



OPEN ACCESS

EDITED BY
Zhi-Qian Zhang,
Southern University of Science and
Technology, China

REVIEWED BY
Ranran Sun,
Zhengzhou University, China
Wei Li,
Shenzhen Longhua District Central
Hospital, China

*CORRESPONDENCE
Changfa Qu,
✉ 2106@hrbmu.edu.cn

SPECIALTY SECTION
This article was submitted to
Pharmacology of Anti-Cancer Drugs,
a section of the journal
Frontiers in Pharmacology

RECEIVED 27 October 2022
ACCEPTED 07 December 2022
PUBLISHED 19 December 2022

CITATION
Liu M, Xiao Q, Yu X, Zhao Y and Qu C
(2022), Characterization of lung
adenocarcinoma based on
immunophenotyping and constructing
an immune scoring model to
predict prognosis.
Front. Pharmacol. 13:1081244.
doi: 10.3389/fphar.2022.1081244

COPYRIGHT
© 2022 Liu, Xiao, Yu, Zhao and Qu. 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.

Characterization of lung adenocarcinoma based on immunophenotyping and constructing an immune scoring model to predict prognosis

Mengfeng Liu¹, Qifan Xiao¹, Xiran Yu¹, Yujie Zhao² and Changfa Qu^{1*}

¹Department of Thoracic Surgery, Harbin Medical University Cancer Hospital, Harbin Medical Sciences University, Harbin, China, ²Regional Marketing Department, YuceBio Technology Co., Shenzhen, China

Background: Lung cancer poses great threat to human health, and lung adenocarcinoma (LUAD) is the main subtype. Immunotherapy has become first line therapy for LUAD. However, the pathogenic mechanism of LUAD is still unclear.

Methods: We scored immune-related pathways in LUAD patients using single sample gene set enrichment analysis (ssGSEA) algorithm, and further identified distinct immune-related subtypes through consistent clustering analysis. Next, immune signatures, Kaplan-Meier survival analysis, copy number variation (CNV) analysis, gene methylation analysis, mutational analysis were used to reveal differences between subtypes. pRRophetic method was used to predict the response to chemotherapeutic drugs (half maximal inhibitory concentration). Then, weighted gene co-expression network analysis (WGCNA) was performed to screen hub genes. Significantly, we built an immune score (IMscore) model to predict prognosis of LUAD.

Results: Consensus clustering analysis identified three LUAD subtypes, namely immune-Enrich subtype (Immune-E), stromal-Enrich subtype (Stromal-E) and immune-Deprived subtype (Immune-D). Stromal-E subtype had a better prognosis, as shown by Kaplan-Meier survival analysis. Higher tumor purity and lower immune cell scores were found in the Immune-D subtype. CNV analysis showed that homologous recombination deficiency was lower in Stromal-E and higher in Immune-D. Likewise, mutational analysis found that the Stromal-E subtype had a lower mutation frequency in TP53 mutations. Difference in gene methylation (ZEB2, TWIST1, CDH2, CDH1 and CLDN1) among three subtypes was also observed. Moreover, Immune-E was more sensitive to traditional chemotherapy drugs Cisplatin, Sunitinib, Crizotinib, Dasatinib, Bortezomib, and Midostaurin in both the TCGA and GSE cohorts. Furthermore, a 6-gene signature was constructed to predicting prognosis, which performed better than TIDE score. The performance of IMscore model was successfully validated in three independent datasets and pan-cancer.

KEYWORDS

LUAD, immune, immunophenotyping, immune scoring model, prognosis

1 Introduction

Lung cancer still serves as the most common malignancy and the foremost cause of cancer death in the world (Thandra et al., 2021). Lung adenocarcinoma (LUAD) as a main subtype of lung cancer belong to the larger group of non-small cell lung cancer, and accounts for about 40% of all lung cancer cases. Moreover, it is strongly associated with smoking (Pintarelli et al., 2019; Spella and Stathopoulos, 2021). Although treatment modalities such as surgery, chemoradiotherapy, targeted therapy, and immunotherapy have been widely used for the treatment of lung cancer, but the prognosis remains dismal (Ladanyi and Pao, 2008; Bronte et al., 2010; Denisenko et al., 2018), with a 5-year survival rate as low as less than 10% (Hirsch et al., 2017).

Cancer evolution is influenced by complex interactions between tumor cells and host immune responses within the tumor microenvironment (Taube et al., 2018). Different immune cell populations are actively involved in tumor immune microenvironment (TIME), however, their relationship is currently unclear in LUAD therapy (Zhang et al., 2020). Immune cells are both positive and negative regulators of cancer progression (Zamarron and Chen, 2011). For example, B-cells exert antitumor functions by enhancing T-cell immunity, stimulating the production of interferon- γ and helping natural killer (NK) cells against tumors (Sorrentino et al., 2011). However, B-cells also suppress immune responses and promote angiogenesis (Schwartz et al., 2016). Considering the importance of TIME in cancers, we attempted to classify LUAD based on immune pathways and compared their characteristics.

Disease prediction models have been widely used to evaluate patient survival and other prognostic indicators in LUAD (Jiang et al., 2020; Jiang et al., 2022). Recently, a novel model related to lactate metabolism for predicting overall survival and immune signature in LUAD was reported and different immune signatures were built (Jiang et al., 2022). Moreover, an increasing number of LUAD prognostic biomarkers have been discovered by analyzing expression profiles from public databases and related clinical information (Jiang et al., 2020).

Due to advances in genetic technology, various molecular subtypes and gene expression studies have been published and popularized for the identification of prognostic biomarkers (Li and Wang, 2021). However, the prognostic manifestations of established biomarkers are controversial and limited. In the current study, we identified and classified three immune subtypes and a six-gene combination capable of predicting survival in patients with LUAD.

2 Materials and methods

2.1 Data collection

To better understand the pathogenesis of LUAD, RNA-Seq data of LUAD with clinical survival and characteristic information were downloaded from TCGA database by using TCGA GDC API. Fragments per kilobase of transcript per million fragments mapped (FPKM) were converted into transcripts per million (TPM), and then Log₂ was transformed for subsequent analysis. In addition, we downloaded the GSE37745, GSE50081, GSE30219, GSE31210 microarray datasets with survival times from the GEO database (Yamauchi et al., 2012; Botling et al., 2013; Rousseaux et al., 2013; Der et al., 2014). For GEO, MINiML formatted family file(s) were downloaded and samples was preprocessed using RMA implemented in affy package. Normalized data values were transformed in log₂ space and used in subsequent analysis.

2.2 Data quality control

To ensure the accuracy of downstream analysis, the RNA-Seq data of TCGA-LUAD without clinical follow-up information or survival time were removed. Next, the data without status were filtered. Then, only the genes with more than one expression in more than 50% of the samples were retained.

For GSE data, firstly, normal tissues, samples without clinical follow-up information, samples without OS data and samples without status were filtered. Then, we converted the probes to symbols according to the annotation file.

2.3 Batch effect processing

The *removeBatchEffect* function of the limma (R package) was used to remove batch effects between different datasets (Ritchie et al., 2015). Principal component analysis (PCA) was used to observe the batch effect (Supplementary Figure S1).

2.4 LUAD classification based on pathway score

In order to explore the molecular typing of LUAD, single sample gene set enrichment analysis (ssGSEA) analysis was used to calculate the score related to immune pathway (Supplementary Table S1) (Gao et al., 2021). Next, ConsensusclusterPlus (R Bioconductor/R package) software

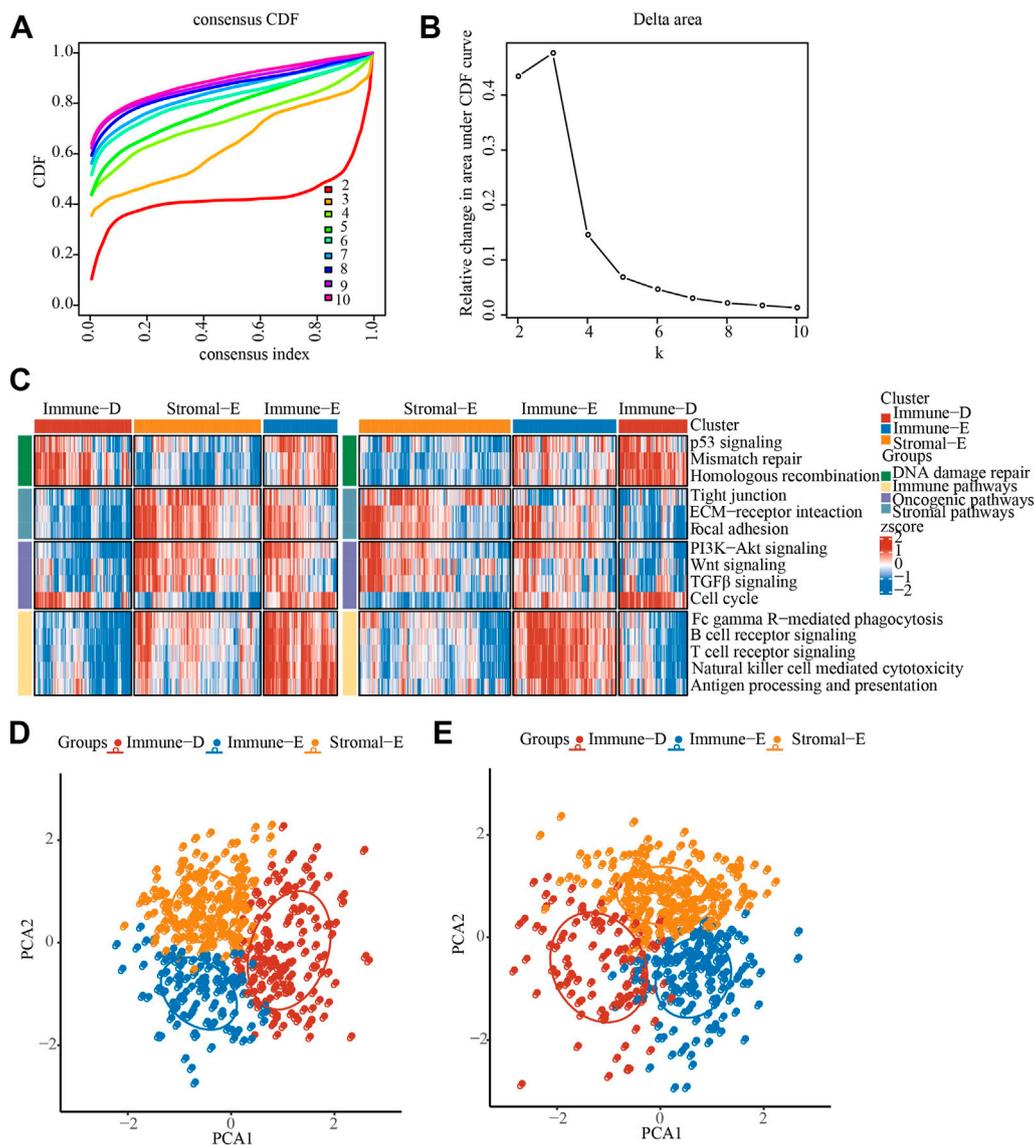


FIGURE 1

Identification of LUAD immunophenotyping. (A) The cumulative distribution function (CDF) of different consensus index. (B) The CDF Delta area curve. (C) The heat map showing the immune pathway score in TCGA and GSE cohorts. (D) PCA analysis in TCGA cohorts, different color means different subtypes. (E) PCA analysis in GSE cohorts, different color means different subtypes.

was used for consensus clustering analysis with “pam” arithmetic and “pearson” distance, and the input was a sample matrix pool of immune pathway scores (Wilkerson and Hayes, 2010). Then, we determined the optimal number of clusters according to the cumulative distribution function (CDF).

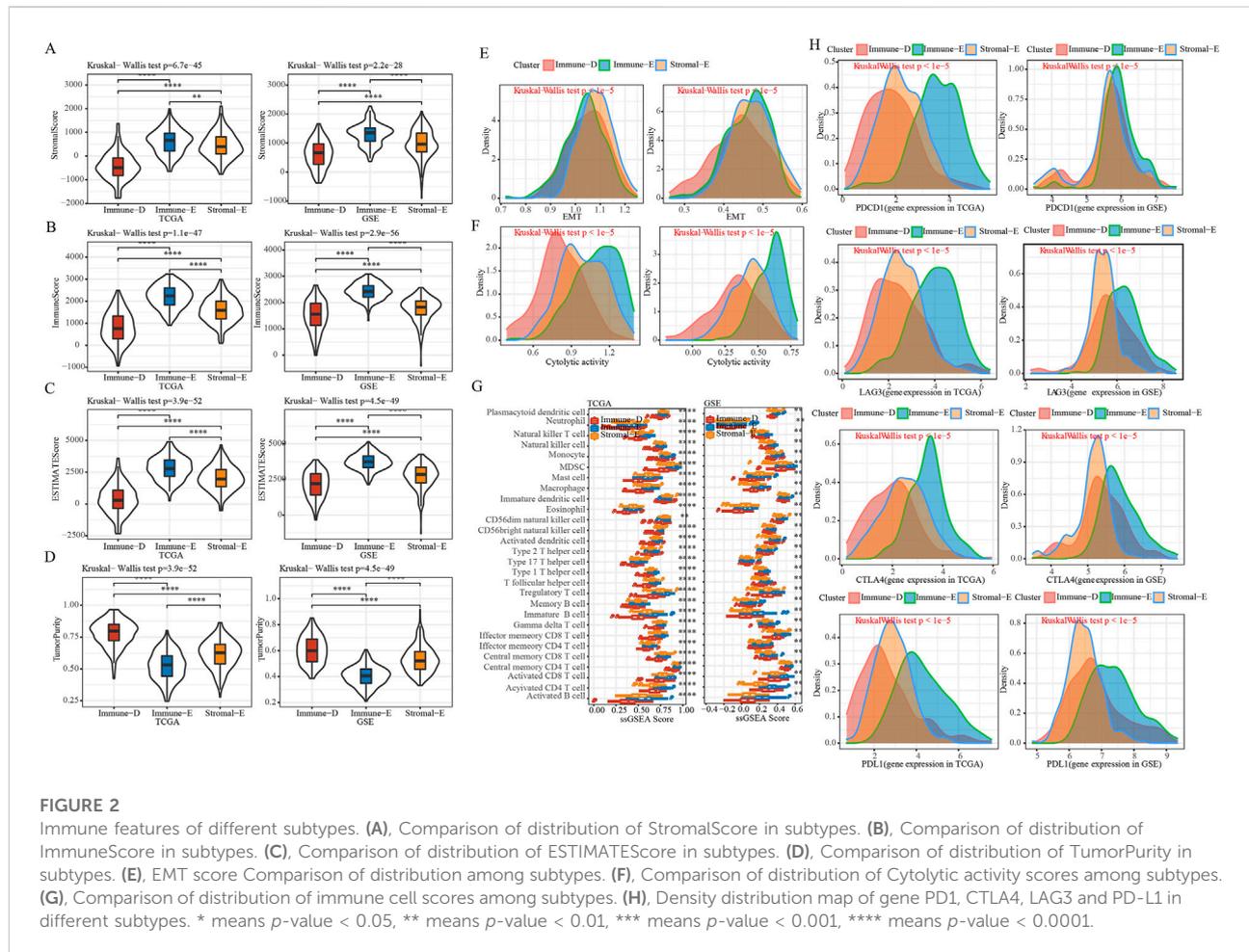
2.5 Gene mutation analysis

The *mutect2* software was applied to perform gene mutation analysis (Benjamin et al., 2019). The tumor promoting genes

were obtained from previous study. The fisher’s test was used to screen genes with significantly high frequency mutations in each subtype, with a threshold of p -value < 0.05. Moreover, the *maftools* software (R package) was used to calculate the tumor mutational burden (TMB) score (Mayakonda et al., 2018).

2.6 Copy number variation (CNV) analysis

GISTIC2.0 was used to analyze the change of CNV (Mermel et al., 2011). If the ratio was greater than 0.2, it was considered as



Gain, if the ratio was less than -0.2 , it was considered as Loss, and the rest was considered as Diploid.

2.7 Methylation analysis

The 450K methylation data of LUAD were used to perform methylation analysis of EMT-promoting genes (Wang and Zhou, 2013), and missing values were imputed using the KNN algorithm of the *impute* software (R package) (Hastie et al., 2011).

2.8 Treatment plan sensitivity analysis

We used the tumor immune dysfunction and exclusion (TIDE) (<https://tide.dfci.harvard.edu/>) algorithm to predict response of immunotherapy (Jiang et al., 2018). The higher the TIDE prediction score, the higher the likelihood of immune escape, and the lower the likelihood that the patient would benefit from immunotherapy. Moreover, IC50 (half maximal inhibitory concentration) analysis was used to

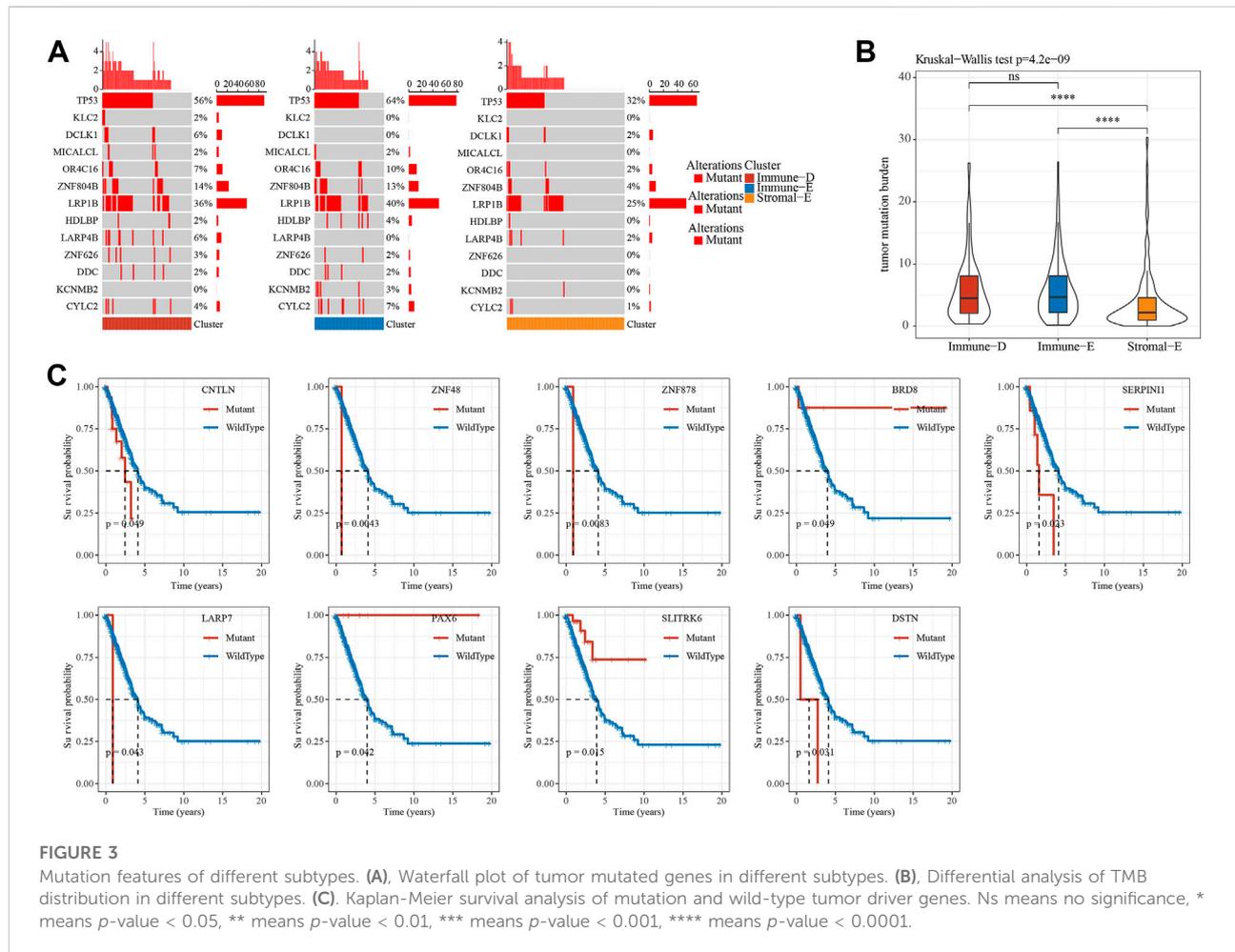
determine the sensitivity of different subtypes to chemotherapy drugs using pRRophetic method.

2.9 Weighted gene co-expression network analysis (WGCNA)

WGCNA was performed by WGCNA (R package) to filter hub genes in the module related to different LUAD subtypes (Langfelder and Horvath, 2008). Hub genes refer to genes that play key roles in modules and are generally closely related to shape. The hub genes were further subjected to KEGG pathway analysis by using clusterProfiler (R Bioconductor/R package). p -value < 0.05 was considered significance.

2.10 IMscore predicted by prognostic model

IMScore model was constructed using univariate Cox regression analysis and LASSO analysis. Finally, six prognostic



genes were obtained, including MARCKS, CDK2, SFN, SSBP1, MRE11 and FZD7. $IMS_{score} = \sum \text{Exp}(i) \cdot \beta(i)$, where i refers to the immune-related prognostic genes, Exp refers to the expression levels of genes, and β refers to the LASSO coefficients.

3 Results

3.1 LUAD classification based on pathway scores

The pathway scores of different samples in the TCGA and GSE cohorts were calculated by ssGSEA algorithm, and then consensus clustering analysis was performed on the pathway scores. The optimal number of clusters was determined according to the CDF, and the CDF Delta area curve was observed. When it was selected as 3, it has a relatively stable clustering result, and finally we choose $k = 3$ to obtain three related subtypes, including immune-Enrich subtype (Immune-E), stromal-Enrich subtype (Stromal-E) and immune-Deprived

subtype (Immune-D) (Figures 1A,B). A heatmap showing the trends in pathway scores for each sample demonstrated a good discrimination between the three different subtypes (Figure 1C). In addition, PCA analysis between different subtypes showed that in both datasets, there were distinct boundaries between different subtypes (Figures 1D,E).

3.2 Analysis of clinical characteristics of different subtypes

Survival analysis showed a better prognosis in the Stromal-E subtype and a worse prognosis for the Immune-D subtype, which was consistent in the TCGA cohort (left) and the GSE cohort (right) (Supplementary Figure S2A). Analysis of differences in the distribution of clinical features among subtypes in the TCGA dataset showed significant differences in age, sex as well as the distribution of T Stage among subtypes (Supplementary Figure S2B). Comparative analysis of the molecular subtypes with the six previously identified immunophenotypes (Thorsson et al., 2018)

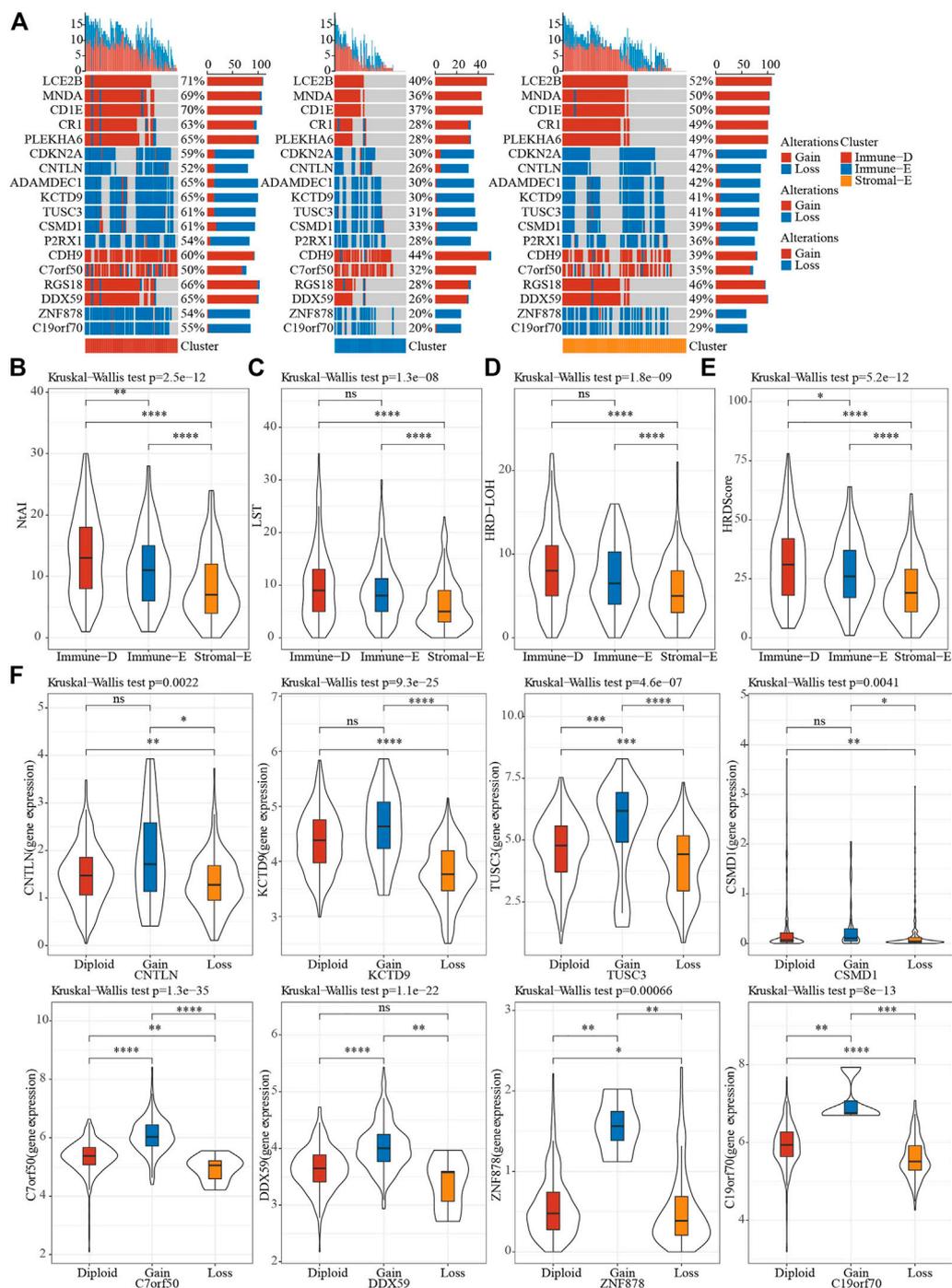
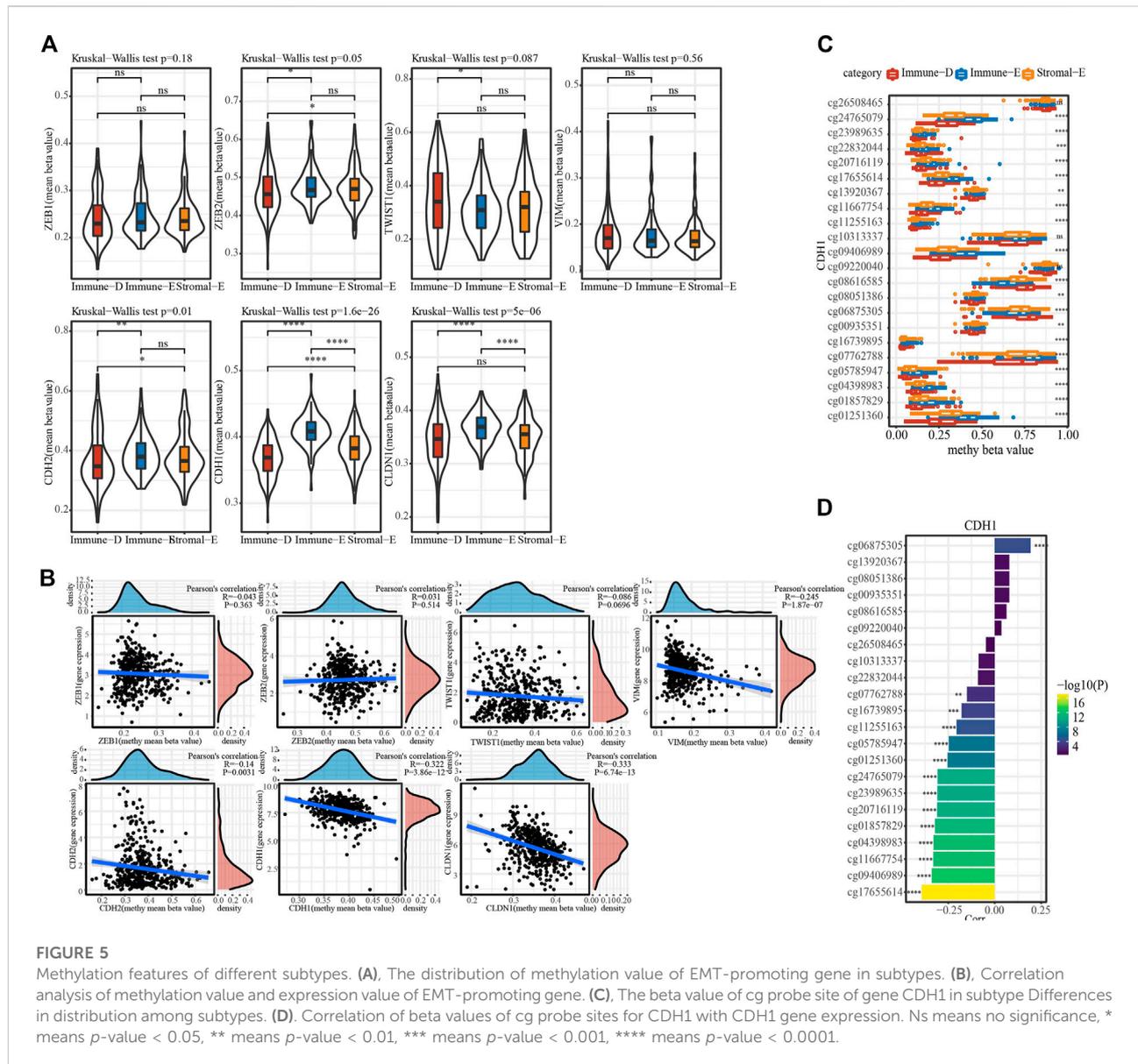


FIGURE 4

CNV features of different subtypes. **(A)**, CNV distribution of tumor driver genes. **(B–E)**, Differences in NtAI, LST, LOH and HRD score of different subtypes, respectively. **(F)**, Differences in gene expression by tumor driver gene CNV groupings. Ns means no significance, * means p -value < 0.05, ** means p -value < 0.01, *** means p -value < 0.001, **** means p -value < 0.0001.

showed that the previously published immunophenotypes were significantly different from our current study, for example, the Immune-E subtype had the highest proportion of the C2 subtype

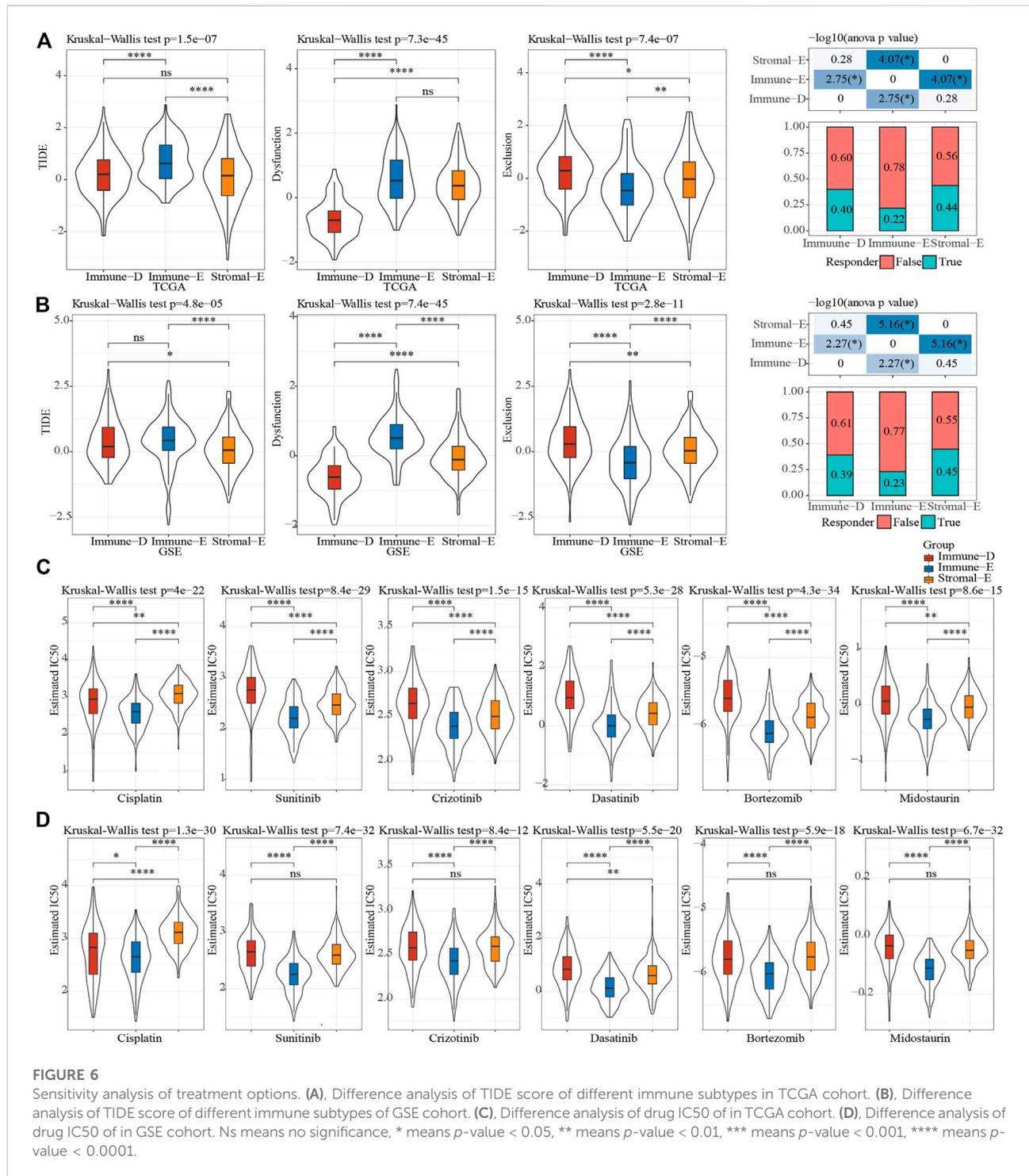
and C3 subtype had the best prognosis and also contributed the largest proportion in Stromal-E subtype. These results supported the reliability of our immunosubtyping (Supplementary Figure S2C)



3.3 Immune characteristics among different LUAD subtypes

In order to further understand the immune characteristics among the various subtypes of LUAD, we first used the *Estimate* software to evaluate the immune scores and tumor purity score of the TCGA and GSE cohorts. The results demonstrated that the immune score, stromal score, and ESTIMATE score were the highest in Immune-E subtype both in the TCGA and GSE cohorts (Figures 2A–C). The opposite was true for the tumor purity score (Figure 2D). Epithelial-mesenchymal Transition (EMT) is closely related to tumor metastasis and recurrence (Wang and Zhou, 2013).

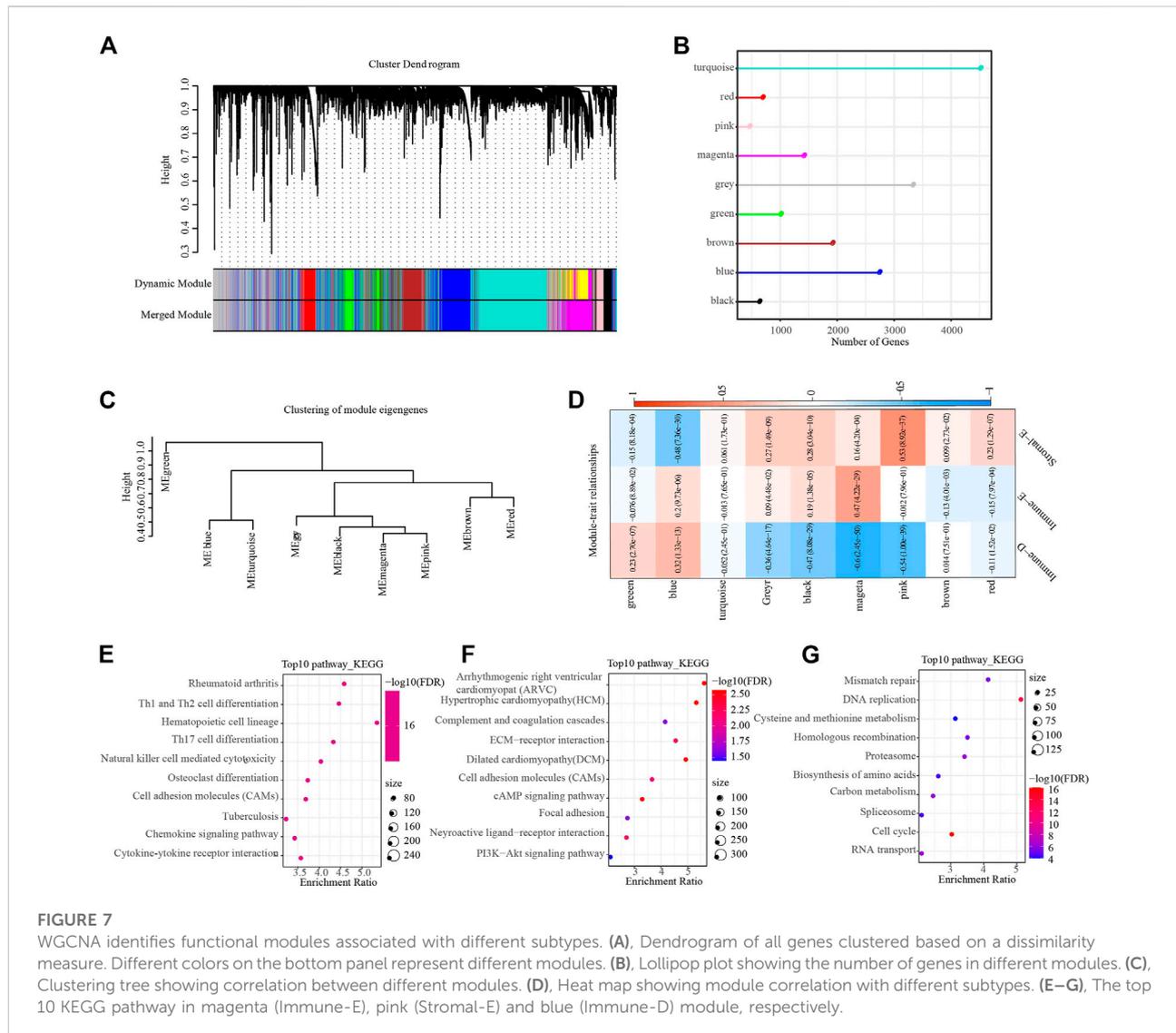
Therefore, we applied the ssGSEA algorithm to evaluate the difference in EMT scores between different subtypes. The results showed that in the TCGA cohort, the Stromal-E subtype had the highest EMT score, and in the GSE cohort Immune-E subtype was the highest (Figure 2E). Cytolytic activity is associated with immunotherapy, and here ssGSEA was used to assess differences in Cytolytic activity score (Rooney et al., 2015). We found that both in TCGA and GSE cohorts, the score was highest in Immune-D subtype and the lowest in Immune-E subtype (Figure 2F). To further understand the status of different immune cells, we assessed immune cell infiltration between different subtypes and found the highest scores in Immune-E, such as



Plasmacytoid dendritic cell, Regulatory T cell and Gamma delta T cell. This was also the same in the TCGA and GSE cohorts (Figure 2G). Moreover, we determine the expression level of PD1, PD-L1, CTLA4 and LAG3, and observed that the expression level was higher in Immune-E subtype (Figure 2H).

3.4 Mutational analysis of tumor driver genes in different subtypes

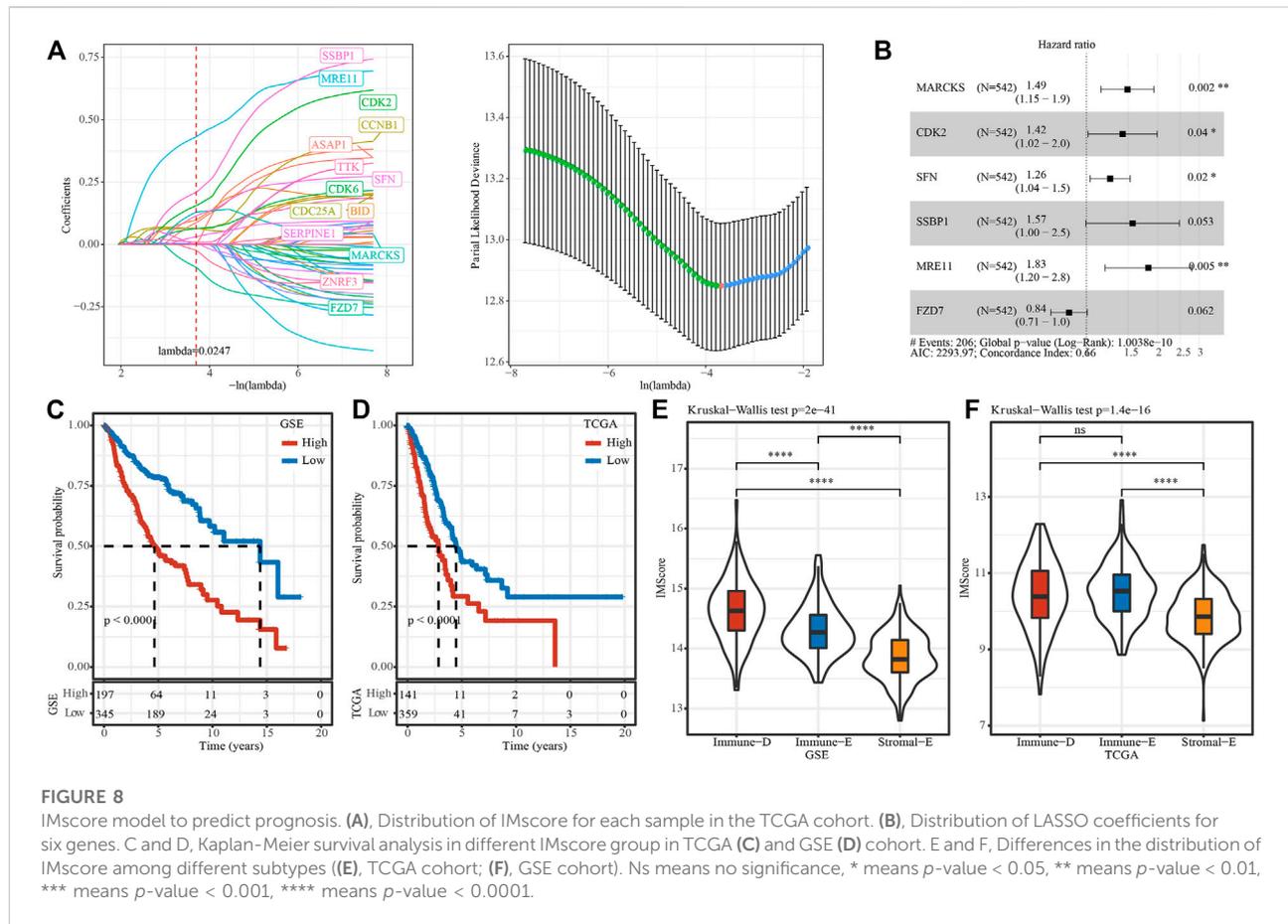
A total of 172 tumor driver genes were obtained from previous study (Gao et al., 2013). Mutation analysis of tumor



driver genes in different subtypes of TCGA dataset found that 13 genes had different mutations in different groups and TP53 had the highest mutation frequency in the Immune-E subtype (Figure 3A). The results of TMB analysis in different subtypes found that TMB in Stroma-E subtype was significantly lower than that in Immune-E and Immune-D subtypes. There was no difference in TMB distribution between Immune-E and Immune-D subtypes in TCGA dataset (Figure 3B). In addition, survival curve (KM) analysis of mutations in driver genes and wild-type samples in TCGA dataset demonstrated that mutations in nine genes were significantly different from wild-type, including CNTLN, ZNF48, ZNF878, BRDS, SERPIN11, LARP7, SLITRK6 and DSTN (Figure 3C). Those analysis indicated that three subtypes may predict mutation status.

3.5 CNV analysis of tumor driver genes in different subtypes

A total of 159 of 172 tumor driver genes had CNV data. To understand CNV in tumor driver genes, GISTIC2 was used to perform CNV analysis. The results showed that the amplification and deletion of 159 driver genes were significant in different subtypes in TCGA dataset, and top 18 genes were visualized. Notably, immune-D had the largest amounts of CNVs (Figure 4A). Homologous recombination deficiency (HRD) is associated with a poorer cancer prognosis (Knijnenburg et al., 2018). Loss of heterozygosity (LOH), LST (large-scale state transitions), NtAI (number of telomeric allelic imbalances) score and HRD score were selected to assess the HRD status in different subtypes, and we found that the above scores were the



lowest in Stromal-E and highest in Immune-D (Figures 4B–E). The expression level of tumor driver genes in the gene CNV grouping showed that the expression of the genes corresponding to the grouping with Gain was higher, while that corresponding to the grouping with Loss was lower (Figure 4F).

3.6 Methylation analysis of genes in different subtypes

The methylation status of genes plays an important role in gene expression. DNA methylation analysis of seven EMT-promoting genes showed that the methylation status of ZEB1 and VIM did not differ among subtypes in TCGA dataset, while the methylation status of ZEB2, TWIST1, CDH2, CDH1 and CLDN1 were significantly different in each subtype (Figure 5A). The correlation analysis between gene expression value and methylation status showed that the gene expression values of VIM, CDH2, CDH1, and CLDN1 were significantly negatively correlated with methylation status (Figure 5B). The distribution of cg sites of CDH1 in different subtypes in TCGA dataset showed that some sites had significant

differences in different subtype groups, and that the highest value in Immune-E was in a hypermethylated state (Figure 5C). From the correlation analysis of CDH1 cg loci and CDH1 gene expression, most of the cg loci and CDH1 gene expression showed a significant negative correlation (Figure 5D).

3.7 Immunotherapy and drug sensitivity of different subtypes

The TIDE (<https://tide.dfci.harvard.edu/>) algorithm was used to assess the potential clinical effects of immunotherapy on different molecular subtypes. In the TCGA cohort, the TIDE score of Immune-E was significantly higher than that of Stromal-E and Immune-D. The immunotherapy effect analysis indicated that the proportion of responses in Immune-E was only 22%, which was much lower than that of Stromal-E and Immune-D (Figure 6A). In GSE dataset, TIDE score was also higher in Immune-E. (Figure 6B). In addition, we also analyzed the response of different subtypes to the traditional chemotherapy drugs Cisplatin, Sunitinib, Crizotinib, Dasatinib, Bortezomib, Midostaurin. The results suggested that all the six drugs were

more sensitive to the Immune-E subtype in both the TCGA and GSE cohorts (Figures 6C, D).

3.8 WGCNA to identify the key gene of subtypes

Hierarchical cluster analysis of cohort samples showed no discrete samples in the TCGA cohorts (Supplementary Figure S3A). In order to ensure that the gene network we constructed conformed to the scale-free distribution, we set the β value to 9 at the time $R^2 > 0.85$ (Supplementary Figures S3A,S3C). After clustering the modules and merging the closer modules into a new module, we acquired a total of nine modules (Figure 7A). The lollipop graph showed the number of genes in each module. It can be observed that the turquoise color module had the largest number of genes, which was more than 4,000 genes (Figure 7B). The heat map of correlation analysis of each subtype and module showed that the Immune-D subtype had the highest positive correlation with the blue module, the Immune-E subtype had the highest positive correlation with the magenta module, and the Stromal-E subtype had the highest positive correlation with the pink module (Figures 7C,D). Further, we performed KEGG pathway analysis on the genes in the above three modules, and the results demonstrated that the genes in the magenta module were closely related to Natural killer cell mediated cytotoxicity, the genes in the pink module were closely related to ECM-receptor interaction, and the genes in the blue module were closely related to Genes were closely related to Mismatch repair, etc. (Figures 7E–G).

3.9 IMscore prognostic model construction

In GSE, six prognostic pathways (Fc gamma R-mediated phagocytosis mediated phagocytosis, p53 signaling, Mismatch repair, Homologous recombination, Wnt signaling, Cell cycle) from 15 pathways were identified by univariate COX analysis. Then, Pearson analysis on the genes in six pathways and six pathways score was used to select the top20 genes. Through univariate COX analysis and LASSO regression analysis, we obtained six genes from 20 genes as related genes affecting prognosis (Figures 8A, B), and the model score was developed based on the following formula: $IMScore = 0.400 \cdot MARCKS + 0.353 \cdot CDK2 + 0.232 \cdot SFN + 0.450 \cdot SSBP1 + 0.604 \cdot MRE11 - 0.169 \cdot FZD7$.

The survminer package was used to find the best cutoff of IMscore and divide the GSE and TCGA data set samples into high IMscore group (high group) and low IMscore group (low group). In the GSE cohort, survival analysis of high- and low-groups found that the prognosis of high groups was significantly worse than that of low groups. High group also had worse

survival outcome in TCGA dataset (Figures 8C, D). In addition, we compared the distribution differences of IMscore among subtypes in different datasets and found that: IMscore in Stromal-E had the lowest in GSE and TCGA, while Immune-D had the highest IMscore in GSE, Immune-E had highest IMscore in TCGA dataset (Figures 8E, F).

3.10 Prediction efficiency of IMscore prognostic model

The survival curve analysis of IMscore model in different cancer types showed that except ESCA, our IMscore had significant differences in high and low IMscore in all cancer types, and that the prognosis of high IMscore was significantly worse than that of low IMscore (Supplementary Figure S4). Furthermore, IMvigor210, GSE91061 and GSE135222 data were used to examine the efficiency of the IMscore model in immunotherapy. To evaluate the efficiency of IMscore, TIDE analysis was used as a control. For IMvigor210 cohort, survival analysis showed significant differences in survival among different groups (IMscore: p -value < 0.0001 ; TIDE, p -value = 0.012) (Supplementary Figures S5A, S5B). For GSE91061 cohort, survival analysis showed significant differences in survival among different IMscore groups (p -value < 0.023) (Supplementary Figures S5C). TIDE group was not statistically significant (p -value = 0.067) (Supplementary Figures S5D). For GSE135222 cohort, survival analysis showed no significant differences in IMscore or TIDE groups (IMscore: p -value = 0.055; TIDE, p -value = 0.051) (Supplementary Figures S5E, S5F). Notably, ROC curve analysis demonstrated that in the above three cohorts, the predictive efficiency of IMscore was higher than that of TIDE.

4 Conclusion

LUAD is one of the most common malignant tumors, with high metastasis rate and strong invasiveness. Its low 5-year survival rate seriously threatens the life and health of human beings (Charloux et al., 1997; Tan et al., 2016). Significantly, early-stage LUAD has been reported to be associated with a higher risk of postoperative recurrence and death (Bittner et al., 2014). The immune system has been shown to play a critical role and even determine different stages of cancer development and progression (Shurin, 2018). Hence, an accurate classification of LUAD patients according to immune characteristics and the identification of LUAD biomarkers have positive significance for the selection of LUAD treatment methods. A previous study reported six subtypes in LUAD, including Wound Healing, IFN- γ Dominant, Inflammatory, Lymphocyte Depleted, Immunologically Quiet, and TGF- β Dominant (Thorsson et al., 2018). However, in this research, three molecular

subtypes classified by immunological features were obtained and defined as Immune-E, Stromal-E and Immune-D.

Kaplan-Meier survival analysis was able to determine survival differences between different subtypes (Goel et al., 2010). We found that Stromal-E subtype had a better prognosis and Immune-D had a worse prognosis. To gain a deeper understanding of the differences in survival between the different subtypes, we used *Estimate* software for immune calculating index scores and tumor purity scores (Yoshihara et al., 2013). The results showed that higher tumor purity and lower immune cell scores were in the Immune-D subtype, which accounted for the poorer prognosis of the subtype.

Recently, immunotherapy has also been developed as a new treatment for LUAD (Saito et al., 2018). Moreover, antibodies against PD-1 and PD-L1 have been reported to be effective in the treatment of various malignancies (Antonia et al., 2017; Bie et al., 2021). Although the biology of the TIME driving these responses was not fully understood, it is critical for the design of immunotherapeutic strategies. We found that the expression level of PD1, PD-L1, CTLA4 and LAG3 was higher in Immune-E subtype. However, as we mentioned above, the survival analysis showed that the Stromal-E subtype had a better prognosis, which suggested the complexity of tumorigenesis mechanisms.

TMB is an emerging tumor biomarker and it is associated with response to PD-1/PD-L1 targeted therapies in lung cancer (Spigel et al., 2016). We found that TMB in the Stromal-E subtype was significantly lower than in the Immune-E and Immune-D subtypes. TP53, known as the guardian of the genome, is one of the most well-known tumor suppressor genes (Surget et al., 2014). Interestingly, the mutation analysis suggested that Stromal-E subtype had a lower TP53 mutation rate of only 32%, while the other two subtypes both exceeded 50%, which explained a better prognosis of Stromal-E subtype. Moreover, high levels of TIDE scores suggested that Immune-E was more likely to occur immune escape, suggesting that the Immune-E subtype had limited benefit from immunotherapy (Jiang et al., 2018). Interestingly, although the Immune-E subtype had little limited benefit from immunotherapy, IC50 analysis of chemotherapeutic agents showed that this subtype was more sensitive to chemotherapeutic agents, including Cisplatin, Sunitinib, Crizotinib, Dasatinib, Bortezomib and Midostaurin.

The above analysis showed a high heterogeneity among different immune subtypes. In order to better understand the differences between different subtypes, we used WGCNA analysis to identify hub genes among each subtype. Extensive analyses showed that WGCNA was an effective method for identifying phenotype-genotype linkages and biomarkers and therapeutic targets (Niemira et al., 2019; Ma et al., 2021; Zhang et al., 2022). Nine modules were obtained, and pink module was highly associated with Stromal-E subtype. Further function

enrichment analysis showed that the hub genes in pink module were involved in ECM-receptor interaction pathway. Tumor progression depends not only on cell-autonomous changes in tumor cells, but also on the changes within the microenvironment (Götte and Kovalszky, 2018). An important feature of the dysregulated lung cancer microenvironment is the altered extracellular matrix (ECM), which can promote tumor angiogenesis, allow tumor cell immune escape, etc. (Mahale et al., 2016). In addition, magenta module was positively correlated with Immune-E subtype and participated in natural killer cell mediated cytotoxicity.

Notably, we developed an IMscore model containing six immune-related genes (MARCKS, CDK2, SFN, SSBP1, MRE11 and FZD7) to predict LUAD prognosis. CCNA2-CDK2 complex have been reported to inhibit LUAD progression (Li et al., 2021). CDK2 also was a biomarker for other cancers and next-generation CDK2 inhibitors play an increasingly pivotal role in the treatment of cancer (Wadler, 2001; Chohan et al., 2015; Tadesse et al., 2020). SFN gene encodes a protein participating in regulating epithelial-mesenchymal interaction (Asdaghi et al., 2012). Aya Shiba-Ishii reported that SFN promoted early progression of LUAD by activating cell proliferation (Shiba-Ishii, 2021). SSBP1, MRE11 and FZD7 have been considered as potential treatment sites of LUAD (Sun et al., 2017; Wang et al., 2017; Zeng et al., 2020). Kaplan-Meier survival analysis showed the high IMscore group had a significantly lower prognosis than the low IMscore group. Then, we detected the IMscore in three immune-related subtypes. As expected, the lowest score of IMscore in Stromal-E subtype was obtained. To explore the efficiency of the IMscore model, we used the expression profile data of the remaining 32 cancer types in the TCGA database for validation, and found that IMscore was significantly different in all cancer types except ESCA.

5 Conclusion

In this study, we used immune-related signaling pathways to classify LUAD and obtained three different subtypes, which had great differences in survival, gene mutation, CNV, gene methylation, etc. We further constructed an IMscore model to predict the prognosis among different subtypes of LUAD, and observed that the IMscore model had a higher efficiency. In addition, data in different cancers further confirmed the validity of the IMscore model.

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 author.

Author contributions

All authors contributed to this present work: ML designed the study, QX acquired the data, XY drafted the manuscript, YZ and CQ revised the manuscript. All authors read and approved the manuscript.

Conflict of interest

YZ was employed by YuceBio Technology Co.,

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

References

- Antonia, S. J., Villegas, A., Daniel, D., Vicente, D., Murakami, S., Hui, R., et al. (2017). Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. *N. Engl. J. Med.* 377 (20), 1919–1929. doi:10.1056/NEJMoa1709937
- Asdaghi, N., Kilani, R. T., Hosseini-Tabatabaei, A., Odemuyiwa, S. O., Hackett, T.-L., Knight, D. A., et al. (2012). Extracellular 14-3-3 from human lung epithelial cells enhances MMP-1 expression. *Mol. Cell. Biochem.* 360 (1), 261–270. doi:10.1007/s11010-011-1065-1
- Benjamin, D., Sato, T., Cibulskis, K., Getz, G., Stewart, C., and Lichtenstein, L. (2019). Calling somatic SNVs and indels with Mutect2. *BioRxiv* 2019, 861054.
- Bie, F., Tian, H., Sun, N., Zang, R., Zhang, M., Song, P., et al. (2021). Comprehensive analysis of PD-L1 expression, tumor-infiltrating lymphocytes, and tumor microenvironment in LUAD: Differences between asians and caucasians. *Clin. Epigenetics* 13 (1), 229–315. doi:10.1186/s13148-021-01221-3
- Bittner, N., Ostoros, G., and Géczi, L. (2014). New treatment options for lung adenocarcinoma-in view of molecular background. *Pathol. Oncol. Res.* 20 (1), 11–25. doi:10.1007/s12253-013-9719-9
- Botling, J., Edlund, K., Lohr, M., Hellwig, B., Holmberg, L., Lambe, M., et al. (2013). Biomarker discovery in non-small cell lung cancer: Integrating gene expression profiling, meta-analysis, and tissue microarray validation. *Clin. Cancer Res.* 19 (1), 194–204. doi:10.1158/1078-0432.CCR-12-1139
- Bronte, G., Rizzo, S., La Paglia, L., Adamo, V., Siragusa, S., Ficorella, C., et al. (2010). Driver mutations and differential sensitivity to targeted therapies: A new approach to the treatment of lung adenocarcinoma. *Cancer Treat. Rev.* 36, S21–S29. doi:10.1016/S0305-7372(10)70016-5
- Charloux, A., Quoix, E., Wolkove, N., Small, D., Pauli, G., and Kreisman, H. (1997). The increasing incidence of lung adenocarcinoma: Reality or artefact? A review of the epidemiology of lung adenocarcinoma. *Int. J. Epidemiol.* 26 (1), 14–23. doi:10.1093/ije/26.1.14
- Chohan, T. A., Qian, H., Pan, Y., and Chen, J.-Z. (2015). Cyclin-dependent kinase-2 as a target for cancer therapy: Progress in the development of CDK2 inhibitors as anti-cancer agents. *Curr. Med. Chem.* 22 (2), 237–263. doi:10.2174/0929867321666141106113633
- Denisenko, T. V., Budkevich, I. N., and Zhivotovsky, B. (2018). Cell death-based treatment of lung adenocarcinoma. *Cell. Death Dis.* 9 (2), 117–214. doi:10.1038/s41419-017-0063-y
- Der, S. D., Sykes, J., Pintilie, M., Zhu, C.-Q., Strumpf, D., Liu, N., et al. (2014). Validation of a histology-independent prognostic gene signature for early-stage, non-small-cell lung cancer including stage IA patients. *J. Thorac. Oncol.* 9 (1), 59–64. doi:10.1097/JTO.000000000000042
- Gao, H., Liang, J., Duan, J., Chen, L., Li, H., Zhen, T., et al. (2021). A prognosis marker SLC2A3 correlates with EMT and immune signature in colorectal cancer. *Front. Oncol.* 11, 638099. doi:10.3389/fonc.2021.638099
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* 6 (269), pl1. doi:10.1126/scisignal.2004088
- Goel, M. K., Khanna, P., and Kishore, J. (2010). Understanding survival analysis: Kaplan-Meier estimate. *Int. J. Ayurveda Res.* 1 (4), 274–278. doi:10.4103/0974-7788.76794
- Götte, M., and Kovalszky, I. (2018). Extracellular matrix functions in lung cancer. *Matrix Biol.* 73, 105–121. doi:10.1016/j.matbio.2018.02.018
- Hastie, T., Tibshirani, R., Narasimhan, B., Chu, G., and Narasimhan, M. B. (2011). Package 'impute'. *biocViews Bioinformatics*.
- Hirsch, F. R., Scagliotti, G. V., Mulshine, J. L., Kwon, R., Curran, W. J., Jr, Wu, Y.-L., et al. (2017). Lung cancer: Current therapies and new targeted treatments. *Lancet* 389 (10066), 299–311. doi:10.1016/S0140-6736(16)30958-8
- Jiang, H., Xu, S., and Chen, C. (2020). A ten-gene signature-based risk assessment model predicts the prognosis of lung adenocarcinoma. *BMC cancer* 20 (1), 782–811. doi:10.1186/s12885-020-07235-z
- Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24 (10), 1550–1558. doi:10.1038/s41591-018-0136-1
- Jiang, Z., Luo, Y., Zhang, L., Li, H., Pan, C., Yang, H., et al. (2022). A novel risk score model of lactate metabolism for predicting over survival and immune signature in lung adenocarcinoma. *Cancers* 14 (15), 3727. doi:10.3390/cancers14153727
- Knijnenburg, T. A., Wang, L., Zimmermann, M. T., Chambwe, N., Gao, G. F., Cherniack, A. D., et al. (2018). Genomic and molecular landscape of DNA damage repair deficiency across the Cancer Genome Atlas. *Cell. Rep.* 23 (1), 239–254. e236. doi:10.1016/j.celrep.2018.03.076
- Ladanyi, M., and Pao, W. (2008). Lung adenocarcinoma: Guiding EGFR-targeted therapy and beyond. *Mod. Pathol.* 21 (2), S16–S22. doi:10.1038/modpathol.3801018
- Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9 (1), 559–613. doi:10.1186/1471-2105-9-559
- Li, L., and Wang, X. (2021). Identification of gastric cancer subtypes based on pathway clustering. *NPJ Precis. Oncol.* 5 (1), 46. doi:10.1038/s41698-021-00186-z
- Li, Z., Zhang, Y., Zhou, Y., Wang, F., Yin, C., Ding, L., et al. (2021). Tanshinone IIA suppresses the progression of lung adenocarcinoma through regulating CCNA2-CDK2 complex and AURKA/PLK1 pathway. *Sci. Rep.* 11 (1), 23681–23712. doi:10.1038/s41598-021-03166-2
- Ma, Q., Geng, K., Xiao, P., and Zeng, L. (2021). Identification and prognostic value exploration of radiotherapy sensitivity-associated genes in non-small-cell lung cancer. *Biomed. Res. Int.* 2021, 5963868. doi:10.1155/2021/5963868
- Mahale, J., Smaguraukaite, G., Brown, K., Thomas, A., and Howells, L. M. (2016). The role of stromal fibroblasts in lung carcinogenesis: A target for chemoprevention? *Int. J. Cancer* 138 (1), 30–44. doi:10.1002/ijc.29447
- Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118

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/fphar.2022.1081244/full#supplementary-material>

- Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhi, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 12 (4), R41–R14. doi:10.1186/gb-2011-12-4-r41
- Niemira, M., Collin, F., Szalkowska, A., Bielska, A., Chwialkowska, K., Reszec, J., et al. (2019). Molecular signature of subtypes of non-small-cell lung cancer by large-scale transcriptional profiling: Identification of key modules and genes by weighted gene co-expression network analysis (WGCNA). *Cancers* 12 (1), 37. doi:10.3390/cancers12010037
- Pintarelli, G., Noci, S., Maspero, D., Pettinichio, A., Dugo, M., De Cecco, L., et al. (2019). Cigarette smoke alters the transcriptome of non-involved lung tissue in lung adenocarcinoma patients. *Sci. Rep.* 9 (1), 13039–13110. doi:10.1038/s41598-019-49648-2
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007
- Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell.* 160 (1–2), 48–61. doi:10.1016/j.cell.2014.12.033
- Rousseaux, S., Debernardi, A., Jacquiau, B., Vitte, A.-L., Vesin, A., Nagy-Mignotte, H., et al. (2013). Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci. Transl. Med.* 5 (186), 186ra66. doi:10.1126/scitranslmed.3005723
- Saito, M., Suzuki, H., Kono, K., Takenoshita, S., and Kohno, T. (2018). Treatment of lung adenocarcinoma by molecular-targeted therapy and immunotherapy. *Surg. Today* 48 (1), 1–8. doi:10.1007/s00595-017-1497-7
- Schwartz, M., Zhang, Y., and Rosenblatt, J. D. (2016). B cell regulation of the anti-tumor response and role in carcinogenesis. *J. Immunother. Cancer* 4 (1), 40–15. doi:10.1186/s40425-016-0145-x
- Shiba-Ishii, A. (2021). Significance of stratifin in early progression of lung adenocarcinoma and its potential therapeutic relevance. *Pathol. Int.* 71 (10), 655–665. doi:10.1111/pin.13147
- Shurin, M. R. (2018). Immunological targets for cancer therapy: New recognition. *Immunotargets Ther.* 7, 83–85. doi:10.2147/ITT.S191821
- Sorrentino, R., Morello, S., Forte, G., Montinaro, A., Vita, G. D., Luciano, A., et al. (2011). B cells contribute to the antitumor activity of CpG-oligodeoxynucleotide in a mouse model of metastatic lung carcinoma. *Am. J. Respir. Crit. Care Med.* 183 (10), 1369–1379. doi:10.1164/rccm.201010-1738OC
- Spella, M., and Stathopoulos, G. T. (2021). Immune resistance in lung adenocarcinoma. *Cancers* 13 (3), 384. doi:10.3390/cancers13030384
- Spigel, D. R., Schrock, A. B., Fabrizio, D., Frampton, G. M., Sun, J., He, J., et al. (2016). *Total mutation burden (TMB) in lung cancer (LC) and relationship with response to PD-1/PD-L1 targeted therapies*. Virginia, United States: American Society of Clinical Oncology.
- Sun, Y., Xu, T., Cao, Y., and Ding, X. (2017). Antitumor effect of miR-27b-3p on lung cancer cells via targeting Fzd7. *Eur. Rev. Med. Pharmacol. Sci.* 21 (18), 4113–4123.
- Surget, S., Khoury, M. P., and Bourdon, J.-C. (2014). Uncovering the role of p53 splice variants in human malignancy: A clinical perspective. *Onco. Targets. Ther.* 7, 57–68. doi:10.2147/OTT.S53876
- Tadesse, S., Anshabo, A. T., Portman, N., Lim, E., Tilley, W., Caldon, C. E., et al. (2020). Targeting CDK2 in cancer: Challenges and opportunities for therapy. *Drug Discov. Today* 25 (2), 406–413. doi:10.1016/j.drudis.2019.12.001
- Tan, Q., Cui, J., Huang, J., Ding, Z., Lin, H., Niu, X., et al. (2016). Genomic alteration during metastasis of lung adenocarcinoma. *Cell. Physiol. Biochem.* 38 (2), 469–486. doi:10.1159/000438644
- Taube, J. M., Galon, J., Sholl, L. M., Rodig, S. J., Cottrell, T. R., Giraldo, N. A., et al. (2018). Implications of the tumor immune microenvironment for staging and therapeutics. *Mod. Pathol.* 31 (2), 214–234. doi:10.1038/modpathol.2017.156
- Thandra, K. C., Barsouk, A., Saginala, K., Aluru, J. S., and Barsouk, A. (2021). Epidemiology of lung cancer. *Contemp. Oncol.* 25 (1), 45–52. doi:10.5114/wo.2021.103829
- Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Yang, T.-H. O., et al. (2018). The immune landscape of cancer. *Immunity* 48 (4), 812–830. e814. doi:10.1016/j.immuni.2018.03.023
- Wadler, S. (2001). Perspectives for cancer therapies with cdk2 inhibitors. *Drug resist. updat.* 4 (6), 347–367. doi:10.1054/drup.2001.0224
- Wang, Y., Hu, L., Zhang, X., Zhao, H., Xu, H., Wei, Y., et al. (2017). Downregulation of mitochondrial single stranded DNA binding protein (SSBP1) induces mitochondrial dysfunction and increases the radiosensitivity in non-small cell lung cancer cells. *J. Cancer* 8 (8), 1400–1409. doi:10.7150/jca.18170
- Wang, Y., and Zhou, B. P. (2013). Epithelial-mesenchymal transition—A hallmark of breast cancer metastasis. *Cancer Hallm.* 1 (1), 38–49. doi:10.1166/ch.2013.1004
- Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170
- Yamauchi, M., Yamaguchi, R., Nakata, A., Kohno, T., Nagasaki, M., Shimamura, T., et al. (2012). Epidermal growth factor receptor tyrosine kinase defines critical prognostic genes of stage I lung adenocarcinoma. *PLoS One* 7, e43923. doi:10.1371/journal.pone.0043923
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-García, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4 (1), 2612–2711. doi:10.1038/ncomms3612
- Zamarron, B. F., and Chen, W. (2011). Dual roles of immune cells and their factors in cancer development and progression. *Int. J. Biol. Sci.* 7 (5), 651–658. doi:10.1155/ijbs.7.651
- Zeng, Y., Jie, X., Wu, B., Wu, G., Liu, L., and Xu, S. (2020). IQGAP3 interacts with Rad17 to recruit the Mre11-Rad50-Nbs1 complex and contributes to radioresistance in lung cancer. *Cancer Lett.* 493, 254–265. doi:10.1016/j.canlet.2020.08.042
- Zhang, F.-L., Zhu, W.-M., He, T.-R., Zhao, Y.-T., Ge, W., Tan, J.-H., et al. (2022). Comparative transcriptomic analysis reveals that TPX2 and AURXA are involved in porcine PCV2 infection. *Gene* 834, 146649. doi:10.1016/j.gene.2022.146649
- Zhang, L., Chen, J., Cheng, T., Yang, H., Li, H., and Pan, C. (2020). Identification of the key genes and characterizations of tumor immune microenvironment in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC). *J. Cancer* 11 (17), 4965–4979. doi:10.7150/jca.42531