



# Exploration of Novel Immunological Terms in Lung Cancer With Large Populations: Implications for Immunotherapy

Yuanshan Yao<sup>†</sup>, Jing Wang<sup>†</sup>, Fuzhi Yang<sup>†</sup> and Wen Gao<sup>\*</sup>

Department of Thoracic Surgery, Shanghai Key Laboratory of Clinical Geriatric Medicine, HuaDong Hospital Affiliated to Fudan University, Shanghai, China

## OPEN ACCESS

### Edited by:

Dinler Amaral Antunes,  
University of Houston, United States

### Reviewed by:

Jin-Zhou Xu,  
Huazhong University of Science and  
Technology, China  
Mustafa Elhag,  
University of Seychelles, Seychelles

### \*Correspondence:

Wen Gao  
gaowen1232021@163.com

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Cancer Immunity  
and Immunotherapy,  
a section of the journal  
Frontiers in Immunology

Received: 20 April 2022

Accepted: 08 June 2022

Published: 30 June 2022

### Citation:

Yao Y, Wang J, Yang F and Gao W  
(2022) Exploration of Novel  
Immunological Terms in Lung Cancer  
With Large Populations: Implications  
for Immunotherapy.  
*Front. Immunol.* 13:924498.  
doi: 10.3389/fimmu.2022.924498

**Background:** Ideal biomarkers to predict the response to immunotherapy in lung cancer are still lacking. Therefore, there is a need to explore effective biomarkers in large populations.

**Objective:** The objective of this study is to explore novel immunological classifications that are associated with immunotherapy response through the ssGSEA algorithm.

**Methods:** Six independent lung cancer cohorts were collected for analysis including The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), and the EMBL-EBI database. The ssGSEA algorithm was performed to extract immune terms. Then, TCGA samples were involved as a training group and other cohorts were used as a validation group. After LASSO and Cox regression, prognostic associated immune terms were extracted and an immune-related risk score (IRS) signature was constructed. Furthermore, the association between IRS signature and clinical data, genome features, stemness indices analysis, tumor immune microenvironment, immunotherapy efficiency, and targeted therapy response was also investigated.

**Results:** A total of 1,997 samples were enrolled in this study including six large lung cancer cohorts. Fifty-four immune terms were calculated through the ssGSEA algorithm in combined cohorts. Then, a nine-immune-term risk score model named IRS signature was established to predict the prognosis in combined cohorts. We classified patients into high-risk and low-risk subgroups according to the cutoff point. Subsequently, analysis of clinical data and genome features indicated that the patients in the high-IRS group tend to have advanced clinical features (clinical stage and T classification), as well as a higher level of copy number variation burden, higher tumor burden mutation, and higher tumor stemness indices. Immune landscape analysis demonstrated that high-IRS groups exhibited lower immune cell infiltration and immune-suppressive state. More importantly, the predicted result of the Tumor Immune Dysfunction and Exclusion analysis showed that high-IRS groups might be more insensitive to immunotherapy. Meanwhile, we have also identified that high-IRS groups were associated with better efficiency of several targeted drugs.

**Conclusion:** To summarize, we identified a novel IRS model based on nine immune terms, which was quantified by the ssGSEA algorithm. This model had good efficacy in predicting overall survival and immunotherapy response in non-small cell lung cancer patients, which might be an underlying biomarker.

**Keywords:** NSCLC, immune terms, IRS signature, prognosis, immunotherapy

## INTRODUCTION

Nowadays, immunotherapy has significantly improved the prognosis of advanced non-small cell lung cancer (NSCLC) (1). Although PD1/L1 inhibitors have changed the treatment landscape of NSCLC, the optimal biomarker to predict the clinical response of immunotherapy is still lacking (2). PD-L1 expression level and tumor mutational burden (TMB) in tumor specimens are two reliable biomarkers in clinical practice so far (3). However, a proportion of PD-L1-negative patients can also benefit from PD1/L1 inhibitors. Thus, it is urgent to explore other effective biomarkers with greater accuracy.

Recently, there are various biomarkers to predict immunotherapy response to lung cancer patients derived from the tumor immune microenvironment, molecular alterations, and serum indexes (4). For example, tumor-infiltrating lymphocytes (T cells and B cells) and the spatial location of these immune cells were already validated to be associated with the benefit of immunotherapy (5, 6). However, controversy still exists. Colt et al. demonstrated that penetration of immune cells into the cancer cells represented a better prognosis compared to those in tumor stroma (7). Nevertheless, Paul et al. revealed that immune cells that were located at the tumor margin exhibited a better response to immunotherapy than those in the stroma or intra-tumor (8). In addition, the prognosis of patients who received immune checkpoint inhibitors was associated with a non-invasive source of pre-treatment serum NLR (neutrophil-to-lymphocyte ratio) (9, 10). Currently, there is still no single satisfactory biomarker to evaluate the efficacy of immunotherapy. Incorporating multiple methods may be the best way (11).

In recent years, immune-related genes have also attracted attention. T cell-inflamed gene expression profile, which consisted of 18 IFN- $\gamma$  associated genes, was also explored and confirmed to exert its clinical prediction value in various types of malignancies (12). Johnson et al. demonstrated that patients with higher HLA-DR obtained better clinical response, prognosis, and progression-free survival compared with lower HLA-DR patients (13). Nevertheless, little is known about the correlation between immune terms and immunotherapy response in NSCLC patients.

Meanwhile, two atypical patterns of treatment responses are particularly correlated with immunotherapy including pseudoprogression and hyperprogression, which occupy a certain amount of patients accepting immunotherapy (14). The underlying molecular mechanism remains unknown so far. In clinical applications even using the somewhat improved immune RECIST (iRECIST) criteria, it is probable that the treatment

effect of several patients is misjudged, which could lead to loss of optimal treatment opportunities (15). Patients with improved or stable clinical symptoms may predict effective immunotherapy. Future work is warranted to refine imaging response criteria and explore potential biomarkers that can make treatment recommendations more clear and standardized.

In this study, we collected the expression data from the publicly accessible dataset with 1,997 NSCLC patients. Then, 54 immune terms were extracted through single-sample gene set enrichment analysis (ssGSEA). We identified nine immune terms with prognostic values and constructed an immune-related risk score (IRS) signature. Moreover, the IRS signature can be used as a biomarker to predict the efficacy of immunotherapy. Meanwhile, several targeted drugs were potential candidates for targeting this IRS signature. We hope that our study can provide a reference for the treatment of NSCLC.

## MATERIALS AND METHODS

### Data Collection and Preprocessing

Six cohorts with complete clinical information and available expression matrix data were collected (TCGA, GSE37745, GSE50081, GSE68465, GSE73403, and E-GEOD-30219) after a comprehensive search in this study. TCGA-LUAD data and TCGA-LUSC data were retrieved from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>). In detail, the expression profile data from TCGA were originally in “FPKM” form and then converted to “TPM” form for higher comparability with microarray data. E-GEOD-30219 data were retrieved from the ArrayExpress database (<https://www.ebi.ac.uk/arrayexpress/>). In addition, four individual Gene Expression Omnibus (GEO) datasets were identified from the GEO dataset (<https://www.ncbi.nlm.nih.gov/geo/>), including GSE37745 (GPL570), GSE50081 (GPL570), GSE68465 (GPL96), and GSE73403 (GPL6480). Three packages in the R environment, namely, sva, cluster, and oompaBase, were used to combine data and reduce the likelihood of batch effects and magnitude harmonization. The codes used were uploaded in the figshare website (<https://figshare.com/articles/software/Code/19995041>).

### Single-Sample Gene Set Enrichment Analysis and Pathway Enrichment Analysis

The R package “clusterProfiler” was used to perform gene set enrichment analysis (GSEA), Gene Oncology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis. The R

package GSVA was used to conduct ssGSEA to evaluate the enrichment scores of 54 immune terms. The Hallmark gene set (MSigDB) was selected as a reference set to explore the difference in the oncogenetic pathways. Immune-related features used for ssGSEA quantification were collected from Genomic Data Commons (<https://gdc.cancer.gov/about-data/publications/panimmune>) and previous studies, which was available at figshare ([https://figshare.com/articles/dataset/Untitled\\_Item/19368023](https://figshare.com/articles/dataset/Untitled_Item/19368023)).

## Calculation of Immune-Related Risk Score Signature for the Combined Patient Cohort

After removing batch effects, TCGA-LUAD and TCGA-LUSC cohorts were selected as the training cohort and other datasets were used as the validation cohort (GSE37745, GSE50081, GSE68465, GSE73403, and E-GEOD-30219). For the training cohort, univariate Cox analysis was performed to identify prognosis-related immune terms with the criteria  $p$ -value  $< 0.05$ . Then, L1-penalized (LASSO) estimation across 1,000 iterations was used for dimension reduction. Immune terms with a frequency higher than 50 times were then selected for multivariate Cox regression analysis and IRS calculation using the following formula: Immune risk score = term1\*coef1 + term2\*coef2 + term3\*coef3 + ... + termN\*coefN. R packages “SimDesign” and “tdROC” were used to get the best cutoff value of risk score in training and validation cohorts. The R packages “survival” and “survivalROC” were used to assess the prognostic value of the signature through Kaplan–Meier survival curve and the ROC curve.

## Features of Tumor Genomics Between Two IRS Groups

TMB means the number of mutations per megabase (mt/Mb). Somatic mutation data retrieved from the cBioPortal website (<http://www.cbioportal.org/datasets>, Lung Adenocarcinoma/Lung Squamous Cell Carcinoma, TCGA, PanCancer Atlas) was used to calculate TMB. The R package “maftools” was used to analyze significantly mutated genes with  $p$ -value  $< 0.05$  between the two IRS groups and the interaction effect of gene mutations. GISTIC\_2.0 (<https://cloud.genepattern.org>) was applied for copy number variation (CNV) analysis. Based on the output files from GISTIC\_2.0, copy number gain burden and loss burden at the focal and arm levels were calculated. One-class logistic regression machine learning (OCLR) algorithm was used to quantify the tumor stemness index, including mRNasi, mDNasi, EREG-mRNasi, and EREG-mDNasi, whose process is available on [https://bioinformatics.fmrp.github.io/PanCanStem\\_Web/](https://bioinformatics.fmrp.github.io/PanCanStem_Web/). ESTIMATEscore, StromalScore, and ImmuneScore of each sample were calculated using the R package “ESTIMATE”. A higher score represents a more considerable amount of the corresponding lower tumor purity and higher stromal and immune cells in TMB.

## Prediction of Immunotherapeutic and Chemotherapeutic Response

The Tumor Immune Dysfunction and Exclusion (TIDE, <http://tide.dfci.harvard.edu>) algorithm was used to predict patient

immunotherapy response. Each patient with a TIDE score  $< -0.2$  was defined as a Responder, and a patient with a TIDE score  $> 0.2$  was defined as a No-responder. Subclass mapping algorithm (<https://cloud.genepattern.org/>) was used to assess similarities in response to immunotherapies between the 47 patients who responded to immunotherapies and two lung cancer patients from the IRS group. The response to chemotherapy drugs of patients was predicted based on the public pharmacogenomics database Genomics of Drug Sensitivity in Cancer (GDSC, <https://www.cancerrxgene.org/>), and the R package “pRRophetic” was used to estimate the half-maximal inhibitory concentration ( $IC_{50}$ ) of twelve commonly used chemo drugs for predicting the sensitivity of chemotherapy drugs.

## Statistical Analysis

All statistical analyses were performed with R version 4.0.2 and its appropriate packages. All statistics were two-sided and statistical significance was defined as  $p$ -value  $< 0.05$ . Data were analyzed with standard statistical tests as appropriate. Spearman correlation was used to estimate the correlations between continuous variables. Independent sample  $t$ -test was used to compare continuous variables with normal distribution. Wilcoxon rank-sum test was used to compare continuous variables with skewed distribution.

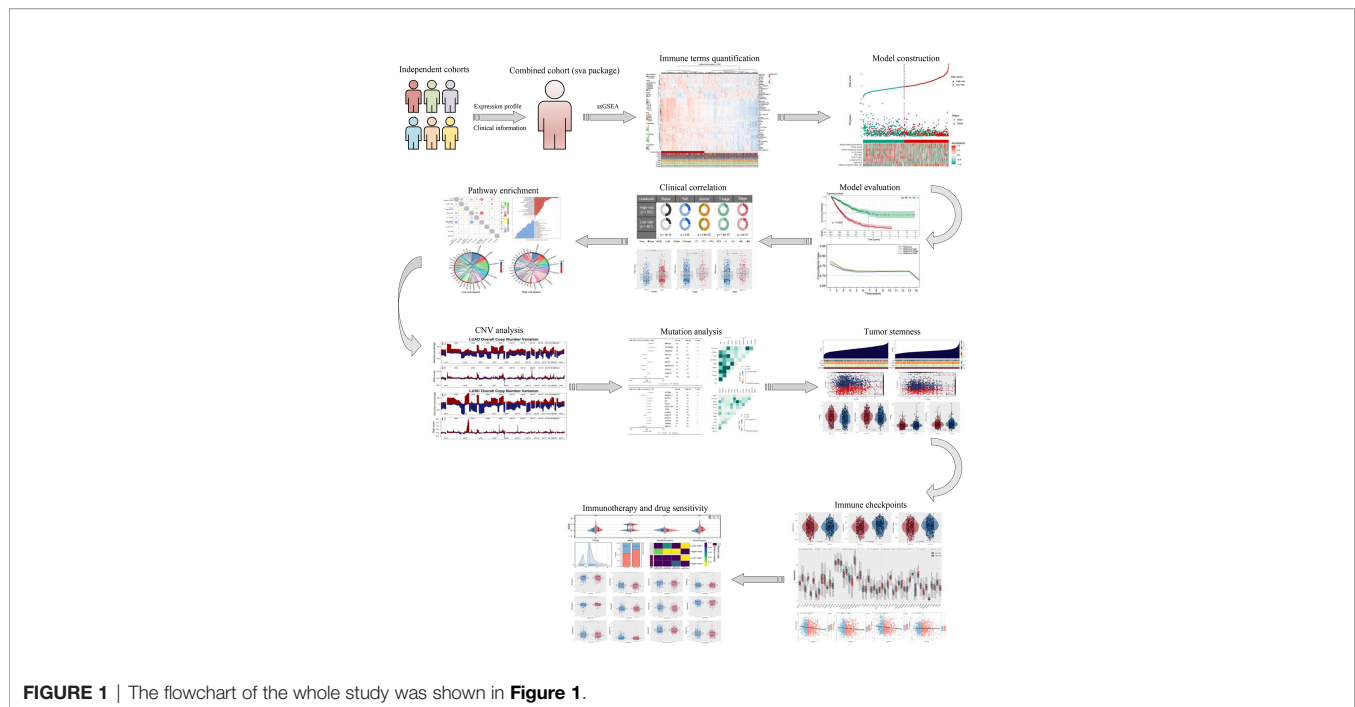
## RESULTS

### Immune Terms' Quantification

The flowchart of the whole study is shown in **Figure 1**. After a comprehensive search of the public database, six individual NSCLC cohorts met our criteria and were finally included in our analysis (TCGA, GSE73403, GSE68465, GSE50081, GSE37745, and E-GEOD-30219). A significant batch difference was observed between these cohorts (**Figure 2A**, comp1: 75.7% variance, comp2: 3.9% variance). Then, the sva package was used to remove the batch effect of these six NSCLC cohorts (**Figure 2B**, comp1: 8.9% variance, comp2: 6.2% variance). The combined expression profile was then used for the 54 immune terms' quantification based on the ssGSEA algorithm, which is shown in **Figure 2C**.

### Identification of the Prognosis-Related Immune Terms

To further identify the immune terms associated with patient prognosis, we first performed a univariate Cox analysis with the threshold of  $p < 0.05$ . Then, iterative LASSO regression was used for high-frequency features selection and 18 prognosis-related terms were retained (**Figures 3A, B**). Nine immune terms were selected for prognosis model construction based on the multivariate Cox analysis, including B-cell receptor score, CD8 T cells, IFNG score, IL13 score, Inflammation promoting, Lymphs PCA, SERUM RESPONSE UP, Tfh cells, and Tgd cells (**Figure 3C**). Among these, IFNG score, SERUM RESPONSE UP, CD8 T cells, and IL13 score were risk factors, yet Tgd cells, Inflammation promoting, Lymphs PCA, Tfh cell,



and B-cell receptor score were protective factors (**Figure 3D** and **Figure S1A**). The model IRS was calculated with the following formula: “IRS = B cell receptor score \* -0.115 + CD8 T cell \* 0.148 + SERUM RESPONSE UP \* 0.225 + IFNG score \* 0.265 + IL13 score \* 0.142 + Inflammation promoting \* -0.249 + Lymphs PCA \* -0.135 + Tfh cells \* -0.128 + Tgd cells \* -0.161”. SimDesign and the tdROC package were used to calculate the optimum cutoff values, and the result showed that the 0.98 and 1.38 were the best cutoff of training cohort and validation cohort, respectively (**Figures S1B,C**). In the training cohort, more death cases could be observed in the high-risk group (**Figure 3E**). The ROC curve indicated a satisfactory prediction efficiency of our model on patients’ OS (**Figure 3F**, 3-year AUC: 0.781, 5-year AUC: 0.780, 8-year AUC: 0.767). Moreover, the Kaplan–Meier survival curve showed that the patients in the high-risk group have a worse prognosis (**Figure 3G**). Meanwhile, time-dependent AUC showed that the risk score combined clinical features tend to have a better prediction efficiency than risk score (**Figure 3H**). The same trend and great prediction efficiency were also observed in the validation cohort (**Figures 3I–L**). Univariate and multivariate Cox analysis showed that the IRS is a risk factor independent of other clinical features (**Figure S4**).

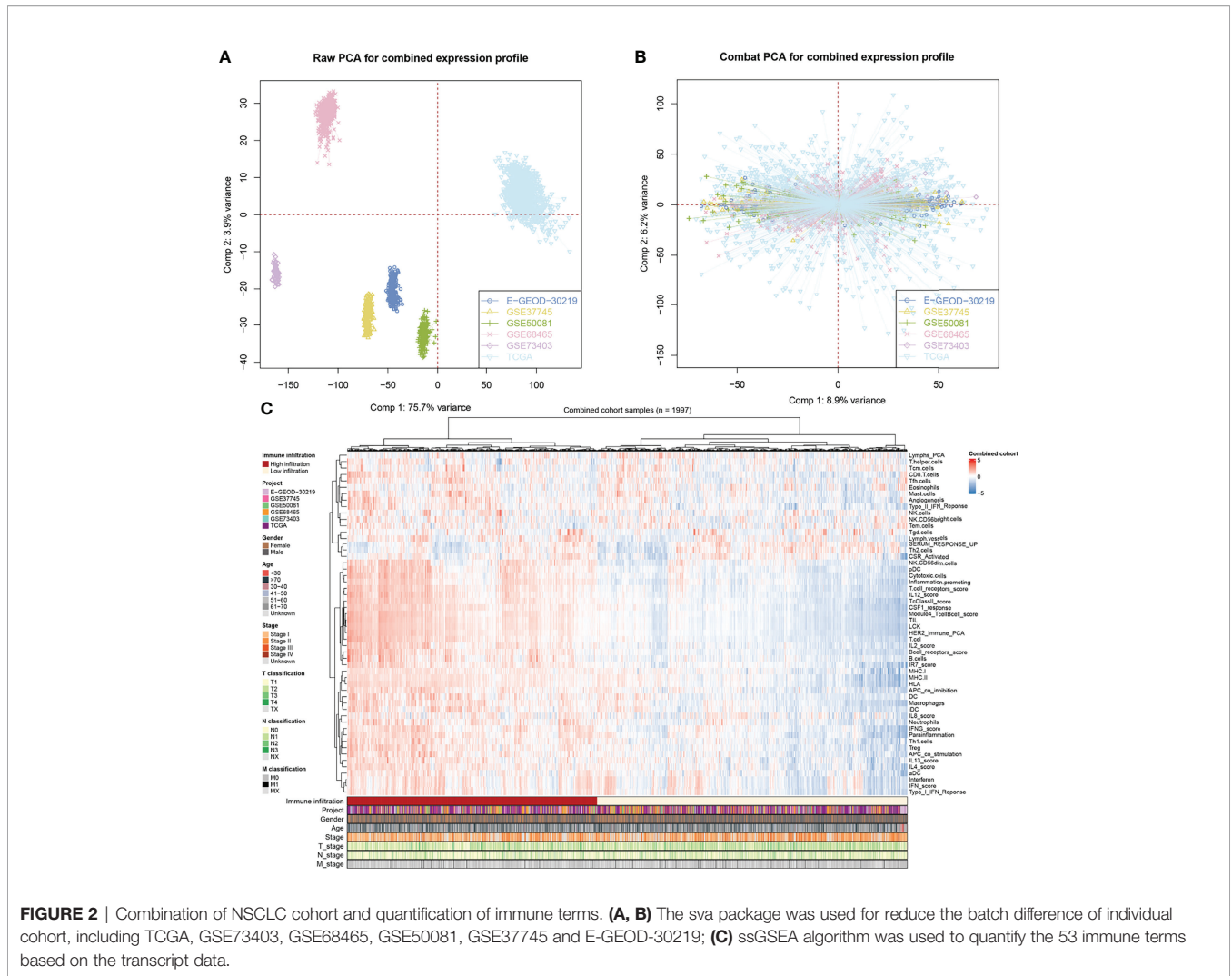
### Clinical Correlation and Biological Role of the IRS Model

We further explored the clinical correlation and biological role of the IRS model to explain its prognosis effect. The result showed that patients in the high-risk group might have adverse clinical features, including T classification and clinical stage (**Figures 4A–D**). Interestingly, the male patients might have a higher IRS than the female patients, indicating their underlying immune

microenvironment difference (**Figure 4B**). The co-expression relationship of nine model immune terms was visualized as a correlation coefficient heatmap to further explore their interactions (**Figure 4E**). GSEA demonstrated that in the high-risk group, the pathway of bile acid metabolism, KRAS signaling, fatty acid metabolism, xenobiotic metabolism, interferon-gamma response, inflammatory response, IL2-STAT5 signaling, and IL6-JAK-STAT3 signaling were activated, yet the hypoxia, P53 pathway, and DNA repair were downregulated (**Figure 4F** and **Figure S2**). GO and KEGG enrichment analysis showed that in the low-risk group, the terms post-translational protein modification (GO:0043687), protein–lipid complex remodeling (GO:0034368), endoplasmic reticulum lumen (GO:0005788), blood microparticle (GO:0072562), lipase inhibitor activity (GO:0055102), immunoglobulin binding (GO:0019865), cholesterol metabolism (hsa04979), and fat digestion and absorption (hsa04975) were enriched (**Figure 4G**). In the high-risk group, the terms cornification (GO:00702680), antimicrobial humoral immune response mediated by antimicrobial peptide (GO:0061844), intermediate filament (GO:0005882), intermediate filament cytoskeleton (GO:0045111), G protein-coupled receptor binding (GO:0001664), *Staphylococcus aureus* infection (hsa05150), and Wnt signaling pathway (hsa04310) were enriched (**Figure 4H**).

### The Correlation Between Genomic Features and IRS in NSCLC

The overview of gene mutation of TCGA-NSCLC is shown in **Figure S3**. Then, the copy number profile in TCGA-LUAD and TCGA-LUSC patients, including gain/loss percentage and gistic score, was characterized (**Figures 5A–D**). The result showed that high-risk patients had a higher degree of genomic instability

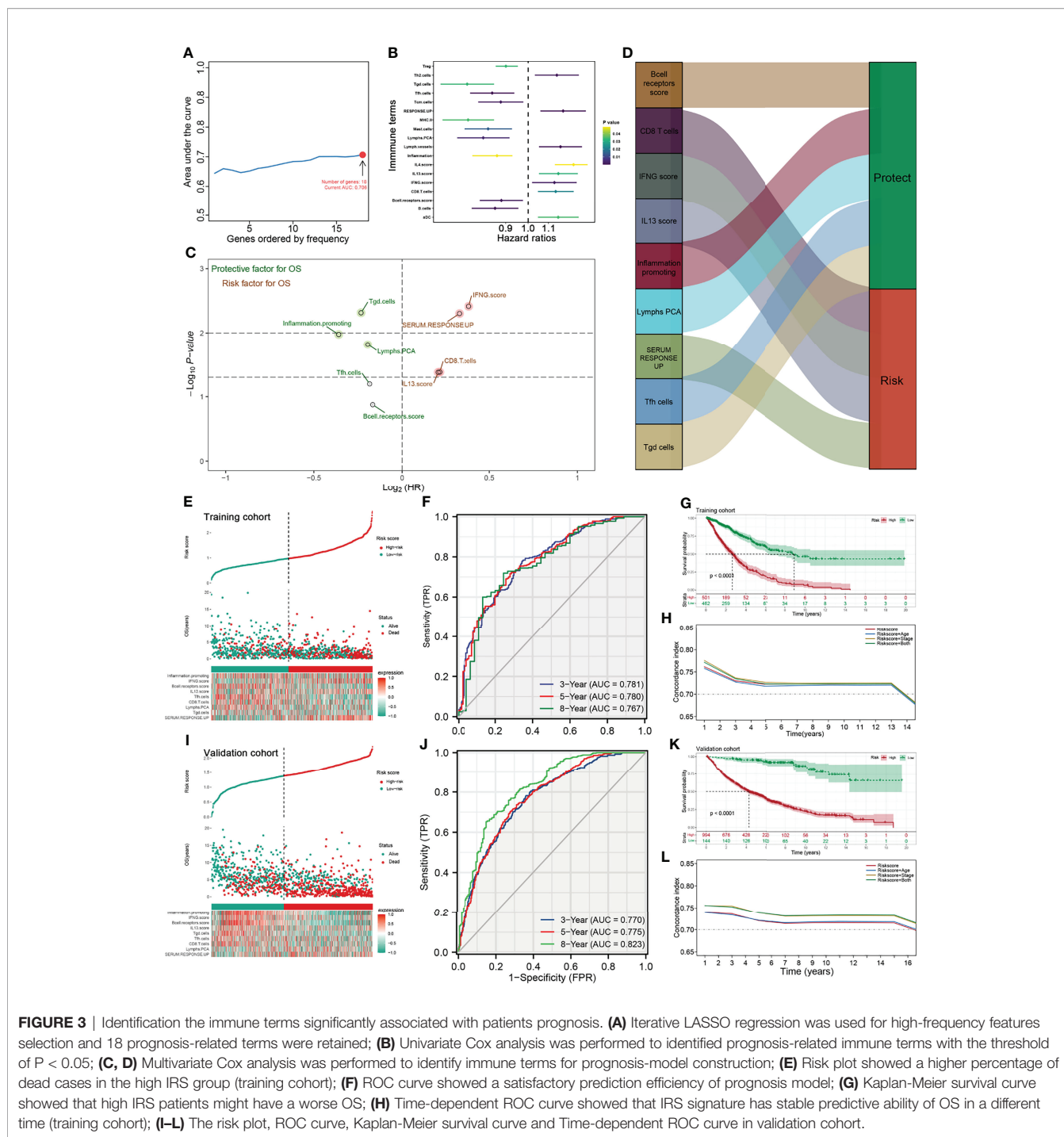


(Figures 5E-H). In detail, CNV burden analysis showed that high-risk patients had a higher level of focal and broad CNV burden in both gain and loss aspects (Figures 5F-H). The overview of the tumor mutation landscape of NSCLC is shown in Figure 6A. Moreover, higher TMB was also observed in the high-risk group (Figure 6B,  $p = 0.021$ ). Furthermore, considering the intrinsic biological differences, we separately explored the mutant gene in TCGA-LUAD and TCGA-LUSC cohorts. The result showed that in the TCGA-LUAD cohort, TEX15, TP53, SMARCA4, NTRK3, and TTN were the most common mutated genes in the high-risk group, while the MMP16, ZFYVE26, FRMPD4, DIDO1, and OBSCN were the opposite (Figure 6C). In the TCGA-LUSC cohort, PKHD1L1, MUC16, and HERC1 were the most common mutated genes in the high-risk group, while the PTPRB, ASTN1, C6, PCLO, KIAA1109, TPTE, LAMA2, DCDC1, LPHN3, and TENM1 were the opposite (Figure 6E). The co-mutation relationship between these genes was then explored for the underlying interaction (Figure 6D, LUAD; Figure 6F, LUSC). Tumor stemness plays an important role in cancer progression.

Therefore, we quantified the tumor stemness in the RNA and DNA level (Figures 7A, B). The result showed that risk score has a significantly positive correlation with mRNasi and EREG-mRNasi, but a negative correlation with mDNasi and EREG-mDNasi (Figure 7C, mRNasi,  $r = 0.21, p < 0.001$ , EREG-mRNasi,  $r = 0.11, p < 0.001$ ; Figure 7D, mDNasi,  $r = -0.13, p < 0.001$ , EREG-mDNasi,  $r = -0.09, p < 0.001$ ). Meanwhile, we found a higher mRNasi and EREG-mRNasi level in the high-risk group (Figures 7E, F). In addition, a lower mDNasi and EREG-mDNasi level was also observed in high-risk patients (Figures 7G, H).

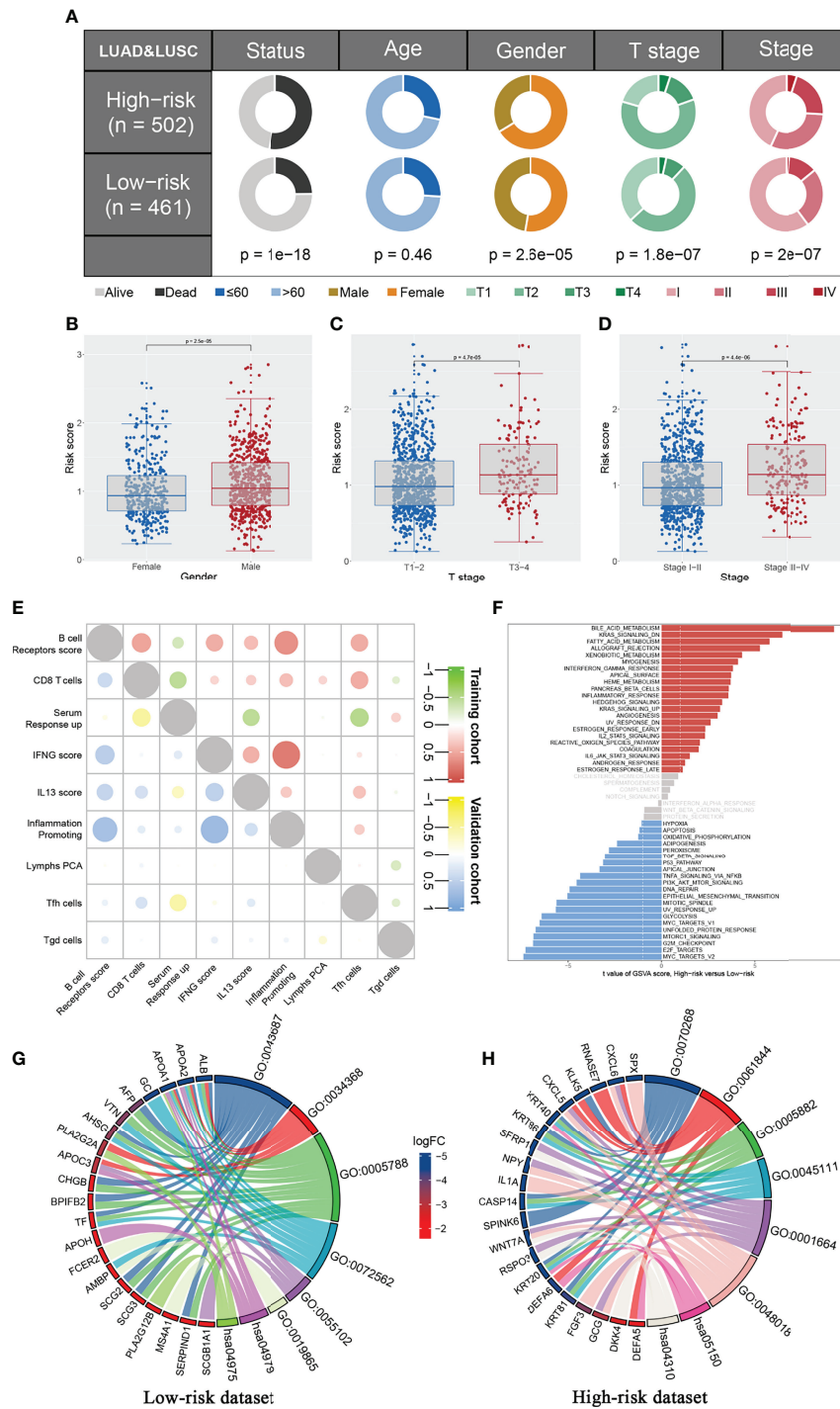
### The IRS Model Is Closely Associated With Immunotherapy Response

The ESTIMATE algorithm was used to quantify the tumor microenvironment score, including ESTIMATEScore, StromalScore, and ImmuneScore. No significant difference in StromalScore was found in high- and low-risk patients (Figure 8A). However, high-risk patients tend to have a lower ImmuneScore and ESTIMATEScore compared with low-risk patients (Figures 8B, C). Recently, immune checkpoints have

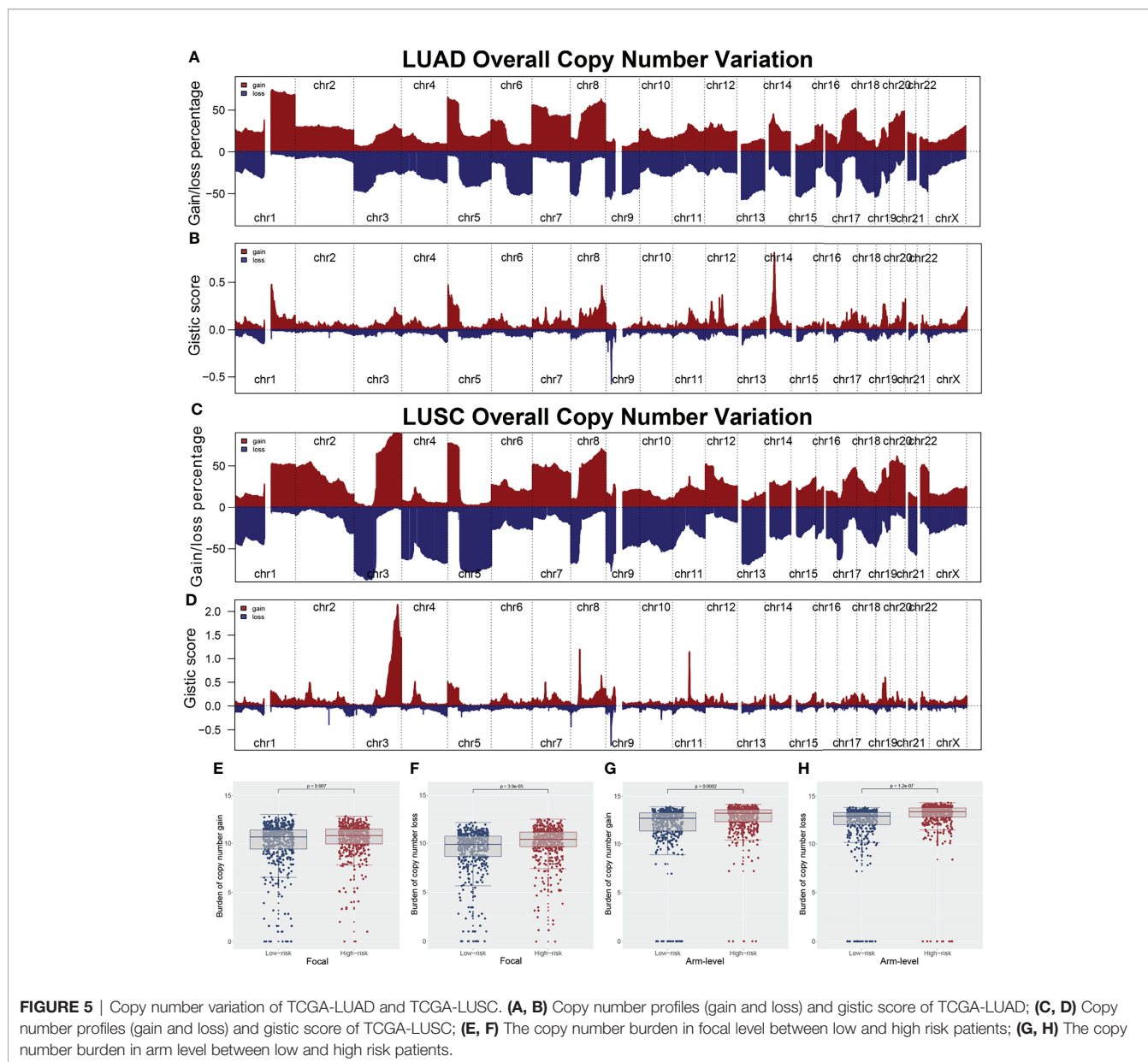


gained a lot of interest among the scientific community. Therefore, we performed a correlation analysis between our IRS model and multiple checkpoint modules. The result showed a significant difference in several immune checkpoint modules between high- and low-IRS patients (Figure 8D). CTLA-4, PD-1, PD-L1, and PD-L2 were the immune checkpoint modules with great attention. We found that high-risk patients might have a higher CTLA-4, PD-1, and PD-L1 level, indicating the underlying immunotherapy response difference between low- and high-risk patients

(Figures 8E-H). TIDE analysis was performed to explore the effect of IRS on immunotherapy response. The result showed that the patients in the high-risk group might have a higher TIDE score (Figure 9A). The patients with TIDE score  $< 0$  were defined as immunotherapy responders, while patients with TIDE score  $> 0$  were defined as immunotherapy non-responders (Figure 9B). We found that the high-risk group had a higher percentage of immunotherapy non-responders (Figure 9C). Based on the submap algorithm obtained from GenePattern, we found that



**FIGURE 4 |** Clinical correlation and biological effect of IRS model. **(A)** The patients in high risk group have a higher percentage of more aggressive clinical features compared with the low risk patients, including T stage and clinical stage; **(B–D)** The IRS level between different group (Male vs. Female, T3-4 vs. T1-2, Stage III-IV vs. Stage I-II); **(E)** Co-expression immune terms of model immune terms; **(F)** GSEA analysis was performed to explore the biological pathway difference between high and low IRS group; **(G, H)** GO and KEGG pathway enrichment in low and high risk patients.



patients in the low-risk group might be more sensitive to PD-1 immunotherapy (**Figure 9D**). Next, drug sensitivity analysis was performed to explore the underlying influence of IRS on common chemotherapy drugs (**Figures 9E–P**). The result showed that high-risk patients might be more sensitive to bosutinib, lapatinib, nilotinib, pazopanib, sunitinib, tipifarnib, temsirolimus, and vorinostat, yet resistant to metformin.

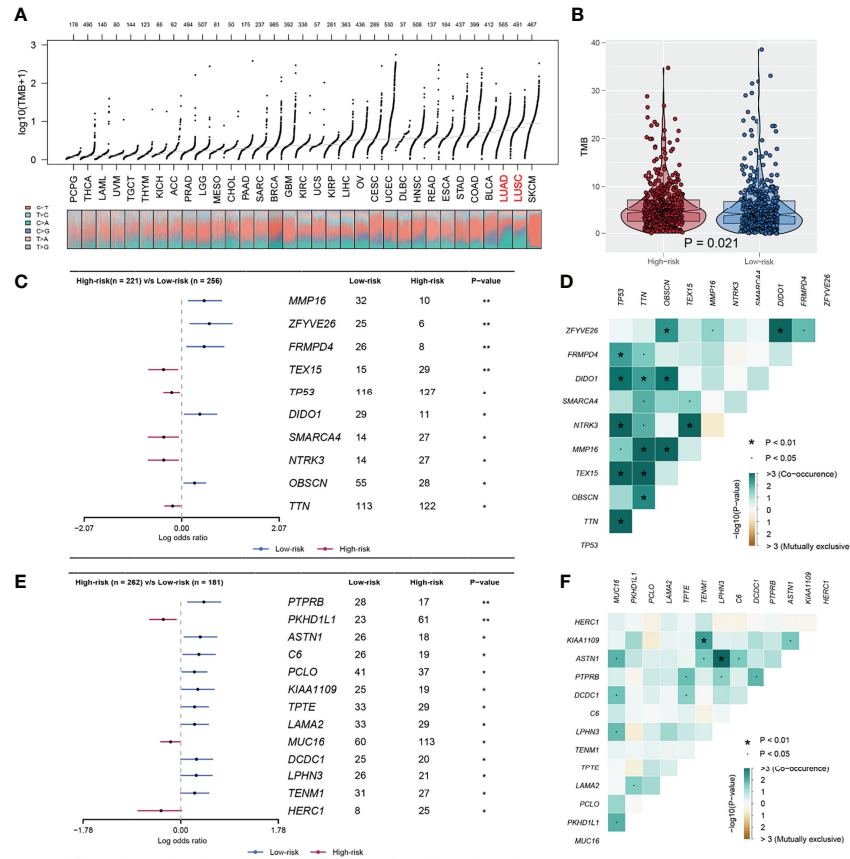
## DISCUSSION

In recent years, PD-1/L1 inhibitors have changed the treatment landscape of advanced lung cancer patients. However, the relatively lower response rate and higher immune-related

adverse occurrence rate are hardly satisfactory (16). Thus, it is necessary to explore another precise predictive biomarker for immunotherapy treatment. With the development of bioinformatics analysis, it is feasible for researchers to find another reliable biomarker in large populations.

In our study, we first enrolled 1,997 samples from six independent lung cancer cohorts that contained RNA-seq and clinical data. Then, the ssGSEA algorithm was performed to probe 54 immune terms in a combined expression profile. Furthermore, TCGA-LUAD combined with TCGA-LUSC samples were involved as a training group and other cohorts were classified as a validation group. Subsequently, nine prognostic-related immune terms were explored through univariate Cox regression analysis, LASSO analysis, and multivariate Cox regression analysis. We





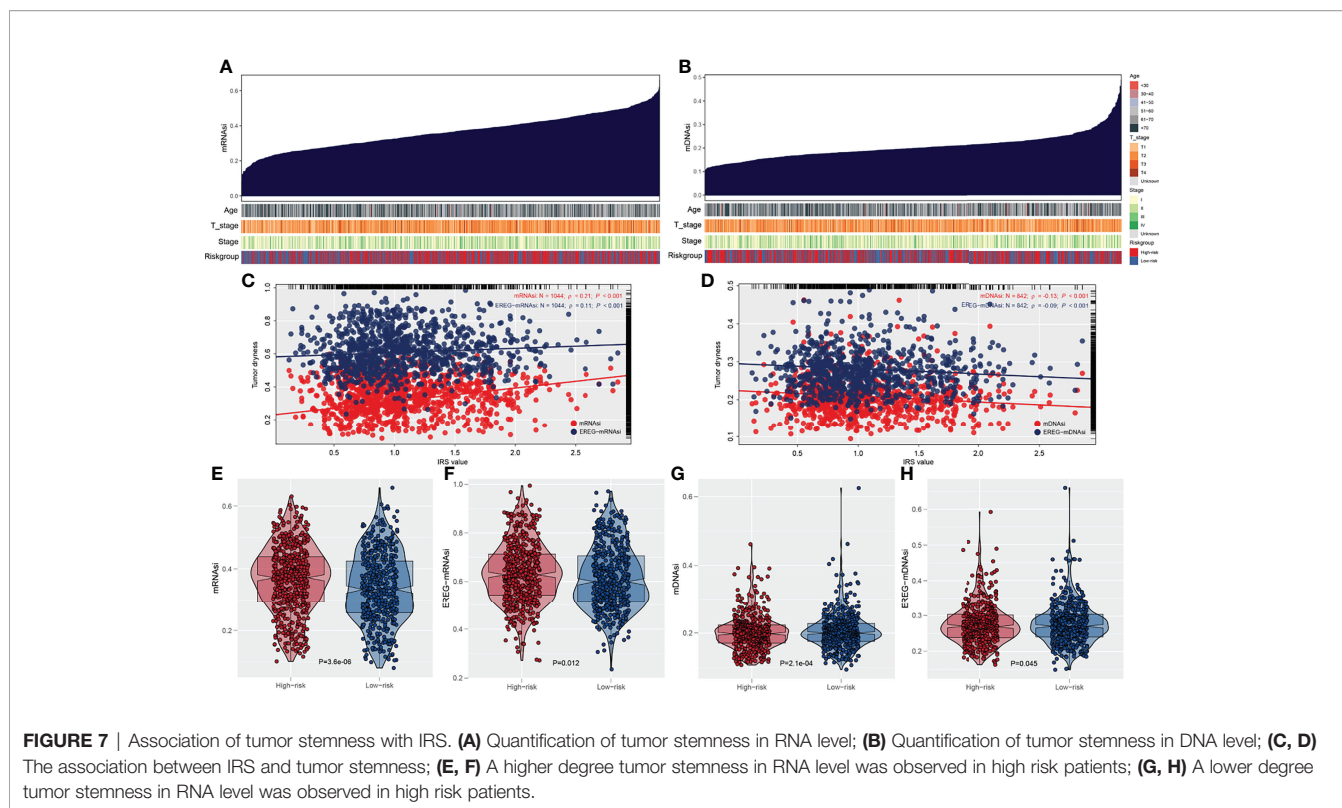
**FIGURE 6 |** The TMB difference between high and low IRS group. **(A)** The overview of TMB in TCGA-pancancer; **(B)** A higher TMB level was observed in high risk patients compared with the low risk patients; **(C, D)** The mutated gene in TCGA-LUAD between high and low risk with  $P < 0.05$ ; **(E, F)** The mutated gene in TCGA-LUSC between high and low risk with  $P < 0.05$ . \* $P < 0.05$ , \*\* $P < 0.01$ .

calculated a prognosis-related immune term signature called IRS based on these nine immune terms. According to the cutoff point, all patients were classified as high-risk and low-risk groups. Our study showed that these immune terms were significantly or marginally significantly associated with the prognosis in NSCLC patients. In addition, high IRS demonstrated aggressive clinical characteristics, worse prognosis, higher level of CNV burden, and higher tumor burden mutation and tumor stemness indices. In addition, the IRS model might effectively predict the response to immunotherapy and targeted therapy in NSCLC patients.

In our study, to enhance trustworthiness and credibility, we collected a total of 1,997 samples from six large lung cancer cohorts. Nine prognosis-associated immune terms were extracted after LASSO and Cox regression analysis. Meanwhile, these nine immune terms that included five protective factors and four risk factors were significantly or marginally significantly associated with the prognosis of lung cancer patients, implying a role in lung cancer progression. For example, Kalli et al. and Papageorgis et al. both indicated that IL13 could promote metastasis of breast cancer cells to lung tissue (17, 18). Moreover, IFN- $\gamma$  and CD8+T cells

servicing as risk factors to prognosis could have surprised us because CD8+T cells and IFN- $\gamma$  secretion were always thought as anti-tumor immune surveillance. We postulated that effector phenotype and exhausted phenotype CD8+T cells both counted through ssGSEA may be the major causes. Exhausted phenotype CD8+T cells represented a dysfunctional state, which negatively regulated the function of tumor cell killing (19). In parallel, IFN- $\gamma$  might enhance endogenous PD-L1 expression and boost tumor metastasis in the tumor immune microenvironment (20). Therefore, it is critically important to facilitate the transformation towards anti-tumor phenotype (21). Recent studies also showed that intratumor-infiltrating B cells inhibited the early-stage lung cancer progression and predicted better immunotherapy response (22, 23). Moreover, the concordance index indicated that the combination of tumor stage and IRS signature could be a strong predictor of overall survival. These nine immune terms provided essential foundation to further explore the underlying tumor microenvironment and genomics information in NSCLC samples.

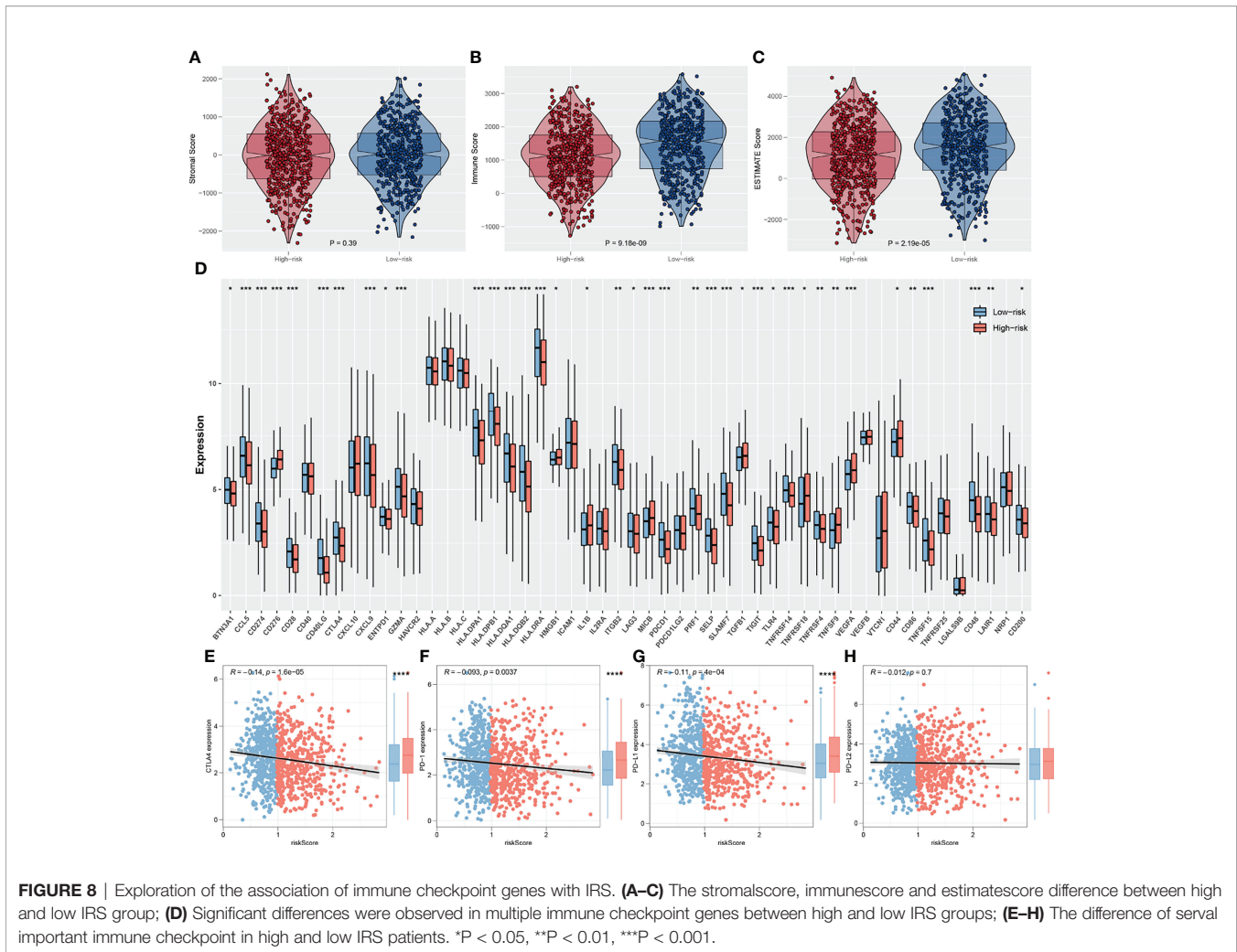
Our study showed that this novel model had good prognostic value in both training and validation groups. Then, we found



that high IRS was correlated with advanced TNM stage, suggesting possible immune suppression. Pathway enrichment analysis displayed significant upregulation of bile acid metabolism, K-RAS signal, and fatty acid metabolism in high-IRS groups. Previous studies already showed that bile acid metabolism and fatty acid metabolism contributed to the tumor invasion (24, 25). Oncogenic K-RAS signaling that has been studied extensively promotes tumor progression in several cancers. Reck et al. indicated that K-RAS gene mutation was correlated with an immunosuppressive landscape through recruiting myeloid-derived suppressor cells (26). This could explain the different immunotherapy response between high- and low-IRS groups. P53 is a classical tumor-suppressive gene and downregulation of P53 was already verified to facilitate cancer progression in various cancers (27). GO and KEGG analysis demonstrated that programmed cell death signaling, G protein-coupled receptor signaling, humoral immune response, and WNT signaling pathway were mostly enriched pathways in high-IRS groups. Shen et al. reported that G protein-coupled estrogen receptor facilitated lung cancer cell metastasis through the NOTCH pathway (28). Our study indicated that high-IRS groups had worse clinical outcomes in combination with the abnormal activation of the mentioned pathways.

In addition, the high-IRS group demonstrated higher copy number alteration burden in the focal and arm level compared with the low-IRS group. This indicated that high-risk patients had a higher degree of genomic instability. It is generally accepted that genetic phenotypic differences caused by copy

number alteration can facilitate tumor initiation and progression (29). Patients with a high level of CNV tend to be associated with decreased levels of immune cell infiltration landscape and elevated tumor proliferation ability (30). Meanwhile, CNVs were verified to be correlated with the prognosis of multiple cancers and the outcome of immunotherapy (31–33). A recent study showed that copy number alteration that can lead to genomic instability was indispensable in the transition from lung carcinoma *in situ* to invasive lung carcinoma (34). Gene mutation analysis revealed that patients with a high risk score possessed a higher TMB, indicating different potential response to immunotherapy. Our study was consistent with Liu et al., which showed that low TMB and low CNV were associated with better survival. In parallel, Liu et al. also revealed that those patients with high TMB and high CNV were relatively immunotherapy resistant (35). In addition, the probability of SMARCA4, NTRK3, and TEX15 mutation frequency was higher in high-risk LUAD groups. Concepcion et al. recently demonstrated that the SMARCA4 mutation type could decrease the chromatin accessibility and promote the progression of LUAD (36). NTRK was a rare genetic mutation in lung cancer and NTRK fusion inhibitor had already been used in clinical application (37). At the same time, there was a higher incidence of HERC1 and PKHD1L1 mutation in the LUSC group. The role of HERC1 and PKHD1L1 in LUSC was not reported and needed further validation. Furthermore, cancer stemness results showed that risk score has a positive correlation with mRNasi and EREG-mRNasi, but a weakly negative correlation with mDNasi

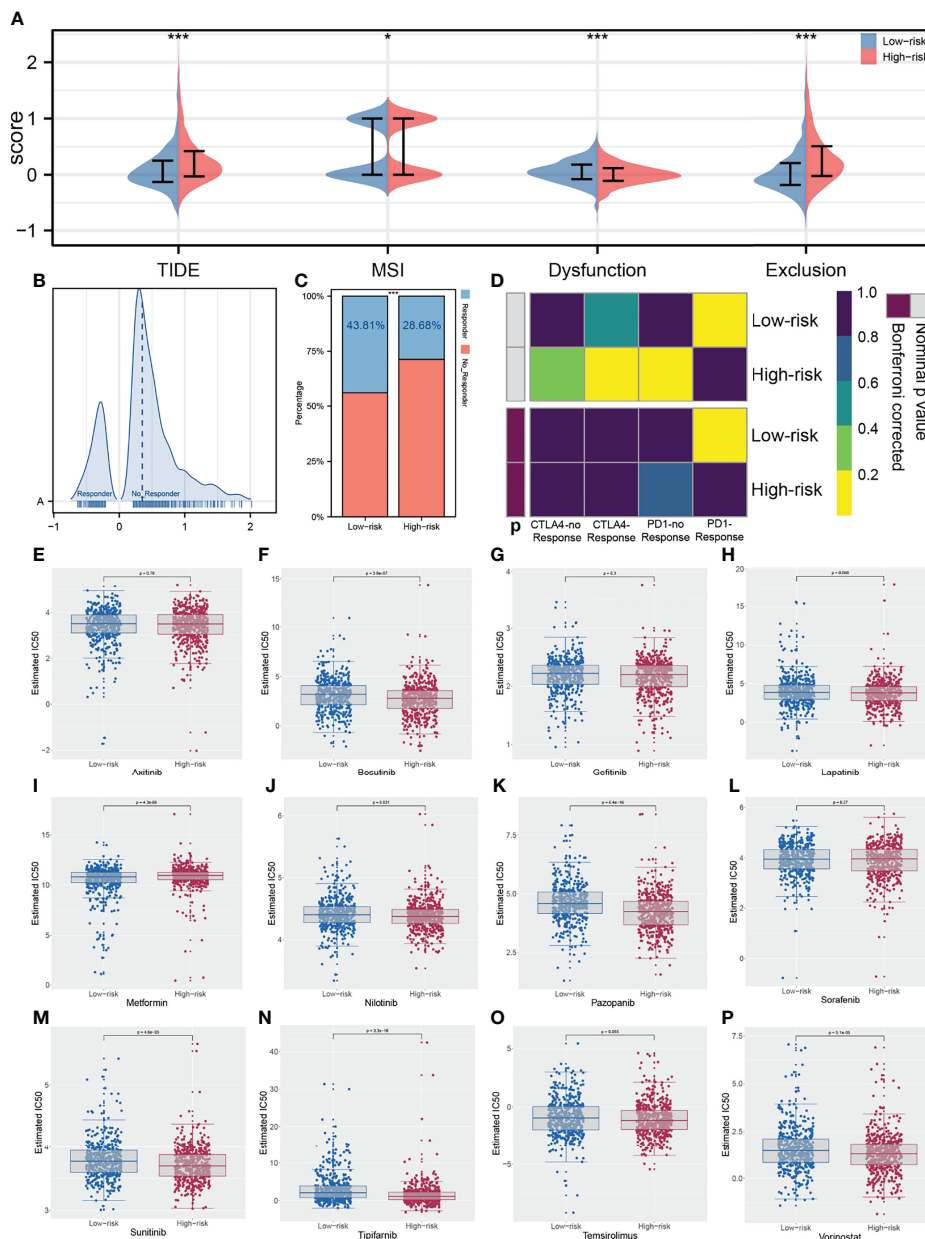


and EREG-mDNAsi (mRNAsi,  $r = 0.21$ , EREG-mRNAsi,  $r = 0.11$ ; mDNAsi,  $r = -0.13$ , EREG-mDNAsi,  $r = -0.09$ ). In the meantime, because mRNAsi was reflective of gene transcription expression, the high-IRS group exhibited higher tumor stemness index compared with the low-risk group (38). The study of Malta et al. showed that a higher tumor stemness index was highly correlated with tumor progression and lower PD-L1 expression level, which was shown in high-risk samples (39).

More importantly, high-IRS group patients had lower immune cell infiltration compared with the low-IRS group through the ESTIMATE algorithm (40). Our study also showed that the high-IRS group harbored a relative low PD-1, PD-L1, and CTLA-4 expression level, which usually means decreased immunotherapy efficacy. Then, results of TIDE and Subclass mapping algorithm showed that the patients with high IRS were not likely to benefit from immunotherapy. This conclusion was consistent with the previous finding. In addition, bosutinib, lapatinib, nilotinib, pazopanib, sunitinib, tipifarnib, temsirolimus, and vorinostat were all potential molecular targeted drugs for the treatment of

high-IRS patients. Most of these drugs were already demonstrated to have an effect on lung cancer cells in recent years. Tan and his partners showed that bosutinib could repress the KRAS mutant lung cancer cells (41). Pazopanib blocked tumor growth and reduced the metastatic sites in lung cancer mouse models (42). Metformin acting as a traditional glucose-lowering agent could not inhibit the progression of lung cancer alone (43). Thus, further development of these agents is still warranted.

Nevertheless, some inevitable limitations should be noticed in our analysis. Firstly, this study was a retrospective analysis and simply using bioinformatics analysis were the largest weaknesses. Furthermore, a prospective sequencing data cohort is needed to support our IRS signature. Secondly, our study consisted only of a few Asian individuals available, which accounted for a small proportion of all cohorts. This situation might limit the clinical application of this IRS signature in China. Thirdly, the open-access data used for analysis were all at the mRNA level, and not at the protein level, which hardly reflects the real situation of lung cancer tissue.



**FIGURE 9 |** Correlation of IRS immunotherapy response rate and drug sensitivity. **(A)** TIDE analysis was performed in enrolled patients; **(B)** The patients with TIDE score < 0 were defined as immunotherapy responders, while > 0 were defined as immunotherapy non-responders; **(C)** The high risk group had a higher percentage of immunotherapy non-responders; **(D)** The low risk group might more sensitive to PD-1 immunotherapy; **(E–P)** Drug sensitivity analysis was performed to explore the underlying influence of IRS on common chemotherapy drugs. \*P < 0.05, \*\*\*P < 0.001.

## CONCLUSION

According to the prognosis-related immune terms quantified through ssGSEA, we established IRS signature to evaluate the outcome of NSCLC patients and respond to immunotherapy. Moreover, results indicated that high-IRS patients are inclined to

bear increased genome instability, cancer CNV, and stemness indices, which mainly accounted for a dismal prognosis and a relatively low efficiency of immunotherapy. In addition, some targeted agents were identified including bosutinib, lapatinib, nilotinib, pazopanib, sunitinib, tipifarnib, temsirolimus, and vorinostat, which might reverse this adversity.

## 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

Conception and design: YY. Administrative support: JW. Provision of study materials: FY. Data analysis and interpretation: WG. Manuscript writing: All authors. Final approval of manuscript: All authors.

## FUNDING

This study was supported by the Shanghai Key Laboratory of Clinical Geriatric Medicine, Shanghai Municipal Key Clinical Specialty (shslczdzk2801) and the Lung Cancer Diagnosis and Treatment Center (H1382).

## REFERENCES

- Hong T, Sun T, Zhang M, Liu X, Yuan Y, Dolo PR, et al. Surgical Perspective in Neoadjuvant Chemotherapy for Stage II-III Non-Small Cell Lung Cancer. *Thorac Cancer* (2021) 12(20):2796–802. doi: 10.1111/1759-7714.14127
- Brozos-Vázquez EM, Díaz-Peña R, García-González J, León-Mateos L, Mondelo-Macia P, Peña-Chilet M, et al. Immunotherapy in Nonsmall-Cell Lung Cancer: Current Status and Future Prospects for Liquid Biopsy. *Cancer Immunol Immunother* (2021) 70(5):1177–88. doi: 10.1007/s00262-020-02752-z
- Shi Y, Lei Y, Liu L, Zhang S, Wang W, Zhao J, et al. Integration of Comprehensive Genomic Profiling, Tumor Mutational Burden, and Pd-L1 Expression to Identify Novel Biomarkers of Immunotherapy in Non-Small Cell Lung Cancer. *Cancer Med* (2021) 10(7):2216–31. doi: 10.1002/cam4.3649
- Kubo T, Shinkawa T, Kikuchi Y, Murata K, Kanaseki T, Tsukahara T, et al. Fundamental and Essential Knowledge for Pathologists Engaged in the Research and Practice of Immune Checkpoint Inhibitor-Based Cancer Immunotherapy. *Front Oncol* (2021) 11:679095. doi: 10.3389/fonc.2021.679095
- Hashemi S, Fransen MF, Niemeijer A, Ben Taleb N, Houda I, Veltman J, et al. Surprising Impact of Stromal TILs on Immunotherapy Efficacy in a Real-World Lung Cancer Study. *Lung Cancer* (2021) 153:81–9. doi: 10.1016/j.lungcan.2021.01.013
- Hendry S, Salgado R, Gevaert T, Russell PA, John T, Thapa B, et al. Assessing Tumor-Infiltrating Lymphocytes in Solid Tumors: A Practical Review for Pathologists and Proposal for a Standardized Method From the International Immuno-Oncology Biomarkers Working Group: Part 2: TILs in Melanoma, Gastrointestinal Tract Carcinomas, Non-Small Cell Lung Carcinoma and Mesothelioma, Endometrial and Ovarian Carcinomas, Squamous Cell Carcinoma of the Head and Neck, Genitourinary Carcinomas, and Primary Brain Tumors. *Adv Anatomic Pathol* (2017) 24(6):311–35. doi: 10.1097/pap.0000000000000161
- Egelston CA, Avalos C, Tu TY, Rosario A, Wang R, Solomon S, et al. Resident Memory Cd8+ T Cells Within Cancer Islands Mediate Survival in Breast Cancer Patients. *JCI Insight* (2019) 4(19):e130000. doi: 10.1172/jci.insight.130000

## ACKNOWLEDGMENTS

We thank all colleagues involved in the study for their contributions.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.924498/full#supplementary-material>

**Supplementary Figure 1** | Kaplan-Meier survival curve of model immune terms, and the best cutoff of IRS in training and validation cohort. **(A)** Kaplan-Meier survival curve of model immune terms; **(B)** The best cutoff of IRS in training and validation cohort.

**Supplementary Figure 2** | GSEA enrichment analysis of IRS. **(A)** GSEA enrichment analysis of IRS model with all the GSEA gene set; **(B)** Representative pathway of each gene set.

**Supplementary Figure 3** | The overview of gene mutation of TCGA-NSCLC.

**Supplementary Figure 4** | Univariate and multivariate Cox regression analysis of IRS and clinical features. **(A)** Univariate Cox regression analysis of IRS and clinical features; **(B)** Multivariate Cox regression analysis of IRS and clinical features.

- 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. doi: 10.1038/nature13954
- Capone M, Giannarelli D, Mallardo D, Madonna G, Festino L, Grimaldi AM, et al. Baseline Neutrophil-To-Lymphocyte Ratio (Nlr) and Derived Nlr Could Predict Overall Survival in Patients With Advanced Melanoma Treated With Nivolumab. *J Immunother Cancer* (2018) 6(1):74. doi: 10.1186/s40425-018-0383-1
- Valero C, Lee M, Hoen D, Weiss K, Kelly DW, Adusumilli PS, et al. Pretreatment Neutrophil-To-Lymphocyte Ratio and Mutational Burden as Biomarkers of Tumor Response to Immune Checkpoint Inhibitors. *Nat Commun* (2021) 12(1):729. doi: 10.1038/s41467-021-20935-9
- Gibney GT, Weiner LM, Atkins MB. Predictive Biomarkers for Checkpoint Inhibitor-Based Immunotherapy. *Lancet Oncol* (2016) 17(12):e542–e51. doi: 10.1016/s1470-2045(16)30406-5
- Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. Ifn- $\gamma$ -Related Mrna Profile Predicts Clinical Response to Pd-1 Blockade. *J Clin Invest* (2017) 127(8):2930–40. doi: 10.1172/JCI91190
- Johnson DB, Estrada MV, Salgado R, Sanchez V, Doxide DB, Opalenik SR, et al. Melanoma-Specific Mhc-Ii Expression Represents a Tumour-Autonomous Phenotype and Predicts Response to Anti-Pd-1/Pd-L1 Therapy. *Nat Commun* (2016) 7:10582. doi: 10.1038/ncomms10582
- Zhou L, Zhang M, Li R, Xue J, Lu Y. Pseudoprogression and Hyperprogression in Lung Cancer: A Comprehensive Review of Literature. *J Cancer Res Clin Oncol* (2020) 146(12):3269–79. doi: 10.1007/s00432-020-03360-1
- Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekas S, et al. Irecist: Guidelines for Response Criteria for Use in Trials Testing Immunotherapeutics. *Lancet Oncol* (2017) 18(3):e143–e52. doi: 10.1016/s1470-2045(17)30074-8
- Song Y, Yan S, Fan W, Zhang M, Liu W, Lu H, et al. Identification and Validation of the Immune Subtypes of Lung Adenocarcinoma: Implications for Immunotherapy. *Front Cell Dev Biol* (2020) 8:550. doi: 10.3389/fcell.2020.00550
- Kalli M, Mpekris F, Wong CK, Panagi M, Ozturk S, Thiagalingam S, et al. Activin A Signaling Regulates Il13 $\alpha$ 2 Expression to Promote Breast Cancer Metastasis. *Front Oncol* (2019) 9:32. doi: 10.3389/fonc.2019.00032
- Papageorgis P, Ozturk S, Lambert AW, Neophytou CM, Tzatsos A, Wong CK, et al. Targeting Il13 $\alpha$ 2 Activates Stat6-Tp63 Pathway to Suppress Breast

- Cancer Lung Metastasis. *Breast Cancer Res: BCR* (2015) 17(1):98. doi: 10.1186/s13058-015-0607-y
19. Kallies A, Zehn D, Utzschneider DT. Precursor Exhausted T Cells: Key to Successful Immunotherapy? *Nat Rev Immunol* (2020) 20(2):128–36. doi: 10.1038/s41577-019-0223-7
  20. Abiko K, Matsumura N, Hamanishi J, Horikawa N, Murakami R, Yamaguchi K, et al. Ifn- $\gamma$  From Lymphocytes Induces Pd-L1 Expression and Promotes Progression of Ovarian Cancer. *Br J Cancer* (2015) 112(9):1501–9. doi: 10.1038/bjc.2015.101
  21. Kurtulus S, Madi A, Escobar G, Klapholz M, Nyman J, Christian E, et al. Checkpoint Blockade Immunotherapy Induces Dynamic Changes in Pd-1(-) Cd8(+) Tumor-Infiltrating T Cells. *Immunity* (2019) 50(1):181–94 e6. doi: 10.1016/j.immuni.2018.11.014
  22. Chen J, Tan Y, Sun F, Hou L, Zhang C, Ge T, et al. Single-Cell Transcriptome and Antigen-Immunoglobulin Analysis Reveals the Diversity of B Cells in Non-Small Cell Lung Cancer. *Genome Biol* (2020) 21(1):152. doi: 10.1186/s13059-020-02064-6
  23. Song P, Li W, Wu X, Qian Z, Ying J, Gao S, et al. Integrated Analysis of Single-Cell and Bulk Rna-Sequencing Identifies a Signature Based on B Cell Marker Genes to Predict Prognosis and Immunotherapy Response in Lung Adenocarcinoma. *Cancer Immunol Immunother* (2022). doi: 10.1007/s00262-022-03143-2
  24. Nie M, Yao K, Zhu X, Chen N, Xiao N, Wang Y, et al. Evolutionary Metabolic Landscape From Preneoplasia to Invasive Lung Adenocarcinoma. *Nat Commun* (2021) 12(1):6479. doi: 10.1038/s41467-021-26685-y
  25. Svensson RU, Parker SJ, Eichner LJ, Kolar MJ, Wallace M, Brun SN, et al. Inhibition of Acetyl-Coa Carboxylase Suppresses Fatty Acid Synthesis and Tumor Growth of Non-Small-Cell Lung Cancer in Preclinical Models. *Nat Med* (2016) 22(10):1108–19. doi: 10.1038/nm.4181
  26. Reck M, Carbone DP, Garassino M, Barlesi F. Targeting Kras in Non-Small-Cell Lung Cancer: Recent Progress and New Approaches. *Ann Oncol* (2021) 32(9):1101–10. doi: 10.1016/j.annonc.2021.06.001
  27. Cho HC, Huang Y, Hung JT, Hung TH, Cheng KC, Liu YH, et al. Puf-A Promotes Cancer Progression by Interacting With Nucleophosmin in Nucleolus. *Oncogene* (2022) 41(8):1155–65. doi: 10.1038/s41388-021-02138-0
  28. Shen Y, Li C, Zhou L, Huang JA. G Protein-Coupled Oestrogen Receptor Promotes Cell Growth of Non-Small Cell Lung Cancer Cells Via Yap1/Qki/Circnotch1/M6a Methylated Notch1 Signalling. *J Cell Mol Med* (2021) 25(1):284–96. doi: 10.1111/jcmm.15997
  29. Christinat Y, Chaskar P, Clement S, Ho L, Charrier M, McKee T, et al. Automated Detection of Arm-Level Alterations for Individual Cancer Patients in the Clinical Setting. *J Mol Diagn* (2021) 23(12):1722–31. doi: 10.1016/j.jmoldx.2021.08.003
  30. Davoli T, Uno H, Wooten EC, Elledge SJ. Tumor Aneuploidy Correlates With Markers of Immune Evasion and With Reduced Response to Immunotherapy. *Science* (2017) 355(6322):eaaf8399. doi: 10.1126/science.aaf8399
  31. Ben-David U, Amon A. Context Is Everything: Aneuploidy in Cancer. *Nat Rev Genet* (2020) 21(1):44–62. doi: 10.1038/s41576-019-0171-x
  32. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. *N Engl J Med* (2017) 376(22):2109–21. doi: 10.1056/NEJMoa1616288
  33. Kim HS, Cha H, Kim J, Park WY, Choi YL, Sun JM, et al. Genomic Scoring to Determine Clinical Benefit of Immunotherapy by Targeted Sequencing. *Eur J Cancer* (2019) 120:65–74. doi: 10.1016/j.ejca.2019.08.001
  34. Teixeira VH, Pipinikas CP, Pennycuik A, Lee-Six H, Chandrasekharan D, Beane J, et al. Deciphering the Genomic, Epigenomic, and Transcriptomic Landscapes of Pre-Invasive Lung Cancer Lesions. *Nat Med* (2019) 25(3):517–25. doi: 10.1038/s41591-018-0323-0
  35. Liu L, Bai X, Wang J, Tang XR, Wu DH, Du SS, et al. Combination of Tmb and Cna Stratifies Prognostic and Predictive Responses to Immunotherapy Across Metastatic Cancer. *Clin Cancer Res* (2019) 25(24):7413–23. doi: 10.1158/1078-0432.Ccr-19-0558
  36. Concepcion CP, Ma S, LaFave LM, Bhutkar A, Liu M, DeAngelo LP, et al. Smarca4 Inactivation Promotes Lineage-Specific Transformation and Early Metastatic Features in the Lung. *Cancer Discovery* (2022) 12(2):562–85. doi: 10.1158/2159-8290.Cd-21-0248
  37. Farago AF, Taylor MS, Doebele RC, Zhu VW, Kummar S, Spira AI, et al. Clinicopathologic Features of Non-Small-Cell Lung Cancer Harboring an Ntrk Gene Fusion. *JCO Precis Oncol* (2018). doi: 10.1200/po.18.00037
  38. Zhang Y, Tseng JT, Lien IC, Li F, Wu W, Li H. Mrnasi Index: Machine Learning in Mining Lung Adenocarcinoma Stem Cell Biomarkers. *Genes (Basel)* (2020) 11(3):257. doi: 10.3390/genes11030257
  39. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine Learning Identifies Stemness Features Associated With Oncogenic Dedifferentiation. *Cell* (2018) 173(2):338–54.e15. doi: 10.1016/j.cell.2018.03.034
  40. Wu J, Li L, Zhang H, Zhao Y, Zhang H, Wu S, et al. A Risk Model Developed Based on Tumor Microenvironment Predicts Overall Survival and Associates With Tumor Immunity of Patients With Lung Adenocarcinoma. *Oncogene* (2021) 40(26):4413–24. doi: 10.1038/s41388-021-01853-y
  41. Tan DS, Haaland B, Gan JM, Tham SC, Sinha I, Tan EH, et al. Bosutinib Inhibits Migration and Invasion Via Ack1 in Kras Mutant Non-Small Cell Lung Cancer. *Mol Cancer* (2014) 13:13. doi: 10.1186/1476-4598-13-13
  42. Zhao H, Yang F, Shen W, Wang Y, Li X, You J, et al. Pazopanib Diminishes Non-Small Cell Lung Cancer (Nslc) Growth and Metastases in Vivo. *Thorac Cancer* (2015) 6(2):133–40. doi: 10.1111/1759-7714.12138
  43. Fatehi Hassanabad A, MacQueen KT. Molecular Mechanisms Underlining the Role of Metformin as a Therapeutic Agent in Lung Cancer. *Cell Oncol (Dordrecht)* (2021) 44(1):1–18. doi: 10.1007/s13402-020-00570-0

**Conflict of Interest:** The 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:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Yao, Wang, Yang and Gao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.