A multi-shRNA vector enhances the silencing efficiency of exogenous and endogenous genes in human cells. *Oncol. Lett.* **13**, 1553–1562 (2017).
50. Kurien, B. T. & Scofield, R. H. Protein blotting: a review. *J. Immunol. Methods* **274**, 1–15 (2003).