

ORIGINAL ARTICLE

OPEN

Comprehensive molecular classification predicted microenvironment profiles and therapy response for HCC

Yihong Chen¹ | Xiangying Deng^{1,2} | Yin Li¹ | Ying Han^{1,2} | Yinghui Peng¹ | Wantao Wu¹ | Xinwen Wang¹ | Jiayao Ma¹ | Erya Hu¹ | Xin Zhou¹ | Edward Shen³ | Shan Zeng^{1,2} | Changjing Cai^{1,2} | Yiming Qin⁴ | Hong Shen^{1,2} 

¹Department of Oncology, Xiangya Hospital, Central South University, Changsha, Hunan, China

²National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, Hunan, China

³Department of Life Science, McMaster University, Hamilton, Ontario, Canada

⁴Department of Spine Surgery and Orthopaedics, Xiangya Hospital, Central South University, Changsha, Hunan, China

Correspondence

Changjing Cai, Department of Oncology, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China.
Email: vccj07@csu.edu.cn

Yiming Qin, Department of Spine Surgery and Orthopaedics, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China.
Email: qinyimingdoctor@163.com

Hong Shen, Department of Oncology, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China.
Email: hongshen2000@csu.edu.cn

Abstract

Background and Aims: Tumor microenvironment (TME) heterogeneity leads to a discrepancy in survival prognosis and clinical treatment response for patients with HCC. The clinical applications of documented molecular subtypes are constrained by several issues.

Approach and Results: We integrated 3 single-cell data sets to describe the TME landscape and identified 6 prognosis-related cell subclusters. Unsupervised clustering of subcluster-specific markers was performed to generate transcriptomic subtypes. The predictive value of these molecular subtypes for prognosis and treatment response was explored in multiple external HCC cohorts and the Xiangya HCC cohort. TME features were estimated using single-cell immune repertoire sequencing, mass cytometry, and multiplex immunofluorescence. The prognosis-related score was constructed based on a machine-learning algorithm. Comprehensive single-cell analysis described TME heterogeneity in HCC. The 5 transcriptomic subtypes possessed different clinical prognoses, stemness characteristics, immune landscapes, and therapeutic responses. Class 1 exhibited an inflamed phenotype with better clinical outcomes, while classes 2 and 4 were characterized by a lack of T-cell infiltration. Classes 5 and 3 indicated an inhibitory tumor immune microenvironment. Analysis of multiple therapeutic cohorts suggested that classes 5 and 3 were sensitive to immune checkpoint blockade and targeted therapy, whereas classes 1 and 2 were more responsive to transcatheter arterial chemoembolization treatment. Class 4 displayed resistance to all conventional HCC therapies. Four potential therapeutic agents and 4 targets were further identified for high prognosis-related score patients with HCC.

Abbreviations: BCR, B cell receptor; CNV, copy number variation; CyTOF, mass cytometry; ICB, immune checkpoint blockade; MDSC, myeloid-derived suppressor cell; PPI, protein-protein interaction network; PRS, prognosis-related score; SCENIC, single-cell regulatory network inference and clustering; scRNA-seq, single-cell RNA sequencing; SSMs, subcluster-specific markers; TACE, transcatheter arterial chemoembolization treatment; TCR, T cell receptor; TME, tumor microenvironment; t-SNE, t-distributed stochastic neighbor embedding.

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.hepjournal.com.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal. Copyright © 2024 The Author(s). Published by Wolters Kluwer Health, Inc.

Conclusions: Our study generated a clinically valid molecular classification to guide precision medicine in patients with HCC.

INTRODUCTION

HCC is the most common type of primary liver cancer and is the third leading cause of cancer-associated death worldwide.^[1] Studies have reported that high tumor microenvironment (TME) heterogeneity may cause discrepancies in survival prognosis and clinical treatment responses of patients with HCC.^[2] Therefore, describing TME heterogeneity contributes to revealing the biological characteristics of HCC and offers insights into individualized treatment for patients with HCC. Developing a reliable classification system for stratifying patients in the context of TME heterogeneity may provide prospective guidance for precision medicine in HCC.

In recent years, emerging single-cell RNA sequencing (scRNA-seq) technology has proven to be a powerful tool for exploring cellular diversity and tumor heterogeneity.^[3] One study revealed intertumoral and intratumoral transcriptomic heterogeneity in HCC and discovered the molecular features correlated with TME reprogramming.^[4] scRNA-seq allows the classification of malignant and nonmalignant cells based on aneuploid copy number profiles and the identification of clonal substructures in different subclusters.^[5] scRNA-seq enables the analysis of tissue heterogeneity at the level of individual cell level and delves into the contributions of distinct cell subclusters to biological functions and pathogenesis.

Several classifications have been developed based on conventional bulk-RNA and proteomic sequence data, and are found to have different biological characteristics and significant survival differences.^[6,7] However, the limited sample size and high classifier complexity have hindered the development of a unified and mature molecular classification. Currently, the treatment of HCC still faces formidable challenges, such as difficulties in prevention and limited treatment targets, etc. In addition, there is still a lack of effective molecular classifications to predict therapeutic efficacy or provide treatment guidance for patients with HCC.

In this study, we used scRNA-seq data to describe the TME landscape of HCC and identified 6 prognosis-related subclusters. Subsequently, we developed 5 transcriptomic subtypes using unsupervised clustering of subcluster-specific markers (SSMs) in the training cohort. These molecular subtypes exhibited distinct clinical outcomes, stemness characteristics, genomic variations, and immune infiltration microenvironments. The therapeutic responses to immune checkpoint blockade (ICB), transcatheter arterial chemoembolization

(TACE), and targeted therapy in patients with HCC with distinct molecular subtypes were explored using publicly available clinical treatment cohorts and the Xiangya HCC cohort. The characteristics of TME were further validated using scRNA-seq and single cell T cell receptor/B cell receptor-sequencing, mass cytometry (CyTOF), and multiplex immunofluorescence. In summary, the integration of scRNA-seq data and bulk RNA-seq cohorts allowed us to generate clinically valid molecular subtypes, providing a theoretical basis for the development of individualized therapeutic approaches for patients with HCC.

METHODS

Publicly available data collection and preprocess

In total, 12 public RNA-seq or microarray data sets with clinical information were collected. Clinical characteristics of cohorts involved in this study are presented in Supplemental Table S1, <http://links.lww.com/HEP/I377>. Approximately 248,478 cells from 3 public scRNA-seq data sets were collected.

Xiangya HCC cohort acquisition

Fresh tumor tissues were surgically collected from 129 patients pathologically diagnosed with HCC at the Xiangya Hospital of Central South University and immediately stored in liquid nitrogen. After surgery, 36 patients received ICB therapy, 40 patients received targeted therapy, and 34 patients received TACE therapy. Treatment responses were evaluated according to mRECIST guidelines.^[8] Detailed clinical information is presented in Supplemental Table S1, <http://links.lww.com/HEP/I377>. We then performed bulk RNA-seq (BGI-Shenzhen, China), scRNA-seq (LC-Bio Technology, Hangzhou), and mass cytometry (Polaris Biology, Shanghai) analysis on 129, 6, and 16 patients of the Xiangya HCC cohort, respectively.

Statistical analysis

All the statistical tests were performed using the R software (v4.1.3). A comparison of continuous variables was performed using the nonparametric tests,

specifically, the Wilcoxon rank-sum test for 2 groups, and the Kruskal-Wallis test for more than 2 groups. A comparison of rates between 2 or more groups was performed using the chi-square test and Fisher exact test. The 95% CI for the population rate was estimated using a binomial distribution. Correlation between continuous variables was measured using Spearman rank order correlation. Survival data analysis included Kaplan-Meier survival analysis and log-rank test to check the statistical significance between groups. For statistical tests with multiple p values, the false discovery rate adjusting method was performed.

Detailed descriptions of methods used in this study are provided in the Supplemental Information, <http://links.lww.com/HEP/I391>.

RESULTS

Single-cell transcriptomics landscape of HCC

First, we integrated 3 scRNA-seq data sets GSE156625, GSE149614, and GSE125449 for analysis after removing batch effects and quality control (Supplemental Figure S1A, <http://links.lww.com/HEP/I369>). Dimensionality reduction clustering of these cells identified 7 major cell types according to canonical marker genes (Figure 1A and Supplemental Figure S1B, <http://links.lww.com/HEP/I369>). Subsequently, 40 distinct subclusters were reclustered, including 6 epithelial subclusters, 11 NK-cell and T-cell subclusters, 1 mast cell subcluster, 3 B-cell subclusters, 8 myeloid subclusters, 6 endothelial subclusters, and 5 fibroblasts subclusters (Figure 1A). For the identification of epithelial subclusters, we inferred copy number variation (CNV) of tumor cells and divided tumor cells into 5 subclusters (CNV_1–5) based on their CNV profiles (Figure 1B). The marker gene expression showed significant transcriptional differences between subclusters (Figure 1C). The cell subclusters were distributed in a tissue-specific manner (Figure 1D).

Identification of prognosis-related subclusters in HCC

The workflow of the identification of prognosis-related subclusters is presented in Figure 2A. After filtration, a total of 1587 SSMs were identified (Supplemental Table S4, <http://links.lww.com/HEP/I380>). Furthermore, the reference signature matrix of these 40 subclusters was acquired based on the expression values of 1587 SSMs across all subcluster cells for subcluster deconvolution (Supplemental Figure S1C, <http://links.lww.com/HEP/I369> and Supplemental Table S4, <http://links.lww.com/HEP/I380>).

We then collected clinical cohorts with RNA-seq and microarray transcriptomic data and separated these cohorts into a combined training cohort, a combined RNA-seq validation cohort, and a combined microarray validation cohort. A summary of the clinical information included in this study is presented in Supplemental Table S1, <http://links.lww.com/HEP/I377>. The collected HCC cohorts were applied for further analysis after removing the batch effect (Supplemental Figure S2A, <http://links.lww.com/HEP/I370>). To calculate the subcluster-specific score, the SSM expression values for each subcluster were z-score normalized and averaged in the combined training cohort. To infer the subcluster percentages, the CIBERSORT algorithm was conducted in a combined training cohort based on the reference signature matrix.

Subsequently, we conducted a univariate Cox regression analysis to screen the prognosis-related subclusters based on either subcluster percentage or subcluster-specific score (Supplemental Table S5, <http://links.lww.com/HEP/I381>). Using the threshold of adjusted p value <0.01 , a total of 6 prognosis-related subclusters were identified, including CNV_2, CNV_4, Treg_2, Mono, pDC, and FB_3 (Figure 2B). CNV_2, Treg_2, Mono, pDC, and FB_3 indicated poor outcomes, while CNV_4 was correlated with improved overall survival (Table 1). The specific expression of SSMs in corresponding prognosis-related subclusters was validated at the single-cell level (Supplemental Figure S1D, <http://links.lww.com/HEP/I369>). We further used CellPhoneDB to analyze the cell-cell communications between subclusters and observed most connections among fibroblasts, endothelial cells, and epithelial cells (Supplemental Figure S1E, <http://links.lww.com/HEP/I369>). Significant ligand-receptor interactions were detected between prognosis-related subclusters (Supplemental Figure S1F, <http://links.lww.com/HEP/I369>).

Construction of transcriptomic classification in HCC

We used unsupervised consensus clustering in the combined training cohort based on the expression of 103 SSMs of prognosis-related subclusters, and 5 transcriptomic subtypes were identified (Figure 2C). The expression profiles of 103 SSMs in 5 transcriptomic subtypes are shown in Supplemental Figure S2B, <http://links.lww.com/HEP/I370>. The 5 subtypes showed significant differences in prognosis (Figure 2D, log-rank test, $p=2e-6$). The subclusters of CNV_2, FB_3, Mono, pDC, and Treg_2 were significantly enriched in class 5, and class 1 was characterized by enrichment of CNV_4 (Figure 2E). Genomic analyses indicated that the *TERT* promoter, *TP53* (encoding p53), and *CTNNB1* (encoding β -catenin) were frequently mutated in HCC.^[9] *CTNNB1* mutation was associated with a more differentiated tumor type and prolonged overall survival,^[10] while the *TERT* promoter and *TP53* mutation demonstrated higher tumor proliferation

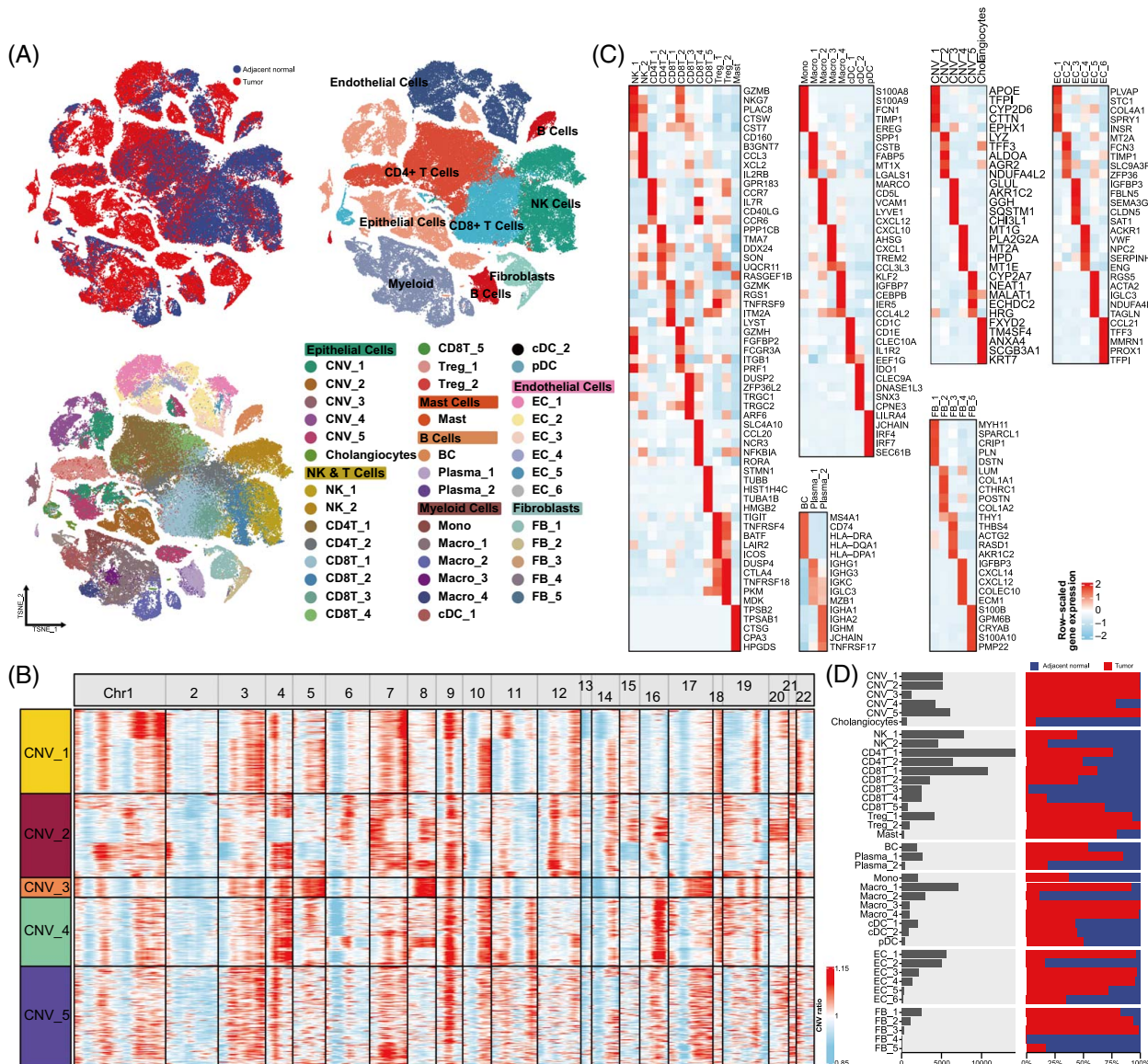


FIGURE 1 Single-cell transcriptomic analysis identified distinct cell subclusters. (A) t-SNE visualization of 248,478 cells annotated in 7 major cell types and 40 subclusters. Cells from different subclusters are marked by color code. (B) The tumor cells were divided into 5 CNV subclusters by applying the k-means clustering based on inferred single-cell CNV profiles. (C) Heatmap of marker gene expression in each cell subcluster. (D) Barplot displaying the number of cells, tissue type (adjacent normal or tumor), and the fraction of cells in each cell subcluster from left to right. Abbreviations: CNV, copy number variation; t-SNE, distributed stochastic neighbor embedding.

activity and a poor prognosis in patients with HCC.^[11] Therefore, we explored the *TERT* promoter mutation in 5 transcriptomic subtypes. In addition, the results of somatic mutation analysis revealed that the mutation frequency of *TP53* was significantly increased in classes 4 and 5 (16.8%, 21.8%, 32.7%, 54.3%, and 50% for classes 1, 2, 3, 4, and 5, respectively. Supplemental Figures S2C, D; <http://links.lww.com/HEP/I370>). In comparison, the mutation of *CTNNB1* was mainly enriched in classes 1 and 2 (Supplemental Figure S2C, <http://links.lww.com/HEP/I370>). We further compared the differences in prognosis and enriched pathways between classes 2 and 4, between classes 3 and 5 (Supplemental Figures S3A–S3D, <http://links.lww.com/HEP/I371>), between classes 1 and 2, and

between classes 3 and 4 (Supplemental Figures S3E–S3G, <http://links.lww.com/HEP/I371>). We found the cancer stemness activity of class 5 and class 4 was higher than class 3 and class 2, respectively (Supplemental Figure S3D, <http://links.lww.com/HEP/I371>).

To validate the survival prediction value of transcriptomic classes, we then developed a machine-learning-based model to predict transcriptomic classes in external validation cohorts. We evaluated the predictive accuracy of our classification system of 8 distinct machine-learning algorithms in the combined training cohort. Finally, support vector machine achieved the best performance with multiclass AUC values of 1 for the training set and 0.938 for the internal

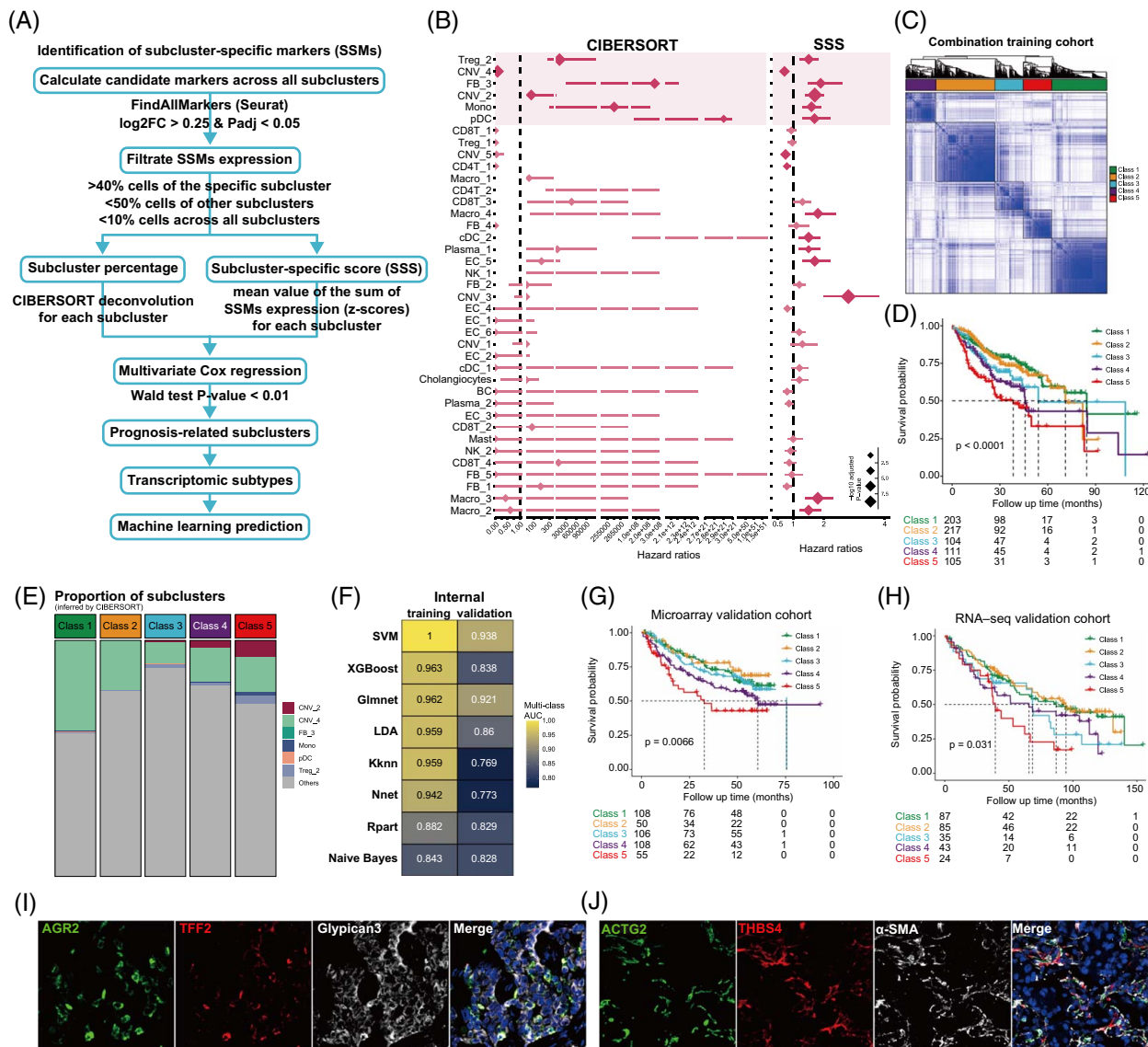


FIGURE 2 Identification of prognosis-related subclusters and transcriptomic subtypes in HCC. (A) Flowchart of screening prognosis-related subclusters and establishing transcriptomic classes in HCC. (B) Univariate Cox regression of subcluster percentages and subcluster-specific scores in the training cohort. (C) Consensus clustering identified 5 transcriptomic subtypes for the training cohort. (D) Kaplan-Meier curves of 5 transcriptomic subtypes in the training cohort. (E) Distribution of prognosis-related subclusters in 5 subtypes. (F) Multiclass AUCs of the internal training set and validation set based on 8 individual machine-learning algorithms. (G) Kaplan-Meier curves comparing the OS of 5 subtypes in the microarray validation cohort. (H) Kaplan-Meier curves showing the OS of 5 subtypes in the RNA-seq validation cohort. (I) Multiplex immunofluorescence staining of 2 CNV_2 SSMs (*TFF2* and *AGR2*). (J) Multiplex immunofluorescence staining of 2 FB_3 SSMs (*THBS4* and *ACTG2*). Abbreviations: CNV, copy number variation; OS, overall survival; SSM, subcluster-specific marker; SSS, subcluster-specific score.

validation set (Figure 2F and Supplemental Materials and Methods, <http://links.lww.com/HEP/I391>). Therefore, support vector machine was chosen to predict transcriptomic classes in external validation cohorts. Kaplan-Meier analysis showed remarkably different overall survival of 5 transcriptomic classes in the external microarray validation cohort (Figure 2G, log-rank test, $p = 0.0066$) and the RNA-seq validation cohort (Figure 2H, log-rank test, $p = 0.031$). CNV_2 markers (*TFF2* and *AGR2*) were highly coexpressed in hepatoma cells (Figure 2I). FB_3 markers (*THBS4* and *ACTG2*) were coexpressed in fibroblasts (Figure 2J).

The higher CNV_2 (*TFF2* and *AGR2*) and FB_3 (*THBS4* and *ACTG2*) expression was correlated with worse prognosis in the combined training cohort and the Xiangya HCC cohort (Supplemental Figures S4A, B, <http://links.lww.com/HEP/I372>).

Cancer stemness characteristics of HCC transcriptomic subtypes

Cancer stemness was closely correlated with cancer progression, prognosis, and treatment resistance.^[12]

TABLE 1 The univariate Cox regression analysis of 6 prognosis-related subclusters based on CIBERSORT proportion and SSS score

Subcluster	CIBERSORT proportion			SSS score		
	HRs (95% CI)	p	FDR	HRs (95% CI)	p	FDR
Treg_2	5800 (280–120,000)	2.20E–08	4.40E–07	1.5 (1.2–1.8)	5.00E–05	0.00017
CNV_4	0.086 (0.034–0.21)	1.20E–07	1.60E–06	0.72 (0.59–0.87)	0.00072	0.00198
FB_3	2.6e+08 (31,000–2.2e+12)	2.70E–05	0.00027	1.9 (1.5–2.6)	3.60E–06	1.98E–05
CNV_2	71 (9.2–540)	4.30E–05	0.00034	1.7 (1.4–2)	2.50E–09	2.75E–08
Mono	260,000 (300–2.2e+08)	3.00E–04	0.002	1.6 (1.3–1.9)	1.20E–05	5.66E–05
pDC	2.9e+21 (8.7e+07–4.9e+32)	0.0014	0.008	1.7 (1.3–2.2)	2.20E–05	9.08E–05

Abbreviations: FDR, false discovery rate; SSS, subcluster-specific score.

First, we analyze the cancer stemness of tumor cells at the single-cell level (CNV subclusters). CNV subclusters were annotated in the t-distributed stochastic neighbor embedding plot (Figure 3A). Non-negative matrix factorization identified 5 gene metaprograms, of which the hub-genes represented a cluster of coexpression genes across all CNV subclusters (Figure 3B and Supplemental Table S6, <http://links.lww.com/HEP/I382>, details see “Method”). Function enrichment analysis revealed metaprograms 3 and 5, which were significantly concentrated in CNV_2, were enriched in pathways of cell cycle, cancer stem cell, and TP53 signaling, as well as pathways of hypoxia, glycolysis, and epithelial mesenchymal transition. Whereas metaprograms 2 and 4, which were mainly upregulated in CNV_4, were enriched in coagulation and immune response (Figure 3B). We conducted scRNA-seq analysis to detect the expression of CD133, EpCAM, and cytokeratin 19 (KRT19), and surprisingly discovered that these well-known HCC stem cell markers were significantly enriched in the CNV_2 cluster (Figure 3C). CNV_2 also had the highest cancer stemness score of cancer stem cell (Palmer, 2012)^[13] and Stem.Sig (Zhang, 2022)^[14] (Figure 3D). Single-cell regulatory network inference and clustering (SCENIC) analysis revealed that regulons of TP53 and MYC, the critical transcription factors driving cancer stemness, were significantly activated in CNV_2 (Figure 3E). Cell trajectory analysis also demonstrated that CNV_2 was populated at the initial position of the developmental tree and had the lowest pseudotime and the highest CytoTRACE score (Figure 3F). All of this evidence indicated that CNV_2 had the strongest stemness among the CNV subclusters.

According to the inferred subcluster percentages in the training cohort, class 5 contained the highest proportion of the CNV_2 subcluster (Figure 3G). Consistently, the cancer stemness score was significantly elevated in class 5 (Figure 3H). Hallmark pathways in 5 transcriptomic subtypes were presented in Supplemental Figure S5A, <http://links.lww.com/HEP/I387>. Next, we reconstructed transcription factor regulatory networks in bulk RNA-seq data from the

training cohort. We noticed distinct regulon expression patterns across 5 transcriptomic subtypes (Supplemental Figure S5B, <http://links.lww.com/HEP/I387>). The significantly activated regulons of class 5, such as PITX2, MYBL2, and FANCB, were associated with cancer stemness, MYC targets, and TP53 signaling pathways (Figure 3I, Supplemental Figure S5C, <http://links.lww.com/HEP/I387>, and Supplemental Table S7, <http://links.lww.com/HEP/I383>). Then, we constructed protein-protein interaction networks (PPI) of stemness-related regulons activating in class 5, where TP53 and MYC were located at the core nodes (Figure 3J). We identified robustly coactivated regulons related to cancer stemness in CNV_2 and class 5, such as PITX2, MYBL2, and ELF3 which were reported (Figure 3K and Supplemental Figure S5D, <http://links.lww.com/HEP/I387>).^[15–17] Immunofluorescence staining confirmed the high expression of EPCAM, TP53, PITX2, and MYBL2 in class 5 compared to class 1 (Figure 3L). We further explored the expression of PITX2 and MYBL2 in HCC cells and discovered that PITX2 and MYBL2 were significantly upregulated in HCCLM3 and Hep3B cell lines (Supplemental Figures S6A, B, <http://links.lww.com/HEP/I373>). Then, we constructed 3 shRNA to knock down PITX2 and MYBL2 in HCCLM3 and Hep3B cell lines, respectively. RT-qPCR and western blot analysis were used to assess the knockdown efficiency of PITX2 and MYBL2 (Supplemental Figures S6C, D, <http://links.lww.com/HEP/I373>). Downregulated PITX2 and MYBL2 decreased the size and number of spheroids formed in the HCCLM3 and Hep3B cells (Figure 3M).

Immune profiles of HCC transcriptomic subtypes

We further analyzed the genomic profiles of 80 immune-related genes (37 immunostimulatory genes, 20 inhibitory immune checkpoints, 12 major histocompatibility complex class I genes and 11 major histocompatibility complex class II genes) (Supplemental Figure S7A, <http://links.lww.com/HEP/I388>). Subsequently, we applied the

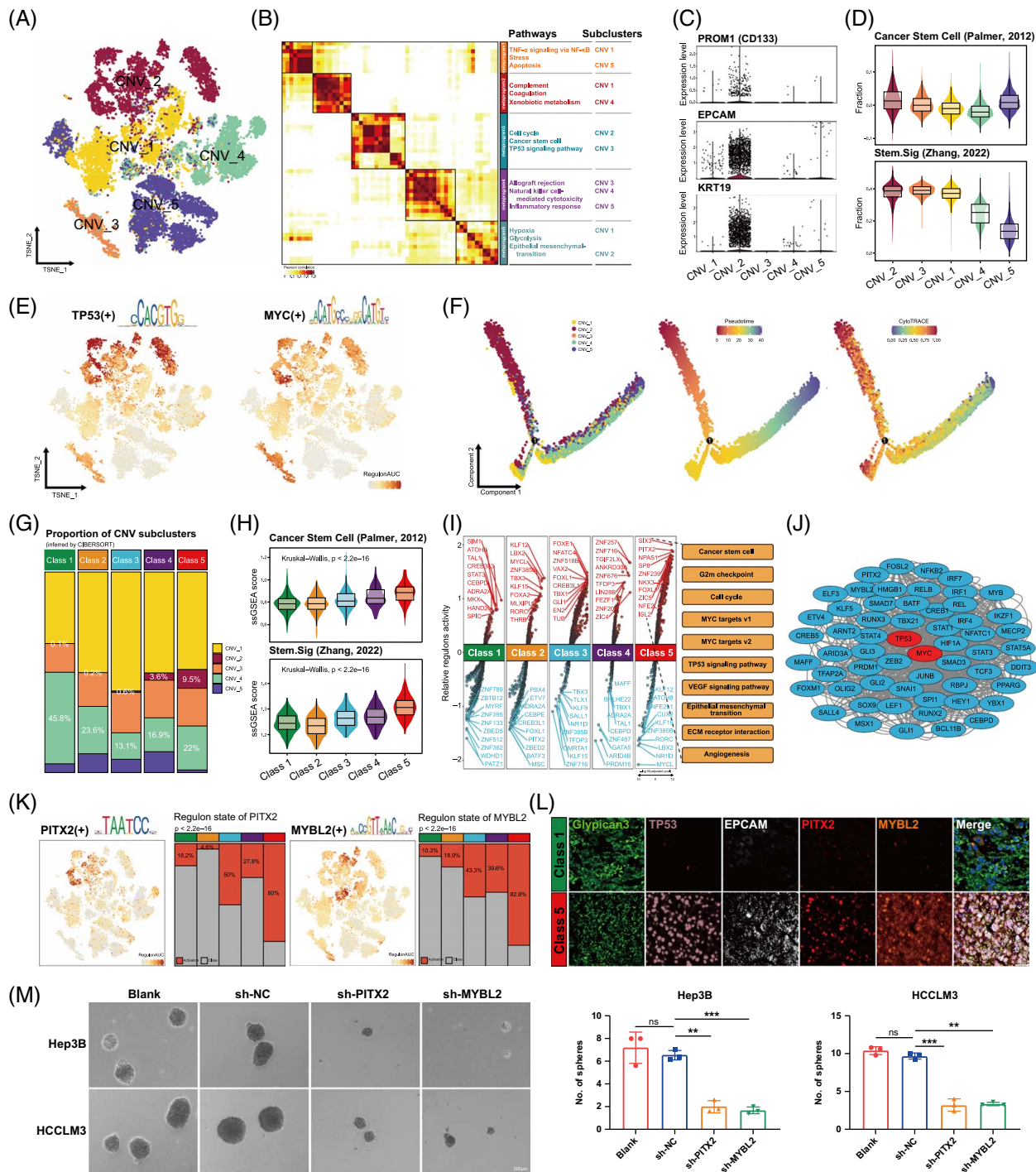


FIGURE 3 Correlation between transcriptomic subtypes and cancer stemness. (A) t-SNE visualization of CNV subclusters in tumor cells. (B) Hierarchical clustering of NMF programs identified 5 prominent metaprograms in epithelial subclusters. (C) The expression of HCC stem cell markers *CD133*, *EPCAM*, and *KRT19* in 5 CNV subclusters. (D) Violin plot displaying the stemness scores of cancer stem cell (Palmer, 2012) and Stem.Sig (Zhang, 2022) across CNV subclusters. (E) t-SNE plot showing the regulon activity of *TP53* and *MYC*. (F) Differentiation trajectory of CNV subclusters (left) displaying pseudotime (middle) and CytoTRACE score (right). Colored dots indicate cells annotated in (A). (G) The proportion of CNV subclusters in distinct transcriptomic classes. (H) Stemness scores of cancer stem cells (Palmer, 2012) and Stem.Sig (Zhang, 2022) for 5 transcriptomic classes. (I) Specific activated or suppressed regulons for 5 transcriptomic classes and the enrichment pathways of activated regulons of class 5 (right). (J) The hub PPI network of stemness-related regulons in class 5. (K) The activities of *PITX2* and *MYBL2* regulons in CNV subclusters (left) and HCC transcriptomic classes (right). (L) Multiplex immunofluorescence staining of stemness regulons including *TP53*, *EPCAM*, *PITX2*, and *MYBL2* in tumor cells of HCC (marked Glypican3). Scale bars = 50 μ m. (M) Sphere-forming assays of Hep3B and HCCLM3 cells after *PITX2* and *MYBL2* knockdown. Representative images are shown, and the number of spheroids is quantified in bar graphs. Scale bars = 200 μ m. ** p < 0.01, *** p < 0.001. Abbreviations: CNV, copy number variation; NMF, non-negative matrix factorization; PPI, protein-protein interaction; t-SNE, t-distributed stochastic neighbor embedding.

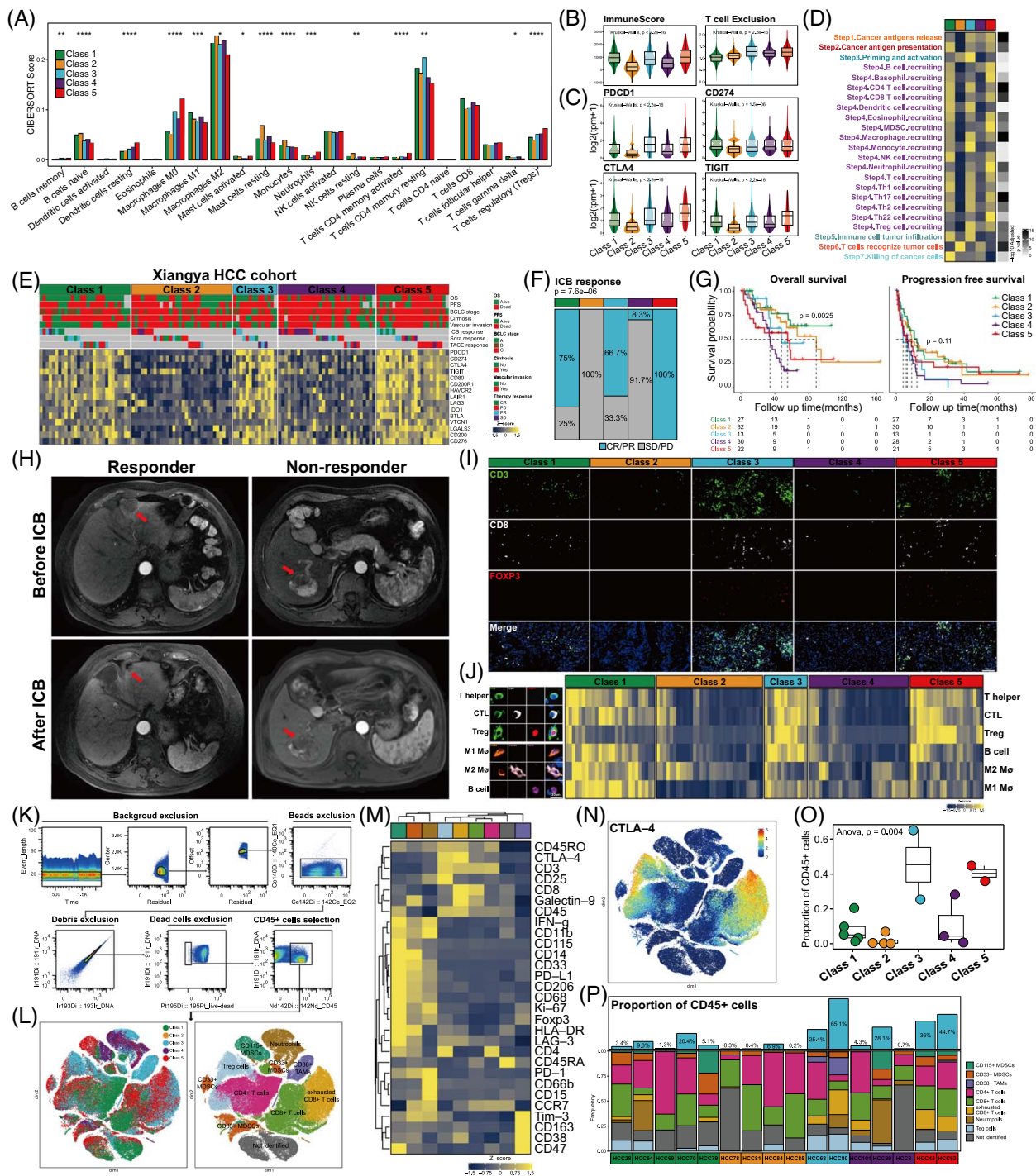


FIGURE 4 Immune characteristics of HCC transcriptomic subtypes in the training cohort and the Xiangya real-world cohort. (A) Comparing the infiltration level of 22 immune cell types in 5 distinct transcriptomic classes. (B) Violin plot of immune score and T-cell exclusion score in 5 transcriptomic classes. (C) Violin plot of the expression of inhibitory immune checkpoint genes *PDCD1*, *CD274*, *CTLA4*, and *TIGIT* in 5 transcriptomic classes. (D) Activities of different cancer immunity cycle steps in transcriptomic subtypes. (E) Heatmap showing the overall status, BCLC stage, cirrhosis, vascular invasion, ICB response, targeted therapy response, TACE response, and the expression of 15 immune checkpoint genes in the transcriptomic classes of the Xiangya HCC cohort. (F) Overall response rate of 5 transcriptomic classes after ICB therapy in the Xiangya real-world cohort (chi-square test, $p = 7.6 \times 10^{-6}$). (G) Kaplan-Meier survival curves displaying the OS (log-rank test, $p = 0.0025$) and PFS (log-rank, $p = 0.11$) of 5 transcriptomic subtypes in the Xiangya real-world cohort. (H) Arterial-phase MRI image displayed the differences in individual immunotherapy responses. The red arrow points to the tumor area. (I) Multiplex immunofluorescence staining in different transcriptomic classes with magnifications of T-helper cells (CD3+ and FOXP3-), cytotoxic T lymphocytes (CTL; CD3+, CD8+, and FOXP3-), and regulatory T cells (Treg; CD3+, CD8-, and FOXP3+). Scale bar = 100 μ m. (J) The heatmap showed the association between the immune cell

infiltration percentage and the transcriptomic classes among the Xiangya HCC cohort. Annotation of immune cells based on canonical markers. (K) Gating strategy for identifying CD45 immune cell populations. (L) Dimplot of transcriptomic subtypes and cell clusters based on CyTOF. (M) A heatmap showing the differential expression of 29 immune markers in the 9 cell clusters. (N) The distribution of immune checkpoint CTLA4 in different cell clusters. (O) The proportion of CD45 immune cells in the 5 HCC transcriptomic subtypes (ANOVA, $p=0.0062$). (P) Frequency diagram of immune cell subclusters in each patient with different transcriptomic subtypes. Abbreviations: BCLC, Barcelona Clinic Liver Cancer; ICB, immune checkpoint blockade; OS, overall survival; TACE, transarterial chemoembolization.

CIBERSORT algorithm and the immune cells gene set curated by Charoentong et al^[18] to investigate the infiltration level of immune cells. Class 5 showed a significantly increased proportion of myeloid-derived suppressor cells (MDSCs) and regulatory T cells ($p<0.001$, Figure 4A and Supplemental Figure S7B, <http://links.lww.com/HEP/I388>). Classes 5 and 3 displayed higher levels of immune score, T-cell exclusion score (Figure 4B), and immunosuppressive markers (such as *PDCD1*, *CD274*, *CTLA4*, and *TIGIT*, Figure 4C). We further explored the steps of the cancer immune cycle, including tumor antigen presentation, immune cell recruitment, tumor cells recognition, and killing.^[19] The activities of most steps, such as cancer antigen presentation, CD8+ T-cell recruiting, and killing of cancer cells, were significantly upregulated in class 1, while these steps were significantly suppressed in classes 2 and 4. Recruitment of MDSCs, neutrophil, Th22, and Treg cells were significantly increased in classes 3 and 5. However, the downregulated activities of T cells recognizing and killing tumor cells were observed in class 5 (Figure 4D). All of these findings indicated that class 1 might be an inflamed phenotype, classes 2 and 4 might be immune-deserted phenotypes, and classes 5 and 3 might be inhibitory tumor immune microenvironments. Consistently, responsiveness of ICB predicted from the TIDE website showed that classes 1, 3, and 5 had higher ICB reactivity (chi-square test, $p=3.9e-06$, Supplemental Figure S7C, <http://links.lww.com/HEP/I388>).

We then validated the immune characteristics and ICB response of transcriptomic subtypes in the Xiangya HCC cohort. As expected, the expression of inhibitory immune checkpoints was higher in classes 3 and 5 compared to classes 2 and 4. The patients in class 5 also presented higher vascular invasion, Barcelona Clinic Liver Cancer stage, and cancer stemness score (Figure 4E and Supplemental Figure S7D, <http://links.lww.com/HEP/I388>).

Surprisingly, the response rate of patients in class 1 (75%; 95% CI: 54.1%–100%), class 3 (66.7%; 95% CI: 9.40%–99.20%), and class 5 (100%; 95% CI: 54.1%–100%) who received ICB therapy was significantly higher than that of patients in class 2 (0%; 95% CI: 0%–41%) and class 4 (8.3%; 95% CI: 0.2%–38.5%) (Figure 4F). These results were consistent with the immune profile of transcriptomic subtypes characterized by the training cohort. Due to the excellent ICB responsiveness of class 5 patients, their overall survival was improved, whereas class 4 patients had the worse prognosis in the Xiangya

real-world cohort (Figure 4G, log-rank test, $p=0.0025$). We did not observe significant PFS differences between the 5 transcriptomic subtypes (Figure 4G, log-rank test, $p=0.11$). The therapy response was assessed through an enhancement MRI scan (Figure 4H). Furthermore, we evaluated the immune cell infiltrations of transcriptomic subtypes, including T helper cells, CTLs, Tregs, B cells, M1, and M2 macrophages, through the corresponding tissue sections (Figures 4I, J). The level of infiltrating immune cells was increased in classes 3 and 5 compared to classes 2 and 4, especially CTLs and Tregs cells.

Subsequently, we conducted CyTOF analysis on 16 collected samples derived from the Xiangya HCC cohort. All CyTOF data were preprocessed, and living single immune cells (CD45+) were retained after gating for further analysis (Figure 4K). Based on the canonical cell markers, we defined 8 major cell clusters, including CD8+ T cells, exhausted CD8+ T cells, Treg cells, CD4+ T cells, CD33+ MDSCs, neutrophils, CD38+ tumor-associated macrophages, and CD115+ tumor-associated macrophages (Figures 4L, M and Supplemental Figure S8A, <http://links.lww.com/HEP/I389>). The expression of immune checkpoint CTLA4 was mainly localized in exhausted CD8+ T cells and Treg cells (Figure 4N). Compared to classes 2 and 4 subtypes, the samples with classes 3 and 5 subtypes were characterized by a high proportion of immune cell infiltration (Figure 4O). The analysis of the density of infiltrating cell clusters in these samples revealed that class 1 was characterized by the presence of CD4+ and CD8+ cells, classes 2 and 4 were characterized by a lack of CD45+ cells. Classes 3 and 5 samples showed higher levels of infiltration by Treg cells, exhausted CD8+ T cells, MDSCs, and tumor-associated macrophages, all of which possessed remarkable immunosuppressive capabilities (Figure 4P). These results are consistent with previous findings.

scRNA-seq and immune repertoire sequencing revealed the immune microenvironment and T/B-cell receptor clonotypes in transcriptomic subtypes

We conducted paired scRNA, scTCR, and scBCR sequencing on the operative specimen of 6 patients with HCC, which were classified as class 1, class 3, class 4, and class 5 based on the corresponding bulk-RNA sequencing. After quality control and dimensionality

reduction clustering, we annotated 10 main cell types (Figures 5A, B and Supplemental Figure S8B, <http://links.lww.com/HEP/I389>). The FB_3 cluster (in fibroblasts) and the CNV_2 cluster (in epithelial cells) were enriched in our identified transcriptomic subtypes (Supplemental Figure S9A, <http://links.lww.com/HEP/I374>). We also observed that the myeloid cells of classes 4 and 5 expressed elevated levels of immunosuppressive macrophage markers (*FOLR2*, *SPP1*, and *SELENOP*).^[20,21] In addition, major histocompatibility complex class II molecules (*HLA-DRB1*, *HLA-DQA1*, and *HLA-DQB1*) were downregulated in class 5, indicating a reduced capacity of tumor antigens presentation from macrophages to T cells in class 5 (Figure 5C). We then explored the phenotype and clonotype characteristics of T cells and B cells in the 4 subtypes. We identified 12,931 T cells as 5 distinct phenotypes according to the expression of canonical phenotypical markers (Figures 5D, E). Faster RNA velocity and higher mRNA expression of *FOXP3* and *CTLA4* were presented in Treg cells (Figure 5F).

We further identified top 10 T cell receptor clonotypes by leveraging paired scTCR and scRNA sequencing (Figure 5G and Supplemental Table S8, <http://links.lww.com/HEP/I384>). Class 3 was mainly characterized by clonotype 1 and clonotype 2, and class 5 was characterized by diverse clonotypes (Figure 5H). The clonal TCRs were mainly distributed in CD8 CTL and exhausted CD8 cells (Figure 5I). T-cell functions of distinct transcriptomic subtypes and T cell receptor clonal expansion were scored. Intriguingly, expanded T cells in class 5 showed an enhancement of cytotoxicity and exhausted signal accompanied by a weakened costimulatory signal, which was the second signal triggering T-cell activation (Figure 5J). This may be a possible reason why class 5 had lack of activity killing the tumor cells, as T cell receptor clonotypes were unable to fully activate and exert their cytotoxic function. A high level of clonal dominance was presented in class 1, class 3, and class 5. Large exhausted CD8 cells and Treg cell proportion existed in class 3 and class 5, whereas class 4 lacked T-cell infiltration and expansion (Figures 5K, L). Researchers reported that exhausted T cells were a major cell type responding to PD1 blockade and the stage of exhausted T cells was significantly correlated with responsiveness of ICB therapy.^[22] We discovered that the feature of precursor-exhausted CD8 T cells (Ly108⁺/CD69⁻) in classes 5 and 3 was increased, which indicated with better efficacy of ICB therapy (Figure 5M). Immunofluorescence staining also revealed the high expression of PD1/PD-L1 in classes 3 and 5 (Figure 5N).

In addition, 4 distinct B cell receptor (BCR) phenotypes were identified according to the expression of canonical phenotypical markers (Supplemental Figures S9B, C, <http://links.lww.com/HEP/I374>). We further integrated B-cell phenotypes and BCR clonotypes (Supplemental Figure S9D, <http://links.lww.com/HEP/I374>) and found that the clonal expansion of BCR mainly existed in IgA⁺ and IgG⁺ plasma cells (Supplemental Figure S9E, <http://links.lww.com/HEP/I374>). The distribution of BCR phenotypes and clonotypes in transcriptomic subtypes were presented in Supplemental Figures S9F, G, <http://links.lww.com/HEP/I374>.

Development of prognosis-related score based on machine learning

First, using the bootstrapping method, we identified 25 SSMs from 103 SSMs of prognosis-related subclusters. Subsequently, we performed a random survival forest to screen the identified SSMs and confirmed 8 genes with the best performance. Finally, a robust prognosis-associated model was constructed by calculating the multivariate Cox regression coefficient of these 8 genes (Figure 6A and Supplemental Table S9, <http://links.lww.com/HEP/I385>). This model performed well in prognostic prediction in the combined training cohort, microarray validation cohort, RNA-seq validation cohort, and Xiangya HCC cohort (Figure 6B). Figure 6C shows the correlation between the transcriptional subtypes and prognosis-related score (PRS) groups. The AUC value of PRS was superior to conventional clinical features, such as age, TNM stage, and Barcelona Clinic Liver Cancer stage, in the TCGA-LIHC, CHCC-HBV, and LIRI-JP cohorts (Figure 6D). The univariate and multivariate Cox regression analyses revealed that PRS was an independent prognostic factor in the TCGA-LIHC, CHCC-HBV, and LIRI-JP cohorts (Figure 6E). These findings confirmed the excellent prognostic value of PRS in patients with HCC.

Next, we analyzed the correlation between PRS and activities of hallmark and PROGENy pathways, cancer stemness scores, and stemness-related regulons. The hallmark pathways of proliferation, DNA repair, *MYC* and *TP53* signaling, PROGENy pathways of hypoxia, *EGFR* and *TP53* signaling, stemness-related regulons, and cancer stemness scores were significantly positively correlated with PRS (Figure 6F). In the TCGA-LIHC and CHCC-HBV cohorts, more advanced clinical stages (including histological grade, Barcelona Clinic Liver Cancer, and TNM staging systems), extensive vascular invasion, and increased AFP levels were found to be correlated with higher PRS. Interestingly, a higher mutation frequency of *TP53* was observed along with the increasing PRS (Figures 6G, H).

We also discovered that PRS was significantly associated with immune checkpoint expression (Figure 6I). The expression of *PDCD1* and *CTLA4* was upregulated in the high PRS group (Figure 6J). CytoF analysis suggested that the proportion of CD45⁺ cells was also elevated in the high PRS group (Figure 6K). Our Xiangya HCC cohort indicated that the patients with higher PRS were more sensitive to ICB

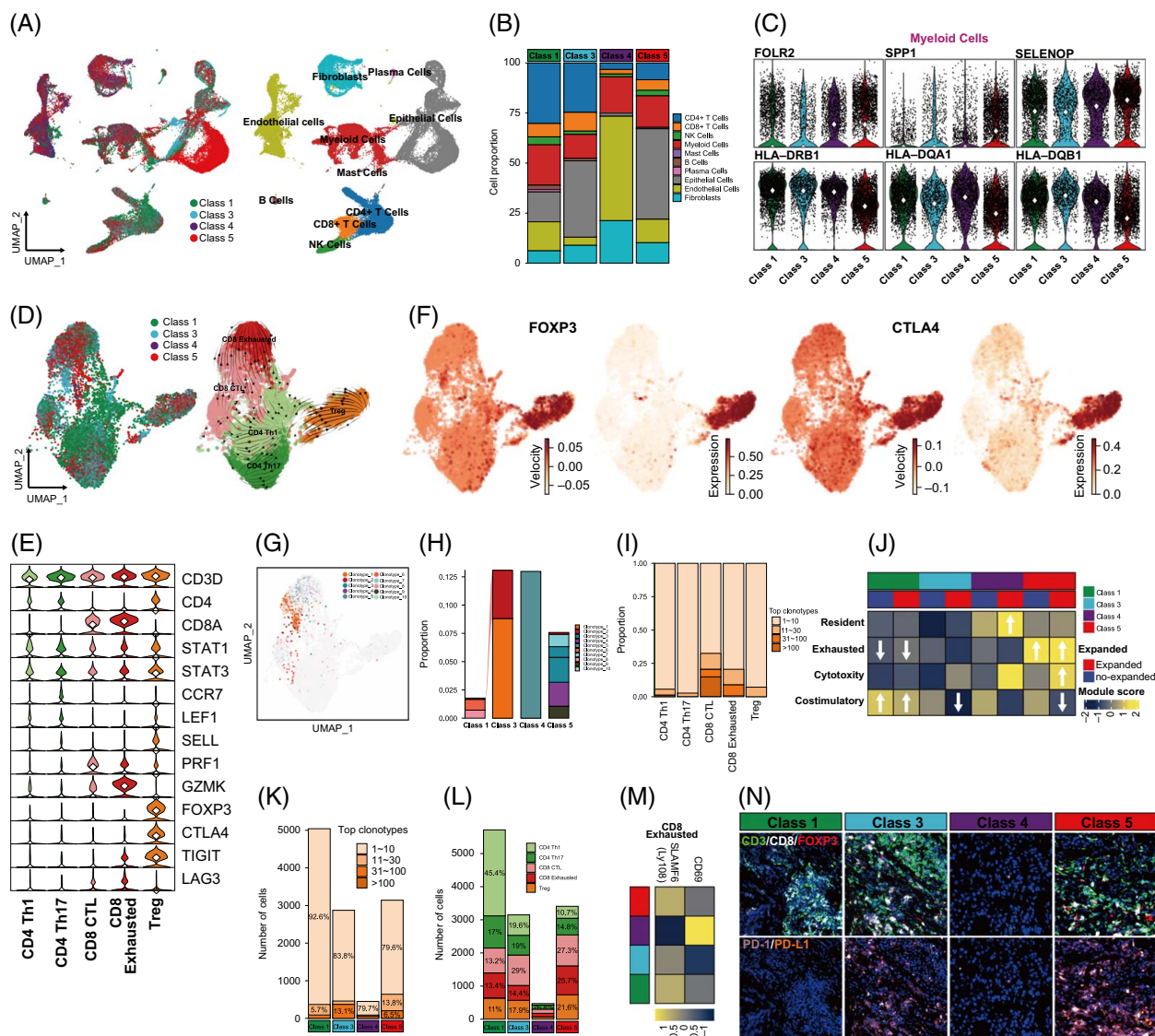
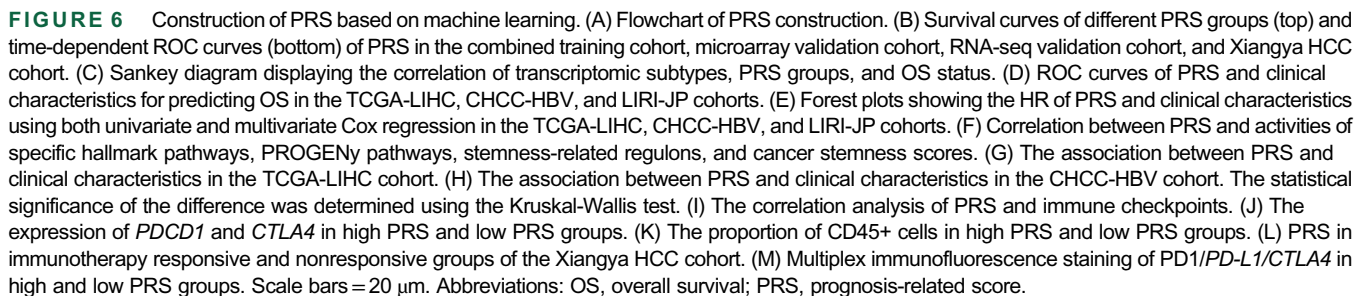


FIGURE 5 Single-cell RNA-seq and scTCR/BCR-seq revealed the immune repertoire of transcriptomic subtypes. (A) UMAP visualization annotated with 10 major cell types in classes 1, 3, 4, and 5. (B) The proportion of 10 major cell types in different transcriptomic subtypes. (C) The expression of immunosuppressive markers (*FOLR2*, *SPP1*, and *SELENOP*) and MHC class II genes (*HLA-DRB1*, *HLA-DQA1*, and *HLA-DQB1*) in myeloid cells across identified transcriptomic subtypes. (D) UMAP visualization of T-cell phenotypes and steady-state RNA velocity of T-cell phenotypes. (E) RNA velocity and expression of *FOXP3* and *CTLA4*. (F) Violin plot displaying the expression of canonical markers in distinct T-cell phenotypes. (G) UMAP plot annotated with TCR clonotypes. (H) The proportion of 10 TCR clonotypes in different transcriptomic subtypes. (I) TCR expansion frequency in distinct T-cell phenotypes. (J) Resident, exhausted, cytotoxicity, and costimulatory scores in T cells of different transcriptomic subtypes. (K) TCR expanding frequency in transcriptomic subtypes. (L) The percentage of T-cell phenotypes in transcriptomic subtypes. (M) The expression pattern of precursor-exhausted CD8 T cells (Ly108+CD69-) and terminal exhausted CD8 T cells (Ly108-CD69+) in different transcriptomic subtypes. (N) Multiplex immunofluorescence staining of CD3/CD8/FOXP3/PD1/PD-L1 in different transcriptomic subtypes. Scale bars = 50 μ m. Abbreviations: BCR, B cell receptor; MHC, major histocompatibility complex; TCR, T cell receptor; UMAP, uniform manifold approximation and projection.

therapy (Figure 6L). The expression of *PD1*, *PD-L1*, and *CTLA4* in the high PRS and low PRS groups of the Xiangya HCC cohort is presented in Figure 6M.

To formulate a clinically relevant quantitative method for predicting the probability of 1-year, 3-year, and 5-year overall survival in patients with HCC, we constructed a nomogram combining PRS and TNM stages, which was also an independent clinical prognostic factor in the training cohort (Supplemental

Figure S10A, <http://links.lww.com/HEP/I375>). The calibration curves of multiple cohorts suggested that the nomogram had a similar performance compared to an ideal model for predicting overall survival probability (Supplemental Figure S10B, <http://links.lww.com/HEP/I375>). The subgroup Cox regression of PRS was presented in the TCGA-LIHC, CHCC-HBV, LIRI-JP cohorts (Supplemental Figures S10C–E, <http://links.lww.com/HEP/I375>).



Exploration of potential drugs and targets for HCC transcriptomic subtypes

To explore whether constructed transcriptomic subtypes had different responses to systemic therapies, 2 treatment cohorts of HCC, including patients who received sorafenib therapy (GSE109211) and TACE (GSE104580), were retrieved. Support vector machine was used to classify the transcriptomic subtypes of these patients. Surprisingly, classes 3 and 5 were much more sensitive to sorafenib compared to other transcriptomic subtypes, with response rates of 77.8% (95% CI: 40%–97.2%) and 81.8% (95% CI: 48.2%–97.7%), respectively (Figure 7A, chi-square test, $p=7.8e-7$). Whereas the sensitivity of TACE therapy declined in class 5 (Figure 7B). We further predicted the IC50 value of sorafenib among 5 subtypes in the combined training cohort based on 2 different databases CGP2016 and CTRP2, which revealed that classes 3 and 5 had the lowest IC50 (Figure 7C). Subsequently, 2 methods were used to screen potentially effective compounds for the high PRS group (Spearman $r < -0.40$), and 6 CGP2016-derived compounds and 7 CTRP2-derived compounds were identified (Figure 7D). All of these compounds had lower IC50 values in the high PRS group (Figure 7E). The CMap score of these compounds on HCC cell lines revealed that Tipifarnib, JW-74, and Etoposide were possibly beneficial to the high PRS group (Figure 7F).

Next, we screened the therapeutic drug targets correlated with PRS. Correlation analysis between the expression of druggable genes and PRS was conducted in the training cohort, and 221 targets were identified with a correlation coefficient > 0.5 ($p < 0.05$, Figures 7G, H). We further confirmed 27 druggable targets based on the correlation analysis between CERES score and PRS in HCC cell lines with a correlation coefficient < -0.5 ($p < 0.05$, Figures 7G, H). Four targets, including ACLY, HK2, PPIA, and PPIH, were identified through both aforementioned analyses (Figure 7G). We further validated the affinity of the predicted small molecular compound to the 4 targets through molecular docking analysis (Supplemental Figure S11, <http://links.lww.com/HEP/I376>). Significant correlations were observed between PRS and the mRNA expression and CERES scores of the 4 intersected targets (Figure 7I). The pathways of hypoxia, MAPK, EGFR, VEGF, and TP53 signaling were remarkably activated in class 5, which also indicates the potential interventional pathways (Figure 7J).

DISCUSSION

TME heterogeneity is the pivotal determinant for the difference in prognosis and drug sensitivity in patients

with advanced HCC.^[23] The widespread application of scRNA sequencing has revealed the characteristics of malignant cells, tumor-associated stromal cells, and tumor immune infiltration, which offer novel insights for tumor classification, tumor initiation and evolution, mechanisms of drug resistance, and the selection of therapeutic agents.^[24] In this study, we integrated the scRNA-seq data of 33 patients with HCC to reveal distinct cell subclusters and observed significant inter-patient and intratumor heterogeneity. CNV analysis showed malignant cell subclusters exhibited anomalous increases or decreases in copy numbers. Six cell subclusters were identified, including CNV_2, CNV_4, FB_3, Mono, pDC, and Treg_2, which played crucial roles in the clinical prognosis for HCC.

Unsupervised clustering further identified 5 HCC transcriptomic subtypes with different overall survival rates. Studies have shown that the exploration of tumor metastasis and recurrence at the transcriptomic level could promote individualized treatment of tumors and accelerate drug development.^[25,26] In addition, researchers considered that HCC could be roughly categorized into 2 major molecular subtypes (proliferation class and nonproliferation class) regardless of the particular nomenclature applied to each class.^[27,28] These HCC classifications constructed by the scholars, as mentioned before, partially explained the differences in patient prognosis or treatment response. Our study integrated single-cell sequencing data from multiple HCC cohorts and constructed transcriptomic subtypes applying specific markers of prognosis-related subclusters, which give enough consideration to the intra-tumoral and interindividual heterogeneity at the single-cell level. The results should be reliable because we validated the prognostic value of these subtypes in multiple bulk RNA-seq and microarray cohorts, as well as real-world large clinical samples.

Our study further described the biological characteristics of the 5 subtypes. Tumor stemness is also an important factor causing HCC heterogeneity and is closely associated with HCC relapse and drug resistance.^[29] Various cancer stemness markers have been identified, including PITX2, MYBL2, MYC, SOX9, ELF3, and KLF5, etc.^[30–32] In this study, the class 5 subtype with the highest proportion of CNV_2 subcluster had a remarkably upregulated cancer stemness score, which may explain the reason for the poorest prognosis in class 5.

Although multiple clinical trials suggest the significant efficacy of immunotherapy in the treatment of malignant tumors, not all patients respond positively to this therapy.^[33] The distinct molecular characteristics of the transcriptomic classes indicate that specific therapies may have different effects on these patients. Only by accurately resolving the types and status of immune cells in TME can a systematic tumor immune profile be obtained, thereby assisting in the discovery of effective

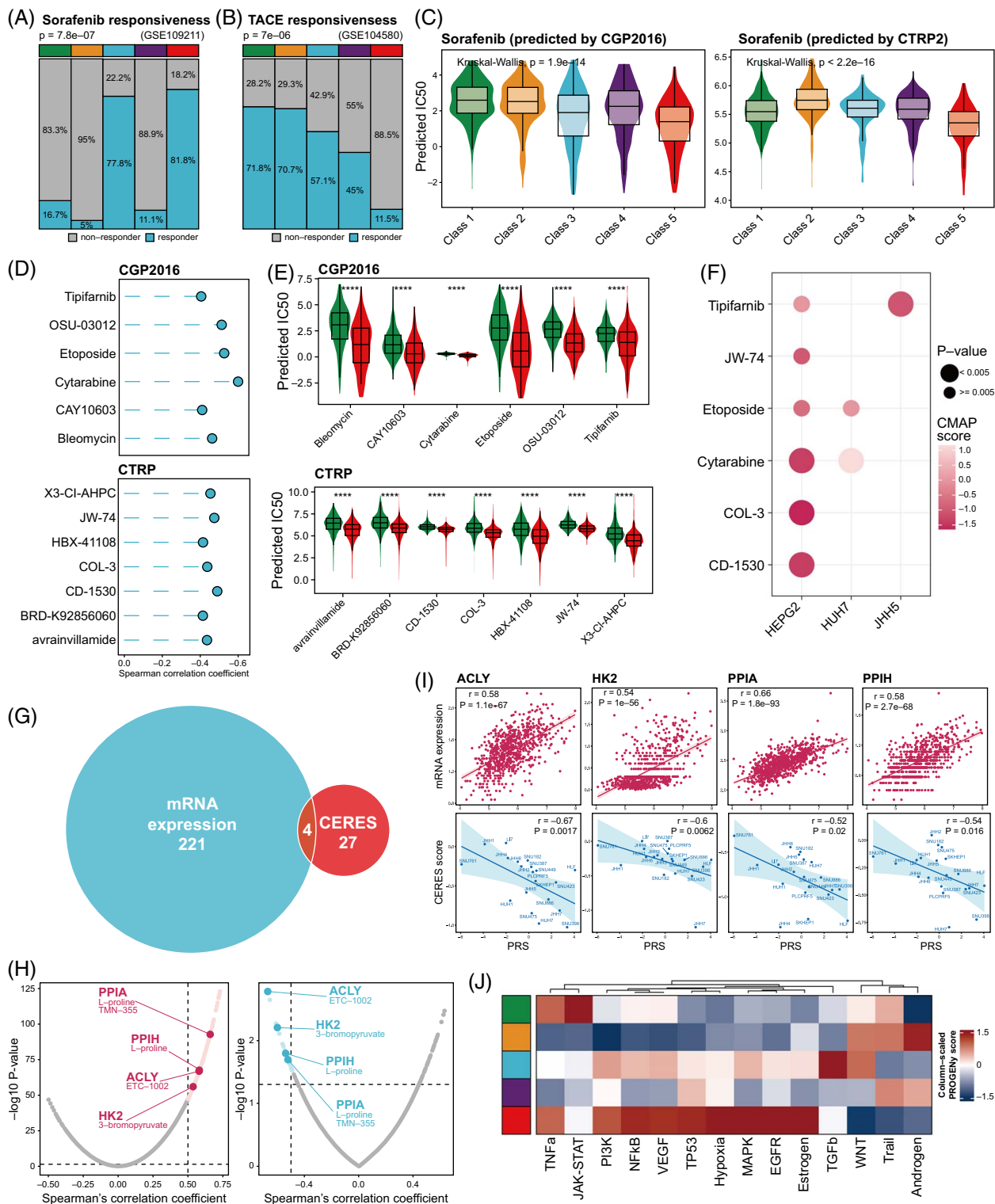


FIGURE 7 Identification of potential drugs, therapeutic targets, and pathways for HCC transcriptomic subtypes and PRS. (A) The response rate of 5 transcriptomic classes after sorafenib treatment (chi-square test, $p = 7.8 \times 10^{-7}$). (B) The response rate of 5 transcriptomic classes after TACE treatment (chi-square test, $p = 7 \times 10^{-6}$). (C) Violin plot showing predicted IC50 value of sorafenib in 5 transcriptomic classes based on CGP2016 (Kruskal-Wallis, $p = 1.9 \times 10^{-14}$) and CTRP2 database (Kruskal-Wallis, $p < 2.2 \times 10^{-16}$). (D) Spearman correlation analysis of PRS and 6 CGP2016-predicted compounds (top) and 7 CTRP-predicted compounds (bottom). (E) Violin plot displaying the differences in IC50 values of 6 CGP2016-predicted compounds and 7 CTRP-predicted compounds between high PRS and low PRS groups. (F) The CMap score of screened compounds in multiple HCC cell lines. (G) Venn diagram for intersecting the druggable targets based on mRNA expression and CERES score. (H) Volcano plot of Spearman correlation analysis between PRS and mRNA expression (left) or CERES score (right) of drug targets. Red and blue dots indicate significant correlations ($p < 0.05$ and Spearman $r > 0.5$ or < -0.5). (I) Scatter plot of PRS with mRNA expression (top) and CERES score (bottom) of 4 intersected targets. (J) PROGENy pathway activities in 5 transcriptomic subtypes. Abbreviations: PRS, prognosis-related score; TACE, transarterial chemoembolization.

tumor immune biomarkers and the development of tumor immunotherapy. The 5 transcriptomic subtypes that we identified demonstrated significant discrepant immune phenotypes. Immune infiltration analysis and the steps of the cancer immune cycle indicated that class 1 might be an inflamed phenotype, classes 2 and 4 might be immune-deserted phenotypes lacking T-cell infiltration, while classes 5 and 3 might be inhibitory tumor immune microenvironment-enriched Treg cells and suppressive immune checkpoints. Our single-cell T cell receptor/B cell receptor-sequencing sequencing and CyTOF analysis further validated these findings. The description of the tumor immune microenvironment in transcriptomic subtypes indicates that classes 1, 3, and 5 might possess positive reactivity to ICB. Notably, our Xiangya real-world HCC cohort showed that the response rate of classes 1, 3, and 5 patients reached 75%, 66.7%, and 100% after receiving ICB therapy, which is consistent with the immune characteristics of these transcriptomic subtypes.

Machine learning has been gradually applied to prognosis research.^[34] In our present study, we generated a risk score (PRS) to predict the overall survival of patients with HCC based on the random survival forest. PRS performed well in prognostic prediction and was correlated with ICB response. In addition, 4 potential therapeutic targets (ACLY, HK2, PPIA, and PPIH) and corresponding compounds were identified for patients with high PRS with HCC. These targets were closely associated with the proliferation, metastasis, metabolism, and immune infiltration of cancer.^[35–37] The evidences suggest the feasibility of developing corresponding agents and targeting proteins for patients with high PRS with HCC.

In conclusion, the comprehensive single-cell analysis offers a high-resolution depiction of the TME heterogeneity in HCC and confirms 6 cellular subclusters of prognostic relevance. Furthermore, 5 transcriptomic subtypes were constructed with different clinical prognosis, stemness characteristics, immune landscape, and therapeutic responses through integrating single-cell data and bulk cohorts. Our work provides a robust classification system for prognostic prediction and personalized medicine. Nevertheless, some limitations of this study should be acknowledged. First, this research was conducted retrospectively, and prospective studies should be designed to further validate the efficacy of categorization. In addition, the cohorts incorporated in this study exhibited variations in etiologies and clinical characteristics. A multicenter, large-sample data set encompassing eligible patients undergoing distinct treatments (targeted therapy, TACE, and immunotherapy) should be further implemented to evaluate clinical effectiveness. Finally, the potential therapeutic targets and agents identified in this study should be subjected to further investigation through experimental validation and clinical trial research.

AUTHOR CONTRIBUTIONS

Changjing Cai, Yiming Qin, and Hong Shen designed and supervised this study. Yihong Chen and Yiming Qin performed the bioinformatics analysis and experiments. Xiangying Deng, Yin Li, Ying Han, Yinghui Peng, Wantao Wu, and Xinwen Wang performed the literature search and collected the data. Jiayao Ma, Erya Hu, Xin Zhou, Edward Shen, and Shan Zeng analyzed the data. Yihong Chen and Yiming Qin wrote and revised the manuscript accordingly. All authors have contributed to the manuscript.

FUNDING INFORMATION

This study was supported by grants from the National Natural Science Foundation of China (nos. 82373275, 81974384, 82173342, and 82203015), the China Postdoctoral Science Foundation (no. 2023JJ40942), 4 projects from the Natural Science Foundation of Hunan Province (nos. 2021JJ3109, 2021JJ31048, 2024JJ6662, and 2023JJ40942), the Natural Science Foundation of Changsha (no. 73201), 2 projects from the CSCO Cancer Research Foundation (no. Y-HR2019-0182 and Y-2019Genecast-043), the Key Research and Development Program of Hainan Province (no. ZDYF2020228 and ZDYF2020125), the Youth Science Foundation of Xiangya Hospital (2023Q01), the Fundamental Research Funds for postgraduate of Hunan province (CX20230379), and the Fundamental Research Funds of Central South University (2021zzts0357).

CONFLICTS OF INTEREST

The authors have no conflicts to report.

ORCID

Hong Shen  <https://orcid.org/0000-0002-6456-8231>

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209–49.
2. Ahmad M, Dhasmana A, Harne PS, Zamir A, Hafeez BB. Chemokine clouding and liver cancer heterogeneity: Does it impact clinical outcomes? *Semin Cancer Biol*. 2022;86:1175–85.
3. Zeng Q, Mousa M, Nadukkandy AS, Franssens L, Alnaqbi H, Alshamsi FY, et al. Understanding tumour endothelial cell heterogeneity and function from single-cell omics. *Nat Rev Cancer*. 2023;23:544–64.
4. Ma L, Hernandez MO, Zhao Y, Mehta M, Tran B, Kelly M, et al. Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer. *Cancer Cell*. 2019;36:418–430.e416.
5. Gao R, Bai S, Henderson YC, Lin Y, Schalck A, Yan Y, et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat Biotechnol*. 2021;39:599–608.
6. Gao Q, Zhu H, Dong L, Shi W, Chen R, Song Z, et al. Integrated proteogenomic characterization of HBV-related hepatocellular carcinoma. *Cell*. 2019;179:561–577.e522.
7. Gong J, Li R, Chen Y, Zhuo Z, Chen S, Cao J, et al. HCC subtypes based on the activity changes of immunologic and

- hallmark gene sets in tumor and nontumor tissues. *Brief Bioinform.* 2021;22:bbaa427.
8. Llovet J M, Lencioni R. mRECIST for HCC: Performance and novel refinements. *J Hepatol.* 2020;72:288–306.
 9. Ally A, Balasundaram M, Carlsen R, Chuah E, Clarke A, Dhalla N, et al. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell.* 2017;169:1327–341.e1323.
 10. Xu C, Xu Z, Zhang Y, Evert M, Calvisi DF, Chen X. β -Catenin signaling in hepatocellular carcinoma. *J Clin Invest.* 2022;132:e154515.
 11. Ng CKY, Dazert E, Boldanova T, Coto-Llerena M, Nuciforo S, Ercan C, et al. Integrative proteogenomic characterization of hepatocellular carcinoma across etiologies and stages. *Nat Commun.* 2022;13:2436.
 12. Lee TKW, Guan X Y, Ma S. Cancer stem cells in hepatocellular carcinoma—From origin to clinical implications. *Nat Rev Gastroenterol Hepatol.* 2022;19:26–44.
 13. Palmer NP, Schmid PR, Berger B, Kohane IS. A gene expression profile of stem cell pluripotentiality and differentiation is conserved across diverse solid and hematopoietic cancers. *Genome Biol.* 2012;13:R71.
 14. Zhang Z, Wang ZX, Chen YX, Wu HX, Yin L, Zhao Q, et al. Integrated analysis of single-cell and bulk RNA sequencing data reveals a pan-cancer stemness signature predicting immunotherapy response. *Genome Med.* 2022;14:45.
 15. Ali SA, Justilien V, Jamieson L, Murray NR, Fields AP. Protein kinase C α drives a NOTCH3-dependent stem-like phenotype in mutant KRAS lung adenocarcinoma. *Cancer Cell.* 2016;29:367–78.
 16. Ward C, Volpe G, Cauchy P, Ptasinska A, Almaghrabi R, Blakemore D, et al. Fine-tuning Mybl2 is required for proper mesenchymal-to-epithelial transition during somatic reprogramming. *Cell Rep.* 2018;24:1496–1511.e1498.
 17. Jiang L, Wang X, Ma F, Wang X, Shi M, Yan Q, et al. PITX2C increases the stemness features of hepatocellular carcinoma cells by up-regulating key developmental factors in liver progenitor. *J Exp Clin Cancer Res.* 2022;41:211.
 18. Charontong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 2017;18:248–62.
 19. Chen D S, Mellman I. Oncology meets immunology: The cancer-immunity cycle. *Immunity.* 2013;39:1–10.
 20. Qi J, Sun H, Zhang Y, Wang Z, Xun Z, Li Z, et al. Single-cell and spatial analysis reveal interaction of FAP(+) fibroblasts and SPP1(+) macrophages in colorectal cancer. *Nat Commun.* 2022;13:1742.
 21. Sharma A, Seow JJW, Dutertre CA, Pai R, Blériot C, Mishra A, et al. Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. *Cell.* 2020;183:377–394.e321.
 22. Miller BC, Sen DR, Al Abosy R, Bi K, Virkud YV, LaFleur MW, et al. Subsets of exhausted CD8(+) T cells differentially mediate tumor control and respond to checkpoint blockade. *Nat Immunol.* 2019;20:326–36.
 23. Giannelli G, Rani B, Dituri F, Cao Y, Palasciano G. Moving towards personalised therapy in patients with hepatocellular carcinoma: The role of the microenvironment. *Gut.* 2014;63:1668–76.
 24. Van de Sande B, Lee JS, Mutasa-Gottgens E, Naughton B, Bacon W, Manning J, et al. Applications of single-cell RNA sequencing in drug discovery and development. *Nat Rev Drug Discov.* 2023;22:496–520.
 25. Ennishi D, Hsi ED, Steidl C, Scott DW. Toward a new molecular taxonomy of diffuse large B-cell lymphoma. *Cancer Discov.* 2020;10:1267–81.
 26. Tan AC, Tan DSW. Targeted therapies for lung cancer patients with oncogenic driver molecular alterations. *J Clin Oncol.* 2022;40:611–25.
 27. Llovet JM, Montal R, Sia D, Finn RS. Molecular therapies and precision medicine for hepatocellular carcinoma. *Nat Rev Clin Oncol.* 2018;15:599–616.
 28. Zucman-Rossi J, Villanueva A, Nault JC, Llovet JM. Genetic landscape and biomarkers of hepatocellular carcinoma. *Gastroenterology.* 2015;149:1226–239.e1224.
 29. Jayachandran A, Dhungel B, Steel JC. Epithelial-to-mesenchymal plasticity of cancer stem cells: Therapeutic targets in hepatocellular carcinoma. *J Hematol Oncol.* 2016;9:74.
 30. Baker SJ, Ma'ayan A, Lieu YK, John P, Reddy MVR, Chen EY, et al. B-myb is an essential regulator of hematopoietic stem cell and myeloid progenitor cell development. *Proc Natl Acad Sci USA.* 2014;111:3122–7.
 31. Karagiannis P, Takahashi K, Saito M, Yoshida Y, Okita K, Watanabe A, et al. Induced pluripotent stem cells and their use in human models of disease and development. *Physiol Rev.* 2019;99:79–114.
 32. Zhang JX, Chen ZH, Xu Y, Chen JW, Weng HW, Yun M, et al. Downregulation of microRNA-644a promotes esophageal squamous cell carcinoma aggressiveness and stem cell-like phenotype via dysregulation of PITX2. *Clin Cancer Res.* 2017;23:298–310.
 33. Zaretsky JM, Garcia-Diaz A, Shin DS, Escuin-Ordinas H, Hugo W, Hu-Lieskovan S, et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N Engl J Med.* 2016;375:819–29.
 34. Swanson K, Wu E, Zhang A, Alizadeh AA, Zou J. From patterns to patients: Advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. *Cell.* 2023;186:1772–91.
 35. Cohen YC, Zada M, Wang SY, Bornstein C, David E, Moshe A, et al. Identification of resistance pathways and therapeutic targets in relapsed multiple myeloma patients through single-cell sequencing. *Nat Med.* 2021;27:491–503.
 36. Xiang W, Lv H, Xing F, Sun X, Ma Y, Wu L, et al. Inhibition of ACLY overcomes cancer immunotherapy resistance via polyunsaturated fatty acids peroxidation and cGAS-STING activation. *Sci Adv.* 2023;9:eadi2465.
 37. Zhang L, Jiang C, Zhong Y, Sun K, Jing H, Song J, et al. STING is a cell-intrinsic metabolic checkpoint restricting aerobic glycolysis by targeting HK2. *Nat Cell Biol.* 2023;25:1208–22.

How to cite this article: Chen Y, Deng X, Li Y, Han Y, Peng Y, Wu W, et al. Comprehensive molecular classification predicted micro-environment profiles and therapy response for HCC. *Hepatology.* 2024;80:536–551. <https://doi.org/10.1097/HEP.0000000000000869>