



Multi-Omics Signatures Identification for LUAD Prognosis Prediction Model Based on the Integrative Analysis of Immune and Hypoxia Signals

Yuqing Lou^{1†}, Qin Shi^{2†}, Yanwei Zhang¹, Ying Qi³, Wei Zhang¹, Huimin Wang¹, Jun Lu^{1,4,5,6*}, Baohui Han^{1,4,5*} and Hua Zhong^{1,5*}

¹Department of Pulmonary Medicine, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China, ²Department of Oncology, Baoshan Branch of Shuguang Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China, ³School of Basic Medical Science, Hangzhou Normal University, Hangzhou, China, ⁴Shanghai Institute of Thoracic Oncology, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China, ⁵Translational Medical Research Platform for Thoracic Oncology, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China, ⁶Department of Bio-bank, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China

OPEN ACCESS

Edited by:

Haipeng Liu,
Tongji University, China

Reviewed by:

Mohammad Shah Alam,
Bangabandhu Sheikh Mujibur
Rahman Agricultural University,
Bangladesh
Yaohui Chen,
Sichuan University, China

*Correspondence:

Jun Lu
lujun512@yahoo.com
Baohui Han
18930858216@163.com
Hua Zhong
eddielong8@hotmail.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Signaling,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 21 December 2021

Accepted: 21 January 2022

Published: 10 March 2022

Citation:

Lou Y, Shi Q, Zhang Y, Qi Y, Zhang W,
Wang H, Lu J, Han B and Zhong H
(2022) Multi-Omics Signatures
Identification for LUAD Prognosis
Prediction Model Based on the
Integrative Analysis of Immune and
Hypoxia Signals.
Front. Cell Dev. Biol. 10:840466.
doi: 10.3389/fcell.2022.840466

Lung adenocarcinoma (LUAD) is the most common histological subtype of lung cancer with heterogeneous outcomes and diverse therapeutic responses. However, the understanding of the potential mechanism behind LUAD initiation and progression remains limited. Increasing evidence shows the clinical significance of the interaction between immune and hypoxia in tumor microenvironment. To mine reliable prognostic signatures related to both immune and hypoxia and provide a more comprehensive landscape of the hypoxia-immune genome map, we investigated the hypoxia-immune-related alteration at the multi-omics level (gene expression, somatic mutation, and DNA methylation). Multiple strategies including lasso regression and multivariate Cox proportional hazards regression were used to screen the signatures with clinical significance and establish an incorporated prognosis prediction model with robust discriminative power on survival status on both the training and test datasets. Finally, combining all the samples, we constructed a robust model comprising 19 signatures for the prognosis prediction of LUAD patients. The results of our study provide a comprehensive landscape of hypoxia-immune related genetic alterations and provide a robust prognosis predictor for LUAD patients.

Keywords: lung adenocarcinoma, multi-omics biomarker, immune, hypoxia, prognosis prediction

INTRODUCTION

Lung cancer is one of the most common and severe types of cancer and presents the leading cause of both incidence and mortality worldwide in both genders (Siegel et al., 2019; Lu et al., 2021a; Lu et al., 2021b; Chu et al., 2021). Lung adenocarcinoma (LUAD) is the most common histological subtype of lung cancer with an increasing incidence over the past few decades (Ferlay et al., 2010; Lou et al., 2020; Lou et al., 2021). Although advances on treatment strategies for LUAD has been made, the overall 5-year survival rate is still at a low level with unoptimistic prognosis (less than 20%) (Wu et al., 2021). Although the routine display of clinicopathologic features by WHO classification and TNM staging system is important for the selection of appropriate treatment (Barth, 2020), these approaches appear to be inadequate due to the heterogeneity among patients.

The tumor microenvironment (TME), consisting of tumor cells, endothelial cells, immune cells, fibroblasts, macrophages, and the extracellular matrix, is a key regulator of carcinogenesis that substantially influences the initiation, development, and progression of LUAD, as well as the response to various therapy approaches (Anderson and Simon, 2020; Fu et al., 2021). Different components of the TME can regulate the development and progression of tumor. The immune cells, which are called tumor-infiltrating lymphocytes or TILs, can detect and destroy abnormal cells and potentially prevent or curb the tumor growth (Lei et al., 2020; Hajiran et al., 2021). In immune cells, reactive oxygen species (ROS) are mediators of several pivotal functions (e.g., phagocytosis, antigen presentation and recognition, cytolysis, as well as phenotypical differentiation) and exert immunosuppressive effects on T and natural killer (NK) cells (Kennel and Greten, 2021). Most recently, hypoxia has been reported as an intrinsic characteristic of solid tumor and plays an important role in cancer progression, angiogenesis, metastasis, and resistance to therapy (Muz et al., 2015). In the tumor microenvironment, the blood vessels and fibroblasts can influence the perfusion and diffusion of O₂, leading to the development of hypoxia in the tissue region (Belcher et al., 2020). Beyond the vasoconstriction, hypoxia recruits bone marrow precursor cells to the lung and affects the behavior of immune cells (Nicolls and Voelkel, 2007). Moreover, hypoxia is one of the reasons for poor therapy efficacy of current anti-angiogenic drugs and was reported to be associated with resistance to PD-1 blockade in squamous cell carcinoma of the head and neck (Jiang et al., 2020; Zandberg et al., 2021). Salem et al. (2018) outlined past and ongoing hypoxia-targeted therapy trials in NSCLC and highlight the potential of hypoxia as a therapeutic target. Therefore, further study on the relationship between hypoxia and immunity in LUAD is required to develop new therapeutic strategies.

In this study, we hypothesized that immune and hypoxia interaction may provide prognostic value in LUAD patients. Based on the expression profiles from The Cancer Genome Atlas (TCGA) portal, we aim to identify the hypoxia and immune status for each sample using the expression of pan-cancer metagenes for 28 immune cell subpopulations and the hypoxia related genes, respectively. Then, we will correlate the hypoxia-immune status with multi-omics genetic alterations to screen the hypoxia-immune biomarkers and finally establish an incorporated prognosis prediction model. The results of this study are expected to provide a more comprehensive hypoxia-immune genome map and may provide a better prognosis predictor for LUAD patients.

RESULTS

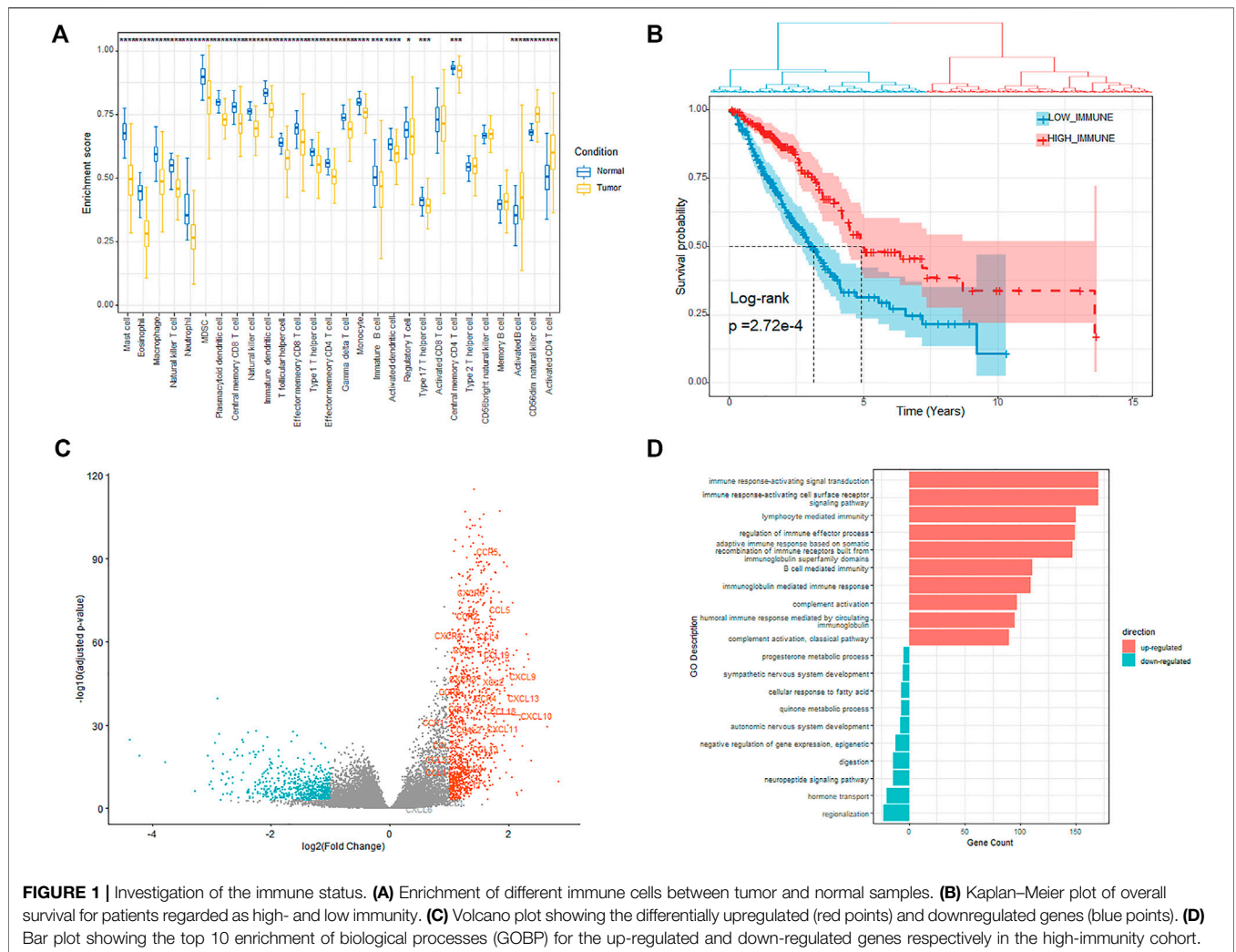
Immune Status and Immune-Related DEGs in LUAD

Based on the immune-related genes (IRGs) for 28 immune cell subpopulations provided in Charoentong et al. (2017) study, we calculated the enrichment scores (ESs) for each of the 569 samples (including 510 tumor samples and 58 normal

samples) using the RNA-seq profile by Gene Set Variation Analysis (GSEA) (Hänzelmann et al., 2013). The results showed that the ESs of 25 in 28 immune cells members were significantly different between the tumor and normal samples. Most of the immune cell members were significantly enriched in the normal samples rather than in tumor samples, except the activate B cell, CD56^{dim} natural killer cell, and activate CD4 T cell (**Figure 1A**). The tumor-infiltrating B cells (TIBs) play a multifaceted dual role in regulating tumor immunity rather than just tumor inhibition or promotion and affect the function of other immune cells such as CD4⁺ T cells and natural killer cells in the tumor microenvironment (Guo and Cui, 2019). We also observed that the enrichment of several immune cell members was also significantly different among different tumor stages (**Supplementary Figure S1**). Based on the ESs profile of the 28 types of immune cells in all the tumor samples, we defined the immune status for the 510 primary tumor samples and divided the related LUAD patients into two groups using hierarchical clustering with “ward.D” agglomeration method, which aims to find compact, spherical clusters by selecting clusters to merge based on the change in the cluster variances (**Figure 1B**), yielding 215 and 295 patients in the two groups, respectively. Survival comparison showed significant differences among the two groups (HR = 0.566, *p*-value = 2.72e-4), and the groups with better prognosis were labeled as IMMUNITY_H and others as IMMUNITY_L.

We next explored the expression alteration between the IMMUNITY_H and IMMUNITY_L cohorts to identify the immune-related DEGs. The genes with fold change larger than 2 and FDR less than 0.001 were regarded as differentially expressed, of which 1,118 and 628 genes were, respectively, up-regulated and down-regulated in the IMMUNITY_H cohort (**Figure 1C**). From the results, we observed that most of the chemotactic factors (e.g., CCR5, CXCR6, and CCL5), which are pivotal mediators of host defense and orchestrate the recruitment of immune cells into sites of infection and inflammation, were significantly up-regulated in the IMMUNITY_H samples.

The functional enrichment analyses of the up-regulated and down-regulated genes were performed using clusterProfile package (Yu et al., 2012). The results showed that the up-regulated genes were enriched in immune-related biological processes such as T cell activation and leukocyte proliferation, which indicated that the up-regulated genes play a positive role in the enhancement of tumor-associated immunity (**Figure 1D**, **Supplementary Table S1**). On the other hand, the down-regulated genes were mainly enriched in nervous system development related biological processes, which indicated that some downregulated genes modulate the activities of immune and tumor cells through affecting nervous system development (**Supplementary Table S1**). For example, PHOX2A/B, which is a paired-like homeodomain transcription factor that participates in specifying the autonomic nervous system, was verified as a tumor suppressor (Wilzén et al., 2009; Pudela et al., 2020). The KEGG pathway enrichment analysis results also showed that the up-regulated genes were mainly enriched in immune-related pathways, while the down-regulated genes were enriched in



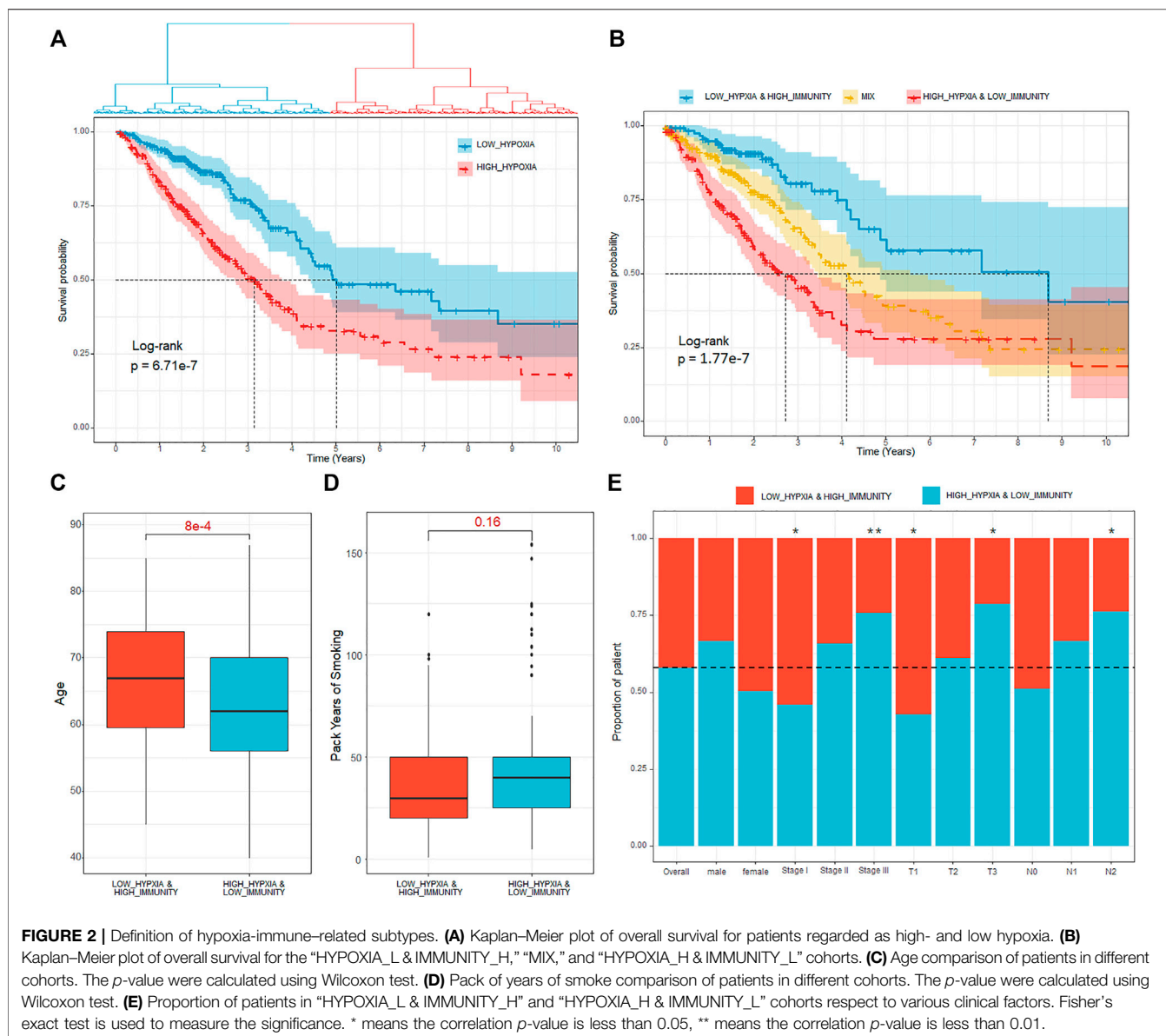
the pathways related with nervous system development and metabolism (**Supplementary Table S1**).

Identification of Hypoxia-Immune-Related Subtype and Associated Prognostic DEGs

To deduce the hypoxia status for each of the samples, we extracted the expression of the 200 hypoxia-related hallmark genes and then handled with UMAP (Uniform Manifold Approximation and Projection). Using the latent variables generated by UMAP, we further divided the patients into two groups (**Figure 2A, Methods and Materials**). There were 249 and 261 patients in the two groups and the survival analysis showed significant difference between the two groups (**Figure 2A**, HR = 2.15, p -value = $6.71e-7$). The patients with better prognosis were assigned to HYPOXIA_L group and others to HYPOXIA_H group. Taking the immune and hypoxia statuses together, we divided the patients into three groups, which are “HYPOXIA_L & IMMUNITY_H” ($n = 124$), “HYPOXIA_H & IMMUNITY_L” ($n = 170$), and “MIX” ($n = 216$). The survival analysis results revealed that the OS times of patients in different groups were

significantly different (**Figure 2B**, HR = 3.33, p -value = $1.77e-7$), and the “HYPOXIA_L & IMMUNITY_H” cohort harbored the best prognosis, while the patients in the “HYPOXIA_H & IMMUNITY_L” yield the worst prognosis.

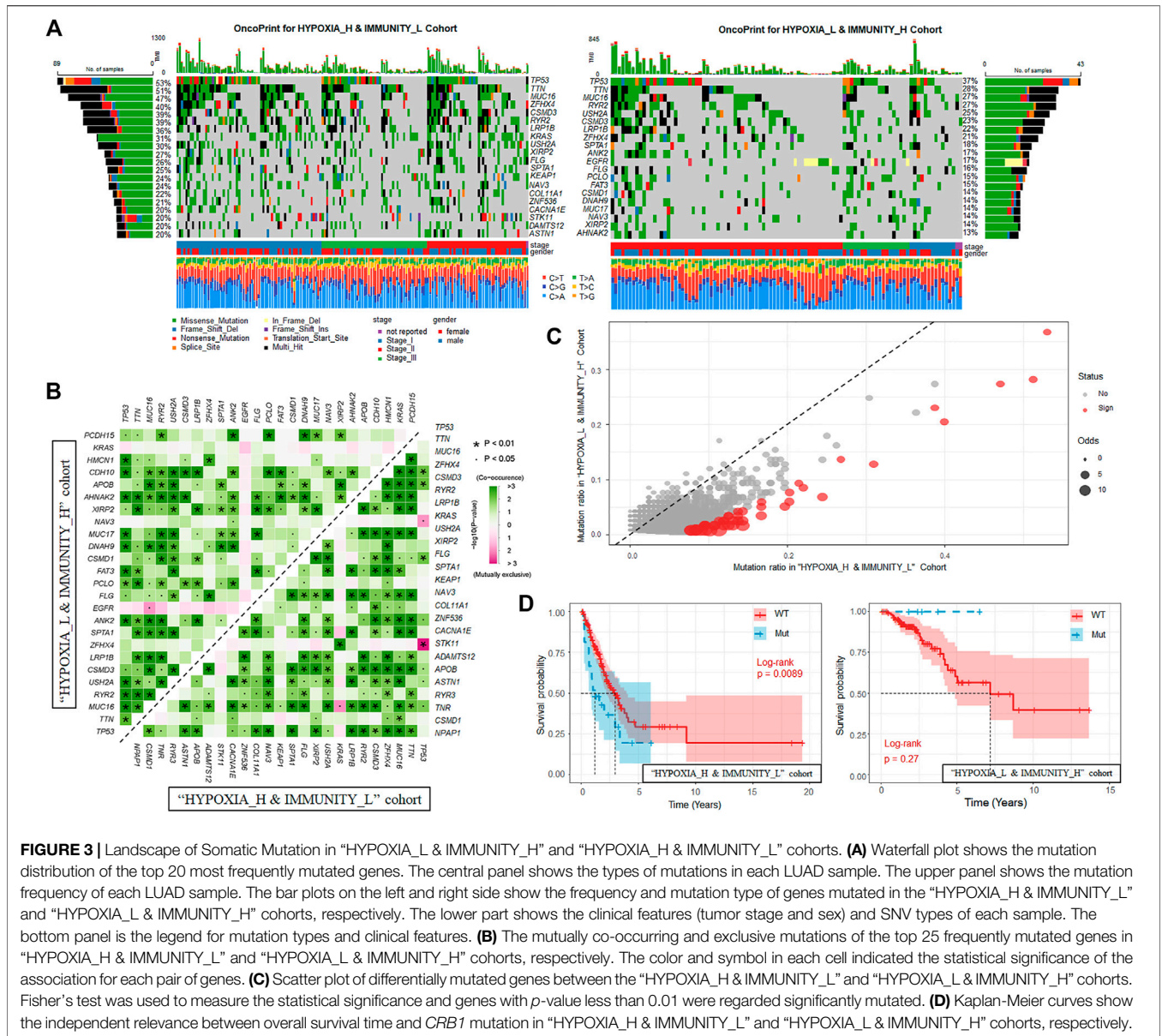
We further investigated the dispersion of various clinical characters (e.g., age, clinical stage, tumor size, lymph node, and distant metastasis) between the cohorts with different hypoxia-immune status. Through Cox proportional hazards regression analysis, we observed that the OS time is independent from age (HR = 1, p -value = 0.64). However, we observed that the patients regarded as “HYPOXIA_H & IMMUNITY_L” were significantly younger than patients in the “HYPOXIA_L & IMMUNITY_H” group (**Figure 2C**, Wilcox test p -value = $8e-4$), which may explain the clinical observation that young lung patients tend to present with advanced disease at diagnosis, resulting in an extremely poor survival (Rocha et al., 1994). Besides that, we also observed that the patients with more packs years of smoke tend to enriched in the high-risk (“HYPOXIA_H & IMMUNITY_L”) cohort (**Figure 2D**). Moreover, we also pay attention to association between the immune-hypoxia status and various clinical



factors, such as gender and clinical stages. Generally, gender was independent from the immune-hypoxia status (**Figure 2E**). For the clinical stage, we observed that the patients in the stage I tended to be in the “HYPOXIA_L & IMMUNITY_H” cohort (Fisher’s exact test, p -value = 0.02, **Figure 2E**), while the patients in the stage III tended to be in the “HYPOXIA_H & IMMUNITY_L” cohort (Fisher’s exact test, p -value = 0.007, **Figure 2E**). As the distant metastases were present in only 4.48% of the patients selected for comparison, we only consider the “N” (Regional lymph nodes) and “T” (Primary tumor) for the TNM dispersion analysis. The results showed that the patients with higher tumor size were significantly enriched in the “HYPOXIA_H & IMMUNITY_L” group, and the patients with more lymph nodes that contain cancer were also significantly enriched in the “HYPOXIA_H & IMMUNITY_L”

group. These results further indicated that the patients in the “HYPOXIA_H & IMMUNITY_L” cohort tend to be high-risk.

The hypoxia-immune-related DEGs were obtained by comparing the expression between the “HYPOXIA_L & IMMUNITY_H” and “HYPOXIA_H & IMMUNITY_L” cohorts, and finally 2,798 DEGs were obtained with fold-change larger than two and adjusted p -value less than 0.001 (**Supplementary Table S2**). The 1,091 genes significantly up-regulated in the “HYPOXIA_H & IMMUNITY_L” cohort where patients yielded worse survival were regarded as risk DEGs (e.g., *GAPDH*, *NTS*, *LDHA*, and *CDH2*), and the 1,707 genes significantly up-regulated in the “HYPOXIA_L & IMMUNITY_H” cohort where patients yielded better outcome were regarded as protective DEGs (e.g., *RCSD1*, *IL16*, *PRB4*, and *VEGFD*).



Comparing Somatic Mutations Between Different Hypoxia-Immune Status

After identifying the gene signatures associated with the hypoxia-immune status, we also tended to explore the alteration at genome level between the “HYPOXIA_L & IMMUNITY_H” and “HYPOXIA_H & IMMUNITY_L” cohorts. The varscan2 results about single-nucleotide variant (SNV), single-nucleotide polymorphism (SNP), insertion (INS), and deletion (DEL) were used in this part. We observed that majority of the genomic variants were missense mutation in both the “HYPOXIA_L & IMMUNITY_H” and “HYPOXIA_H & IMMUNITY_L” cohorts (around 85%), while for most types, the samples in the “HYPOXIA_H & IMMUNITY_L” cohort harbored a

significantly larger number of variants than those in the “HYPOXIA_L & IMMUNITY_H” (Supplementary Table S3, Supplementary Figure S2). The ratios between transversion (Tv) and transition (Ti) in all SNVs were approximately 2:1 and remained stable in both cohorts. Moreover, we also observed that the TMB of the patients in the “HYPOXIA_H & IMMUNITY_L” cohort was significantly larger than that of patients in the “HYPOXIA_L & IMMUNITY_H” (Wilcox test p -value = 1.47×10^{-7}), which also indicated that the “HYPOXIA_H & IMMUNITY_L” is high-risk status.

In the “HYPOXIA_H & IMMUNITY_L” cohort, 181 genes were mutated in more than 10% of the samples while only 44 genes met this criterion in the “HYPOXIA_L & IMMUNITY_H”

cohort, of which there was an overlap of 42 genes. The top 20 most frequently mutated genes in the corresponding cohorts were shown in **Figure 3A**. From the results, we observed that TP53, TTN, and MUC16 rank among the top 3 most frequently mutated genes in the corresponding cohorts. These genes were reported to be interactive and regulated various tumor associated biological processes (Huang et al., 2021; Wu et al., 2021; Zhao et al., 2021; Zhou et al., 2021). We next investigated the co-occurring and exclusive mutation of the top 25 frequently mutated genes (**Figure 3B**). Compared with the pervasive co-occurrence landscape (280 cases), there were only four unique cases in the two cohorts (KRAS-TP53: p -value = 0.013, KRAS-TNR: p -value = 0.011, TP53-STK11: p -value = 1.87×10^{-4} , and MUC16-EGFR: p -value = 0.013) exhibiting mutually exclusive mutations, which suggests their probably redundant effect in the same pathway and selective advantages between them to keep more than one copy of the mutations. To extract the signatures at the somatic genome level, we applied Fisher's test to identify the differentially mutated genes between the two cohorts, and finally 54 genes were regarded as significantly differentially mutated (p -value < 0.01, **Figure 3C**). From the results, we found that the genes mutated more frequently in the "HYPOXIA_H & IMMUNITY_L" cohort than in the "HYPOXIA_L & IMMUNITY_H" cohort. To verify the same mutation may exert distinct impacts on the survival time of patients grouped in different cohorts, we divided the patients in both the "HYPOXIA_L & IMMUNITY_H" and "HYPOXIA_H & IMMUNITY_L" cohorts into "wt" and "mut" groups. The survival analysis results showed that several genes can divide the patients into two groups with significantly different OS times in one cohort, while they cannot in the other cohort (**Supplementary Table S4**). For example, the OS times of the patients in the "HYPOXIA_H & IMMUNITY_L" cohorts with and without *CRB1* mutation were significantly different (HR = 3.09, p -value = 5.11×10^{-6}), while no such significant difference was observed in "HYPOXIA_L & IMMUNITY_H" (**Figure 3D**). And TPR showed the opposite result (**Figure 3D**).

Comparing DNA Methylation Level Between Different Hypoxia-Immune Status

Aberrations in DNA methylation system have an important role in human disease, and DNA methylation patterns are globally disrupted in cancer, with genome-wide hypomethylation and gene-specific hypermethylation events occurring simultaneously in the same cell (Robertson, 2005). In this section, we aimed to identify and compare the effects of DNA methylation patterns in different hypoxia-immune cohorts using Illumina Infinium 450k DNA methylation data. Only the patients grouped into "HYPOXIA_L & IMMUNITY_H" or "HYPOXIA_H & IMMUNITY_L" cohorts were considered. After preprocessing, 264 samples in which no more than 20% probes have missing beta values were used to detect the differential methylation probes (DMPs) using ChAMP (Morris et al., 2014; Tian et al., 2017) (see *Methods and Materials*). Finally, 2,082 hypoxia-immune-related DMPs were identified with the criterion of absolute $\Delta\beta$ larger than 0.15 and adjusted p -value less than 0.05 (**Figure 4A**,

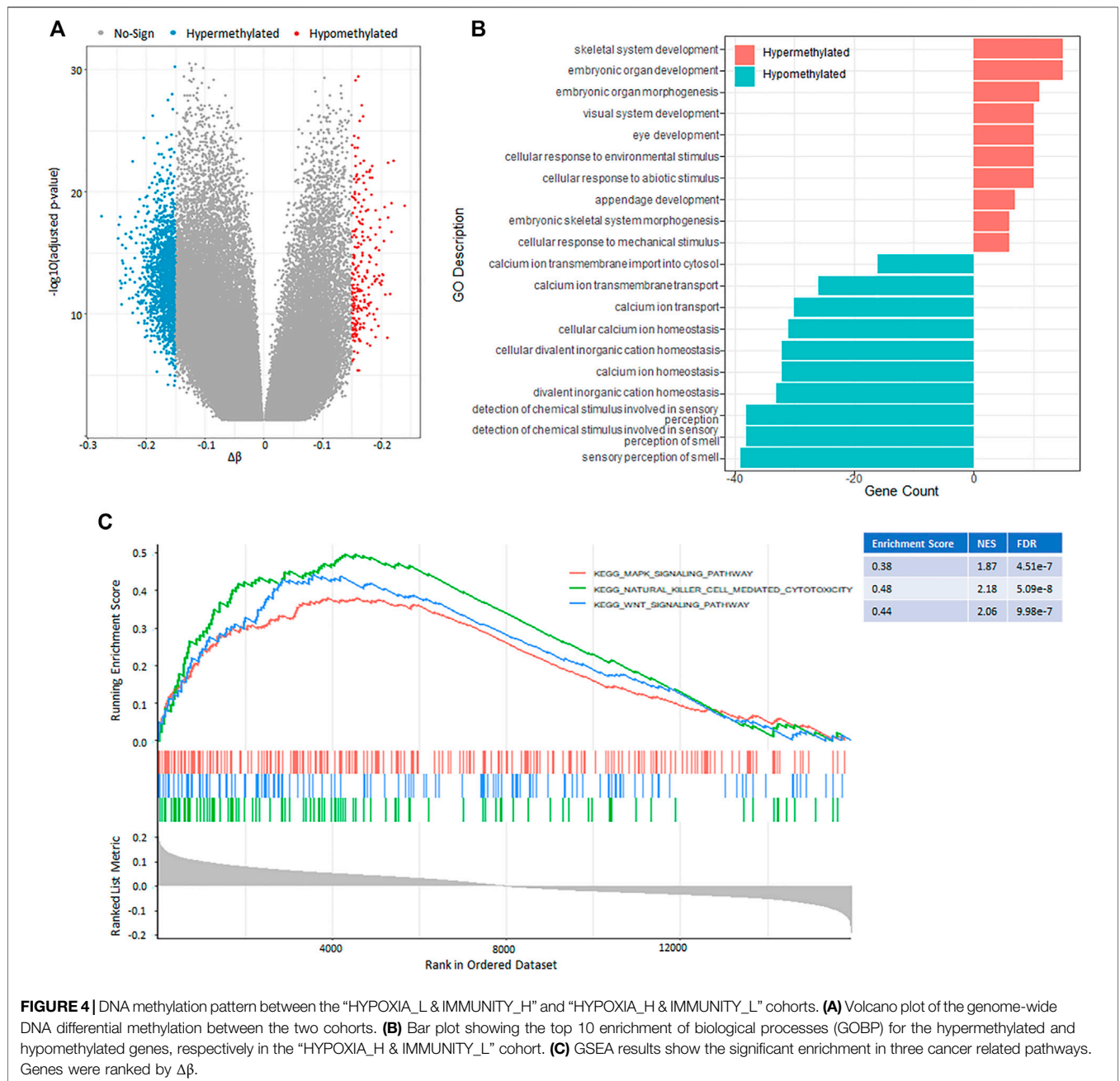
Supplementary Table S5). Compared with "HYPOXIA_L & IMMUNITY_H" cohort, 1844 (88.57%) hypomethylated positions involving 520 genes were identified in the "HYPOXIA_H & IMMUNITY_L" cohort, while only 238 (11.43%) positions related to 128 genes were significantly hypermethylated. These results indicated that the "HYPOXIA_H & IMMUNITY_L" cohort tends to have hypomethylated positions overall. Only 3 genes (*ZC3H12D*, *XKR6*, *DIP2C*) contain both hypermethylated and hypomethylated positions. Among these 520 hypomethylated genes in the "HYPOXIA_H & IMMUNITY_L" cohort, 29 and 23 genes were significantly upregulated and downregulated, respectively. In contrast, there were only 4 upregulated and 5 downregulated genes among the hypermethylated genes.

The functional enrichment analysis results revealed that the hypomethylated genes mainly involved in sensory perception, ion transport, and ion homeostasis, while the hypermethylated genes play potential roles in development and cellular response (**Figure 4B**, **Supplementary Table S6**). The gene set enrichment analysis (GSEA) of these DMP-associated genes showed that hypermethylated genes with highly positive beta difference have more essential contributions to various cancer related pathways such as natural killer cell mediated cytotoxicity, Wnt signal pathway, and Mapk signal pathway (**Figure 4C**, **Supplementary Table S7**).

Prognostic Prediction Using Multi-Omics Signatures

To obtain a more comprehensive and robust model for the prognostic prediction, we integrated the multi-omics genetic signatures obtained above. At the transcriptome level, a total of 1,091 up-regulated and 1,707 down-regulated genes were identified in the "HYPOXIA_H & IMMUNITY_L" cohort. At the genome level, 181 and 44 frequently mutated genes were identified in the "HYPOXIA_H & IMMUNITY_L" and "HYPOXIA_H & IMMUNITY_L" cohorts, respectively. At the DNA methylation level, 1,163 out of 2,208 DMPs locating at the region of 645 annotated genes were differentially methylated between the "HYPOXIA_H & IMMUNITY_L" and "HYPOXIA_H & IMMUNITY_L" cohorts. Furthermore, we refined the hypoxia-immune-related prognostic signatures with significant effect on the overall survival time of patients from these genetic alterations based on univariate Cox proportional hazards model. After that, 336 items composed of 230 DEGs, 9 mutations, and 97 DMPs were selected. Considering the large number of significant signatures and possible interaction among them, we applied LASSO Cox regression model to evaluate the extent to which signatures contributes to predicting survival. Under the optimal parameter $\ln(\lambda) = -3.3$ (**Figure 5A**), we reserved 39 signatures (27 DEGs, 8 mutations, and 4 DMPs) to establish the multivariate Cox proportional hazards regression model with stepwise method.

For the lack of the matching multi-omics data from other sources, we randomly divided TCGA samples into a training (70%, $n = 295$) and independent test set (30%, $n = 126$), and this process was repeated 5 times. The results showed that the



performance of the trained models is satisfied with the average concordance index (C-index) equal to 0.816. Next, the risk score for each sample was calculated based on the established models, which has great discriminative power on survival status. The average AUC values of 1-, 3-, and 5-year prognosis prediction on training sets reached 0.841, 0.86, and 0.853 (Figure 5B). With regard to the prediction on the test sets, the performance exhibited a similar performance with the average AUC value of 1-, 3-, and 5-year survival equal to 0.788, 0.755, and 0.805 (Figure 5B). Moreover, the samples were classified into high-risk and low-risk cohorts by median-risk score. Kaplan-Meier survival analysis showed that the high-risk cohorts had a poorer overall

survival compared with the low-risk cohort (p -value < 0.001, Supplementary Figure S3).

The above results revealed the great robustness and validity of the strategy of model construction, and we further combined all TCGA samples and generated an overall prediction model comprised of 19 signatures, including 11 DEGs, 7 mutations, and 1 DMPs (Figure 5C), from which we found some signatures, such as DEGs FSIP2, LINC01697, FAM83A, and ADM, seemed to be statistically insignificant initially (p -value > 0.05) but are likely associated with other signatures and outcome. The contributions of these 19 signatures on the overall model are listed in Table 1.

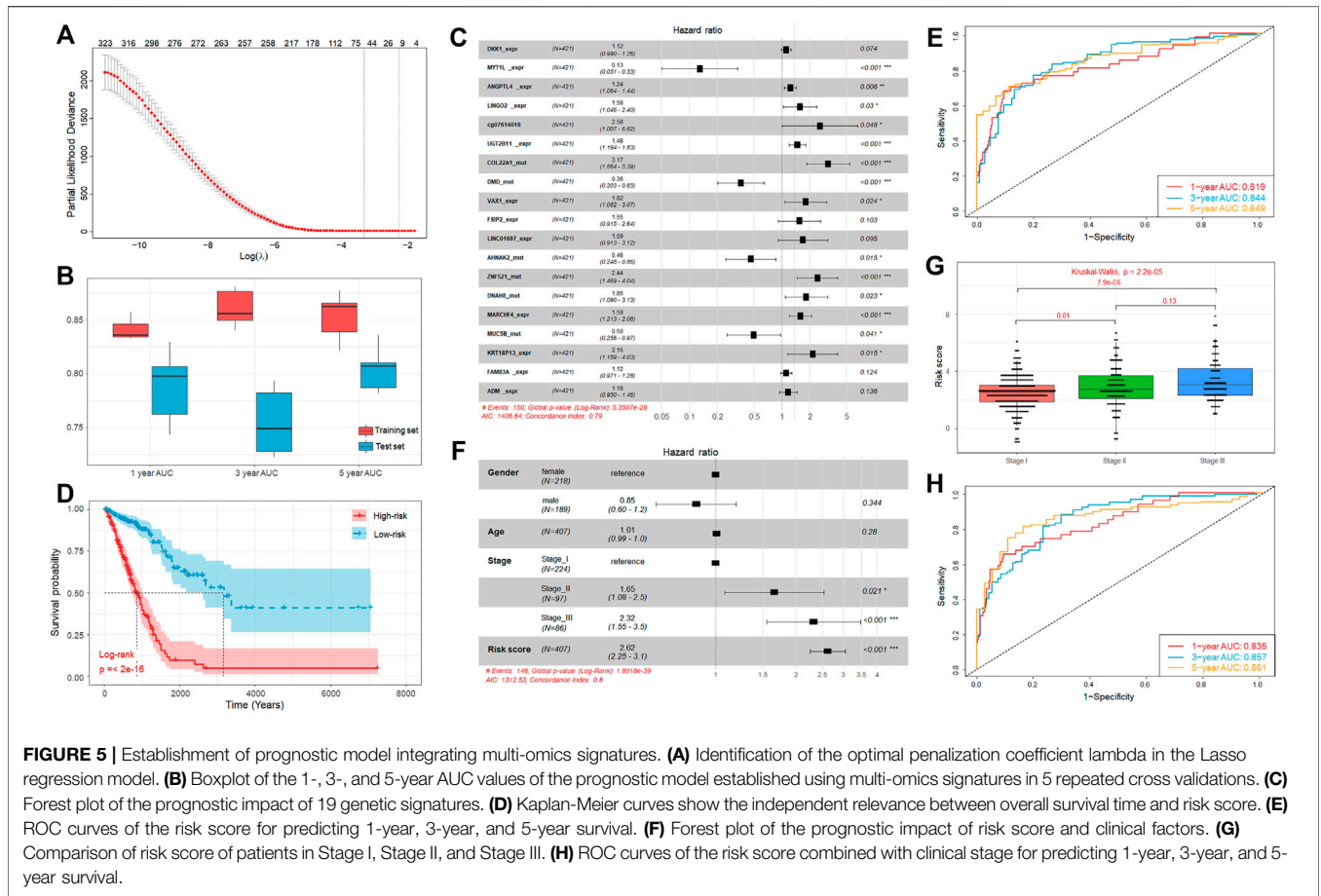


TABLE 1 | Nineteen signatures associated with overall survival by multivariate Cox regression analysis.

ID	Coefficient	HR	HR_95L	HR_95H	p-value
DKK1_expr	0.108858	1.115004	0.989575	1.25633	0.0738
MYT1L_mutation	-2.04102	0.129897	0.050676	0.332959	2.14E-05
ANGPTL4_expr	0.213808	1.238385	1.064096	1.441223	0.005732
LINGO2_expr	0.459297	1.582961	1.045753	2.396135	0.029895
cg07614018	0.948601	2.582095	1.007391	6.618301	0.048234
UGT2B11_expr	0.389956	1.476916	1.194457	1.82617	0.000317
COL22A1_mutation	1.153178	3.168244	1.863983	5.38512	2.04E-05
DMD_mutation	-1.01334	0.363003	0.203103	0.648791	0.000626
VAX1_expr	0.599708	1.821587	1.08203	3.066624	0.024032
FSIP2_expr	0.441211	1.554588	0.915153	2.640809	0.102679
LINC01697_expr	0.523275	1.687546	0.9127	3.120206	0.095184
AHNAK2_mutation	-0.77174	0.462206	0.248079	0.861156	0.015066
ZNF521_mutation	0.890135	2.435457	1.469187	4.037234	0.000557
DNAH8_mutation	0.613942	1.8477	1.089994	3.132124	0.022609
MARCHF4_expr	0.463818	1.590134	1.212916	2.084666	0.000788
MUC5B_mutation	-0.69462	0.499262	0.256239	0.972773	0.041243
KRT18P13_expr	0.77069	2.161258	1.158634	4.031502	0.015399
FAM83A_expr	0.109214	1.115401	0.970587	1.281821	0.123754
ADM_expr	0.163852	1.17804	0.949734	1.461228	0.13603

In brief, the mutation of MYT1L, DMD, AHNAK2, and MUC5B have significantly positive contribution to a better prognosis, while the others played opposite roles. Moreover,

similar to our above observation, the overall survival time of the high-risk cohort is significantly shorter than that of low-risk cohort (Figure 5D). Besides that, we also observed that high

discriminative power of risk score for the 1-year, 3-year, and 5-year survival rates, according to the respective AUC value, were 0.819, 0.844, and 0.849 (Figure 5E). To further prove that integrating multi-omic characteristics can provide more robust prognostic prediction than using single-omic characteristics, we adopted the same strategy as mentioned above for each type of omic data. The results showed that no single-omic characteristics can provide a stronger model than the integrated model (Supplementary Figure S4).

In addition to the genetic alteration, we also consider some clinical factors which may also have prognosis ability, such as stage, gender, and age. We found that the clinical stage was significantly associated with the overall survival time, but the gender and age were not (Figure 5F). We tested the association between different clinical factors and the risk score, and we found that the risk scores of patients in Stage III and Stage II were significantly larger than those of patients in Stage I (Figure 5G). Combining these clinical factors with risk score, we established an incorporated model and the results showed that the prognosis ability can be improved through incorporating the risk score with stage information (C-index = 0.803). Besides that, the incorporated model can also achieve a better performance on its 1-year (AUC = 0.835), 3-year (AUC = 0.857), and 5-year (AUC = 0.861) survival predictions (Figure 5H). Hence, the multi-omics signatures comprising the above 19 genetic alteration can produce an accurate prognostic prediction and the risk score calculated based on these multi-omics signatures can be regarded as an independent prognostic indicator.

METHODS AND MATERIALS

Patient Cohorts and Data Preparation

The Cancer Genome Atlas (TCGA) cohort consisted of 510 LUAD primary solid tumor samples and 58 normal control samples of RNA-seq profiles, 561 LUAD samples of WES data, and 455 profiles of the Illumina 450 k DNA methylation array, downloaded from the UCSC Xena browser (<http://xena.ucsc.edu>) (Goldman et al., 2020). To eliminate the error caused by the quantitative mRNA abundance of FPKM in multiple samples, we convert FPKM to TPM for standardization (Li et al., 2010).

Immunity Status Definition

The expression data was transformed into Transcripts per Million. The immune-related genes (IRGs) for 28 immune cell subpopulations were obtained from Charoentong et al. (2017) study. Using the normalized gene expression data, we calculated the enrichment score (ES) of the 28 immune cell types for each sample through Gene Set Variation Analysis (GSEA) (Hänzelmann et al., 2013). Using the ES profile, the LUAD samples of expression profiles were classified into two groups by hierarchical clustering, and survival analysis was performed according to the groups. Patients with good prognosis were assigned to IMMUNITY_H group and the others to IMMUNITY_L group. The ESTIMATE algorithm was used to generate an immune score through estimating the proportion of different infiltrating stromal and immune cells. The estimated scores

between the high-immunity and low-immunity cohorts were compared using the Mann-Whitney U-test. By mapping the sample IDs of the RNA-seq profiles, we also constructed immune-related cohorts for the LUAD samples of both the WES and DNA methylation profiles.

Identification of Hypoxia-Immune-Related Subtypes

The hypoxia related genes (HRGs) were downloaded from the molecular marker database (MSigDB v7.4). To deduce the hypoxia status, we applied the algorithm of uniform manifold approximation and projection (UMAP), which can be used for general non-linear dimension reduction, to reduce the dimension of the HRG expression profile, and the latent variable was used to cluster the patients into two groups using hierarchical clustering with “ward.D” agglomeration method. The survival analysis was performance on both the groups and the patients with poorer prognosis were regarded as HYPOXIA_L group and the others as HYPOXIA_H. Combining the immune status, we further divided the patients into three groups, which were “HYPOXIA_L & IMMUNITY_H,” “HYPOXIA_L & IMMUNITY_H,” and “MIX” groups.

Multi-Omics Data Analysis and Prognosis Prediction Model Construction

In this study, we aimed to investigate the difference in gene expression, somatic mutations, and DNA methylation between the “HYPOXIA_L & IMMUNITY_H” and “HYPOXIA_L & IMMUNITY_H” cohort, respectively. The differential gene expression analysis between the two cohorts were identified using DESeq2 package (Anders et al., 2013) with the RNA-seq raw count data. Genes with fold change larger than two and adjusted *p*-value less than 0.001 were regarded as significantly differentially expressed. For Somatic mutations, the SNVs, SNPs, and INDELS were detected using VarScan2 (Koboldt et al., 2012; Lu et al., 2019a), and the results were downloaded from the UCSC Xena browser. Fisher’s exact test was used to identify the differential mutation patterns, and genes with adjusted *p*-value less than 0.01 were regarded as differentially mutated genes. The co-occurrence and mutually exclusive mutations were identified using the CoMEt algorithm (Leiserson et al., 2015). For DNA methylation, we first filtered the samples with more than 20% missing values and the remaining miss values were imputed using “*champ.impute*” function in R package “ChAMP.” The differentially methylation probes were also identified using the “*champ.DMP*” function in R package “ChAMP.”

As the survival time and status were used to evaluated the prognosis of LUAD patients, we only used cases with an overall survival record to construct the prognosis prediction model based on gene expression, DNA methylation, and somatic mutation data, respectively. For the multi-omic integration-based model, only the cases with all three types of omic-data were used. The univariate Cox proportional hazards regression was first used to assess the individual effect of every alteration and features with *p*-value less than 0.05 were retained. After that, we used the lasso

regression method to further filter the less informative features. At last, the multivariate Cox proportion hazards regression with a stepwise procedure was used to reduce the redundant variables, and formed the final prognosis prediction model. The risk score for each sample was calculated based on the coefficients provided by the model. The C-index which was used to evaluate the performance of the prognosis prediction model was calculated using R package “survival.” The 1-, 3-, and 5-year receiver operating characteristic (ROC) curve were generated using the R package “timeROC.”

Functional Enrichment Analysis

GO analysis was performed as our previous studies (Lu et al., 2019b; Lu et al., 2019c; Lu et al., 2019d). The analysis of DEG and DMP-associated genes and GSEA were performed using the R package “clusterProfile.” The GSEA plot was generated using the R package “enrichplot.”

DISCUSSION

Considering the heterogeneous outcomes and diverse therapeutic responses of LUAD, it is essential to establish a robust predictor to evaluate the risk and prognosis of patients. Immune and hypoxia were reported to play a critical role in the tumor initiation and progression, and significantly associated with prognosis. However, comprehensive analysis of genetic alterations related to both the immune and hypoxia remains far from satisfactory. In this study, we take into account the immune and hypoxia signal and perform a multi-omics integrative analysis.

Based on the expression profile of immune and hypoxia related genes, we defined the immune and hypoxia status for each patient, respectively. As expected, the immune status is positively associated with prognosis, while the hypoxia status is the opposite. Combining both the immune and hypoxia status, the patients were further divided into three cohorts, which are “HYPOXIA_H & IMMUNITY_L,” “HYPOXIA_H & IMMUNITY_L,” and “MIX.” After that, we investigated the genetic alterations between the first two cohorts from the gene expression, DNA methylation, and somatic mutation layers and gave the signatures accordingly. With multi-step strategy, we refined 19 signatures and established a robust model stratify patients with different risks and prognosis.

Although our study provided a more comprehensive landscape of the hypoxia-immune genome map, there are several limits that need deeper investigation. As our model contains three types of omic data, including RNA-seq, WES, and DNA methylation array data, we cannot find other datasets to validate our model, which is the main limitation of our study. Moreover, as we all know, LUAD is very complex and heterogeneous, and it is difficult to cover all variations among the patients. Even so, with the rapid development of biological technologies, more researchers adopt multi-omics strategy to resolve the initiation and progression of LUAD. Besides that,

our comprehensive characterization of genetic alteration from different omic layers between the “HYPOXIA_H & IMMUNITY_L” and “HYPOXIA_H & IMMUNITY_L” status can also be solely used. Moreover, our prognostic prediction model can potentially exhibit compelling clinical value that may lead to the improvement of overall survival and even for the development of new therapeutic strategies for LUAD patients.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Cancer Genome Atlas. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HZ, BH, and JL conceived the study. Data analysis were performed by YL, QS, YZ, YQ, WZ, HW, and JL. Figures and tables were generated by YL, QS, and YZ, and the manuscript was written by YL. The manuscript was revised by JL, BH, and HZ.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 82104944); the foundation of Chinese society of clinical oncology (No. Y-2019AZZD-0355 and Y-QL2019-0125); the foundation of Shanghai Chest Hospital (No. 2020YNJCM06 and 2019YNJCM11 and YJXT20190102); the National Multi-disciplinary Treatment Project for Major Diseases (No. 2020NMDTP); the Natural Science Foundation of Zhejiang Province (No. LY20C060002), the Shanghai Leading Talents Program (2013), the program of system biomedicine innovation center from Shanghai Jiao Tong University (No. 15ZH4009 and YG2021QN121); the key program of translational medicine from Shanghai Jiao Tong University School of Medicine (No. 15ZH1008).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2022.840466/full#supplementary-material>

REFERENCES

- Anders, S., McCarthy, D. J., Chen, Y., Okoniewski, M., Smyth, G. K., Huber, W., et al. (2013). Count-based Differential Expression Analysis of RNA Sequencing Data Using R and Bioconductor. *Nat. Protoc.* 8, 1765–1786. doi:10.1038/nprot.2013.099
- Anderson, N. M., and Simon, M. C. (2020). The Tumor Microenvironment. *Curr. Biol.* 30, R921–R925. doi:10.1016/j.cub.2020.06.081
- Barth, M. J. (2020). Xliv-Xliv. *NCCN Clin. Pract. Guidel. Oncol. J Natl Compr Canc Ne* 18.
- Belcher, D. A., Lucas, A., Cabrales, P., and Palmer, A. F. (2020). Polymerized Human Hemoglobin Facilitated Modulation of Tumor Oxygenation Is Dependent on Tumor Oxygenation Status and Oxygen Affinity of the Hemoglobin-Based Oxygen Carrier. *Sci. Rep.* 10, 11372. doi:10.1038/s41598-020-68190-0
- Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cel Rep.* 18, 248–262. doi:10.1016/j.celrep.2016.12.019
- Chu, T., Lu, J., Bi, M., Zhang, H., Zhuang, W., Yu, Y., et al. (2021). Equivalent Efficacy Study of QL1101 and Bevacizumab on Untreated Advanced Non-squamous Non-small Cell Lung Cancer Patients: a Phase 3 Randomized, Double-Blind Clinical Trial. *Cancer Biol. Med.* doi:10.20892/j.issn.2095-3941.2020.0212
- Ferlay, J., Shin, H.-R., Bray, F., Forman, D., Mathers, C., and Parkin, D. M. (2010). Estimates of Worldwide burden of Cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* 127, 2893–2917. doi:10.1002/ijc.25516
- Fu, T., Dai, L.-J., Wu, S.-Y., Xiao, Y., Ma, D., Jiang, Y.-Z., et al. (2021). Spatial Architecture of the Immune Microenvironment Orchestrates Tumor Immunity and Therapeutic Response. *J. Hematol. Oncol.* 14, 98. doi:10.1186/s13045-021-01103-4
- Goldman, M. J., Craft, B., Hastie, M., Repčeka, K., McDade, F., Kamath, A., et al. (2020). Visualizing and Interpreting Cancer Genomics Data via the Xena Platform. *Nat. Biotechnol.* 38, 675–678. doi:10.1038/s41587-020-0546-8
- Guo, F. F., and Cui, J. W. (2019). The Role of Tumor-Infiltrating B Cells in Tumor Immunity. *J. Oncol.* 2019, 2592419. doi:10.1155/2019/2592419
- Hajiran, A., Chakiryan, N., Aydin, A. M., Zemp, L., Nguyen, J., Laborde, J. M., et al. (2021). Reconnaissance of Tumor Immune Microenvironment Spatial Heterogeneity in Metastatic Renal Cell Carcinoma and Correlation with Immunotherapy Response. *Clin. Exp. Immunol.* 204, 96–106. doi:10.1111/cei.13567
- Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSEA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *Bmc Bioinformatics* 14, 7. doi:10.1186/1471-2105-14-7
- Huang, Y.-J., Cao, Z.-F., Wang, J., Yang, J., Wei, Y.-J., Tang, Y.-C., et al. (2021). Why MUC16 Mutations lead to a Better Prognosis: A Study Based on the Cancer Genome Atlas Gastric Cancer Cohort. *Wjcc* 9, 4143–4158. doi:10.12998/wjcc.v9.i17.4143
- Jiang, X., Wang, J., Deng, X., Xiong, F., Zhang, S., Gong, Z., et al. (2020). The Role of Microenvironment in Tumor Angiogenesis. *J. Exp. Clin. Cancer Res.* 39, 204. doi:10.1186/s13046-020-01709-5
- Kennel, K. B., and Greten, F. R. (2021). Immune Cell - Produced ROS and Their Impact on Tumor Growth and Metastasis. *Redox Biol.* 42, 101891. doi:10.1016/j.redox.2021.101891
- Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., et al. (2012). VarScan 2: Somatic Mutation and Copy Number Alteration Discovery in Cancer by Exome Sequencing. *Genome Res.* 22, 568–576. doi:10.1101/gr.129684.111
- Lei, X., Lei, Y., Li, J.-K., Du, W.-X., Li, R.-G., Yang, J., et al. (2020). Immune Cells within the Tumor Microenvironment: Biological Functions and Roles in Cancer Immunotherapy. *Cancer Lett.* 470, 126–133. doi:10.1016/j.canlet.2019.11.009
- Leiserson, M. D., Wu, H. T., Vandin, F., and Raphael, B. J. (2015). CoMet: a Statistical Approach to Identify Combinations of Mutually Exclusive Alterations in Cancer. *Genome Biol.* 16, 160. doi:10.1186/s13059-015-0700-7
- Li, B., Ruotti, V., Stewart, R. M., Thomson, J. A., and Dewey, C. N. (2010). RNA-seq Gene Expression Estimation with Read Mapping Uncertainty. *Bioinformatics* 26, 493–500. doi:10.1093/bioinformatics/btp692
- Lou, Y., Xu, J., Zhang, Y., Lu, J., Chu, T., Zhang, X., et al. (2020). Chemotherapy Plus EGFR-TKI as First-Line Treatment Provides Better Survival for Advanced EGFR-Positive Lung Adenocarcinoma Patients: Updated Data and Exploratory *In Vitro* Study. *Targ Oncol.* 15, 175–184. doi:10.1007/s11523-020-00708-y
- Lou, Y., Xu, J., Zhang, Y., Zhang, W., Zhang, X., Gu, P., et al. (2021). Akt Kinase LANCL2 Functions as a Key Driver in EGFR-Mutant Lung Adenocarcinoma Tumorigenesis. *Cell Death Dis* 12, 170. doi:10.1038/s41419-021-03439-8
- Lu, J., Zhong, H., Chu, T., Zhang, X., Li, R., Sun, J., et al. (2019). Role of Anlotinib-Induced CCL2 Decrease in Anti-angiogenesis and Response Prediction for Non-small Cell Lung Cancer Therapy. *Eur. Respir. J.* 53. doi:10.1183/13993003.01562-2018
- Lu, J., Shi, Q., Zhang, L., Wu, J., Lou, Y., Qian, J., et al. (2019). Integrated Transcriptome Analysis Reveals KLK5 and L1CAM Predict Response to Anlotinib in NSCLC at 3rd Line. *Front. Oncol.* 9, 886. doi:10.3389/fonc.2019.00886
- Lu, J., Xu, W., Qian, J., Wang, S., Zhang, B., Zhang, L., et al. (2019). Transcriptome Profiling Analysis Reveals that CXCL2 Is Involved in Anlotinib Resistance in Human Lung Cancer Cells. *BMC Med. Genomics* 12, 38. doi:10.1186/s12920-019-0482-y
- Lu, J., Zhang, Y., Lou, Y., Yan, B., Zou, B., Hu, M., et al. (2021). ctDNA-Profilin-Based UBL Biological Process Mutation Status as a Predictor of Atezolizumab Response Among TP53-Negative NSCLC Patients. *Front. Genet.* 12, 723670. doi:10.3389/fgene.2021.723670
- Lu, J., Zhong, H., Wu, J., Chu, T., Zhang, L., Li, H., et al. (2019). Circulating DNA-Based Sequencing Guided Anlotinib Therapy in Non-Small Cell Lung Cancer. *Adv. Sci.* 6, 1900721. doi:10.1002/advs.201900721
- Lu, J., Zhong, R., Lou, Y., Hu, M., Yang, Z., Wang, Y., et al. (2021). TP53 Mutation Status and Biopsy Lesion Type Determine the Immunotherapeutic Stratification in Non-small-cell Lung Cancer. *Front. Immunol.* 12, 732125. doi:10.3389/fimmu.2021.732125
- Morris, T. J., Butcher, L. M., Feber, A., Teschendorff, A. E., Chakravarthy, A. R., Wojdacz, T. K., et al. (2014). ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics* 30, 428–430. doi:10.1093/bioinformatics/btt684
- Muz, B., de la Puente, P., Azab, F., and Azab, A. K. (2015). The Role of Hypoxia in Cancer Progression, Angiogenesis, Metastasis, and Resistance to Therapy. *Hp* 3, 83–92. doi:10.2147/hp.s93413
- Nicolls, M. R., and Voelkel, N. F. (2007). Hypoxia and the Lung: Beyond Hypoxic Vasoconstriction. *Antioxid. Redox Signaling* 9, 741–743. doi:10.1089/ars.2007.1574
- Pudela, C., Balyasny, S., and Applebaum, M. A. (2020). Nervous System: Embryonal Tumors: Neuroblastoma. *Atlas Genet. Cytogenet. Oncol. Haematol.* 24, 284–290. doi:10.4267/2042/70771
- Robertson, K. D. (2005). DNA Methylation and Human Disease. *Nat. Rev. Genet.* 6, 597–610. doi:10.1038/nrg1655
- Rocha, M. P., Fraire, A. E., Guntupalli, K. K., and Greenberg, S. D. (1994). Lung Cancer in the Young. *Cancer Detect. Prev.* 18, 349–355.
- Salem, A., Asselin, M. C., Reymen, B., Jackson, A., Lambin, P., West, C. M. L., et al. (2018). Targeting Hypoxia to Improve Non-small Cell Lung Cancer Outcome. *J. Natl. Cancer Inst.* 110. doi:10.1093/jnci/djx160
- Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer Statistics, 2019. *CA A. Cancer J. Clin.* 69, 7–34. doi:10.3322/caac.21551
- Tian, Y., Morris, T. J., Webster, A. P., Yang, Z., Beck, S., Feber, A., et al. (2017). ChAMP: Updated Methylation Analysis Pipeline for Illumina BeadChips. *Bioinformatics* 33, 3982–3984. doi:10.1093/bioinformatics/btx513
- Wilzén, A., Nilsson, S., Sjöberg, R. M., Kogner, P., Martinsson, T., and Abel, F. (2009). The Phox2 Pathway Is Differentially Expressed in Neuroblastoma Tumors, but No Mutations Were Found in the Candidate Tumor Suppressor Gene PHOX2A. *Int. J. Oncol.* 34, 697–705. doi:10.3892/ijo.00000196
- Wu, J., Li, L., Zhang, H., Zhao, Y., Zhang, H., Wu, S., et al. (2021). A Risk Model Developed Based on Tumor Microenvironment Predicts Overall Survival and Associates with Tumor Immunity of Patients with Lung Adenocarcinoma. *Oncogene* 40, 4413–4424. doi:10.1038/s41388-021-01853-y
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* 16, 284–287. doi:10.1089/omi.2011.0118
- Zandberg, D. P., Menk, A. V., Velez, M., Normolle, D., DePeaux, K., Liu, A., et al. (2021). Tumor Hypoxia Is Associated with Resistance to PD-1 Blockade in

Squamous Cell Carcinoma of the Head and Neck. *J. Immunother. Cancer* 9. doi:10.1136/jitc-2020-002088

Zhao, F., Xu, N., Yang, J., Li, B., Shi, J., Zheng, Y., et al. (2021). Identification of an Immune Subtype Predicting Survival Risk and Immune Activity in Hepatocellular Carcinoma. *Aging (Albany NY)* 13. doi:10.18632/aging.202953

Zhou, C., Wang, Y., Wang, Y., Lei, L., Ji, M.-H., Zhou, G., et al. (2021). Predicting Lung Adenocarcinoma Prognosis with a Novel Risk Scoring Based on Platelet-Related Gene Expression. *Aging* 13, 8706–8719. doi:10.18632/aging.202682

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

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

Copyright © 2022 Lou, Shi, Zhang, Qi, Zhang, Wang, Lu, Han and Zhong. 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.