

RESEARCH

Open Access



# Comprehensive analyses of a CD8<sup>+</sup> T cell infiltration related gene signature with regard to the prediction of prognosis and immunotherapy response in lung squamous cell carcinoma

Liang Chen<sup>1</sup>, Yiming Weng<sup>1</sup>, Xue Cui<sup>1</sup>, Qian Li<sup>1</sup>, Min Peng<sup>1\*</sup> and Qibin Song<sup>1\*</sup>

\*Correspondence:  
mpeng320@whu.edu.cn;  
qibinsong@whu.edu.cn

<sup>1</sup>Cancer Center, Renmin  
Hospital of Wuhan University,  
Wuhan 430060, China

## Abstract

Lung squamous cell carcinoma (LUSC) is associated with a worse prognosis than other histological subtypes of non-small cell lung cancer. Due to the vital role of CD8<sup>+</sup> T cells in anti-tumor immunity, the characterization of CD8<sup>+</sup> T cell infiltration-related (CTLIR) gene signature in LUSC is worthy of in-depth exploration. In our study, tumor tissues of LUSC patients from Renmin Hospital of Wuhan University were stained by multiplex immunohistochemistry to evaluate the density of infiltrated CD8<sup>+</sup> T cells and explore the correlation with immunotherapy response. We found that the proportion of LUSC patients who responded to immunotherapy was higher in the high density of CD8<sup>+</sup> T cell infiltration group than in the low density of CD8<sup>+</sup> T cell infiltration group. Subsequently, we collected bulk RNA-sequencing data from The Cancer Genome Atlas (TCGA) database. The abundance of infiltrating immune cells in LUSC patients was analyzed by using CIBERSORT algorithm, and weighted correlation network analysis was performed to identify the co-expressed gene modules related to CD8<sup>+</sup> T cells. We then developed a prognostic gene signature based on CD8<sup>+</sup> T cell co-expressed genes and calculated the CTLIR risk score, which stratified LUSC patients into high-risk and low-risk groups. With univariate and multivariate analyses, the gene signature was identified as an independent prognostic factor in LUSC patients. The overall survival of LUSC patients in the high-risk group was significantly shorter than that of the low-risk group in the TCGA cohort, which was validated in Gene Expression Omnibus datasets. We analyzed immune cell infiltration in the tumor microenvironment and found fewer CD8<sup>+</sup> T cells and more regulatory T cell infiltration in the high-risk group, which is characterized as an immunosuppressive phenotype. Furthermore, the LUSC patients in the high-risk group were predicted to have a better response to immunotherapy than those in the low-risk group when treated with PD-1 and CTLA4 inhibitors. In conclusion, we performed a comprehensive molecular analysis of the CTLIR gene signature in LUSC and constructed a risk model for LUSC patients to predict prognosis and immunotherapy response.

**Keywords:** Lung squamous cell carcinoma, CD8<sup>+</sup> T cell, Prognosis, Immunotherapy



## Introduction

Over the past few decades, lung cancer has been the leading cause of cancer death worldwide [1]. Lung squamous cell carcinoma (LUSC), which is a distinct histological subtype of non-small cell lung cancer (NSCLC), accounts for approximately 30% of all lung cancer cases [2]. Despite the rapid development of therapeutic approaches, the prognosis of LUSC remains poor, and the 5-year overall survival (OS) remains below 20% [3]. The efficacy of conventional treatment approaches, which include surgical resection, chemotherapy, and radiotherapy, to improve the prognosis of LUSC has reached its limits [4]. Additionally, oncogenic mutations and alterations approved for targeted therapies rarely occur in LUSC [5, 6].

Recently, immunotherapy has become a promising therapeutic strategy for LUSC [7]. Tumor cells exhibit various mechanisms to evade tumor immunosurveillance and suppress anti-tumor immune responses during cancer development and progression. Dysfunction of immune checkpoint molecules is a major mechanism underlying tumor immune evasion. Immune checkpoint blockade (ICB) helps the immune system recognize and attack tumor cells [8]. Immune checkpoint inhibitors targeting PD-1/PD-L1, and CTLA-4 have been approved for the treatment of advanced LUSC [9]. Although ICB has revolutionized cancer therapeutics, they are not effective for all patients. Among the cancer types for which ICB has proven efficacy, potent and durable response has been limited to a subgroup of patients, with several patients demonstrating a lack of initial response to the treatment of ICB. In fact, only one-third of patients respond to ICB in most cancer types [10]. Previous studies have shown that the objective response rate (ORR) to pembrolizumab is approximately 20% in advanced NSCLC [11, 12]. Several biological prognostic and predictive factors have been identified; however, no single biomarker can perfectly discriminate between responders and non-responders. Therefore, there is an urgent need to develop biological prognostic and predictive factors to identify patients who are most likely to respond to ICB therapy.

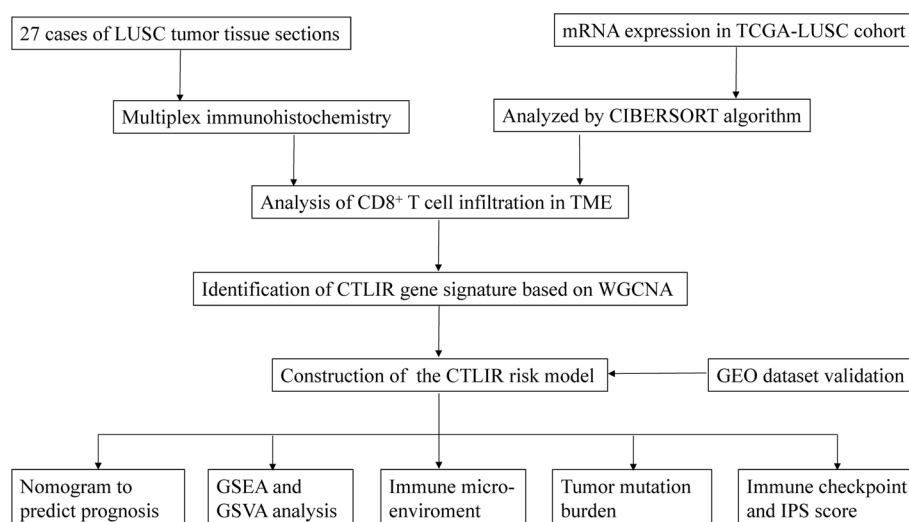
Various types of innate and adaptive immune cells reside within or infiltrate the tumor microenvironment [13]. The dynamic crosstalk between these immune cells and tumor cells defines the immune status of the tumor and can promote or hinder the tumor response to ICB [14]. CD8<sup>+</sup> T cells play a central role in mediating anti-tumor immunity and eliminating tumor cells by recognizing tumor-associated antigens present in major histocompatibility complex class I [15]. CD8<sup>+</sup> effector T cells exert an anti-tumor immune response through the release of cytolytic factors and the induction of apoptosis in tumor cells. The presence of CD8<sup>+</sup> T cells at tumor margins and within the tumor prior to treatment with checkpoint inhibitors is associated with a stronger response to immunotherapy [16]. Increasing evidence has implicated the effect of CD8<sup>+</sup> T cell infiltration on tumor prognosis [17]. Studies have shown that the infiltration of CD8<sup>+</sup> T cells correlates with better prognosis in lung cancer [18]. Based on the abundance of CD4<sup>+</sup> and CD8<sup>+</sup> T cells and their penetration into the tumor, tumor immune profiles can be classified into “cold” or “hot” tumors or more precisely into the immune-inflamed phenotype, the immune-excluded phenotype, or the immune-desert phenotype, which are different in immunotherapy response [19]. Therefore, it is important to clarify the characteristics of the CD8<sup>+</sup> T cell infiltration-related (CTLIR) gene signature in LUSC to provide a guidance for the administration of immunotherapy.

Herein, we identified the co-expression genes related to CD8<sup>+</sup> T cells infiltration through weighted correlation network analysis (WGCNA) and established a risk model as a robust prognostic biomarker and predictive factor for immunotherapy response to enable informed decision-making.

## Materials and methods

### Multiplex immunohistochemistry (mIHC) of tumor tissues from LUSC patients

An overview of the study design is shown in Fig. 1. A total of 27 LUSC samples with pathological sections were collected from the tissue bank in Renmin Hospital of Wuhan University. Formalin-fixed paraffin-embedded blocks were retrieved, and 4  $\mu\text{m}$ -thick sections were prepared for mIHC staining. The tissue sections were immunostained using the PerkinElmer OPAL 7-Colour Manual IHC kit (NEL811001KT) and the following anti-human antibodies: CD8 (Dako, 1:100) and cytokeratin (CK, Dako, 1:200). Cell nuclei were stained with DAPI. Tumor islets and stroma were distinguished using CK staining. Briefly, tissue sections were baked at 65  $^{\circ}\text{C}$  for 1 h, followed by deparaffinization cycles. Microwave treatment was performed for antigen retrieval with EDTA buffer (pH 9.0). The tissue sections were blocked with antibody-blocking buffer for 12 min at room temperature. The sections were then incubated with primary antibodies overnight at 4  $^{\circ}\text{C}$  and then with secondary antibodies for 10 min at room temperature. Tyramide signal amplification was performed using Opal TSA Plus (1:10). Multiplex staining was performed by repeating the staining steps in a series. Fluorescent sections were scanned using the Vectra Polaris 3.0 (Akoya Biosciences, Marlborough, MA, United States of America) with 40 $\times$  magnification (Plan APO 40 $\times$ /NA 0.75, 0.25  $\mu\text{m}/\text{pixel}$ ) and auto-estimated exposure times [20]. The study was approved by the Ethics Committee of Renmin Hospital of Wuhan University (protocol code: WDRY2022-K041) on May 23, 2022.



**Fig. 1** Flow chart of the study design. *LUSC* lung squamous cell carcinoma, *TCGA* The Cancer Genome Atlas, *TME* tumor microenvironment, *CTLIR* CD8<sup>+</sup> T cell infiltration-related, *WGCNA* weighted correlation network analysis, *GEO* Gene Expression Omnibus, *GSEA* gene set enrichment analysis, *GSVA* gene set variation analysis, *IPS* immunophenoscore

### Data collection from public databases

Gene expression profiles and clinical data of LUSC patients were downloaded from the Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>) and Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>) databases. For The Cancer Genome Atlas (TCGA)-LUSC cohort, fragments per kilobase million values were transformed into transcripts per million values. Two expression datasets (GSE30219 and GSE37745) were downloaded from the GEO database. The gene expression information and clinical data of LUSC patients in these datasets were merged and then normalized using the “sva” R package. Only patients with complete relevant information were included in the study. The basic information of LUSC patients from The Cancer Genome Atlas (TCGA) and GEO databases in this study is documented in Additional file 5: Table S1 and Additional file 6: Table S2.

### Analysis of the abundance of infiltrated immune cells

CIBERSORT is an approach for characterizing cell subsets of interest in high-dimensional genomic data derived from bulk tissue samples and can estimate the proportions of 22 distinct immune cell types with gene expression profiles [21]. In this study, we analyzed the abundance of tumor-infiltrating immune cells in LUSC using the CIBERSORT algorithm.

### Identification of CTLIR gene signature based on weighted correlation network analysis (WGCNA)

WGCNA can be used to identify modules of highly correlated genes, summarize such clusters using the module eigengene (ME) or an intramodular hub gene, relate modules to one another and external sample traits using eigengene network methodology, and calculate module membership measures [22]. In our study, WGCNA was used to identify co-expressed gene modules closely related to multiple types of immune cells. Key modules were determined by calculating the Pearson’s correlation coefficient and *P* values of ME and immune cell types. We constructed a gene network and identified modules using the one-step network construction function of the “WGCNA” R package.

### Construction of the CTLIR risk model

Univariate Cox regression analysis was used to identify the prognostic value of the genes in a specific module. Least Absolute Shrinkage and Selection Operator (LASSO) regression was applied for further selection, and the selected genes were put into multivariate Cox regression analysis to construct a risk model for LUSC. Finally, the CTLIR risk score was calculated based on gene expression and multivariate Cox regression coefficient. The formula is as follows:

$$\text{Risk score} = \sum_i^n \text{Exp}_i * \beta_i$$

where *n*, *Exp<sub>i</sub>*, and *β<sub>i</sub>* represent the number of genes, gene expression levels, and regression coefficients, respectively.

### Evaluation of the prognostic value based on the CTLIR risk model

LUSC patients were divided into low- and high-risk groups based on the median risk scores. We utilized the “survival” and “survminer” R packages to draw the

Kaplan–Meier survival curve. We performed univariate and multivariate Cox regression analyses to verify whether the CTLIR risk score was an independent prognostic factor for LUSC patients. The results of univariate and multivariate Cox regression were acquired and visualized using the “forestplot” R package. Additionally, we assessed the accuracy of our risk model in predicting prognosis by applying receiver operating characteristic (ROC) curve and calculating the area under the curve (AUC) values.

#### **Identification of differentially expressed genes (DEGs) and functional enrichment analysis**

DEGs were identified by using the “limma” R package for subsequent analysis ( $P < 0.05$ ,  $|\log_{2}FC| > 1$ ). Gene set enrichment analysis (GSEA) is a calculation method for determining the statistical significance of the previously defined gene set as well as the concordant heterogeneities [23]. We explored the different biological functions between the low- and high-risk groups annotated using the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. The GSEA was performed using the “clusterProfiler” R package. False discovery rate value  $< 0.05$  and normalized enriched score  $> 1$  were considered as the significance thresholds.

The gene set variation analysis (GSVA) is an approach for gene set enrichment to unsupervised estimate pathway activity variations among a certain population [24]. We explored the different activities of pathways related to tumorigenesis and progression in the low- and high-risk groups using GSVA. The analysis was realized by the “GSVA” R package, while  $P < 0.05$  and  $|\log_{2}FC| > 1$  were considered statistically significant. The “clusterProfiler” package was used for functional annotation, and the gene set file (c2.cp.kegg.v7.2. symbols.gmt) was obtained from the MSigDB database (<https://www.gsea-msigdb.org>) [25].

#### **Correlation analysis of the CTLIR risk score with infiltrating immune cells**

The Spearman correlation analysis was used to calculate the correlation coefficient between the CTLIR risk score and 22 types of tumor-infiltrating immune cells. The Estimation of Stromal and Immune cells in Malignant Tumors using Expression data (ESTIMATE) algorithm was used to assess the immune score, stromal score, estimate score, and tumor purity of each LUSC sample [26].

#### **Analysis of tumor mutation burden (TMB)**

The somatic mutation data of LUSC patients in TCGA were downloaded from the Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>). Somatic mutation data sorted in the form of mutation annotation format were analyzed and then used to calculate TMB using the “maftools” R package [27]. Significantly mutated genes in low- and high-risk groups were analyzed by the “maftools” R package. The difference in mutation frequency between the low- and high-risk groups was evaluated using the Chi-square test, and  $P$  value  $< 0.05$  was considered significant.

### Prediction of therapeutic response

The immune checkpoint inhibitor immunophenoscore (IPS) profile of LUSC patients was downloaded from the Cancer Immunome Database (TCIA, <https://tcia.at/home>) [28]. The IPS is a good predictor of the responsiveness to anti-PD-1 and anti-CTLA-4 therapies. In general, a higher IPS indicates a better response to immunotherapy. The IPS is calculated based on representative cell-type gene expression z-scores, with a scale ranging from 0 to 10. In addition, we investigated the predictive capacity of our risk model for the response to chemotherapy and molecular targeted therapy. The 50% inhibiting concentration (IC50) value of the 138 drugs was inferred using the “pRRophetic” R package [29].

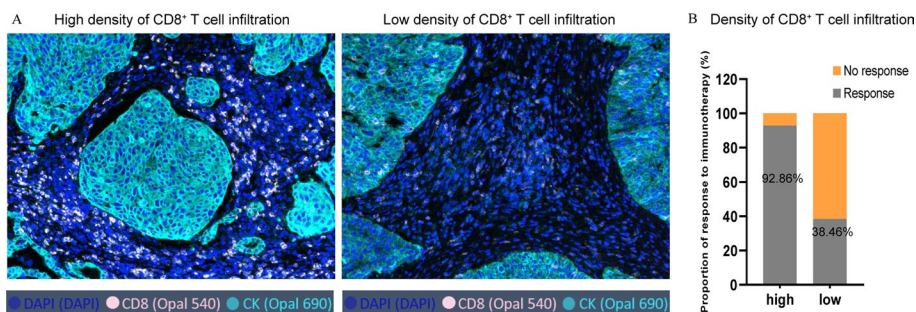
### Statistical analysis

Statistical analysis was conducted using R (v4.1.1) software (<https://www.r-project.org/>). Differences between the two independent groups were evaluated using nonparametric or chi-square tests. The Kaplan–Meier method was used to generate survival curves, and the log-rank test was used to determine statistically significant differences between groups. In our study, the *P* value was two-sided, and  $P < 0.05$  was considered statistically significant.

## Results

### CD8<sup>+</sup> T cell infiltration is related to immunotherapy response

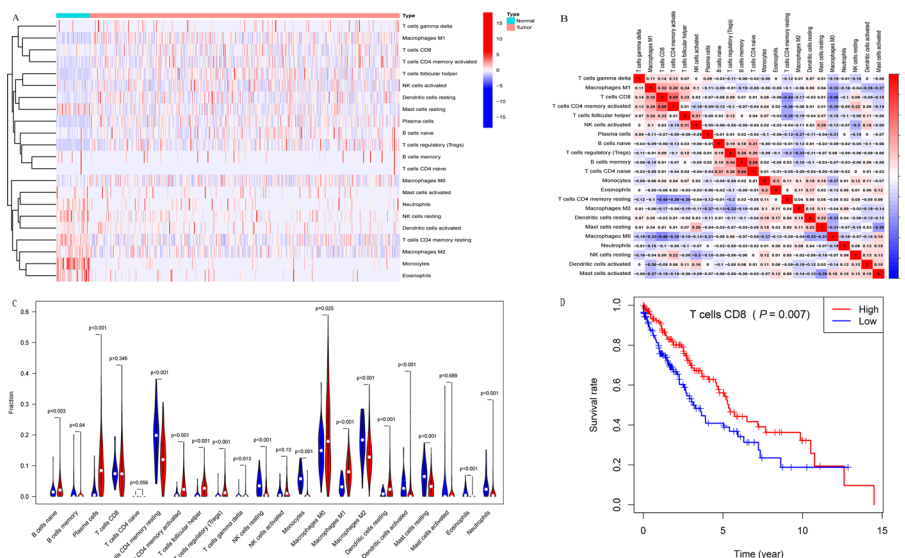
We collected pretreatment tumor tissue sections from 27 patients with LUSC and analyzed the number of infiltrating CD8<sup>+</sup> T cells in tumor tissues using mIHC staining. We then divided the LUSC patients into low density and high density of CD8<sup>+</sup> T cell infiltration groups based on the median number of CD8<sup>+</sup> T cell counts in the tumor tissues (Fig. 2A and Additional file 7: Table S3). On subsequent treatment with PD-1 inhibitors, LUSC patients were considered responsive to immunotherapy if the progression-free survival (PFS) was longer than 4 months. In contrast, LUSC patients were regarded as having no response to immunotherapy if their PFS did not exceed 4 months. Finally, we found that the proportion of LUSC patients responsive to immunotherapy was higher in the high density of CD8<sup>+</sup> T cell infiltration group than in the low density of CD8<sup>+</sup> T cell infiltration group (92.86% vs. 38.46%,  $P < 0.001$ ; Fig. 2B).



**Fig. 2** CD8<sup>+</sup> T cell infiltration is related to immunotherapy response. **A** Representative CD8 staining in tumor tissues of lung squamous cell carcinoma (LUSC) with high or low density of CD8<sup>+</sup> T cell infiltration. **B** Proportion of response to immunotherapy in the high and low density of CD8<sup>+</sup> T cell infiltration groups

### The landscape of infiltrating immune cells in LUSC

A total of 502 LUSC tumor tissues and 49 adjacent normal tissues in the TCGA database were included to analyze the abundance of infiltrating immune cells. The CIBERSORT algorithm was used to identify the fraction of infiltrating immune cells (n=22) in LUSC patients (Additional file 8: Table S4). The abundance of immune cell subtypes in each LUSC normal and tumor tissue is shown in the heatmap (Fig. 3A). Overall, high proportions of monocytes and memory resting CD4<sup>+</sup> T cells were infiltrated in normal tissues, whereas the types of immune cells infiltrated in tumor tissues were diverse. Subsequently, we investigated the correlation between different types of infiltrating immune cells in the tumor tissues. Our results showed that the abundance of CD8<sup>+</sup> T cells positively correlated with activated memory CD4<sup>+</sup> T cells (Cor=0.59), follicular helper T cells (Cor=0.23), gamma delta T cells (Cor=0.14) and macrophages M1 (Cor=0.32) (Fig. 3B). We also found that there were high abundance of CD8<sup>+</sup> T cells, CD4<sup>+</sup> memory resting T cells, and macrophages M0/M1/M2 both in normal and tumor tissues (Fig. 3C). Compared to normal tissues, immune cells, including plasma cells (P<0.001), activated memory CD4<sup>+</sup> T cells (P<0.001), regulatory T (Treg) cells (P<0.001) and macrophages M1 (P<0.001), were more abundant in tumor tissues (Fig. 3C). Although the results showed that the median abundance of CD8<sup>+</sup> T cells was not significant between normal and tumor tissues (P=0.346), the abundance of CD8<sup>+</sup> T cell infiltration was significantly different in various tumor tissues (Fig. 3C and Additional file 1: Fig. S1). Moreover, we found that a higher abundance of infiltrating CD8<sup>+</sup> T cells in tumor tissues was associated with a better prognosis (P=0.007, Fig. 3D). These results indicate the feasibility of risk stratification according to the CTLIR gene signature in LUSC patients.

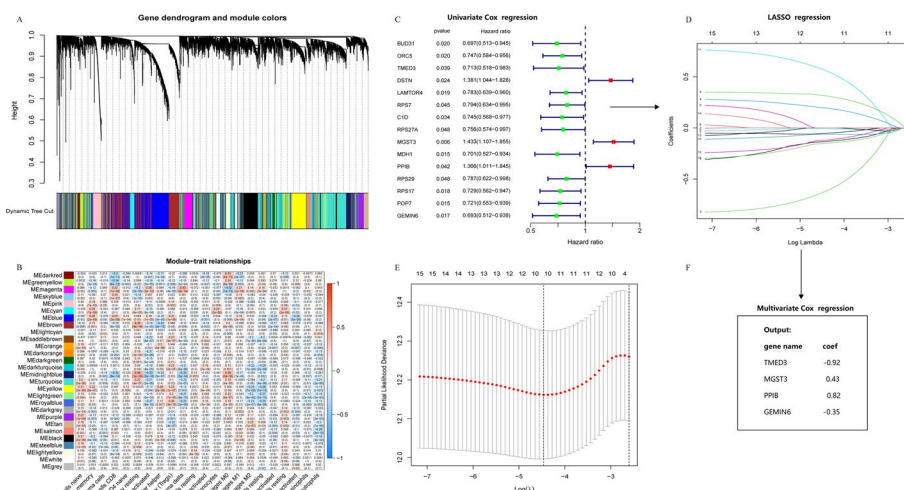


**Fig. 3** Characteristics of infiltrating immune cells in lung squamous cell carcinoma (LUSC). **A** The heat-map of 22 immune cell abundance in normal and tumor tissues of LUSC. **B** Correlation analysis of all 22 immune infiltrating cells in tumor tissues of LUSC. **C** Differences between the immune infiltrating cells in normal and tumor tissues of LUSC. **D** Survival analysis of infiltrated CD8<sup>+</sup> T cells in tumor tissues of LUSC

### Characteristics of the CTLIR risk model

A total of 502 tumor tissues from LUSC patients were enrolled in the WGCNA analysis. The scale-free network condition was satisfied when  $\beta = 5$ , and the scale independence value achieved 0.9 and lower mean connectivity (Additional file 2: Fig. S2A). Genes with similar expression were grouped into the same module by hierarchical clustering and a height of 0.25 was set as the clipping height threshold. Finally, 28 qualified modules were obtained after merging modules with high similarity (Fig. 4A and Additional file 2: Fig. S2B). We observed that the magenta module had the strongest positive correlation with CD8<sup>+</sup> T cell infiltration (Cor = 0.25,  $P = 1e-08$ ) (Fig. 4B). Therefore, the magenta module was identified as the most clinically significant module in the subsequent analyses.

We evaluated the prognostic value of genes in the magenta module using univariate Cox regression analysis (Fig. 4C). Fifteen prognosis-associated genes were further subjected to LASSO and multivariate Cox regression analyses. The most appropriate tuning parameter  $\lambda$  for LASSO regression was 0.03 when the partial likelihood binomial deviance reached its minimum value (Fig. 4D, E). Our results showed that the expression of *MGST3*, *TMED3*, *PP1B*, and *GEMIN6* was an independent prognostic factor of LUSC (Fig. 4F). Thus, four genes were included in the CTLIR risk model. Corresponding coefficients were obtained using multivariate Cox regression. Finally, the risk model was established as follows, CTLIR risk score =  $\text{Exp}(\text{MGST3}) \times 0.43 - \text{Exp}(\text{TMED3}) \times 0.92 + \text{Exp}(\text{PP1B}) \times 0.82 - \text{Exp}(\text{GEMIN6}) \times 0.35$ .



**Fig. 4** The CD8<sup>+</sup> T cell infiltration-related gene signature in lung squamous cell carcinoma (LUSC). **A** Cluster dendrogram of the median absolute deviation (MAD) top 5000 genes. **B** Pearson's correlation coefficients and corresponding  $P$  value between module eigengenes (ME) and the 22 types of immune cells. **C** Analysis of prognosis-associated genes in the selected module. **D** Least Absolute Shrinkage and Selection Operator (LASSO) coefficient profiles of the 15 prognosis-associated genes. **E** Partial likelihood deviance of variables revealed by the LASSO regression model. The red dots represented the partial likelihood of deviance values, the gray lines represented the standard error (SE), and the two vertical dotted lines on the left and right represented optimal values by minimum criteria and 1-SE criteria, respectively. **F** The identified gene signature and the coefficients by multivariate Cox regression analysis



### Prognosis value of the CTLIR risk model for LUSC

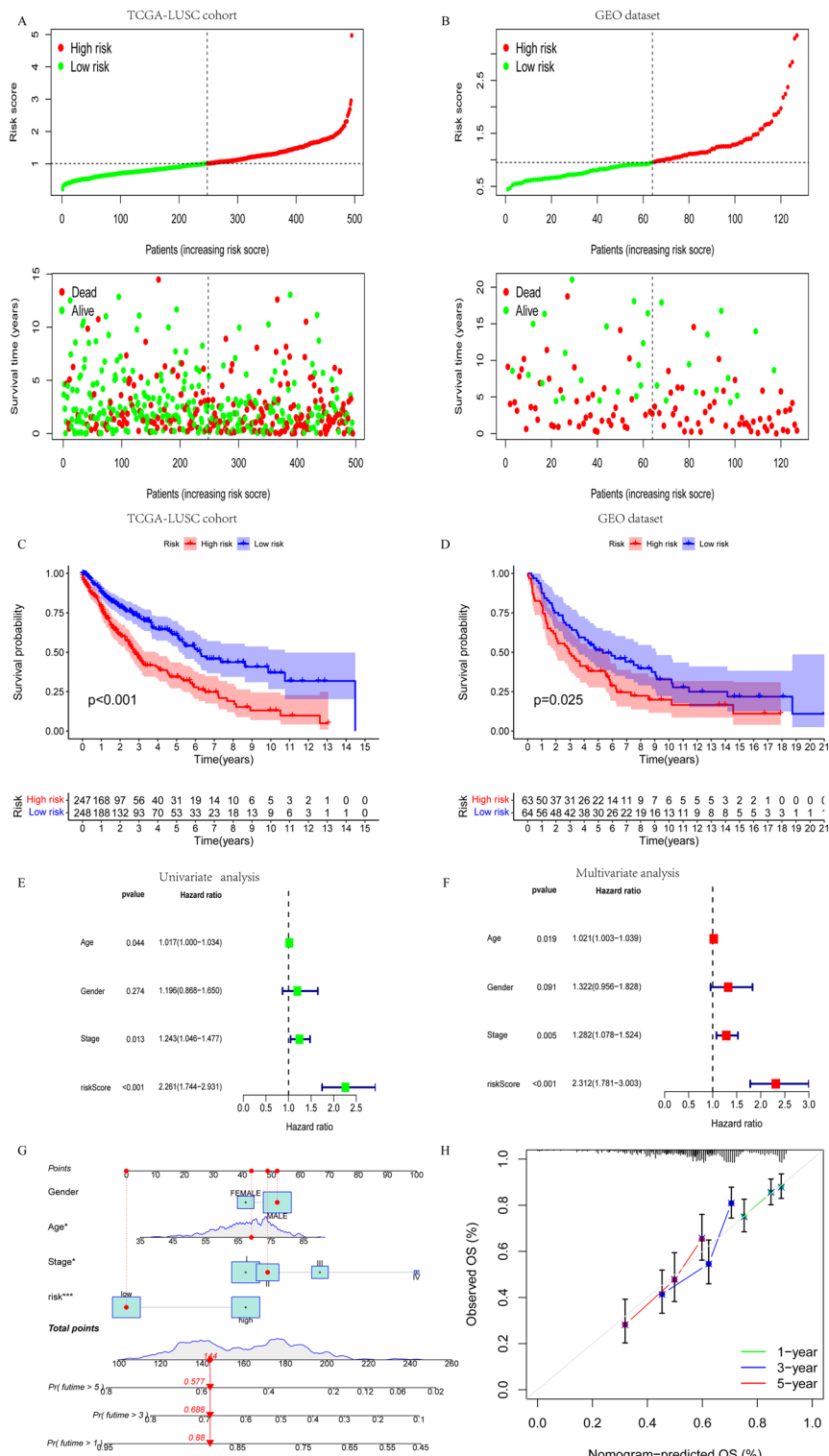
We first ranked the CTLIR risk scores of LUSC patients and analyzed their survival status distribution. Considering the median risk score as the cutoff value, LUSC patients were divided into low- and high-risk groups in both the TCGA-LUSC training cohort and the GEO validation dataset. Our results showed that the proportion of deaths was significantly higher in the high-risk group (Fig. 5A, B). Kaplan–Meier survival curves showed that the OS of LUSC patients in the high-risk group was significantly shorter than that in the low-risk group (log-rank  $P < 0.001$ , Fig. 5C), and the result was validated in the GEO dataset (log-rank  $P = 0.025$ , Fig. 5D). Additionally, an ROC analysis was performed to validate the reliability of this risk model. Based on our risk model, the AUC value of 1-year survival was 0.593, which was the highest value when compared with other clinicopathological factors (Additional file 3: Fig. S3A). The AUC values for 3-year and 5-year survival were 0.654, and 0.675, which demonstrated the good predictive accuracy of the risk model (Additional file 3: Fig. S3B).

We also performed univariate and multivariate Cox regression analyses to explore the effect of risk score and other clinicopathological factors on the prognosis of LUSC patients. The clinicopathological factors included age, gender, and American Joint Committee on Cancer stage. Both univariate ( $P < 0.001$ , Fig. 5E) and multivariate Cox regression ( $P < 0.001$ , Fig. 5F) analyses indicated that the CTLIR risk score was an independent prognostic factor for LUSC patients. Furthermore, we constructed a nomogram generated using the CTLIR risk score and other important clinicopathological traits as a quantitative tool to predict the survival probability of individual LUSC patients (Fig. 5G). Calibration curves revealed ideal consistency of the nomogram in predicting 1-year, 3-year, and 5-year survival rates (Fig. 5H). In summary, these results indicate that our risk model has a certain degree of applicability in predicting the prognosis of LUSC patients.

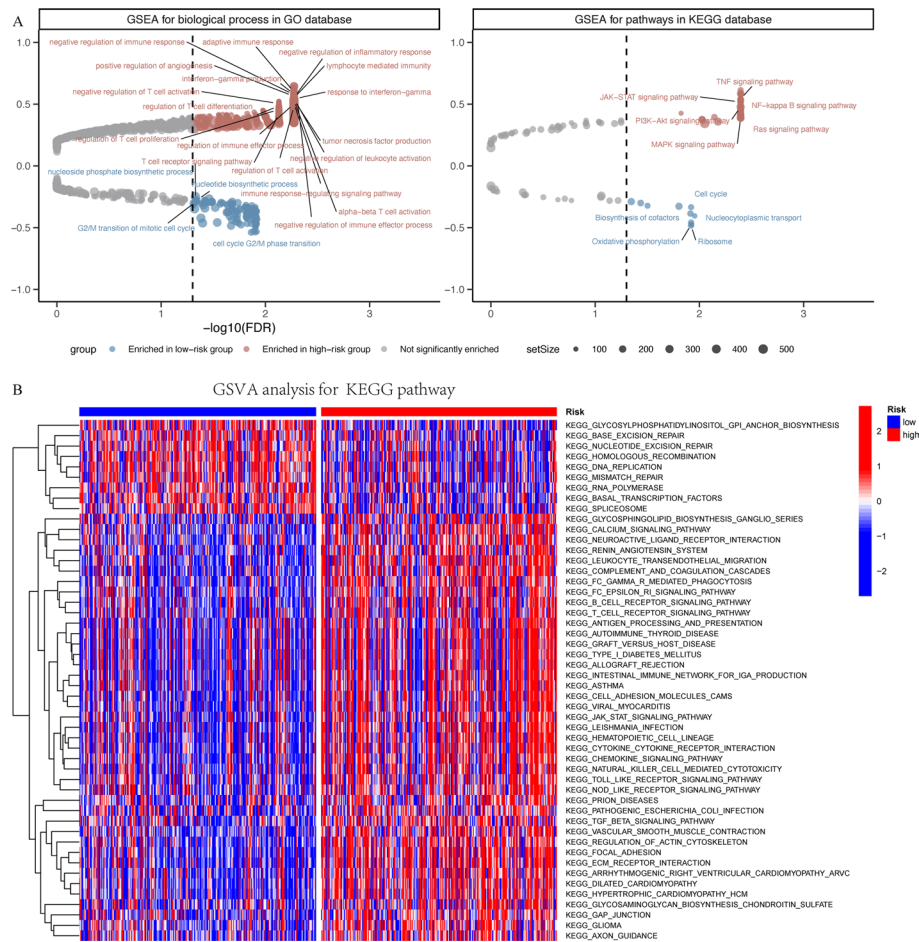
### GSEA and GSVA between low- and high-risk groups

To explore the potential mechanisms leading to different outcomes between the low- and high-risk groups, we performed GSEA with annotations of the GO and KEGG gene sets (Additional file 9: Table S5). According to our results, the gene sets that promote tumor progression, including the JAK-STAT signaling pathway (hsa04630), NF-kappa B signaling pathway (hsa04064), PI3K-Akt signaling pathway (hsa04151), Ras signaling pathway (hsa04014) and MAPK signaling pathway (hsa04010), were significantly enriched in the high-risk group, while the process nucleotide biosynthesis (GO:0009165), regulation of DNA replication (GO:0006275), and cell cycle checkpoint (GO:0000075) were enriched in the low-risk group. Additionally, biological processes associated with the suppression of immune responses, such as negative regulation of immune response (GO:00507770), negative regulation of immune effector processes (GO:0002698), and negative regulation of T cell activation (GO:0050868), were also enriched in the high-risk group (Fig. 6A).

In addition, we conducted GSVA for KEGG pathways based on the differentially expressed genes between the low- and high-risk groups (Additional file 10: Table S6) [30–32]. As shown in our results, the DNA replication and mismatch repair pathways were significantly enriched in the low-risk group. However, many pathways that promote



**Fig. 5** Construction and evaluation of the risk model. **A, B** Risk score distribution and survival status of lung squamous cell carcinoma (LUSC) patients from The Cancer Genome Atlas (TCGA)-LUSC cohort and Gene Expression Omnibus (GEO) dataset. **C, D** Kaplan–Meier survival analysis of risk score for overall survival (OS) in the TCGA-LUSC cohort and GEO dataset. **E, F** Univariate and multivariate analysis of clinical traits and the risk score for OS. **G** Nomogram for predicting the probability of 1-year, 3-year, and 5-year survival in LUSC. **H** Calibration plots of the nomogram

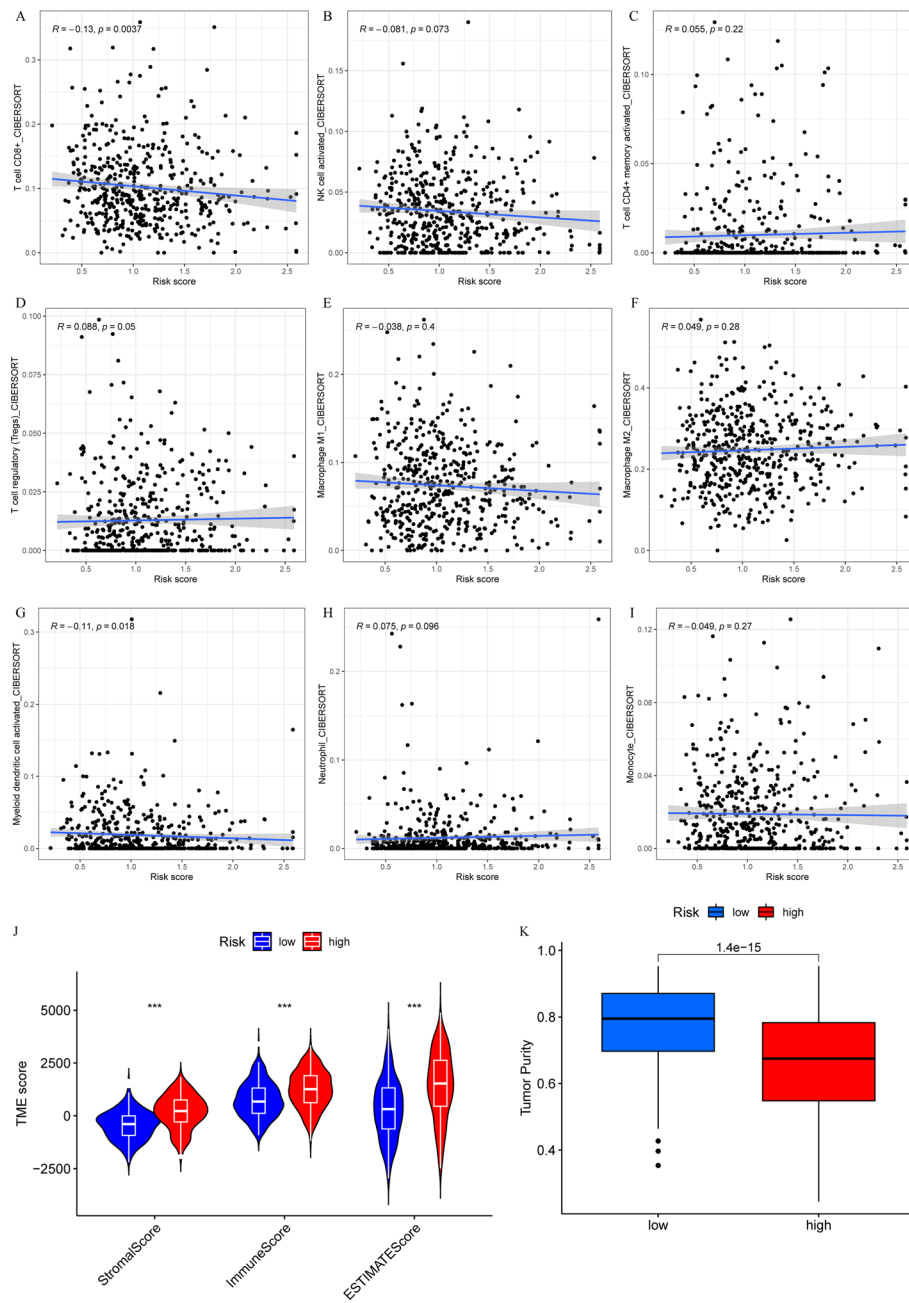


**Fig. 6** Function enrichment analysis in low- and high-risk groups. **A** Gene set enrichment analysis (GSEA) with annotations of the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets. **B** Gene set variation analysis (GSVA) for KEGG pathways

tumor progression, such as JAK/STAT signal pathway and TGF- $\beta$  signal pathway, were significantly enriched in the high-risk group (Fig. 6B). Based on the above analysis, we characterized the high-risk group as a pro-tumorigenic phenotype.

### Correlations between the CTLIR risk score and infiltrating immune cells

To clarify the characteristics of tumor microenvironment (TME), we explored the correlation between the risk score and infiltrating immune cells. Our results indicate that the CTLIR risk score was negatively associated with CD8<sup>+</sup> T cells ( $P=0.0037$ ) and activated dendritic cells ( $P=0.018$ ) and positively associated with Treg cells ( $P=0.05$ ). However, no correlations were found for other major immune cells, such as activated NK cells ( $P=0.073$ ), activated memory CD4<sup>+</sup> T cells ( $P=0.22$ ), macrophages M1 ( $P=0.4$ ), macrophages M2 ( $P=0.28$ ), neutrophils ( $P=0.096$ ), or monocytes ( $P=0.27$ ), with the CTLIR risk score (Fig. 7A–I). Compared with the low-risk group, the high-risk group had higher immune, stromal, and ESTIMATE scores ( $P<0.001$ ; Fig. 7J). Moreover, the tumor purity of the high-risk group was lower than that of the low-risk group ( $P<0.001$ ;



**Fig. 7** A–I Correlation analysis between the risk score and infiltrating immune cells. J Stromal score, immune score, Estimation of Stromal and Immune cells in Malignant Tumors using the expression data (ESTIMATE) score, and K tumor purity in the low- and high-risk groups. \*\*\* $P < 0.001$

Fig. 7K). Taken together, these results suggested that the TME in the high-risk group tended to have an immunosuppressive phenotype.

### TMB features in low- and high-risk groups

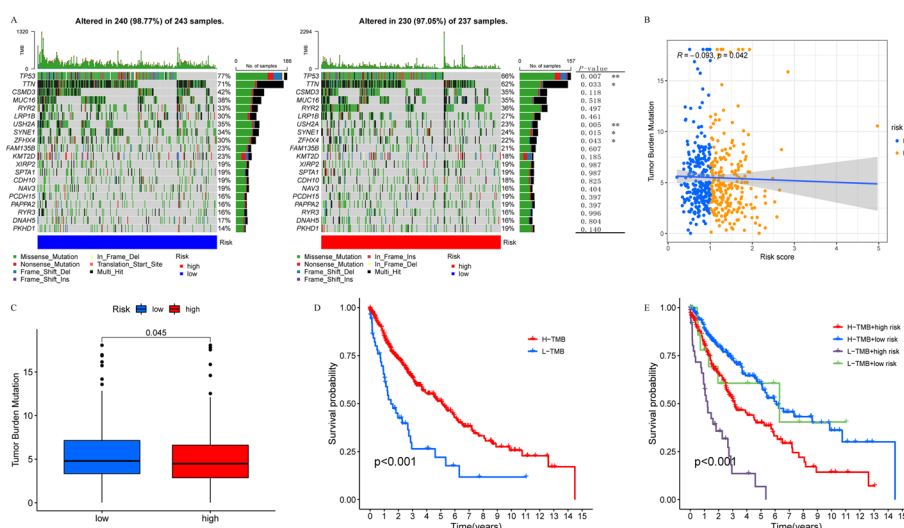
To investigate the correlations between the risk score and TMB, we first explored the differences in the somatic mutation status between the high- and low-risk groups. The

proportions of mutated samples in the two groups were similar (98.77% vs. 97.05%), with TP53 being the most commonly mutated gene in the two groups. We also observed that the mutation frequencies of genes, including TP53, TTN, USH2A, SYNE1, and ZFH4 were significantly different between the two groups ( $P < 0.05$ ; Fig. 8A). Furthermore, we found that the risk score was negatively correlated with the TMB levels (Fig. 8B). The TMB was significantly higher in the low-risk group than in the high-risk group (Fig. 8C), and a high TMB was associated with a better survival outcome (Fig. 8D). In the subgroup analysis, we found that the low-risk and high-TMB subgroups had the most favorable survival, while the high-risk and low-TMB subgroups had the worst prognosis ( $P < 0.001$ ; Fig. 8E). Therefore, our study suggests that the low-risk group is likely to be more immunogenic than the high-risk group.

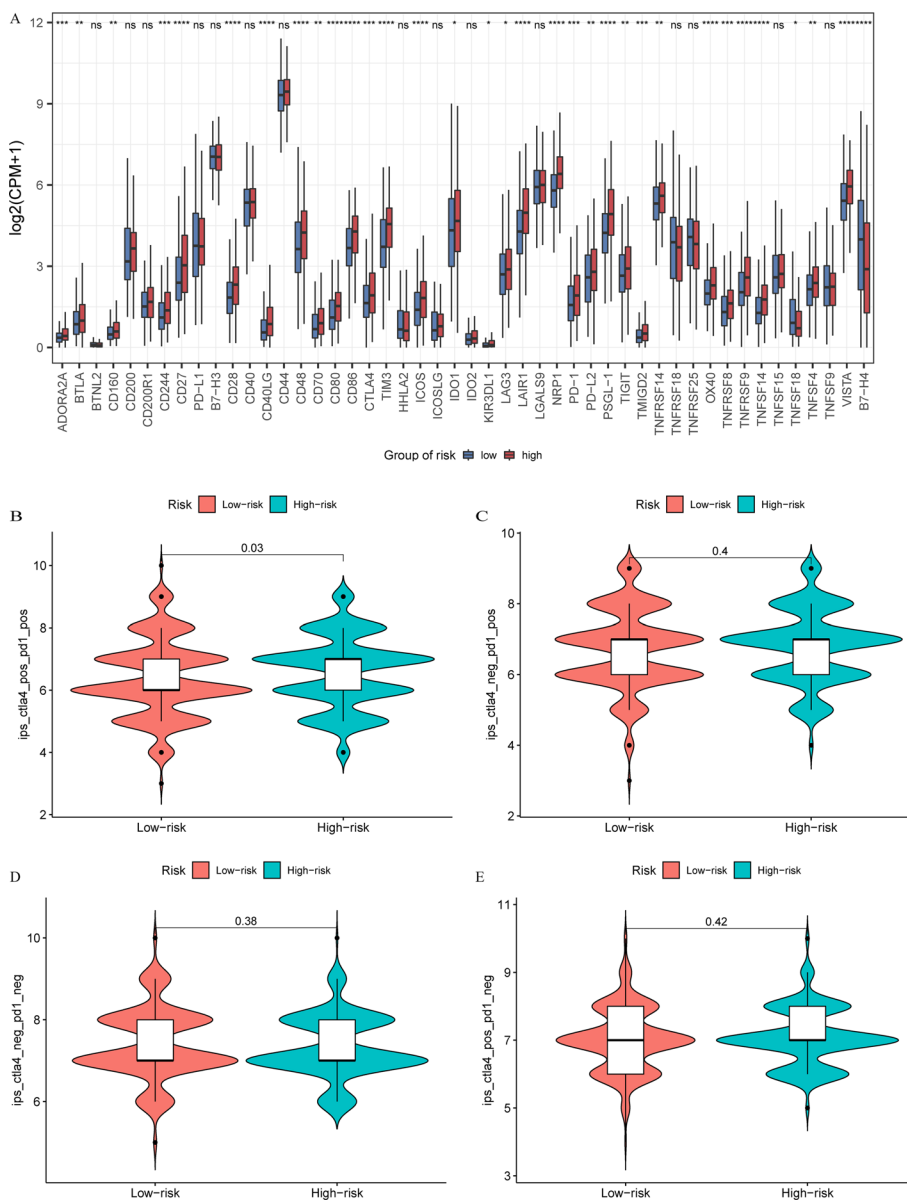
### Significance of the CTLIR risk model in predicting immunotherapy response

To explore the capability of the risk model to predict immunotherapy response in LUSC patients, we compared the gene expression of immune checkpoints in the two groups. Our results showed that the gene expression of 33 immune checkpoints increased in the high-risk group, indicating that immune checkpoint inhibitors may have a pharmacological effect in the high-risk group. It is worth noting that several important immune checkpoints, such as PD-1 and CTLA4, achieved significantly higher expression levels in the high-risk group than in the low-risk group, whereas PD-L1 expression did not differ between the two groups (Fig. 9A).

We obtained the IPS of LUSC patients from the TCIA database to predict their response to immunotherapy (Additional file 11: Table S7). The results showed that the high-risk group had higher IPS-CTLA-4-positive and PD-1-positive scores, which were more sensitive to the combination of CTLA-4 and PD-1 immunotherapy than the low-risk group ( $P = 0.03$ ; Fig. 9B). However, we did not find a difference in the IPS



**Fig. 8** Characteristics of tumor somatic mutation in lung squamous cell carcinoma (LUSC). **A** Waterfall plot of tumor somatic mutation in low- and high-risk groups. **B** Correlation analysis between tumor mutation burden (TMB) and risk score. **C** The TMB level in low- and high-risk groups. **D** Survival analysis of low- and high-TMB groups. **E** Survival analysis stratified by risk score and TMB



**Fig. 9** **A** Gene expressions of immune checkpoints in low- and high-risk groups. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ , ns $P > 0.05$ . **B–E** The immunophenoscore (IPS) in low- and high-risk groups

for other types of immunotherapy between the low- and high-risk groups (Fig. 9C–E). Additionally, we explored the differences in drug sensitivity of potential small-molecule compounds for the treatment of LUSC between the low- and high-risk groups and found that the low-risk group was more sensitive to commonly used chemotherapeutic agents and molecular targeted agents, such as etoposide, vinorelbine, erlotinib, and gefitinib (Additional file 4: Fig. S4A–D). In summary, our results indicate that patients with LUSC in the low-risk group were more likely to be sensitive to chemotherapy and targeted therapy, whereas patients in the high-risk group may have a better response to immunotherapy.

## Discussion

LUSC is an especially challenging disease and is associated with a worse prognosis than other histological subtypes of NSCLC [33]. Treatment approaches for NSCLC include surgery, radiotherapy, chemotherapy, immunotherapy, and molecular targeted therapy, either alone or in combination [34]. Immunotherapy, either as a monotherapy or in combination, has become the standard of care in frontline settings for advanced squamous lung cancer [35]. Although ICB can induce long-term remission, most patients fail to achieve a durable clinical response [36]. This has led to considerable effort to identify robust and sensitive risk models for prognosis and ICB response prediction.

Analysis of the tumor-infiltrating immune cell subtypes revealed that CD8<sup>+</sup> T cells had the strongest positive impact on patients' survival. The positive prognostic value of CD8<sup>+</sup> T cells was confirmed in >18,700 patients across 17 solid cancer types [37]. Tissue microarrays from 335 resected stage I to IIIA NSCLC showed that high densities of CD8<sup>+</sup> T cells in the stroma were independent positive prognostic indicators for resected NSCLC patients. This results suggest that CD8<sup>+</sup> T cells mediate a strong anti-tumor immune response in NSCLC [38]. A previous study investigated whether tumor-infiltrating immune cells in biopsy specimens could be used to predict the clinical outcomes of stage IV NSCLC patients. The results showed that patients with more tumor-infiltrating CD8<sup>+</sup> T cells in the cancer nests had significantly better OS (MST, 388 days vs. 256 days;  $P=0.007$ ) [39]. However, another study conducted by Wakabayashi et al. indicated that CD4<sup>+</sup> T cells in the cancer stroma, but not CD8<sup>+</sup> T cells in cancer cell nests, are associated with a favorable prognosis in NSCLC [40]. It is worth noting that the increased densities of CD8<sup>+</sup> T cells might be associated with more advanced tumors. Consistently, we found that LUSC patients with a high level of CD8<sup>+</sup> T cell infiltration had a favorable prognosis.

Next, we identified a gene signature that was highly co-expressed with CD8<sup>+</sup> T cell infiltration and constructed a risk model to evaluate the prognosis of LUSC patients. In our study, a risk model was established based on four CD8<sup>+</sup> T cell infiltration-related genes associated with the prognosis of LUSC patients, including *TMED3*, *MGST3*, *PPIB*, and *GEMIN6*. According to univariate and multivariate Cox analyses, *TMED3* and *PPIB* had negatively correlated coefficients, and the gene expression level was negatively correlated with the risk score. *MGST3* and *GEMIN6* had positively correlated coefficients, and the gene expression was positively correlated with the risk score. Transmembrane emp24 protein transport domain containing 3 (*TMED3*) belongs to the *TMED* family and plays a role in intracellular protein transport [41]. A previous study showed that the expression of *TMED3* in tumor cells promotes the progression and development of LUSC [42]. However, we found that high expression of *TMED3* in LUSC tissues was associated with good survival, which indicates that *TMED3* expressed in the TME has a protective effect on survival. Peptidylprolyl cis–trans isomerase B (*PPIB*) belongs to the cyclophilin family, which participates in protein folding, secretion, and post-translational modification processes [43]. *PPIB* is associated with malignant progression in gastric cancer, hepatocellular carcinoma, pancreatic cancer, and head and neck squamous cell cancer and is considered a candidate biomarker for these cancers [44]. In our study, LUSC patients with high expression of *PPIB* acquired survival benefits, suggesting that the functions of *PPIB* may be diverse in different types of cancer. Microsomal

glutathione S-transferase 3 (MGST3) is a member of the glutathione S-transferase family, which is involved in detoxification [45]. A previous study showed that overexpression of MGST3 significantly promotes the progression of esophageal squamous cell carcinoma [46]. Gem nuclear organelle-associated protein 6 (GEMIN6) is a component of the survival of the motor neuron (SMN) complex. GEMIN6 is upregulated in LUSC tumor tissues and high expression of GEMIN6 is correlated with poor clinical outcomes [47]. We consistently found that the *MGST3* and *GEMIN6* gene expression is associated with poor survival in LUSC patients.

Based on the risk model, we calculated the CTLIR risk score of patients with LUSC and classified them into low- and high-risk groups with distinct clinical outcomes. Univariate and multivariate Cox analyses revealed that the CTLIR risk score was an independent prognostic factor for LUSC patients. Our results suggest that the LUSC patients in the high-risk group had a worse prognosis than those in the low-risk group. Therefore, we explored the differential molecular mechanisms between low- and high-risk groups. The signaling pathways that promote tumor tumorigenesis and progression, such as JAK-STAT signaling pathway [48], NF-kappa B signaling pathway [49], PI3K-Akt signaling pathway [50], Ras signaling pathway [51], and MAPK signaling pathway [52], were activated in the high-risk group, which may explain why the high-risk group experienced a worse prognosis.

Although the density of immune cell infiltration varies in different tumors, all immune cell components, innate immune cells (macrophages, neutrophils, dendritic cells, innate lymphoid cells, myeloid-derived suppressor cells, and NK cells), and adaptive immune cells (T cells and B cells) are found in the TME [53]. We further explored the association between the CTLIR risk score and immune cells. Our results showed that the CTLIR risk score was negatively correlated with CD8<sup>+</sup> T cell infiltration but positively correlated with Treg cells. The immunosuppressive microenvironment is an important mechanism by which tumor cells escape immune attacks and promote tumor progression. Immune cells with suppressive phenotypes promote tumor escape. These cell types include tumor-associated macrophages, tumor-associated neutrophils, myeloid-derived suppressor cells, and Treg cells [54]. Treg cells suppress antitumor immunity by inhibiting the killing of tumor cells by antigen-specific CD8<sup>+</sup> T cells [55]. Moreover, Treg cells orchestrate memory T cell quiescence by suppressing effector and proliferation programs through the inhibitory receptor CTLA-4 [56]. In consistency with the infiltration of immunosuppressive cells in the TME, GSEA also showed that the biological process associated with the suppression of the immune response was enriched in the high-risk group. Hot tumors are usually characterized by extensive CD8<sup>+</sup> T cell infiltration within the tumor core and at the invasive margin, thereby promoting anti-tumor immunity; cold tumors may show myeloid cell infiltration, but they uniformly show a limited number or absence of CD8<sup>+</sup> T cells [57]. This indicates that the high-risk group may have characteristics of cold tumors. This tumor profile likely reflects the inefficient generation of anti-tumor immunity, which may lead to a poor prognosis.

Tumors are characterized by a high mutational burden, which is thought to increase the occurrence of tumor-associated antigens, thereby promoting immune recognition by tumor cells. Our results indicated that the CTLIR risk score was negatively correlated with the TMB and that patients with a high TMB had better survival. It is plausible



that a higher TMB could act as a prognostic factor for better outcomes, regardless of treatment type. A previous study evaluated the relationship between the TMB and OS in 1415 immunotherapy-naïve patients with diverse advanced malignancies. These results demonstrate that TMB may be a useful prognostic biomarker in immunotherapy-naïve patients, with a protective effect at a higher TMB [58].

Tumor cells can create an immunosuppressive microenvironment by activating immune checkpoint pathways, thus enabling immune escape [59]. Several immune checkpoints have been identified and studied in tumors over recent decades, including but not limited to PD-1 and CTLA-4 [60]. Conventionally, PD-L1 expressed on the surface of tumor cells interacts with PD-1 expressed on the surface of T cells to induce inhibitory signaling [61]. CTLA-4 competes with CD28, a TCR co-stimulatory receptor, to bind ligands, such as CD80 and CD86, which prevents CD28-mediated T cell activation [62]. CD8<sup>+</sup> T effector cells are thought to be the major type of immune cell affected by the PD-1/PD-L1 checkpoint pathway. In contrast, CTLA-4 predominantly regulates the activity of CD4<sup>+</sup> T cells, including the effector and Treg cell subtypes [63]. We found the expression of PD-1 and CTLA-4 in the high-risk group was higher than that in the low-risk group. These findings indicate that immunotherapy is a promising strategy for treating high-risk patients with LUSC.

ICB therapy harnesses the immune system to destroy tumor cells by relieving effector cell dysfunction and inhibiting suppressive immune cell populations [64]. Tumors with high expression of PD-L1 or a high TMB may be more sensitive to ICB monotherapy, whereas combination therapy may be required for tumors with fewer infiltrated CD8<sup>+</sup> T cells and a lower TMB [35]. In our study, we found that LUSC patients in the high-risk group had fewer infiltrating CD8<sup>+</sup> T cells and a lower TMB than those in the low-risk group. However, high-risk LUSC patients were predicted to have a better response to CTLA-4 blocker in combination with PD-1 blocker. Combining different treatment approaches is a promising strategy for overcoming tumor resistance and sensitizing cold tumors to more effective immunotherapies. As observed in experimental models, combining anti-PD-1 and anti-CTLA-4 monotherapies may result in higher and more durable responses in tumors [65]. Clinical trials have revealed that patients with NSCLC treated with nivolumab plus ipilimumab have a tolerable safety profile and show encouraging clinical activity, characterized by a high response rate and durable response [66, 67]. Based on the above results, we believe that our risk model could be an optional tool for predicting the immunotherapy response in LUSC.

This study had several limitations. First, the stability of our risk model was tested and validated using only GEO datasets. Second, the data were primarily derived from public databases and were retrospective. Prospective cohort studies are required to prove the reliability of our risk model. Moreover, biological experiments should be conducted on the discovered gene signature for in-depth characterization of the mechanisms underlying immune responsive regulation.

## Conclusions

In summary, we comprehensively analyzed the CTLIR gene signature in LUSC and constructed a risk model for LUSC patients to predict prognosis and immunotherapy response. LUSC patients identified as a high-risk group based on our risk model

exhibited a worse prognosis and better response to anti-PD-1 in combination with anti-CTLA-4 immunotherapy than the low-risk group. Our research provides new insight for the identification of prognostic biomarkers and prediction of the immunotherapy response in LUSC, which might be helpful for clinical decision-making.

#### Abbreviations

LUSC	Lung squamous cell carcinoma
NSCLC	Non-small cell lung cancer
CTLIR	CD8 <sup>+</sup> T cell infiltration-related
mIHC	Multiplex immunohistochemistry
TCGA	The Cancer Genome Atlas
GEO	Gene Expression Omnibus
WGCNA	Weighted correlation network analysis
OS	Overall survival
PFS	Progression-free survival
TME	Tumor microenvironment
ICB	Immune checkpoint blockade
LASSO	Least Absolute Shrinkage and Selection Operator
GSEA	Gene set enrichment analysis
GSEA	Gene set enrichment analysis
GSA	Gene set variation analysis
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
TMB	Tumor mutation burden
IPS	Immunophenoscore
TCIA	The Cancer Immunome Database

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12859-023-05302-3>.

**Additional file 1: Fig. S1.** The bar charts of 22 immune cells in normal and tumor tissues of LUSC.

**Additional file 2: Fig. S2.** The scale independence and clustering of module eigengenes in WGCNA. (A) Analysis of the scale-free network coefficient R-squared for the soft threshold ( $\beta$ ) and the mean connectivity for the soft threshold. (B) The hierarchical clustering of gene expressions.

**Additional file 3: Fig. S3.** ROC analysis in LUSC patients. (A) ROC analysis of risk score and clinicopathologic factors. (B) ROC analysis for 1-year, 3-year, and 5-year survival based on the CTLIR risk score.

**Additional file 4: Fig. S4.** The differences in drug sensitivity between low and high-risk groups. (A) etoposide, (B) vinorelbine, (C) erlotinib, (D) gefitinib.

**Additional file 5: Table S1.** The clinical information of LUSC patients from TCGA database.

**Additional file 6: Table S2.** The clinical information of LUSC patients from GEO database.

**Additional file 7: Table S3.** The number of infiltrated CD8<sup>+</sup> T cells in LUSC tumor tissues.

**Additional file 8: Table S4.** The CIBERSORT results of LUSC patients from TCGA database.

**Additional file 9: Table S5.** GSEA analysis with annotations of GO and KEGG gene sets in low and high-risk groups.

**Additional file 10: Table S6.** GSEA analysis for KEGG pathways in low and high-risk groups.

**Additional file 11: Table S7.** The IPS score of LUSC patients from TCIA database.

#### Acknowledgements

We would like to thank Editage ([www.editage.cn](http://www.editage.cn)) for English language editing.

#### Author contributions

LC, MP and QS conceived and designed the study. LC, YW and XC contributed to data collection, bioinformatics analysis and interpretation. QL helped with software application. LC and MP completed original draft preparation and writing for the manuscript. QS revised the manuscript and supervised the research. All authors contributed to the article and approved the submitted version.

#### Funding

This work was supported by the National Natural Science Foundation of China (Grant No.81770169), Wu Jieping Medical Foundation Project (320.6750.2021-02-127), Beijing Science and Innovation Medical Development Foundation (KC2021-JX-0186-18), China primary healthcare foundation (cphcf-2022-183) and Renmin hospital cross innovation support project (JCRCFZ-2022-025).

**Availability of data and materials**

The primary data used to support the findings of this study are included within the supplementary information files. The datasets can be found in Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>) and the Gene Expression Omnibus (GEO) (GSE30219 and GSE37745, <https://www.ncbi.nlm.nih.gov/geo/>) databases.

**Declarations****Ethics approval and consent to participate**

Data obtained from studies approved by Renmin Hospital of Wuhan University Ethics Committee. Protocol code: WDRY2022-K041.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

Received: 27 January 2023 Accepted: 24 April 2023

Published online: 06 June 2023

**References**

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–49. <https://doi.org/10.3322/caac.21660>.
- Houston KA, Henley SJ, Li J, White MC, Richards TB. Patterns in lung cancer incidence rates and trends by histologic type in the United States, 2004–2009. *Lung Cancer*. 2014;86(1):22–8. <https://doi.org/10.1016/j.lungcan.2014.08.001>.
- Lu T, Yang X, Huang Y, Zhao M, Li M, Ma K, et al. Trends in the incidence, treatment, and survival of patients with lung cancer in the last four decades. *Cancer Manag Res*. 2019;11:943–53. <https://doi.org/10.2147/CMAR.S187317>.
- Gandara DR, Hammerman PS, Sos ML, Lara PN Jr, Hirsch FR. Squamous cell lung cancer: from tumor genomics to cancer therapeutics. *Clin Cancer Res*. 2015;21(10):2236–43. <https://doi.org/10.1158/1078-0432.CCR-14-3039>.
- Zhao W, Choi YL, Song JY, Zhu Y, Xu Q, Zhang F, et al. ALK, ROS1 and RET rearrangements in lung squamous cell carcinoma are very rare. *Lung Cancer*. 2016;94:22–7. <https://doi.org/10.1016/j.lungcan.2016.01.011>.
- Pan Y, Wang R, Ye T, Li C, Hu H, Yu Y, et al. Comprehensive analysis of oncogenic mutations in lung squamous cell carcinoma with minor glandular component. *Chest*. 2014;145(3):473–9. <https://doi.org/10.1378/chest.12-2679>.
- Miller JF, Sadelain M. The journey from discoveries in fundamental immunology to cancer immunotherapy. *Cancer Cell*. 2015;27(4):439–49. <https://doi.org/10.1016/j.ccell.2015.03.007>.
- Mahoney KM, Rennert PD, Freeman GJ. Combination cancer immunotherapy and new immunomodulatory targets. *Nat Rev Drug Discov*. 2015;14(8):561–84. <https://doi.org/10.1038/nrd4591>.
- Socinski MA, Obasaju C, Gandara D, Hirsch FR, Bonomi P, Bunn PA Jr, et al. Current and emergent therapy options for advanced squamous cell lung cancer. *J Thorac Oncol*. 2018;13(2):165–83. <https://doi.org/10.1016/j.jtho.2017.11.111>.
- Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell*. 2017;168(4):707–23. <https://doi.org/10.1016/j.cell.2017.01.017>.
- Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Five-year survival and correlates among patients with advanced melanoma, renal cell carcinoma, or non-small cell lung cancer treated with nivolumab. *JAMA Oncol*. 2019;5(10):1411–20. <https://doi.org/10.1001/jamaoncol.2019.2187>.
- Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med*. 2015;372(21):2018–28. <https://doi.org/10.1056/NEJMoa1501824>.
- Gajewski TF, Schreiber H, Fu YX. Innate and adaptive immune cells in the tumor microenvironment. *Nat Immunol*. 2013;14(10):1014–22. <https://doi.org/10.1038/ni.2703>.
- Morad G, Helmink BA, Sharma P, Wargo JA. Hallmarks of response, resistance, and toxicity to immune checkpoint blockade. *Cell*. 2021;184(21):5309–37. <https://doi.org/10.1016/j.cell.2021.09.020>.
- Fu C, Jiang A. Dendritic cells and CD8 T cell immunity in tumor microenvironment. *Front Immunol*. 2018;9:3059. <https://doi.org/10.3389/fimmu.2018.03059>.
- Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature*. 2014;515(7528):568–71. <https://doi.org/10.1038/nature13954>.
- Wu J, Zhu Y, Luo M, Li L. Comprehensive analysis of pyroptosis-related genes and tumor microenvironment infiltration characterization in breast cancer. *Front Immunol*. 2021;12:748221. <https://doi.org/10.3389/fimmu.2021.748221>.
- Fridman WH, Pages F, Sautès-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer*. 2012;12(4):298–306. <https://doi.org/10.1038/nrc3245>.
- Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature*. 2017;541(7637):321–30. <https://doi.org/10.1038/nature21349>.
- Stack EC, Wang C, Roman KA, Hoyt CC. Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods*. 2014;70(1):46–58. <https://doi.org/10.1016/j.jymeth.2014.08.016>.
- Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol*. 2018;1711:243–59. [https://doi.org/10.1007/978-1-4939-7493-1\\_12](https://doi.org/10.1007/978-1-4939-7493-1_12).
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform*. 2008;9:559. <https://doi.org/10.1186/1471-2105-9-559>.

23. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–50. <https://doi.org/10.1073/pnas.0506580102>.
24. Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform*. 2013;14:7. <https://doi.org/10.1186/1471-2105-14-7>.
25. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417–25. <https://doi.org/10.1016/j.cels.2015.12.004>.
26. Yoshihara K, Shahmoradgoli M, Martinez 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. <https://doi.org/10.1038/ncomms3612>.
27. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747–56. <https://doi.org/10.1101/gr.239244.118>.
28. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. 2017;18(1):248–62. <https://doi.org/10.1016/j.celrep.2016.12.019>.
29. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS ONE*. 2014;9(9):e107468. <https://doi.org/10.1371/journal.pone.0107468>.
30. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30. <https://doi.org/10.1093/nar/28.1.27>.
31. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci*. 2019;28(11):1947–51. <https://doi.org/10.1002/pro.3715>.
32. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. 2023;51(D1):D587–92. <https://doi.org/10.1093/nar/gkac963>.
33. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ Jr, Wu YL, et al. Lung cancer: current therapies and new targeted treatments. *Lancet*. 2017;389(10066):299–311. [https://doi.org/10.1016/S0140-6736\(16\)30958-8](https://doi.org/10.1016/S0140-6736(16)30958-8).
34. Alexander M, Kim SY, Cheng H. Update 2020: management of non-small cell lung cancer. *Lung*. 2020;198(6):897–907. <https://doi.org/10.1007/s00408-020-00407-5>.
35. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature*. 2018;553(7689):446–54. <https://doi.org/10.1038/nature25183>.
36. Kim TK, Vandsemb EN, Herbst RS, Chen L. Adaptive immune resistance at the tumour site: mechanisms and therapeutic opportunities. *Nat Rev Drug Discov*. 2022;21(7):529–40. <https://doi.org/10.1038/s41573-022-00493-5>.
37. Bruni D, Angell HK, Galon J. The immune contexture and Immunoscore in cancer prognosis and therapeutic efficacy. *Nat Rev Cancer*. 2020;20(11):662–80. <https://doi.org/10.1038/s41568-020-0285-7>.
38. Al-Shibli KI, Donnem T, Al-Saad S, Persson M, Bremnes RM, Busund LT. Prognostic effect of epithelial and stromal lymphocyte infiltration in non-small cell lung cancer. *Clin Cancer Res*. 2008;14(16):5220–7. <https://doi.org/10.1158/1078-0432.CCR-08-0133>.
39. Kawai O, Ishii G, Kubota K, Murata Y, Naito Y, Mizuno T, et al. Predominant infiltration of macrophages and CD8(+) T Cells in cancer nests is a significant predictor of survival in stage IV nonsmall cell lung cancer. *Cancer*. 2008;113(6):1387–95. <https://doi.org/10.1002/cncr.23712>.
40. Wakabayashi O, Yamazaki K, Oizumi S, Hommura F, Kinoshita I, Ogura S, et al. CD4+ T cells in cancer stroma, not CD8+ T cells in cancer cell nests, are associated with favorable prognosis in human non-small cell lung cancers. *Cancer Sci*. 2003;94(11):1003–9. <https://doi.org/10.1111/j.1349-7006.2003.tb01392.x>.
41. Connolly DJ, O'Neill LA, McGettrick AF. The GOLD domain-containing protein TMED1 is involved in interleukin-33 signaling. *J Biol Chem*. 2013;288(8):5616–23. <https://doi.org/10.1074/jbc.M112.403899>.
42. Xie A, Xu X, Kuang P, Zhang L, Yu F. TMED3 promotes the progression and development of lung squamous cell carcinoma by regulating EZR. *Cell Death Dis*. 2021;12(9):804. <https://doi.org/10.1038/s41419-021-04086-9>.
43. Wu J, Zhang W, Xia L, Feng L, Shu Z, Zhang J, et al. Characterization of PPIB interaction in the P3H1 ternary complex and implications for its pathological mutations. *Cell Mol Life Sci*. 2019;76(19):3899–914. <https://doi.org/10.1007/s00018-019-03102-8>.
44. Xin W, Zhao C, Jiang L, Pei D, Zhao L, Zhang C. Identification of a novel epithelial-mesenchymal transition gene signature predicting survival in patients with HNSCC. *Pathol Oncol Res*. 2021;27:585192. <https://doi.org/10.3389/pore.2021.585192>.
45. Uno Y, Murayama N, Kunori M, Yamazaki H. Characterization of microsomal glutathione S-transferases MGST1, MGST2, and MGST3 in cynomolgus macaque. *Drug Metab Dispos*. 2013;41(9):1621–5. <https://doi.org/10.1124/dmd.113.052977>.
46. Liu X, Hong R, Du P, Yang D, He M, Wu Q, et al. The metabolic genomic atlas reveals potential drivers and clinically relevant insights into the etiology of esophageal squamous cell carcinoma. *Theranostics*. 2022;12(14):6160–78. <https://doi.org/10.7150/thno.70814>.
47. Lin J, Liu B, Zhang Y, Lv L, Cheng D, Zhang W, et al. Gemin6 promotes c-Myc stabilisation and non-small cell lung cancer progression via accelerating AURKB mRNA maturation. *Clin Transl Med*. 2022;12(4):e811. <https://doi.org/10.1002/ctm2.811>.
48. Owen KL, Brockwell NK, Parker BS. JAK-STAT signaling: a double-edged sword of immune regulation and cancer progression. *Cancers (Basel)*. 2019. <https://doi.org/10.3390/cancers11122002>.
49. Chen W, Li Z, Bai L, Lin Y. NF-kappaB in lung cancer, a carcinogenesis mediator and a prevention and therapy target. *Front Biosci (Landmark Ed)*. 2011;16(3):1172–85. <https://doi.org/10.2741/3782>.
50. Tan AC. Targeting the PI3K/Akt/mTOR pathway in non-small cell lung cancer (NSCLC). *Thorac Cancer*. 2020;11(3):511–8. <https://doi.org/10.1111/1759-7714.13328>.
51. Young A, Lyons J, Miller AL, Phan VT, Alarcon IR, McCormick F. Ras signaling and therapies. *Adv Cancer Res*. 2009;102:1–17. [https://doi.org/10.1016/S0065-230X\(09\)02001-6](https://doi.org/10.1016/S0065-230X(09)02001-6).

52. Ciuffreda L, Incani UC, Steelman LS, Abrams SL, Falcone I, Curatolo AD, et al. Signaling intermediates (MAPK and PI3K) as therapeutic targets in NSCLC. *Curr Pharm Des.* 2014;20(24):3944–57. <https://doi.org/10.2174/13816128113196660763>.
53. Hinshaw DC, Shevde LA. The tumor microenvironment innately modulates cancer progression. *Cancer Res.* 2019;79(18):4557–66. <https://doi.org/10.1158/0008-5472.CAN-18-3962>.
54. Philip M, Schietinger A. CD8(+) T cell differentiation and dysfunction in cancer. *Nat Rev Immunol.* 2022;22(4):209–23. <https://doi.org/10.1038/s41577-021-00574-3>.
55. Budhu S, Schaefer DA, Li Y, Toledo-Crow R, Panageas K, Yang X, et al. Blockade of surface-bound TGF-beta on regulatory T cells abrogates suppression of effector T cell function in the tumor microenvironment. *Sci Signal.* 2017;10(494):eaak9702. <https://doi.org/10.1126/scisignal.aak9702>.
56. Kalia V, Penny LA, Yuzefpolskiy Y, Baumann FM, Sarkar S. Quiescence of memory CD8(+) T cells is mediated by regulatory T cells through inhibitory receptor CTLA-4. *Immunity.* 2015;42(6):1116–29. <https://doi.org/10.1016/j.immuni.2015.05.023>.
57. Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discov.* 2019;18(3):197–218. <https://doi.org/10.1038/s41573-018-0007-y>.
58. Riviere P, Goodman AM, Okamura R, Barkauskas DA, Whitchurch TJ, Lee S, et al. High tumor mutational burden correlates with longer survival in immunotherapy-naïve patients with diverse cancers. *Mol Cancer Ther.* 2020;19(10):2139–45. <https://doi.org/10.1158/1535-7163.MCT-20-0161>.
59. Motz GT, Coukos G. Deciphering and reversing tumor immune suppression. *Immunity.* 2013;39(1):61–73. <https://doi.org/10.1016/j.immuni.2013.07.005>.
60. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer.* 2012;12(4):252–64. <https://doi.org/10.1038/nrc3239>.
61. Lin H, Wei S, Hurt EM, Green MD, Zhao L, Vatan L, et al. Host expression of PD-L1 determines efficacy of PD-L1 pathway blockade-mediated tumor regression. *J Clin Investig.* 2018;128(2):805–15. <https://doi.org/10.1172/JCI96113>.
62. Wei SC, Duffy CR, Allison JP. Fundamental mechanisms of immune checkpoint blockade therapy. *Cancer Discov.* 2018;8(9):1069–86. <https://doi.org/10.1158/2159-8290.CD-18-0367>.
63. Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat Rev Cancer.* 2016;16(5):275–87. <https://doi.org/10.1038/nrc.2016.36>.
64. Yang Y. Cancer immunotherapy: harnessing the immune system to battle cancer. *J Clin Investig.* 2015;125(9):3335–7. <https://doi.org/10.1172/JCI83871>.
65. Curran MA, Montalvo W, Yagita H, Allison JP. PD-1 and CTLA-4 combination blockade expands infiltrating T cells and reduces regulatory T and myeloid cells within B16 melanoma tumors. *Proc Natl Acad Sci U S A.* 2010;107(9):4275–80. <https://doi.org/10.1073/pnas.0915174107>.
66. Hellmann MD, Rizvi NA, Goldman JW, Gettinger SN, Borghaei H, Brahmer JR, et al. Nivolumab plus ipilimumab as first-line treatment for advanced non-small-cell lung cancer (CheckMate 012): results of an open-label, phase 1, multicohort study. *Lancet Oncol.* 2017;18(1):31–41. [https://doi.org/10.1016/S1470-2045\(16\)30624-6](https://doi.org/10.1016/S1470-2045(16)30624-6).
67. Hellmann MD, Paz-Ares L, Bernabe Caro R, Zurawski B, Kim SW, Carcereny Costa E, et al. Nivolumab plus ipilimumab in advanced non-small-cell lung cancer. *N Engl J Med.* 2019;381(21):2020–31. <https://doi.org/10.1056/NEJMoa1910231>.

## Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

