

RESEARCH

Open Access



# Hypoxia within tumor microenvironment characterizes distinct genomic patterns and aids molecular subtyping for guiding individualized immunotherapy

Run Shi<sup>1</sup>, Jing Sun<sup>2,3</sup>, Hanyu Zhou<sup>1</sup>, Tong Hu<sup>1</sup>, Zhaojia Gao<sup>4</sup>, Xin Wang<sup>5</sup>, Minglun Li<sup>6,7</sup>, Zhaokai Zhou<sup>8</sup> and Yongqian Shu<sup>1,9\*</sup>

<sup>1</sup>Run Shi, Jing Sun, Hanyu Zhou and Tong Hu contributed equally to this work.

\*Correspondence: Yongqian Shu  
shuyongqian@csc.org.cn

Full list of author information is available at the end of the article

## Abstract

Assessing the hypoxic status within the tumor microenvironment (TME) is crucial for its significant clinical relevance in evaluating drug resistance and tailoring individualized strategies. In this study, we proposed a robust pan-cancer hypoxic quantification method utilizing multiple public databases, diverse bioinformatics, and statistical methods. All tumor samples were classified into four subtypes: non-hypoxic/TME<sup>high</sup> (C1), hypoxic/TME<sup>high</sup> (C2), non-hypoxic/TME<sup>low</sup> (C3), and hypoxic/TME<sup>low</sup> (C4). We systematically analyzed multi-omics data and single-cell RNA-sequencing (scRNA-seq) data to reveal distinct immune landscape patterns and genomic characteristics among the four subtypes across pan-cancer. Furthermore, we employed multiple machine learning approaches to construct a hypoxic-TME model to enhance the predictive accuracy of immunotherapy response. Additionally, drug repositioning was implemented for cancer patients predicted as non-responders to immunotherapy. A pan-cancer analysis identified PDK1 as a hub gene linking tumor hypoxia, glycolysis, and immunotherapy resistance. In vivo experimental validation further confirmed that targeting PDK1 could improve the response to immunotherapy. Overall, our study may offer valuable insights for integrating hypoxic-TME classification into tumor staging and providing personalized strategies for cancer patients.

**Keywords** Hypoxia, Tumor microenvironment, Molecular subtyping, Multi-omics, Precision medicine, Immunotherapy

## Introduction

Tumor hypoxia, defined by an insufficient oxygen supply within the tumor microenvironment (TME), results directly from the rapid proliferation of tumors that exceeds the rate of new blood vessel formation [1]. Recognized as a critical characteristic of numerous solid tumors, hypoxia triggers an array of molecular and cellular changes that remodel the TME [2]. This remodeling confers upon cancer cells advanced survival

mechanisms, fostering the development of increasingly aggressive tumor phenotypes [3, 4]. Consequences of tumor hypoxia include impaired immune response, metabolic reprogramming, activation of cancer stem cells, enhanced tumor vascularization, increased invasion and metastasis, genomic instability, and a reduction in cell proliferation [5, 6]. Therefore, evaluating the hypoxic status of a tumor is of great importance, as it holds significant clinical relevance for assessing drug resistance and tailoring individualized treatment strategies.

However, we have to admit that defining hypoxia status in tumors is challenging due to the variation in oxygen levels across different types of tissues. Numerous diagnostic methods for tumor hypoxia have been developed, ranging from direct techniques such as oxygen electrodes and phosphorescence quenching, to physiological approaches including photoacoustic tomography and near-infrared spectroscopy/tomography, as well as the identification of hypoxia-derived gene signatures and markers [7–9]. However, these methods are not readily scalable for application to a large volume of patient samples, such as those from pan-cancer studies.

On the other hand, TME constitutes a multifaceted network comprising immune cells, stromal cells, extracellular matrix components, and signaling molecules [10, 11]. The interactions between these cellular entities within TME are critical determinants of tumor progression and therapy efficacy [12]. Two key issues merit further investigation: firstly, the development of a standardized method for quantifying hypoxia levels across diverse tumor types is imperative; and secondly, the implementation of tumor hypoxia status within TME attributes into the molecular classification of cancers could greatly improve the precision of treatments. An integrated approach to incorporate tumor hypoxia assessments and TME characteristics into molecular classifications is anticipated to refine existing models, more accurately capturing tumor behavior and its response to therapeutic interventions. Such enriched molecular classifications, which include hypoxia and TME components, may differentiate tumor subpopulations with similar genetic profiles but distinct physiological responses. Considering that hypoxic cells typically exhibit resistance to cancer therapy due to adaptive survival mechanisms under low-oxygen conditions, a comprehensive classification framework encompassing hypoxic factors and TME components will better aid clinicians to tailor treatment choices, stratify patients by risk, and identify novel immunotherapeutic targets [13].

In this study, to address the above issues, we developed a robust pan-cancer hypoxic quantification method utilizing multiple public databases and various bioinformatics and statistical methods. We classified all tumor samples into four subtypes according to their hypoxic and TME score: non-hypoxic/TME<sup>high</sup> (C1), hypoxic/TME<sup>high</sup> (C2), non-hypoxic/TME<sup>low</sup> (C3), hypoxic/TME<sup>low</sup> (C4). Upon further analysis of these four subtypes, we observed varying levels of immune infiltration and distinct gene mutation patterns among the four subtypes. Subsequently, we conducted in-depth analysis of the multi-omics data and single-cell RNA-sequencing (scRNA-seq) data to uncover the genomic variability and the distinct features of the immune landscape among the four subtypes of lung adenocarcinoma (LUAD) and stomach adenocarcinoma (STAD). In addition, we employed multiple machine learning approaches to construct a hypoxic-TME model predicting the accuracy of immunotherapy response and performed individualized target-drug analysis and drug repositioning for STAD patients who are predicted as non-responders to immunotherapy. Ultimately, a pan-cancer analysis identified PDK1

as a hub gene that connects tumor hypoxia, glycolysis, and resistance to immunotherapy. Experimental validation *in vivo* confirmed that targeting PDK1 can enhance the response to tumor immunotherapy. In conclusion, our objective was to systematically develop a hypoxic-TME classifier for precise classification and therapeutic response prediction through the integration of hypoxia and TME. Our study may contribute to a better understanding of tumor-specific biology based on an integrated assessment of hypoxic TME, presenting meaningful implications for clinical practice.

## Materials and methods

### Data collection and processing

The transcriptomic profiles of various cell lines or sample fragments under hypoxic and normoxic conditions were retrieved from the Gene Expression Omnibus (GEO) database. Comprehensive multi-omics data, encompassing transcriptomic, copy number variation (CNV), and somatic mutation information for the TCGA pan-cancer cohort, which includes 32 solid tumor types and associated sample metadata, were acquired from the TCGA portal (<https://portal.gdc.cancer.gov/>). For standardized comparison of RNA-seq data, we converted fragments per kilobase of exon per million mapped reads (FPKM) to transcripts per million (TPM) values. CNV data were analyzed employing the GISTIC 2.0 algorithm, while the “maftools” R package was used to analyze the mutation data, and the “Sigminer” package was utilized to extract mutational signatures from whole exome sequencing (WES) data [14].

To develop an integrated hypoxic-TME scoring model for predicting the response to immunotherapy, we incorporated several cohorts and validated the model’s predictive performance. For training, we used a STAD-ICB cohort (cohort id: PRJEB25780; sample size: 78) [15]. A mUC-ICB cohort (cohort id: GSE176307; sample size: 89) served as the first validation cohort [16], and a total of 405 melanoma patients from five cohorts (cohort ids: GSE78220 [17], GSE91061 [18], Nathanson-2017 [19], phs000452 [20], and PRJEB23709 [21]) who underwent ICB therapy composed the secondary validation cohort. The function “Combat” in R package “sva” was used to eliminate batch effects if multiple independent datasets were merged for further analysis.

### Single-cell RNA-sequencing (scRNA-seq) analysis

We analyzed the association between hypoxic TME and a variety of clinical characteristics or specimen types using three scRNA-seq datasets for primary lung adenocarcinoma (LUAD) (GSE171145 [22]), metastatic LUAD (GSE143423, not published), and stomach adenocarcinoma (STAD) (GSE167297 [23]) with the “Seurat” package. The detailed workflow of scRNA-seq analysis was described in a previously published article [24]. The “CellChat” package was applied to characterize intercellular communications across diverse cell types. Furthermore, we conducted a single-cell trajectory analysis using the “SCP” R package, employing its default functions. This analysis ordered cells along a developmental trajectory, which was then partitioned into varying branch points to reflect cellular lineage divergence.

### Estimation of immune infiltration and evaluation of immune features

The ESTIMATE algorithm was applied to estimate the immune, stromal, and TME scores for tumor samples [25], and the Cibersort was used to quantify the relative

proportions of 22 tumor-infiltrating immune cell types based on transcriptome profiling data [26]. Data of clonal and sub-clonal fraction of TCGA samples were obtained from The Cancer Immunome Atlas (TCIA) [27]. The aneuploidy score, FGA, HRD score, CTA score, TCR richness, BCR richness, non-silent mutation, SNV neoantigen, indel neoantigen were obtained from a pan-cancer immune landscape study [28]. The cytolytic activity score (CYT) was defined as the geometric mean of GZMA and PRF1 expression [29].

#### **Selection of an optimal model from multiple machine learning approaches**

In the STAD-ICB cohort, we utilized the expression profiles of TME (immune, stromal, and epithelial components), ECM, and hypoxic markers to explore a range of 12 machine learning algorithms with diverse parameters. These algorithms included RF, GBM, plsRglm, NaiveBayes, Enet (alpha ranging from 0.1 to 0.9), glmBoost, Lasso, Step-glm (including forward, backward, and bidirectional steps), LDA, XGBoost, and SVM. A 10-fold cross-validation strategy was implemented to identify the most effective model with the highest AUC for predicting ICB response. Additionally, we compared the predictive accuracy of the machine learning model, determined by the AUC value, with that of documented predictors such as IFNG signature score, PD-1 and PD-L1 expression, T-GEP, TIDE score, CD8 signature, CYT score, and TLS signature in a testing mUC-ICB cohort.

#### **Analyses of cancer cell line data**

The CERES scores, derived from genome-scale CRISPR knockout screens of 18,333 genes across cancer cell lines, were obtained from the Dependency Map (DepMap) portal (<https://depmap.org/portal/>). The CERES score quantifies the extent to which a particular cancer cell line (CCL) is reliant on a gene of interest, with lower scores indicate greater essentiality of the gene for cell growth and survival in that CCL [30]. Gene expression data and their corresponding drug IC50 values for CCLs were collected from the Cancer Therapeutics Response Portal (CTRP v.2.0). The CTRP 2.0 contains the sensitivity data for 481 compounds in 835 CCLs [31]. The area under the dose-response curve (AUC) value was used as a measure of drug sensitivity, and a higher AUC value indicates insensitivity to the drug. The inbuilt function “calcPhenotype” of the “pRRophetic” package was used to predict the sensitivity of each sample to different compounds [32].

#### **Animal experiments**

In the animal study,  $2 \times 10^6$  Lewis lung carcinoma (LLC) cells were subcutaneously injected into the right flank of five-week-old C57BL/6 mice to establish a subcutaneous tumor model. Upon reaching a tumor volume of approximately 80–100 mm<sup>3</sup>, mice with tumors were randomly assigned to one of four groups ( $n=5$  per group). Treatments included administration of a mouse anti-PD-1 monoclonal antibody (mAb) (BioXcell, BE0146, 100 µg per mouse), an IgG isotype control (BioXcell, BE0089, 100 µg per mouse), a PDK1 inhibitor (MCE, GSK2334470, 100 mg/kg), or a combination of the PDK1 inhibitor and the anti-PD-1 mAb. All treatments were delivered via intraperitoneal injection. Body weight and tumor dimensions were recorded every two to three days. Tumor volumes were calculated using the formula:  $0.5 \times (\text{length} \times \text{width}^2)$ . All



experimental protocols involving animals were conducted in accordance with guidelines and approved by the Animal Experiment Ethical Committee of Nanjing Medical University.

#### **Flow cytometry analysis**

C57BL/6 mice were sacrificed to obtain fresh tumor specimens, which were then enzymatically digested for 2 h, and then strained through 70- $\mu\text{m}$  filters to procure single-cell suspensions. After washing, cells were re-suspended in 40% Percoll and centrifuged over a 70% Percoll gradient to isolate lymphocytes. Subsequently, cells were stimulated for 4 h using a leukocyte activation cocktail (BD Pharmingen) to prepare for intracellular cytokine staining. Anti-Cd45, anti-Cd3e, and anti-Cd8a monoclonal antibodies were purchased from BD Pharmingen, and intracellular granzyme B (Gzmb) was stained using an anti-Gzmb antibody from BioLegend. Analysis of the stained cells was conducted using a BD FACS Canto II flow cytometer, and the data were processed with FlowJo software.

#### **Additional statistical and bioinformatics analyses**

The “limma” package was used to identify differentially expressed genes (DEGs) between different groups. The weighted gene co-expression network analysis (WGCNA) was employed to construct a scale-free co-expression network using the “WGCNA” package and to identify a gene module which is mostly correlated with hypoxic status [33]. A single-sample gene set enrichment analysis (ssGSEA) algorithm was applied to calculate the enrichment scores of specific cancer hallmarks (or features, pathways etc.) for each sample or cell based on the scRNA-seq transcriptome data and corresponding gene signatures using the R package “gsva”. Multivariate logistical regression analysis was performed to evaluate the adjusted odds ratio (OR) of each variable for immunotherapy response. By integrating the RF model with clinicopathological variables, a decision-making tree was constructed to improve the predictive accuracy for immunotherapy response. The concordance index (C-index) or receiver operating characteristic (ROC) analysis was used to test the discriminative capacity. Pearson or Spearman correlation analysis was performed to evaluate the correlation between two variables. Student’s t-test or Mann-Whitney U-test was used to analyze the difference between two groups. Two-sided p value less than 0.05 was considered statistically significant. All analyses were performed in the GraphPad Prism 8.0 and R 4.1.0 software.

## **Results**

### **Development of a gene expression signature-based hypoxic scoring classifier for pan-cancer**

Initially, we utilized the GSE30979 dataset for training in order to develop a gene expression signature for quantifying hypoxic status. This dataset comprises transcriptome data from 20 non-small cell lung cancer (NSCLC) fragments that were ex vivo cultured under hypoxic or normoxic conditions for three days. Upon conducting PCA analysis, significant batch effects were observed across samples from distinct patients (Fig. S1A), and we employed the “sva” package to remove batch effects (Fig. S1B). Applying an optimal soft threshold ( $\beta=5$ ) (Fig. S2), we utilized the WGCNA algorithm to construct a scale-free co-expression network (Fig. S3A). Among the identified gene modules, the blue module exhibited the most significant correlation with hypoxic status ( $r=0.98$ ,  $p=1e-14$ ;

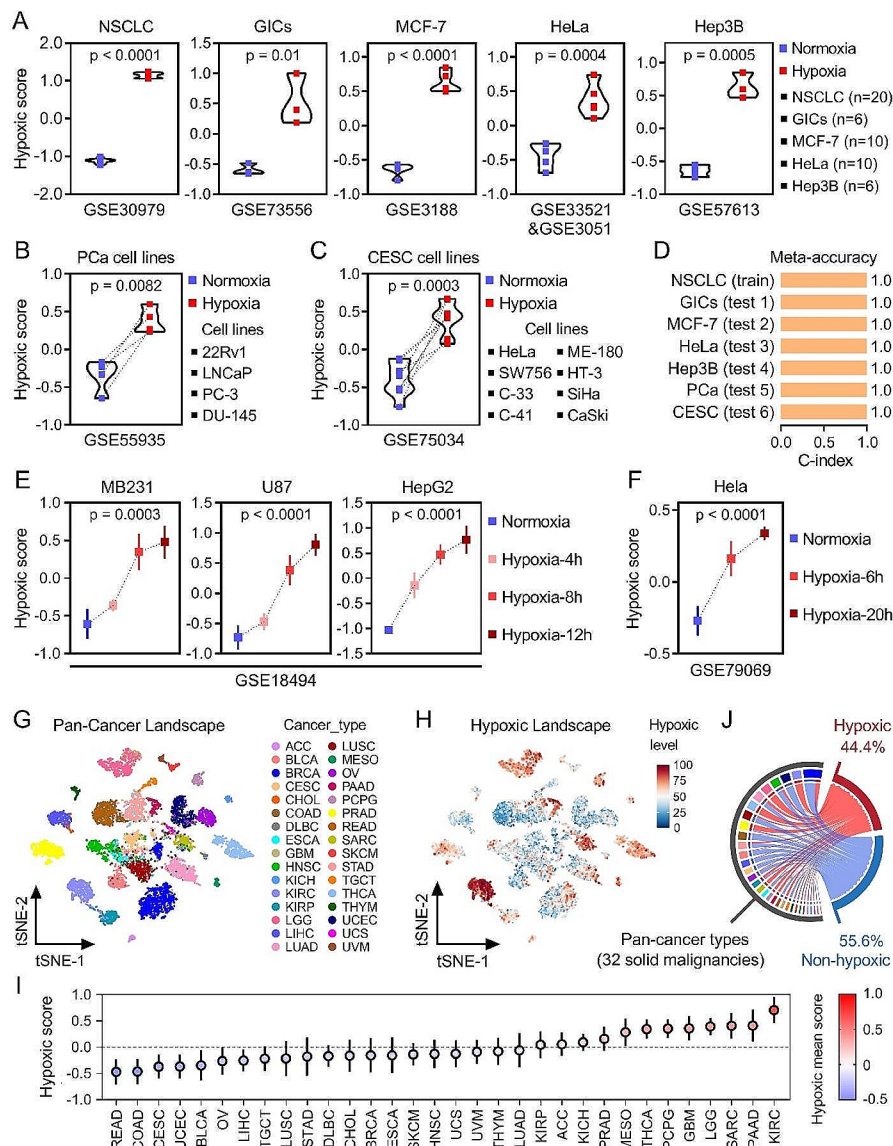
Fig. S3B). Subsequently, the blue module was considered as a hypoxic module. The relationship between module membership (MM) and gene significance (GS) is depicted in Fig. S3C. Following a filtering threshold of  $MM > 0.9$  and  $GS > 0.9$ , we identified 279 genes as the hub genes associated with hypoxia (details in supplementary Table 1). Reactome enrichment analysis showed that the 279 genes are mainly enriched in “Cell cycle” (Fig. S3D). As expected, we observed all 279 genes were involved in the differentially expressed genes (DEGs) associated with hypoxia and normoxia (Fig. S3E). Additionally, a pie chart depicted the distribution of upregulated genes under hypoxic or normoxic conditions (Fig. S3F). Subsequently, a formula was established to quantify the hypoxic status for each sample, calculated as the difference between the ssGSEA score of the hypoxic signature (118 genes) and normoxic signature (161 genes) using the “gsva” algorithm (Fig. S3G).

Upon applying the hypoxic formula to multiple datasets, we observed a significant elevation in hypoxic scores in various cancer cells cultured under hypoxic conditions, including glioma-initiating cells (GICs), MCF-7, HeLa, and Hep3B cells (Fig. 1A). Additionally, distinctly high hypoxic scores were noted in four prostate cancer cell lines (PCa, Fig. 1B) and eight cervical squamous cancer cell lines (CESC, Fig. 1C) under hypoxic culture. A summary of the C-index of the hypoxic formula in all the aforementioned datasets was depicted in Fig. 1D, demonstrating 100% accuracy when the cut-off value was set at 0. That is to say, a hypoxic score greater than 0 indicates hypoxic status, while a score lower than 0 indicates normoxic status. In addition, the hypoxic score gradually increases from normoxic to hypoxic status with extended culture time in MB231, U87, and HepG2 cells in GSE18494 (Fig. 1E), and HeLa cells in GSE79069 (Fig. 1F).

The tSNE reduction plot displayed over 10,000 samples representing 32 solid cancer types from the TCGA-PanCancer cohort (Fig. 1G), with the hypoxic scores mapped onto the corresponding samples using varying color gradients to indicate levels of hypoxia (Fig. 1H). The distribution of hypoxic scores for each cancer type is illustrated in Fig. 1I. Notably, READ and COAD exhibited the lowest hypoxic levels, while KIRC displayed the highest hypoxic scores among all 32 cancer types (Fig. 1I). With the cut-off value of 0, all samples were categorized into hypoxic and non-hypoxic class, and a chord diagram shows the affiliation between all tumor samples and the hypoxic and non-hypoxic groups (Fig. 1J). The hypoxic group accounts for 44.4% of all solid tumors, with the remaining being the non-hypoxic group.

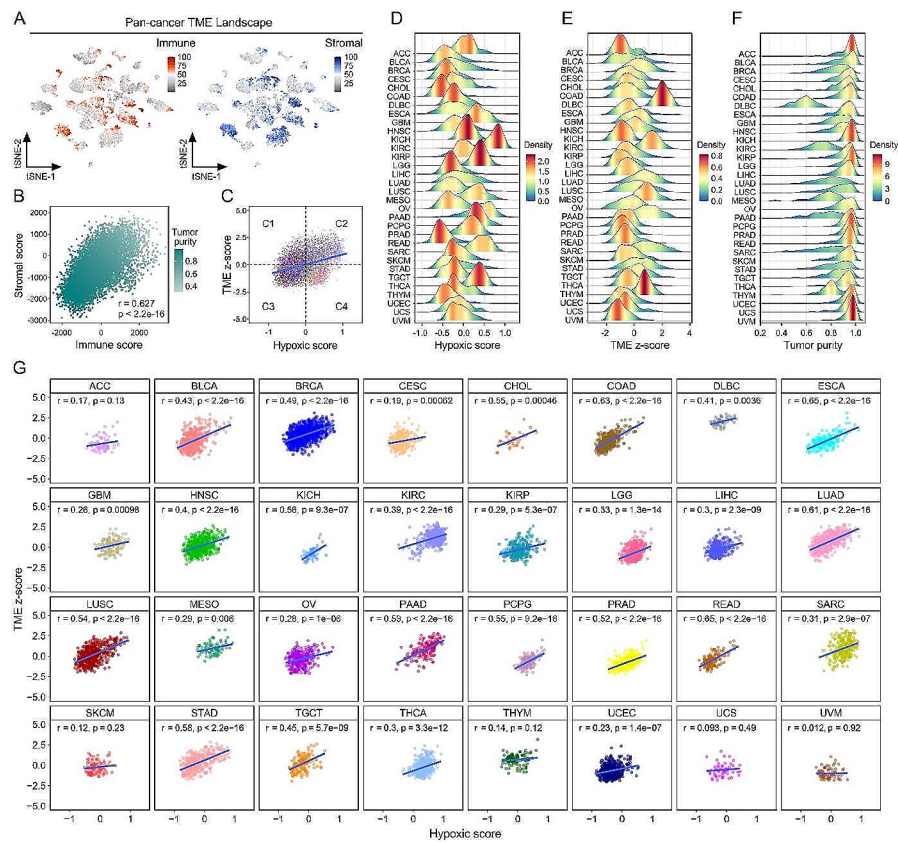
### **The hypoxia score, in conjunction with the TME score, categorizes all samples into four subtypes**

Furthermore, using the ESTIMATE algorithm, we computed the immune, stromal, and tumor purity for each sample and mapped the immune and stromal scores on the tSNE plot to illustrate the TME landscape of all the solid cancer samples (Fig. 2A). As depicted in Fig. 2B, the immune score exhibited a significantly positive correlation with the stromal score but a negative correlation with tumor purity in the pan-cancer cohort. Additionally, as shown in Fig. 2C, based on the medians of hypoxic and TME scores, all the samples were classified into four subtypes: non-hypoxic/TME<sup>high</sup> (C1), hypoxic/TME<sup>high</sup> (C2), non-hypoxic/TME<sup>low</sup> (C3), hypoxic/TME<sup>low</sup> (C4). In detail, ridge plots depicted the distributions of hypoxic scores (Fig. 2D), TME z-scores (Fig. 2E), and tumor purity (Fig. 2F) for each cancer type. Furthermore, to explore the relationship between hypoxic



**Fig. 1** Validation of the hypoxic scoring classifier in pan-cancer. **(A)** Significant elevations of hypoxic scores were observed in various cancer cells cultured under hypoxic conditions, including glioma-initiating cells (GICs), MCF-7, HeLa, and Hep3B cells. **(B)** High hypoxic scores were observed in four prostate cancer cell lines under hypoxic culture. **(C)** High hypoxic scores were observed in eight cervical squamous cancer cell lines under hypoxic culture. **(D)** A summary of the 100% C-indexes of the hypoxic formula in all the datasets. **(E)** The hypoxic score gradually increases from normoxic to hypoxic status with extended culture time in MB231, U87, and HepG2 cells in GSE18494, **(F)** and similar results were observed in the HeLa cells in GSE79069. **(G)** A tSNE reduction plot displayed over 10,000 samples representing 32 solid cancer types from the TCGA-PanCancer cohort. **(H)** The hypoxic scores are mapped onto the corresponding samples using varying color gradients to indicate levels of hypoxia. **(I)** The distribution of hypoxic scores for each cancer type. **(J)** A chord diagram shows the affiliation between all tumor samples and the hypoxic and non-hypoxic groups. The hypoxic group accounts for 44.4% of all solid tumors, with the remaining (55.6%) being the non-hypoxic group

scores and TME scores in pan-cancer, we illustrated the correlation in the 32 solid cancer types, finding positive correlations between hypoxic and TME scores in most solid cancer types, such as LUAD and STAD, while no significant correlation was observed in a minority of cancer types such as ACC and SKCM (Fig. 2G).

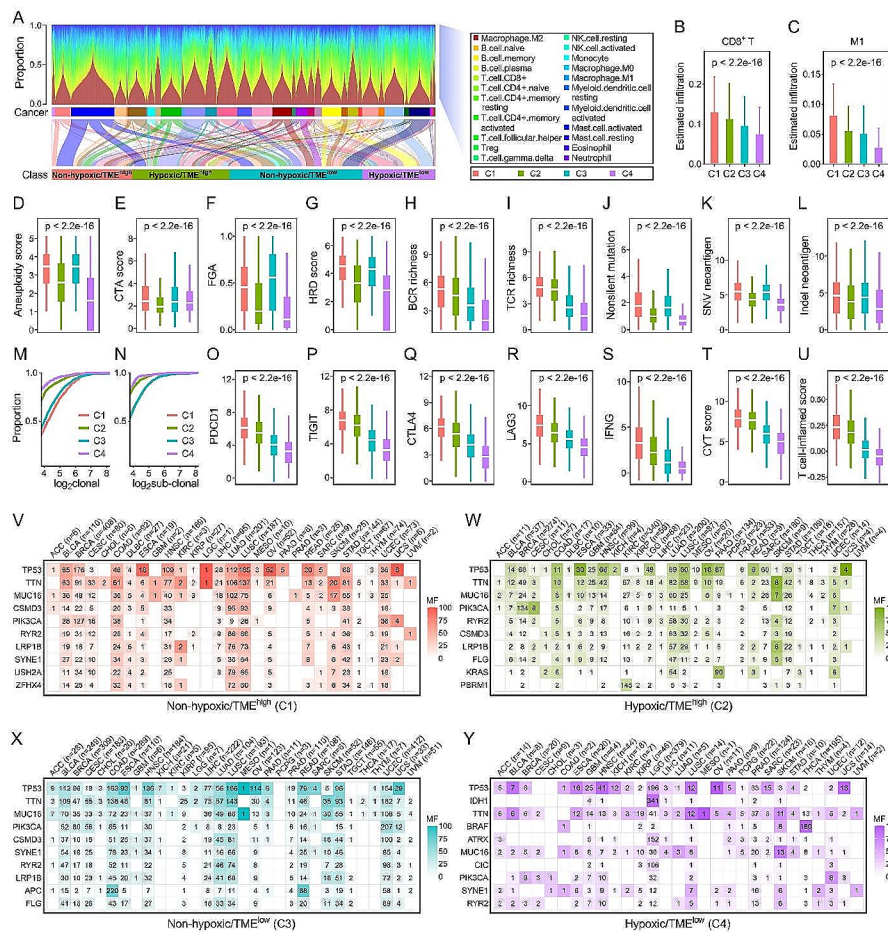


**Fig. 2** Pan-cancer samples are classified into four subtypes based on the combination of the hypoxic score and TME score. **(A)** The immune and stromal scores were calculated and mapped on the tSNE plot to illustrate the TME landscape of all the solid cancer samples. **(B)** The immune score exhibited a significantly positive correlation with the stromal score but a negative correlation with tumor purity in the pan-cancer cohort. **(C)** Based on the medians of hypoxic and TME scores, all the samples were classified into four subtypes: non-hypoxic/TME<sup>high</sup> (C1), hypoxic/TME<sup>high</sup> (C2), non-hypoxic/TME<sup>low</sup> (C3), hypoxic/TME<sup>low</sup> (C4). **(D - F)** The ridge plots depicted the distributions of hypoxic scores, TME z-scores, and tumor purity for each cancer type. **(G)** The correlation between the hypoxic score and TME score in each solid cancer type

### Distinct immune landscapes and mutational patterns were observed among the four hypoxic-TME subtypes

The Cibersort algorithm was utilized to estimate the abundance of 22 immune cell types in each tumor sample. A combination of a stacked bar plot and Sankey plot was used to depict the distribution and affiliation of the 22 immune cell types to the four hypoxic-TME subtypes (Fig. 3A). It was observed that the infiltrating abundance of CD8<sup>+</sup> T and M1 cells gradually decreases from C1 to C4 subtype (Fig. 3B and C). Additionally, several immune characteristics, including aneuploidy score (Fig. 3D), CTA score (Fig. 3E), FGA (Fig. 3F), HRD score (Fig. 3G), BCR richness (Fig. 3H), TCR richness (Fig. 3I), non-silent mutation (Fig. 3J), SNV neoantigen (Fig. 3K), and indel neoantigen (Fig. 3L), were compared among the subtypes, revealing that C4 samples exhibited the lowest levels in the majority of these parameters. Moreover, distinct clonal and subclonal counts were observed among the subtypes (Fig. 3M and N). The comparison of the four main immune checkpoints (PDCD1, TIGIT, CTLA4, and LAG3; Fig. 3O – 3R), along with IFNG (Fig. 3S) among the four subtypes, revealed a gradual decrease from C1 to C4 with significant differences. Additionally, the CYT score (Fig. 3T) and the T cell-inflamed score (Fig. 3U) demonstrated similar trends.



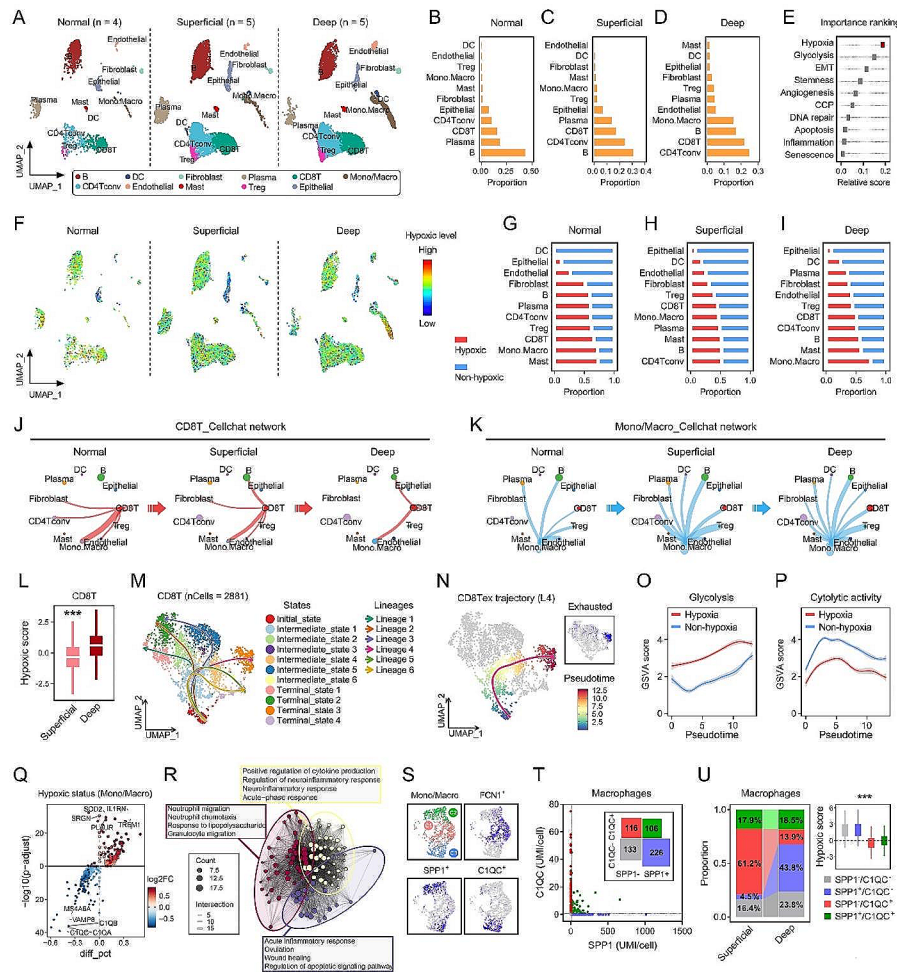


**Fig. 3** Distinct immune landscapes and mutational patterns were observed among the four hypoxic-TME subtypes. (A) A combination of a stacked bar plot and Sankey plot was used to depict the distribution and affiliation of the 22 immune cell types to the four hypoxic-TME subtypes in pan-cancer. (B & C) The infiltrating abundance of CD8<sup>+</sup> T and M1 cells gradually decreases from C1 to C4 subtype. (D - L) Several immune characteristics, including aneuploidy score, CTA score, FGA, HRD score, BCR richness, TCR richness, non-silent mutation, SNV neoantigen, and indel neoantigen were compared among the subtypes, revealing that C4 samples exhibited the lowest levels in the majority of these parameters. (M & N) Distinct clonal and subclonal counts were observed among the subtypes. (O - S) The comparison of the four main immune checkpoints (PDCD1, TIGIT, CTLA4, and LAG3), along with IFNG among the four subtypes, revealed a gradual decrease from C1 to C4 with significant differences. (T - U) The CYT score and the T cell-inflamed score across the four subtypes demonstrated similar trends. (V - Y) The mutational patterns of the four hypoxic-TME subtypes

The mutational patterns of the four subtypes were further investigated, consistently revealing TP53 as the most frequently mutated gene among the subtypes (Fig. 3V – 3Y). Of significance, IDH1 mutation ranks second in the C4 subtype (Fig. 3Y), resulting from the major occupation of LGG in C4 and suggesting C4 as a “cold” subtype with the lowest tumor immunogenicity. More details of the frequently mutated genes in each subtype could be found in Fig. S4A - S4D. These findings indicate that the distinct immune landscapes and mutational patterns observed among the four hypoxic-TME subtypes could be pivotal in reflecting diverse responses to immunotherapy.

### STAD progression is accompanied by tumor hypoxia with an expansion of SPP1<sup>+</sup>/C1QC<sup>-</sup> macrophages

We further investigated the correlation between the hypoxic TME and tumor progression using a STAD scRNA-seq dataset, comprising 4 normal samples, 5 superficial tumor samples, and 5 deep tumor samples (Fig. 4A). The cell types were annotated in



**Fig. 4** STAD progression is accompanied by tumor hypoxia with an expansion of SPP1<sup>+</sup>/C1QC<sup>-</sup> macrophages. **(A)** A STAD scRNA-seq dataset comprising 4 normal samples, 5 superficial tumor samples, and 5 deep tumor samples was used for further investigation. **(B - D)** The relative proportions of various cell types involved in STAD TME were visualized in bar plots for each sample category. **(E)** The random forest algorithm identified hypoxia as the most significant characteristic in STAD progression among various cancer hallmarks. **(F)** The hypoxic score was mapped to each cell using color gradients to indicate levels of hypoxia. **(G - I)** The proportion of hypoxic and non-hypoxic for each cell type was illustrated in stacked bar plots within each sample category. **(J - K)** The cellular interactions of CD8<sup>+</sup> T cells and mono/macrophages with other cells were depicted, demonstrating a gradual attenuation of CD8<sup>+</sup> T cell communications from normal to superficial to deep tumors. **(L)** A significant increase of hypoxia score was observed in the deep tumors compared to the superficial tumors within the CD8<sup>+</sup> T cell population. **(M)** A total of 2,881 CD8<sup>+</sup> T cells were utilized for a single-cell trajectory analysis. **(N)** Among six main developmental lineages, lineage 4 represents the Tex developmental phenotype. **(O & P)** CD8<sup>+</sup> T cells in the hypoxic status unveiled significantly higher glycolytic activity and lower cytolysis activity compared to non-hypoxic cells. **(Q)** Mono/macrophages were extracted for further study, and the differentially expressed genes (DEGs) correlated with hypoxic status were displayed in a volcano plot. **(R)** All the DEGs were enriched in three terms, mainly labeled as “cytokine production”, “neutrophil migration”, and “acute inflammatory response”. **(S)** Upon higher resolution, all mono/macrophages were segregated into three subclusters. **(T)** All macrophages were categorized into four subclasses: SPP1<sup>-</sup>/C1QC<sup>-</sup>, SPP1<sup>+</sup>/C1QC<sup>-</sup>, SPP1<sup>-</sup>/C1QC<sup>+</sup>, and SPP1<sup>+</sup>/C1QC<sup>+</sup>. **(U)** A substantial expansion of SPP1<sup>+</sup>/C1QC<sup>-</sup> macrophages was observed in deep cancers compared to superficial ones. \*\*\*  $p < 0.001$

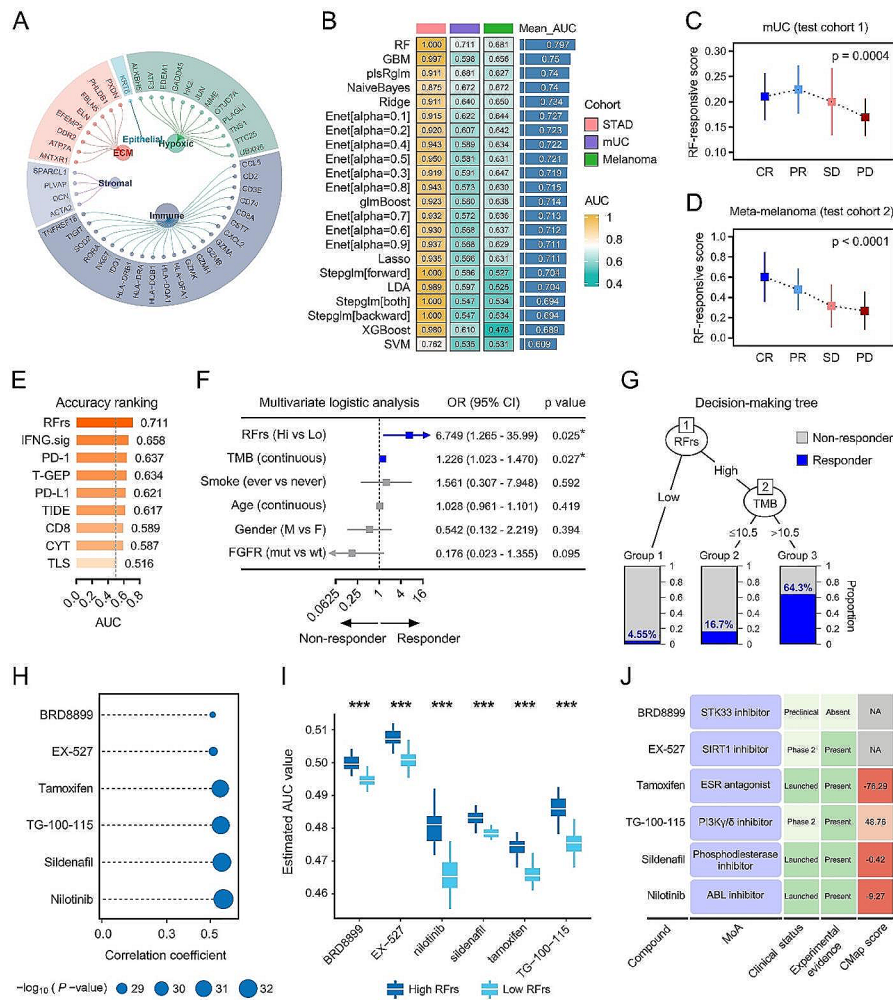


different sample categories within UMAP plots, and the relative proportions of these cells were visualized in bar plots for each sample category (Fig. 4B–4D). The random forest algorithm identified hypoxia as the most significant characteristic in STAD progression among various cancer hallmarks (Fig. 4E). Furthermore, the ssGSEA algorithm was utilized to calculate the hypoxic score, which was then mapped to each cell using color gradients to indicate levels of hypoxia (Fig. 4F). The proportion of hypoxic and non-hypoxic for each cell type was illustrated in stacked bar plots within each sample category, and we observed that the highest hypoxic levels in mono/macrophages in the deep tumor samples (Fig. 4G–4I). Additionally, the cellular interactions of CD8<sup>+</sup> T cells and mono/macrophages with other cells were depicted, demonstrating a gradual attenuation of CD8<sup>+</sup> T cell communications from normal to superficial to deep tumors (Fig. 4J), while the communications of mono/macrophages were progressively strengthened (Fig. 4K). A significant increase of hypoxia score was observed in the deep tumors compared to the superficial tumors within the CD8<sup>+</sup> T cell population (Fig. 4L). Subsequently, a total of 2,881 CD8<sup>+</sup> T cells were utilized for a single-cell trajectory analysis, which revealed six main developmental lineages (Fig. 4M), with lineage 4 representing the developmental trajectory showcasing the highest ssGSEA scores of the exhausted (Tex signature) phenotype (Fig. 4N). Further analysis of glycolytic and cytolytic activity in CD8<sup>+</sup> T cells with hypoxic and non-hypoxic status unveiled significantly higher glycolytic activity (Fig. 4O) and lower cytolytic activity (Fig. 4P) in hypoxic CD8<sup>+</sup> T cells compared to non-hypoxic cells.

Mono/macrophages were extracted for further study, and the differentially expressed genes (DEGs) correlated with hypoxic status were displayed in a volcano plot (Fig. 4Q). All the DEGs were enriched in three terms, mainly labeled as “cytokine production”, “neutrophil migration”, and “acute inflammatory response”, reflecting the alterations in anti-tumor-related immune responses (Fig. 4R). Upon higher resolution, all mono/macrophages were segregated into three subclusters. C1 exhibited a high expression of FCN, indicative of the monocyte population, while C2 had a mixed expression of SPP1 and C1QC, suggesting the macrophage population (Fig. 4S). Furthermore, all macrophages were categorized into four subclasses: SPP1<sup>-</sup>/C1QC<sup>-</sup>, SPP1<sup>+</sup>/C1QC<sup>-</sup>, SPP1<sup>-</sup>/C1QC<sup>+</sup>, and SPP1<sup>+</sup>/C1QC<sup>+</sup>, as shown in Fig. 4T. Notably, there was a substantial expansion of SPP1<sup>+</sup>/C1QC<sup>-</sup> macrophages in deep cancers compared to superficial ones (Fig. 4U), and the SPP1<sup>+</sup>/C1QC<sup>-</sup> tumor-associated macrophages (TAM) has been reported to promote cancer metastasis and therapeutic resistance [34].

### **Selection of a machine learning model based on hypoxic-TME signatures for predicting immunotherapy response**

Considering the close relationship between immunotherapy response and hypoxic TME, we integrated a set of TME markers and the hypoxic gene signature to develop an optimal model for predicting immunotherapy response. The expression profiles of 1 epithelial marker, 4 stromal markers, 22 immune markers, 12 hypoxic markers, and 8 ECM markers were input into 12 machine learning algorithms with diverse parameters to select the optimal model (Fig. 5A). The random forest (RF) model demonstrated the highest mean value of AUCs (mean AUC=0.797) in the training and two validation cohorts (STAD, mUC, and meta-melanoma ICB-therapy cohorts) compared to other models (Fig. 5B). Furthermore, the RF model-responsive score (RFrs) progressively



**Fig. 5** Development of a machine learning model based on hypoxic-TME signatures for predicting immunotherapy response. **(A)** The expression profiles of 1 epithelial marker, 4 stromal markers, 22 immune markers, 12 hypoxic markers, and 8 ECM markers were input into 12 machine learning algorithms with diverse parameters to select the optimal model. **(B)** The random forest (RF) model demonstrated the highest mean value of AUCs (mean AUC = 0.797) in the training and two validation cohorts (STAD, mUC, and meta-melanoma ICB-therapy cohorts) compared to other models. **(C & D)** The RF model-responsive score (RFrs) progressively decreased from CR/PR to SD to PD patients in the mUC and meta-melanoma cohorts who received ICB therapy. **(E)** RFrs exhibited the highest AUC of 0.711 compared to other conventional predictors in the mUC-ICB cohort. **(F)** Multivariate logistic regression analysis in the mUC-ICB cohort identified RFrs and tumor mutational burden (TMB) as the two independent significant variables influencing ICB response. **(G)** An integrated decision-making tree was constructed based on RFrs and TMB, and the mUC-ICB cohort was categorized into three groups, each with significantly different proportions of responders (4.55% responder in group 1, 16.7% in group 2, and 64.3% in group 3). **(H)** The CTRP database was used to assess the AUC values of 481 compounds for each STAD sample, and the Pearson correlation coefficients were calculated between all the AUCs and the RFrs in the STAD-ICB cohort. Six compounds with a correlation > 0.5 (nilotinib, sildenafil, TG-100-115, tamoxifen, EX-527, and BRD8899) were shown. **(I)** The estimated AUCs of all the six compounds were notably lower in the low RFrs group. **(J)** The MoAs, clinical status, experimental evidence, and cMap scores of the six candidate compounds was provided for the six compounds. \*\*\*  $p < 0.001$

decreased from CR/PR to SD to PD patients in the mUC and meta-melanoma cohorts who received ICB therapy (Fig. 5C and D). Additionally, RFrs exhibited the highest AUC of 0.711 compared to other conventional predictors in the mUC-ICB cohort (Fig. 5E). Multivariate logistic regression analysis in the mUC-ICB cohort identified RFrs and tumor mutational burden (TMB) as the two independent significant variables influencing ICB response (Fig. 5F). Subsequently, an integrated decision-making tree was

constructed based on RFRs and TMB using the “rpart” package, categorizing the mUC-ICB cohort into three groups, each with significantly different proportions of responders (4.55% responder in group 1, 16.7% in group 2, and 64.3% in group 3; Fig. 5G).

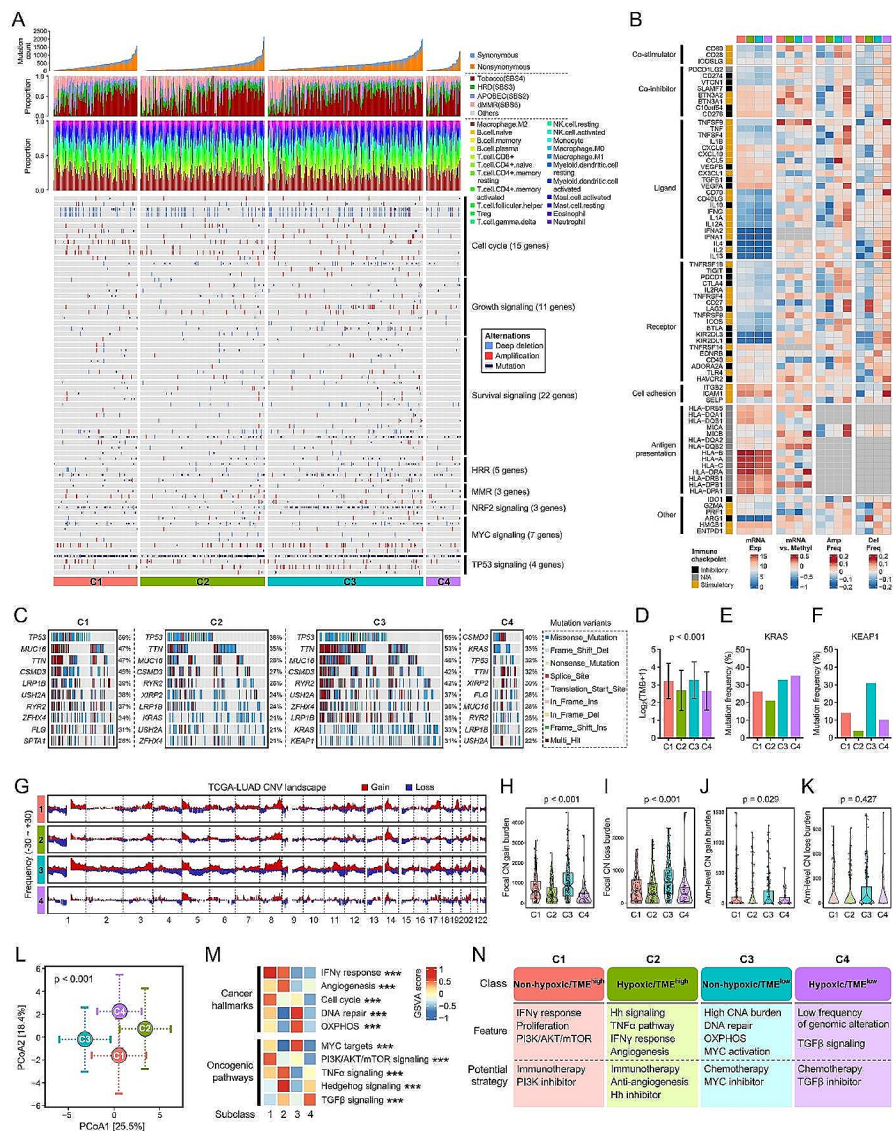
In order to identify potential targets and therapeutic agents for STAD patients with low RFRs who did not respond to ICB therapy, we conducted a screening of the CTRP database and assessed the AUC values of 481 compounds for each STAD sample using the inbuilt “calcPhenotype” function of the “pRRophetic” package. Next, we calculated the Pearson correlation coefficients between all the AUCs and the RFRs in the STAD-ICB cohort. Six compounds with a correlation  $>0.5$  (nilotinib, sildenafil, TG-100-115, tamoxifen, EX-527, and BRD8899) were showcased in Fig. 5H, revealing that the estimated AUCs of all the six compounds were notably lower in the low RFRs group (Fig. 5I). Furthermore, we conducted an exhaustive literature review to assess the clinical status and experimental evidence of the six candidate compounds, and the cMap score of each compound was also provided (Fig. 5J).

#### **Distinct genomic alterations were observed in the four hypoxic-TME subtypes in LUAD**

We comprehensively analyzed the genomic alterations in the four hypoxic-TME subtypes in LUAD. The mutation burden, mutational signatures, distribution of infiltrating immune cells, and alterations in representative oncogenic pathways were depicted in Fig. 6A. Understanding the expression of immune-related genes and regulatory mechanisms within different hypoxic TME states is crucial. Therefore, we examined the expression of these genes, as well as somatic copy-number alterations (SCNAs) and epigenetic mechanisms. Our findings revealed considerable variation in the expression of most immune-related genes across hypoxic-TME subtypes, potentially highlighting their influence on shaping the hypoxic TME within the TCGA-LUAD cohort (Fig. 6B). Furthermore, the mutational oncoplots for the four hypoxic-TME subtypes were depicted in Fig. 6C. Notably, we found that the tumor mutation burden (TMB) was significantly decreased in C2 and C4 compared to C1 and C3 (Fig. 6D). In addition, the mutation frequency of KRAS is shown as the lowest in C2 (Fig. 6E), while the highest KEAP1 mutation in C3 (Fig. 6F).

The copy number variation (CNV) landscape across the 22 chromosomes within the four subtypes was analyzed and illustrated in Fig. 6G. Our findings revealed that C3 demonstrated the highest focal gain and loss burden (Fig. 6H), whereas C4 displayed the lowest burden (Fig. 6I). A landscape of focal SCNA levels in the four subtypes could be found in Fig. S4E. Moreover, there were no significant differences in the arm-level burden observed among the four subtypes (Fig. 6J and K).

The ssGSEA algorithm quantified various cancer hallmarks and oncogenic pathways for each sample, with the subsequent principal coordinate analysis (PCoA) highlighting significant differences in these hallmarks and pathways across the four subtypes (Fig. 6L). These distinctions, including the top five hallmarks such as IFNG response, and the top five pathways such as MYC targets among the subtypes, are visually represented in a heatmap (Fig. 6M). Based on these findings, a summary outlining specific features and potential therapeutic strategies for the four subtypes was provided in Fig. 6N. In summary, C1 patients may benefit from immunotherapy and PI3K inhibitors, C2 patients may benefit from immunotherapy, anti-angiogenesis, and Hedgehog inhibitors, C3



**Fig. 6** Distinct genomic alterations were observed in the four hypoxic-TME subtypes in LUAD. **(A)** A comprehensive analysis of the mutation burden, mutational signatures, infiltrating immune cells, and alterations in representative oncogenic pathways among the four hypoxic-TME subtypes. **(B)** Considerable variation in the expression, CNV, and epigenetic alterations of most immune-related genes across hypoxic-TME subtypes. **(C)** The mutational oncoplots for the four hypoxic-TME subtypes. **(D)** The tumor mutation burden (TMB) was significantly decreased in C2 and C4 compared to C1 and C3 ( $p < 0.001$ ). **(E)** The mutation frequency of KRAS in the four subtypes. **(F)** The mutation frequency of KEAP1 in the four subtypes. **(G)** The CNV landscape across the 22 chromosomes within the four subtypes was analyzed and illustrated. **(H)** C3 demonstrated the highest focal gain and loss burden. **(I)** C4 displayed the lowest focal gain and loss burden. **(J & K)** There were no significant differences in the arm-level burden observed among the four subtypes. **(L)** PCA analysis highlights significant differences in these hallmarks and pathways across the four subtypes. **(M)** The top five hallmarks such as IFNγ response, and the top five pathways such as MYC targets among the subtypes, are visually represented in a heatmap. **(N)** A summary of specific features and potential therapeutic strategies for the four subtypes. \*\*\*  $p < 0.001$



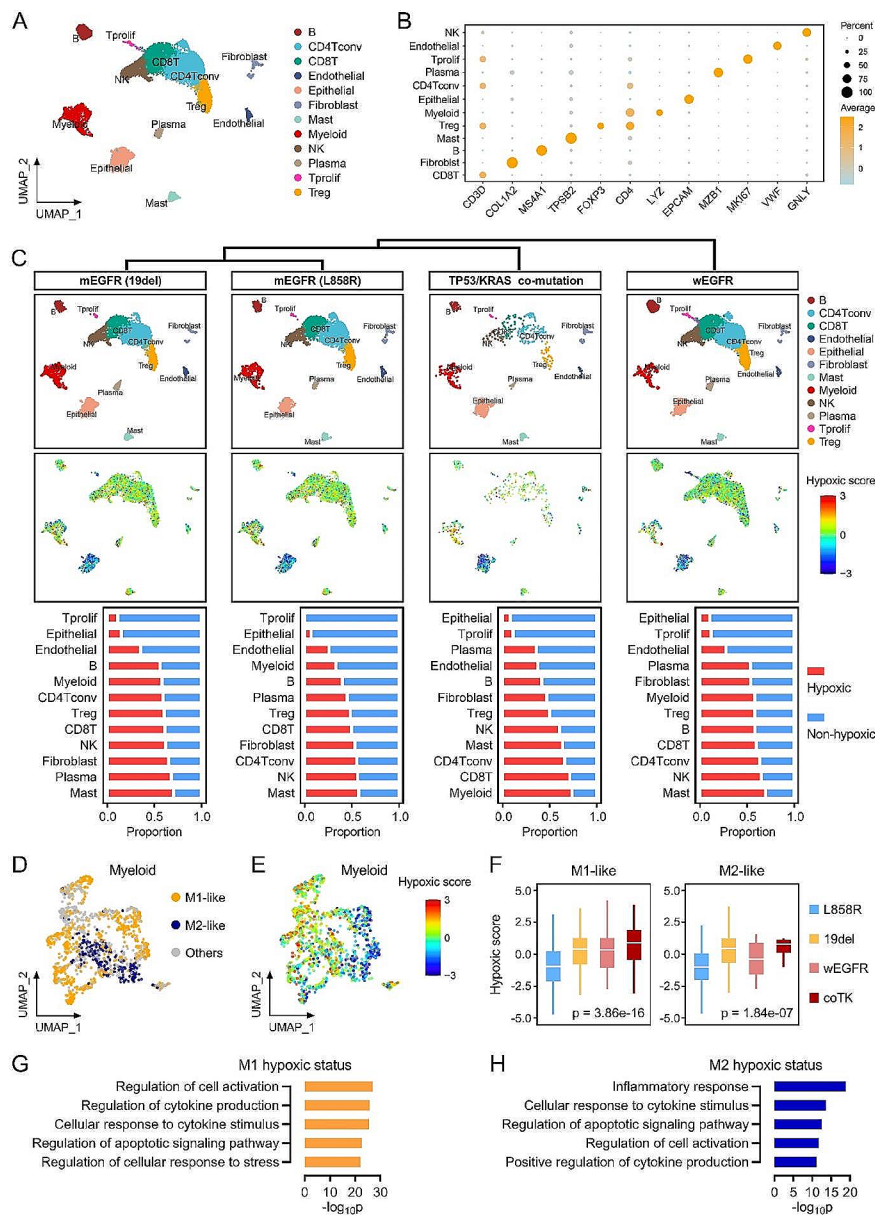
patients may benefit from chemotherapy and MYC inhibitors, and C4 patients may benefit from chemotherapy and TGF $\beta$  inhibitors.

### **Hypoxic TME is closely correlated with mutation status in LUAD**

To investigate the association between the hypoxic tumor microenvironment (TME) and mutation status in LUAD, we analyzed a scRNA-seq dataset comprising 9 samples with four different mutation types (EGFR 19del, L858R, TP53/KRAS co-mutation, and EGFR-wild type). A UMAP plot was generated to depict the distribution of distinct cell types within the LUAD TME (Fig. 7A), and specific markers for each cell type were visualized in a dot plot (Fig. 7B). Additionally, UMAP plots were used to display the distribution of various cell types in samples with different mutation types. Furthermore, the hypoxic score was computed for each cell, and the proportion of hypoxic and non-hypoxic cells for each cell type was represented in stacked bar plots within each mutation type. Notably, myeloid cells in the TP53/KRAS co-mutation samples exhibited the highest hypoxic proportion (Fig. 7C). Subsequently, myeloid cells were extracted, and M1-like and M2-like macrophages were identified using the ssGSEA algorithm based on corresponding M1 and M2 gene signatures (Fig. 7D). Concurrently, hypoxic scores for all myeloid cells were quantified and mapped on the UMAP plot (Fig. 7E). The analysis revealed notable variations in hypoxic scores among different mutation types in both M1-like and M2-like macrophages (Fig. 7F). DEGs between hypoxic and non-hypoxic cells were identified, and we found that the hypoxic status of M1-like macrophages primarily impacted cell activation and cytokine production (Fig. 7G), while the hypoxic status of M2-like macrophages mainly affected the inflammatory response and cellular response to cytokine stimulus (Fig. 7H). These findings suggested a significant association between the hypoxic TME and mutation types in LUAD, influencing specific cellular biological processes such as promoting or suppressing inflammatory effects in macrophages.

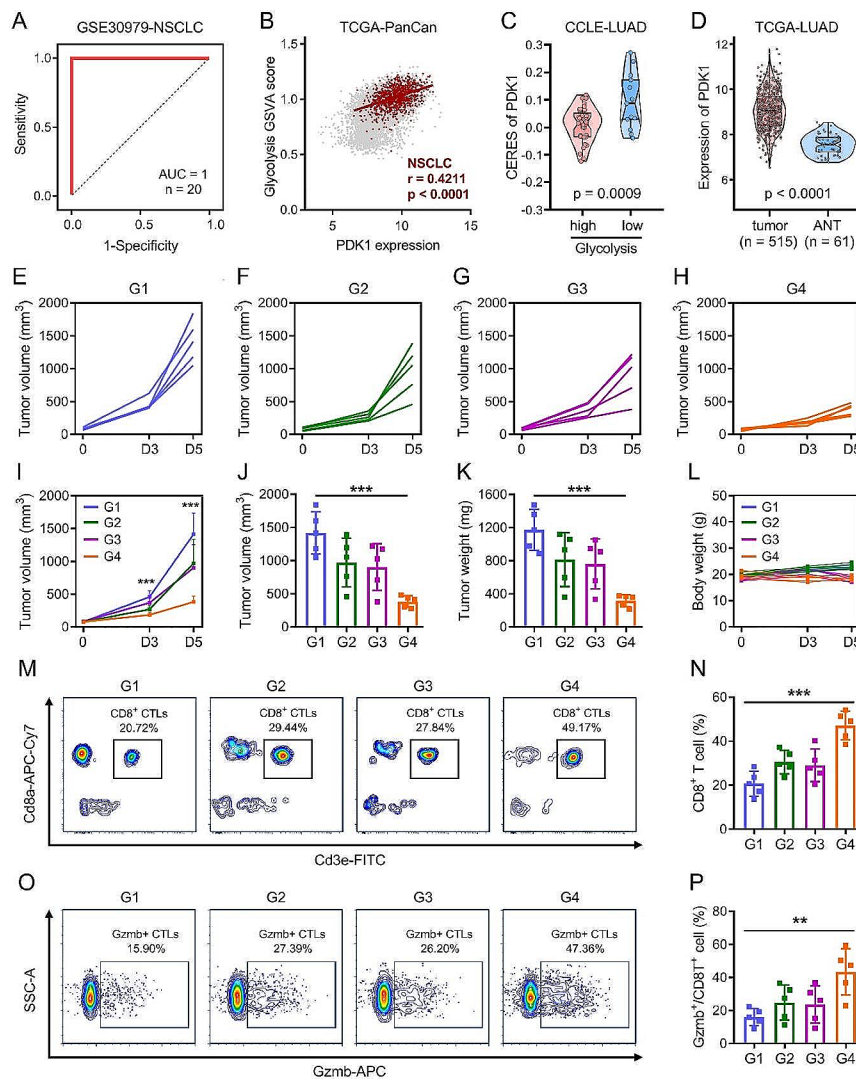
### **Identification of PDK1 as a hypoxic-related target to enhance the efficiency of immunotherapy**

In the hypoxic gene signature, we observed a significant correlation between PDK1 expression and glycolysis in the majority of solid cancers (Fig. S5A). It is well acknowledged that glycolysis could attenuate the efficacy of immunotherapy. A protein-protein interaction (PPI) network was constructed for PDK1 (Fig. S5B), revealing a substantial and positive correlation between PDK1 and various hypoxic and glycolysis markers (Fig. S5C). Using the GSE30979 dataset, we demonstrated that PDK1 expression could accurately discriminate hypoxic samples from normoxic ones (AUC=1,  $n=20$ ; Fig. 8A). Furthermore, a positive correlation ( $r=0.4211$ ,  $p<0.0001$ ) between PDK1 expression and glycolysis ssGSEA score was observed in the TCGA-NSCLC cohort (Fig. 8B). In CCLE-LUAD cell lines, PDK1 demonstrated significantly lower CERES in glycolysis-high cells compared to glycolysis-low ones (Fig. 8C). Additionally, PDK1 expression was significantly upregulated in LUAD samples compared to adjacent normal tissues (Fig. 8D). Given that brain metastases (mBrain) from LUAD tend to form a hypoxic TME compared to primary tumors, we utilized a LUAD-mBrain scRNA-seq dataset to analyze the relationship between PDK1 and glycolysis in the hypoxic TME. Dissection of various cell types with corresponding markers from the mBrain samples was depicted in a UMAP



**Fig. 7** Hypoxic TME is closely correlated with mutation status in LUAD. **(A)** A LUAD scRNA-seq dataset comprising 9 samples with four different mutation types (EGFR 19del, L858R, TP53/KRAS co-mutation, and EGFR-wild type) was used for further study. **(B)** Specific markers for each cell type were visualized in a dot plot. **(C)** UMAP plots were used to display the distribution of various cell types in samples with different mutation types. The hypoxic score was computed for each cell, and the proportion of hypoxic and non-hypoxic cells for each cell type was represented in stacked bar plots within each mutation type. **(D)** Myeloid cells were extracted, and M1-like and M2-like macrophages were identified using the ssGSEA algorithm based on corresponding M1 and M2 gene signatures. **(E)** Hypoxic scores for all myeloid cells were quantified and mapped on the UMAP plot. **(F)** Notable variations in hypoxic scores among different mutation types in both M1-like and M2-like macrophages. **(G)** DEGs between hypoxic and non-hypoxic cells were identified, and we found that the hypoxic status of M1-like macrophages primarily impacted cell activation and cytokine production. **(H)** The hypoxic status of M2-like macrophages mainly affected the inflammatory response and cellular response to cytokine stimulus





**Fig. 8** Identification of PDK1 as a hypoxic-related target to enhance the efficiency of immunotherapy. **(A)** The PDK1 expression could accurately discriminate hypoxic samples from normoxic ones in the GSE30979 dataset ( $AUC = 1$ ,  $n = 20$ ). **(B)** A positive correlation ( $r = 0.4211$ ,  $p < 0.0001$ ) between PDK1 expression and glycolysis ssGSEA score was observed in the TCGA-NSCLC cohort. **(C)** In CCLE-LUAD cell lines, PDK1 demonstrated significantly lower CERES in glycolysis-high cells compared to glycolysis-low ones. **(D)** PDK1 expression was significantly upregulated in LUAD samples compared to adjacent normal tissues. **(E - K)** The combination therapy group, involving PDK1 inhibition and PD-1 blockade, exhibited the most notable reduction in tumor burden compared to the other groups. **(L)** No significant body weight loss or other common toxic effects were observed across the four groups. **(M & N)** The combination of PD-1 mAb and PDK1 inhibition led to the most substantial enrichment of CD8<sup>+</sup> T cells within the tumors. **(O & P)** Tumors from the combination therapy group particularly showed a significant higher proportion of Gzmb<sup>+</sup>CD8<sup>+</sup> T cells. \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

plot (Fig. S5D), revealing that PDK1 was enriched only in a subset of malignant LUAD cells (highlighted in a dotted circle; Fig. S5E). Notably, the hypoxic score (Fig. S5F) and glycolytic activity (Fig. S5G) were also prominently manifested in the same area. These findings collectively suggest a substantial correlation between PDK1 and glycolytic activity in the hypoxic TME, uncovering its potential as a therapeutic target for LUAD.

To assess the immune regulatory impact of PDK1 in vivo, we employed the mouse LUAD cell line LLC to establish subcutaneous xenografts in immunocompetent C57BL/6 mice. The mice were then subjected to treatment with anti-PD-1 monoclonal

antibody (mAb), an IgG isotype control, a PDK1 inhibitor, and a combination of the PDK1 inhibitor and the anti-PD-1 mAb. The experimental procedure is depicted in the flowchart shown in Fig. S6. Our results indicated that PDK1 inhibition resulted in restricted tumor growth when compared to the control group, a pattern consistent with PD-1 mAb treatment. Additionally, the combination therapy group, involving PDK1 inhibition and PD-1 blockade, exhibited the most notable reduction in tumor burden compared to the other groups (Fig. 8E and K). No significant body weight loss or other common toxic effects were observed across the four groups (Fig. 8L). Tumor samples were harvested for further analysis upon completion of the experiment. Notably, the combination of PD-1 mAb and PDK1 inhibition led to the most substantial enrichment of CD8<sup>+</sup> T cells within the tumors (Fig. 8M and N), with the tumors from the combination therapy group particularly displaying enriched Gzmb<sup>+</sup>CD8<sup>+</sup> T cells (Fig. 8O and P). These findings demonstrated that PDK1 inhibition could promote the infiltration of CD8<sup>+</sup> T cells, thereby enhance the anti-tumor effects induced by PD-1 therapy.

## Discussion

Hypoxia occurs in tumor regions with inadequate oxygen supply and poses a significant barrier to effective cancer treatment. The distribution of hypoxia levels is highly heterogeneous, varying from almost non-hypoxic to severe and anoxic level. These individual hypoxia levels trigger a range of biological responses that hamper treatment effectiveness. Therefore, accurately quantifying the level of hypoxia can significantly improve the development of personalized treatment strategies for cancer patients with varying degrees of hypoxia and indicates great clinical importance. In this study, to address the above-mentioned issue, we initially developed a robust pan-cancer hypoxic quantification method utilizing multiple public databases and various bioinformatics and statistical methods, and the accuracy rate reaches 100% for all the cell lines or tumor fragments cultured under hypoxia or normoxia. Compared to other diagnostic methods for tumor hypoxia, our hypoxic scoring approach appears to be simpler and more efficient: by applying the ssGSEA algorithm with two signatures (hypoxic and normoxic) to RNA-sequencing data, it is possible to determine the hypoxic status of a specific tumor specimen accurately. Utilizing our methodology, the TCGA pan-cancer specimens with transcriptome profiling data could be categorized as either hypoxic or non-hypoxic.

It has recently become evident that hypoxia plays a crucial role in modulating the interactions among cancer cells, stromal, and immune cells, thereby influencing the composition and structure of the tumor microenvironment. That is to say, tumor hypoxia is not a simple shortage of oxygen but a catalyst for transformation within the tumor microenvironment: reshaping cellular interactions, metabolic processes, and the immune landscape. We classified all tumor samples into four subtypes according to their hypoxic and TME score: non-hypoxic/TME<sup>high</sup> (C1), hypoxic/TME<sup>high</sup> (C2), non-hypoxic/TME<sup>low</sup> (C3), hypoxic/TME<sup>low</sup> (C4), and comprehensively analyzed the distinct patterns of immune infiltration and landscape, mutational signatures, and genomic variability across the four hypoxic-TME subtypes, facilitating the understanding of the hypoxic TME molecular subtypes and genomic features. Interestingly, a significant feature of C4 is the high frequency of IDH1 mutation (a driver gene in LGG), accompanied by the lowest levels of activated CD8<sup>+</sup> cells infiltration and other immune features. These findings implied that distinct hypoxic-TME subtypes correspond to varied

tumor immunogenicity, leading to differing effectiveness of immunotherapeutic therapy. Through comprehensive analysis of genomic alterations and oncogenic pathways in the TCGA-LUAD cohort, we identified that C1 and C2 patients are likely to benefit from immunotherapy, whereas the strategy of chemotherapy in conjunction with pathway-specific inhibitors is more appropriate for C3 or C4 patients. Notably, C3 patients exhibit a high burden of copy number variations (CNV) and a high rate of DNA repair, yet this subset demonstrates low potential of the IFNG response. From the scRNA-seq datasets, we observed that the hypoxic status was influenced by various mutational types such as EGFR and TP53/KRAS co-mutation in LUAD, and that hypoxia could also impact the cytolytic activity of CD8<sup>+</sup> T cells and the biological processes of TAM1 or TAM2, including cytokine production and inflammatory response.

Insights into the impacts of tumor hypoxia on the TME could unveil the delicate balance and immense plasticity of cancer cells in adapting to challenging conditions. In light of the strong correlation between hypoxic status and TME signatures with immunotherapy effectiveness, we developed a machine learning-based hypoxic-TME scoring model to predict ICB response in cancer patients. Through comparison with 10 conventional machine learning approaches, the random forest algorithm emerged as the optimal model, with the responsive score it produced outperforming documented ICB biomarkers in testing cohorts. Employing machine learning approaches for molecular subtyping is believed to significantly enhance the prediction of clinical outcomes and therapeutic responses, deserving a broad range of application prospects in clinical practice.

Metabolic reprogram is another hallmark response to hypoxic stress, with tumor cells exhibiting a preference for glycolysis even in the presence of oxygen, a phenomenon known as the Warburg effect. This metabolic shift not only ensures continuance of energy supply but also contributes to the acidification of the tumor microenvironment, which in turn facilitates tumor aggressiveness especially drug resistance. In this study, we identified a key gene, PDK1, functioning as a link between hypoxia and glycolysis. Subsequently, we utilized the mouse LUAD cell line LLC to establish subcutaneous xenografts in immunocompetent C57BL/6 mice and examined the immune modulatory effect of PDK1 *in vivo*. Our observations revealed that inhibition of PDK1 could enhance the infiltration of CD8<sup>+</sup> T cells, thereby augmenting the anti-tumor effects elicited by PD-1 therapy. These findings suggest that targeting essential glycolysis-related factors such as PDK1 involved in the hypoxic TME may effectively enhance the anti-tumor immune response induced by PD-1 therapy.

Our study acknowledges several limitations. While we carefully utilized various multi-omics databases for rigorous validation and implemented the Combat algorithm to remove batch effects, it is important to recognize that sampling bias stemming from tumor genetic heterogeneity and the challenges of cross-platform integration could only be alleviated but not completely eliminated. Additionally, further clinical studies in a real world are needed to elucidate the clinical relevance and generalizability of our hypoxic-TME model.

In this study, we are pleasantly surprised to discover that machine learning as a powerful tool for predicting immunotherapeutic outcomes in cancer patients due to its ability to analyze complex data patterns and generate precise predictions. By leveraging machine learning algorithms, as detailed in our study, clinicians can integrate various patient data, particularly genomic data, to construct precise predictive models that

facilitate the guidance of personalized immunotherapy. While filled with anticipation, we must also admit that the machine learning approaches achieving clinical translation truly requires collective efforts and a test of time.

In summary, our study may offer valuable insights for integrating hypoxic-TME classification into tumor staging and providing personalized strategies for cancer patients.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40537-024-00945-2>.

Supplementary Material 1

Supplementary Material 2

## Acknowledgements

We acknowledge all the other colleagues in our department for their helpful discussions and comments.

## Author contributions

R.S. and Y.S. conceived and designed the study. R.S. and J.S. performed the bioinformatics and statistical analyses. H.Z., T.H., and Z.G. participated in the cell and animal experiments. X.W. helped to guide the flow cytometry analysis. Z.Z. and M.L. guided data analysis and assisted with language improvement. All authors have read and approved the final manuscript.

## Funding

This work was supported by the grants from Jiangsu Provincial Medical Innovation Center of Jiangsu Province Capability Improvement Project through Science, Technology and Education (No. CXZX202204 to Yongqian Shu), Natural Science Foundation of Jiangsu Province (No. BK20220732 to Run Shi), Jiangsu Funding Program for Excellent Postdoctoral Talent (No. 2022ZB427 to Run Shi), and Postdoctoral International Exchange Program (No. YJ20220433 to Run Shi).

## Data availability

The data and codes that support the findings in this study are available from the corresponding author on reasonable request.

## Declarations

### Ethics approval and consent to participate

All the animal experiments were conducted in accordance with protocols approved by the Research Ethics Committee of Nanjing Medical University.

### Consent for publication

Not applicable.

### Competing interests

No potential conflict of interest was reported by the authors.

### Author details

<sup>1</sup>Department of Oncology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

<sup>2</sup>Department of Endocrinology, Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China

<sup>3</sup>The First Clinical Medical College of Nanjing University of Chinese Medicine, Nanjing, China

<sup>4</sup>Department of Thoracic Surgery, The Affiliated Changzhou No. 2 People's Hospital of Nanjing Medical University, Changzhou, China

<sup>5</sup>Department of Oncology, Jiangsu Cancer Hospital, The Affiliated Cancer Hospital of Nanjing Medical University, Jiangsu Institute of Cancer Research, Nanjing, China

<sup>6</sup>Department of Radiation Oncology, LMU University Hospital, LMU Munich, Munich, Germany

<sup>7</sup>Department of Radiation Oncology, Lueneburg Municipal Hospital, Lueneburg, Germany

<sup>8</sup>Department of Urology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

<sup>9</sup>Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, China

Received: 17 February 2024 / Accepted: 31 May 2024

Published online: 07 June 2024

## References

1. Harris AL. Hypoxia—a key regulatory factor in tumour growth. *Nat Rev Cancer*. 2002;2(1):38–47.
2. Vaupel P, Mayer A. Hypoxia in cancer: significance and impact on clinical outcome. *Cancer Metastasis Rev*. 2007;26(2):225–39.

3. Semenza GL. Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy. *Trends Pharmacol Sci.* 2012;33(4):207–14.
4. Rankin EB, Giaccia AJ. Hypoxic control of metastasis. *Science.* 2016;352(6282):175–80.
5. Bhandari V, Hoey C, Liu LY, Lalonde E, Ray J, Livingstone J, Lesurf R, Shiah YJ, Vujcic T, Huang X, et al. Molecular landmarks of tumor hypoxia across cancer types. *Nat Genet.* 2019;51(2):308–18.
6. Lee KE. Hypoxia as a regulator of tumor stroma and metastasis. *Am J Physiol Cell Physiol.* 2023;324(1):C10–3.
7. Walsh JC, Lebedev A, Aten E, Madsen K, Marciano L, Kolb HC. The clinical importance of assessing tumor hypoxia: relationship of tumor hypoxia to prognosis and therapeutic opportunities. *Antioxid Redox Signal.* 2014;21(10):1516–54.
8. Ye Y, Hu Q, Chen H, Liang K, Yuan Y, Xiang Y, Ruan H, Zhang Z, Song A, Zhang H, et al. Characterization of Hypoxia-associated molecular features to Aid Hypoxia-targeted therapy. *Nat Metab.* 2019;1(4):431–44.
9. Shi R, Bao X, Unger K, Sun J, Lu S, Manapov F, Wang X, Belka C, Li M. Identification and validation of hypoxia-derived gene signatures to predict clinical outcomes and therapeutic responses in stage I lung adenocarcinoma patients. *Theranostics.* 2021;11(10):5061–76.
10. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med.* 2013;19(11):1423–37.
11. Joyce JA, Fearon DT. T cell exclusion, immune privilege, and the tumor microenvironment. *Science.* 2015;348(6230):74–80.
12. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, Coussens LM, Gabrilovich DI, Ostrand-Rosenberg S, Hedrick CC, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med.* 2018;24(5):541–50.
13. Wilson WR, Hay MP. Targeting hypoxia in cancer therapy. *Nat Rev Cancer.* 2011;11(6):393–410.
14. Wang S, Li H, Song M, Tao Z, Wu T, He Z, Zhao X, Wu K, Liu XS. Copy number signature analysis tool and its application in prostate cancer reveals distinct mutational processes and clinical outcomes. *PLoS Genet.* 2021;17(5):e1009557.
15. Kim ST, Cristescu R, Bass AJ, Kim KM, Odegaard JJ, Kim K, Liu XQ, Sher X, Jung H, Lee M, et al. Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer. *Nat Med.* 2018;24(9):1449–58.
16. Rose TL, Weir WH, Mayhew GM, Shibata Y, Eulitt P, Uronis JM, Zhou M, Nielsen M, Smith AB, Woods M, et al. Fibroblast growth factor receptor 3 alterations and response to immune checkpoint inhibition in metastatic urothelial cancer: a real world experience. *Br J Cancer.* 2021;125(9):1251–60.
17. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, Berent-Maoz B, Pang J, Chmielowski B, Cherry G, et al. Genomic and transcriptomic features of response to Anti-PD-1 therapy in metastatic melanoma. *Cell.* 2016;165(1):35–44.
18. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, Hodi FS, Martin-Algarra S, Mandal R, Sharfman WH, et al. Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell.* 2017;171(4):934–e949916.
19. Nathanson T, Ahuja A, Rubinsteyn A, Aksoy BA, Hellmann MD, Miao D, Van Allen E, Merghoub T, Wolchok JD, Snyder A, et al. Somatic mutations and Neoepitope Homology in Melanomas treated with CTLA-4 blockade. *Cancer Immunol Res.* 2017;5(1):84–91.
20. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, Sucker A, Hillen U, Foppen MHG, Goldinger SM, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science.* 2015;350(6257):207–11.
21. Gide TN, Quek C, Menzies AM, Tasker AT, Shang P, Holst J, Madore J, Lim SY, Velickovic R, Wongchenko M, et al. Distinct Immune cell populations define response to Anti-PD-1 monotherapy and Anti-PD-1/Anti-CTLA-4 combined Therapy. *Cancer Cell.* 2019;35(2):238–55. e236.
22. Yang L, He YT, Dong S, Wei XW, Chen ZH, Zhang B, Chen WD, Yang XR, Wang F, Shang XM et al. Single-cell transcriptome analysis revealed a suppressive tumor immune microenvironment in EGFR mutant lung adenocarcinoma. *J Immunother Cancer* 2022, 10(2).
23. Jeong HY, Ham IH, Lee SH, Ryu D, Son SY, Han SU, Kim TM, Hur H. Spatially distinct reprogramming of the Tumor Microenvironment based on Tumor Invasion in diffuse-type gastric cancers. *Clin Cancer Res.* 2021;27(23):6529–42.
24. Shi R, Wang X, Wu Y, Xu B, Zhao T, Trapp C, Wang X, Unger K, Zhou C, Lu S, et al. APOBEC-mediated mutagenesis is a favorable predictor of prognosis and immunotherapy for bladder cancer patients: evidence from pan-cancer analysis and multiple databases. *Theranostics.* 2022;12(9):4181–99.
25. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun.* 2013;4:2612.
26. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015;12(5):453–7.
27. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 2017;18(1):248–62.
28. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, et al. The Immune Landscape of Cancer. *Immunity.* 2018;48(4):812–e830814.
29. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell.* 2015;160(1–2):48–61.
30. Pacini C, Dempster JM, Boyle I, Goncalves E, Najgebauer H, Karakoc E, van der Meer D, Barthorpe A, Lightfoot H, Jaaks P, et al. Integrated cross-study datasets of genetic dependencies in cancer. *Nat Commun.* 2021;12(1):1661.
31. Basu A, Bodycombe NE, Cheah JH, Price EV, Liu K, Schaefer GI, Ebright RY, Stewart ML, Ito D, Wang S, et al. An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell.* 2013;154(5):1151–61.
32. Gleeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS ONE.* 2014;9(9):e107468.
33. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.
34. Liu Y, Zhang Q, Xing B, Luo N, Gao R, Yu K, Hu X, Bu Z, Peng J, Ren X, et al. Immune phenotypic linkage between colorectal cancer and liver metastasis. *Cancer Cell.* 2022;40(4):424–e437425.

## Publisher's Note

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