frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 08 July 2024
DOI 10.3389/fimmu.2024.1416751

OPEN ACCESS

EDITED BY
Xuefeng Niu,
First Affiliated Hospital of Guangzhou Medical University, China

REVIEWED BY
Longchao Liu,
Chinese Academy of Sciences (CAS), China
Jing Jin,
First Affiliated Hospital of Guangzhou Medical University, China

*CORRESPONDENCE
Chongming Jiang
chongming.jiang@terasaki.org

RECEIVED 13 April 2024
ACCEPTED 20 June 2024
PUBLISHED 08 July 2024

CITATION
Shen A, Garrett A, Chao C-C, Liu D, Cheng C, Wang Z, Qian C, Zhu Y, Mai J and Jiang C (2024) A comprehensive meta-analysis of tissue resident memory T cells and their roles in shaping immune microenvironment and patient prognosis in non-small cell lung cancer.
Front. Immunol. 15:1416751.
doi: 10.3389/fimmu.2024.1416751

A comprehensive meta-analysis of tissue resident memory T cells and their roles in shaping immune microenvironment and patient prognosis in non-small cell lung cancer

Aidan Shen[1], Aliesha Garrett[1], Cheng-Chi Chao[2], Dongliang Liu[3], Chao Cheng[4], Zhaohui Wang[1], Chen Qian[5], Yangzhi Zhu[1], Junhua Mai[6] and Chongming Jiang[1]*

[1]Department of Precision Medicine, Terasaki Institute for Biomedical Innovation, Los Angeles, CA, United States, [2]Department of Pipeline Development, Biomap, Inc., San Francisco, CA, United States, [3]Michael E. DeBakey Department of Surgery, Baylor College of Medicine, Houston, TX, United States, [4]Department of Medicine, Baylor College of Medicine, Houston, TX, United States, [5]Department of Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States, [6]Department of Nanomedicine, Houston Methodist Research Institute, Houston, TX, United States

Tissue-resident memory T cells (T_{RM}) are a specialized subset of long-lived memory T cells that reside in peripheral tissues. However, the impact of T_{RM}-related immunosurveillance on the tumor-immune microenvironment (TIME) and tumor progression across various non-small-cell lung cancer (NSCLC) patient populations is yet to be elucidated. Our comprehensive analysis of multiple independent single-cell and bulk RNA-seq datasets of patient NSCLC samples generated reliable, unique T_{RM} signatures, through which we inferred the abundance of T_{RM} in NSCLC. We discovered that T_{RM} abundance is consistently positively correlated with CD4+ T helper 1 cells, M1 macrophages, and resting dendritic cells in the TIME. In addition, T_{RM} signatures are strongly associated with immune checkpoint and stimulatory genes and the prognosis of NSCLC patients. A T_{RM}-based machine learning model to predict patient survival was validated and an 18-gene risk score was further developed to effectively stratify patients into low-risk and high-risk categories, wherein patients with high-risk scores had significantly lower overall survival than patients with low-risk. The prognostic value of the risk score was independently validated by the Cancer Genome Atlas Program (TCGA) dataset and multiple independent NSCLC patient datasets. Notably, low-risk NSCLC patients with higher T_{RM} infiltration exhibited enhanced T-cell immunity, nature killer cell activation, and other TIME immune responses related pathways, indicating a more active immune profile benefitting from immunotherapy. However, the T_{RM} signature revealed low T_{RM} abundance and a lack of prognostic association among lung squamous cell carcinoma patients in contrast to adenocarcinoma, indicating that the two

NSCLC subtypes are driven by distinct TIMEs. Altogether, this study provides valuable insights into the complex interactions between T_{RM} and TIME and their impact on NSCLC patient prognosis. The development of a simplified 18-gene risk score provides a practical prognostic marker for risk stratification.

KEYWORDS

tissue resident memory T cell, non-small-cell lung cancer, prognosis, tumor immune microenvironment, machine learning

Introduction

Non-small cell lung cancer (NSCLC) accounts for ~85% of lung tumors in adults and is a leading cause of death. Various immune cell populations are present within the NSCLC tumor-immune microenvironment (TIME) (1). Among them, tissue-resident memory T (T_{RM}) cells are a unique subset of T cells that permanently reside within tissues (2–5). Associated with cell surface markers including CD69, ITAG1 (CD49a), and ITGAE (CD103), T_{RM} are non-circulating memory T-cells residing in various tissues that provide an intrinsic defense system against antigens.

T_{RM} are characterized by the expression of tissue-specific homing molecules and immune exhaustion markers (6). T_{RM} cells could play a critical role in anti-tumor immune responses by either directly attacking cancer cells or indirectly promoting the recruitment of activated cytotoxic T cells to the tumor site (5, 7–9). In addition, T_{RM} cells exhibit transcriptional programs associated with tissue-resident memory and display characteristics of tumor neoantigen-specific T cells (10). Targeting T_{RM} cells for potential enhancement of immunotherapies has also been proposed (11). Studies have demonstrated that TRM cells can synergize with checkpoint inhibitors to improve anti-tumor responses. Researchers have explored strategies like adoptive T_{RM} cell transfer, inducing T_{RM} cell accumulation within tumors using cytokines like IL-33, promoting T cells to express homing receptors for tumor localization, and combining T_{RM} cell-targeting approaches with cancer vaccines (12, 13). These strategies aim to harness the localized tumor surveillance and rapid response capabilities of T_{RM} cells, potentially leading to improved efficacy of immunotherapies and durable anti-tumor immunity (13–15). However, how to evaluate the T_{RM} abundance in the TIME of NSCLC patients, and the role of T_{RM} in the TIME to affect the tumor progression and patient prognosis are still unclear. Therefore, a comprehensive understanding of how T_{RM} cells shape the NSCLC TIME and a robust gene signature for assessing T_{RM}-related influence and prognosis across independent patient cohorts is much needed.

In this study, we comprehensively analyzed all publicly available single-cell datasets to extract T_{RM}-related signatures representative of T_{RM} abundance in the tumor milieu. A systematic evaluation of available NSCLC single-cell and bulk RNA-seq datasets revealed the relationship of the T_{RM} with various immune cell populations in TIME and with patient prognosis. A universal 18-gene risk score derived from T_{RM} signatures across independent datasets stratified low- and high-risk NSCLC patients, distinguishing their survival. Among NSCLC patients, the T_{RM} signature has significant prognostic value for lung adenocarcinoma but not lung squamous cell carcinoma, suggesting that distinct TIME may drive different therapeutic responses in the two lung cancer subtypes.

Materials and methods

Data utilized in this study

The level 3 TCGA RNAseq data and clinical information involving the lung adenocarcinoma (LUAD, n=513) and lung squamous cell carcinoma (LUSC, n=501) datasets were obtained from TCGA on FireBrowse (gdac.broadinstitute.org/). TCGA MAF files for gene mutation analyses were obtained from https://gdc.cancer.gov/about-data/publications/pancanatlas. All genes in which non-silent mutations occurred were considered to be mutated. Total mutation burden was represented as the sum all non-silent mutations in a given TCGA sample. Macrophage regulation scores, leukocyte and lymphocyte infiltration scores, and IFNγ response and TGFβ response scores for TCGA-LUAD and TCGA-LUSC samples were downloaded as a Supplementary File (Supplementary Table 1) from prior work (16). The gene expression of 745 NSCLC patients and related clinical information data were collected from the GSE67639 of the open Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67639).

Curation of immune-related genes (IRGs)

Immune-related genes (IRGs) were obtained from Supplementary Table 6 by Charoentong et al. (17). All genes from immune cells were collected (marker genes attributed to cancer cells were excluded) and combined into a single list of 783 IRGs genes.

Immune cell inference

Immune infiltration scores of six immune cells were calculated using Binding Association with Sorted Expression (BASE) (18), a rank-based gene set enrichment method. Previous publications have detailed and validated immune cell infiltration using this method (18–23). BASE uses immune cell-specific weight profiles and patient gene expression data to infer immune cell infiltration for each patient and immune cell type. The BASE orders genes for a patient's gene expression profile from high to low expression and then uses weights from each immune cell weight profile to weigh the patient's gene expression values. BASE calculates two running sums, one representing the cumulative distribution of the patient's weighted gene expression values (foreground function) and another representing the cumulative distribution of the patient's complementary weighted (1-weight) gene expression values (background function). In the presence of a high amount of infiltrate from a specific immune cell type, the foreground function increases quickly, as the highly expressed genes in a patient's profile tend to be the ones with high immune cell weights, while the background function increases slowly. The maximal absolute difference between the foreground and background functions represents the immune infiltration level and, after a normalization procedure, results in the final immune infiltration score (19, 21). Similarly, BASE was used to calculate single cell-based T_{RM} scores using T_{RM} signatures (see next section).

Generation of T_{RM} signatures

NSCLC single cell RNA-seq datasets from human NSCLC were obtained from previous publications (10, 24). Cluster annotations were also obtained from these publications. For each NSCLC cluster, a list of marker genes was provided by identifying genes that are over-expressed in the corresponding cluster compared to all the other clusters. These cluster-specific marker gene sets were used as T_{RM} signatures. In total, 20 human T_{RM} signatures were defined, including 10 CD8+ sources and 10 generalized T_{RM} signature. The signature gene expression and the proportion of cells expressing these genes in the $T_{RM}8$, $T_{RM}9$, and $T_{RM}12$ signatures were shown in the sub tables of the Supplementary Table 2. Given a NSCLC gene expression dataset, the BASE algorithm was used to calculate sample-specific T_{RM} scores for each signature. The T_{RM} signatures were represented as gene sets without assigning weights to the member genes. In this case, the BASE algorithm degenerated into a method like single-sample Gene Set Enrichment Analysis (ssGSEA) (25). A high T_{RM} score indicates that the corresponding T_{RM} cells are strongly infiltrated into the tumor.

Principal component analysis

Principal component analysis (PCA) was performed using the prcomp R function. Principal component coordinates for each sample were extracted using the factoextra R package (https://github.com/kassambara/factoextra). The percentage associated with each principal component (PC) in PCA is calculated based on the amount of variance that the component accounts for in the original dataset. Mathematically, if there are n principal components with eigenvalues λ1, λ2,…, λn, the percentage for the first principal component (PC1) is calculated as:

$$PC1\ percentage\ = (\lambda 1\ /(\lambda 1\ +\ \lambda 2\ +\ldots +\ \lambda n))*100$$

This percentage represents the proportion of the total variance in the original data that is accounted for by the first principal component. The same calculation is applied to the other PCs as well. The percentage for the second PC (PC2) is calculated as:

$$PC2\ percentage = (\lambda 2\ /(\lambda 1\ +\ \lambda 2\ +\ldots +\lambda n))*100$$

And so on for the remaining components. The percentages for all the PCs should sum up to 100%, as they represent the decomposition of the total variance in the dataset.

Principal component 1 (PC1) is the first principal component, and it accounts for the largest possible amount of variance in the dataset. A high PC1 percentage suggests that a significant portion of the variation in the dataset can be captured by this single component, which simplifies the interpretation of the data and allows for visualization in a lower-dimensional space. PC1 was used to represent T_{RM} infiltration.

Estimation of stromal and immune scores

The gene expression data of LUAD and LUSC tissues in derivation population were downloaded from the Genomic Data Commons (GDC, available at: http://potal.gdc.cancer.gov/) Data Portal. The FPKM (fragments per kilobase of exon per million reads mapped) method was used to quantify gene expression. The expression matrix for estimating the stromal and immune scores was normalized by the ESTIMATE algorithm. Stromal and immune scores were calculated by performing single-sample gene set enrichment analysis. These scores formed the basis for the Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data (ESTIMATE) score (26).

Lasso Cox regression

The TCGA-LUAD dataset was randomly divided into a training and testing set with a 1:1 ratio. The training set was analyzed to identify potential prognostic genes and both the testing set and the entire set were used for validation. First, univariate Cox-proportional hazards regression analysis was used to evaluate the association between the overall survival and the gene expression of the gene set, which including the T_{RM} signatures' genes and 783 IRGs. Genes with a p-value of< 0.05 based on the log-rank test were selected as candidate genes. Second, most minor absolute shrinkage and selection operator (Lasso) Cox regression analysis from the R glmnet package was employed to screen the genes most associated with overall survival in a multivariate model, which resulted in 18

genes (ABAT, AHSA1, BTN2A2, CCL20, CD109, CD200R1, CD70, CLEC17A, FST, GNG7, HSPA4, HVCN1, KIR2DL1, LTK, NEFL, RDX, and SIK1). These 18 genes composed the final risk score, which is described as follows:

$$Riskscore = \sum_{i=0}^{n} \beta_i x_i$$

where βi refers to the coefficients of each gene and x_i represents the expression value of the gene.

Survival analysis

For univariate and multivariate survival analyses, Cox proportional hazards models were calculated using the "coxph" function from the R "survival" package. Survival curves were visualized using Kaplan-Meier curves using the "survfit" function from the R "survival" package. Median immune cell infiltration scores were used to stratify patients into "high" and "low" groups for univariate analyses. For multivariate analyses, an infiltration score of 0 was used as separator to stratify patients into "high" and "low" groups. Differences in survival distributions in each Kaplan-Meier plot were calculated using a log-rank test using the "survdiff" function from the R "survival" package.

Enrichment pathway analysis

The R package fgsea, version 1.26.0 (27), was used to perform GSEA with hallmark pathways from the Human Molecular Signatures Database (MSigDB) (28) to investigate which hallmark pathways were significantly (adjust P value< 0.05).

Statistical analyses

The Spearman correlation coefficient (SCC) was reported for all correlation analyses as the assumptions underlying the Pearson correlation (i.e., normal distribution, homoscedasticity or linearity) were not met. SCC was calculated using the R function cor and significance was assessed using cor.test. The sensitivity and specificity of the diagnostic and prognostic prediction models were analyzed by the ROC curve and quantified based on the area under the ROC curve (AUC). All statistical tests were two-sided and p-values< 0.05 were considered statistically significant. All statistical analyses were performed using R software (version 4.2.0).

Data availability

All data available in this study is publicly available. These data can be found at: gdac.broadinstitute.org/, https://gdc.cancer.gov/about-data/publications/pancanatlas. caintergator.nci.nih.gov, https://cgga.org.cn.

Results

Custom T_{RM} signatures representing the T_{RM} abundance in NSCLC from patient single-cell data

Associated with cell surface markers including CD69, CD49a, and CD103, T_{RM} are non-circulating memory T-cells residing in various tissues that provide an intrinsic defense system against antigens. Developed from circulating effector memory T cells in response to an antigen, T_{RM} undergo rapid proliferation upon reactivation with dual capability of both effector memory T cells and memory T cells.

Binding Association with Sorted Expression (BASE), a rank-based gene set enrichment analysis method (18–23), was performed in conjunction with T_{RM} signatures. Twenty T_{RM} signatures, based on gene expression, were isolated from T_{RM} cell clusters to compile data from various independent human NSCLC single-cell RNA-sequencing (scRNA-seq) cohorts (10, 24) (Figure 1A). The twenty T_{RM} cell cluster gene sets are shown in Supplementary Table 3. This process generated a manageable form of estimated abundance data, validated in our previous studies (14, 20). Patients in TCGA-LUAD were separated into high and low T_{RM} abundance groups (Figure 1B). The twenty T_{RM} signatures have high correlation with each other (Figure 1C).

Significant signatures were chosen for further analysis, and individual signature survival analysis was conducted using Lifelines KaplanMeierFitter to visualize the results. Three T_{RM} signatures (T_{RM}8, T_{RM}9, and T_{RM}12) are most likely correlated with NSCLC patients' prognosis. Higher T_{RM} abundance was correlated with higher survival in the selected T_{RM} signatures (Hazard Ratio<1.0, Supplementary Figure 1A). When isolating the three signatures into two high and low T_{RM} infiltration groups, the outcome is depicted in Figure 1D showing the two groups are divided clearly and distinctly. The selected signatures had positive correlations with key immune cells like T cells, monocytes, memory B cells, naïve B cells, and especially strong correlations with natural killer (NK) cells and CD8+ T cells, in multiple independent cohorts (Figure 1E, Supplementary Figure 2A). Some kinds of these immune cells are always beneficial to patient prognosis, such as the CD8+ T cells (Supplementary Figure 1C). These T_{RM} signatures have very high correlations with each other in multiple independent NSCLC cohorts (Figure 1F, Supplementary Figure 2B).

T_{RM} abundance is associated with the expression of immune checkpoint and stimulatory genes and immune regulatory pathways

Principal component analysis for dimensionality reduction captured the variance present in the highly correlated T_{RM} signatures. The first principal component (PC1) was highly

10.3389/fimmu.2024.1416751

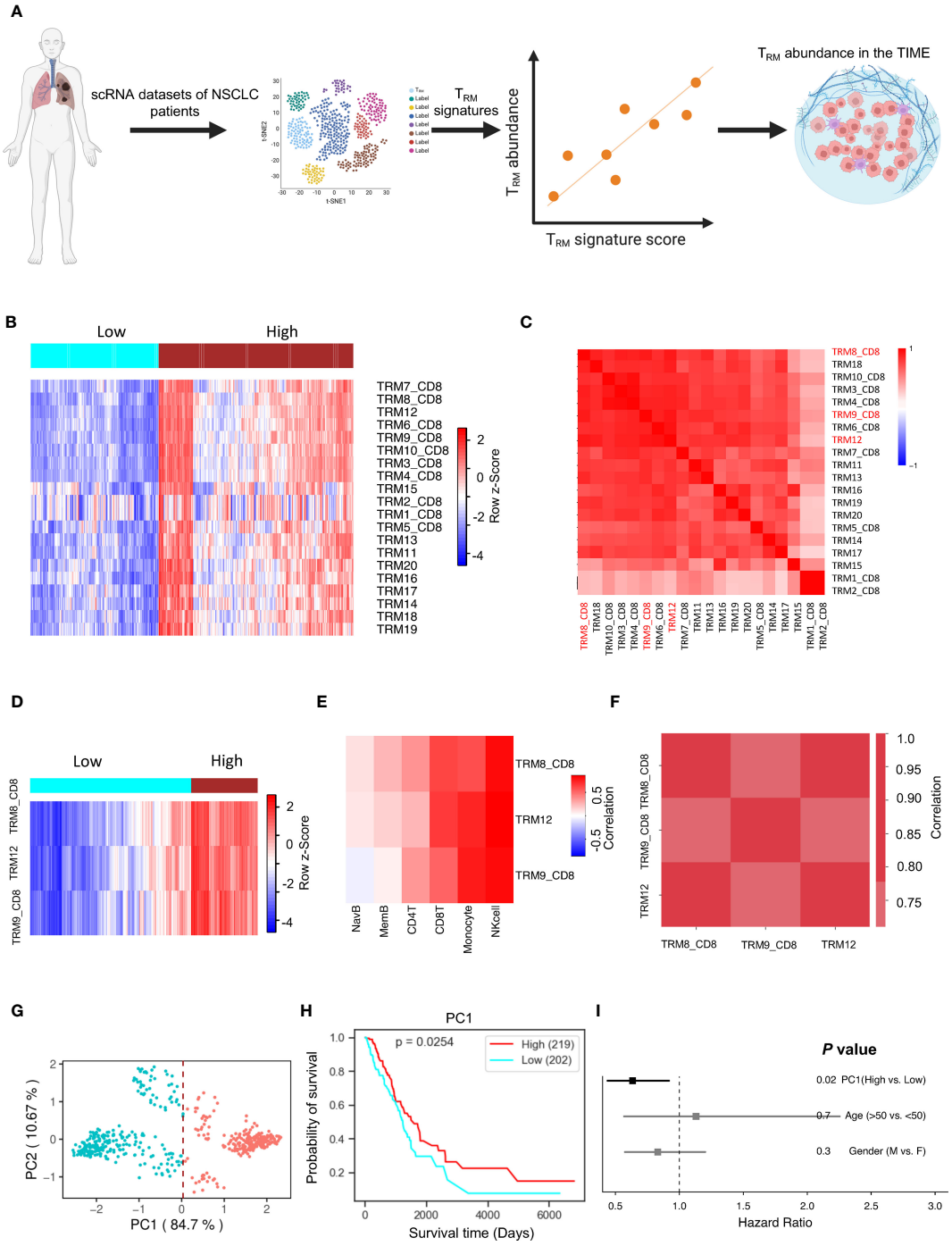

FIGURE 1
T_{RM} cell abundance is positively associated with NSCLC prognosis. The PC1 score can represent T_{RM} signatures, and it can represent the T_{RM} cell proportion in the NSCLC. **(A)** Utilizing multiple independent single-cell RNA-seq data from human NSCLC samples, we crafted 20 distinct NSCLC T_{RM} signatures reflective of T_{RM} infiltration. **(B)** The infiltration distribution of the 20 T_{RM} signatures in patients. **(C)** Heatmap of 20 T_{RM} signatures we found and the correlation of them. **(D)** The infiltration distribution of the T_{RM}8, T_{RM}9, and T_{RM}12 signatures in patients. **(E)** The correlation of T_{RM}8, T_{RM}9, and T_{RM}12 signatures with mainly immune cells (natural killer (NK) cells, CD8+ T cells, monocytes, memory B cells, naïve B cells, and CD4+ T cells) in patients. **(F)** The correlation of the selected T_{RM} signatures with each other. **(G)** Principal Component Analysis (PCA) on the expression of the selected T_{RM} signatures in NSCLC patients. **(H)** Kaplan-Meier plot showing the association between overall survival and the first principle component (PC1) in NSCLC. **(I)** Forest plot depicting hazard ratios of univariate Cox regression models evaluating the association between overall survival and several clinical variables. Figure 1A created with BioRender.com.

correlated with all T_{RM} signatures and captured 84.7% of the variation in patients (Figure 1G, Supplementary Figure 1C), and is better associated with patient prognosis (Figure 1H) than conventional clinical variables including age and gender (Figure 1I). The PC1 could be suitable to represent the T_{RM} abundance (14). These results were then validated in an independent NSCLC dataset, GSE67639 (Supplementary Figures 2C, D).

10.3389/fimmu.2024.1416751

The patients with high PC1 values also had hotter TIME, with higher immune scores, ESTIMATE scores, stromal cell scores, and lower tumor purity, which were evaluated by using the ESTIMATE algorithm (26) (Figures 2A–D, respectively). PC1 is positively correlated with various types of immune cells, immune checkpoint and stimulatory genes (Supplementary Table 4), which were identified and reported in the previous studies (29–32), respectively (Figures 2E, F). Furthermore, PC1 shows positive correlations with many important immune-related pathways, like leukocyte infiltration, lymphocyte infiltration, TCR richness, TCR Shannon, macrophage regulation, stromal cell infiltration, and IFN-γ response (Figures 2G–L). These results corroborate the derived T_{RM} signatures and their association with key immune markers, suggesting that tumors with higher T_{RM} abundance may have more active TIME.

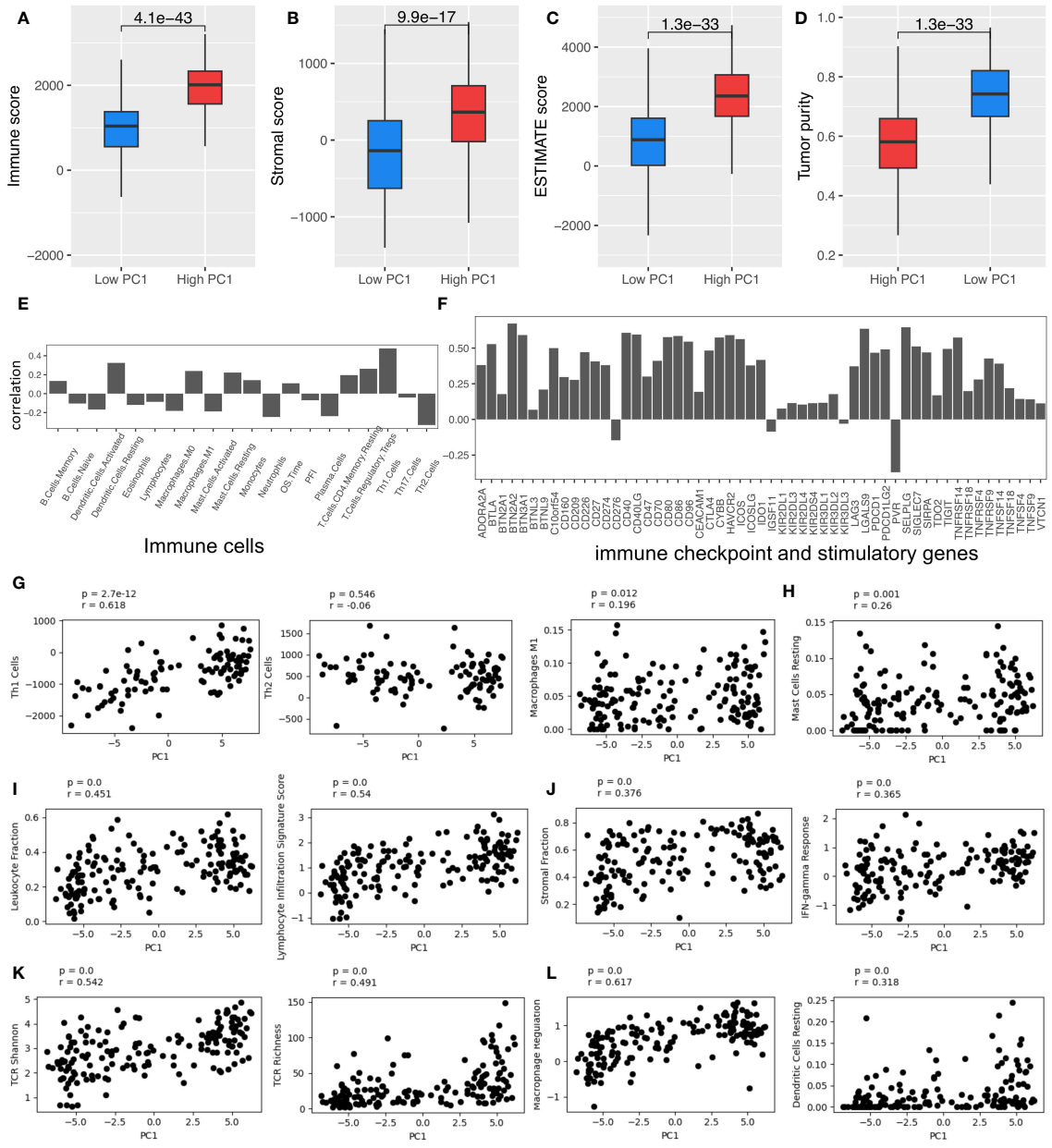

FIGURE 2
T_{RM} cell abundance is associated with the expression of immune checkpoint and stimulatory genes and immune regulatory pathways. **(A)** Immune score; **(B)** Stromal score; **(C)** Estimate score; **(D)** Tumor purity. **(E)** The Spearman correlation coefficient (SCC) between PC1 and immune cells. **(F)** SCC between PC1 and immune checkpoint and stimulatory genes expressed in NSCLC. **(G)** SCC between PC1 and Th1, Th2, and macrophages M1 cells. **(H)** SCC between PC1 and mast cells resting. **(I)** SCC between PC1 and leukocyte and lymphocyte infiltration. **(J)** SCC between PC1 and immune infiltration score. **(K)** SCC between PC1 and TCR Shannon and richness. **(L)** SCC between PC1 and macrophage regulation and dendritic cell (DC) resting. LUAD, lung adenocarcinoma.

A prognostic tool of the risk score from an 18-gene panel for lung adenocarcinoma

Abundance data based on the T_{RM} signatures is formed from a large number of genes composing a final infiltration score. The Lasso Cox regression model identified 18 genes significantly associated with patient survival (Figures 3A, B).

A risk score based on these 18 genes was calculated in the form of $riskscore = \sum_{i=0}^{n} \beta_i x_i$, where corresponds to the weight of each gene and $x_i$corresponded to the expression of that gene in the patients' cancer tissues. The weights of the 18 genes are shown in Supplementary Table 5. A higher weight represented more prognostic significance for that gene. Weighted gene expressions were coalesced into a risk score for each patient (Figure 3A). Each of the 18 genes is correlated with patient survival (Figure 3B), the overall risk score provides a much significant p-value as the low-risk patients had much higher survival than the high-risk patients according to Kaplan-Meier analysis of patients in TCGA-LUAD (p value =3e-11, HR =3.5, Figure 3C, left). The prognostic significance of the risk score was further supported by the independent dataset, GSE67639, where survival once again significantly favors the low-risk patients (Figure 3C, right). More low-risk patients have a high T_{RM} abundance and have survived (Supplementary Figures 3A, B). The predicted AUC reached 0.79, 0.77, and 0.84 in 1, 5, and 10 years, respectively (Figure 3D). Taken together, lower risk patients had significantly higher survival; the 18-gene risk score is a strong independent prognostic risk factor for patients with NSCLC (Figure 3).

Performance of the risk model with respect to clinicopathological factors

The risk model proves effective with respect to patient cohorts separated into male patients and female patients (Figures 4A, B), age over and under 50 (Figures 4C, D), high and low tumor stage progression (Figures 4E, F), and TNM cancer staging (T stages in Figures 4G and H, N stages in Figures 4I and J, and M stages in Figures 4K and L, respectively). T is assigned based on the extent of involvement at the primary tumor site, N for the extent of involvement in regional lymph nodes, and M for distant spread. Furthermore, in the multivariate model, the risk score maintained its significance and significantly outperformed the other clinical variables, including gender, age, stage, and stages T, N, and M (Figure 4M). Therefore, the 18-gene risk model remains an effective prognostic tool when weighed against current clinicopathological factors for patient prognosis.

The risk score correlated with immune cell infiltration and regulatory pathways

The risk model correlated with immune cell infiltration and relevant immune pathways (Figure 5). Low-risk patients have markedly higher infiltration of lymphocytes, mast cells, and memory B cells (p value = 2.0-e4, 1.5e-4, 0.012, respectively), whereas high-risk patients have higher infiltration of neutrophil and macrophages M2 (p value = 0.002 and p value = 0.003, respectively), suggesting more beneficial TIME characteristics for patients with low-risk scores

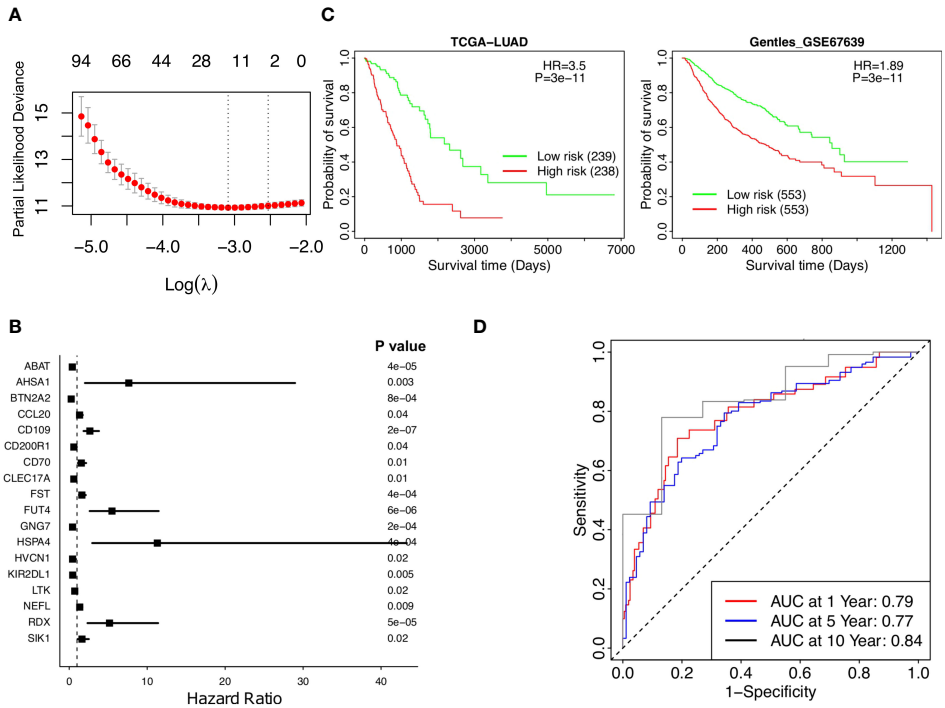

FIGURE 3
Stratified survival analysis of the 18-gene risk score model and Kaplan-Meier survival analysis for the patients in independent datasets by the 18-gene risk score model. **(A)** Development of a T_{RM} risk score for NSCLC patients by Lasso Cox regression analysis. **(B)** The forest plot of the 18 genes in the risk model. **(C)** Patients in the TCGA-LUAD and GSE67639 cohorts. **(D)** Time-dependence of NSCLC in 1, 5, and 10 years, respectively.

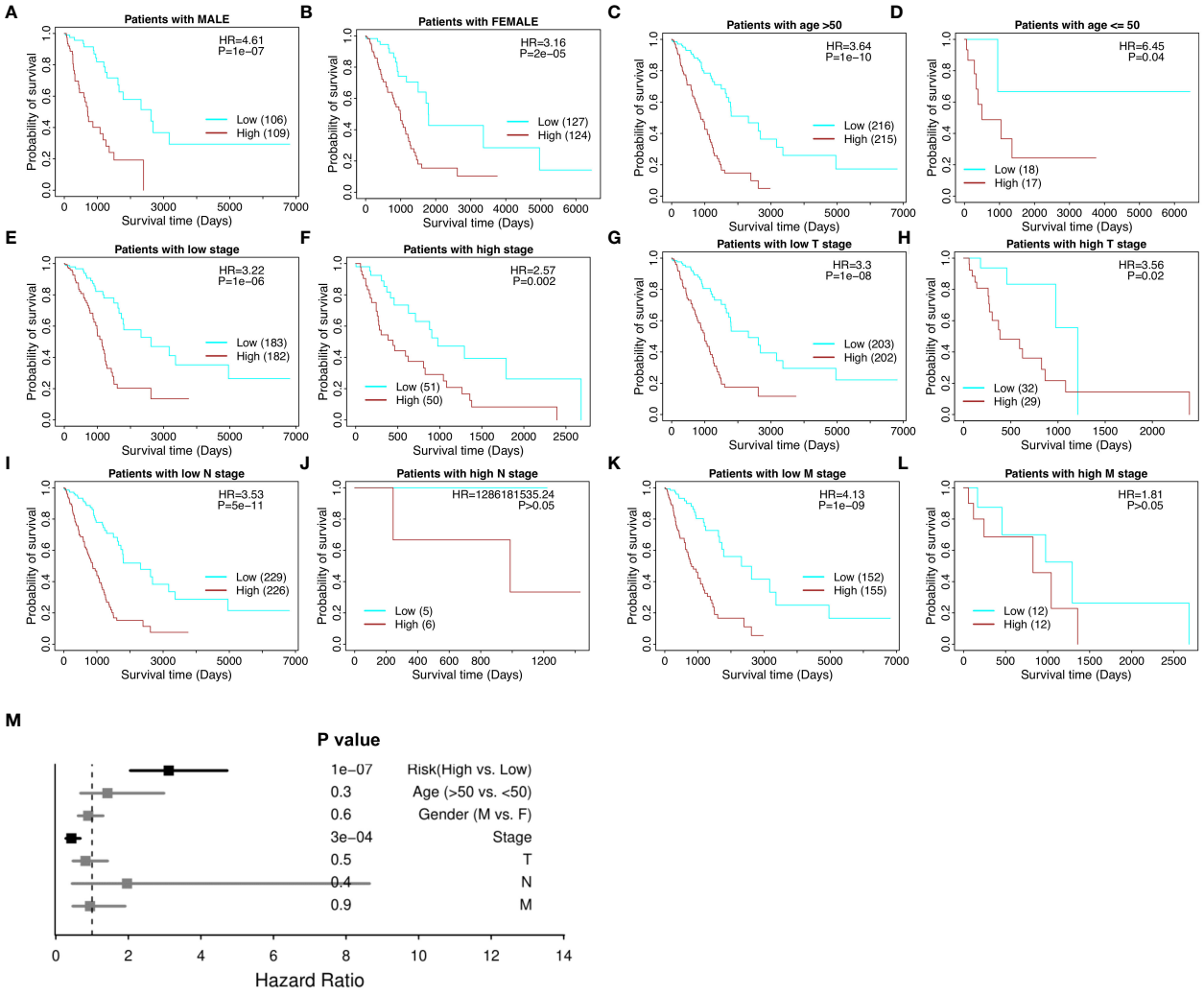

FIGURE 4
Stratified survival analysis of the 18-gene risk score model in clinicopathological factors. **(A)** The risk model in male patients. **(B)** The risk model in female patients. **(C)** The risk model in the elderly (age > 50). **(D)** The risk model in the young (age ≤ 50). **(E)** The risk model in low tumor stage patients. **(F)** The risk model in high tumor stage patients. For the TNM cancer staging system, TNM stands for Tumor, Nodes, and Metastasis. T is assigned based on the extent of involvement at the primary tumor site, N for the extent of involvement in regional lymph nodes, and M for distant spread. **(G)** The risk model in low T stage patients. **(H)** The risk model in high T stage patients. **(I)** The risk model in low N stage patients. **(J)** The risk model in high N stage patients. **(K)** The risk model in low M stage patients. **(L)** The risk model in high M stage patients. **(M)** Multivariate independent prognosis analysis in NSCLC cohort.

(Figure 5A). Notably, low-risk scores correspond to a high number of Th17 cells, but a lower number of Th2 cells and macrophages M2 infiltration (Figures 5B–D, respectively), whereas high-risk scores are positively correlated with TGF-beta response, wound healing, mast cell activation, and tumor proliferation (Figures 5E-H, respectively) and negatively correlated with patients' overall survival time and progression-free interval (PF1) time (Figures 5I, J, respectively).

The patients with low risk scores and high T_{RM} abundance exhibit enrichment for active immune pathways

The 18-gene risk score was further evaluated against various Gene Ontology biological processes (GOBPs) by the gene set enrichment analysis (GSEA) and the associated Molecular Signatures Database (MSigDB) (28). The NSCLC patients with low risk scores and high T_{RM} abundance are associated with the upregulation of various T cells, natural killer cells, and lymphocytes related immune pathways (Figure 6A). Especially, these patients showed a similar correlation with pathways such as T cell and lymphocyte chemotaxis (Figures 6B, C), affirming the connection between PC1, risk, and patient prognosis.

Differential T_{RM} abundance and its prognostic value in lung adenocarcinoma compared with lung squamous cell carcinoma

Lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) are two main subtypes of NSCLC. The T_{RM}

10.3389/fimmu.2024.1416751

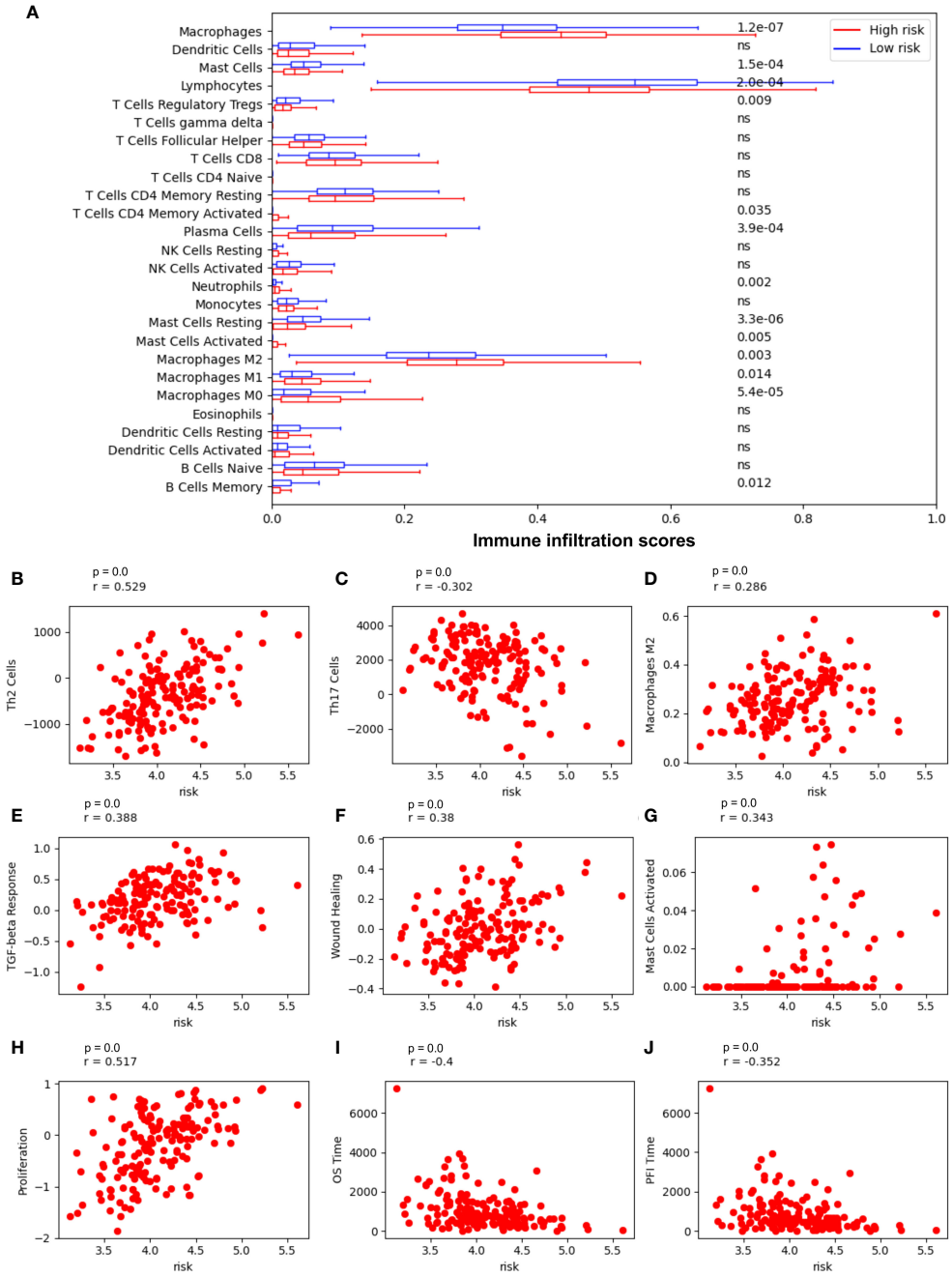

FIGURE 5
Risk model is most associated with immune cells in NSCLC. **(A)** Immune cell infiltration in low-risk vs high-risk patients. **(B)** The Spearman correlation coefficient (SCC) between risk score and Th2 cells; **(C)** SCC between risk score and Th17 cells; **(D)** SCC between risk score and macrophages M2 cells; **(E)** SCC between risk score and the transforming growth factor beta (TGFβ) response; **(F)** SCC between risk score and wound Healing; **(G)** SCC between risk score and mast cells activated; **(H)** SCC between risk score and tumor proliferation; **(I)** SCC between risk score and overall survival (OS) time; **(J)** SCC between risk score and Progression-Free Interval (PFI) time.

abundance are positively correlated with various anti-tumor immune cells in both subtypes, but its correlation with B cell mediated immunity pathways are noticeably different between LUAD and LUSC (Figure 7A, Supplementary Figure 4). In contrast to LUAD (Figure 3C), neither the T_{RM} PC1 nor the 18-gene risk score was able to distinguish any difference in survival among LUSC patients (Figures 7B, C). Furthermore, the T_{RM} score distinguished prognostic differences for both smoker or non-smoker populations respectively in LUAC (Figure 7D) but not in LUSC (Figure 7E). The T_{RM} marker genes expression seem to be much lower in LUSC than in LUAD patients (Figure 7F). The immune score, stromal score, and ESTIMATE score (26), indicators of how hot or cold the TIME is, are also lower in LUCS than LUAD (p = 1e-18, 2.5e-11, and 6.1e-17, respectively, Figures 7G).

10.3389/fimmu.2024.1416751

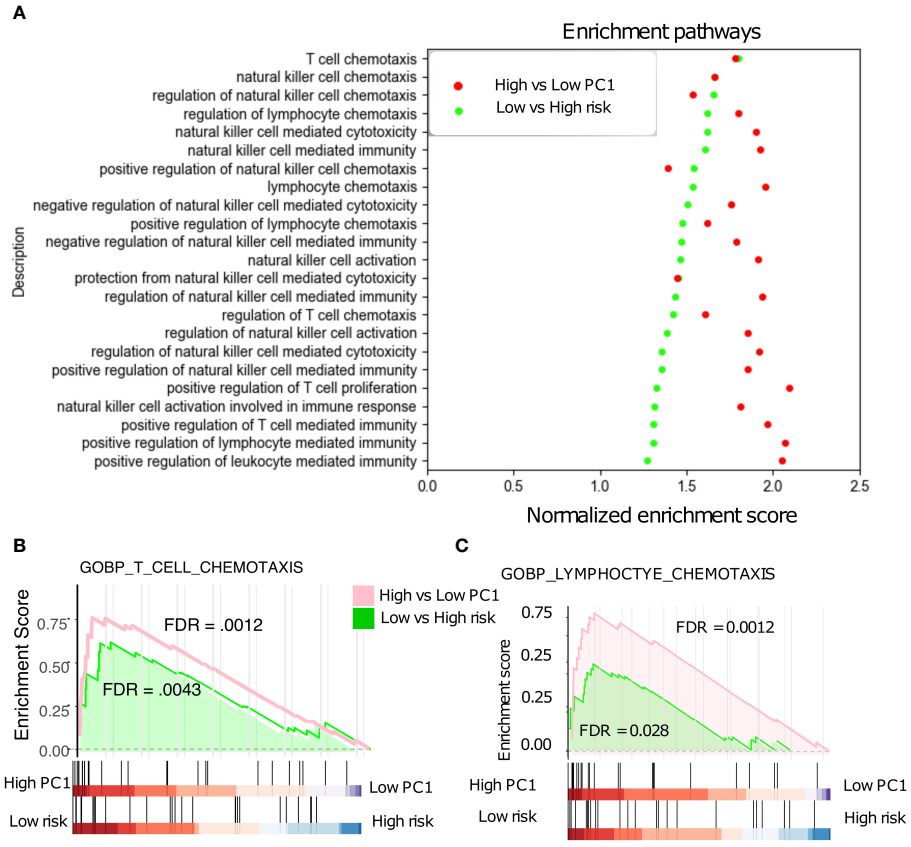

FIGURE 6
Gene Set Enrichment Analysis (GSEA). **(A)** Low-risk patients with significant up-regulated T cell, NK cell, and Lymphocyte -related pathways in the NSCLC. **(B)** T cell chemotaxis pathways are significantly up-regulated in low- vs high- risk patients. **(C)** Lymphocyte chemotaxis pathways are also significantly up-regulated in low- vs high- risk patients.

Discussion

T_{RM} are a specialized population of T cells that reside in peripheral tissues, especially in the lung and skin (33, 34). Our comprehensive meta-analysis generated signatures of T_{RM} abundance in NSCLC patients using available single-cell RNA-sequencing (scRNA-seq) data. We provide evidence that T_{RM} signatures are indicative of prognosis and immune responses in NSCLC. A higher T_{RM} abundance was correlated with higher survival and better prognostic outcome in NSCLC patients. Furthermore, the T_{RM} signatures demonstrated strong correlations with the presence of immune cells such as the CD8+ T cells and NK cell in the TIME, which are known to impact patient prognosis (35). Higher T_{RM} abundance in the TIME is associated with higher degree of immune infiltration and 'hotter' TIME (Figure 3). Infiltration by leukocytes, lymphocytes, stromal, and DC cells are positively correlated with T_{RM} whereas Th2 cells and M2 macrophages are negatively correlated with T_{RM} in NSCLC patients (Figure 2).

An 18-gene risk score for lung adenocarcinoma prognosis was developed, which are associated with T cell functions and demonstrate significant associations with patient survival (36–49). The risk model has better prognostic associations than various clinicopathological factors, such as the gender, age, and stages. Similar patterns and immune regulation results were observed in low- vs high- risk patients and high- vs low- T_{RM} abundance patients (Figure 6). The roles of most genes in activation, metabolism, regulation, inflammation of T cells, and other immune functions, have been solidly established in literature (36–49). For example, ABAT-dependent mitochondrial anaplerosis is critical for T cell-mediated inflammation (36), AHSA1 is involved in the T-cell activation pathway and related pathways (37), and BTN2A2 can inhibit the proliferation of CD4 and CD8 T-cells activated by anti-CD3 antibodies, T-cell metabolism, IL2, and IFN-γ secretion (42). CCL20 is responsible for the chemotaxis of dendritic cells (DC), effector/memory T-cells, and B-cells (43). CD109 could activate T cells (44). CD200R1 might play an important role in immunoregulation, which suppress T cell function and inflammation through DC apoptosis and polarization of macrophages toward M2 subtype (45). Tregs stably-expressing CD70 will lost their regulatory functions but activates cytolytic T cells instead (46). FUT4 was reported to be involved in PD-1-related immunosuppression and could affect operable lung adenocarcinoma patient survival. The overexpression of HVCN1 on CD8+ T cells could enhance adoptive T cell transfer immunotherapy (47), and KIR2DL1 plays a unique opposite function in $CD4^+$ T cells when interacting with SHP-2 and/or SHP-1 proteins (48). The other 8 genes, CLEC17A,

10.3389/fimmu.2024.1416751

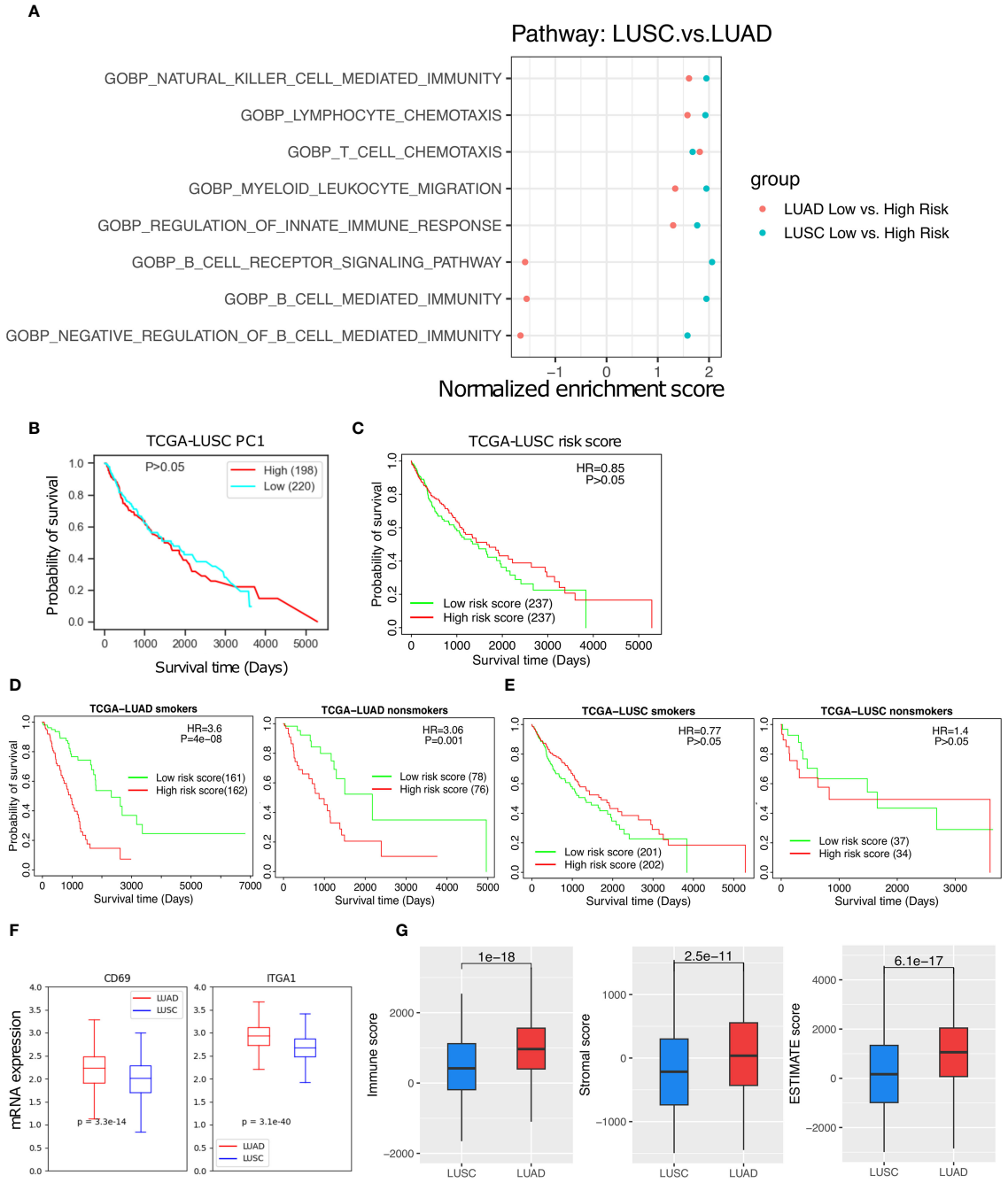

FIGURE 7
The T_{RM} abundance difference between the LUAD and LUSC. **(A)** The T cell, NK cell, and immune response related pathways are significantly up-regulated in low- vs high- risk of LUSC and LUAD patients. **(B)** High PC1 vs low PC1 LUSC patient survival. **(C)** High risk vs low risk in LUSC patient survival. **(D)** High risk vs low risk in LUAD smoker patients and non-smoker patients. **(E)** High risk vs low risk in LUSC smoker patients and LUSC non-smoker patients. **(F)** T_{RM} marker genes expression difference between the LUAD and LUSC patients. **(G)** Immune score, Stromal score, and Estimate score, respectively.

FST, GNG7 (49), HSPA4 (38), LTK (39), NEFL, RDX (40), and SIK1 (41), have been associated with general immune or tumor suppression pathways (38–41, 49). The correlation between immune response and some of the genes in the 18-gene list like CCL20, CD109 and CD200R1 were also identified in other studies (50–52). Given the significant correlations with T_{RM} observed in the current multi-omics study, those genes related to tumor immunity and T cell interaction that have not been well-studied will be applied for future biological investigation. We compared the results of PC1, risk model, and the risk model ROC using $T_{RM8,9,12}$ and $T_{RM9,12}$, respectively. There are no significant differences between them. The performance of the risk model based on $T_{RM8,9,12}$ (AUC = 0.79, 0.77, and 0.84 in 1, 5, and 10 years, respectively) is better than $T_{RM9,12}$ (AUC = 0.77, 0.73, and 0.74 in 1, 5, and 10 years, respectively), for the NSCLC patient prognosis (Figure 3, Supplementary Figure 5).

Surprisingly, the T_{RM} risk model was strongly predictive of prognosis and survival in lung adenocarcinoma but not in lung squamous cell carcinoma at all, although both are categorized as NSCLC. The potential cause could be that the TIME of lung squamous cell carcinoma has lower T_{RM} abundance and a colder TIME than lung adenocarcinoma as our analytical scores indicate. Understanding the nuanced roles of T_{RM} in TIME in lung cancer subtypes may aid the efficacy of emerging treatment approaches (53–55).

Altogether, the study highlights the importance of T_{RM} in the TIME and their potential as a prognostic tool for NSCLC. In addition to cancer, our analytical method may also be applied to understanding the potential role of T_{RM} in other immune-related diseases, such as rheumatoid arthritis, systemic lupus erythematosus, type 1 diabetes, multiple sclerosis, psoriasis, inflammatory bowel disease, autoimmune thyroid disease, etc. (56–58).

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

CJ: Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology. AS: Writing – review & editing, Writing – original draft, Visualization, Formal analysis, Data curation. AG: Writing – review & editing, Investigation. C-CC: Writing – review & editing, Investigation. DL: Writing – review & editing. CC: Writing – review & editing, Investigation, Visualization. ZW: Writing – review & editing, Conceptualization, Investigation. CQ: Writing – review & editing, Conceptualization, Validation. YZ: Writing – review & editing, Investigation. JM: Writing – review & editing, Investigation.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work is supported by the National Institutes of Health, United States (NIH) R01 DK119795 and R35 GM122465.

Acknowledgments

We would like to thank to Dr. Xiling Shen for his valuable discussions and critical feedback. Thank Xiuying Li and Sheng Chang for their useful suggestions.

Conflict of interest

Author C-CC was employed by the company Biomap, Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1416751/full#supplementary-material

References

1. Mirhadi S, Tam S, Li Q, Moghal N, Pham N-A, Tong J, et al. Integrative analysis of non-small cell lung cancer patient-derived xenografts identifies distinct proteotypes associated with patient outcomes. *Nat Commun*. (2022) 13:1811. doi: 10.1038/s41467-022-29444-9

2. Corgnac S, Boutet M, Kfoury M, Naltet C, Mami-Chouaib F. The emerging role of CD8+ Tissue resident memory T (TRM) cells in antitumor immunity: A unique functional contribution of the CD103 integrin. *Front Immunol*. (2018) 9:1904. doi: 10.3389/fimmu.2018.01904

3. Yang L, He Y-T, Dong S, Wei X-W, Chen Z-H, Zhang B, et al. Single-cell transcriptome analysis revealed a suppressive tumor immune microenvironment in EGFR mutant lung adenocarcinoma. *J Immunother Cancer*. (2022) 10:e003534. doi: 10.1136/jitc-2021-003534

4. Oja AE, Piet B, van der Zwan D, Blaauwgeers H, Mensink M, de Kivit S, et al. Functional heterogeneity of CD4+ Tumor-infiltrating lymphocytes with a resident memory phenotype in NSCLC. *Front Immunol*. (2018) 9:2654. doi: 10.3389/fimmu.2018.02654

5. Marceaux C, Weeden CE, Gordon CL, Asselin-Labat M-L. Holding our breath: the promise of tissue-resident memory T cells in lung cancer. *Transl Lung Cancer Res*. (2021) 10:2819–29. doi: 10.21037/tlcr-20-819

6. Amsen D, van Gisbergen KPJM, Hombrink P, van Lier RAW. Tissue-resident memory T cells at the center of immunity to solid tumors. *Nat Immunol*. (2018) 19:538–46. doi: 10.1038/s41590-018-0114-2

7. Aramini B, Masciale V, Samarelli AV, Dubini A, Gaudio M, Stella F, et al. Phenotypic, functional, and metabolic heterogeneity of immune cells infiltrating non–small cell lung cancer. *Front Immunol*. (2022) 13:959114. doi: 10.3389/fimmu.2022.959114

8. Yang G, Cai S, Hu M, Li C, Yang L, Zhang W, et al. Spatial features of specific CD103+CD8+ tissue-resident memory T cell subsets define the prognosis in patients with non-small cell lung cancer. *J Transl Med*. (2024) 22:27. doi: 10.1186/s12967-023-04839-4

9. Nizard M, Roussel H, Diniz MO, Karaki S, Tran T, Voron T, et al. Induction of resident memory T cells enhances the efficacy of cancer vaccine. *Nat Commun*. (2017) 8:15221. doi: 10.1038/ncomms15221

10. Caushi JX, Zhang J, Ji Z, Vaghasia A, Zhang B, Hsiue EHC, et al. *Nature*. (2021) 596:126–32. doi: 10.1038/s41586-021-03752-4

11. Dumauthioz N, Labiano S, Romero P. Tumor resident memory T cells: New players in immune surveillance and therapy. *Front Immunol*. (2018) 9:2076. doi: 10.3389/fimmu.2018.02076

12. Menares E, Gálvez-Cancino F, Cáceres-Morgado P, Ghorani E, López E, Díaz X, et al. Tissue-resident memory CD8+ T cells amplify anti-tumor immunity by triggering antigen spreading through dendritic cells. *Nat Commun*. (2019) 10:4401. doi: 10.1038/s41467-019-12319-x

13. Okla K, Farber DL, Zou W. Tissue-resident memory T cells in tumor immunity and immunotherapy. *J Exp Med*. (2021) 218:1–14. doi: 10.1084/jem.20201605

14. Jiang C, Chao C-C, Li J, Ge X, Shen A, Jucaud V, et al. Tissue-resident memory T cell signatures from single-cell analysis associated with better melanoma prognosis. *iScience*. (2024) 27:109277. doi: 10.1016/j.isci.2024.109277

15. Lange J, Rivera-Ballesteros O, Buggert M. Human mucosal tissue-resident memory T cells in health and disease. *Mucosal Immunol*. (2022) 15:389–97. doi: 10.1038/s41385-021-00467-7

16. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang T-H, et al. The immune landscape of cancer. *Immunity*. (2018) 48:812–830.e14. doi: 10.1016/j.immuni.2018.03.023

17. Charoentong P, Angelova M, Charoentong P, Finotello F, Angelova M, Mayer C, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *CellReports*. (2017) 18:248–62. doi: 10.1016/j.celrep.2016.12.019

18. Cheng C, Yan X, Sun F, Li LM. Inferring activity changes of transcription factors by binding association with sorted expression profiles. *BMC Bioinf*. (2007) 8:1–12. doi: 10.1186/1471-2105-8-452

19. Varn FS, Tafe LJ, Amos CI, Cheng C. Computational immune profiling in lung adenocarcinoma reveals reproducible prognostic associations with implications for immunotherapy. *Oncoimmunology*. (2018) 7:1–10. doi: 10.1080/2162402X.2018.1431084

20. Schaafsma E, Jiang C, Nguyen T, Zhu K, Cheng C. Microglia-based gene expression signature highly associated with prognosis in low-grade glioma. *Cancers (Basel)*. (2022) 14:4802. doi: 10.3390/cancers14194802

21. Varn FS, Wang Y, Mullins DW, Fiering S, Cheng C. Systematic pan-cancer analysis reveals immune cell interactions in the tumor microenvironment. *Cancer Res*. (2017) 77:1271–82. doi: 10.1158/0008-5472.CAN-16-2490

22. Cheng C, Nguyen TT, Tang M, Wang X, Jiang C, Liu Y, et al. Immune infiltration in tumor and adjacent non-neoplastic regions co-determines patient clinical outcomes in early-stage lung cancer. *J Thorac Oncol*. (2023) 18:1184–98. doi: 10.1016/j.jtho.2023.04.022

23. Schaafsma E, Jiang C, Cheng C. B cell infiltration is highly associated with prognosis and an immune-infiltrated tumor microenvironment in neuroblastoma. *J Cancer metastasis Treat*. (2021) 7:0–11. doi: 10.20517/2394-4722.2021.72

24. Clarke J, Panwar B, Madrigal A, Singh D, Gujar R, Wood O, et al. Single-cell transcriptomic analysis of tissue-resident memory T cells in human lung cancer. *J Exp Med*. (2019) 216:2128–49. doi: 10.1084/jem.20190249

25. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL. Gene set enrichment analysis : A knowledge-based approach for interpreting genome-wide. *Proc Natl Acad Sci U.S.A*. (2005) 102:15545–50. doi: 10.1073/pnas.0506580102

26. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. (2013) 4:2612. doi: 10.1038/ncomms3612

27. Korotkevich G, Sukhov V, Budin N, Atryomov MN, Sergushichev A. Fast gene set enrichment analysis. *bioRxiv. bioRxiv*. (2021). doi: 10.1101/060012

28. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. (2011) 27:1739–40. doi: 10.1093/bioinformatics/btr260

29. Hu F, Liu C-J, Liu L, Zhang Q, Guo A. Expression profile of immune checkpoint genes and their roles in predicting immunotherapy response. *Brief Bioinform*. (2021) 22:1–12. doi: 10.1093/bib/bbaa176

30. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*. (2012) 12:252–64. doi: 10.1038/nrc3239

31. Campbell KS, Purdy AK. Structure/function of human killer cell immunoglobulin-like receptors: lessons from polymorphisms, evolution, crystal structures and mutations. *Immunology*. (2011) 132:315–25. doi: 10.1111/imm.2011.132.issue-3

32. Afrache H, Gouret P, Ainouche S, Pontarotti P, Olive D. The butyrophilin (BTN) gene family: from milk fat to the regulation of the immune response. *Immunogenetics*. (2012) 64:781–94. doi: 10.1007/s00251-012-0619-z

33. Szabo PA, Miron M, Farber DL. Location, location, location: Tissue resident memory T cells in mice and humans. *Sci Immunol*. (2019) 4:546. doi: 10.1126/sciimmunol.aas9673

34. Shin H, Iwasaki A. Tissue-resident memory T cells. *Immunol Rev*. (2013) 255:165–81. doi: 10.1111/imr.12087

35. Luoma AM, Suo S, Williams HL, Sharova T, Sullivan K, Manos M, et al. Molecular pathways of colon inflammation induced by cancer immunotherapy. *Cell*. (2020) 182:655–671.e22. doi: 10.1016/j.cell.2020.06.001

36. Kang S, Liu L, Wang T, Cannon M, Lin P, Fan TW-M, et al. GAB functions as a bioenergetic and signalling gatekeeper to control T cell inflammation. *Nat Metab*. (2022) 4:1322–35. doi: 10.1038/s42255-022-00638-1

37. Shao J, Wang L, Zhong C, Qi R, Li Y. AHSA1 regulates proliferation, apoptosis, migration, and invasion of osteosarcoma. *BioMed Pharmacother*. (2016) 77:45–51. doi: 10.1016/j.biopha.2015.11.008

38. Shang B-B, Chen J, Wang Z-G, Liu H. Significant correlation between HSPA4 and prognosis and immune regulation in hepatocellular carcinoma. *PeerJ*. (2021) 9:e12315. doi: 10.7717/peerj.12315

39. Cooper AJ, Sequist LV, Johnson TW, Lin JJ. LTK fusions: A new target emerges in non-small cell lung cancer. *Cancer Cell*. (2022) 40:23–5. doi: 10.1016/j.ccell.2021.12.012

40. Tokunou M, Niki T, Saitoh Y, Imamura H, Sakamoto M, Hirohashi S. Altered expression of the ERM proteins in lung adenocarcinoma. *Lab Invest*. (2000) 80:1643–50. doi: 10.1038/labinvest.3780174

41. Feng S, Wei F, Shi H, Chen S, Wang B, Huang D, et al. Roles of salt−inducible kinases in cancer (Review). *Int J Oncol*. (2023) 63:118. doi: 10.3892/ijo.2023.5566

42. He X, Hu R, Luo P, Gao J, Yang W, Li J, et al. BTN2A2 protein negatively regulates T cells to ameliorate collagen-induced arthritis in mice. *Sci Rep*. (2021) 11:19375. doi: 10.1038/s41598-021-98443-5

43. Wasilko DJ, Johnson ZL, Ammirati M, Che Y, Griffor MC, Han S, et al. Structural basis for chemokine receptor CCR6 activation by the endogenous protein ligand CCL20. *Nat Commun*. (2020) 11:3031. doi: 10.1038/s41467-020-16820-6

44. Mii S, Enomoto A, Shiraki Y, Taki T, Murakumo Y, Takahashi M. CD109: a multifunctional GPI-anchored protein with key roles in tumor progression and physiological homeostasis. *Pathol Int*. (2019) 69:249–59. doi: 10.1111/pin.12798

45. Kotwica-Mojzych K, Jodłowska-Jędrych B, Mojzych M. CD200:CD200R interactions and their importance in immunoregulation. *Int J Mol Sci*. (2021) 22:1602. doi: 10.3390/ijms22041602

46. Arroyo Hornero R, Georgiadis C, Hua P, Trzupek D, He L-Z, Qasim W, et al. CD70 expression determines the therapeutic efficacy of expanded human regulatory T cells. *Commun Biol*. (2020) 3:375. doi: 10.1038/s42003-020-1097-8

47. Navarro F, Casares N, Martín-Otal C, Lasarte-Cía A, Gorraiz M, Sarrión P, et al. Overcoming T cell dysfunction in acidic pH to enhance adoptive T cell transfer immunotherapy. *Oncoimmunology*. (2022) 11:2070337. doi: 10.1080/2162402X.2022.2070337

48. Fourmentraux-Neves E, Jalil A, Da Rocha S, Pichon C, Chouaib S, Bismuth G, et al. Two opposite signaling outputs are driven by the KIR2DL1 receptor in human CD4+ T cells. *Blood*. (2008) 112:2381–9. doi: 10.1182/blood-2007-12-127779

49. Zheng J, Zhang W, Zhang J. Establishment of a new prognostic risk model of GNG7 pathway-related molecules in clear cell renal cell carcinoma based on immunomodulators. *BMC Cancer*. (2023) 23:864. doi: 10.1186/s12885-023-11265-8

50. Yue Y, Tao J, An D, Shi L. Exploring the role of tumor stemness and the potential of stemness-related risk model in the prognosis of intrahepatic cholangiocarcinoma. *Front Genet*. (2022) 13:1089405. doi: 10.3389/fgene.2022.1089405

51. Ma J, Tibbitt CA, Georén SK, Christian M, Murrell B, Cardell L-O, et al. Single-cell analysis pinpoints distinct populations of cytotoxic CD4+ T cells and an IL-10 +CD109+ TH2 cell population in nasal polyps. *Sci Immunol*. (2021) 6:eabg6356. doi: 10.1126/sciimmunol.abg6356

52. Hu P, Ma J, Chen J. A systematic and comprehensive analysis of T cell exhaustion related to therapy in lung adenocarcinoma tumor microenvironment. *Front Pharmacol*. (2023) 14:1126916. doi: 10.3389/fphar.2023.1126916

53. Zheng MZM, Wakim LM. Tissue resident memory T cells in the respiratory tract. *Mucosal Immunol*. (2022) 15:379–88. doi: 10.1038/s41385-021-00461-z

54. Jiang C, Li J, Zhang W, Zhuang Z, Liu G, Hong W, et al. Potential association factors for developing effective peptide-based cancer vaccines. *Front Immunol*. (2022) 13:931612. doi: 10.3389/fimmu.2022.931612

55. Rotrosen E, Kupper TS. Assessing the generation of tissue resident memory T cells by vaccines. *Nat Rev Immunol*. (2023) 23:655–65. doi: 10.1038/s41577-023-00853-1

56. Yenyuwadee S, Sanchez-Trincado Lopez JL, Shah R, Rosato PC, Boussiotis VA. The evolving role of tissue-resident memory T cells in infections and cancer. *Sci Adv*. (2022) 8:eabo5871. doi: 10.1126/sciadv.abo5871

57. Sasson SC, Gordon CL, Christo SN, Klenerman P, Mackay LK. Local heroes or villains: tissue-resident memory T cells in human health and disease. *Cell Mol Immunol*. (2020) 17:113–22. doi: 10.1038/s41423-019-0359-1

58. Yuan R, Yu J, Jiao Z, Li J, Wu F, Yan R, et al. The roles of tissue-resident memory T cells in lung diseases. *Front Immunol*. (2021) 12:710375. doi: 10.3389/fimmu.2021.710375