



## OPEN ACCESS

## EDITED BY

Nan Zhang,  
Harbin Medical University, China

## REVIEWED BY

Kui Zhang,  
The University of Chicago,  
United States  
Kewa Gao,  
University of California, Davis,  
United States

## \*CORRESPONDENCE

Huan Xu  
joyxu1027@163.com  
Zhixin Li  
saklizhixin@163.com

## SPECIALTY SECTION

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

RECEIVED 18 August 2022

ACCEPTED 12 September 2022

PUBLISHED 27 September 2022

## CITATION

Gu X, Wei S, Li Z and Xu H (2022)  
Machine learning reveals two  
heterogeneous subtypes to assist  
immune therapy based on lipid  
metabolism in lung adenocarcinoma.  
*Front. Immunol.* 13:1022149.  
doi: 10.3389/fimmu.2022.1022149

## COPYRIGHT

© 2022 Gu, Wei, Li and Xu. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Machine learning reveals two heterogeneous subtypes to assist immune therapy based on lipid metabolism in lung adenocarcinoma

Xuyu Gu<sup>1</sup>, Shiyu Wei<sup>2</sup>, Zhixin Li<sup>3\*</sup> and Huan Xu<sup>2\*</sup>

<sup>1</sup>School of Medicine, Southeast University, Nanjing, China, <sup>2</sup>Department of Anesthesiology, Shanghai Pulmonary Hospital, School of Medicine, Tongji University, Shanghai, China, <sup>3</sup>Department of Thoracic Surgery, Shanghai Pulmonary Hospital, School of Medicine, Tongji University, Shanghai, China

**Background:** Lipid metabolism pivotally contributes to the incidence and development of lung adenocarcinoma (LUAD). The interaction of lipid metabolism and tumor microenvironment (TME) has become a new research direction.

**Methods:** Using the 1107 LUAD records from the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases, a comprehensive exploration was performed on the heterogeneous lipid metabolism subtypes based on lipid metabolism genes (LMGs) and immune-related genes (LRGs). The clinical significance, functional status, TME interaction and genomic changes of different subtypes were further studied. A new scoring system, lipid-immune score (LIS), was developed and validated.

**Results:** Two heterogeneous subtypes, which express more LMGs and show the characteristics of tumor metabolism and proliferation, are defined as lipid metabolism phenotypes. The prognosis of lipid metabolism phenotype is poor, and it is more common in patients with tumor progression. Expressing more IRGs, enrichment of immunoactive pathways and infiltration of effector immune cells are defined as immunoactive phenotypes. The immunoactive phenotype has a better prognosis and stronger anti-tumor immunity and is more sensitive to immunotherapy. In addition, KEAP1 is a driving mutant gene in the lipid metabolism subtype. Finally, LIS was developed and confirmed to be a robust predictor of overall survival (OS) and immunotherapy in LUAD patients.

**Conclusion:** Two heterogeneous subtypes of LUAD (lipid metabolism subtype and immune activity subtype) were identified to evaluate prognosis and immunotherapy sensitivity. Our research promotes the understanding of the interaction between lipid metabolism and TME and offers a novel direction for clinical management and precision therapy aimed to LUAD patients.

## KEYWORDS

lung adenocarcinoma, machine learning, tumor microenvironment, lipid metabolism, immunotherapy

## Introduction

As the most frequent malignancy, lung cancer causes the highest cancer-related deaths around the world (1). Lung cancer can appear in different histological types among which, non-small cell lung cancer (NSCLC) is the most common type with about 85% proportion of all lung cancer patients (2). Lung adenocarcinoma (LUAD) is the most abundant subtype of NSCLC, accounting for about 55% (3). LUAD is a heterogeneous disease with different clinical prognosis and drug response results. It is worth noting that despite the great progress in clinical diagnostic methods and multimodal treating approaches, the 5-year overall survival (OS) rate of patients with advanced lung cancer has remained very low (4). Therefore, LUAD patients are still in an urgent need for new early diagnosis and clinical intervention methods.

During cancer occurrence and progression, the immune system and the tumor cells are in complex interaction. On one hand, demand for local nutrition and oxygen highly increases due to fast proliferation of tumor cells. On the other hand, the same reason causes poor local vascularization, resulting in acidosis and hypoxia in the tumor microenvironment (TME) as well as local glucose deficient (5–7). Eventually, lipids existing in the TME begin to be used main as the alternative source of energy in both tumor tissues and immune cells to compensate for the energy shortages (8). Lipids also contribute to the biofilm formation, supplying the biomass production, and mediating some complex signaling pathways contributing to the growth and migration of cancer cells (9). In addition, the affinity of cancerous cells for lipids and cholesterol increases, directly leading to lipid accumulation in the TME and developing malignancy in the tumor tissues (10). Although, the lipid metabolic reprogramming and dysfunction as well as its dual impact in the TME and immune responses to tumor is not exactly recognized yet. Such further elaboration is essentially required for developing specific treatments based on the anti-tumor immune responses.

The present study aimed to survey the crosstalk between lipid metabolism and tumor immune response in LUAD patients and identified two heterogeneous subtypes (lipid metabolism subtype and immune activity subtype). These two subtypes show specific differences in clinical outcomes, biological functions, immune infiltration and genomic variation. In addition, a lipid-immune score (LIS) was developed and validated, which shows significant advantages

in predicting prognosis and immunotherapy response. In conclusion, our work strengthens the understanding of the complex role between lipid metabolism and immune system in LUAD and provides a new perspective and reference for the accurate prediction and immunotherapy of LUAD patients.

## Methods

### Data extraction

Transcriptome RNA-seq data (HT-seq FPKM), mutation data (mutect2 tool), copy number variation (CNV), and their corresponding clinical information (from Cancer Genome Atlas-lung adenocarcinoma, tcga-LUAD queue) were obtained (<https://portal.gdc.cancer.gov/repository>). After excluding patients who lost follow-up and clinical information, 492 LUAD samples were collected. These data were used as discovery queues after Transcripts per million (TPM) standardization. In addition, three independent data sets from the GEO database were collected, including GSE30219 from GPL570 platform, GSE42127 from GPL6884 platform, and GSE72094 from GPL15048 platform. In order to prevent the batch effect of chips, three GEO data sets were combined through the combat function of the “sva” package and the data were log<sub>2</sub> standardized (11). Finally, a total of 615 GEO meta queues containing LUAD samples with complete clinical information were used as external validation queues. Finally, two immunotherapy cohorts were collected to verify the model's prognostic power: NSCLC cohort GSE135222 receiving Programmed Death-1(PD-1) treatment, including 27 patients (12) and Imvigor210, a cohort of advanced urothelial carcinoma cases undergoing anti-Programmed Cell Death-Ligand 1 (PD-L1) immunotherapy, including 298 patients (13).

### Identification of lipid and immune subtypes of LUAD

Lipid metabolism genes (LMGs) were obtained from the Molecular Signatures Database (MSigDB) (<http://www.gsea-msigdb.org/gsea/index.jsp>), containing 1426 LMGs (14). The immune-related genes (IRGs) were obtained from the “ImmPort” database (<https://www.immport.org/resources>) (15). It contains a total of 1638 IRGs defined as functional and

immune related. A detailed list in [Table S1](#) to indicate the lipid genes and immune genes we used. First, LMGs and IRGs with independent prognostic efficacy were evaluated by univariate Cox regression analysis, and candidate genes were identified according to the threshold of  $p < 0.05$ . According to the transcriptional map of candidate genes, consensus clustering was conducted in the discovery queue and validation queue through the ConsensusClusterPlus package (16). Pam unsupervised clustering algorithm was adopted in this analysis, and 1000 iterations were carried out based on Euclidean distance. Eighty per cent of the samples were randomly selected in each iteration. The number of clusters was set to 2–5, and the optimal cluster number was jointly determined using the consensus matrix and cumulative distribution function (CDF).

## Functional enrichment and immune infiltration analysis

Significant differentially expressed genes (DEGs) between subgroups were identified by ‘limma’ package in R program according to the threshold of False Discovery Rate (FDR)  $< 0.05$  and fold change (FC)  $> 2$ . The functional enrichment of DEGs was achieved using metascape ([www.metascape.org/](http://www.metascape.org/)) database. Gene Set Enrichment Analysis (GSEA) was conducted among subgroups and significantly altered pathways were selected using Kyoto Encyclopedia of Genes and Genomes (KEGG) by  $p < 0.05$ .

Based on the previously published molecular markers, ssGSEA analysis was performed using the ‘gsva’ package in R program to evaluate the biological pathway activity of the samples which included angiogenesis, epithelial-mesenchymal transition (EMT), myoid inflammation, and molecular markers of other immune related pathways (17–20). Molecular markers of hypoxia were collected from MSigDB (14). Detailed pathway gene markers were displayed in [Table S2](#). The relative infiltration abundance of 22 different immune cell types was evaluated by ‘cibersoft’ package in R program (21). The immune activity and tumor purity of tumor samples were evaluated by Estimate algorithm (22). Finally, the microsatellite instability (MSI) score, indel neoantigens and SNV neoantigens of samples from the study of Thorsson et al. (23).

## Analysis of the genome variation map between subgroups

The mutation data was processed with ‘maftools’ package in R package. First, the total number of mutations in the sample was measured, and then, the genes with the minimum mutation number  $> 30$  were identified. The difference of mutation frequency of high-frequency mutation genes between the two

subgroups was compared using the chi square test and visualized with maftools (24). CNV data were processed by Gistic 2.0 software. Based on the threshold of 0.2, significantly amplified and deleted chromosome segments were identified, and CNV differences on chromosome arms were evaluated. The CNV results were visualized by ‘ggplot2’ R package.

## Constructing lipid-immune score

DEGs contained in all cohorts were selected for further analysis based on the above identified DEGs between the two subtypes. Univariate Cox regression analysis revealed the prognostic value of these genes. Subsequently, genes with statistical significance ( $p < 0.05$ ) were incorporated into the Cox proportional hazard model with Least absolute shrinkage and selection operator (Lasso) penalty, and 300 iterative searches were carried out to find the most robust model. In order to prevent over fitting, five-cross validation was set up. The model with the highest frequency in 300 iterations was used as the final prognostic model and the lipid-immune score (LIS) was generated according to the formula:  $LIS = \sum_i \text{Coefficient}(mRNA_i) \times \text{Expression}(mRNA_i)$ . The ‘survcomp’ package in R program was used to calculate the consistency C index and evaluate the prognostic value of the risk score (RS) in the training and verification sets. The higher C index indicates the more accurate prognostic power of the model (25). The high-risk group and low-risk group were divided based on their median FRS, and the prognostic value of the risk model was calculated using Kaplan-Meier (KM) survival curve, univariate and multivariate Cox regression, time-dependent ROC curve (tROC), and subgroup analysis system.

## Predicting immunotherapy response

According to previous studies, the immunophenoscore (IPS) of the sample was calculated. Briefly, IPS is calculated from transcriptomic data of representative genes for different immunophenotypes and normalizes the final result to 0–10. Samples were positively weighted according to effective immune cells and negatively weighted according to suppressive immune cells, and then applied Z-score averaged. Z-score  $\geq 3$  was defined as IPS10 and Z-score  $\leq 0$  was defined as IPS0. The higher the IPS, the better the immunotherapy response (26). The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm (<http://tide.dfci.harvard.edu>) was applied to predict the patients’ response to the anti-PD-1 and anti-CTLA-4 treatments (27–30). Finally, the predictive power of LIS was evaluated in two external immunotherapy cohorts (GSE135222 and Imvigor210).

## Statistical analysis

Pearson chi square or Fisher exact tests were applied to compare categorical variables. The continuous variables were compared between the two groups by Wilcoxon rank sum test. The KM curve was drawn by 'survminer' package and the tROC analysis was carried out by 'survivalROC' package both in R program. The univariate and multivariate Cox regression was completed by 'survival' package in R program. The 'rms' package in R was used to draw nomograms and calibration curves, and decision curve analysis (DCA) was carried out through DCA package (31). The ROC curve used to predict immunotherapy was performed by the 'pROC' package. Two tailed  $p < 0.05$  was considered statistically significant unless otherwise specified.

## Results

### Parsing LMGs and IRGs in LUAD

The design of our study is shown in Figure S1. Univariate Cox regression analysis displayed 155 LMGs and IRGs with prognostic value ( $p < 0.05$ ). The forest map showed the prognostic candidate genes of top15 (Figure 1A). Detailed Cox results are provided in Table S3. Figure 1B summarizes the mutation of top15 candidate genes. Specifically, the mutation type is single nucleotide mutation, and the genes with the highest mutation frequency are VEGFC (24%) and tnfrsf11a (10%). The waterfall diagram shows their mutation map in the tcga-LUAD cohort (Figure 1C). The histogram summarizes the CNV of the top15 candidate genes in tcga-LUAD, and the results show that

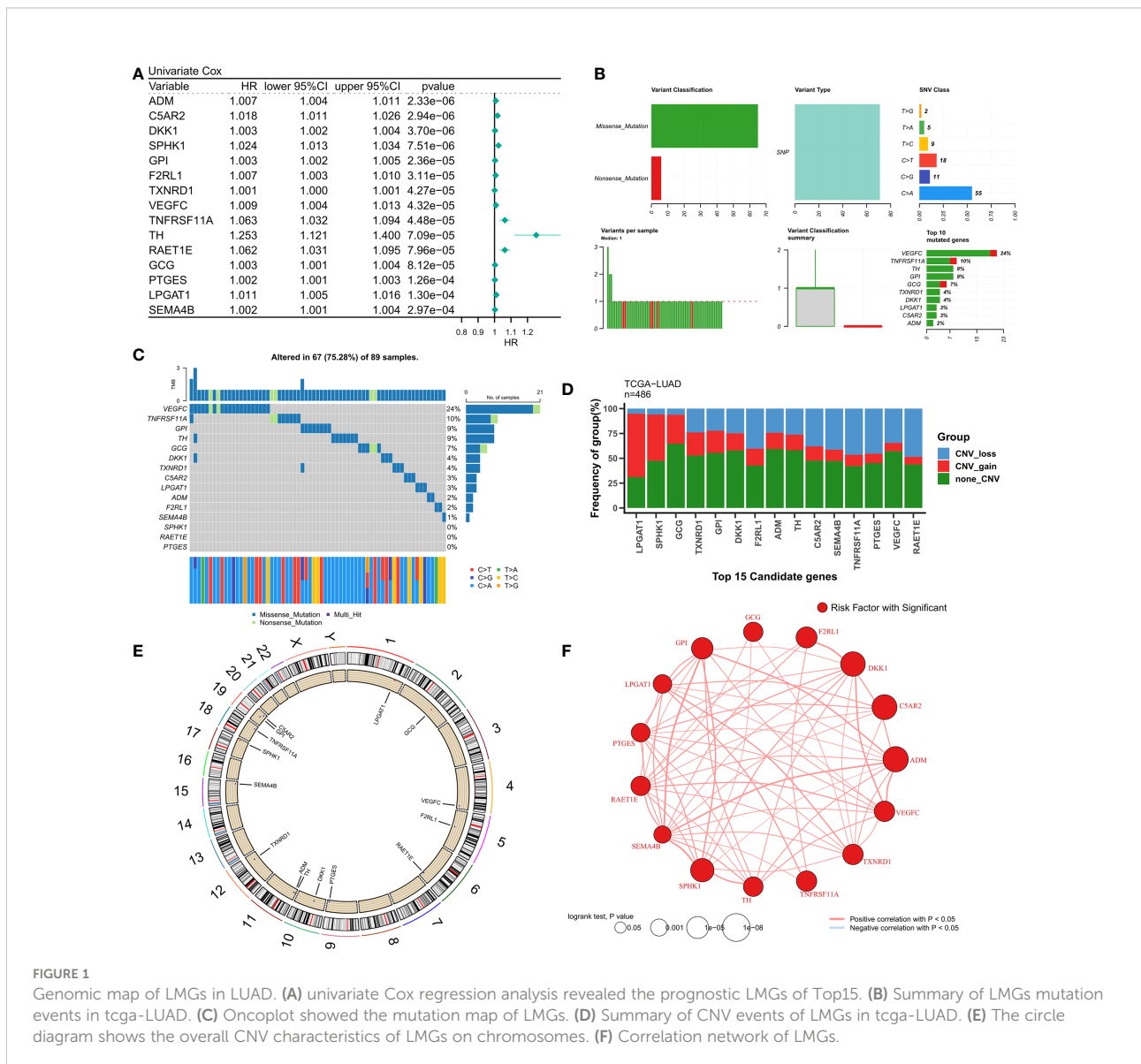


FIGURE 1

Genomic map of LMGs in LUAD. (A) univariate Cox regression analysis revealed the prognostic LMGs of Top15. (B) Summary of LMGs mutation events in tcga-LUAD. (C) OncoPrint showed the mutation map of LMGs. (D) Summary of CNV events of LMGs in tcga-LUAD. (E) The circle diagram shows the overall CNV characteristics of LMGs on chromosomes. (F) Correlation network of LMGs.

they have a wide range of CNV events. Lpgat1 was the most amplified gene, and raet1e was the most deleted gene (Figure 1D). The circle chart shows the overall CNV of the top15 candidate gene on the chromosome (Figure 1E). Finally, the interaction of top15 candidate genes was analyzed, and the correlation network showed that they were highly positively correlated (Figure 1F).

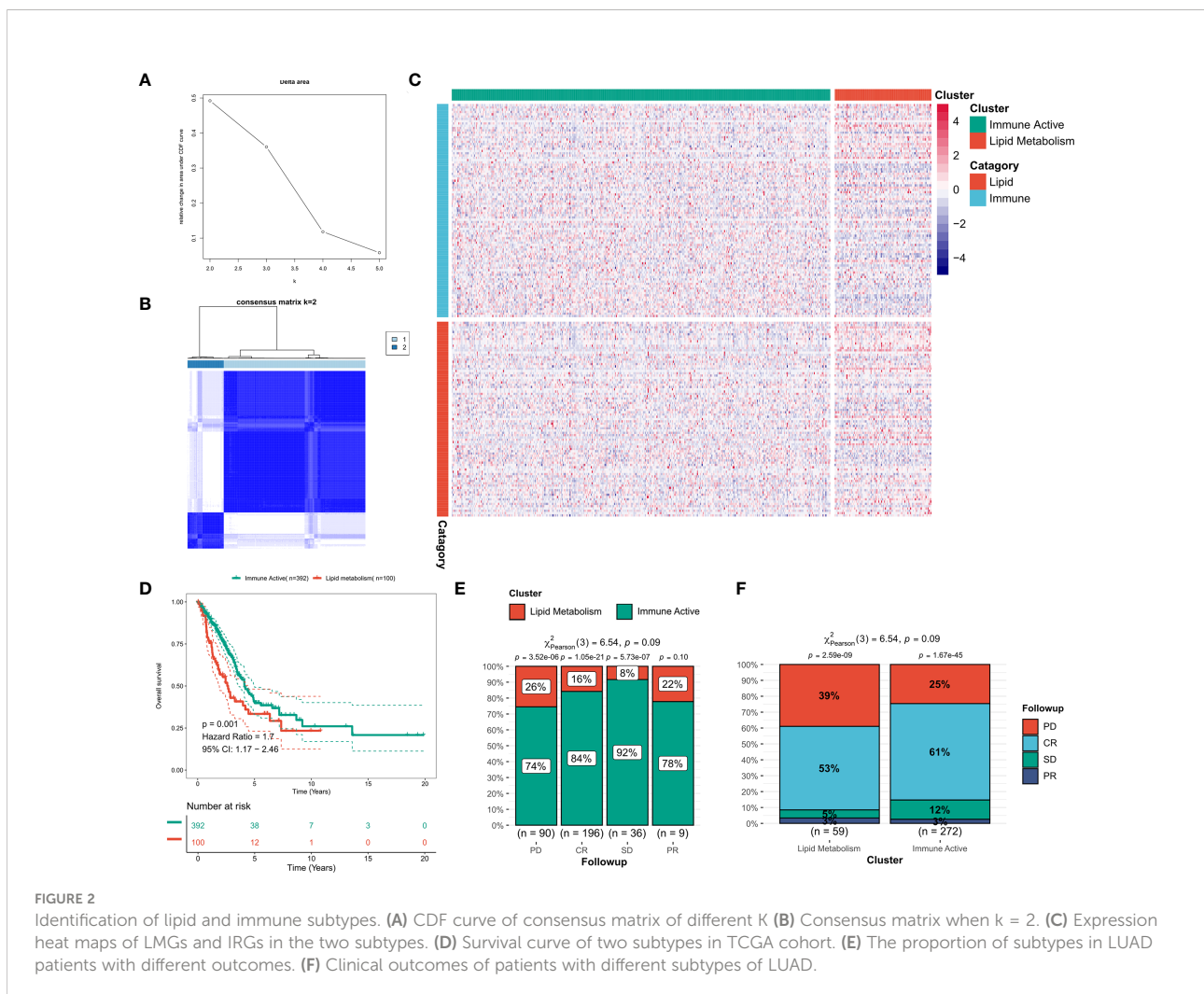
## Identification of lipid and immune subtypes

Consensus clustering was performed on the discovery queue and GEO meta queue from tcga-LUAD by ConsensusClusterPlus. According to the CDF curve of consensus score, k = 2 was found to be the best choice (Figure 2A, Figure S2A). The consensus matrix also confirmed this result (Figure 2B, Figure S2B). Based on the transcriptional profiles of candidate LMGs and IRGs, lipid metabolism subtypes and immune activity subtypes were defined

(Figure 2C, Figure S2C). IRGs were significantly increased in immunoactive subtypes, while LMGs were significantly increased in lipid metabolism subtypes. According to the survival analysis, lipid metabolism subtypes in the cohort was significantly worse compared to that of immune activity subtypes (p = 0.001, Figure 2D). A worse clinical outcome of lipid metabolism subtypes was confirmed in the validation cohort (p < 0.001, Figure S2D). In addition, the tcga-LUAD cohort had more detailed clinical follow-up information. There was a significant increase in patients with disease progression in the lipid metabolism subtype (Figures 2E, F).

## Biological function difference between two subtypes

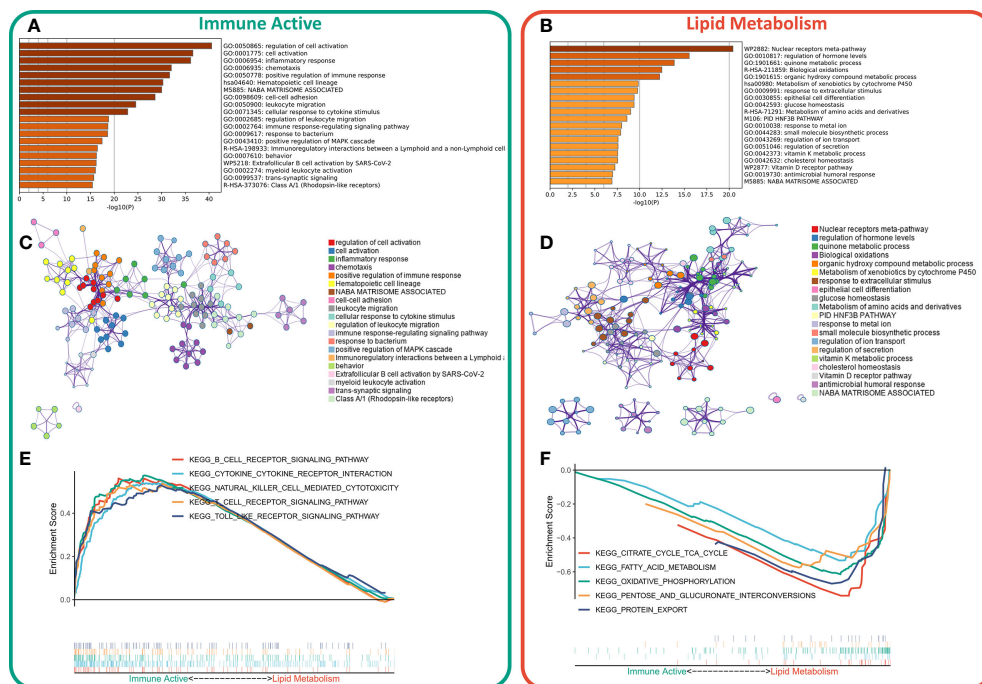
First, the DEGs between the two subtypes were identified by limma package. According to the threshold of FDR < 0.05 and FC > 2, a total of 1597 DEGs were identified, of which



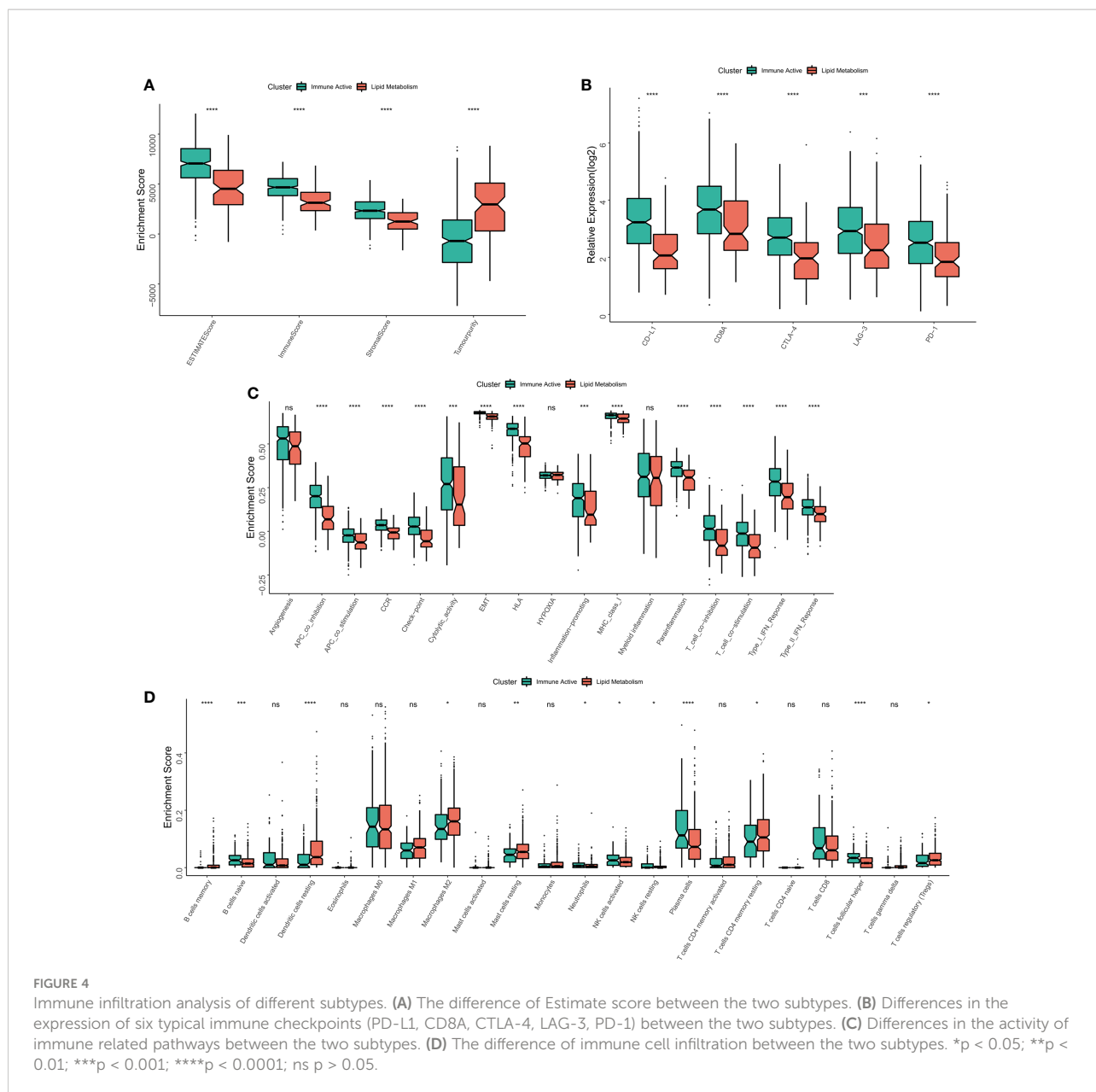
1233 were up-regulated in immunocompetent subtypes and 362 were up-regulated in lipid metabolism subtypes. Detailed results are provided in Table S4. Based on the functional enrichment analysis, the up-regulated genes in immunocompetent subtypes mainly regulate cell activation, inflammatory response, cell adhesion and lymphocyte migration (Figure 3A), and Figure 3C showed the functional interaction network of immunocompetent subtypes. The up-regulated genes in lipid metabolism subtypes mainly regulate biological oxidation, epithelial cell differentiation and glucose homeostasis (Figure 3B). Figure 3D shows the functional interaction network of lipid metabolism subtypes. GSEA analysis showed that the pathways enriched in immunocompetent subtypes were mainly B-cells receptor, T-cells receptor, Toll-like receptor signal pathway and NK-cells killing activity (Figure 3E). The pathways enriched in lipid metabolism subtypes were fatty acid metabolism, protein secretion and TCA cycle pathway (Figure 3F). In conclusion, these results confirm that the immunocompetent subtype has stronger antitumor immune activity, while the tumor cells of the lipid metabolism subtype have stronger metabolic and proliferative activity, which may lead to the difference in prognosis between the two.

### Difference of immune infiltration between two subtypes

The immune infiltration degrees were systematically compared between the two subtypes. First, the estimate algorithm showed that the immune activity subtype had a higher immune score, while the lipid metabolism subtype had a higher tumor purity (Figure 4A), which was confirmed in the validation queue (Figure S3A). The expression differences between five classical immune checkpoints and therapeutic targets (PD-L1, CD8A, CTLA-4, LAG-3, PD-1) were then examined. The results showed that the five checkpoints were significantly up-regulated in immunocompetent subtypes (Figure 4B), and the validation cohort (Figure S3B). Through ssGSEA algorithm, we found that except for myeloid inflammation, other immune pathways were up-regulated in the immune activity subgroups. In particular, the activity of EMT pathway in the immunocompetent pathway was also up-regulated (Figure 4C). Similar results were observed in the validation cohort. It is worth noting that the activity of angiogenesis pathway in lipid metabolism subtypes was up-regulated in the validation cohort (Figure S3C). Finally, cibersort results showed that NK cells, plasma cells and natural B cells increased in immunocompetent subtypes, while Tregs increased



**FIGURE 3** Functional enrichment of different subtypes. (A) Functional enrichment of characteristic genes of immune activity subtypes. (B) Functional enrichment of lipid metabolism subtype characteristic genes. (C) Functional network of immunocompetent subtypes. (D) Functional network of lipid metabolism subtypes. (E) KEGG pathway enriched in immunocompetent subtypes. (F) KEGG pathway enriched in lipid metabolism subtypes.

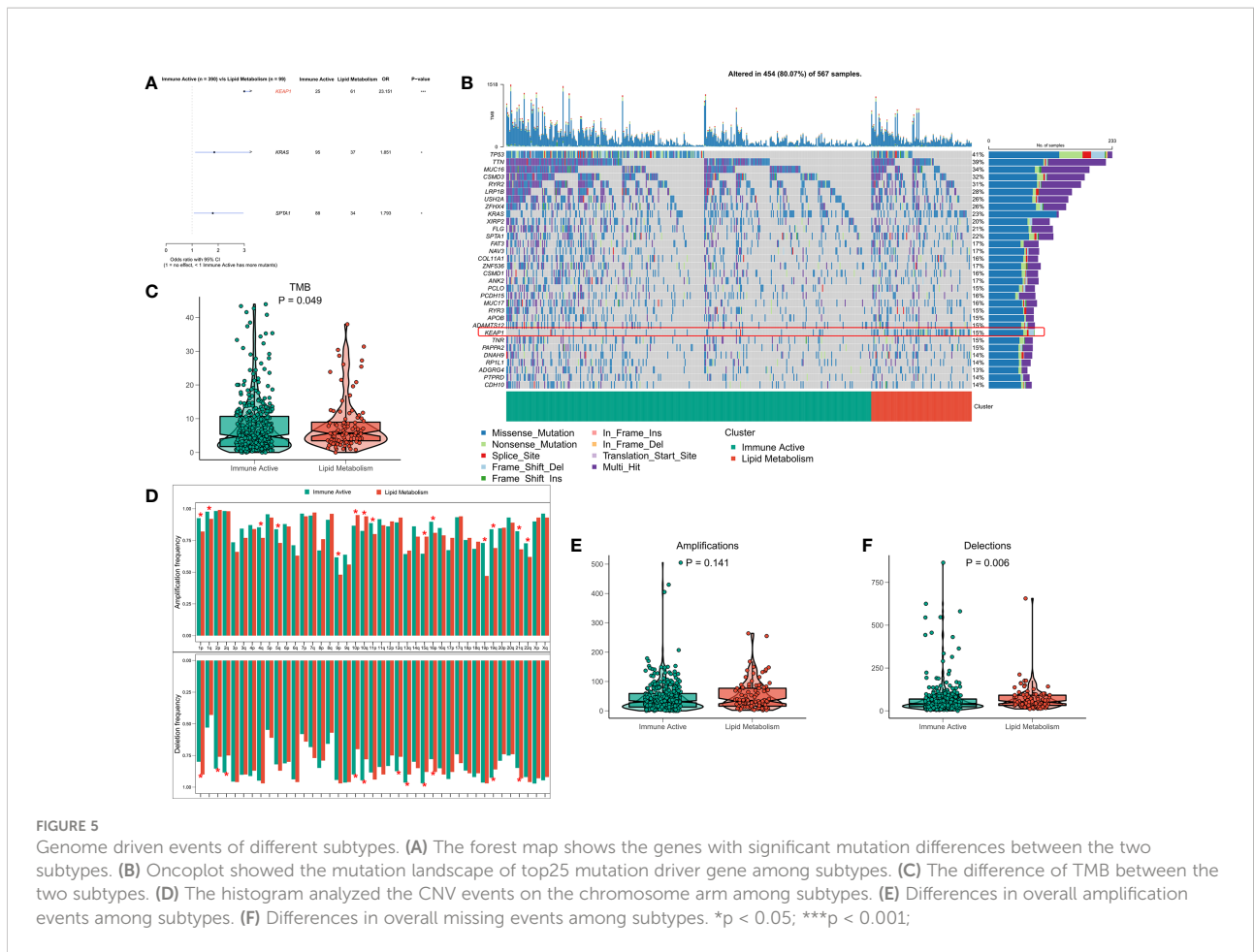


in lipid metabolism subtypes (Figure 4D). In addition, higher Tregs in lipid metabolism subtypes were also confirmed in the validation cohort (Figure S3D). In conclusion, these results convey that the immunoactive subtype has more antitumor immune activity and effector immune cells, while the lipid metabolism subtype is inhibited by higher Tregs infiltration.

### Analysis of genome changes among subtypes

The original mutation data were processed with maftools package. Chi square test showed that the mutation frequency of

KEAP1, KRAS and SPTA1 in lipid metabolism subtypes increased, especially KEAP1 (Figure 5A). The waterfall diagram shows the mutation map difference of a total of 32 high-frequency mutant genes between the two subtypes (Figure 5B). The TMB of each patient was calculated, and the results showed that the lipid metabolism subtype had a higher TMB, but the difference between the two subtypes was not significant (Figure 5C). CNV leads to chromosome variation in another way. We then evaluated the correlation between FRS and CNV and found that the amplification and deletion levels of immunoactive subtypes were significantly higher at the chromosome arm level (Figure 5D). The box diagram showed no significant difference in the total number of chromosome



amplification between the two subtypes (Figure 5E), and the number of chromosome deletions in the lipid metabolism subtype increased significantly (Figure 5F).

### Immunoactive subtypes that are more sensitive to immunotherapy

The functional differences and immune landscape among subgroups suggest that patients with immunoactive subtypes may have better immune treatment response. According to the literature, better immunotherapeutic efficacy is in close relation to the increase in the number of neoantigens (32, 33). Therefore, we first evaluated the difference in the number of neoantigens between the two subtypes, and the results showed that the immunoactive subtypes have more SNV neoantigens and Indel neoantigens (Figures 6A, B). Recent studies have shown that MSI score is expected to become a new predictor of immunotherapy (34). However, there is no significant difference in MSI score between the two subtypes (Figure 6C). IPS can systematically evaluate the activity of effector immune cells and immune treatment response of patients. The discovery queue showed

that IPS of immunoactive subtypes was significantly higher than that of lipid metabolism subtypes (Figure 6D), and the response rate of immunoactive subtypes to immunotherapy predicted by TIDE algorithm was higher than that of lipid metabolism subtypes (Figure 6E). Although there was no significant difference in IPS between the two subtypes in the validation cohort, the immunoactive subtypes in the validation cohort also had a higher response to immunotherapy (Figures 6E, F). In conclusion, our results suggest that immunoactive subtypes are more sensitive to immunotherapy.

### Constructing and validating LIS

First, 1597 DEGs were analyzed by univariate Cox regression to identify the prognostic valuable DEGs. According to the threshold of  $p < 0.001$ , a total of 88 DEGs with prognostic significance were identified. Then, these 88 DEGs were recruited for Lasso regression to simplify the model. After 300 iterations, the model with 22 DEGs was the most stable showing a suitable efficacy in the training queue as well as the validation queue (C index  $> 0.6$ , Figure 7A). According to the best  $\lambda$  (0.02631), the



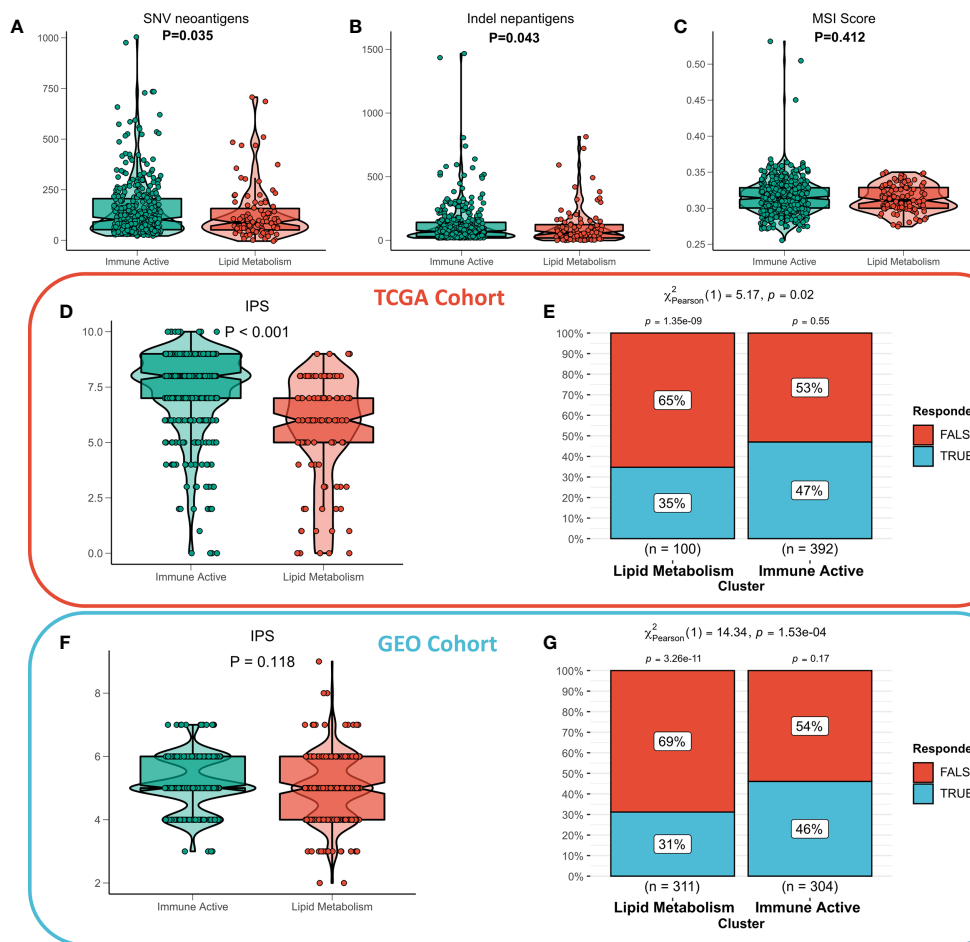


FIGURE 6

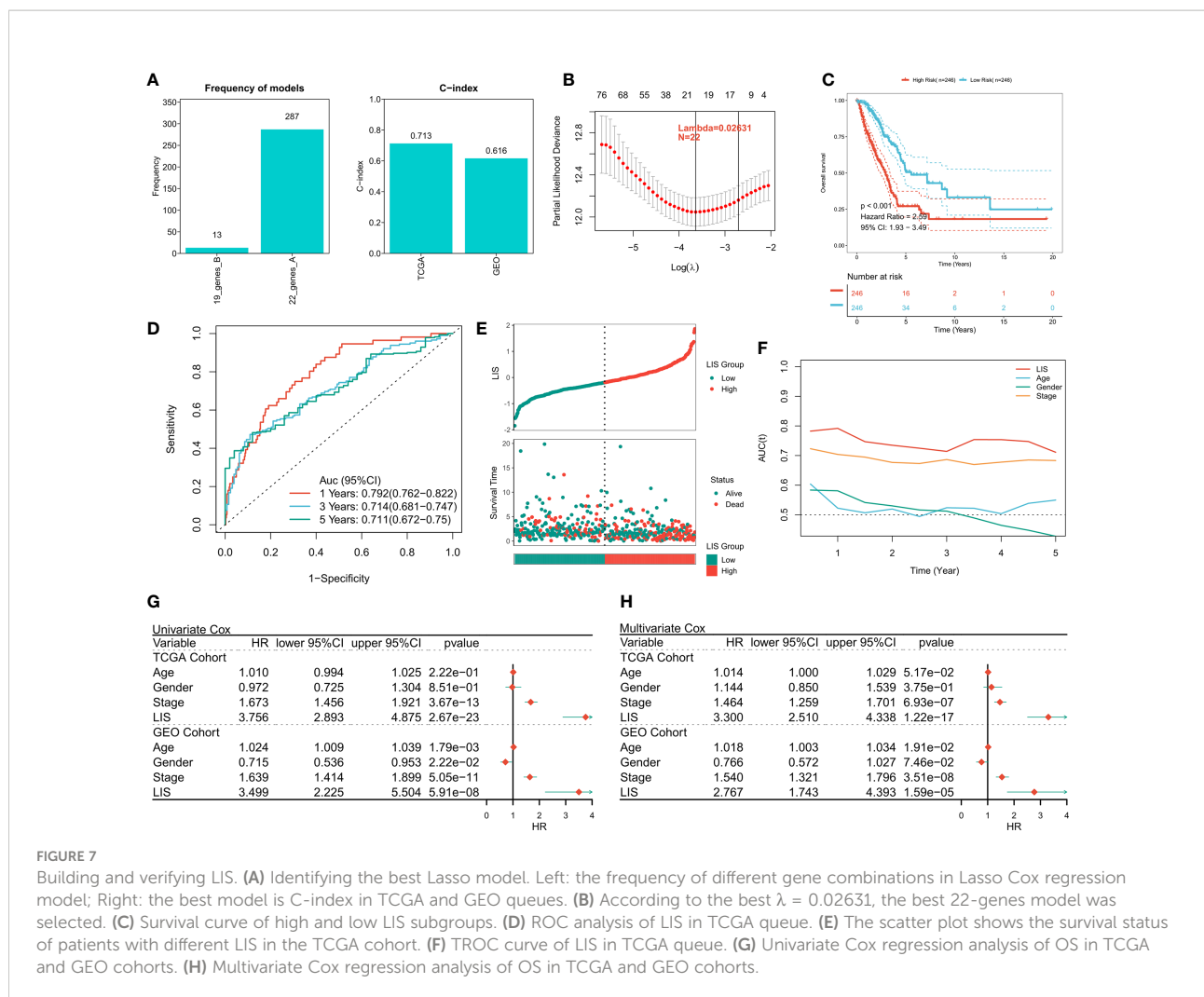
Immunotherapeutic sensitivity of different subtypes. (A) Differences in SNV neoantigens between the two subtypes. (B) The difference of indel neoantigens between the two subtypes. (C) The difference of MSI score between the two subtypes. (D) IPS differences between subtypes in TCGA cohort and (F) GEO cohort. The TIDE algorithm predicts the immune treatment responses of different subtypes in (E) TCGA queue and (G) GEO queue.

best model of 22 genes was identified (Figure 7B), and LIS was generated according to the formula:  $LIS = \sum_i \text{Coefficient}(mRNA_i) \times \text{Expression}(mRNA_i)$ , detailed coefficients of 22 LIS genes can be found in Table S5. According to the survival analysis, patients with high LIS showed significantly less survival rate compared to the patients with low LIS ( $p < 0.001$ , Figure 7C), which was confirmed in the validation cohort ( $p < 0.001$ , Figure S4A). Based on the ROC analysis, the AUC values of the model in 1 year, 3 years, and 5 years were 0.792, 0.714, and 0.711, respectively (Figure 7D). In the external validation queue, LIS also had satisfactory prediction efficiency, specifically, 0.68 in 1 year, 0.69 in 3 years, 0.69 in 5 years and 0.71 in 8 years (Figure S4B). Figure 7E shows that the survival status of patients with high LIS were significantly worse compared that of patients with low LIS, and similar results were observed in the validation cohort (Figure S4C). TROC analysis showed that LIS was the

best predictor of OS (Figure 7F), and the effectiveness of LIS and Stage was equivalent in the validation queue (Figure S4D). Finally, univariate Cox regression confirmed that LIS was an independent prognostic indicator in both training and validation sets ( $p < 0.0001$ , Figure 7G). Multivariate Cox regression showed that LIS was still an independent prognosticator for OS in the training and validation cohorts after correcting for other factors ( $p < 0.0001$ , Figure 7H).

## Quantifying the risk of individual LUAD patients

Subgroup analysis showed that LIS in the training cohort showed excellent predictive ability in different clinical subgroups except patients in stage 3 and stage 4 ( $p < 0.001$ , Figure 8A). In the

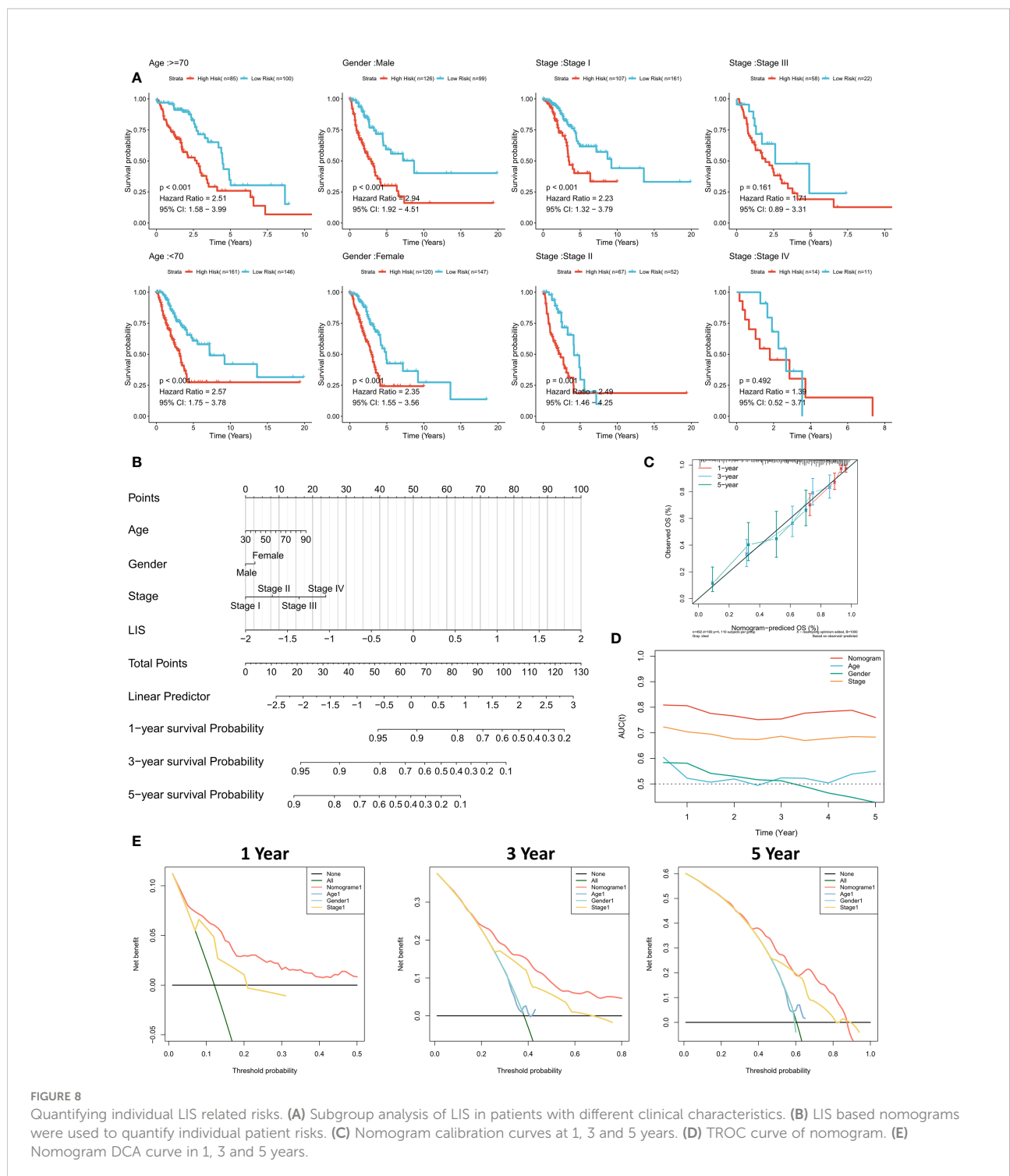


validation cohort, LIS was able to distinguish patients with poor survival except for patients in stage 2-4 ( $p < 0.05$ , Figure S4E). These results suggest that LIS shows better performance in predicting early LUAD patients. For better quantifying of the death risk in individual LUAD patients, nomograms were constructed based on LIS (Figure 8B). Nomogram correction curve shows that nomogram model has good stability and accuracy in 1, 3 and 5 years (Figure 8C). TROC analysis showed that compared with clinical characteristics, nomogram model was the best predictor (Figure 8D). DCA was then performed to calculate the decision-making benefits of nomogram model. The results showed that nomogram was suitable for risk assessment of LUAD patients in 1, 3 and 5 years (Figure 8E).

### LIS in predicting immunotherapy

First, TIDE was used to evaluate the difference of immunotherapy response between patients with high LIS and

patients with low LIS. According to the results, patients with low LIS showed to be more benefitting from immunotherapy (Figure 9A, Figure S5A). Then five widely used immunotherapy biomarkers were calculated, including MDSC, MSI score, IFNG, CD8 and CD274. In the training cohort and validation cohort, LIS provided higher accuracy in predicting immunotherapy (Figure 9B, Figure S5B). Then, two immunotherapy cohorts were included to further study whether LIS could predict patients' response to immunotherapy. Consistent with the above, patients with high LIS showed worse survival in these two immunotherapy cohorts (Figures 9C, D). Finally, the relationship between LIS and neoantigens and TMB in Imvigor210 cohort was evaluated. The results showed that LIS had no strong correlation with neoantigens and TMB. However, patients with low LIS had higher neoantigens (Figures 9E, F). Overall, our study strongly confirms LIS as a prognosis factor for OS and immunotherapeutic response of patients, and is superior to widely used biomarkers.

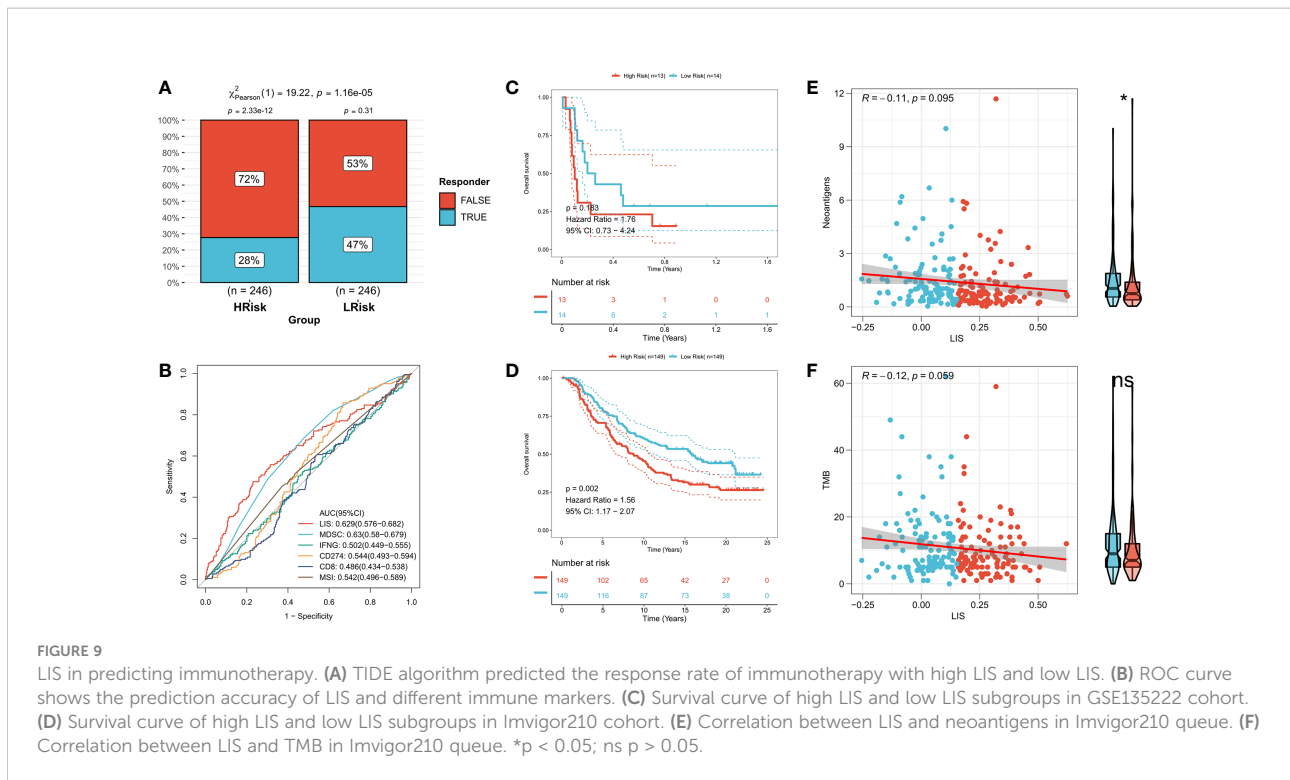


**FIGURE 8** Quantifying individual LIS related risks. **(A)** Subgroup analysis of LIS in patients with different clinical characteristics. **(B)** LIS based nomograms were used to quantify individual patient risks. **(C)** Nomogram calibration curves at 1, 3 and 5 years. **(D)** TROC curve of nomogram. **(E)** Nomogram DCA curve in 1, 3 and 5 years.

## Discussion

Lung cancer is the main cause of cancer-related death and LUAD is the most common histological subtype with the most patients at the advanced stage on the initial diagnosis (35, 36). Although a variety of targeted therapies and new chemotherapeutic drugs have been

approved, the OS of advanced patients is still not ideal (4). Lipid metabolism has long been reported as the main energy source of cancer cells and is involved in the incidence and development of cancer (8). Recently, the dual regulation of lipid metabolism on immune response in TME has attracted extensive attention and has become a promising target for targeted therapy (6, 10). Our study



identified and verified two heterogeneous subtypes in LUAD, one of which was an effector immune cell with more expression of IRGs, enrichment of immunoactive pathways and high abundance, which was defined as an immunoactive subtype. Another kind of suppressive immune cells expressing more LMGs and high abundance, showing the characteristics of tumor metabolism and proliferation, was defined as lipid metabolism subtype. We verified the stability and repeatability of the two subtypes in a GEO meta cohort. These two subtypes also showed heterogeneity in genome driven events, clinical outcomes, and immunotherapy responses. In addition, a robust prognostic feature was proposed based on these two subtypes: LIS. Further analysis showed that LIS shows a leading advantage in predicting the immunotherapy of LUAD patients. These results promote the understanding of the interaction between lipid metabolism and TME and offer a new direction for clinical management and precision treatment of LUAD patients. These two subtypes showed different clinical characteristics. The survival of immunoactive phenotype was significantly better than that of lipid metabolism phenotype, and patients with more disease progression were in lipid metabolism phenotype. Functional enrichment indicated that metabolic related pathways and cell-cycle related pathways were enriched in the lipid metabolism phenotype, while effector immune cell receptor signaling pathways and immune related pathways were enriched in the immunoactive phenotype. In addition, immune infiltration analysis also suggested that there was higher effector immune cell infiltration in the immunoactive phenotype, and more tumor cells and inhibitory immune cells in the lipid metabolism subtype. These results suggest the hypermetabolism and proliferation

of tumors in the lipid metabolism subtype and explain the worse survival rate and tumor progression of patients with this subtype (37). More effector immune cells and stronger immune activity in the immunoactive phenotype play an anti-tumor role, resulting in better survival and tumor remission of patients (38).

Next, in order to elaborate the molecular characteristics of the two subtypes, the genomic alterations of the two subtypes were compared. In general, there was a higher TMB in the lipid metabolism subtype. It is worth noting that the mutation frequency of KEAP1 gene in the lipid subtype was significantly increased compared to that in the immunoactive phenotype. KEAP1 is an essential regulator of cell homeostasis and antioxidant stimulation (39). Studies have reported that this mutation is common in NSCLC with close correlation to higher tumor growth and invasiveness (40). Additionally, tumors bearing KEAP pathway mutations have been reported in preclinical and clinical studies, which have stronger resistance to traditional treatment methods, such as chemotherapy, radiotherapy, and targeted therapy (41–43). In addition, we found that the amplification and deletion levels of immunoactive subtypes were significantly higher at the chromosome arm level, and the deletion levels of lipid metabolism subtypes were higher in general. The contradictory results may suggest that CNV does not seem to be playing a pivotal role in regulating the differences between subtypes. In general, the genomic changes of these two subtypes are mainly mediated by gene mutations, especially KEAP1, which may contribute to the heterogeneous response of subtypes to tumor

treatment, leading to different clinical outcomes. In addition, KEAP1 may also be a new target for drug development and clinical treatment of LUAD.

Finally, a prognostic feature called LIS was developed and validated in the TCGA cohort, GEO meta cohort, and two external immunotherapy cohorts. High LIS is an independent negative prognostic factor for OS, and subgroup analysis showed that LIS showed stronger performance in predicting early LUAD patients. Considering the heterogeneity of subtypes in immunotherapy, we also evaluated the effectiveness of LIS in predicting immunotherapy. The results showed that LIS also showed high accuracy in the immunotherapy cohort. In addition, LIS also showed better accuracy than commonly used biomarkers (MDSC, MSI score, IFNG, CD8 and CD274). Finally, we found that patients with low LIS may have more neoantigens, which may lead to stronger immunotherapy sensitivity in patients with low LIS. In conclusion, our results suggest that LIS is not only a robust prognostic marker, but also a promising predictive marker of immunotherapy.

We admit that our research also has some defects. First, we only use Bulk-seq data without considering the heterogeneity between cells. Secondly, the sequenced samples came from tumor tissue, which may lead to the fact that LIS is not suitable for peripheral blood samples, and the clinical application is limited. Finally, although we used algorithms and mature immunotherapy cohorts to evaluate the sensitivity of the two subtypes to immunotherapy, prospective clinical research cohorts are still needed for validation. In conclusion, our work identified and validated heterogeneous lipid metabolism subtypes and immune activity subtypes in LUAD, which showed heterogeneity in clinical outcomes, biological functions, immune infiltration, and genome driven events. In addition, we have developed a feature called LIS, which can be used as a reliable prognostic biomarker for predicting OS and immunotherapy response. These results promote the understanding of the interaction between lipid metabolism and TME and offer a new direction for clinical management and precision therapy of LUAD patients.

## Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

## Author contributions

XG performed the data analyses and wrote the manuscript. SW contributed significantly to analysis and manuscript preparation. ZL helped perform the analysis with constructive

discussions. HX contributed to the conception of the study. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the Programs of Shanghai Pulmonary Hospital (No. fkxr1902 and No.fkxr1904).

## Acknowledgments

We greatly appreciate the analytical data provided by the TCGA and GEO databases.

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.1022149/full#supplementary-material>

### SUPPLEMENTARY FIGURE 1

The workflow of this study. In this study, we aimed to survey the crosstalk between lipid metabolism and tumor immune response in LUAD patients and identified two heterogeneous subtypes (lipid metabolism subtype and immune activity subtype). These two subtypes show specific differences in clinical outcomes, biological functions, immune infiltration and genomic variation. In addition, a lipid-immune score (LIS) was developed and validated, which shows significant advantages in predicting prognosis and immunotherapy response. In conclusion, our work strengthens the understanding of the complex role between lipid metabolism and immune system in LUAD and provides a new perspective and reference for the accurate prediction and immunotherapy of LUAD patients.

## SUPPLEMENTARY FIGURE 2

Validation of immune and lipid subtypes. (A) CDF curve of consensus matrix of different K in GEO queue. (B) Consensus matrix when k = 2 in GEO queue. (C) Expression heat maps of LMGs and IRGs in two subtypes in the GEO cohort. (D) Survival curve of two subtypes in GEO cohort.

## SUPPLEMENTARY FIGURE 3

Verification of immune infiltration in subtypes. (A) Differences in Estimate scores between the two subtypes in the GEO cohort. (B) Differences in the expression of six typical immune checkpoints (PD-L1, CD8A, CTLA-4, LAG-3, PD-1) between the two subtypes in the GEO cohort. (C) Differences in immune related pathway activity between the two subtypes in the GEO cohort. (D) Differences in immune cell infiltration between the two subtypes in the GEO cohort.

## SUPPLEMENTARY FIGURE 4

External verification of LIS. (A) Survival curves of high and low LIS subgroups in the GEO cohort. (B) ROC analysis of LIS in GEO queue. (C) The scatter plot shows the survival status of different LIS patients in the GEO cohort. (D) TROC curve of LIS in GEO queue. (E) Subgroup analysis of LIS in patients with different clinical characteristics in GEO cohort.

## SUPPLEMENTARY FIGURE 5

External verification of LIS. (A) The TIDE algorithm in GEO queue predicted the response rate of immunotherapy with high LIS and low LIS. (B) ROC curves in the GEO cohort showed the predictive accuracy of LIS and different immune markers.

## References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
- Cheng TY, Cramb SM, Baade PD, Youlden DR, Nwogu C, Reid ME. The international epidemiology of lung cancer: Latest trends, disparities, and tumor characteristics. *J Thorac Oncol: Off Publ Int Assoc Study Lung Cancer* (2016) 11(10):1653–71. doi: 10.1016/j.jtho.2016.05.021
- Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature* (2018) 553(7689):446–54. doi: 10.1038/nature25183
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA: Cancer J Clin* (2018) 68(1):7–30. doi: 10.3322/caac.21442
- Warburg O. On the origin of cancer cells. *Sci (New York NY)* (1956) 123(3191):309–14. doi: 10.1126/science.123.3191.309
- Yu W, Lei Q, Yang L, Qin G, Liu S, Wang D, et al. Contradictory roles of lipid metabolism in immune response within the tumor microenvironment. *J Hematol Oncol* (2021) 14(1):187. doi: 10.1186/s13045-021-01200-4
- Pavlova NN, Thompson CB. The emerging hallmarks of cancer metabolism. *Cell Metab* (2016) 23(1):27–47. doi: 10.1016/j.cmet.2015.12.006
- Wen YA, Xing X, Harris JW, Zaytseva YY, Mitov MI, Napier DL, et al. Adipocytes activate mitochondrial fatty acid oxidation and autophagy to promote tumor growth in colon cancer. *Cell Death Dis* (2017) 8(2):e2593. doi: 10.1038/cddis.2017.21
- Yao CH, Fowle-Grider R, Mahieu NG, Liu GY, Chen YJ, Wang R, et al. Exogenous fatty acids are the preferred source of membrane lipids in proliferating fibroblasts. *Cell Chem Biol* (2016) 23(4):483–93. doi: 10.1016/j.chembiol.2016.03.007
- Corn KC, Windham MA, Rafat M. Lipids in the tumor microenvironment: From cancer progression to treatment. *Prog Lipid Res* (2020) 80:101055. doi: 10.1016/j.plipres.2020.101055
- Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinf (Oxford England)* (2012) 28(6):882–3. doi: 10.1093/bioinformatics/bts034
- Jung H, Kim HS, Kim JY, Sun JM, Ahn JS, Ahn MJ, et al. DNA Methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat Commun* (2019) 10(1):4278. doi: 10.1038/s41467-019-12159-9
- Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGF $\beta$  attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* (2018) 554(7693):544–8. doi: 10.1038/nature25501
- Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinf (Oxford England)* (2011) 27(12):1739–40. doi: 10.1093/bioinformatics/btr260
- Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, et al. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Sci Data* (2018) 5:180015. doi: 10.1038/sdata.2018.15
- Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic: J Integr Biol* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118
- McDermott DF, Huseni MA, Atkins MB, Motzer RJ, Rini BI, Escudier B, et al. Clinical activity and molecular correlates of response to atezolizumab alone or in combination with bevacizumab versus sunitinib in renal cell carcinoma. *Nat Med* (2018) 24(6):749–57. doi: 10.1038/s41591-018-0053-3
- Gibbons DL, Creighton CJ. Pan-cancer survey of epithelial-mesenchymal transition markers across the cancer genome atlas. *Dev dyn: an Off Publ Am Assoc Anat* (2018) 247(3):555–64. doi: 10.1002/dvdy.24485
- 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
- Yu T, Tan H, Liu C, Nie W, Wang Y, Zhou K, et al. Integratively genomic analysis reveals the prognostic and immunological characteristics of pyroptosis and ferroptosis in pancreatic cancer for precision immunotherapy. *Front Cell Dev Biol* (2022) 10:826879. doi: 10.3389/fcell.2022.826879
- Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* (2015) 12(5):453–7. doi: 10.1038/nmeth.3337
- Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-García W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* (2013) 4:2612. doi: 10.1038/ncomms3612
- Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. *Immunity* (2018) 48(4):812–30.e14. doi: 10.1016/j.immuni.2018.03.023
- 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. doi: 10.1101/gr.239244.118
- Schröder MS, Culhane AC, Quackenbush J, Haibe-Kains B. Survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinf (Oxford England)* (2011) 27(22):3206–8. doi: 10.1093/bioinformatics/btr511
- 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. doi: 10.1016/j.celrep.2016.12.019
- Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med* (2018) 24(10):1550–8. doi: 10.1038/s41591-018-0136-1
- Fu J, Li K, Zhang W, Wan C, Zhang J, Jiang P, et al. Large-Scale public data reuse to model immunotherapy response and resistance. *Genome Med* (2020) 12(1):21. doi: 10.1186/s13073-020-0721-z
- Wang C, Gu X, Zhang X, Zhou M, Chen Y. Development and validation of an E2F-related gene signature to predict prognosis of patients with lung squamous cell carcinoma. *Front Oncol* (2021) 11:756096. doi: 10.3389/fonc.2021.756096
- Wang Y, Tan H, Yu T, Chen X, Jing F, Shi H. Potential immune biomarker candidates and immune subtypes of lung adenocarcinoma for developing mRNA vaccines. *Front Immunol* (2021) 12:755401. doi: 10.3389/fimmu.2021.755401
- Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inf decis mak* (2008) 8:53. doi: 10.1186/1472-6947-8-53
- Desrichard A, Snyder A, Chan TA. Cancer neoantigens and applications for immunotherapy. *Clin Cancer res: an Off J Am Assoc Cancer Res* (2016) 22(4):807–12. doi: 10.1158/1078-0432.Ccr-14-3175

33. Hu Z, Ott PA, Wu CJ. Towards personalized, tumour-specific, therapeutic vaccines for cancer. *Nat Rev Immunol* (2018) 18(3):168–82. doi: 10.1038/nri.2017.131
34. Abida W, Cheng ML, Armenia J, Middha S, Autio KA, Vargas HA, et al. Analysis of the prevalence of microsatellite instability in prostate cancer and response to immune checkpoint blockade. *JAMA Oncol* (2019) 5(4):471–8. doi: 10.1001/jamaoncol.2018.5801
35. Yin N, Liu Y, Khor A, Wang X, Thompson EA, Leitges M, et al. Protein kinase  $\text{ct}$  and  $\text{wnt}/\beta$ -catenin signaling: Alternative pathways to Kras/Trp53-driven lung adenocarcinoma. *Cancer Cell* (2019) 36(2):156–67.e7. doi: 10.1016/j.ccell.2019.07.002
36. Justilien V, Ali SA, Jamieson L, Yin N, Cox AD, Der CJ, et al. Ect2-dependent rRNA synthesis is required for KRAS-TRP53-Driven lung adenocarcinoma. *Cancer Cell* (2017) 31(2):256–69. doi: 10.1016/j.ccell.2016.12.010
37. Vander Heiden MG, DeBerardinis RJ. Understanding the intersections between metabolism and cancer biology. *Cell* (2017) 168(4):657–69. doi: 10.1016/j.cell.2016.12.039
38. Madden MZ, Rathmell JC. The complex integration of T-cell metabolism and immunotherapy. *Cancer Discovery* (2021) 11(7):1636–43. doi: 10.1158/2159-8290.Cd-20-0569
39. Yamamoto M, Kensler TW, Motohashi H. The KEAP1-NRF2 system: a thiol-based sensor-effector apparatus for maintaining redox homeostasis. *Physiol Rev* (2018) 98(3):1169–203. doi: 10.1152/physrev.00023.2017
40. Hellyer JA, Padda SK, Diehn M, Wakelee HA. Clinical implications of KEAP1-NFE2L2 mutations in NSCLC. *J Thorac Oncol: Off Publ Int Assoc Study Lung Cancer* (2021) 16(3):395–403. doi: 10.1016/j.jtho.2020.11.015
41. Jeong Y, Hellyer JA, Stehr H, Hoang NT, Niu X, Das M, et al. Role of KEAP1/NFE2L2 mutations in the chemotherapeutic response of patients with non-small cell lung cancer. *Clin Cancer Res: an Off J Am Assoc Cancer Res* (2020) 26(1):274–81. doi: 10.1158/1078-0432.Ccr-19-1237
42. Binkley MS, Jeon YJ, Nesselbush M, Moding EJ, Nabet BY, Almanza D, et al. KEAP1/NFE2L2 mutations predict lung cancer radiation resistance that can be targeted by glutaminase inhibition. *Cancer Discovery* (2020) 10(12):1826–41. doi: 10.1158/2159-8290.Cd-20-0282
43. Hellyer JA, Stehr H, Das M, Padda SK, Ramchandran K, Neal JW, et al. Impact of KEAP1/NFE2L2/CUL3 mutations on duration of response to EGFR tyrosine kinase inhibitors in EGFR mutated non-small cell lung cancer. *Lung Cancer (Amsterdam Netherlands)* (2019) 134:42–5. doi: 10.1016/j.lungcan.2019.05.002