The Prognostic Model Based on Tumor Cell Evolution Trajectory Reveals a Different Risk Group of Hepatocellular Carcinoma

Haoren Wang[1,2†], Shizhe Yu[1,3†], Qiang Cai[1], Duo Ma[1], Lingpeng Yang[1], Jian Zhao[1], Long Jiang[1], Xinyi Zhang[1] and Zhiyong Yu[1]*

[1] Department of Hepatobiliary Surgery, The Affiliated Hospital of Yunnan University, Kunming, China, [2] Department of Oncology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, [3] Department of Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

Hepatocellular carcinoma (HCC) is one of the leading causes of cancer-related death worldwide, and heterogeneity of HCC is the major barrier in improving patient outcome. To stratify HCC patients with different degrees of malignancy and provide precise treatment strategies, we reconstructed the tumor evolution trajectory with the help of scRNA-seq data and established a 30-gene prognostic model to identify the malignant state in HCC. Patients were divided into high-risk and low-risk groups. C-index and receiver operating characteristic (ROC) curve confirmed the excellent predictive value of this model. Downstream analysis revealed the underlying molecular and functional characteristics of this model, including significantly higher genomic instability and stronger proliferation/progression potential in the high-risk group. In summary, we established a novel prognostic model to overcome the barriers caused by HCC heterogeneity and provide the possibility of better clinical management for HCC patients to improve their survival outcomes.

Keywords: hepatocellular carcinoma, single-cell transcriptomics, tumor heterogeneity, tumor evolution, copy number aberration, prognosis, cell state transition, genomic diversity

INTRODUCTION

Hepatocellular carcinoma (HCC), with more than one million new cases annually and a 5-year survival rate <20% in most countries, is the fastest growing malignancy both in terms of incidence and mortality (Allemani et al., 2018; Bray et al., 2018). Despite the clinical efficacy of systemic therapies such as tyrosine kinase inhibitors (TKIs) and immune checkpoint inhibitors (ICIs), primary and secondary drug resistance is inevitable and ultimately leads to treatment failure (Llovet et al., 2008, 2018; Abou-Alfa et al., 2018; Pinter et al., 2021). This ability of HCC to adapt to pharmacologic pressures can be described as tumor evolution and can be attributed to the heterogeneity of HCC (Amirouchene-Angelozzi et al., 2017), which refers to the different genetic or epigenetic alterations within the same lesion (intratumor heterogeneity) or in different lesions in the same patient (inter-tumor heterogeneity). Thus, the understanding of the potential mechanisms underlying HCC heterogeneity and its impact on therapeutic intervention is paramount for treatment success and overall survival (OS; Llovet et al., 2021).

The traditional bulk RNA-seq only provides the average number of genes expressed in a pooled population of cells and cannot detect the wide transcriptome heterogeneity in cell populations (Wang et al., 2020). Thus, researchers previously classified cells by purpose-related features and focused on a set of genes, ignoring the continuity of the tumor evolution process (Bidkhori et al., 2018; Chaudhary et al., 2018; Long et al., 2019b; Sarathi and Palaniappan, 2019; Sun et al., 2020; Zhang et al., 2020; Zhu et al., 2020). However, with the advent of single-cell RNA sequencing (scRNA-seq), a novel technology that allows transcriptomic analyses of individual cells, researchers can explore the heterogeneity and plasticity of tumor cells, which can result in early recurrence and drug resistance in the process of tumor evolution on single cell resolution. Given the large number of cells, we can reasonably hypothesize that the sequencing results include every distinct point of the dynamic process (Losic et al., 2020; Marjanovic et al., 2020; Rajewsky et al., 2020; Sun et al., 2021). Several studies have profiled the single-cell landscape of tumor generation and progression (Kim et al., 2018; Durante et al., 2020; Losic et al., 2020; Marjanovic et al., 2020). Furthermore, scRNA-seq is a promising tool that can facilitate individualized therapy owing to its ability to define cell subsets with potential treatment targets.

Here, we reconstructed the evolution trajectory of tumor cells with the help of scRNA-seq data and established a prognostic model to classify different risk groups of HCC. Our findings provide a strategy for precision medicine on the basis of tumor heterogeneity, and we also identified a wide range of potential therapeutic targets, thus improving the survival of patients with HCC.

MATERIALS AND METHODS

Data Sources

The normalized gene-level RNA-Seq data and clinical information for 364 patients LIHC-TCGA cohorts were downloaded from UCSC Xena[1] with R package UCSC Xena Tools (Wang and Liu, 2019). To obtain 258 patients LIRI-JP validation set, RNA-seq data, and related clinic pathological data were downloaded from the ICGC website[2] (Zhang J. et al., 2019). The scRNA-seq barcode sequences and raw gene expression matrix were downloaded from the CNP0000650 (Sun et al., 2021). Mutation data that contained somatic variants were stored in Mutation Annotation Format (MAF) form and were downloaded from Genomic Data Commons (GDC).[3]

Processing of Single-Cell RNA-Seq Data
Dimension Reduction and Unsupervised Clustering
Single-cell RNA sequencing data were processed for dimension reduction and unsupervised clustering by following the workflow in Seurat (v4.0.2) (Butler et al., 2018). In brief, first, the read counts for each cell were divided by the total counts

for that cell and multiplied by the scale factor (10,000), and then natural-log transformed. A principal component analysis (PCA) matrix with 50 components were calculated to reveal the main axes of variation and the data were denoised by using "Run PCA" function with default parameter. For visualization, the dimensionality of each dataset was further reduced using Uniform Manifold Approximation and Projection (UMAP) implemented in "Run UMAP" function (Becht et al., 2019). We retained cell clustering based on the original study. The cluster-specific marker genes were identified by using the "Find All Markers" function with MAST algorithm (Finak et al., 2015).

Define Subpopulations of Aneuploid Tumor Cells
TPM gene expression matrix was extracted from the Seurat object as recommended in the "prepare the read count input file" section (CopyKAT). For each patient, normal reference T cells and malignant cells were selected and identified from the annotated clusters as determined above. Quality control filtering was performed to select the highest quality cells by only including malignant cells with at least five genes in each chromosome to calculate DNA copy numbers. We extracted aneuploid cells that are considered as tumor cells in aneuploid tumors to define two copy number subpopulations of single tumor cells using default parameters in CopyKAT (Gao et al., 2021).

Construct Tumor Cell Evolution Trajectory
Malignant cells were identified from the annotated clusters as determined above. This resulted in six high-quality malignant clusters to use for this analysis. Single-cell pseudo-time trajectories were constructed with Monocle 2 (2.10.1) (Qiu et al., 2017). Genes for trajectory inference were selected using the "dispersion table" function to calculate a smooth function describing how variance in each gene's expression across cells varies according to the mean. Only genes with mean expression greater than or equal to 0.1 were used for the analysis. The "reduce Dimension" function was utilized with the DDRTree reduction method with default parameters. Results were visualized using the "plot cell trajectory" and "plot complex cell trajectory" functions and annotated with cell type, subclones, and calculated cell states. Once the pseudo-space trajectory was defined, we used the Tradeseq (Trajectory Differential Expression analysis for Sequencing data) R package to select genes that were differentially expressed along the trajectory (Van den Berge et al., 2020). Association Test function was used to test whether the average gene expression is significantly changing along pseudotime. The top 500 gene upregulated genes and Top 500 downregulated genes decrease along the inferred pseudo-time trajectory with a q-value less than 0.01 were separated with hierarchical clustering using the "plot multiple ranches heatmap" function with num clusters = 3 and "branches" set to the terminal branchpoints for aneuploid tumor cells.

Development and Validation of the Tumor Evolution Signature for Hepatocellular Carcinoma
Select differential genes based on single cell tumor evolution trajectory to reduce the impact of non-tumor cells. The cases from the TCGA LIHC datasets were used as the training cohort

to establish the LASSO model. Univariate analysis and logRank test were used to identify genes with prognostic ability. For the genes with prognostic ability, Cox proportional hazards model (iteration = 1,000) with a lasso penalty was used to find the best gene model utilizing an R package called "glmnet" (Friedman et al., 2010). The best gene model was used to establish the tumor evolution signature. The risk score for each patient was calculated with the LASSO model weighting coefficient as follows:

$$riskscores = \sum_{i=1}^{n} Coef_j * X_j$$

In this formula, n represents the number of key genes, $Coef_j$ is the LASSO coefficient of Gene j, and X_j is the normalized expression value of Gene j (**Supplementary Table 4**). Then, the concordance (c)-index proposed by Harrell[24] was applied to validate the predictive ability of the signature in all datasets, by using the "survcomp" R package (Haibe-Kains et al., 2008). The larger c-index indicated the more accurate predictive ability of the model.

Survival Analysis

To verify the trend of this tumor evolution trajectory, Kaplan–Meier (K–M) analysis was performed. The top 10 end-genes were extracted and the potential prognostic significance of these genes was assessed with the LIHC data from GEPIA2.[4] The K–M survival curves were also generated to graphically demonstrate the OS to the high-risk group and low-risk group, which were stratified by the tumor evolution signature. The R package called "survminer" was utilized to perform the survival analysis, and the optimal cutoff was ascertained by the "surv_cutpoint" function.

Somatic Mutation and Copy-Number Aberration Analysis

Mutation comment file (MAF) of TCGA-LIHC cohort was downloaded from the GDC client. Differential analysis and visualization of somatic mutations was done using Maftools (The Cancer Genome Atlas Research Network et al., 2013; Mayakonda et al., 2018). This difference between high- and low-risk group was detected using function "mafComapre," which performs Fisher's exact test on all genes between two groups to detect differentially mutated genes.

Composite copy number profiles were generated to highlight differences between high- and low- risk group. Segment file of TCGA-LIHC cohort was downloaded from FIREHOSE and samples were further divided into high- and low-risk groups. Then we ran the GISTIC 2.0 pipeline to generate discrete copy number data file. Chromosomes reference objects were from the "BSgenome.Hsapiens.UCSC.hg19" R package.

As in the previous study, a non-synonymous mutation from the TCGA database was used as the raw mutation count, and it was divided by 38 MB to quantify TMB (Chalmers et al., 2017). The samples were sorted according to the value of the median TMB from low to high.

Bioinformatics Analyses

Gene Enrichment Analysis (GSEA) was further used to investigate the functional enrichment of risk score-associated genes using the R package "clusterProfiler" (Yu et al., 2012). The Benjamini–Hochberg method was used to adjust nominal p-values (false discovery rate, FDR) for multiple testing.

The Maftools package was used to illustrate the respective mutation profiling of the two risk group levels by waterfall plot, and differentially mutated genes were identified by using the "mafCompare" function where genes mutated in greater than 5% of LIHC samples in the TCGA cohort were considered (Mayakonda et al., 2018).

Statistical Analysis

Student's *t-test* was conducted to make statistical comparison. The "pheatmap" R package was applied to generate heatmaps. Survival analysis was completed using Kaplan–Meier method, and the prediction performance of the risk model was evaluated using receiver operating characteristic (ROC) *via* "time-ROC" R package. Multivariate COX regression analyses were used to investigate the prognostic value of risk-score. Hazard ratio (HR) and 95% confidence intervals (CI) for each variable were also calculated where needed. A value of $p < 0.05$ was defined as statistically significant difference. All of our analyses were conducted using R software version 4.0.2.[5]

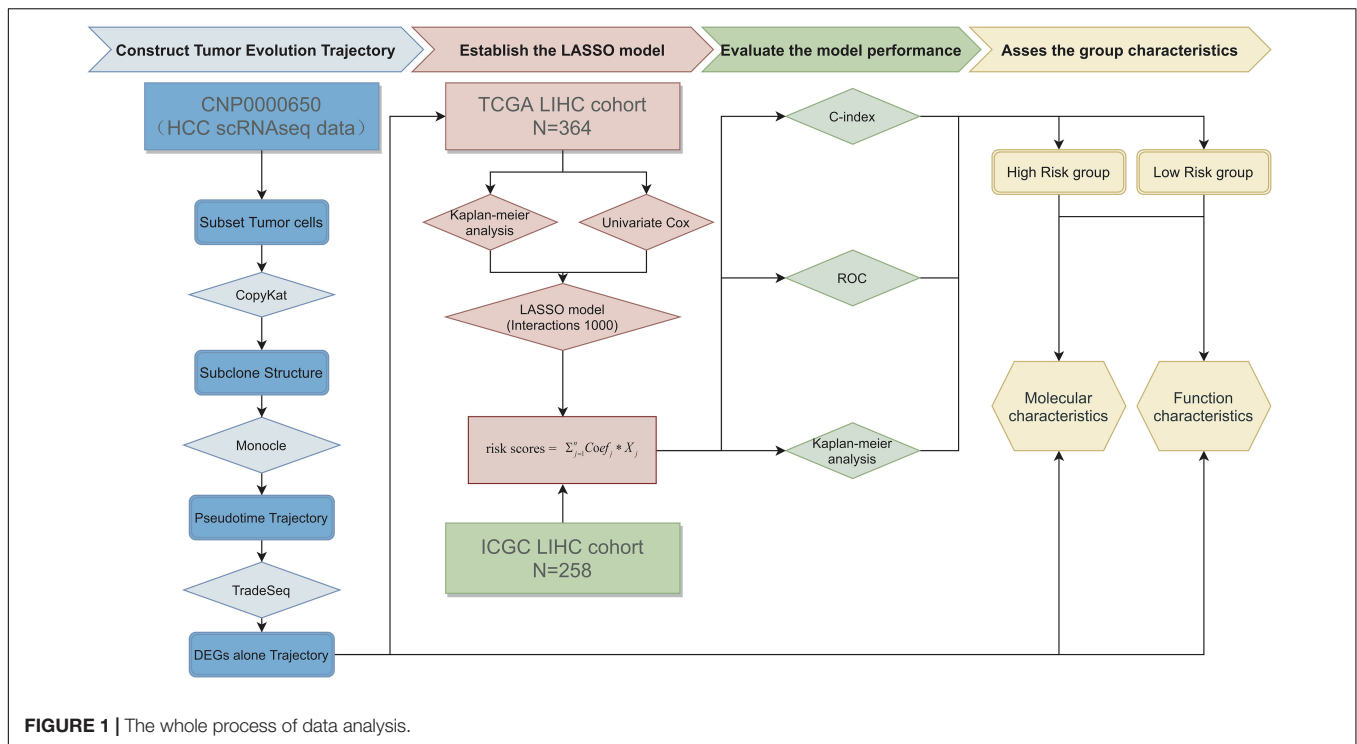The whole process of data analysis is depicted in **Figure 1**.

RESULTS

Classification of the Malignant Cell Clusters With Single-Cell RNA Sequencing Data

Unsupervised dimensionality reduction and graph-based clustering analysis were performed with the data from CNP0000650, and 24 clusters (**Figure 2A**) were visualized by the UMAP method (Becht et al., 2019; Sun et al., 2021). The immune cells mainly consisted of myeloid-derived cells, T cells, B cells, plasma cells, and natural killer (NK) cells, while non-immune cells included endothelial cells, hepatic stellate cells, apparently normal epithelial cells, and HCC malignant cells. To contrast the difference among different patients, we classified the cells by patient origin (**Figure 2B**), and the result showed that tumor cells contained obvious heterogeneity, while non-tumor cells kept homogeneous, proving that the differences between tumor cell clusters are mainly due to the tumor heterogeneity, rather than batch effects between samples. No normal liver cells were detected, likely because of the technical limitations, resulting in no comparison between normal and malignant liver cells.

We further extracted the varied tumor cell clusters for the analysis of tumor heterogeneity (**Figure 2C**), and two major sub-clones were defined by the clustered heat maps of single cell copy number profiles (**Figure 2D**). Compared with sub-clone 2 (green), the heatmap showed that sub-clone 1 (red)

---

[4]http://gepia2.cancer-pku.cn/#index

[5]https://www.r-project.org/

FIGURE 1 | The whole process of data analysis.

contained more CNAs, implying that sub-clone 1 might be more malignant (**Figure 2E**).

Reconstruction of Tumor Cells Progression Trajectory

To determine the relationship between malignant cell clusters, we performed single-cell trajectory analysis with scRNA-seq data using Monocle (Qiu et al., 2017). As is well known, genomic mutations accumulate over time in the process of tumor evolution (Turajlic et al., 2019; Craig et al., 2020; Losic et al., 2020). Thus, we defined the cell cluster with a lower CNA burden as the root, while the cell cluster with a higher CNA burden was defined as the end of the trajectory. We noticed that sub-clone 2, comprising cells with obvious liver characteristics, was concentrated at the beginning of the trajectory, while sub-clone 1, comprising cells with less specificity of origins, was concentrated at the end of the trajectory, indicating that the trajectory model fits the process of tumor evolution well (**Figures 3A,B**).

During the process of transition of tumor cells from a lower to higher malignant state, some genes are silenced, while others become newly active. We used Tradeseq, a powerful generalized additive model framework based on the negative binomial distribution, to interpret the within lineage differential expression (**Figure 3C** and **Supplementary Table 1**). To verify the malignant trend of this trajectory, we extracted the top 10 downregulated genes (root genes) and top 10 upregulated genes (end genes) for Kaplan–Meier (K-M) analysis. The high expression of the top 10 root genes represented a benign prognosis, while the high expression of the top 10 end genes represented a poor prognosis (**Figures 3F,G**). Furthermore, along the trajectory, liver

characteristics such as *ALB* (**Figure 3D**) were gradually lost, while stemness and malignant marker genes such as *TWIST1* (**Figure 3E**), gradually increased, suggesting that the malignant state was advancing.

The gene set variation analysis (GSVA; Hänzelmann et al., 2013) was used to further analyze the underlying biological processes along the trajectory. In the Molecular Signature Database (MSigDB) "hallmark" collection of major biological categories (Liberzon et al., 2015), the upregulated genes of sub-clone 1 were enriched in the tumor-promoting pathway ("Wnt/β-catenin signaling") and proliferation pathway ("G2M checkpoint," "E2F Targets," "MYC Targets"), while the downregulated genes of sub-clone 1 were enriched in the tumor-suppressor pathway ("P53 pathway") and essential liver function pathway ("Complement," "Fatty acid metabolism," "Adipogenesis") (**Figure 3H**), which were consistent with the characteristics of cells that progressed from the lower malignant state to the higher malignant state (Maley et al., 2017; Losic et al., 2020; Marjanovic et al., 2020; Llovet et al., 2021).

Establishment of the 30-Gene Prognostic Model

Although the top 10 genes had a certain predictive effect, we preferred optimizing gene combination to obtain a better prognostic model. With the selection criteria of $p < 0.01$, the intersection of univariate Cox regression analysis and K-M analysis identified 200 credibly survival-related genes. We used TCGA data as the training cohort and ICGC data as the external validation cohort (The Cancer Genome Atlas Research Network et al., 2013; Zhang J. et al., 2019). Lasso-penalized Cox
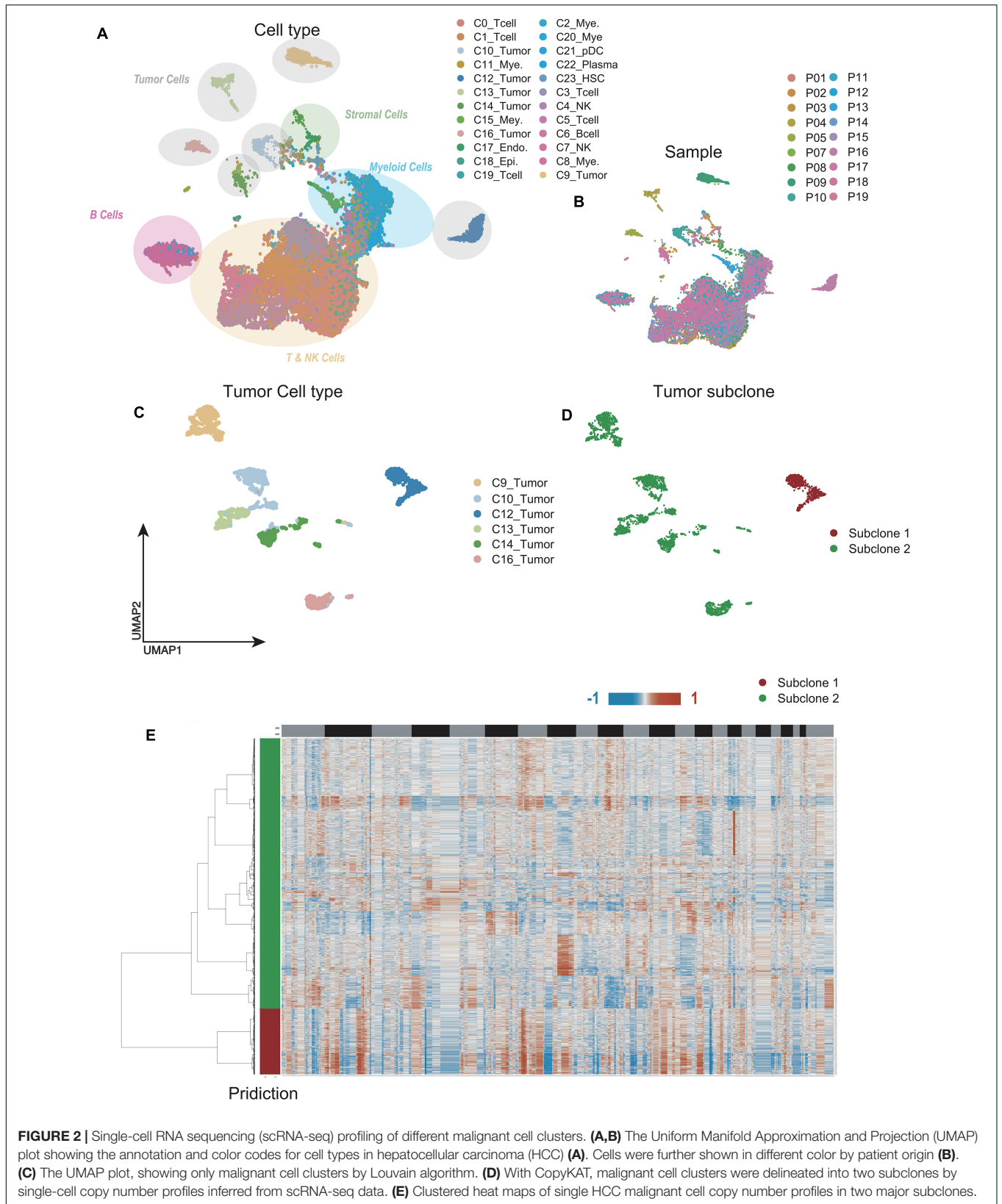
FIGURE 2 | Single-cell RNA sequencing (scRNA-seq) profiling of different malignant cell clusters. **(A,B)** The Uniform Manifold Approximation and Projection (UMAP) plot showing the annotation and color codes for cell types in hepatocellular carcinoma (HCC) **(A)**. Cells were further shown in different color by patient origin **(B)**. **(C)** The UMAP plot, showing only malignant cell clusters by Louvain algorithm. **(D)** With CopyKAT, malignant cell clusters were delineated into two subclones by single-cell copy number profiles inferred from scRNA-seq data. **(E)** Clustered heat maps of single HCC malignant cell copy number profiles in two major subclones.
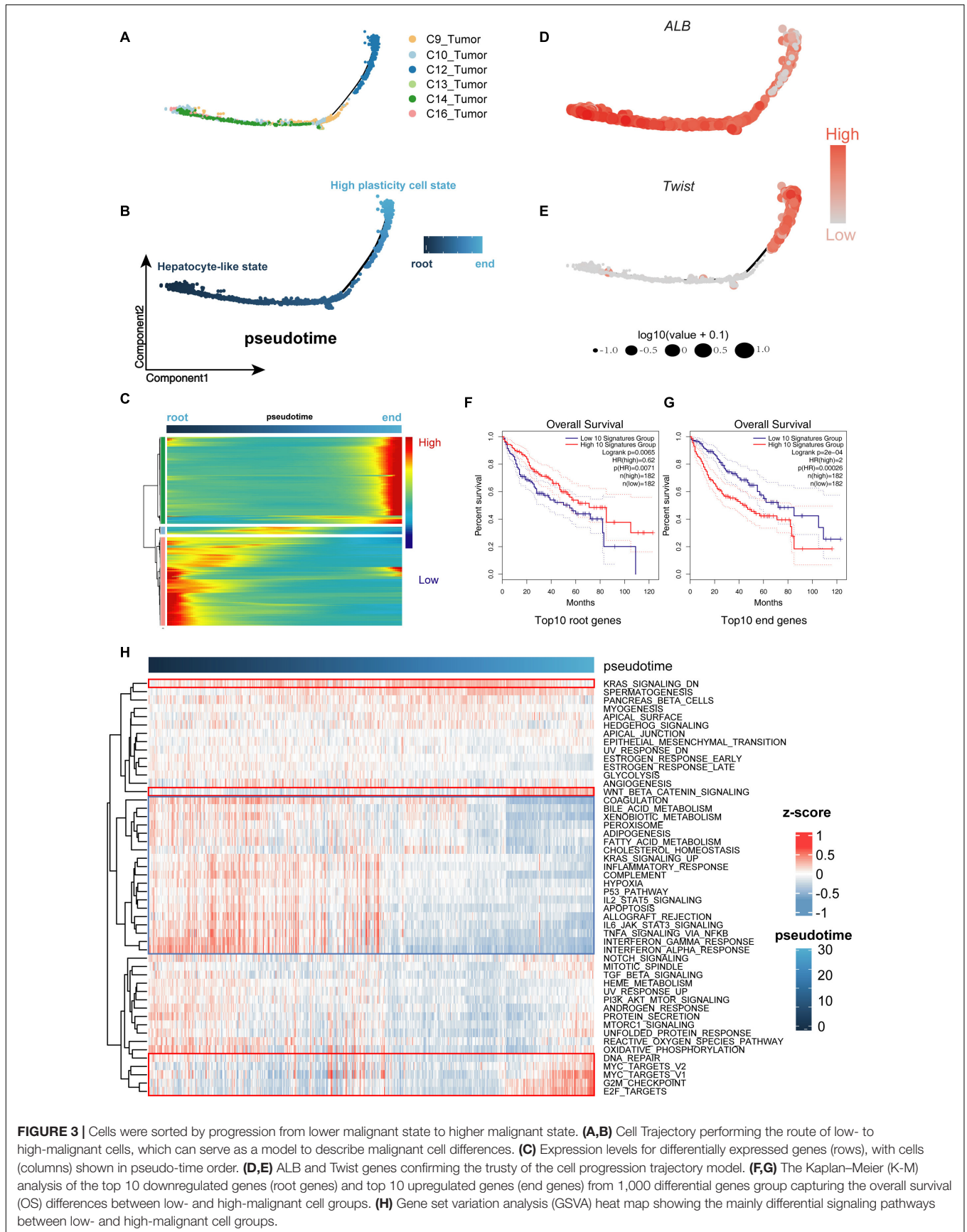
FIGURE 3 | Cells were sorted by progression from lower malignant state to higher malignant state. (A,B) Cell Trajectory performing the route of low- to high-malignant cells, which can serve as a model to describe malignant cell differences. (C) Expression levels for differentially expressed genes (rows), with cells (columns) shown in pseudo-time order. (D,E) ALB and Twist genes confirming the trusty of the cell progression trajectory model. (F,G) The Kaplan–Meier (K-M) analysis of the top 10 downregulated genes (root genes) and top 10 upregulated genes (end genes) from 1,000 differential genes group capturing the overall survival (OS) differences between low- and high-malignant cell groups. (H) Gene set variation analysis (GSVA) heat map showing the mainly differential signaling pathways between low- and high-malignant cell groups.

analysis was subsequently performed 1,000 times in the TCGA training cohort with 10-fold cross-validation to evaluate and eliminate variables that contributed less to the model, and a 30-gene signature with the most powerful predictive features were selected (**Figures 4A,C,D**). To validate the credibility of this model, C-index was assessed in the TCGA training cohort and ICGC validation cohort, which was confirmed as 0.79 and 0.73, respectively (**Figure 4B**), suggesting that our model had favorable efficacy for predicting prognosis (Harrell, 1982). Based on the 30-gene prognostic model, TCGA and ICGC samples were clustered into high-risk and low-risk groups, and the OS time of patients in the high-risk group was remarkably decreased (**Figures 4E,F**).

Evaluation of the Prognostic Model in TCGA Cohort and ICGC Cohort

The K–M analysis and time-dependent ROC was used to assess the prognostic capacity of the 30-gene prognostic model in the TCGA cohort and ICGC cohort, respectively. The K–M analysis illustrated that patients in the low-risk group had significantly longer OS than those in the high-risk group, both in the TCGA cohort (**Figure 5A**) and the ICGC cohort (**Figure 5B**). The area under the ROC curve (AUC) for the 1-, 3-, and 5-year OS was 0.843, 0.848, and 0.824 in the TCGA cohort, while it was 0.77, 0.796, and 0.774 in the ICGC cohort (**Figures 5C,D**), indicating that this 30-gene prognostic model had high sensitivity and specificity for survival prediction.

Gender, age, stage, vascular invasion, bile duct invasion, fibrosis, and the risk score of the prognostic model were included in the multivariate Cox regression model, and the risk score was revealed to be independent predictor for OS, showing splendid predictive performance ability, with HR: 6.432, 95% CI: 4.095–10.103, $p < 0.001$ in TCGA cohort, and HR: 3.751, 95% CI: 1.419–9.918, $p < 0.001$ in ICGC cohort. Taken together, the 30-gene prognostic model was completely reliable for the precise prediction of OS in HCC (**Figure 5E**).

Comparison of Genomic Aberrations in Different Risk Groups

Increasing mutation frequency is a typical feature of human cancer. To identify the divergence of genomic aberrations between the high-risk and low-risk groups in the TCGA database, CNAs data were downloaded from the GDC portal and analyzed with GISTIC 2.0, with which the high-risk group had an obviously higher genomic aberration burden than the low-risk group (**Figure 6A**). Meanwhile, the top 20 genes with high genomic mutation frequency in the high-risk and low-risk groups were constructed by Maftools (Mayakonda et al., 2018; **Figures 6B,C**). To analyze the discrepancy between the high-risk and low-risk groups, the differentially mutated gene type and frequency were compared by Fisher's exact tests. The results showed three significantly differential genes—*TP53* (47 versus 19%), *OBSCN* (18 versus 5%), and *RB1* (11 versus 2%) (**Figure 6D**). The six genes most recurrently mutated were *TP53* (47 versus 19%), *TTN* (31 versus 27%), *CTNNB1* (26 versus 28%), *MUC16* (19 versus 16%), *OBSCN* (18 versus 5%), and *ALB* (12 versus 13%) (**Figure 6E** and **Supplementary Table 2**). These

findings suggested that some mutated genes such as *TP53* and *OBSCN* that were notably different compared with the high-risk and low-risk groups could be related to the malignant progression and can continuously accumulate mutations over time; whereas, the others such as *CTNNB1* and *ALB*, which remained stable from the low-risk state to high-risk state, likely contribute to the essential neoplastic process rather than malignant progression.

The tumor mutation burden (TMB) is also considered an essential factor impacting on the occurrence and progression of the tumor. The distribution plot shows TMB distribution of different cancer types (**Figure 6F**). The three types including low-risk HCC, all HCC samples (liver hepatocellular carcinoma, LIHC), and high-risk HCC were apparently distinct from each other, and the TMB gradually increased from low-risk type to high-risk type, which suggested that our model had excellent distinguishing capability.

Gene Enrichment Analysis of the 30-Gene Prognostic Model

To explore the underlying molecular mechanisms of this prognostic model, we conducted GSEA to compare the low-risk group with the high-risk group in TCGA cohorts (Yu et al., 2012). In the MsigDB "hallmark" collection of major biological categories, proliferative signaling pathways ("E2F targets," "G2M checkpoint," "KARS signaling"), and the invasion and metastasis-related signaling pathways ("EMT" and "myogenesis") were dramatically increased in the high-risk group (**Figure 7B**). This was consistent with the data at the single-cell level. Notably, the low-risk group was enriched in inflammation-related gene pathways such as "inflammatory response," "interferon-α response," and "interferon-γ response," which suggested that the secretion of inflammatory factors might originate from the tumor cells, indicating a strong proinflammation potential in the early tumorigenesis stage (**Figure 7A**). A similar phenomenon had been reported in melanoma, breast cancer, and colorectal cancer, where the expression of immune-related genes was also presented by tumor cells and could likely be an independent influence on their prognostic differences (Sconocchia et al., 2014; Forero et al., 2016; Buetow et al., 2019; McCaw et al., 2019; Jin et al., 2020).

DISCUSSION

Hepatocellular carcinoma is one of the leading causes of cancer-related mortality worldwide. Previous studies have proved that the heterogeneity, which was thought to be evolutionarily selected for increasing fitness of tumor cells, might be the major barrier for improving patients outcome (Chaudhary et al., 2018; Llovet et al., 2018, 2021; Long et al., 2019b,a; Yang et al., 2019; Craig et al., 2020). Thus, there is a critical need to stratify HCC patients accurately on the basis of heterogeneity and provide precise treatment strategies.

In this study, we analyzed CNAs in tumor cells using CopyKat and classified them into two major subclones, wherein subclone 1 had a higher CNA burden than that of subclone 2. Then, we defined the cells in hepatocyte-like state as the root and cells in high-plasticity state as the end, and
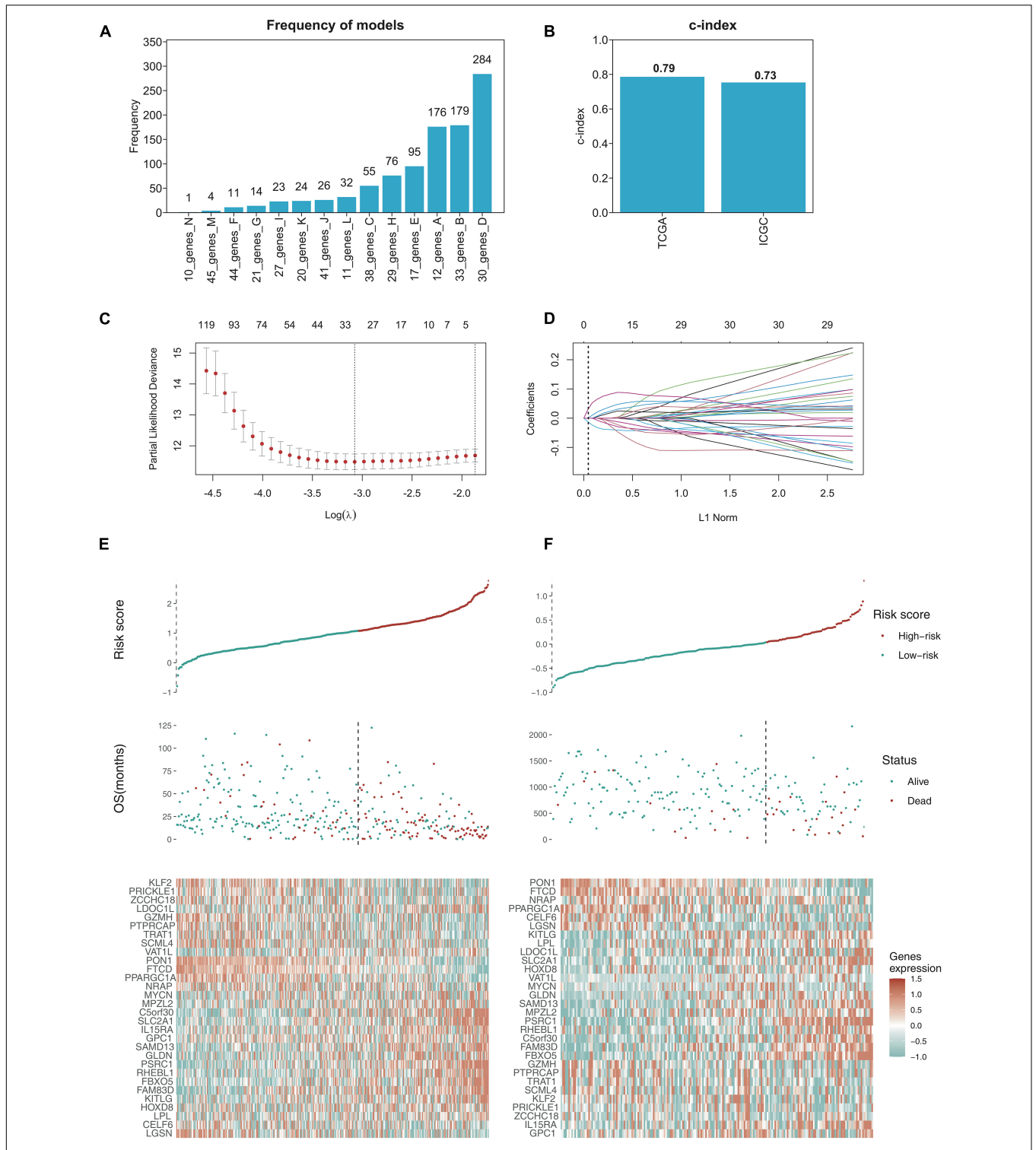
FIGURE 4 | Establishing the 30-gene prognostic model with LASSO regression analysis. **(A)** LASSO regression analysis performed the frequency of different gene combination models and finally determined the 30-gene signature for OS prediction. **(B)** C-index of 30-gene prognostic model was 0.79 in TCGA training cohort, while 0.73 in ICGC validation cohort. **(C)** LASSO coefficient profiles of the gene features. **(D)** Ten-time cross-validation for tuning parameter selection in the LASSO model. **(E,F)** The risk score distribution and survival status distribution of 30-gene prognostic model in TCGA training cohort and ICGC validation cohort, and the heat map of gene expression are shown below with color, red (high) and green (low).
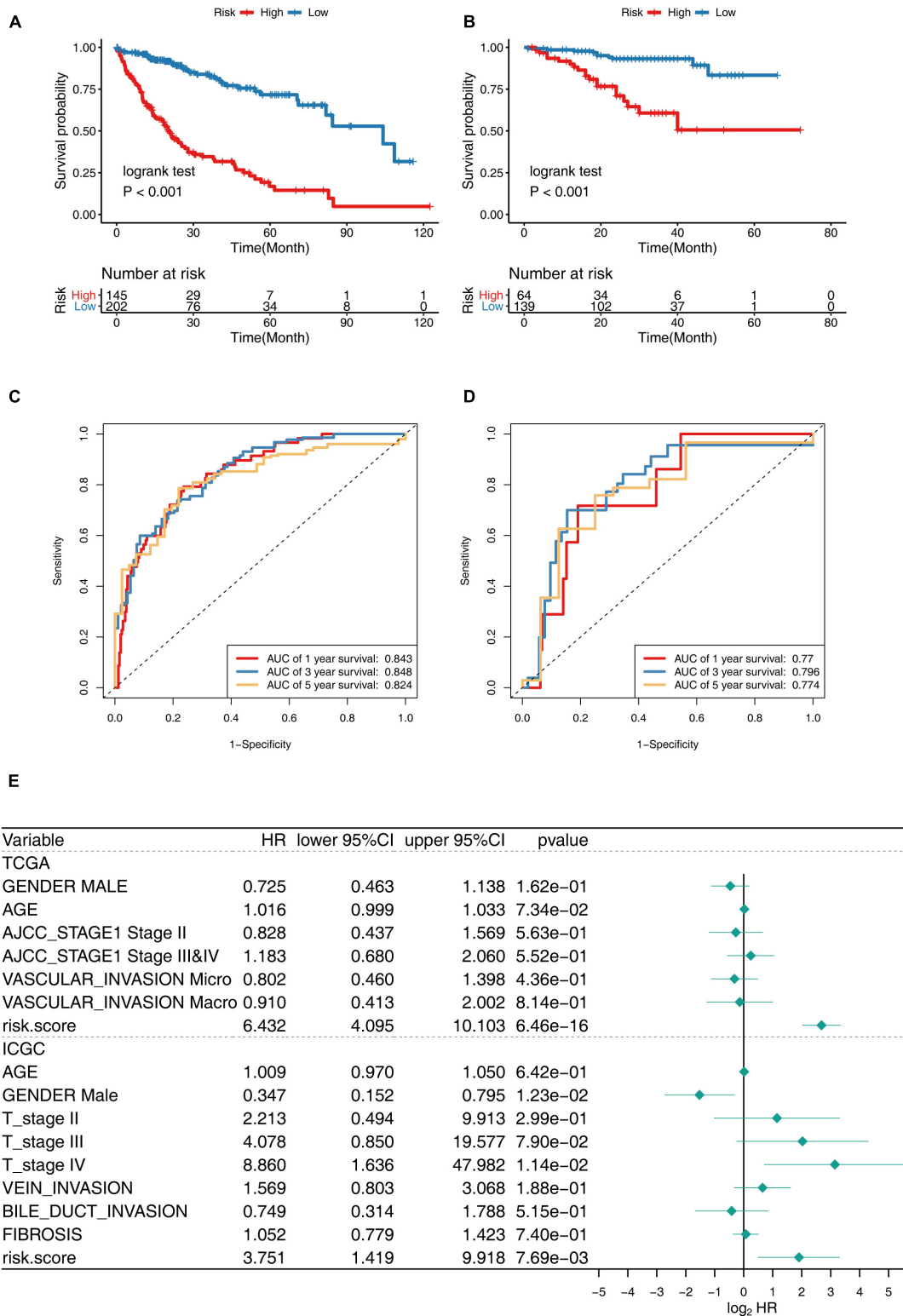
FIGURE 5 | Prognostic performance of 30-gene signature in TCGA Training Cohort and ICGC Validation Cohort. **(A,B)** K–M survival curve for risk score in TCGA training cohort **(A)** and ICGC validation cohort **(B)**. **(C,D)** Receiver operating characteristic (ROC) curve of the 30-gene prognostic model in TCGA cohort **(C)** and ICGC cohort **(D)**. **(E)** Multivariate Cox regression analysis of clinical parameters and prognostic model for OS.
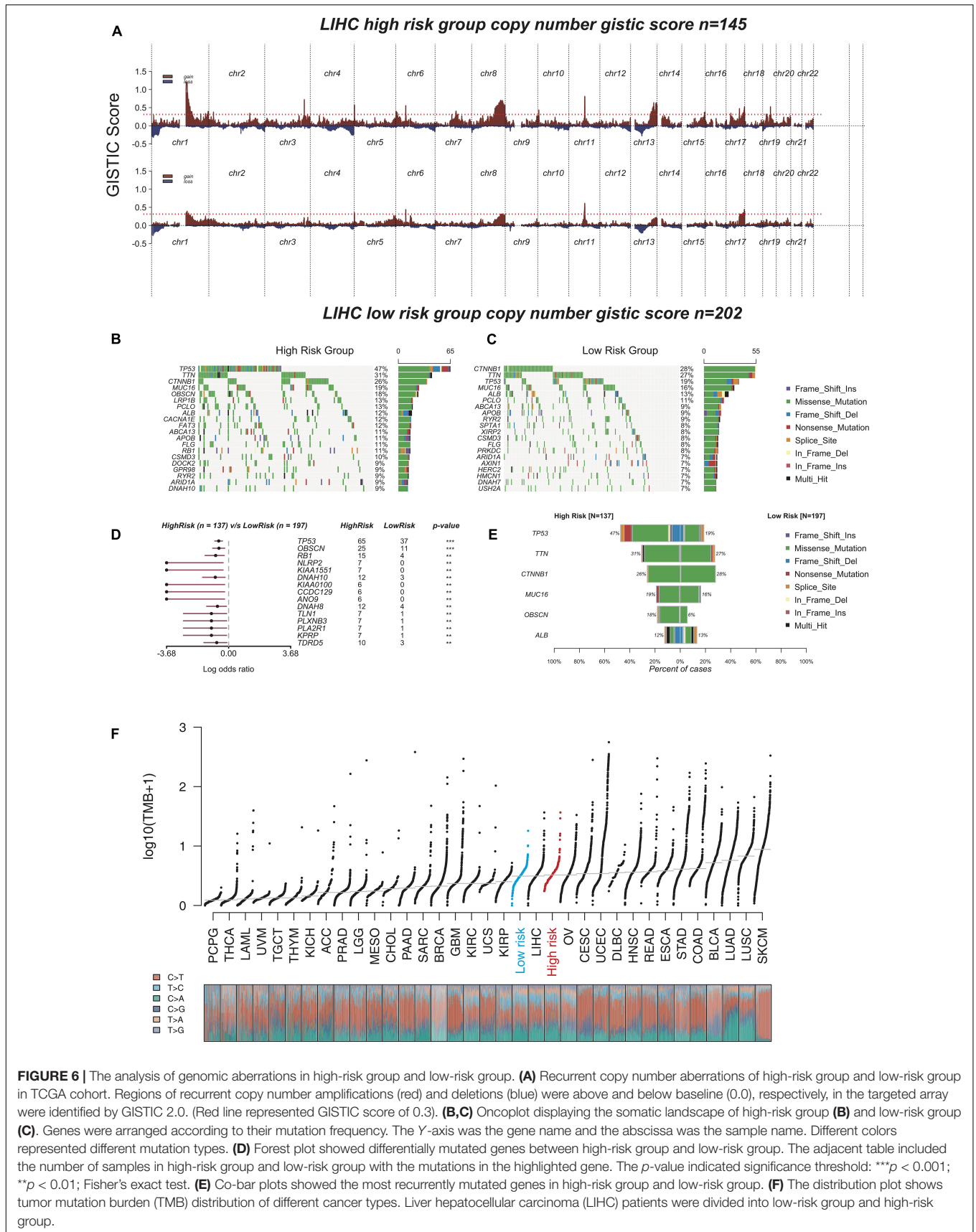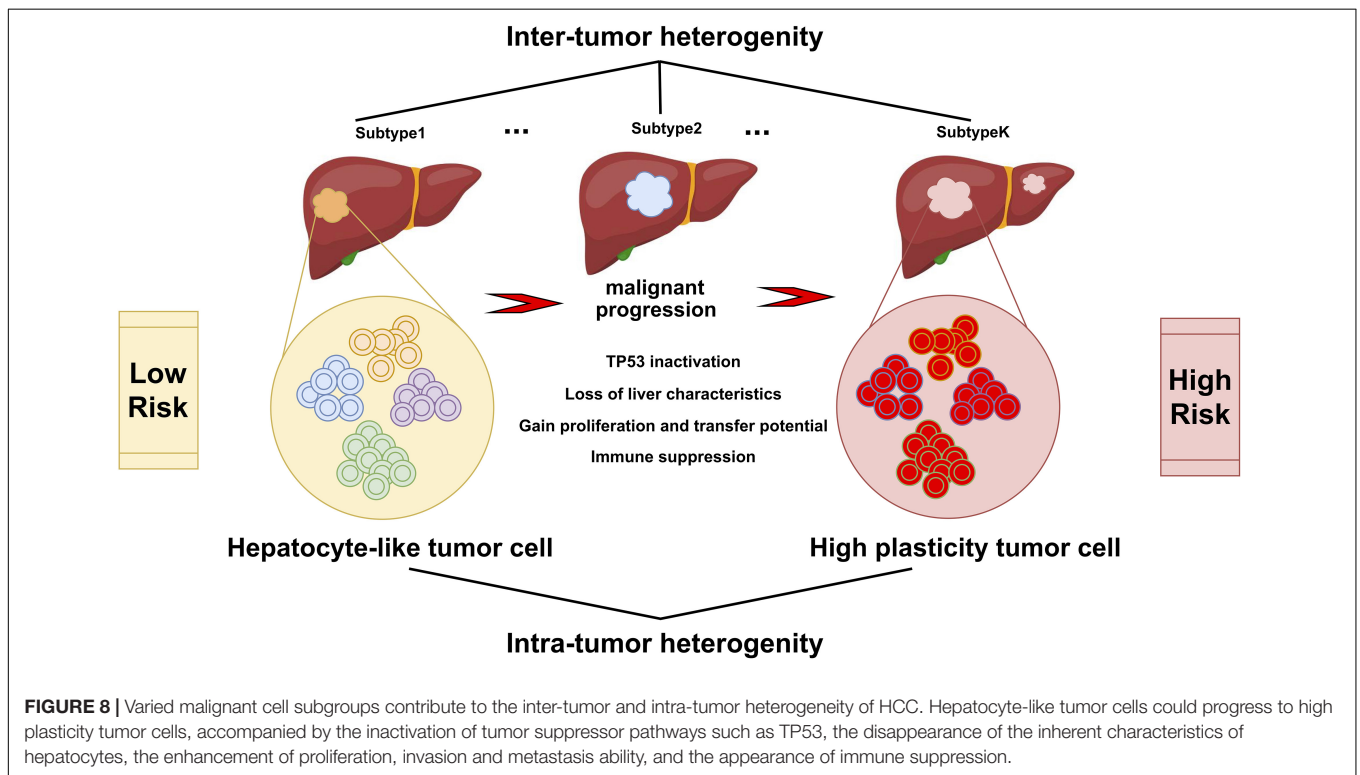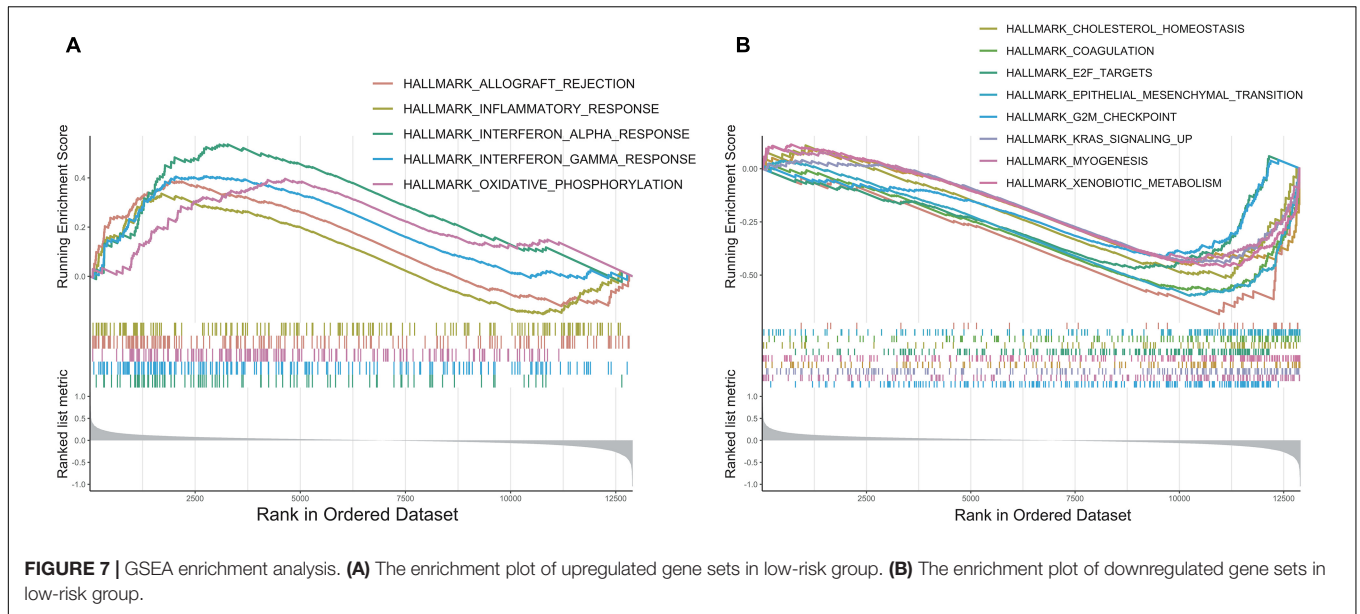
FIGURE 6 | The analysis of genomic aberrations in high-risk group and low-risk group. **(A)** Recurrent copy number aberrations of high-risk group and low-risk group in TCGA cohort. Regions of recurrent copy number amplifications (red) and deletions (blue) were above and below baseline (0.0), respectively, in the targeted array were identified by GISTIC 2.0. (Red line represented GISTIC score of 0.3). **(B,C)** Oncoplot displaying the somatic landscape of high-risk group **(B)** and low-risk group **(C)**. Genes were arranged according to their mutation frequency. The Y-axis was the gene name and the abscissa was the sample name. Different colors represented different mutation types. **(D)** Forest plot showed differentially mutated genes between high-risk group and low-risk group. The adjacent table included the number of samples in high-risk group and low-risk group with the mutations in the highlighted gene. The p-value indicated significance threshold: ***$p < 0.001$; **$p < 0.01$; Fisher's exact test. **(E)** Co-bar plots showed the most recurrently mutated genes in high-risk group and low-risk group. **(F)** The distribution plot shows tumor mutation burden (TMB) distribution of different cancer types. Liver hepatocellular carcinoma (LIHC) patients were divided into low-risk group and high-risk group.

FIGURE 7 | GSEA enrichment analysis. **(A)** The enrichment plot of upregulated gene sets in low-risk group. **(B)** The enrichment plot of downregulated gene sets in low-risk group.

FIGURE 8 | Varied malignant cell subgroups contribute to the inter-tumor and intra-tumor heterogeneity of HCC. Hepatocyte-like tumor cells could progress to high plasticity tumor cells, accompanied by the inactivation of tumor suppressor pathways such as TP53, the disappearance of the inherent characteristics of hepatocytes, the enhancement of proliferation, invasion and metastasis ability, and the appearance of immune suppression.

reconstructed the tumor evolution trajectory with the help of scRNA-seq data. Consistent with previous studies of other tumors (Zhang L. et al., 2019; Marjanovic et al., 2020), GSVA analysis revealed that along the trajectory, cells gradually lost their intrinsic characteristics and transformed into a high plasticity state. Based on this evolutionary trajectory, we further constructed a 30-gene prognostic model. C-index and multivariate analysis confirmed that compared with the other three existing prognostic models (Bidkhori et al., 2018;

Chaudhary et al., 2018; Liang et al., 2020; Sun et al., 2020; Zhang et al., 2020; Deng et al., 2021), this model possessed high predictive efficacy and accuracy. Finally, we also performed GSEA analysis to explore the underlying biological mechanisms of this model.

To investigate the heterogeneity and potential progression trajectory of HCC cells on single-cell resolution, high-quality scRNA-seq data were necessary. Thus, we reviewed associated studies of liver cancer published in recent years to identify the

most suitable single-cell dataset (**Supplementary Table 3**; Zheng et al., 2018; Ma et al., 2019; Zhang Q. et al., 2019; Losic et al., 2020; Sharma et al., 2020; Sun et al., 2021). Two key points need to be considered, namely, the number of tumor cells and the quality of sequencing data. Dissociating tissues is one of the difficulties of single-cell sequencing. Epithelial cells require more stringent dissociation conditions than immune cells and need to be enriched with FACs. However, in early studies, few authors noticed this point, making their results got a large proportion of immune cells and stromal cells instead of tumor cells. This was also the reason why some HCC studies focused on the immune microenvironment (Ma et al., 2019; Ramachandran et al., 2019; Massalha et al., 2020; Sun et al., 2021). To obtain enough cells, we narrowed the range into the data of Sun and Sharma (Sharma et al., 2020; Sun et al., 2021). Regarding the quality of sequencing data, a major factor is the sequencing platform. The plate-based SMART-seq2 full-length method provides in-depth coverage for a smaller number of cells, but the droplet-based 10× Genomics Chromium approach captures cells on a larger scale but with the limitation of inadequate gene coverage. The gene capturing rate of 20 cells by SMART-seq2 was comparable with that of 1,000 cells by 10× (Ding et al., 2019; Zhang Q. et al., 2019). Unfortunately, at least half of the tumor cells from the data of Sharma et al. (2020) based on the 10× platform, could not meet the input threshold of CopyKat, which would introduce a large bias in the downstream analysis (Sharma et al., 2020; Gao et al., 2021). To explore a more refined dynamic change process, we finally chose the data of Sun et al. (2021) for the downstream analysis.

As our model was based on tumor heterogeneity and the trajectory was highly similar to the natural process of tumor evolution, genes included in this model and the underlying biological mechanisms were complicated. Some upregulated genes were found to be associated with cell proliferation and progression. The upregulation of *GPC1* had been reported to be dramatically correlated with the reduced OS time for HCC patients (Wang et al., 2021). *MYCN*, a member of the *Myc* family, was positively correlated with the recurrence of *de novo* HCC (Qin et al., 2018). Furthermore, *EVA1* expression was significantly increased in HCC and was also associated with a poor prognosis and recurrence in these patients. Overexpression of *EVA1* promoted cell growth, invasion, and migration *in vitro*, while knockdown of *EVA1* expression inhibited proliferation and migration *in vitro* (Ni et al., 2020). Some downregulated genes were considered to function as tumor suppressor genes, such as *PPARGC1A*, also known as *PGC-1α*, a master regulator

of mitochondrial biogenesis and oxidative phosphorylation. A previous study had reported that low levels of *PPARGC1A* expression were correlated with poor survival, vascular invasion, and large tumor size (Huang et al., 2020; Zuo et al., 2021). *PRICKLE1* has been reported to be a negative regulator of the Wnt/beta-catenin signaling pathway and is a putative tumor suppressor gene in HCC (Chan et al., 2006).

Our study has some limitations. First, RNA-Seq detected more of the content specific to Affymetrix and Illumina microarrays than either of the microarray platforms on the same samples, and many of the feature genes included in our analysis were not detected on the microarray platform (Zhao et al., 2014). Thus, we only used cohorts based on RNA-seq platform. Second, our retrospective findings need to be further validated in prospective research, and complex mechanisms involved in the progression of liver cancer cells still need to be further explored.

In conclusion, this study integrated the scRNA-seq data and bulk multi-omics data to reconstruct the tumor evolution trajectory and establish a novel prognostic model to clarify different risk groups of HCC, which might help in clinical decision making for individual treatment and improve patient outcomes (**Figure 8**).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repositories and accession numbers can be found in section "Materials and Methods."

AUTHOR CONTRIBUTIONS

HW, SY, and ZY contributed to conception and design of the study. SY, HW, and QC performed the statistical analysis. HW wrote the first draft of the manuscript. DM, LY, XZ, and JZ wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.737723/full#supplementary-material

REFERENCES

Abou-Alfa, G. K., Meyer, T., Cheng, A.-L., El-Khoueiry, A. B., Rimassa, L., Ryoo, B.-Y., et al. (2018). Cabozantinib in patients with advanced and progressing hepatocellular carcinoma. *N. Engl. J. Med.* 379, 54–63. doi: 10.1056/NEJMoa1717002

Allemani, C., Matsuda, T., Di Carlo, V., Harewood, R., Matz, M., Nikšić, M., et al. (2018). Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet* 391, 1023–1075. doi: 10.1016/S0140-6736(17)33326-3

Amirouchene-Angelozzi, N., Swanton, C., and Bardelli, A. (2017). Tumor evolution as a therapeutic target. *Cancer Discov.* 7, 805–817. doi: 10.1158/2159-8290.CD-17-0343

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I. W. H., Ng, L. G., et al. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* 37, 38–44. doi: 10.1038/nbt.4314

Bidkhori, G., Benfeitas, R., Klevstig, M., Zhang, C., Nielsen, J., Uhlen, M., et al. (2018). Metabolic network-based stratification of hepatocellular carcinoma reveals three distinct tumor subtypes. *Proc. Natl. Acad. Sci. U.S.A.* 115, E11874–E11883. doi: 10.1073/pnas.1807305115

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and

mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492

Buetow, K. H., Meador, L. R., Menon, H., Lu, Y.-K., Brill, J., Cui, H., et al. (2019). High GILT expression and an active and intact MHC Class II antigen presentation pathway are associated with improved survival in melanoma. *J. I.* 203, 2577–2587. doi: 10.4049/jimmunol.1900476

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36, 411–420. doi: 10.1038/nbt.4096

Chalmers, Z. R., Connelly, C. F., Fabrizio, D., Gay, L., Ali, S. M., Ennis, R., et al. (2017). Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* 9:34. doi: 10.1186/s13073-017-0424-2

Chan, D. W., Chan, C.-Y., Yam, J. W. P., Ching, Y.-P., and Ng, I. O. L. (2006). Prickle-1 negatively regulates Wnt/beta-catenin pathway by promoting Dishevelled ubiquitination/degradation in liver cancer. *Gastroenterology* 131, 1218–1227. doi: 10.1053/j.gastro.2006.07.020

Chaudhary, K., Poirion, O. B., Lu, L., and Garmire, L. X. (2018). Deep learning–based multi-omics integration robustly predicts survival in liver cancer. *Clin. Cancer Res.* 24, 1248–1259. doi: 10.1158/1078-0432.CCR-17-0853

Craig, A. J., von Felden, J., Garcia-Lezana, T., Sarcognato, S., and Villanueva, A. (2020). Tumour evolution in hepatocellular carcinoma. *Nat. Rev. Gastroenterol. Hepatol.* 17, 139–152. doi: 10.1038/s41575-019-0229-4

Deng, T., Hu, B., Jin, C., Tong, Y., Zhao, J., Shi, Z., et al. (2021). A novel ferroptosis phenotype-related clinical-molecular prognostic signature for hepatocellular carcinoma. *J. Cell Mol. Med.* 25, 6618–6633. doi: 10.1111/jcmm.16666

Ding, J., Adiconis, X., Simmons, S. K., Kowalczyk, M. S., Hession, C. C., Marjanovic, N. D., et al. (2019). Systematic comparative analysis of single cell RNA-sequencing methods. *Genomics* 38, 737–746. doi: 10.1101/632216

Durante, M. A., Rodriguez, D. A., Kurtenbach, S., Sanchez, M. I., Decatur, C. L., Snyder, H., et al. (2020). Single-cell analysis reveals new evolutionary complexity in uveal melanoma. *Nat. Commun.* 11:496.

Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A. K., et al. (2015). MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* 16:278. doi: 10.1186/s13059-015-0844-5

Forero, A., Li, Y., Chen, D., Grizzle, W. E., Updike, K. L., Merz, N. D., et al. (2016). Expression of the MHC class II pathway in triple-negative breast cancer tumor cells is associated with a good prognosis and infiltrating lymphocytes. *Cancer Immunol. Res.* 4, 390–399. doi: 10.1158/2326-6066.CIR-15-0243

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Soft.* 33, 1–22. doi: 10.18637/jss.v033.i01

Gao, R., Bai, S., Henderson, Y. C., Lin, Y., Schalck, A., Yan, Y., et al. (2021). Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat. Biotechnol.* 39, 599–608. doi: 10.1038/s41587-020-00795-2

Haibe-Kains, B., Desmedt, C., Sotiriou, C., and Bontempi, G. (2008). A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all? *Bioinformatics* 24, 2200–2208. doi: 10.1093/bioinformatics/btn374

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* 14:7. doi: 10.1186/1471-2105-14-7

Harrell, F. E. (1982). Evaluating the yield of medical tests. *JAMA* 247, 2543–2546. doi: 10.1001/jama.1982.03320430047030

Huang, Y., Chen, S., Qin, W., Wang, Y., Li, L., Li, Q., et al. (2020). A novel RNA binding protein-related prognostic signature for hepatocellular carcinoma. *Front. Oncol.* 10:580513. doi: 10.3389/fonc.2020.580513

Jin, S., Li, R., Chen, M.-Y., Yu, C., Tang, L.-Q., Liu, Y.-M., et al. (2020). Single-cell transcriptomic analysis defines the interplay between tumor cells, viral infection, and the microenvironment in nasopharyngeal carcinoma. *Cell Res.* 30, 950–965. doi: 10.1038/s41422-020-00402-8

Kim, C., Gao, R., Sei, E., Brandt, R., Hartman, J., Hatschek, T., et al. (2018). Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing. *Cell* 173, 879.e13–893.e13. doi: 10.1016/j.cell.2018.03.041

Liang, J., Wang, D., Lin, H., Chen, X., Yang, H., Zheng, Y., et al. (2020). A novel ferroptosis-related gene signature for overall survival prediction in patients

with hepatocellular carcinoma. *Int. J. Biol. Sci.* 16, 2430–2441. doi: 10.7150/ijbs.45050

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The molecular signatures database hallmark gene set collection. *Cell Syst.* 1, 417–425. doi: 10.1016/j.cels.2015.12.004

Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., et al. (2021). Hepatocellular carcinoma. *Nat. Rev. Dis. Primers* 7:6. doi: 10.1038/s41572-020-00240-3

Llovet, J. M., Montal, R., Sia, D., and Finn, R. S. (2018). Molecular therapies and precision medicine for hepatocellular carcinoma. *Nat. Rev. Clin. Oncol.* 15, 599–616. doi: 10.1038/s41571-018-0073-4

Llovet, J. M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J.-F., et al. (2008). Sorafenib in advanced hepatocellular carcinoma. *N. Engl. J. Med.* 359, 378–390. doi: 10.1056/NEJMoa0708857

Long, J., Chen, P., Lin, J., Bai, Y., Yang, X., Bian, J., et al. (2019a). DNA methylation-driven genes for constructing diagnostic, prognostic, and recurrence models for hepatocellular carcinoma. *Theranostics* 9, 7251–7267. doi: 10.7150/thno.31155

Long, J., Wang, A., Bai, Y., Lin, J., Yang, X., Wang, D., et al. (2019b). Development and validation of a TP53-associated immune prognostic model for hepatocellular carcinoma. *EBioMedicine* 42, 363–374. doi: 10.1016/j.ebiom.2019.03.022

Losic, B., Craig, A. J., Villacorta-Martin, C., Martins-Filho, S. N., Akers, N., Chen, X., et al. (2020). Intratumoral heterogeneity and clonal evolution in liver cancer. *Nat. Commun.* 11:291. doi: 10.1038/s41467-019-14050-z

Ma, L., Hernandez, M. O., Zhao, Y., Mehta, M., Tran, B., Kelly, M., et al. (2019). Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer. *Cancer Cell* 36, 418.e6–430.e6. doi: 10.1016/j.ccell.2019.08.007

Maley, C. C., Aktipis, A., Graham, T. A., Sottoriva, A., Boddy, A. M., Janiszewska, M., et al. (2017). Classifying the evolutionary and ecological features of neoplasms. *Nat. Rev. Cancer* 17, 605–619. doi: 10.1038/nrc.2017.69

Marjanovic, N. D., Hofree, M., Chan, J. E., Canner, D., Wu, K., Trakala, M., et al. (2020). Emergence of a high-plasticity cell state during lung cancer evolution. *Cancer Cell* 38, 229.e13–246.e13. doi: 10.1016/j.ccell.2020.06.012

Massalha, H., Bahar Halpern, K., Abu-Gazala, S., Jana, T., Massasa, E. E., Moor, A. E., et al. (2020). A single cell atlas of the human liver tumor microenvironment. *Mol. Syst. Biol.* 16:e9682. doi: 10.15252/msb.20209682

Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28, 1747–1756. doi: 10.1101/gr.239244.118

McCaw, T. R., Li, M., Starenki, D., Cooper, S. J., Liu, M., Meza-Perez, S., et al. (2019). The expression of MHC class II molecules on murine breast tumors delays T-cell exhaustion, expands the T-cell repertoire, and slows tumor growth. *Cancer Immunol. Immunother.* 68, 175–188. doi: 10.1007/s00262-018-2262-5

Ni, Q., Chen, Z., Zheng, Q., Xie, D., Li, J.-J., Cheng, S., et al. (2020). Epithelial V-like antigen 1 promotes hepatocellular carcinoma growth and metastasis via the ERBB-PI3K-AKT pathway. *Cancer Sci.* 111, 1500–1513. doi: 10.1111/cas.14331

Pinter, M., Scheiner, B., and Peck-Radosavljevic, M. (2021). Immunotherapy for advanced hepatocellular carcinoma: a focus on special subgroups. *Gut* 70, 204–214. doi: 10.1136/gutjnl-2020-321702

Qin, X.-Y., Suzuki, H., Honda, M., Okada, H., Kaneko, S., Inoue, I., et al. (2018). Prevention of hepatocellular carcinoma by targeting MYCN-positive liver cancer stem cells with acyclic retinoid. *Proc. Natl. Acad. Sci. U.S.A.* 115, 4969–4974. doi: 10.1073/pnas.1802279115

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., et al. (2017). Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* 14, 979–982. doi: 10.1038/nmeth.4402

Rajewsky, N., Almouzni, G., Gorski, S. A., Aerts, S., Amit, I., et al. (2020). LifeTime and improving European healthcare through cell-based interceptive medicine. *Nature* 587, 377–386. doi: 10.1038/s41586-020-2715-9

Ramachandran, P., Dobie, R., Wilson-Kanamori, J. R., Dora, E. F., Henderson, B. E. P., Luu, N. T., et al. (2019). Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* 575, 512–518. doi: 10.1038/s41586-019-1631-3

Sarathi, A., and Palaniappan, A. (2019). Novel significant stage-specific differentially expressed genes in hepatocellular carcinoma. *BMC Cancer* 19:663. doi: 10.1186/s12885-019-5838-3

Sconocchia, G., Eppenberger-Castori, S., Zlobec, I., Karamitopoulou, E., Arriga, R., Coppola, A., et al. (2014). HLA Class II antigen expression in colorectal

carcinoma tumors as a favorable prognostic marker. *Neoplasia* 16, 31–42. doi: 10.1593/neo.131568

Sharma, A., Seow, J. J. W., Dutertre, C.-A., Pai, R., Blériot, C., Mishra, A., et al. (2020). Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. *Cell* 183, 377.e21–394.e21. doi: 10.1016/j.cell.2020.08.040

Sun, X., Yu, S., Zhang, H., Li, J., Guo, W., and Zhang, S. (2020). A signature of 33 immune-related gene pairs predicts clinical outcome in hepatocellular carcinoma. *Cancer Med.* 9, 2868–2878. doi: 10.1002/cam4.2921

Sun, Y., Wu, L., Zhong, Y., Zhou, K., Hou, Y., Wang, Z., et al. (2021). Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. *Cell* 184, 404.e16–421.e16. doi: 10.1016/j.cell.2020.11.041

The Cancer Genome Atlas Research Network, Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., et al. (2013). The cancer genome atlas pan-cancer analysis project. *Nat. Genet.* 45, 1113–1120. doi: 10.1038/ng.2764

Turajlic, S., Sottoriva, A., Graham, T., and Swanton, C. (2019). Resolving genetic heterogeneity in cancer. *Nat. Rev. Genet.* 20, 404–416. doi: 10.1038/s41576-019-0114-6

Van den Berge, K., Roux de Bézieux, H., Street, K., Saelens, W., Cannoodt, R., Saeys, Y., et al. (2020). Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.* 11:1201. doi: 10.1038/s41467-020-14766-3

Wang, J.-Y., Wang, X.-K., Zhu, G.-Z., Zhou, X., Yao, J., Ma, X.-P., et al. (2021). Distinct diagnostic and prognostic values of Glypicans gene expression in patients with hepatocellular carcinoma. *BMC Cancer* 21:462. doi: 10.1186/s12885-021-08104-z

Wang, S., and Liu, X. (2019). The UCSCXenaTools R package: a toolkit for accessing genomics data from UCSC Xena platform, from cancer multi-omics to single-cell RNA-seq. *JOSS* 4:1627. doi: 10.21105/joss.01627

Wang, Y., Mashock, M., Tong, Z., Mu, X., Chen, H., Zhou, X., et al. (2020). Changing technologies of RNA sequencing and their applications in clinical oncology. *Front. Oncol.* 10:447. doi: 10.3389/fonc.2020.00447

Yang, J. D., Hainaut, P., Gores, G. J., Amadou, A., Plymoth, A., and Roberts, L. R. (2019). A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat. Rev. Gastroenterol. Hepatol.* 16, 589–604. doi: 10.1038/s41575-019-0186-y

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Zhang, B., Tang, B., Gao, J., Li, J., Kong, L., and Qin, L. (2020). A hypoxia-related signature for clinically predicting diagnosis, prognosis and immune microenvironment of hepatocellular carcinoma patients. *J. Transl. Med.* 18:342. doi: 10.1186/s12967-020-02492-9

Zhang, J., Bajari, R., Andric, D., Gerthoffert, F., Lepsa, A., Nahal-Bose, H., et al. (2019). The international cancer genome consortium data portal. *Nat. Biotechnol.* 37, 367–369. doi: 10.1038/s41587-019-0055-9

Zhang, L., He, X., Liu, X., Zhang, F., Huang, L. F., Potter, A. S., et al. (2019). Single-cell transcriptomics in medulloblastoma reveals tumor-initiating progenitors and oncogenic cascades during tumorigenesis and relapse. *Cancer Cell* 36, 302.e7–318.e7. doi: 10.1016/j.ccell.2019.07.009

Zhang, Q., He, Y., Luo, N., Patel, S. J., Han, Y., Gao, R., et al. (2019). Landscape and dynamics of single immune cells in hepatocellular carcinoma. *Cell* 179, 829.e7–845.e7. doi: 10.1016/j.cell.2019.10.003

Zhao, S., Fung-Leung, W.-P., Bittner, A., Ngo, K., and Liu, X. (2014). Comparison of RNA-Seq and microarray in transcriptome profiling of activated T Cells. *PLoS One* 9:e78644. doi: 10.1371/journal.pone.0078644

Zheng, H., Pomyen, Y., Hernandez, M. O., Li, C., Livak, F., Tang, W., et al. (2018). Single-cell analysis reveals cancer stem cell heterogeneity in hepatocellular carcinoma. *Hepatology* 68, 127–140. doi: 10.1002/hep.29778

Zhu, Y., Wang, R., Chen, W., Chen, Q., and Zhou, J. (2020). Construction of a prognosis-predicting model based on autophagy-related genes for hepatocellular carcinoma (HCC) patients. *Aging* 12, 14582–14592. doi: 10.18632/aging.103507

Zuo, Q., He, J., Zhang, S., Wang, H., Jin, G., Jin, H., et al. (2021). PPARγ Coactivator-1α suppresses metastasis of hepatocellular carcinoma by inhibiting warburg effect by PPARγ-Dependent WNT/β-Catenin/Pyruvate dehydrogenase kinase isozyme 1 axis. *Hepatology* 73, 644–660. doi: 10.1002/hep.31280