Check for updates

Analysis on heterogeneity of hepatocellular carcinoma immune cells and a molecular risk model by integration of scRNA-seq and bulk RNA-seq

Xiaorui Liu[1†], Jingjing Li[1†], Qingxiang Wang[2], Lu Bai[1],
Jiyuan Xing[1], Xiaobo Hu[1], Shuang Li[3]* and Qinggang Li[1]*

[1]Department of Infection, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China,
[2]Department of physical examination&Blood collection Xuchang Blood Center, Xuchang, China,
[3]Bioinformatics R&D Department, Hangzhou Mugu Technology Co., Ltd, Hangzhou, China

Background: Studies have shown that hepatocellular carcinoma (HCC) heterogeneity is a main cause leading to failure of treatment. Technology of single-cell sequencing (scRNA) could more accurately reveal the essential characteristics of tumor genetics.

Methods: From the Gene Expression Omnibus (GEO) database, HCC scRNA-seq data were extracted. The FindCluster function was applied to analyze cell clusters. Autophagy-related genes were acquired from the MSigDB database. The ConsensusClusterPlus package was used to identify molecular subtypes. A prognostic risk model was built with the Least Absolute Shrinkage and Selection Operator (LASSO)−Cox algorithm. A nomogram including a prognostic risk model and multiple clinicopathological factors was constructed.

Results: Eleven cell clusters labeled as various cell types by immune cell markers were obtained from the combined scRNA-seq GSE149614 dataset. ssGSEA revealed that autophagy-related pathways were more enriched in malignant tumors. Two autophagy-related clusters (C1 and C2) were identified, in which C1 predicted a better survival, enhanced immune infiltration, and a higher immunotherapy response. LASSO−Cox regression established an eight-gene signature. Next, the HCCDB18, GSA14520, and GSE76427 datasets confirmed a strong risk prediction ability of the signature. Moreover, the low-risk group had enhanced immune infiltration and higher immunotherapy response. A nomogram which consisted of RiskScore and clinical features had better prediction ability.

Conclusion: To precisely assess the prognostic risk, an eight-gene prognostic stratification signature was developed based on the heterogeneity of HCC immune cells.

KEYWORDS

HCC, ScRNA-seq, riskscore, autophagy, molecular subtypes

Introduction

Primary liver cancer is a malignancy with a high degree of histological and biological heterogeneity and has therefore become a major public health problem (1). According to the Global Cancer Statistics Report, liver hepatocellular carcinoma (LIHC) was the sixth highest cancer worldwide in 2020, and its mortality rose to the third highest, accounting for 4.7% of all cancer cases and 830,000 deaths and 8.3% of all cancer deaths in the same period (2). In China, primary LIHC is the fourth most common malignant tumor and its mortality rate ranks the second in China due to historical factors, population region, and health conditions (3). Therefore, there is a strong clinical need for more effective strategies such as developing new therapeutic targets, biomarkers, and therapies for treating HCC, which remain key unmet needs for the treatment of hepatocellular cell carcinoma (HCC).

Tumor heterogeneity can be manifested as different pathological types, tumor stages, and differentiation degrees in clinical practice as well as different genomes and transcriptomes at the molecular level, which eventually lead to different sensitivities to chemotherapy and therapeutic drugs, thereby bringing great difficulties to cancer treatment (4, 5). In 2011, Navin et al. used mononuclear whole-genome amplification technology to amplify and sequence a total of 200 single nuclei in two cases of breast cancer *in situ* tissues and one case of liver metastasis tissues and analyzed their copy number changes to explain the population structure and evolution process of tumor cells (6). Patel et al. isolated 430 single cells from five brain tumor patients, conducted transcriptome analysis, and found that the internal changes in gene expression during different transcription processes were related to oncogene signaling, proliferation, immune response, and hypoxia (7). Dalerba et al. analyzed the gene expression of single cells in colonic epithelial carcinoma *in situ* tissues and normal colon tissues and observed that colon cancer tissues contained different subsets of cells and their transcripts were different from those in normal colon tissues (8).

Autophagy is a conserved lysosome-dependent pathway that degrades organelles, macromolecules, and cytoplasmic proteins in a dynamic, multistep process. Evidence suggests that autophagy has a role in tumor suppression in HCC (9, 10). For example, systemic Atg5-null mice and liver-specific Atg7−/− mice would develop benign hepatic adenomas (11), but Beclin-1 haploid deficiency could induce spontaneous HCC formation (12).

Based on the above analyses, we identified the immune heterogeneity of HCC from public databases applying bioinformatics methods and screened autophagy-related genes associated with the prognosis of HCC. Furthermore, the classification and prognostic signature of the autophagy-related genes in HCC samples were analyzed to further supplement the prognostic markers of HCC and provide new insights for clinical targeted drug therapy.

Material and methods

HCC data of public databases

A sum of 360 HCCs in TCGA dataset (TCGA-LIHC), 242 samples in the GSE14520 dataset, 10 HCCs in the single-cell sequencing database (GSE149614), and 389 samples in the HCCDB18 dataset were acquired from The Cancer Genome Atlas (TCGA) (13), Gene Expression Omnibus (GEO) (14), and Hepatocellular Carcinoma Database (HCCDB) (15).

Autophagy-related genes were obtained from the MSigDB (https://www.gsea-msigdb.org/gsea/index.jsp) database.

Data control

The Seurat package (16) was used to set the expression of each gene in at least three cells (each cell expressing at least 250 genes) to filter a single cell. The proportion of rRNA and mitochondria was further calculated by PercentageFeatureSet function, and genes expressed in each cell were more than 200 but fewer than 6,000, the percentage of mitochondria was fewer than 25%, and the UMI of each cell was at least greater than 1,000. To screen highly variable genes, the FindVariableFeatures function was used, followed by scaling and PCA dimensionality reduction for all genes using the ScaleData function.

Cell type annotation

The cells were clustered using FindNeighbor and FindCluster (Dim = 20, Resolution = 0.1). The FindAllMarker function was conducted to select marker genes. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation was performed using R Package ClusterProfiler (17).

Analysis of autophagy-related pathways

Autophagy-related pathways were obtained from GSEA (http://www.gsea-msigdb.org/gsea/index.jsp) and analyzed using the ssGSEA of GSVA package (18).

Clustering

According to the standard of p< 0.05, oxidative stress-related genes with prognosis of HCC were filtered *via* univariate Cox survival analysis using coxph function of the R package. Then, molecular subtypes were performed on each TCGA-LIHC dataset sample *via* the ConsensusClusterPlus 1.52.0 (19). Pam arithmetic and "spearman" distance were utilized to complete 500 bootstraps

with every bootstrap containing ≥80% of TCGA-LIHC dataset specimens. Cluster number k was between 2 and 10, and the optimum k was identified as per cumulative distribution function (CDF) and AUC. Survival curves (K–M curves) between molecular subtypes were then analyzed for difference. In addition, differences in the distribution of clinical characteristics between molecular subtypes were compared and the chi-square test was conducted, and p<0.05 indicated statistical significance.

Mutation analysis

A waterfall plot was generated to explore the detailed single-nucleotide variant (SNV) characteristics between molecular subtypes *via* the "mutect2" (20) function in R software.

Construction and evaluation of a prognostic risk model for HCC

LASSO–Cox regression was conducted in the glmnet package in R language to select prognostic genes (21). By penalized maximum likelihood, glmnet fits generalized linear and similarity models. The regularization path is the calculation of the LASSO or elastic net penalty on the value (on a logarithmic scale) of the regularization parameter lambda (22). The genes included in the model and the optimal value of the penalty coefficient λ were determined through running a 1,000-time 10-fold cross-validation probability. Subsequently, coefficients of prognostic genes were extracted by Cox multivariate regression analysis, and the gene expression levels were used to calculate the risk score as the survival risk score of each patient by the following formula:

$$RiskScore = \sum_{k=0}^{n} \beta i \times Expi$$

where βi is the Cox hazard ratio coefficient of mRNA and Expi represents the gene expression level. TCGA-LIHC samples were divided into high-risk and low-risk groups according to the risk score, with the median risk score as the threshold. At the same time, GSE14520 and HCCDB18 were used to analyze the robustness and effectiveness of the prognostic risk model. Kaplan–Meier (K–M) curves combined with the log-rank test were applied to analyze survival differences among different risk groups. The timeROC package was employed to determine the area under the receiver operating characteristic curve (AUC) to predict the 5-, 4-, 3-, 2-, and 1-year survival rates, respectively.

Nomogram

To further evaluate the predictive efficacy of the risk score model, a nomogram was constructed by combining other clinicopathological characteristics of HCC patients (including family history, TNM stage, age, histological grade, gender, etc.) using the RMS R package (23). The predictive accuracy of the nomogram was assessed by calculating the C-index, which quantifies the degree of agreement between the actual and predicted survival rates. The abscissa was the 1-, 3-, and 5-year survival probability of each patient predicted according to the nomogram, and the ordinate was the actual 1-, 3-, and 5-year survival probability of each patient, with the 45° line representing the optimal prediction.

Cell-type identification using estimating relative subsets of RNA transcripts

Cell-type Identification using Estimating Relative Subsets of RNA Transcripts (CIBERSORT) analyses were utilized to compare diversities in different immunocytes in molecular subtypes. Wilcox test analyses were completed to identify the difference of 22 kinds of infiltrating immunocyte score among molecular subtypes. The "ggplot2" package (24) was used to visualize the distributional status of the diversities in 22 kinds of infiltration immunocytes.

Estimate

The R software Estimation of Stromal and Immune cells in Malignant Tumors using Expression data (ESTIMATE) arithmetic (25) was utilized to compute overall stroma level (StromalScore), immunocyte infiltration (ImmuneScore), and combination (ESTIMATEScore) of samples in the TCGA-LIHC cohort using Wilcox test analysis for analyzing the differences between molecular subtypes.

Tumor immune dysfunction and exclusion

The Tumor Immune Dysfunction and Exclusion (TIDE) (26, 27) algorithm (http://tide.dfci.harvard.edu) evaluated three cell types that limit T-cell invasion into tumors, including IFNG, myeloid suppressor cells (MDSC), and M2 subtypes of tumor-associated macrophages (TAM.M2), dysfunction of tumor infiltration cytotoxic T lymphocytes (CTL) (Dysfunction), and exclusion of CTL by immunosuppressive factors (Exclusion).

Drug sensitivity analysis

pRRophetic (28) was used to predict the sensitivity of erlotinib, sunitinib, paclitaxel, VX-680, TAE684, and crizotinib to IC50. We used Sangerbox for assisting data analysis (29).

Results

Definition of cell clusters and dimensionality reduction

The flowchart of this work is shown in Figure S1. The "Seurat" function and PercentageFeatureSet function were performed to screen 33,117 cells from the scRNA-seq dataset. The UMI was obviously correlated with numbers of mRNAs (Figure S2A). Figures S2B, C display the samples before and after quality control. PCA revealed that most HCC patients were distributed by cluster (Figure S2D). We conducted the "ScaleData" function to scale all genes extracted from the scRNA-seq dataset (GSE149614) and performed PCA dimensionality reduction to find anchor points. Finally, 11 clusters were found based on FindNeighbors and FindClusters functions.

Cell annotation of 11 clusters was performed in terms of classical markers of immune cells. Macrophage (cluster 1; markers: CD163, CD68, CD14), T cells (clusters 2, 8; markers: CD2, CD3D, CD3E, and CD3G), B cells (cluster 9; markers: CD19, CD79A, and MS4A1), plasma cells (cluster 4; markers: CD79A and JSRP1), mast cells (cluster 10; markers: TPSAB1 and CPA3), fibroblasts (cluster 7; markers: ACTA2, PDGFRB, and NOTCH3), endothelial cells (cluster 6; markers: PECAM1), and hepatocellular carcinomas (cluster 0, 3, 5; GPC3, CD24, and MDK) were clustered according to immune cell markers (Figure S3).

An overview of the single cells from 10 samples in the GSE149614 dataset are listed in Figure 1A. All the cells were classified into 11 clusters (Figure 1B). Eleven clusters were labeled as eight cell types by immune cell marker genes (Figure 1C). The top five genes with the most prominent contributions are shown in Figure 1D. Next, the CNV of eight cell types was predicted using the CopyCat package to identify 13,617 malignant (tumor) cells and 19,500 no_malignant (normal) cells (Figure 1E). Moreover, we calculated the proportion of malignant and no_malignant cells in 10 samples (Figure 1F).

Analysis of autophagy pathways

ssGSEA showed that autophagy pathways were activated in malignant tumors (Figure 2A). Next, ssGSEA of five autophagy pathways in TCGA-LIHC dataset showed that three autophagy pathways scored lower in tumors compared to the normal ones (Figure 2B), which was the opposite as shown in Figure 2A. Thus, we further analyzed the autophagy pathway scores in grade 1 to grade 4 (Figure S4). Here, with the increase in tumor grade, autophagy pathway scores were continuously decreased, indicating that the body was abnormal and the cells had

autophagy; however, as the disease deteriorated, autophagy was gradually weakened.

Identification of autophagy-related clusters

Based on the results in Figure 2B, 253 genes in the intersection of TCGA-LIHC dataset and three autophagy pathways were analyzed by limma to obtain 38 differentially expressed genes. Three hundred sixty samples in TCGA-LIHC were clustered using the ConsensusClusterPlus package based on 38 genes. According to cumulative distribution function (CDF) and delta area (Figures 3A, B), two clusters (Clust1 and Clust2) were obtained when k = 2 (Figure 3C). K–M analysis demonstrated that HCC patients in Clust2 tended to have a shorter survival than those in Clust1 in TCGA dataset (Figure 3D) and HCCDB18 dataset (Figure 3E). Clust2 had more women, a higher T stage and clinical stage, and a poorer tumor grade (Figure 4).

Genome analysis and functional enrichment analysis

Molecular characteristics of TCGA-LIHC were obtained from a previous study (30). Clust1 presented a lower aneuploidy score, homologous recombination defects, and fraction altered (Figure 5A). The differences in gene mutation between clusters were analyzed; the top 10 genes are shown in Figure 5B. TP53 and TTN had obvious differences between two clusters (Figure 5B). GSEA showed that pathways such as cell cycle were activated in Clust1 (Figure 6A), and five of 10 pathways associated with tumorigenesis had higher scores in Clust2 (Figure 6B).

Analysis of the tumor immune microenvironment

Firstly, CIBERSORT analysis indicated that seven of 22 immune cells had a significant difference between two clusters (Figure 7A). Then, ESTIMATE analysis showed that Clust1 had higher scores of StromalScore, ImmuneScore, and ESTIMATEScore (Figure 7B). We then evaluated the 47 immune check gene expressions, and 18 immune checkpoint genes had obviously high expressions in Clust2 than those in Clust1 (Figure 7C). Next, the scores of Toll-like receptor signaling pathway, natural killer cell-mediated cytotoxicity, antigen processing, and presentation were calculated using ssGSEA, and here it has been observed that the Toll-like receptor score and NK cytotoxicity score were higher in Clust1 than those in Clust2 (Figure 7D). TIDE was lower in Clust1 than in Clust2 (Figure 7E), suggesting that patients in Clust1 were

FIGURE 1

Definition of cell clusters. **(A)** t-SNE of 10 samples in the GSE149614 dataset. **(B)** t-SNE of 11 cell subgroups. **(C)** t-SNE of eight cell types. **(D)** Top five genes that made the most significant contribution. **(E)** t-SNE of malignant and no_malignant. **(F)** The proportion of malignant and no_maliganant in 10 samples in the GSE149614 dataset.
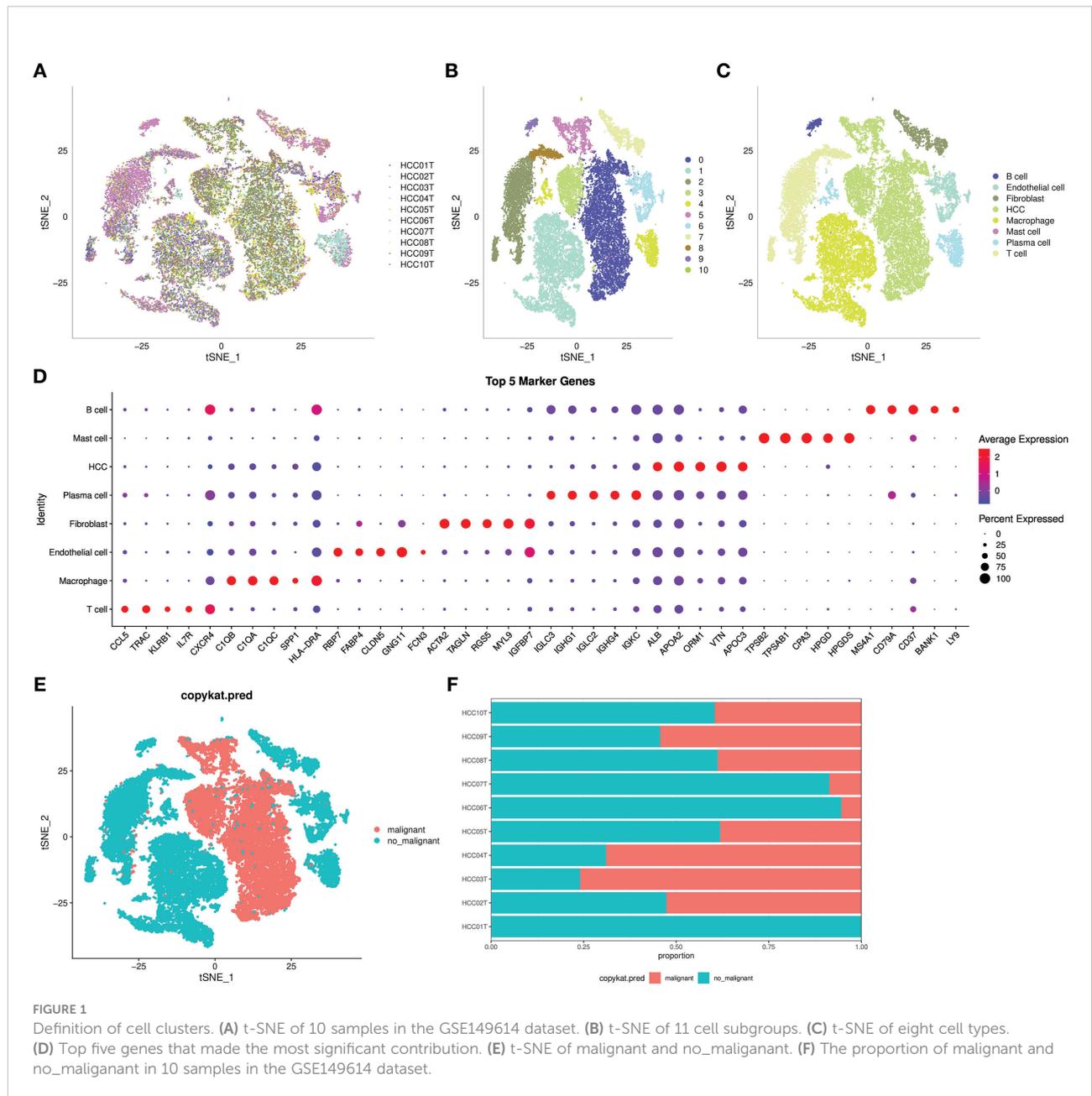
more likely to benefit from immunotherapy and more patients in Clust1 responded to immunotherapy (Figure 7F).

Identification of hub autophagy related genes and RiskScore

A sum of 344 differentially expressed genes were identified between Clust1 and Clust2 (Figure 8A). Next, univariate Cox survival analysis determined 137 genes associated with prognosis

(Figure 8B). The LASSO–Cox regression module was used to build a prognostic signature based on the expression matrix of the 137 genes. Subsequently, we identified an eight-gene signature module according to the optimal λ value (Figures 8C, D). RiskScore of HCC patients based on eight genes (Figure 8E) was calculated using the following formula:

$$RiskScore = 0.011*RACGAP1 - 0.024*HAO2 - 0.055*OGDHL$$
$$+ 0.122*ZWINT - 0.069*CFHR3 - 0$$
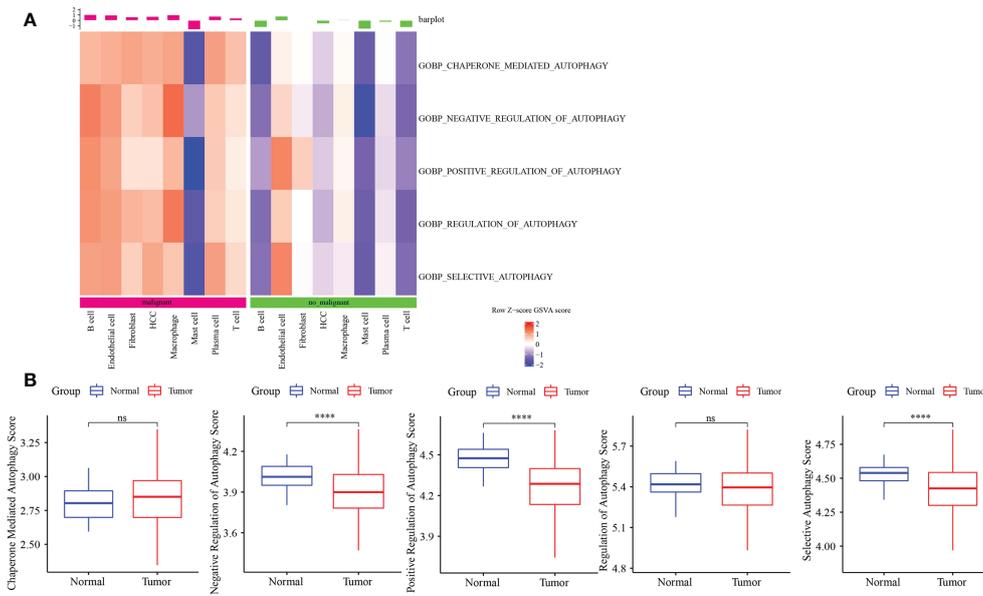
FIGURE 2

Autophagy-related pathway analysis. **(A)** Five autophagy-related pathways were activated in malignant. **(B)** The difference analysis of autophagy-related pathways scores in normal and tumor. ****p<0.0001, ns, no significance.
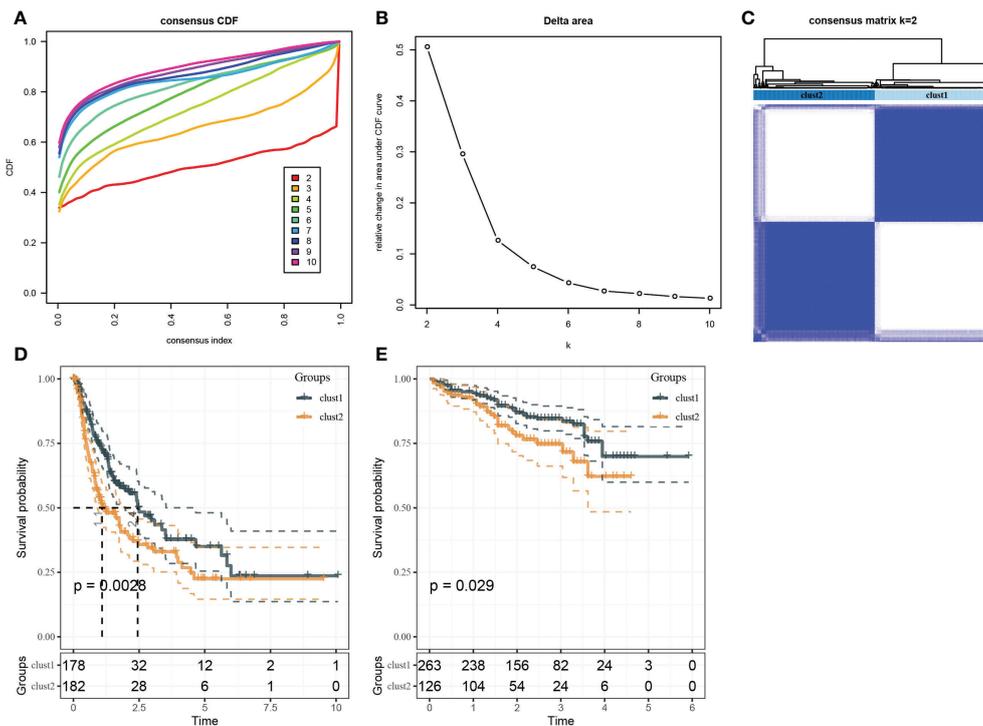
FIGURE 3

Identification of molecular subtypes. **(A)** Cumulative distribution function. **(B)** Delta area. **(C)** Heatmap of sample clustering when k = 2. **(D)** K−M survival analysis of Clust1 and Clust2 in TCGA-LIHC dataset. **(E)** K−M survival analysis of Clust1 and Clust2 in the HCCDB18 dataset.
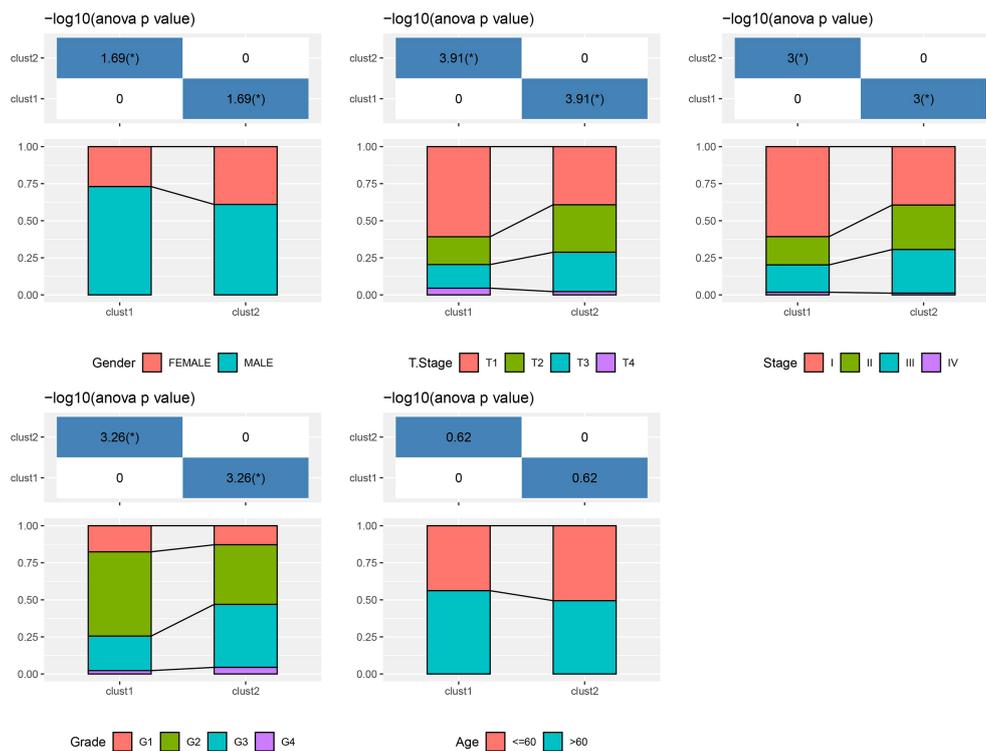
FIGURE 4
The distribution of clinical features, including gender, T stage, stage, grade and age, in Clust1 and Clust2.

$$044*CYP2C9 + 0.07*SFN - 0.005*SPP2$$

Validation of the prognostic model

The median RiskScore was the cutoff in classifying the samples into low-risk (RiskScore< median) and high-risk (RiskScore > median) groups. ROC and survival analyses were performed in TCGA-LIHC (Figure 9A), HCCDB18 (Figure 9B), GSE14520 (Figure 9C), and GSE76427 datasets (Figure 9D). The results revealed that the accuracy of the model was higher in predicting the 1-, 2-, 3-, 4-, and 5-year survival rates in the above datasets, as all values of the area under the curve (AUC) were greater than 0.6. Results of the Kaplan–Meier survival analysis showed an overall survival higher in the low-risk group than the high-risk group. Female patients, those with a clinically advanced stage, younger samples (<=60), and clust2 showed a higher RiskScore (Figure 10).

Analysis of immune infiltration and immunotherapy

CIBERSORT analysis indicated that 11 of 22 immune cells were significantly higher in the low group that those in the high group

(Figure 11A). ESTIMATE analysis showed that the low group had higher StromalScore and ESTIMATEScore (Figure 11B). Moreover, 31 immune checkpoint genes had obviously high expressions in the high group than in the low group (Figure 11C). Moreover, TIDE, IFNG, MDSC, Exclusion, and TAM.M2 were lower in the low group than in the high group, but Dysfunction was higher in the low group (Figure 11D), suggesting that the low group was more likely to benefit from immunotherapy. IC50 of erlotinib, sunitinib, paclitaxel, VX-680, TAE684, and crizotinib were higher in the high group, suggesting that patients in the high group were more sensitive to those drugs (Figure 11E).

Nomogram

Univariate and multivariate Cox analyses indicated that Stage and RiskScore were independent prognostic factors (Figures 12A, B). Next, we constructed a prognostic nomogram based on Stage and RiskScore to predict the 1-, 3-, and 5-year overall survival of HCC patients (Figure 12C). The calibration curve proved that the prognostic nomogram was reliable and accurate (Figure 12D). The results of the AUC indicated that among other clinical variables, the RiskScore and

FIGURE 5

Genome analysis. **(A)** The analysis of aneuploidy score, homologous recombination defects, fraction altered, number of segments, and non-silent mutation rate in Clust1 and Clust2. **(B)** Top 15 mutation genes in Clust1 and Clust2. ****p<0.0001, ns, no significance.
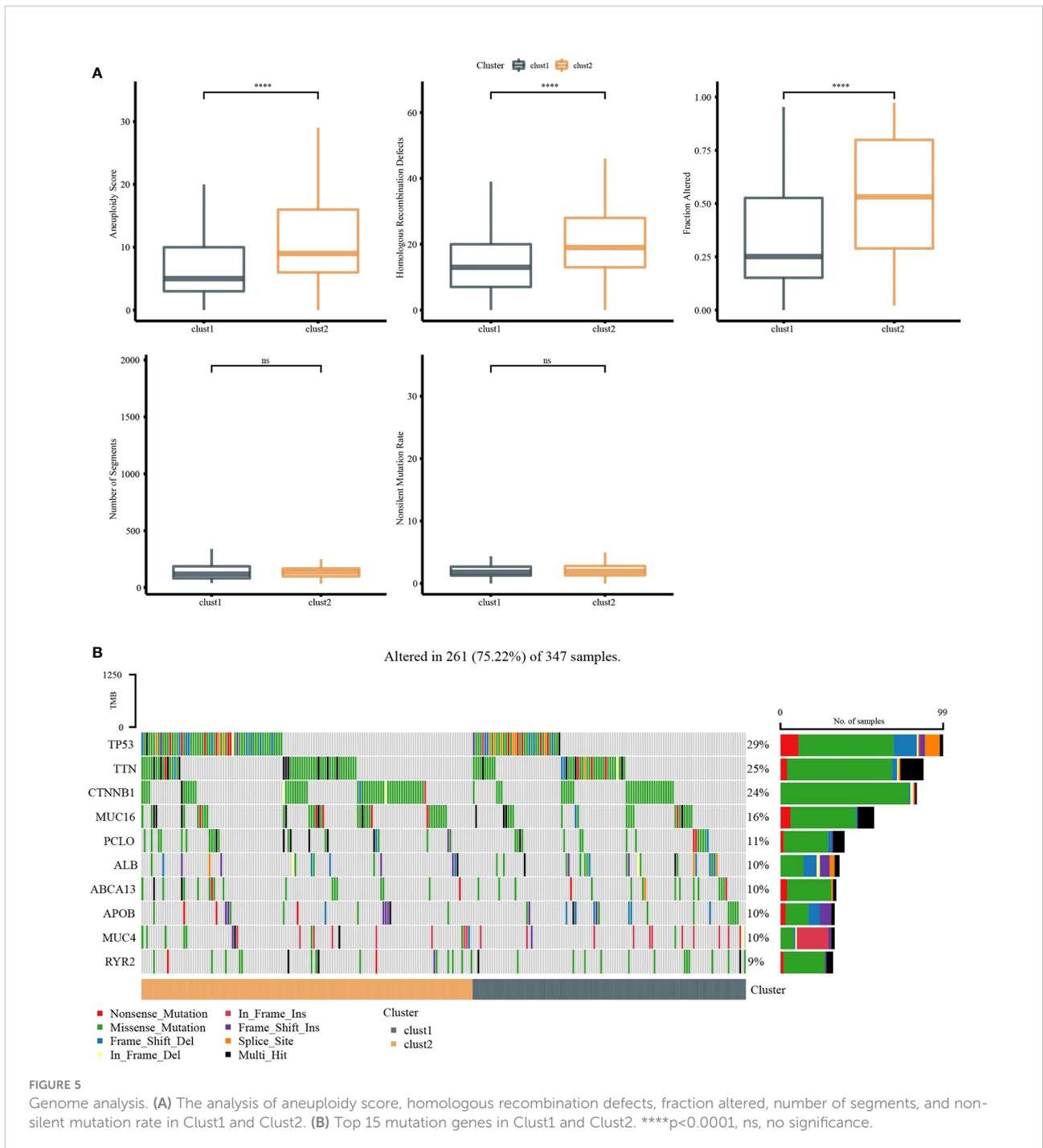
Nomogram served as accurate prognostic indicators in clinical decision-making process (Figure 12E).

Discussion

Although there are several methods to treat HCC, the effectiveness of treatment is limited by a late diagnosis at an advanced stage, which is usually accompanied by a high rate of disease recurrence (31). Therefore, identifying patients with high risk is crucial for doctors to determine the use of aggressive treatments. The TNM staging system developed by AJCC is a standard system for evaluating the prognosis of HCC patients (32, 33). However, clinical outcomes of patients with the same TNM stage could be greatly different sometimes. The current study investigated the prognostic significance of the autophagy

FIGURE 6
Functional enrichment analysis. **(A)** GSEA demonstrated that pathways, such as cell cycle, were activated in Clust1. **(B)** Five of 10 pathways associated with tumorigenesis had higher scores in Clust2. **p<0.01, ***p<0.001, ****p<0.0001, ns, no significance.

FIGURE 7
Analysis of immune infiltration. **(A)** Analysis of 22 immune cells using CIBERSORT. **(B)** Analysis of immune infiltration using ESTIMATE. **(C)** The expression levels of 42 immune check genes between Clust1 and Clust2. **(D)** The differences in Toll-like receptor signaling pathway score, natural killer cell-mediated cytotoxicity score, and antigen processing and presentation score between Clust1 and Clust2. **(E)** The differences in TIDE between Clust1 and Clust2. **(F)** Responses to immunotherapy between Clust1 and Clust2. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns, no significance.

FIGURE 8

Identification hub autophagy-related genes. **(A)** Volcano plot of differentially expressed genes identified from autophagy-related pathways between Clust1 *vs.* Clust2. **(B)** Volcano plot of differentially expressed genes identified using univariate Cox analysis. **(C)** Lambda trajectory of differentially expressed genes. **(D)** Confidence interval under lambda. **(E)** Eight hub genes, including three risk genes and five protective genes, were obtained.
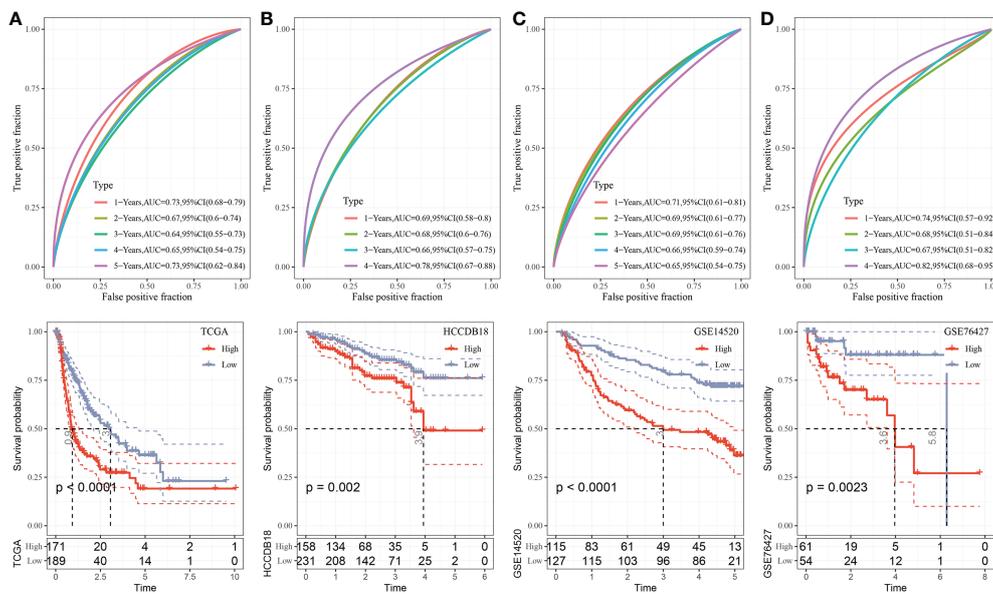
FIGURE 9

Validation of RiskScore. **(A)** ROC and K−M survival analysis of RiskScore in TCGA-LIHC dataset. **(B)** ROC and K−M survival analysis of RiskScore in the HCCDB18 dataset. **(C)** ROC and K−M survival analysis of RiskScore in the GSE14520 dataset. **(D)** ROC and K−M survival analysis of RiskScore in the GSE76427 dataset.
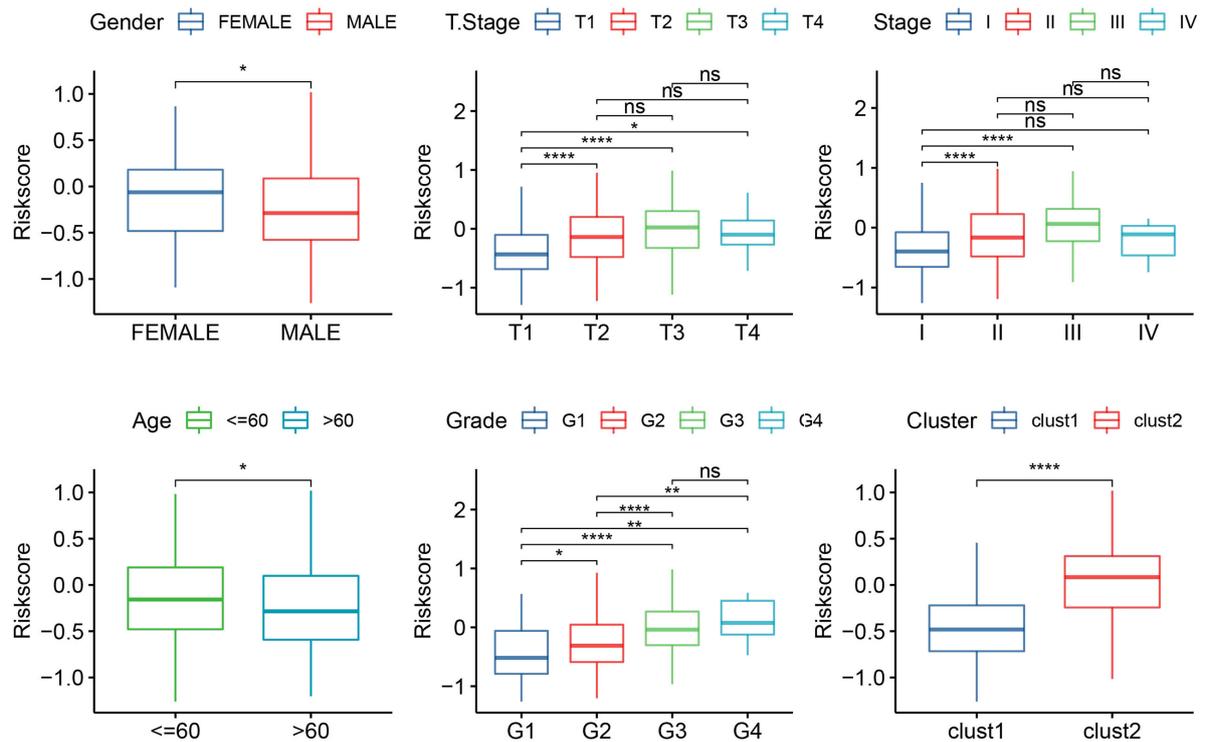
FIGURE 10
The RiskScore analysis in clinical features, including gender, T stage, stage, grade, age, and clusters. *p<0.05, **p<0.01, ****p<0.0001, ns, no significance.

score and developed a new nomogram model for predicting the OS of HCC patients. The autophagy score as an independent prognostic factor plays a key role in HCC survival. After internal verification, we found that nomogram prediction results were better than the TNM staging system.

Targeting autophagy is a new strategy in cancer immunotherapy (34). A study has shown that autophagy is associated with adaptive and innate immune responses and that it could be induced by immune receptors such as nucleotide oligomerization domain-like receptors (NLRs) and Toll-like receptors (35). Autophagy is involved in the lymphocyte development and antigen presentation, making autophagy a potential target in improving cancer immunotherapy (36). It has been further demonstrated that autophagy could facilitate cancer cells to effectively evade immune responses and immune surveillance. Thus, to prevent immune escape of cancer cells, targeting autophagy has gradually become a novel immunotherapeutic strategy (34). In this study, the autophagy score was developed to predict the response to immunotherapy, which further suggested that the autophagy score may be used to

differentiate clinical patients and select more suitable treatment options.

A study reported a three-gene signature for predicting the survival outcome of HCC using scRNA and bulk RNA (37). Single-cell sequencing identifies three hub genes in HCC (38). However, in this study, we combined scRNA-seq and bulk-RNA to analyze HCC. The RiskScore model had eight genes, namely, HAO2, RACGAP1, OGDHL, ZWINT, CFHR3, CYP2C9, SFN, and SPP2. In human HCC tissues, several studies reported an obviously lower HAO2 expression, which was associated with a worse HCC prognosis (39); moreover, HAO2 was overexpressed and HCC cell invasion, migration, and proliferation were inhibited (40, 41). RACGAP1 is frequently overexpressed and associated with shorter survival time of HCC patients (42, 43). Silencing OGDHL would induce lipogenesis and affect sorafenib's chemosensitization effect on HCC cells (44). ZWINT upregulation showed a significant association with unfavorable survivals and clinicopathological features of HCC patients (45). CFHR3 overexpression was correlated with a favorable prognosis for HCC patients (46). CYP2C9 is
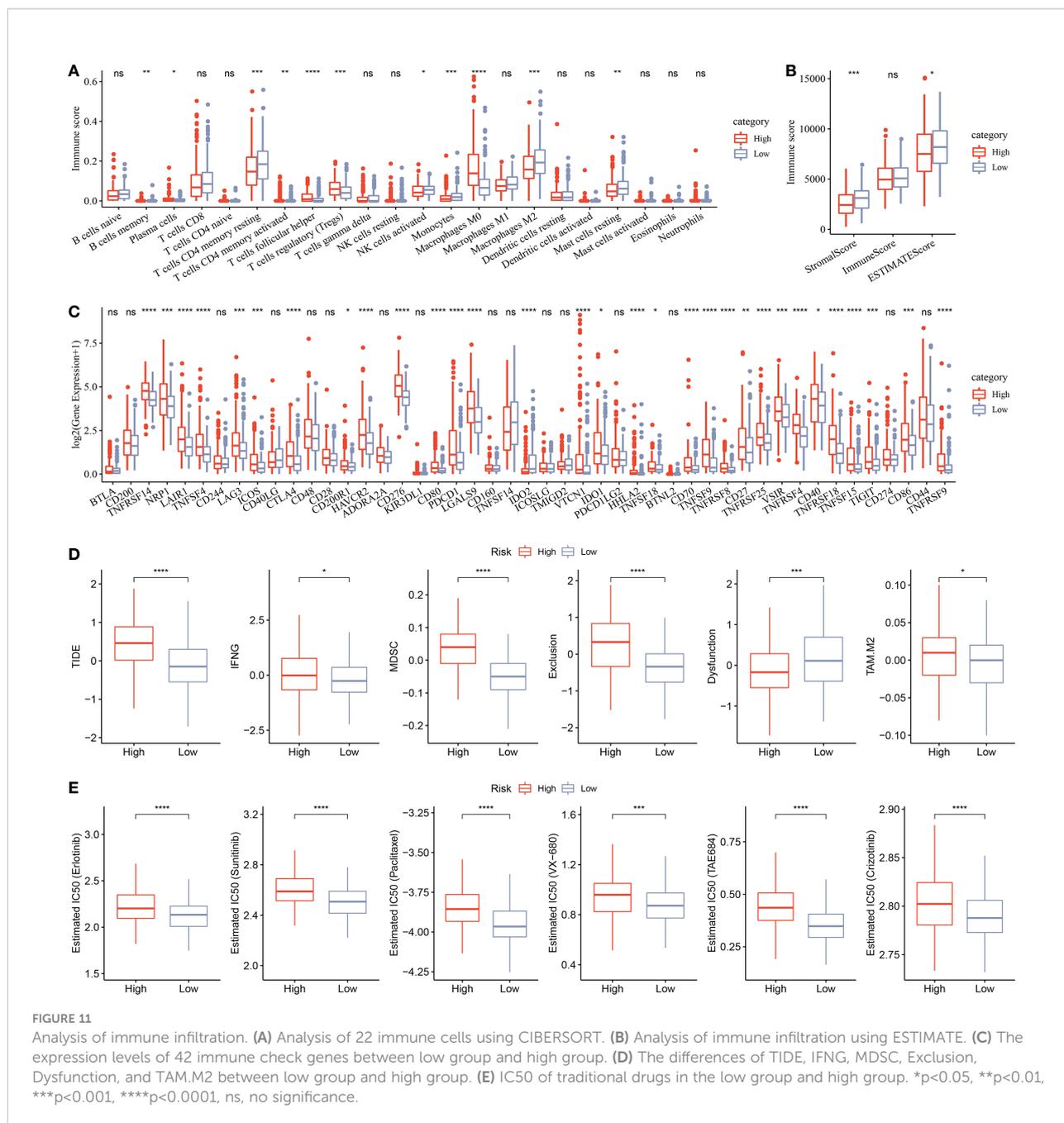
FIGURE 11

Analysis of immune infiltration. **(A)** Analysis of 22 immune cells using CIBERSORT. **(B)** Analysis of immune infiltration using ESTIMATE. **(C)** The expression levels of 42 immune check genes between low group and high group. **(D)** The differences of TIDE, IFNG, MDSC, Exclusion, Dysfunction, and TAM.M2 between low group and high group. **(E)** IC50 of traditional drugs in the low group and high group. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns, no significance.

involved in the metabolism of many carcinogens and drugs and is downregulated in HCC (47). SFN significantly inhibited the proliferation of HepG2 cells (48). So far, there was no study discussing the potential functions of SPP2 in HCC, but results indicated that SPP2 was associated with the stability of spliceosome and chromatin (49). Hence, we speculated that SPP2 dysregulation may lead to carcinogenesis and disease progression.

Some limitations existed in this study. Firstly, the sample lacked clinical follow-up information; thus, factors such as the

presence of other health conditions were not considered during the identification of the biomarkers. Secondly, the results obtained by bioinformatics analysis alone were not convincing enough, which requires further experimental verification. In addition, the molecular processes and signaling pathways obtained from TCGA cases alone are not sufficient and should be confirmed by further studies.

In conclusion, we constructed a prognostic prediction model which consisted of HAO2, RACGAP1, OGDHL, ZWINT, CFHR3, CYP2C9, SFN, and SPP2 for HCC, which provided
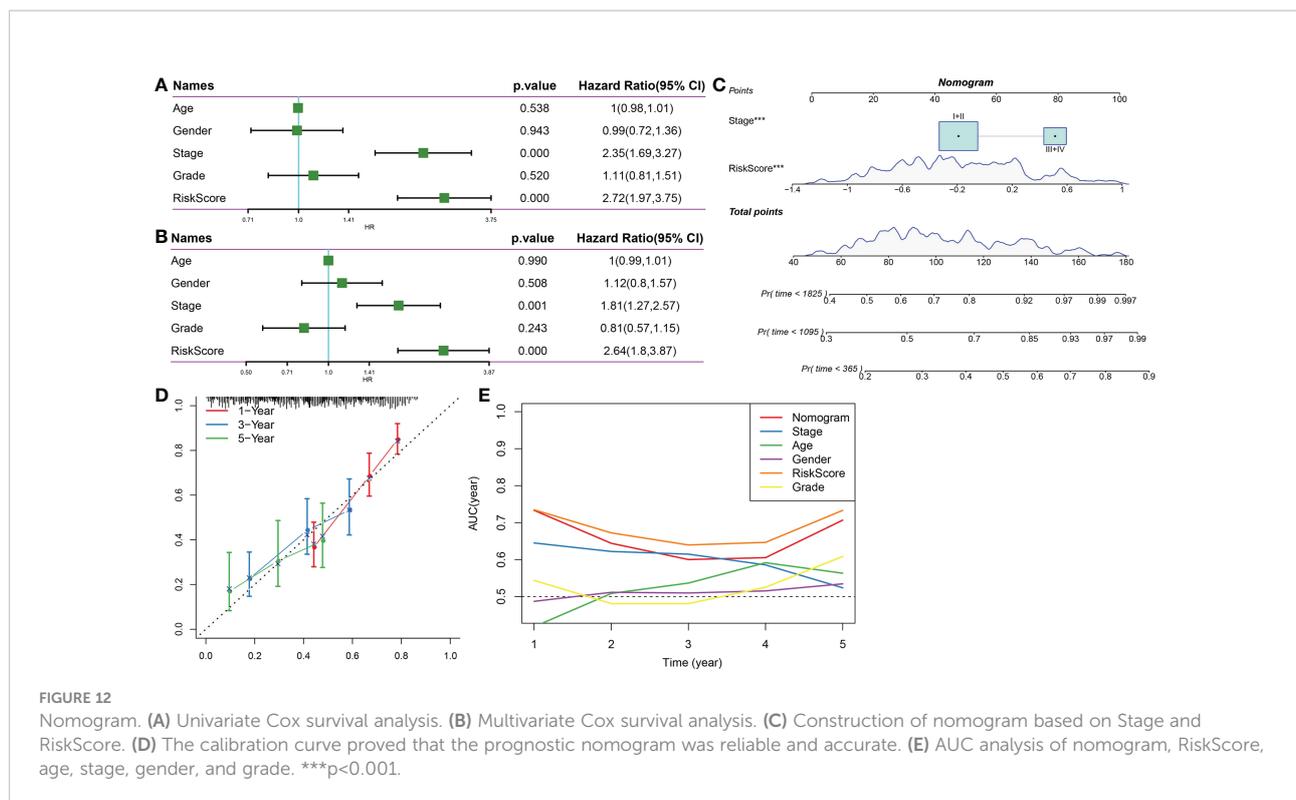
FIGURE 12

Nomogram. **(A)** Univariate Cox survival analysis. **(B)** Multivariate Cox survival analysis. **(C)** Construction of nomogram based on Stage and RiskScore. **(D)** The calibration curve proved that the prognostic nomogram was reliable and accurate. **(E)** AUC analysis of nomogram, RiskScore, age, stage, gender, and grade. ***p<0.001.

new ideas for the prognostic treatment of HCC patients. However, the main limitation of our results was that our study was conducted in a public database, and further experiments are needed for in-depth study.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

XL and JL conducted the statistical analyses of the data and prepared the draft manuscript. QW and LB edited the manuscript. JX, XH, SL, and QL provided critical comments to the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The present study was supported by the Henan Medical Science and Technology Key Project: The Effect of Hyperammonemia on Hepatocyte Regeneration and Metabolism After Liver Failure and Its Clinical Treatment (20170219).

Conflict of interest

Author SL was employed by Hangzhou Mugu Technology Co., Ltd.

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2022.1012303/full#supplementary-material

References

1. Feng M, Pan Y, Kong R, Shu S. Therapy of primary liver cancer. *Innovation (Cambridge (Mass))* (2020) 1(2):100032. doi: 10.1016/j.xinn.2020.100032

2. 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(3):209–49. doi: 10.3322/caac.21660

3. Zhou M, Wang H, Zeng X, Yin P, Zhu J, Chen W, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: A systematic analysis for the global burden of disease study 2017. *Lancet (London England)* (2019) 394 (10204):1145–58. doi: 10.1016/S0140-6736(19)30427-1

4. Navin NE. Tumor evolution in response to chemotherapy: Phenotype versus genotype. *Cell Rep* (2014) 6(3):417–9. doi: 10.1016/j.celrep.2014.01.035

5. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet* (2014) 46(3):225–33. doi: 10.1038/ng.2891

6. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, et al. Tumour evolution inferred by single-cell sequencing. *Nature* (2011) 472(7341):90–4. doi: 10.1038/nature09807

7. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Sci (New York NY)* (2014) 344(6190):1396–401. doi: 10.1126/science.1254257

8. Dalerba P, Kalisky T, Sahoo D, Rajendran PS, Rothenberg ME, Leyrat AA, et al. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol* (2011) 29(12):1120–7. doi: 10.1038/nbt.2038

9. Yazdani HO, Huang H, Tsung A. Autophagy: Dual response in the development of hepatocellular carcinoma. *Cells* (2019) 8(2):91. doi: 10.3390/cells8020091

10. Akkoç Y, Gözüaçık D. Autophagy and liver cancer. *Turkish J Gastroenterol Off J Turkish Soc Gastroenterol* (2018) 29(3):270–82. doi: 10.5152/tjg.2018.150318

11. Takamura A, Komatsu M, Hara T, Sakamoto A, Kishi C, Waguri S, et al. Autophagy-deficient mice develop multiple liver tumors. *Genes Dev* (2011) 25 (8):795–800. doi: 10.1101/gad.2016211

12. Yue Z, Jin S, Yang C, Levine AJ, Heintz N. Beclin 1, an autophagy gene essential for early embryonic development, is a haploinsufficient tumor suppressor. *Proc Natl Acad Sci United States America* (2003) 100(25):15077–82. doi: 10.1073/pnas.2436255100

13. Hutter C, Zenklusen JC. The cancer genome atlas: Creating lasting value beyond its data. *Cell* (2018) 173(2):283–5. doi: 10.1016/j.cell.2018.03.042

14. Toro-Domínguez D, Martorell-Marugán J, López-Domínguez R, García-Moreno A, González-Rumayor V, Alarcón-Riquelme ME, et al. ImaGEO: Integrative gene expression meta-analysis from GEO database. *Bioinf (Oxford England)* (2019) 35(5):880–2. doi: 10.1093/bioinformatics/bty721

15. Lian Q, Wang S, Zhang G, Wang D, Luo G, Tang J, et al. HCCDB: A database of hepatocellular carcinoma expression atlas. *Genomics Proteomics Bioinf* (2018) 16(4):269–75. doi: 10.1016/j.gpb.2018.07.003

16. Zhai Y, Li G, Li R, Chang Y, Feng Y, Wang D, et al. Single-cell RNA-sequencing shift in the interaction pattern between glioma stem cells and immune cells during tumorigenesis. *Front Immunol* (2020) 11:581209. doi: 10.3389/fimmu.2020.581209

17. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an r package for comparing biological themes among gene clusters. *Omics J Integr Biol* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118

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

19. Wilkerson MD, Hayes DN. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26 (12):1572–3. doi: 10.1093/bioinformatics/btq170

20. Pei S, Liu T, Ren X, Li W, Chen C, Xie Z. Benchmarking variant callers in next-generation and third-generation sequencing analysis. *Briefings Bioinf* (2021) 22(3):bbaa148. doi: 10.1093/bib/bbaa148

21. Tibshirani R. The lasso method for variable selection in the cox model. *Stat Med* (1997) 16(4):385–95. doi: 10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3

22. Goeman JJ. L1 penalized estimation in the cox proportional hazards model. *Biometrical J Biometrische Z* (2010) 52(1):70–84. doi: 10.1002/bimj.200900028

23. Liu TT, Li R, Huo C, Li JP, Yao J, Ji XL, et al. Identification of CDK2-related immune forecast model and ceRNA in lung adenocarcinoma, a pan-cancer analysis. *Front Cell Dev Biol* (2021) 9:682002. doi: 10.3389/fcell.2021.682002

24. Ito K, Murphy D. Application of ggplot2 to pharmacometric graphics. *CPT: pharmacometrics Syst Pharmacol* (2013) 2(10):e79. doi: 10.1038/psp.2013.56

25. Yang P, Chen W, Xu H, Yang J, Jiang J, Jiang Y, et al. Correlation of CCL8 expression with immune cell infiltration of skin cutaneous melanoma: Potential as a prognostic indicator and therapeutic pathway. *Cancer Cell Int* (2021) 21(1):635. doi: 10.1186/s12935-021-02350-8

26. Fu J, Li K, Zhang W, Wan C, Zhang J, Jiang P, et al. Large-Scale public data reuse to model immunotherapy response and resistance. *Genome Med* (2020) 12 (1):21. doi: 10.1186/s13073-020-0721-z

27. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med* (2018) 24(10):1550–8. doi: 10.1038/s41591-018-0136-1

28. Geeleher P, Cox N, Huang RS. pRRophetic: an r package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PloS One* (2014) 9(9):e107468. doi: 10.1371/journal.pone.0107468

29. Shen W, Song Z, Xiao Z, Huang M, Shen D, Gao P, et al. Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* (2022) 1(3):e36. doi: 10.1002/imt2.36

30. Jager LE. Hazards in the plating industry. *Occup Health Rev* (1966) 18(2):3–10.

31. Camp RL, Dolled-Filhart M, Rimm DL. X-Tile: A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res an Off J Am Assoc Cancer Res* (2004) 10(21):7252–9. doi: 10.1158/1078-0432.CCR-04-0713

32. Kamarajah SK, Frankel TL, Sonnenday C, Cho CS, Nathan H. Critical evaluation of the American joint commission on cancer (AJCC) 8th edition staging system for patients with hepatocellular carcinoma (HCC): A surveillance, epidemiology, end results (SEER) analysis. *J Surg Oncol* (2018) 117(4):644–50. doi: 10.1002/jso.24908

33. Abdel-Rahman O. Assessment of the discriminating value of the 8th AJCC stage grouping for hepatocellular carcinoma. *HPB Off J Int Hepato Pancreato Biliary Assoc* (2018) 20(1):41–8. doi: 10.1016/j.hpb.2017.08.017

34. Yao C, Ni Z, Gong C, Zhu X, Wang L, Xu Z, et al. Rocaglamide enhances NK cell-mediated killing of non-small cell lung cancer cells by inhibiting autophagy. *Autophagy* (2018) 14(10):1831–44. doi: 10.1080/15548627.2018.1489946

35. Liu G, Pei F, Yang F, Li L, Amin AD, Liu S, et al. Role of autophagy and apoptosis in non-Small-Cell lung cancer. *Int J Mol Sci* (2017) 18(2):367. doi: 10.3390/ijms18020367

36. Zhu J, Wang M, Hu D. Development of an autophagy-related gene prognostic signature in lung adenocarcinoma and lung squamous cell carcinoma. *PeerJ* (2020) 8:e8288. doi: 10.7717/peerj.8288

37. Lu J, Chen Y, Zhang X, Guo J, Xu K, Li L. A novel prognostic model based on single-cell RNA sequencing data for hepatocellular carcinoma. *Cancer Cell Int* (2022) 22(1):38. doi: 10.1186/s12935-022-02469-2

38. Wang H, Fu Y, Da BB, Xiong G. Single-cell sequencing identifies the heterogeneity of CD8+ T cells and novel biomarker genes in hepatocellular carcinoma. *J healthcare Eng* (2022) 2022:8256314. doi: 10.1155/2022/8256314

39. Mattu S, Fornari F, Quagliata L, Perra A, Angioni MM, Petrelli A, et al. The metabolic gene HAO2 is downregulated in hepatocellular carcinoma and predicts metastasis and poor survival. *J Hepatol* (2016) 64(4):891–8. doi: 10.1016/j.jhep.2015.11.029

40. Li Y, Zhang M, Li X, Wang Y, Wang Y, Li Y, et al. Hydroxyacid oxidase 2 (HAO2) inhibits the tumorigenicity of hepatocellular carcinoma and is negatively regulated by miR-615-5p. *J Immunol Res* (2022) 2022:5003930. doi: 10.1155/2022/5003930

41. Zhuo H, Xia J, Zhang J, Tang J, Han S, Zheng Q, et al. CircASPH promotes hepatocellular carcinoma progression through methylation and expression of HAO2. *Front Oncol* (2022) 12:911715. doi: 10.3389/fonc.2022.911715

42. Yang XM, Cao XY, He P, Li J, Feng MX, Zhang YL, et al. Overexpression of rac GTPase activating protein 1 contributes to proliferation of cancer cells by reducing hippo signaling to promote cytokinesis. *Gastroenterology* (2018) 155 (4):1233–49.e22. doi: 10.1053/j.gastro.2018.07.010

43. Wang SM, Ooi LL, Hui KM. Upregulation of rac GTPase-activating protein 1 is significantly associated with the early recurrence of human hepatocellular carcinoma. *Clin Cancer Res an Off J Am Assoc Cancer Res* (2011) 17(18):6040–51. doi: 10.1158/1078-0432.CCR-11-0557

44. Dai W, Xu L, Yu X, Zhang G, Guo H, Liu H, et al. OGDHL silencing promotes hepatocellular carcinoma by reprogramming glutamine metabolism. *J Hepatol* (2020) 72(5):909–23. doi: 10.1016/j.jhep.2019.12.015

45. Lin T, Zhang Y, Lin Z, Peng L. ZWINT is a promising therapeutic biomarker associated with the immune microenvironment of hepatocellular carcinoma. *Int J Gen Med* (2021) 14:7487–501. doi: 10.2147/IJGM.S340057

46. Liu J, Li W, Zhao H. CFHR3 is a potential novel biomarker for hepatocellular carcinoma. *J Cell Biochem* (2020) 121(4):2970–80. doi: 10.1002/jcb.29551

47. Yu D, Green B, Marrone A, Guo Y, Kadlubar S, Lin D, et al. Suppression of CYP2C9 by microRNA hsa-miR-128-3p in human liver cells and association with hepatocellular carcinoma. *Sci Rep* (2015) 5:8534. doi: 10.1038/srep08534

48. Zou X, Qu Z, Fang Y, Shi X, Ji Y. Endoplasmic reticulum stress mediates sulforaphane-induced apoptosis of HepG2 human hepatocellular carcinoma cells. *Mol Med Rep* (2017) 15(1):331–8. doi: 10.3892/mmr.2016.6016

49. Warkocki Z, Schneider C, Mozaffari-Jovin S, Schmitzová J, Höbartner C, Fabrizio P, et al. The G-patch protein Spp2 couples the spliceosome-stimulated ATPase activity of the DEAH-box protein Prp2 to catalytic activation of the spliceosome. *Genes Dev* (2015) 29(1):94–107. doi: 10.1101/gad.253070.114