Discover Oncology

Analysis

Effects of ER-phagy regulatory genes on the microenvironment of hepatocellular carcinoma: a comprehensive analysis

Rongchang Zhao[1] · Dan Ding[2] · Minhui Bao[2] · Yan Ding[1] · Rongjie Ding[1] · Jun Liu[1] · Yu Li[1] · Chunrong Zhu[3]

Received: 16 January 2025 / Accepted: 9 May 2025
Published online: 17 May 2025

Abstract
The relationships between gene regulatory functions and hepatocellular carcinoma (HCC) occurrence and progression are constantly being clarified. However, tumour microenvironment complexity has hindered the classification of the role of genes. A comprehensive analysis to further clarify gene functions could provide additional benefits to HCC patients. In the present study, we combined single-cell sequencing data, Mendelian randomization, and bioinformatics analysis for comprehensive analysis. After the study was completed we found that T cell, dendritic cell (DC), macrophage and monocyte contents and the interaction between immune cells in the HCC microenvironment differed between the microvascular invasion-positive (MVI+) and microvascular invasion-negative (MVI-) groups. Mendelian randomization analysis indicated that causal relationships between several endoplasmic reticulum autophagy (ER-phagy) genes and T cell, DC, macrophage and monocyte contents. Single-cell sequencing data were used to validate the association of these genes with immune cells in the microenvironment. Based on the above results, we preliminarily elucidated the potential role of ER autophagy in the HCC microenvironment. Furthermore, a prognostic model was constructed using these causal association genes, which could accurately predict the prognosis and survival of HCC patients.

Keywords Hepatocellular carcinoma · Single-cell · MVI · ER-phagy · Mendelian randomization analysis

1 Introduction

According to the latest report on cancer mortality and morbidity in 185 countries released by the World Health Organization, there are approximately 841,000 new cases of liver cancer and 781,000 deaths from liver cancer worldwide each year, and liver cancer ranks first among digestive system tumours in men, accounting for 10.2% of related deaths [1]. These data suggest that breakthroughs in the clinical treatment of liver cancer patients are urgently needed. However, because most patients have advanced-stage disease or metastases at the time of diagnosis, they have lost the opportunity for radical treatment. Moreover, the disease can develop rapidly, the prognosis is extremely poor, and the mortality rate is high [2]; even if some types of radiotherapy or chemotherapy are administered, the curative effect is still not satisfactory, and drug resistance and relapse readily occur. Survival rates are low, with 5-year survival rates ranging from only 3% to 5% [3]. To enhance the overall survival of patients with liver cancer, it is imperative to implement more personalized

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12672-025-02649-2.

✉ Minhui Bao, yaq1395117@163.com | [1]Department of Oncology, Taixing People's Hospital, Taixing, China. [2]Department of Intensive Care Unit, Taixing People's Hospital, Taixing, China. [3]Department of Oncology, The First Affiliated Hospital of Soochow University, Suzhou, China.

and precision-based therapeutic approaches. The ultimate developmental trajectory of a tumor is contingent upon the dynamic alterations in its microenvironment, which encompass not only the metabolic profile variations but also the functional and quantitative modifications of the interacting cellular components. A comprehensive understanding of the distinct survival microenvironments of tumor cells at various progression stages, coupled with the identification of underlying factors contributing to poor prognostic outcomes, is essential for developing more individualized treatment strategies.

Hepatocellular carcinoma, the most prevalent form of liver cancer, is characterized by microvascular invasion when malignant cells infiltrate the minute vascular structures of the liver, including branches of the portal vein, hepatic artery, and bile ducts. MVI serves as direct pathological evidence of the aggressive behavior of liver cancer and constitutes a significant prognostic factor contributing to the poor long-term survival rates observed in HCC patients [4, 5]. The emergence of metastatic lesions during tumor progression indicates a progressive deterioration of the body's immune surveillance system, with the occurrence of MVI being no exception. The immune microenvironment, composed of various immune cells, plays a particularly crucial role in this pathological process. Immune cells can enter the tumour microenvironment from the blood, altering the composition and proportion of immune cells in the immune microenvironment. When MVI exists, it is worth in-depth exploration which key genes affect the distribution ratio of these immune cells and the changes of the tumor itself. MVI and endoplasmic reticulum stress (ERS) exhibit a significant correlation in the progression of hepatocellular carcinoma through the regulation of angiogenesis, immune evasion, and metabolic reprogramming [6]. ERS constitutes a critical component in maintaining endoplasmic reticulum homeostasis. It denotes the stress response elicited by the accumulation of unfolded/misfolded proteins or the disruption of calcium ion homeostasis within the endoplasmic reticulum under internal and external environmental pressures, which activates the unfolded protein response (UPR) [7]. However, if the UPR fails to restore homeostasis, endoplasmic reticulum autophagy is initiated to eliminate redundant endoplasmic reticulum structures and deleterious proteins [8]. Consequently, it can be inferred that endoplasmic reticulum stress and endoplasmic reticulum autophagy are mutually influential, and the specific association between endoplasmic reticulum autophagy and MVI warrants further investigation. The relationships between endoplasmic reticulum stress and the UPR and tumours have been described by many scholars [9–11]. ER-phagy, serving as the ultimate homeostatic regulatory mechanism of the endoplasmic reticulum, has been systematically investigated by Jacob E. Corn's research team. Through comprehensive genome-wide screening utilizing HCT116 CRISPRi cells, the research group successfully identified the regulatory genes associated with this process [12]. Interference with ER-phagy dependent on UFMylation was found to lead to the accumulation of misfolded proteins by inducing the UPR via IRE1α signalling. Taken together, these findings suggest that ER-phagy also plays an important role in the occurrence and development of tumours. In HCC, extensive research has elucidated that ER-phagy regulates sorafenib-induced ferroptosis through the FAM134B pathway [13]. Furthermore, the induction of ER-phagy has been shown to mitigate endoplasmic reticulum stress, thereby preventing HCC progression [14]. Moreover, ER-phagy is intricately associated with immune regulation within the HCC tumor microenvironment [15]. These findings collectively underscore the pivotal role of ER-phagy in HCC pathogenesis. Nevertheless, significant knowledge gaps persist regarding its interaction with the immune microenvironment. For instance, immune cells may undergo ER-phagy influenced by tumor cells or other metabolic factors, potentially altering the proportion of specific cell types. Given the critical role of immune cells as delineated above, comprehensive research is imperative to elucidate these underlying mechanisms. Mendelian randomization analyses use genetic variants strongly associated with exposure as instrumental variables to assess the causal relationship between exposure factors and outcomes.However, direct analysis of causal relationships between gene expression and traits is difficult. Single nucleotide polymorphisms (SNPs) represent variations of individual nucleotides (A, T, C, or G) at specific genomic loci. These genetic variations can be employed as independent variables in linear regression analyses, with mRNA expression levels serving as dependent variables. Through the utilization of extensive individual data points as samples, this methodology facilitates the elucidation of associations between specific SNP loci and their corresponding mRNA expression profiles [16]. Expression quantitative trait loci (eQTLs) are regions on chromosomes that specifically regulate mRNA and protein expression levels, and their levels are proportional to the mRNA and protein expression levels. Therefore, Mendelian randomization analysis combining SNPs and eQTLs can be used to determine the causal associations between gene expression and some traits.For example, it can be used to analyse the relationship between ER-phagy regulatory genes (ER-phagy REGs) and HCC. In addition to tumour cells, immune cells are equally important for treatment, as mentioned above. These immune cells can also theoretically undergo ER-phagy in response to stimulus signals. Previously, it was difficult to extract immune cells from the microenvironment for analysis, but with the continuous development of single-cell technology, this problem has been simplified. The advent of single-cell RNA sequencing has made it possible to perform specific analyses of cell populations at the single-cell level. This

approach involves analysis of the contents of biological macromolecules such as RNA/DNA, proteins, and metabolites that are released from individual cells upon rupture of the cell membrane and can reveal the changes that make each cell type unique [17, 18]. Therefore, this technology can be applied to identify immune cells in the microenvironment. Based on these findings, we can stratify HCC patients based on their MVI status, systematically analyze the differential characteristics of the immune microenvironment among distinct subgroups, and identify causal genes associated with tumor progression. These insights will provide innovative strategies for comprehensive patient evaluation in subsequent clinical management and facilitate the development of more precise therapeutic interventions.

Building upon the previously established theoretical framework. In this study, we first performed single-cell analysis of MVI + and MVI- samples from HCC patients to identify differences in immune cell features and cell communication signals in the microenvironment. ER-phagy REGs, HCC and immune cells were subsequently analysed via Mendelian randomization. The obtained causal genes were combined with the single-cell analysis results of HCC for further screening of the genes that are correlated with the number of immune cells in the tumour microenvironment. Next, these genes were analysed in clusters on the basis of the data of HCC patients from The Cancer Genome Atlas (TCGA) and used to group these patients. We constructed a prognostic model on the basis of the differentially expressed genes between the groups and found that the model had good predictive performance. Validation with International Cancer Genome Consortium (ICGC) data suggested that the predictive performance of the model was better in Asian HCC patients than in other populations.

2 Materials and methods

2.1 Preliminary single-cell analysis

Single-cell data for HCC in GSE242889 were obtained from the GEO database [19]. The Seurat package is used to control the quality of these data. First, the red blood cells are filtered. Then, nFeature_RNA > 200 & nFeature_RNA < 5000 & percent.mt < 20 & nCount_RNA < 100,000 were set to screen the data of each sample in R. Harmony was used to integrate the data and distinguish between MVI + and MVI- samples. In addition, the de-batch effect was also performed. The RunTSNE function is used to cluster cells. The SingleR package was subsequently used to annotate the cell clusters according to the integrated data to obtain information on immune cells in the microenvironment.

2.2 Mendelian randomization analysis

The labels of immune cells with quantitative differences in the microenvironment were further searched in the Genome-wide association study (GWAS) database. Two hundred genes related to the regulation of ER-phagy were obtained from a previously published study [12]. Through the application of the TwoSampleMR package in R, the eQTLs of ER-phagy REGs from the GWAS database were subsequently used as instrumental variables, and the numbers of immune cells and HCC cells were subsequently used as outcome variables for batch analysis to obtain causal genes. The HCC cell number was retrieved, and the eQTLs of causal genes and HCC were analysed via reverse Mendelian randomization to determine the direction of the influence of genes on HCC. To obtain independent SNPs, we performed clustering to remove SNPs in linkage disequilibrium (LD) ($r\ 2 < 0.001$). Traits were used at the genome-wide level ($p < 5 \times 10 - 8$). The instrument strength of each SNP was assessed via F statistics = (β/SE) 2. An F statistic > 10 suggests that a variable can be used as a suitable analytical tool variable [20].

2.3 Analysis and integration of gene and single-cell data

The immune cells in the tumour microenvironment were subdivided according to the consensus marker genes and the genes identified in previous studies. Cell communication was analysed through the CellChat program package. The samples were then differentiated to identify differences in immune cell content between MVI + and MVI- patients and differences in cell communication. The correlations between the genes related to immune cell causality obtained via Mendelian randomization analysis and the contents of the above immune cell subsets were subsequently analysed. The ER-phagy REGs with causal relationships between the tumour microenvironment and immune cells were subjected to further screening. The correlations between the genes related to immune cell causality obtained via Mendelian randomization analysis and the contents of the above immune cell subsets were subsequently analysed. Furthermore, ER-phagy

REGs that have a causal relationship between the tumour microenvironment and immune cells were identified. GEPIA database (http://gepia.cancer-pku.cn/) was used to perform prognostic analysis of some genes.

2.4 Extraction of gene expression data and clustering grouping

Gene expression and clinicopathological data of HCC patients were downloaded from the TCGA database (https://cancergenome.nih.gov/). Similarly, appropriate datasets containing HCC patient gene expression and clinical data were downloaded from the ICGC database (https://icgc.org/content/icgc-home-0). The k-means clustering algorithm was used to group data obtained from the TCGA database on the basis of patterns of genes associated with the proportion of immune cells. Differences in clinical features (T, N, M, stage, grade, gender, and age) between clusters and differences in gene expression are shown in heatmaps. Subsequently, differential expression analysis of the genes and analysis of survival differences between different groups were carried out (set the parameters to logFCfilter = 1, fdrFilter = 0.0001).

2.5 Prognostic model construction and performance validation

Least absolute shrinkage and selection operator (LASSO) analysis of the differential genes mentioned above, and suitable genes were identified via Cox regression analysis to construct a prognostic model, which was defined as the training dataset. The model calculation formula was as follows: βgene1 × Expressiongene1 + β gene2 × Expressiongene2 + βgene3 × Expressiongene3 + … + βgenen × Expressiongenen. Here, β refers to the coefficient value in the multifactor Cox analysis, and Expressiongene refers to the expression level of a gene in patients with liver cancer. Next, receiver operating characteristic (ROC) curves were drawn for the prediction model, and area under the ROC curve (AUC) values were calculated for 1-, 3-, and 4-year survival. Finally, the predictive performance of the model was verified with gene expression and survival data from the Asian HCC population from the ICGC database, which was defined as the test dataset, and the 1-, 3-, and 4-year survival AUC values were also calculated.

2.6 Comprehensive analysis of the risk score, and clinicopathological features

In the training dataset, univariate Cox regression analysis and multivariate Cox regression analysis were used to analyse the correlations of age, gender, stage, and the risk score with prognosis. These clinicopathological features were then integrated with the risk score to construct a nomogram to improve clinical utility. The nomogram performance was also verified.

3 Results

3.1 *Differences in the immune microenvironment between MVI + and MVI- patients*

The dataset encompasses single-cell sequencing samples derived from five patients, including three MVI + and two MVI- cases, which were subjected to comprehensive analysis. Subsequent to data integration and batch effect correction (Supplementary Material 1), an extensive analytical procedure was implemented. Utilizing cellular clustering analysis, we identified 26 distinct clusters within the tumor microenvironment of both MVI + and MVI- samples (Fig. 1A). After automatic annotation by singleR, the contents of T cells, DCs, macrophages and monocytes among immune cells, not among all liver cells, were significantly different (Fig. 1B). Data quality control is provided in Supplementary Materials 2. The expression of various immune cell marker genes is shown in Supplementary Materials 3.

3.2 ER-phagy REGs and T cells

After the T-cell data were extracted, 10 clusters were obtained when the subgroups were subdivided (Fig. 2A and B), and the number of these clusters varied among samples with different MVI conditions (Fig. 2C). After marker gene annotation [21], CD8 + effector T cells, CD4 + memory T cells, CD4 + Tfh cells, CD4 + effector T cells, γδ T cells and CD8 + naive T cells were identified (Fig. 2D). There were also differences in the numbers of these cells between MVI- and MVI + patients (Fig. 2E). The ebi-a-GCST90001603 label was obtained after T cells were retrieved from the GWAS database, which contains information on a total of 3653 samples containing 15,195,758 SNPs. A study with Mendelian randomization analysis of

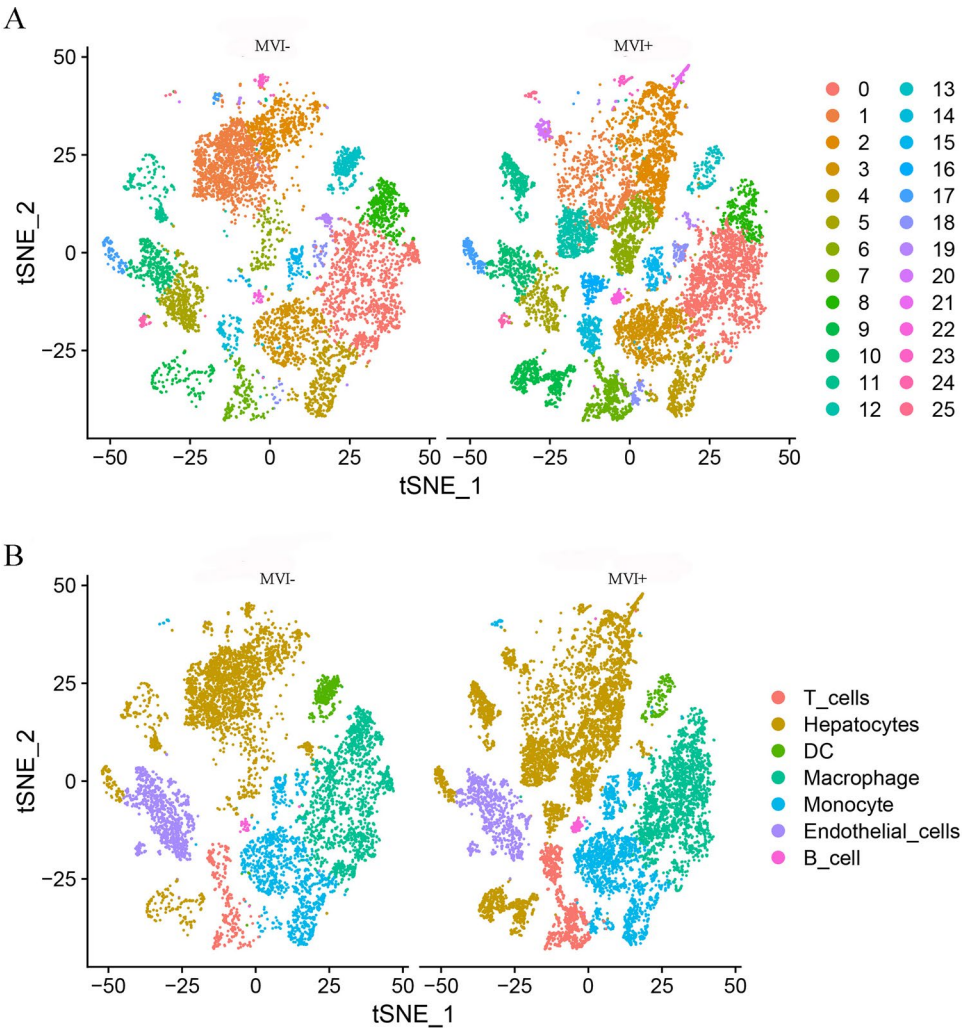


Fig. 1 Microenvironmental differences between MVI + and MVI- samples. **A** The content of 26 clusters was correlated with the status of MVI. **B** The content of identified immune cells (T cells, DCs, macrophages and monocytes) differed between MVI + and MVI- samples

ER-phagy REGs revealed that a total of 8 genes have a causal relationship with the number of T cells (Fig. 3). The results of the cell communication analysis suggested that there were also differences among the different samples. CD8 effector T cell interacted with macrophages, monocytes, Endothelial cell, DC and other CD8 effector T cells in MVI- microenvironment (Fig. 4A), and it was most closely associated with macrophages with 13 ligand-receptor sites and least associated with DC with 4 ligand-receptor sites (Fig. 4B). CD8 effector T cell interacted more closely with CD4 + T cells and B cells in MVI + microenvironment (Fig. 4C), and it has six receptor ligand binding sites with CD4 + Tfh, four with CD4 + memory T cell, and two with B cell (Fig. 4D). Subsequent single-cell analysis showed that RPS5 was positively correlated with the contents of CD4 + Tfh cells, CD4 + memory T cells and CD8 + effector T cells (Fig. 4E, F and J); GNAS was negatively correlated with the content of CD4 + effector T cells and positively correlated with that of CD8 + effector T cells (Fig. 4G and K); RSRC2 and FOXK2 were negatively correlated with the content of CD4 + memory T cells (Fig. 4H and I); and ATG10 was positively correlated with the content of CD8 + effector T cells (Fig. 4L).

3.3 ER-phagy REGs and macrophages

There is no label directly related to the number of macrophages in the GWAS database, only the labels of Macrophage receptor MARCO levels (ebi-a-GCST90010150, 1,301 samples, 18,166,693 SNPs) and macrophage mannose receptor 1 (prot-a-1937,3,301 samples, 10,534,735 SNPs). These two types of data and the eQTL data of the ER-phagy genes were analysed via Mendelian randomization, and a total of 18 causally associated genes were obtained (Fig. 5A and B). The macrophages in the HCC microenvironment can be further divided into 9 clusters (Fig. 6A). According to the annotation of marker genes provided in the literature [22–25], four types of cells, FOLR2 + macrophage, TREM2 + SPP1 + macrophage, Kupffer cell, and MT1G + macrophage, were distinguished, and their contents in MVI + and MVI- samples were significantly

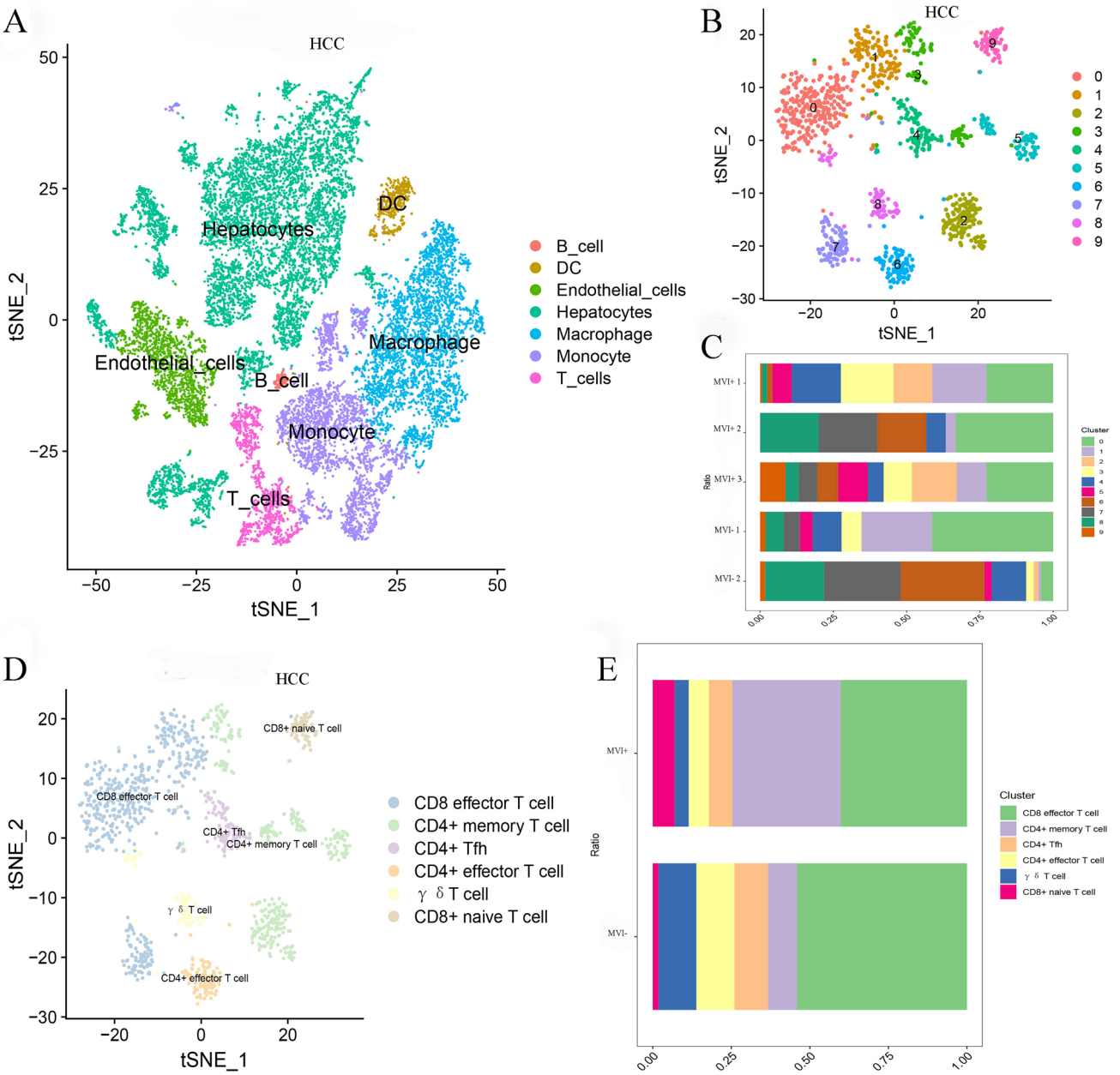


Fig. 2 Distribution of T cell types in the HCC microenvironment. **A** Presentation of each cell type in all HCC samples. **B** T cell data were further divided into 10 clusters after extraction. **C** Ten clusters showed differences in content in 3 MVI + samples and 2 MVI- samples. **D** T cell subsets are CD8 effector T cell, CD4 + memory T cell, CD4 + Tfh, CD4 + effector T cell, γδ T cell and CD8 + naive T cell. **E** Changes in the number of 6 T cell subsets existed in the MVI + samples relative to the MVI- samples

different (Fig. 6B and C). The results of cell communication analysis suggested that FOLR2 + macrophage only did not interact with B cells in MVI- samples (Fig. 6D), while only did not interact with MT1G + macrophages in MVI + samples (Fig. 6F). Moreover, its effect on TREM2 + SPP1 + macrophages was strongest in both MVI- (Fig. 6E) and MVI + samples (Fig. 6G). The results after single cell analysis suggest that DMXL2 (P = 5.50e—05) (Fig. 6 H), PTEB (P = 7.04e—03) (Fig. 6 I), PSMB5 (P = 0.03) (Fig. 6 J), DOCK2 (P = 2.31e—03) (Fig. 6 K) and NMD3 (P = 0.03) (Fig. 6 L) were negatively correlated with FOLR2 + macrophages. And NDUFS2 (P = 0.03) (Fig. 6M), ASCC3 (P = 1.76e—03) (Fig. 6N) and CHAF1 (P = 0.03) (Fig. 6O) were positively correlated with the TREM2 + SPP1 + macrophage cell content.

exposure	method	nsnp	pval	OR (95% CI)
RPS5	Wald ratio	1	0.0046957655	0.8615 (0.7769 − 0.9553)
GNAS	MR Egger	6	0.4967395566	1.1065 (0.8484 − 1.4430)
GNAS	Weighted median	6	0.2566909374	0.9246 (0.8075 − 1.0587)
GNAS	Inverse variance weighted	6	0.0450022752	0.8842 (0.7840 − 0.9973)
GNAS	Simple mode	6	0.4911195520	0.9266 (0.7577 − 1.1331)
GNAS	Weighted mode	6	0.3200488590	0.9266 (0.8092 − 1.0610)
RSRC2	Wald ratio	1	0.0403254132	1.7041 (1.0238 − 2.8364)
CCNC	Wald ratio	1	0.0497682988	0.6229 (0.3882 − 0.9995)
MMP24	Inverse variance weighted	2	0.0105377512	1.3290 (1.0687 − 1.6527)
FOXK2	Inverse variance weighted	2	0.0394059988	1.5850 (1.0226 − 2.4568)
ATG10	MR Egger	3	0.5315446802	0.9313 (0.7983 − 1.0864)
ATG10	Weighted median	3	0.0321459978	0.8909 (0.8015 − 0.9902)
ATG10	Inverse variance weighted	3	0.0189887277	0.8837 (0.7970 − 0.9799)
ATG10	Simple mode	3	0.2099004812	0.8911 (0.7872 − 1.0087)
ATG10	Weighted mode	3	0.1802862789	0.8911 (0.7969 − 0.9964)
NDUFS5	MR Egger	5	0.1186425682	1.1632 (1.0146 − 1.3335)
NDUFS5	Weighted median	5	0.0007545344	1.1512 (1.0607 − 1.2495)
NDUFS5	Inverse variance weighted	5	0.0008124960	1.1369 (1.0546 − 1.2256)
NDUFS5	Simple mode	5	0.1381376731	1.1468 (0.9918 − 1.3261)
NDUFS5	Weighted mode	5	0.0286363968	1.1647 (1.0652 − 1.2735)

Fig. 3 Results of Mendelian randomization analysis of ER-phagy REGs and total number of T cells in blood

3.4 ER-phagy REGs and monocytes

The monocytes were searched in the GWAS database to obtain the label ebi-a-GCST90018967, which corresponds to information on 349,856 samples and 19,052,644 SNPs. A total of 28 causally related genes were obtained via Mendelian randomization analysis (Fig. 7). The monocyte data were divided into seven clusters (Fig. 8A). Further annotation revealed three conventional subgroups (classical, intermediate, and nonclassical) [26–29] and CD14 (low), CD16 (low), CD11c (low), CD14 (low), CD16 (low), and CD11c (bright) special subgroups (Fig. 8B). There were some differences in the content of these cells between the MVI + samples and the MVI- samples (Fig. 8C). Monocyte subsets did not interact with B cells only in MVI- samples (Fig. 8D), and it was most closely associated with macrophages with 12 ligand-receptor sites and least associated with Hepatocytes with 1 ligand-receptor sites (Fig. 8E). Monocyte subsets interacted with all other cells identified in the MVI + sample (Fig. 8F), and it was most closely associated with intermediate monocyte with 7 ligand-receptor sites and least associated with Hepatocytes and T cell with 2 ligand-receptor sites (Fig. 8G). RPS5 (P = 2.54e–04) (Fig. 8H), NUCKS1 (P = 2.14e–11) (Fig. 8I), RPS5 (P = 2.54e–04) (Fig. 8J), EI24 (P = 9.45e–04) (Fig. 8K), DFFA (P = 0.01) (Fig. 8L), and CHAF1A (P = 0.04) (Fig. 8M) were positively correlated with classical monocyte. EPG5 was negatively correlated with classical monocyte (P = 0.03) (Fig. 8N). MRPL16 (P = 5.30e–04) (Fig. 8O), RPS5 (P = 1.01e–24) (Fig. 8Q) and TARS2 (P = 0.03) (Fig. 8R) were negatively correlated with CD14(low)-CD16(low)-CD11c(bright) monocyte. RAB5C was positively correlated with CD14(low)-CD16(low)-CD11c(bright) monocyte (P = 9.34e—04) (Fig. 8P).

3.5 ER-phagy REGs and DCs

After searching DC data, the label ebi-a-GCST90001458 was obtained, with information on a total of 3387 samples containing 15,136,664 SNPs. Five genes correlated with the absolute count of myeloid dendritic cell and four genes correlated with the absolute count of plasmacytoid dendritic cells were identified via Mendelian randomization analysis (Fig. 9A and B). The DC subgroup data were reclustered, and 9 clusters were obtained (Fig. 10A). The distribution patterns

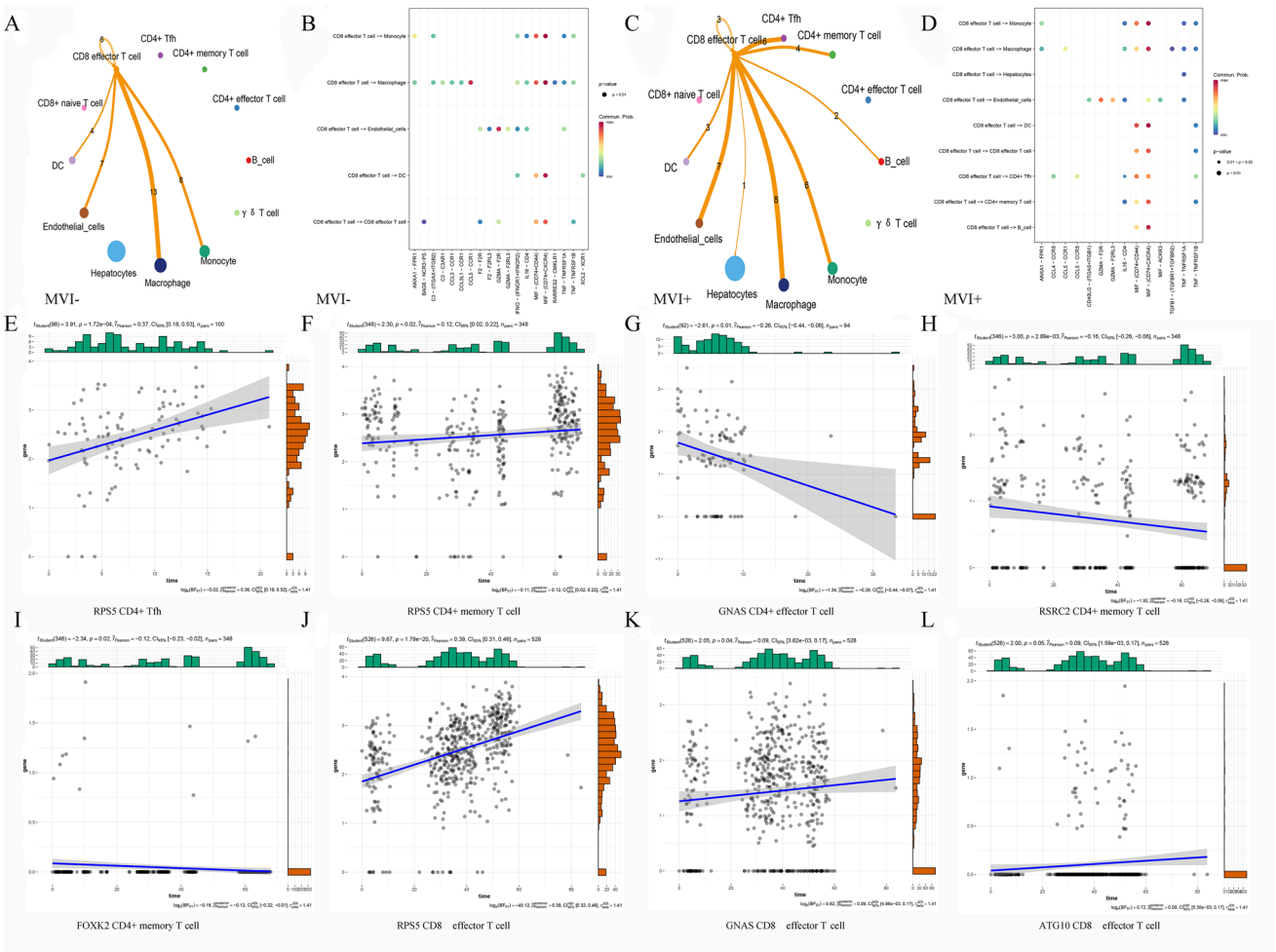


Fig. 4 Cellular communication changes and validation of causal genes in T cells. **A** Network diagram of the associated interactions between T cell subsets and other cells in MVI- samples. **B** Ligand-receptor sites information for interactions between T cell subsets and with other cells in MVI- samples. **C** Network diagram of the associated interactions between T cell subsets and other cells in MVI + samples. **D** Ligand-receptor sites information for interactions between T cell subsets and with other cells in MVI + samples. **F** RPS5 was positively correlated with CD4 + memory T cell content. **G** GNAS was negatively correlated with CD4 + effectorT cell content. **H** RSRC2 was negatively correlated with CD4 + memory T cell content. **I** FOXK2 was negatively correlated with CD4 + memory T cell content. (J) RPS5 was positively correlated with CD8 effector T cell content. **K** GNAS was positively correlated with CD8 effector T cell content. **L** ATG10 was positively correlated with CD8 effector T cell content

of these cellular clusters demonstrate significant heterogeneity among distinct HCC specimens (Fig. 10B). DC-c3-CLEC9A and DC-c4-LAMP3 were obtained after annotation on the basis of marker genes in previous literature [22, 30–32]. These cells differed in content between MVI + and MVI- patients (Fig. 10C and D). In addition, DC-c4-LAMP3 interact closely with B cells in MVI + patients, whereas DC-c4-LAMP3 interact closely with hepatocytes in MVI- patients (Fig. 10E, F, G and H). UBE2J1 was positively correlated with DC-c4-LAMP3 levels (P = 0.02) (Fig. 10I) in the tumour microenvironment. All the above genes are shown in Table 1 for correlation.

3.6 ER-phagy REGs and HCC

The GWAS search label for hepatocellular carcinoma was bbj-a-158, and information was available on 197,611 samples and 8,885,115 SNPs. Six genes with causal relationships with HCC were obtained via Mendelian randomization analysis, and the P values of the reverse analysis were all greater than 0.05 (Fig. 11A and B). Among these genes, five were identified

A

exposure	method	nsnp	pval	OR (95% CI)
FUCA2	Wald ratio	1	0.0436307386	0.6897 (0.4807 − 0.9894)
NUCKS1	Inverse variance weighted	2	0.0324043381	0.8269 (0.6948 − 0.9842)
PSMB5	MR Egger	3	0.8800388785	1.0829 (0.4775 − 2.4560)
PSMB5	Weighted median	3	0.0008455260	1.4308 (1.1593 − 1.7658)
PSMB5	Inverse variance weighted	3	0.0007318577	1.4379 (1.1646 − 1.7752)
PSMB5	Simple mode	3	0.1195774824	1.5206 (1.1121 − 2.0791)
PSMB5	Weighted mode	3	0.0991674023	1.4201 (1.1235 − 1.7950)
RBM28	Inverse variance weighted	2	0.0429819886	1.2720 (1.0076 − 1.6056)
ASCC3	MR Egger	3	0.3561243680	1.8831 (0.8661 − 4.0942)
ASCC3	Weighted median	3	0.0006216041	1.7738 (1.2774 − 2.4630)
ASCC3	Inverse variance weighted	3	0.0004102568	1.7750 (1.2911 − 2.4403)
ASCC3	Simple mode	3	0.1790559696	1.7618 (1.0206 − 3.0411)
ASCC3	Weighted mode	3	0.0925473727	1.7726 (1.2277 − 2.5594)
DOCK2	MR Egger	4	0.3580898766	0.7462 (0.4596 − 1.2116)
DOCK2	Weighted median	4	0.0495005600	0.7843 (0.6154 − 0.9995)
DOCK2	Inverse variance weighted	4	0.0471930768	0.7992 (0.6406 − 0.9972)
DOCK2	Simple mode	4	0.2906674962	0.8107 (0.5879 − 1.1180)
DOCK2	Weighted mode	4	0.1422008670	0.7832 (0.6148 − 0.9977)
ENPP3	MR Egger	4	0.5252549201	0.8506 (0.5613 − 1.2891)
ENPP3	Weighted median	4	0.0505911495	0.7791 (0.6067 − 1.0006)
ENPP3	Inverse variance weighted	4	0.0159890188	0.7445 (0.5856 − 0.9465)
ENPP3	Simple mode	4	0.2430112121	0.6741 (0.3955 − 1.1489)
ENPP3	Weighted mode	4	0.1897236172	0.8009 (0.6191 − 1.0362)
C21orf91	Wald ratio	1	0.0392409790	1.6642 (1.0254 − 2.7009)
TMEM170A	Inverse variance weighted	2	0.0476132286	0.8209 (0.6753 − 0.9979)
NMD3	Wald ratio	1	0.0417018721	1.3652 (1.0118 − 1.8422)
GDAP2	MR Egger	4	0.9317626858	0.9515 (0.3473 − 2.6064)
GDAP2	Weighted median	4	0.2848482675	1.4253 (0.7445 − 2.7290)
GDAP2	Inverse variance weighted	4	0.0463228734	1.7241 (1.0089 − 2.9464)
GDAP2	Simple mode	4	0.2931507557	1.5623 (0.7854 − 3.1076)
GDAP2	Weighted mode	4	0.3201818808	1.4510 (0.7853 − 2.6811)

B

exposure	method	nsnp	pval	OR (95% CI)
DMXL2	MR Egger	6	0.22131702	1.1473 (0.9525 − 1.3819)
DMXL2	Weighted median	6	0.02538749	1.1574 (1.0182 − 1.3156)
DMXL2	Inverse variance weighted	6	0.02313056	1.1507 (1.0194 − 1.2989)
DMXL2	Simple mode	6	0.74185736	1.0507 (0.7953 − 1.3882)
DMXL2	Weighted mode	6	0.07447296	1.1748 (1.0208 − 1.3519)
TMTC4	Inverse variance weighted	2	0.04056590	1.0862 (1.0036 − 1.1757)
ABHD8	Wald ratio	1	0.02733908	1.2589 (1.0261 − 1.5446)
PREB	Wald ratio	1	0.01849167	1.3658 (1.0537 − 1.7702)
MS4A3	MR Egger	9	0.84047826	1.0721 (0.5580 − 2.0599)
MS4A3	Weighted median	9	0.11888744	1.2070 (0.9528 − 1.5289)
MS4A3	Inverse variance weighted	9	0.04141068	1.2088 (1.0074 − 1.4505)
MS4A3	Simple mode	9	0.33226416	1.2101 (0.8424 − 1.7381)
MS4A3	Weighted mode	9	0.26937054	1.2061 (0.8851 − 1.6435)
NDUFS2	Wald ratio	1	0.02049479	0.8286 (0.7067 − 0.9714)
CHAF1A	Wald ratio	1	0.03804687	1.1350 (1.0070 − 1.2793)

Fig. 5 Results of Mendelian randomization analysis of ER-phagy REGs and macrophage in blood. **A** Macrophage receptor MARCO levels and ER-phagy REGs. **B** Macrophage mannose receptor 1 and ER-phagy REGs

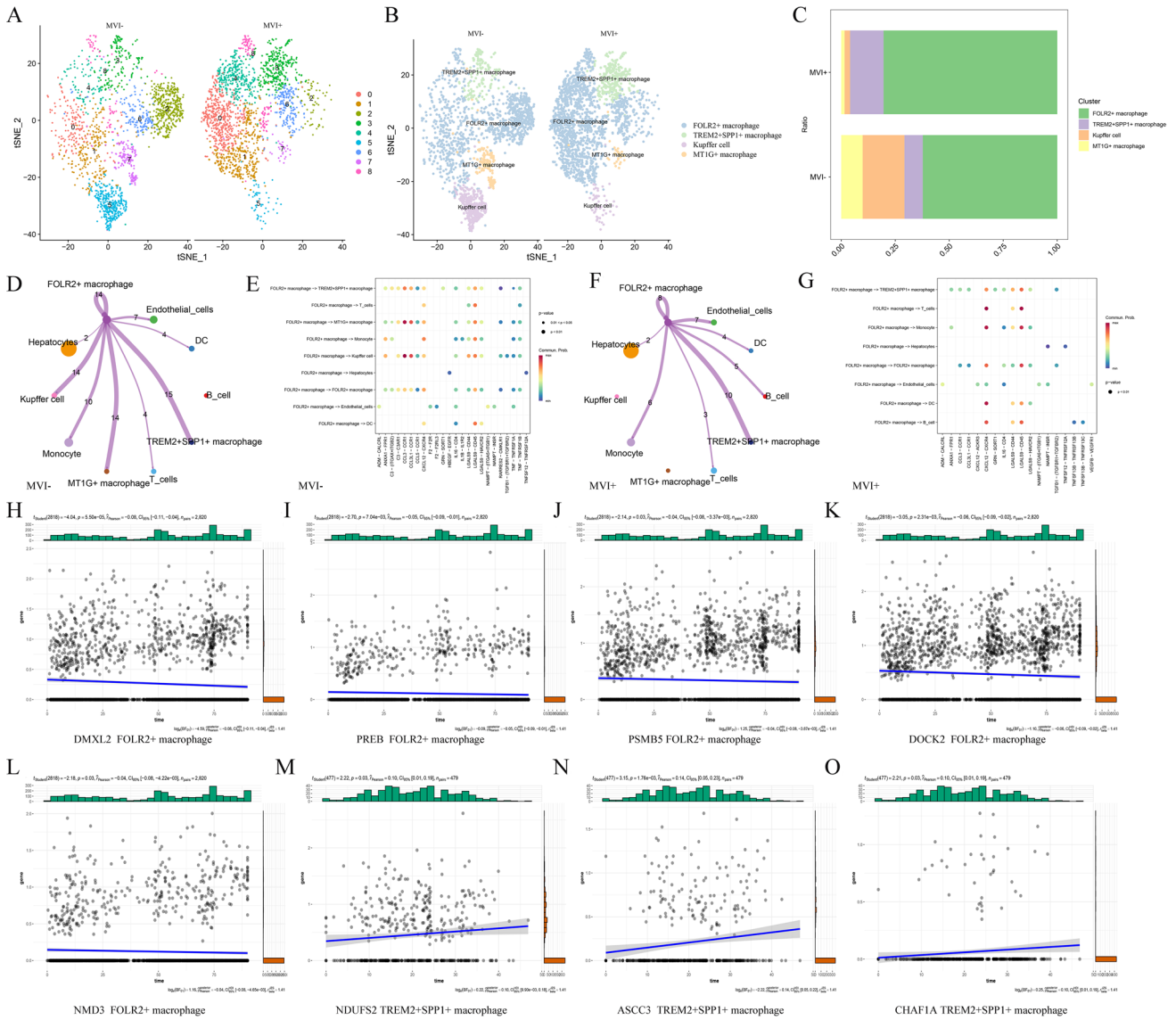


Fig. 6 Changes in cellular content and interactions, and validation of causal genes in Macrophage. **A** Nine clusters of cells identified as macrophages were extracted. **B** Annotation of macrophage subsets. **C** Comparison of the proportions of each macrophage subset in different MVI states. **D** Network diagram of the associated interactions between macrophage subsets and other cells in MVI- samples. **E** Ligand-receptor sites information for interactions between macrophage subsets and with other cells in MVI- samples. **F** Network diagram of the associated interactions between macrophage subsets and other cells in MVI + samples. **G** Ligand-receptor sites information for interactions between macrophage subsets and with other cells in MVI + samples. **H** DMXL2, **I** PTEB, (**J**) PSMB5, (**K**) DOCK2 and (**L**) NMD3 were negatively correlated with the FOLR2 + macrophage cell content. **M** NDUFS2, (**N**) ASCC3 and (**O**) CHAF1 were positively correlated with the TREM2 + SPP1 + macrophage cell content

in the GEPIA database as having survival-related correlations, among which three demonstrated statistically significant associations between their expression levels and patient prognosis (Fig. 11C). The genes delineated in Table 1 facilitate the stratification of 377 HCC patient samples obtained from the TCGA database into two distinct molecular subgroups, as demonstrated in Fig. 12A. There were significant differences in survival time (Fig. 12B), gene expression and clinicopathological features (T, N, M, stage, grade, gender, and age) between the two groups (Fig. 12C).

3.7 Accuracy and performance of the prognostic model

The gene expression and clinical data of 273 HCC samples from an Asian population were obtained from the ICGC. These data and differential gene expression data of the TCGA clustering group were used to perform univariate Cox analysis, and

exposure	method	nsnp	pval	OR (95% CI)
NUCKS1	Inverse variance weighted	2	6.665018e−08	0.9652 (0.9528 − 0.9777)
RPS5	Wald ratio	1	3.077404e−04	1.0179 (1.0081 − 1.0277)
CHERP	Wald ratio	1	8.021517e−03	0.9394 (0.8970 − 0.9838)
CAPN3	Inverse variance weighted	2	3.105714e−02	1.0557 (1.0050 − 1.1091)
POLE2	Wald ratio	1	3.170199e−02	0.9589 (0.9229 − 0.9963)
PTCD1	MR Egger	3	6.038146e−01	1.0216 (0.9637 − 1.0829)
PTCD1	Weighted median	3	1.153605e−02	1.0249 (1.0055 − 1.0446)
PTCD1	Inverse variance weighted	3	9.001678e−03	1.0249 (1.0062 − 1.0439)
PTCD1	Simple mode	3	1.543307e−01	1.0265 (1.0033 − 1.0502)
PTCD1	Weighted mode	3	1.661717e−01	1.0241 (1.0020 − 1.0466)
RAB5C	Wald ratio	1	1.379917e−11	1.1750 (1.1213 − 1.2313)
CCNC	Wald ratio	1	4.110366e−03	0.9388 (0.8992 − 0.9802)
NR2C1	Wald ratio	1	6.857140e−03	0.9762 (0.9594 − 0.9934)
AARS2	Wald ratio	1	7.110050e−07	0.9201 (0.8903 − 0.9509)
ABHD8	Wald ratio	1	3.090859e−05	1.0393 (1.0206 − 1.0582)
CHMP2A	Wald ratio	1	2.260769e−02	1.0476 (1.0065 − 1.0903)
THOC6	Inverse variance weighted	2	1.148238e−03	1.0698 (1.0272 − 1.1142)
TARS2	Wald ratio	1	2.254241e−05	0.9337 (0.9046 − 0.9638)
INTS7	Inverse variance weighted	2	1.089137e−02	0.9613 (0.9325 − 0.9910)
EI24	Inverse variance weighted	2	2.118426e−02	1.0242 (1.0036 − 1.0453)
EPG5	Inverse variance weighted	2	4.891429e−03	1.0308 (1.0093 − 1.0529)
MED21	MR Egger	4	3.895593e−01	1.0945 (0.9305 − 1.2873)
MED21	Weighted median	4	1.581195e−02	1.0214 (1.0040 − 1.0391)
MED21	Inverse variance weighted	4	6.308099e−03	1.0242 (1.0068 − 1.0419)
MED21	Simple mode	4	1.560025e−01	1.0184 (0.9993 − 1.0380)
MED21	Weighted mode	4	9.026475e−02	1.0210 (1.0043 − 1.0379)
NDUFS2	Wald ratio	1	3.295266e−04	0.9764 (0.9637 − 0.9892)
DFFA	Wald ratio	1	2.780690e−02	1.0147 (1.0016 − 1.0279)
ASB16	Wald ratio	1	4.110366e−03	0.9521 (0.9207 − 0.9846)
MRPL16	Wald ratio	1	4.829645e−03	1.0794 (1.0235 − 1.1384)
CHAF1A	Wald ratio	1	3.708548e−04	1.0204 (1.0091 − 1.0318)
GP9	MR Egger	10	6.173999e−01	1.0087 (0.9763 − 1.0421)
GP9	Weighted median	10	5.193021e−03	1.0134 (1.0040 − 1.0230)
GP9	Inverse variance weighted	10	4.129318e−02	1.0215 (1.0008 − 1.0426)
GP9	Simple mode	10	1.312017e−02	1.0533 (1.0191 − 1.0887)
GP9	Weighted mode	10	4.368641e−02	1.0112 (1.0018 − 1.0206)
LIPT2	Wald ratio	1	1.147201e−04	1.0853 (1.0411 − 1.1314)
RPS6KB2	Inverse variance weighted	2	3.987280e−03	0.9666 (0.9446 − 0.9892)
GDI1	Wald ratio	1	5.569422e−101	0.5702 (0.5415 − 0.6004)
TECPR1	MR Egger	6	7.549818e−01	0.9929 (0.9522 − 1.0353)
TECPR1	Weighted median	6	3.921377e−02	1.0094 (1.0005 − 1.0185)
TECPR1	Inverse variance weighted	6	2.942181e−02	1.0098 (1.0010 − 1.0187)
TECPR1	Simple mode	6	1.448379e−01	1.0157 (0.9979 − 1.0338)
TECPR1	Weighted mode	6	1.022408e−01	1.0093 (1.0002 − 1.0184)

Fig. 7 Results of Mendelian randomization analysis of ER-phagy REGs and total number of monocytes in blood

Fig. 8 Changes of macrophage types in the HCC microenvironment. **A** Seven clusters of cells identified as monocytes were extracted. **B** ▶ Distribution of monocyte subsets in different MVI states. **C** The proportion of the identified subpopulation cells in MVI + sample was significantly different from that in MVI- sample. **D** Interactions between monocyte subsets and other cells in MVI- samples. **E** Ligand-receptor sites information for interactions between monocyte subsets and with other cells in MVI- samples. **F** Interactions between monocyte subsets and other cells in MVI + samples. **G** Ligand-receptor sites information for interactions between monocyte subsets and with other cells in MVI + samples. **H** RPS5, (**I**) NUCKS1, (**J**) NDUFS2, (**K**) EI24, (**L**) DFFA and (**M**) CHAF1A were positively correlated with the content of the classical subgroup cells. (N) EPG5 was negatively correlated with classical monocyte. (**O**) MRPL16 and (**P**) RPS5 were positively correlated with the content of CD14 (low) CD16 (low) CD11c (bright) subgroup cells. (**Q**) RPS5 was negatively correlated with CD14 (low) -CD16 (low) -CD11c (bright) monocyte. (**R**) TARS2 was negatively correlated with CD14 (low) -CD16 (low) -CD11c (bright) monocyte

a total of 338 genes were found to be associated with prognosis (Supplementary Material 4). These genes were subjected to further screening via LASSO regression analysis (Fig. 13A and B). Multivariate Cox proportional hazards regression analysis of these genes was then performed to obtain coefficients (Table 2). Using TCGA data as the training set, the HCC patients were divided into a high-risk group and a low-risk group on the basis of the calculated risk score, and the survival time was different between the groups ($p < 0.001$) (Fig. 13C). The AUC values for 1-, 3- and 4-year survival in the training dataset were 0.794, 0.744 and 0.745, respectively (Fig. 13E). Similar results were obtained with the ICGC test data, and there was a difference in survival time between the high-risk group and the low-risk group ($p < 0.001$) (Fig. 13D). The AUC values for 1-, 3-, and 4-year survival were 0.760, 0.737 and 0.732, respectively (Fig. 13F).

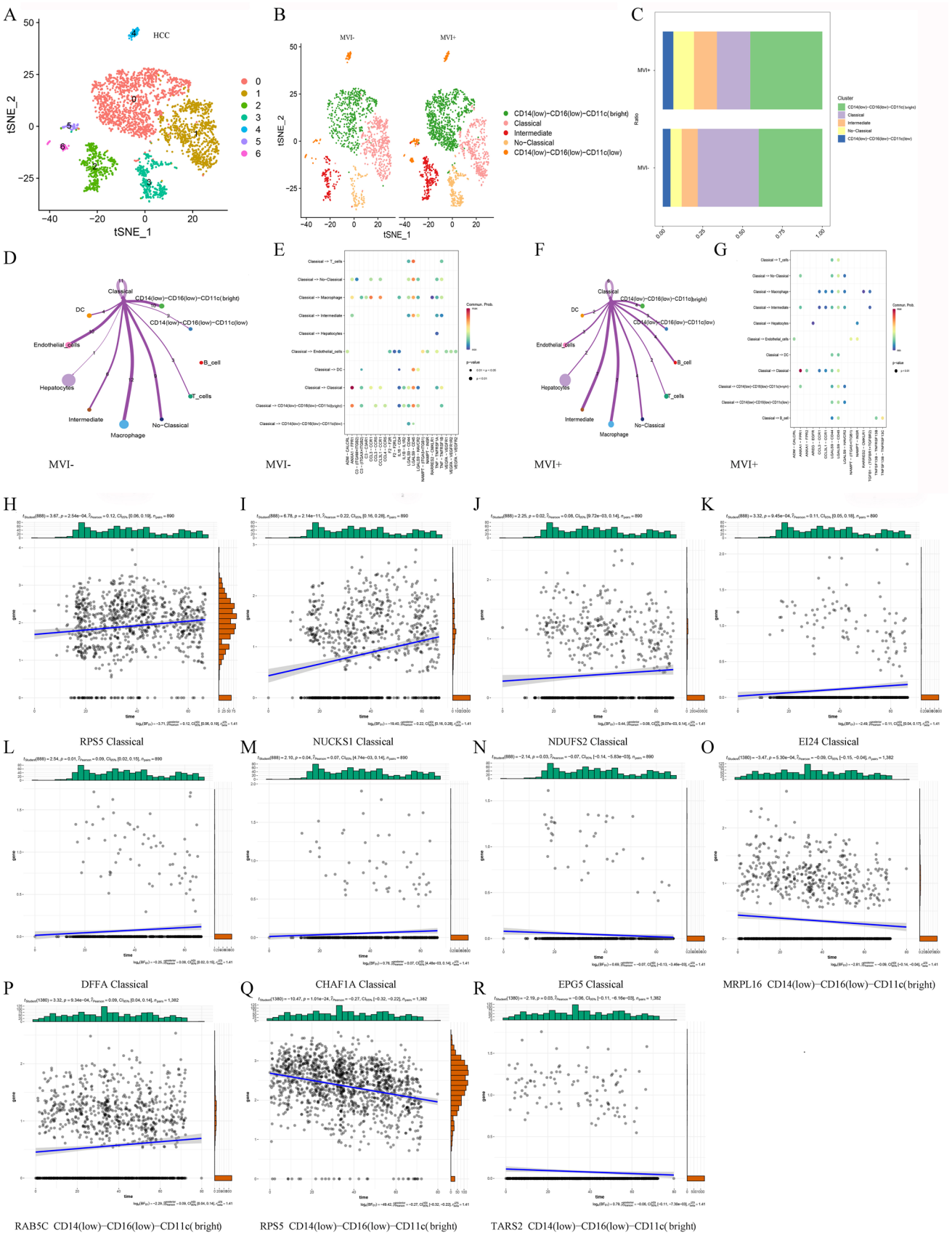
3.8 3.8 Analysis of relationships between clinicopathological features and the risk score and construction of a nomogram

Univariate Cox analysis of the relationships between the risk score and clinicopathological characteristics (including age, gender, grade, and stage) in the training set suggested that stage ($p < 0.001$, HR = 1.672, 95% CI [1.354–2.064]) and the risk score ($p < 0.001$, HR = 3.829, 95% CI [2.775–5.283]) were associated with survival (Fig. 13G); the multivariate Cox analysis results suggested that stage ($p = 0.006$, HR = 1.374, 95% CI [1.098–1.719]) and the risk score ($p < 0.001$, HR = 3.347, 95% CI [2.387–4.693]) were associated with survival (Fig. 13H). The nomogram incorporating these clinicopathological factors and the risk score was able to predict 1-, 3- and 4-year survival, as the predicted values were consistent with the actual results (Fig. 13I, J). Compared with the prognostic models in previous studies, the prediction performance was excellent [33–35] (Fig. 13K).

4 Discussion

In this study, our objectives were to elucidate the disparities in the immune cell landscape and cellular interactions between MVI + and MVI- HCC patients, investigate the impact of ER-phagy REGs on HCC, and assess the prognostic significance of these effect genes. The methods used to study the differences in the distribution of immune cells were single-cell analysis of selected data from the GEO database, automatic annotation via software packages, and manual annotation of immune cell subpopulations using marker genes; we further differentiated between MVI- and MVI + samples and performed differential analysis of cell content. The study of the influence of genes on HCC was conducted from two aspects. The first aspect pertains to the direct impact on HCC. Causal genes were identified through Mendelian randomization analysis, along with the relationship between gene expression and prognosis. The second aspect involved identifying causal genes associated with the proportion of immune cells in the blood using Mendelian randomization analysis within the GWAS database, followed by single-cell analysis to validate the effects of these genes on the content of immune cells in the HCC microenvironment. Consequently, these genes also lead to changes in the proportion of immune cells after affecting the function of ER-phagy following the emergence of MVI + tumors. Additionally, the analysis of immune cell interaction between MVI + and MVI- revealed distinct characteristics for each. Finally, a model with exceptional predictive performance was developed by integrating data from public databases.

After complete data analysis, we found that the HLTF, RPS5, SHOC2, MS4A3, TMC4, and TRAPPC12 genes have a certain effect on the occurrence of HCC. Among them, the HLTF, RPS5 and SHOC genes increase the risk of HCC, whereas the MS4A3, TMC4 and TRAPPC12 genes reduce the risk of HCC (Fig. 11). Moreover, after reverse Mendelian randomization analysis, the effects were found to occur in a single direction. Previous studies have shown that the HLTF gene can affect the proliferation and progression of liver cancer by acting on the p62/mTOR axis [36, 37]. RPS5 has also been investigated as a therapeutic target for liver cancer [38]. However, the mechanisms by which SHOC2, MS4A3, TMC4 and


A
HCC
tSNE_2
tSNE_1
0
1
2
3
4
5
6
B
MVI-
MVI+
CD14(low)−CD16(low)−CD11c(bright)
Classical
Intermediate
No−Classical
CD14(low)−CD16(low)−CD11c(low)
C
MVI+
MVI-
Ratio
Cluster
D
Classical
CD14(low)−CD16(low)−CD11c(bright)
DC
CD14(low)−CD16(low)−CD11c(low)
Endothelial cells
B_cell
Hepatocytes
T_cells
Intermediate
Macrophage
No−Classical
MVI-
E
MVI-
F
MVI+
G
MVI+
H
RAB5C CD14(low)−CD16(low)−CD11c(bright)
I
RPS5 CD14(low)−CD16(low)−CD11c(bright)
J
TARS2 CD14(low)−CD16(low)−CD11c(bright)
K
MRPL16 CD14(low)−CD16(low)−CD11c(bright)
L
RPS5 Classical
M
NUCKS1 Classical
N
NDUFS2 Classical
O
EI24 Classical
P
DFFA Classical
Q
CHAF1A Classical
R
EPG5 Classical

A

exposure	method	nsnp	pval	OR (95% CI)
CCNC	Wald ratio	1	0.009913639	0.5234 (0.3199 − 0.8561)
THOC6	Wald ratio	1	0.021872793	0.4391 (0.2173 − 0.8875)
TC2N	MR Egger	4	0.598607944	0.9575 (0.8347 − 1.0984)
TC2N	Weighted median	4	0.017001076	0.9203 (0.8596 − 0.9853)
TC2N	Inverse variance weighted	4	0.008724955	0.9175 (0.8603 − 0.9785)
TC2N	Simple mode	4	0.186607814	0.9212 (0.8383 − 1.0123)
TC2N	Weighted mode	4	0.098308708	0.9204 (0.8594 − 0.9857)
UBE2J1	MR Egger	7	0.174001470	1.2301 (0.9521 − 1.5895)
UBE2J1	Weighted median	7	0.038675899	1.1718 (1.0083 − 1.3619)
UBE2J1	Inverse variance weighted	7	0.032902534	1.1636 (1.0124 − 1.3375)
UBE2J1	Simple mode	7	0.214831514	1.1778 (0.9346 − 1.4842)
UBE2J1	Weighted mode	7	0.081377247	1.1778 (1.0104 − 1.3729)
MRPL33	Wald ratio	1	0.047135693	1.4403 (1.0046 − 2.0650)

B

exposure	method	nsnp	pval	OR (95% CI)
ARSA	MR Egger	3	0.433373939	1.0816 (0.9550 − 1.2249)
ARSA	Weighted median	3	0.015790773	1.1147 (1.0206 − 1.2175)
ARSA	Inverse variance weighted	3	0.010608449	1.1210 (1.0270 − 1.2237)
ARSA	Simple mode	3	0.350687685	1.0838 (0.9511 − 1.2352)
ARSA	Weighted mode	3	0.174011247	1.1141 (1.0059 − 1.2340)
ABHD8	Wald ratio	1	0.040561059	0.7919 (0.6334 − 0.9900)
TARS2	Wald ratio	1	0.009559066	0.6205 (0.4325 − 0.8902)
TC2N	MR Egger	4	0.533305169	0.9519 (0.8364 − 1.0834)
TC2N	Weighted median	4	0.050352749	0.9374 (0.8787 − 1.0001)
TC2N	Inverse variance weighted	4	0.037392857	0.9376 (0.8825 − 0.9963)
TC2N	Simple mode	4	0.334740461	0.9504 (0.8712 − 1.0368)
TC2N	Weighted mode	4	0.174904512	0.9380 (0.8738 − 1.0069)

Fig. 9 Results of Mendelian randomization analysis of ER-phagy REGs and total number of DCs in blood. **A** Myeloid Dendritic Cell and ER-phagy REGs. **B** Plasmacytoid Dendritic Cell and ER-phagy REGs

TRAPPC12 affect liver cancer have not been clarified. The SHOC2-encoded protein functions as a crucial scaffold protein that facilitates the interaction between downstream signal transducers within the RAS/ERK MAP kinase signaling cascade. Perturbations in the RAS/ERK MAP pathway have been demonstrated to induce circadian rhythm dysregulation and hepatocarcinogenesis [39], suggesting that alterations in SHOC2 functionality may potentially contribute to hepatocarcinogenesis through this molecular pathway. MS4A3, a member of the tetraspanin superfamily, remains understudied in the context of solid tumors. Although MS4A3 has been shown to enhance chemosensitivity in lung cancer via the THAP1/EGFR pathway [40], its potential association with hepatocellular carcinoma warrants more comprehensive investigation. TMC4 has been implicated in nonalcoholic fatty liver disease (NAFLD), which is a well-established risk factor for hepatocellular carcinoma [41–43]. TRAPPC12, initially characterized as a mediator of ER-to-Golgi vesicular trafficking [44], remains to be fully elucidated in terms of its potential oncogenic functions.

The effects of immune cells on tumours are complex [45]. In this study, the role and influence of ER-phagy in liver cancer were further investigated by analysing the correlation between ER-phagy REGs and the contents of these immune cells. An association with ER-Phagy would require an increase in cell number accompanied by a decrease in the expression of positively regulated genes and an increase in the expression of negatively regulated genes by ER-Phagy. The opposite result was observed when the number of cells was reduced. In terms of T cells, CD4 + memory and CD8 + naive cell numbers were increased in MVI + samples, whereas other T-cell types were generally decreased (Fig. 2E). These findings indicate that the levels of CD4 + memory T cells and CD8 + naive T cells exhibit a positive correlation with tumor MVI, characterized by suppressed intracellular ER-phagy and a proliferative state of immune cells. Conversely, other T cell subsets may demonstrate enhanced ER-phagy activity or undergo alternative forms of programmed cell death, resulting in their quantitative reduction. These observations potentially reflect the active participation of these immune cells in tumor progression resistance. Notably, such immune microenvironmental characteristics have been documented across various tumor types. For instance, CD4 + memory cell infiltration levels are higher in colon cancer tissues than in control normal tissues did [46]; a high enrichment score of CD4 + memory cells in breast cancer was found in more aggressive tissues [47]; and patients with a lower percentage of CD4 + memory resting T cells in gastric cancer tended to have a better prognosis [48]. As the content of CD8 + effector cells decrease, CD8 + naive cell activation and proliferation may be induced, but further studies are needed to confirm this phenomenon. In terms of cell communication, we found that

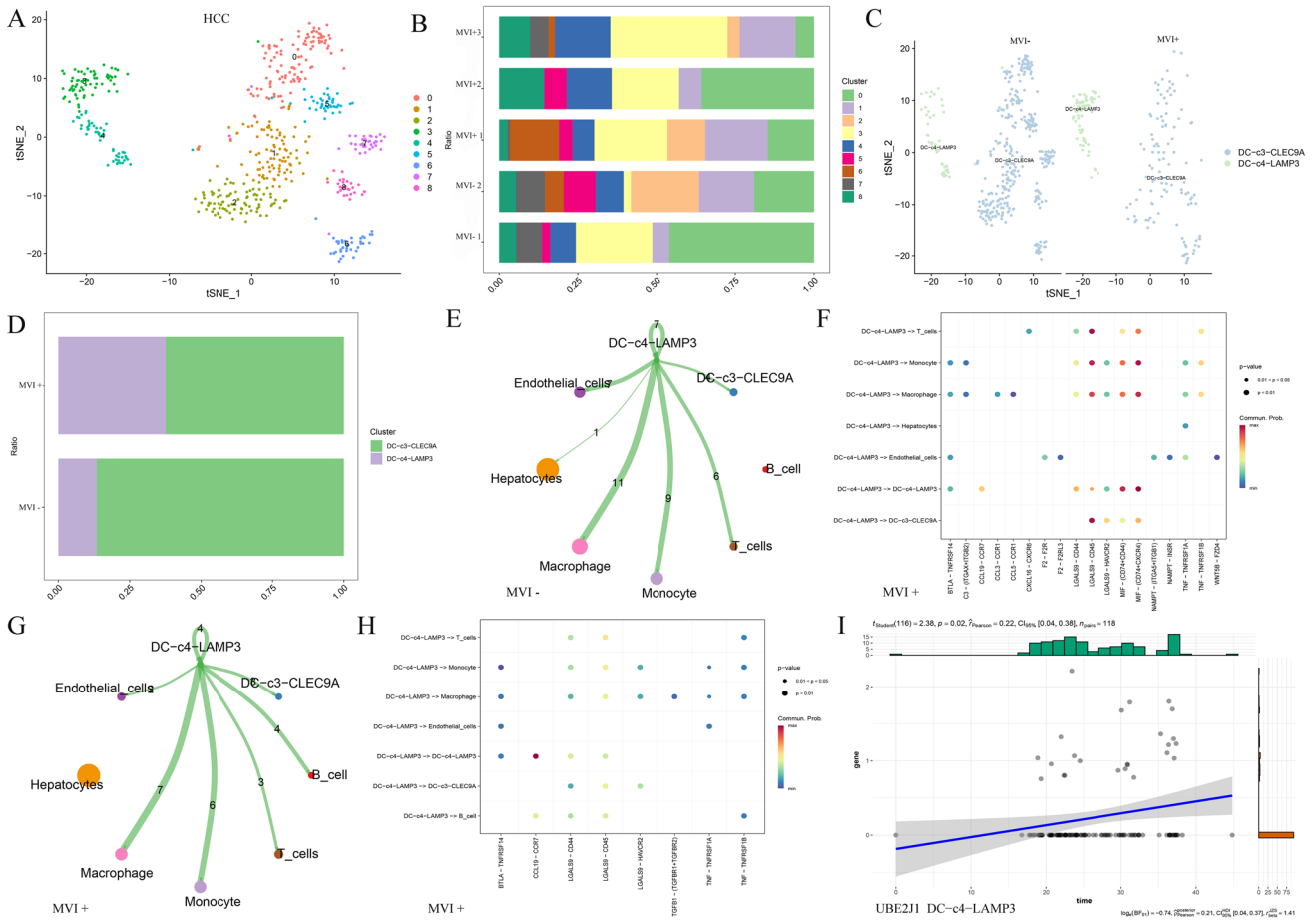


Fig. 10 Changes in cellular content and interactions, and validation of causal genes in DCs. **A** Nine clusters of cells identified as DC were extracted. **B** Distribution of clusters in different MVI samples. **C** Annotation of DC cell subsets. **D** Distribution of DC subsets in different MVI states. **E** Interactions between DC subsets and other cells in MVI- samples. **F** Ligand-receptor sites information for interactions between DC subsets and with other cells in MVI- samples. **G** Interactions between DC subsets and other cells in MVI + samples. **H** Ligand-receptor sites information for interactions between DC subsets and with other cells in MVI + samples. **I** UBE2J1 was positively correlated with DC-c4-LAMP3 cell content

CD4 + T cells, CD8 + T cells and B cells interact more closely in MVI + samples (Fig. 4C), which suggests that the body has a comprehensive function of regulating the immune system to defend against tumour cells during tumour progression. As can be seen in Table 1, RPS5 in CD4 + Tfh cells, RSRC2 and FOXK2 in CD4 + memory T cells, RPS5 in CD8 + effector T cells, GNAS in CD4 + effector T cells fit the logical association described previously. This suggests that these genes may be the main regulatory genes in the development of ER-phagy in these immune cells. Among them, RPS5 has been confirmed to be correlated with CD8 + T-cell infiltration (44). GNAS-PKA signalling is a pancancer carcinogenic pathway that inhibits the antitumour T-cell response [49]. ATG10 can be used as the core gene of CD8 + T-cell failure to predict the prognosis of liver cancer [50]. However, the relationships among RSRC2, FOXK2 and T cells have not been published thus far. For macrophages, the macrophage clusters in this study were divided into four subgroups by referring to the latest literature on subgroup classification. The MVI + samples exhibited elevated levels of FOLR2 + macrophages and TREM2 + SPP1 + macrophages. TREM2 + SPP1 + macrophages secreted SPP1, which activated cancer-associated fibroblasts (CAFs) and induced extracellular matrix (ECM) remodeling, thereby establishing a physical barrier that facilitated tumor cell invasion and promoted microvascular invasion [51]. Conversely, FOLR2 + macrophages were predominantly localized around tumor blood vessels and secreted vascular stabilizing factors (e.g., CXCL10 and ANGPT1), which inhibited abnormal vascular proliferation and leakage, thereby reducing the likelihood of tumor cell invasion through compromised blood vessels. Given their opposing effects on MVI, these two macrophage populations mutually inhibit and interact

Table 1 Potential correlation between altered cellular content and the regulation of ER-phagy

Cell	Content(MVI+)	Causal association Gene	Gene Function (Knockdown) [18]
CD4+Tfh cell	Decrease	RPS5 (P)	Enhances ER-phagy
CD4+memory T cell	Increase	RPS5(P)	/
		RSRC2(N)	Enhances ER-phagy
		FOXK2(N)	Enhances ER-phagy
CD8+effector T cell	Decrease	RPS5(P)	/
		GNAS(P)	Represses ER-phagy
		ATG10(P)	Represses ER-phagy
CD4+effector T cell	Decrease	GNAS(N)	/
FOLR2+macrophage	Increase	DOCK2(N)	Enhances ER-phagy
		NMD3(N)	Enhances ER-phagy
		PREB(N)	Represses ER-phagy
		PSMB5(N)	Represses ER-phagy
		DMXL2(N)	Enhances ER-phagy
TREM2+SPP1+macrophage	Increase	CHAF1A(P)	Represses ER-phagy
		NDUFS2(P)	Represses ER-phagy
		ASCC3(P)	
Classical	Decrease	RPS5(P)	/
		NUCKS1(P)	Enhances ER-phagy
		NDUFS2(P)	/
		EI24(P)	Represses ER-phagy
		DFFA(P)	Represses ER-phagy
		CHAF1A(P)	/
		EPG5(N)	Represses ER-phagy
CD14 (low) CD16 (low) CD11c (bright)	Increase	MRPL16(N)	Represses ER-phagy
		RAB5C(P)	Represses ER-phagy
		RPS5(N)	/
		TARS2(N)	Represses ER-phagy
DC-c4-LAMP3	Increase	UBE2J1(P)	Enhances ER-phagy

P Positive correlation with cell content; *N* Negative correlation with cell content

with each other during the MVI process, resulting in the most robust cellular interactions In terms of monocytes, CD14 (low) CD16 (low)—CD11c (bright) cells were more abundant in MVI+samples; compared with classical types, subgroup interactions occur more closely among these cells. CD14 and CD16 are important cell markers for distinguishing the different morphologies of monocytes. There are three subtypes: CD14++CDI6- (classical), CD14++CD16+(intermediate) and CD14+CD16++(nonclassical). Cells with low expression of both CD14 and CD16 (FCGR3A) were found in this study, which could be further distinguished by high and low expression of CD11c. The expression of CD11c (ITGAX) is usually seen in "intermediate", which suggests that CD14 (low) -CD16 (low) -CD11c (bright) cell might be intermediate with special morphology and CD14 (low) -CD16 (low) -CD11c (low) cells might be the other two types with special morphology. This needs to be confirmed by further experimental research. Among the two identified dendritic cell (DC) subsets, DC-C3-CLEC9a exhibits a significant reduction in MVI+cases. As a pivotal subset responsible for cross-presenting antigens, it potentially enhances anti-tumor immune responses through the activation of $CD8^+$ T cells. The diminution of this subset may attenuate anti-tumor T cell activity, thereby indirectly contributing to the development of MVI. Conversely, LAMP3+DCs, which potentially impair anti-tumor immune responses by mediating immunosuppressive signals (e.g., PD-L1/IL-10) to establish a tumor-promoting microenvironment [22], demonstrate an increased presence in MVI+cases, consistent with established theoretical frameworks. Furthermore, ER-phagy regulatory genes, which influence these cellular alterations, warrant further investigation as potential therapeutic targets. In the future, flow cytometry, qRT-PCR and other methods will be carried out to verify these genes more comprehensively.

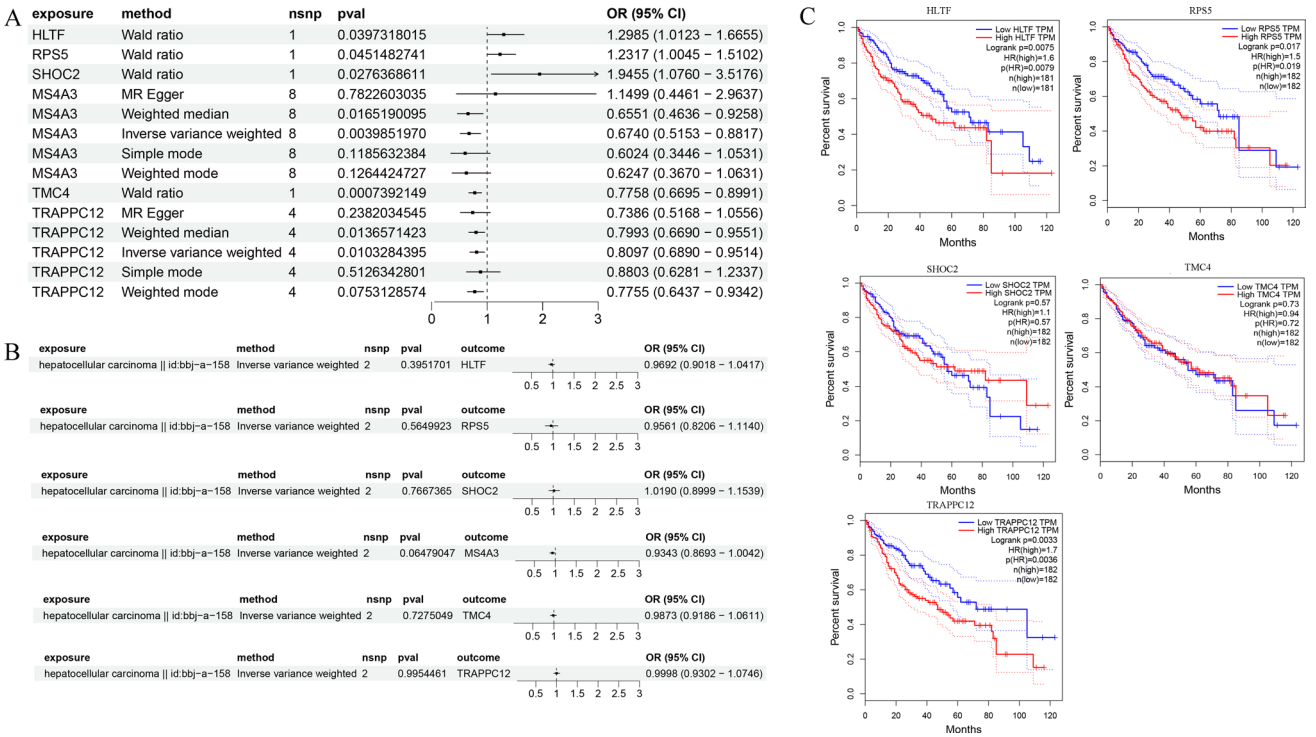


Fig. 11 Results of Mendelian randomization analysis of ER-phagy REGs and HCC. **A** Conventional Mendelian randomization analysis. **B** Reverse Mendelian randomization analysis. **C** Differences in survival time between high and low gene expression groups

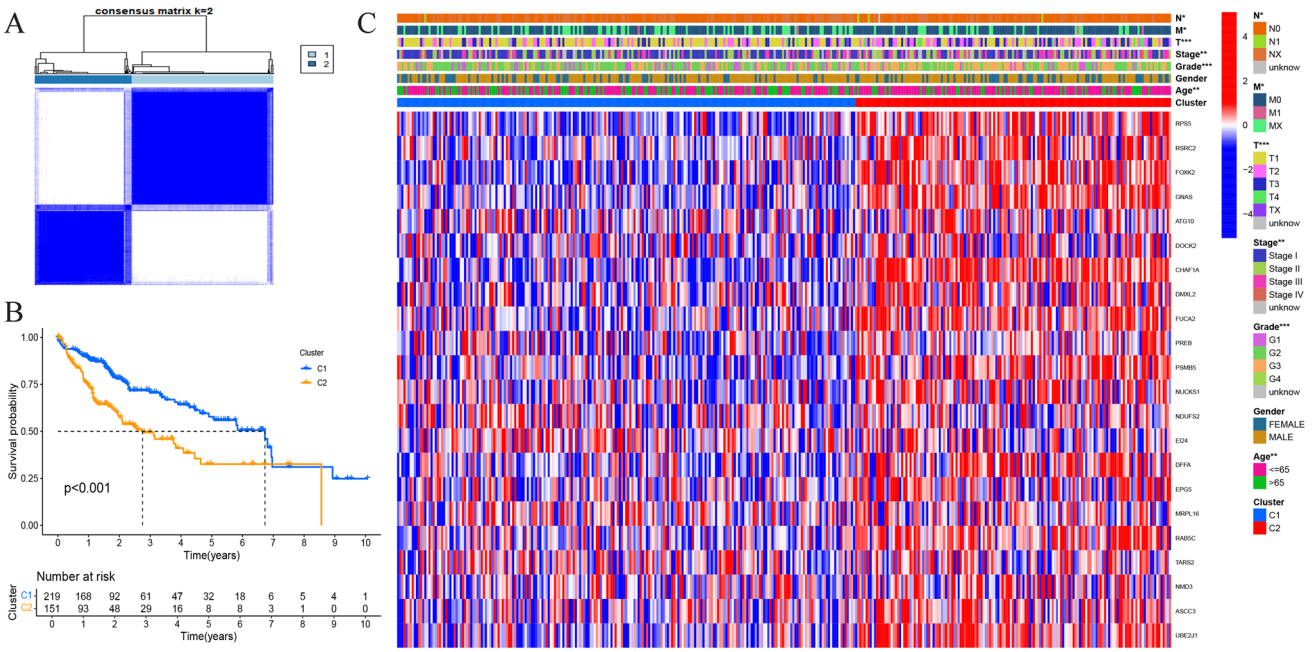


Fig. 12 In-depth analysis of genes combined with database. **A** Two groups were obtained by gene cluster analysis of HCC data in TCGA database. **B** Two groups were obtained by gene cluster analysis of HCC data in TCGA database ($P < 0.001$). **C** The clinicopathological features and expression levels of 30 genes were different between the two groups

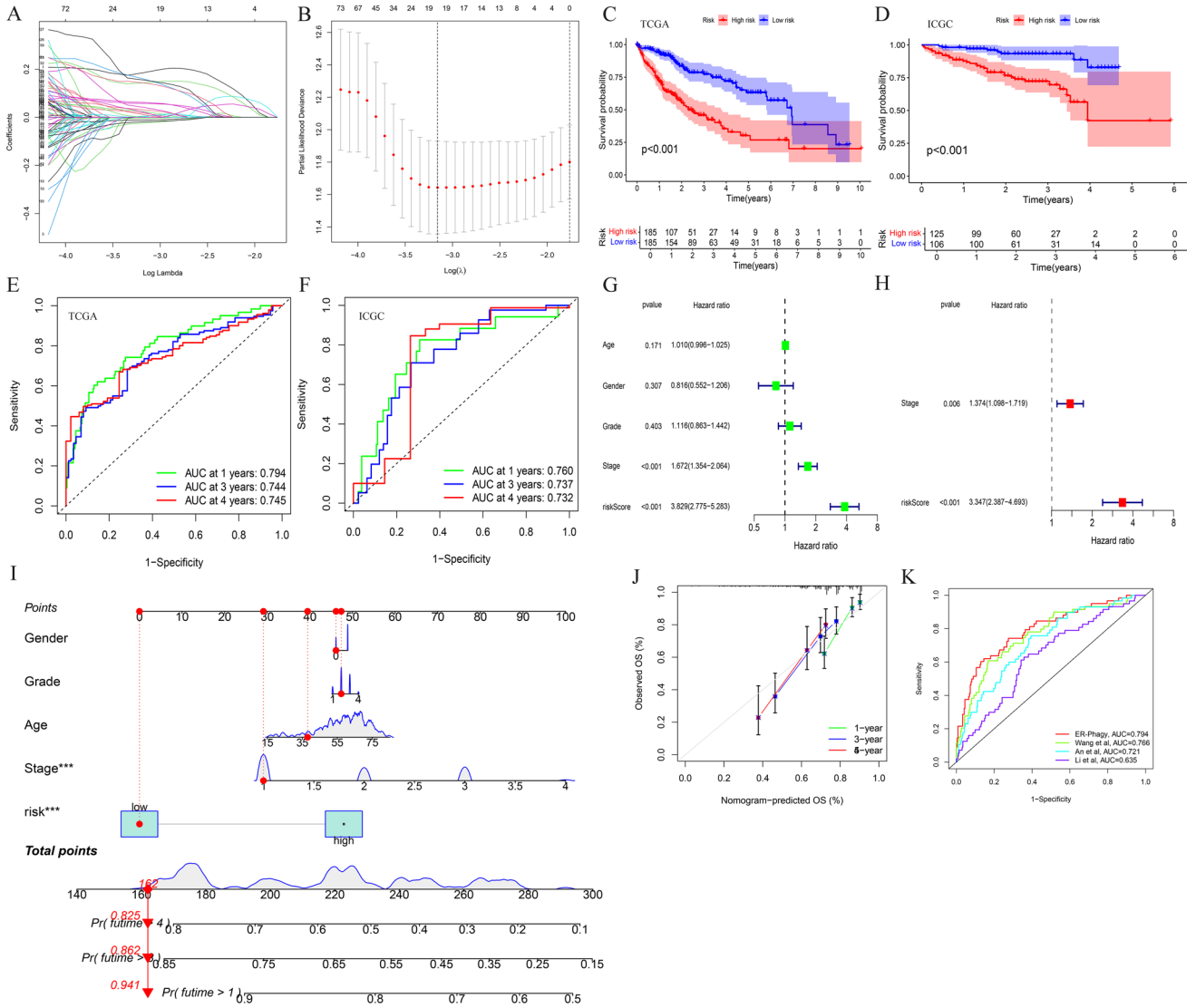


Fig. 13 Parameters and performance evaluation of prognostic models. **A** Cluster map of LASSO analysis of differentially expressed genes between groups. **B** Screening plot of prognostic factors from LASSO analysis. **C** Differences in survival time between risk score groups in the prognostic model in the TCGA database ($P < 0.001$). **D** Differences in survival time between risk score groups in the prognostic model in the ICGC database ($P < 0.001$). **E** Performance evaluation of prognostic model in TCGA database. **F** Performance evaluation of prognostic model in ICGC database. **G** Univariate COX analysis results of risk score, Stage, grade, gender, and age were obtained in TCGA database. **H** Results of multivariate COX analysis of risk score and Stage, in TCGA database. **I** Nomogram of risk score, Stage, grade, gender and age combinations. **J** Performance prediction of the nomogram. (K) Comparison of prediction performance with other models

5 Conclusion

We integrated various analytical methods, from causal relationship analysis to in-depth analysis of the microenvironment, and ultimately found that ER-phagy-related regulatory genes play important roles in liver cancer. We also distinguished the impact of these genes on liver cancer and immune cells. We subsequently used these impact genes to construct a model that can predict patient outcomes and the effectiveness of drug therapy. However, this study also has certain limitations; for example, the study did not involve experimental studies to elucidate the specific mechanisms of these genes. To address this limitation, we will integrate gene characteristics and resources for further in-depth research.

Table 2 Multivariate Cox proportional hazards regression analysis

Gene	Coefficient value
CBX2	0.18707
ANXA10	−0.02322
S100A9	0.05788
MSC	0.02299
ADH4	−0.01095
UCK2	0.14359
MYCN	0.15818
PFN2	0.00554
MAGEA6	0.01663
NEIL3	0.00315
CCDC34	0.04047
HMMR	0.15439
POF1B	0.01033
SLC1A7	0.00105
PAGE1	0.06156
KRT17	0.03102
CFHR3	−0.00404
HAVCR1	0.14221
TKTL1	0.01998
CYP2C9	−0.01865

Acknowledgements Thanks to AJE and Charlesworth Group for the language polishing and all participants for their contributions.

Author Contributions RCZ was responsible for designing the review protocol, writing the protocol and report, conducting the search, screening potentially eligible studies, extracting and analysing data, interpreting results, updating reference lists and creating'Summary of findings' tables. DD was responsible for designing the review protocol and screening potentially eligible studies. MHB contributed to data extraction and provided feedback on the report. YD, RJD, JL, YL and CRZ provided feedback on the report.

Funding This research received no external funding.

Data availability The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Declarations

Competing interests The authors declare no competing interests.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018 GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424. https://doi.org/10.3322/caac.21.
2. Zeng MS, Ye HY, Guo L, Peng WJ, Lu JP, Teng GJ, et al. Gd-EOB-DTPA-enhanced magnetic resonance imaging for focal liver lesions in Chinese patients: a multicenter, open-label, phase III study. Hepatobiliary Pancreat Dis Int. 2013;12(6):607–16. https://doi.org/10.1016/s1499-3872(13)60096-x.
3. Yu SJ. A concise review of updated guidelines regarding the management of hepatocellular carcinoma around the world: 2010–2016. Clin Mol Hepatol. 2016;22(1):7–17. https://doi.org/10.3350/cmh.2016.22.1.7.

4. Erstad DJ, Tanabe KK. Prognostic and therapeutic implications of microvascular invasion in hepatocellular carcinoma. Ann Surg Oncol. 2019;26(5):1474–93. https://doi.org/10.1245/s10434-019-07227-9.
5. Shindoh J, Kobayashi Y, Kawamura Y, Akuta N, Kobayashi M, Suzuki Y, et al. Microvascular invasion and a size cutoff value of 2 cm predict long-term oncological outcome in multiple hepatocellular carcinoma: reappraisal of the american joint committee on cancer staging system and validation using the surveillance, epidemiology, and end-results database. Liver Cancer. 2020;9(2):156–66. https://doi.org/10.1159/000504193.
6. Zhang T, Zhao F, Hu Y, Wei J, Cui F, Lin Y, et al. Environmental monobutyl phthalate exposure promotes liver cancer via reprogrammed cholesterol metabolism and activation of the IRE1α-XBP1s pathway. Oncogene. 2024;43(30):2355–70. https://doi.org/10.1038/s41388-024-03086-1.
7. Chen X, Cubillos-Ruiz JR. Endoplasmic reticulum stress signals in the tumour and its microenvironment. Nat Rev Cancer. 2021;21(2):71–88. https://doi.org/10.1038/s41568-020-00312-2.
8. Qi Z, Chen L. Endoplasmic reticulum stress and autophagy. Adv Exp Med Biol. 2019;1206:167–77. https://doi.org/10.1007/978-981-15-0602-4_8.
9. Cubillos-Ruiz JR, Bettigole SE, Glimcher LH. Tumorigenic and immunosuppressive effects of endoplasmic reticulum stress in cancer. Cell. 2017;168(4):692–706. https://doi.org/10.1016/j.cell.2016.12.004.
10. Mohamed E, Cao Y, Rodriguez PC. Endoplasmic reticulum stress regulates tumor growth and anti-tumor immunity: a promising opportunity for cancer immunotherapy. Cancer Immunol Immunother. 2017;66(8):1069–78. https://doi.org/10.1007/s00262-017-2019-6.
11. Salvagno C, Mandula JK, Rodriguez PC, Cubillos-Ruiz JR. Decoding endoplasmic reticulum stress signals in cancer cells and antitumor immunity. Trends Cancer. 2022;8(11):930–43. https://doi.org/10.1016/j.trecan.2022.06.006.
12. Liang JR, Lingeman E, Luong T, Ahmed S, Muhar M, Nguyen T, et al. A genome-wide ER-phagy screen highlights key roles of mitochondrial metabolism and ER-resident UFMylation. Cell. 2020;180(6):1160-77.e20. https://doi.org/10.1016/j.cell.2020.02.017.
13. Liu Z, Ma C, Wang Q, Yang H, Lu Z, Bi T, et al. Targeting FAM134B-mediated reticulophagy activates sorafenib-induced ferroptosis in hepatocellular carcinoma. Biochem Biophys Res Commun. 2022;589:247–53. https://doi.org/10.1016/j.bbrc.2021.12.019.
14. He L, Li H, Li C, Liu ZK, Lu M, Zhang RY, Wu D, Wei D, Shao J, Liu M, Wei HL, Zhang C, Wang Z, Kong L, et al. HMMR alleviates endoplasmic reticulum stress by promoting autophagolysosomal activity during endoplasmic reticulum stress-driven hepatocellular carcinoma progression. Cancer Commun (Lond). 2023;43(9):981–1002. https://doi.org/10.1002/cac2.12464.
15. Qian H, Chao X, Williams J, Fulte S, Li T, Yang L, et al. Autophagy in liver diseases: a review. Mol Aspects Med. 2021;82: 100973. https://doi.org/10.1016/j.mam.2021.100973.
16. Li L, Zhang X. Zhao H. eQTL. Methods Mol Biol. 2012;871:265–79. https://doi.org/10.1007/978-1-61779-785-9_14.
17. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods. 2009;6(5):377–82. https://doi.org/10.1038/nmeth.1315.
18. Gawad C, Koh W, Quake SR. Single-cell genome sequencing: current state of the science. Nat Rev Genet. 2016;17(3):175–88. https://doi.org/10.1038/nrg.2015.16.
19. Li K, Zhang R, Wen F, Zhao Y, Meng F, Li Q, et al. Single-cell dissection of the multicellular ecosystem and molecular features underlying microvascular invasion in HCC. Hepatology. 2024;79(6):1293–309. https://doi.org/10.1097/HEP.0000000000000673.
20. Baum, C.F., Schaffer, M.E. and Stillman, S. Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing. 465–506 (The Stata Journal, 2007)
21. Dong, C. & Martinez, G. J. T cells: the usual subsets. (Nature Reviews Immunology, 2010).
22. Zhang Q, He Y, Luo N, Patel SJ, Han Y, Gao R, et al. Landscape and dynamics of single immune cells in hepatocellular carcinoma. Cell. 2019;179(4):829-845.e20. https://doi.org/10.1016/j.cell.2019.10.003.
23. Sun Y, Wu L, Zhong Y, Zhou K, Hou Y, Wang Z, et al. Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. Cell. 2021;184(2):404-421.e16. https://doi.org/10.1016/j.cell.2020.11.041.
24. Nalio Ramos R, Missolo-Koussou Y, Gerber-Ferder Y, Bromley CP, Bugatti M, Núñez NG, et al. Tissue-resident FOLR2+ macrophages associate with CD8+ T cell infiltration in human breast cancer. Cell. 2022;185(7):1189-1207.e25. https://doi.org/10.1016/j.cell.2022.02.021.
25. Bill R, Wirapati P, Messemaker M, Roh W, Zitti B, Duval F, et al. CXCL9:SPP1 macrophage polarity identifies a network of cellular programs that control human cancers. Science. 2023;381(6657):515–24. https://doi.org/10.1126/science.ade2292.
26. Auffray C, Sieweke MH, Geissmann F. Blood monocytes: development, heterogeneity, and relationship with dendritic cells. Annu Rev Immunol. 2009;27:669–92. https://doi.org/10.1146/annurev.immunol.021908.132557.
27. Guilliams M, Ginhoux F, Jakubzick C, Naik SH, Onai N, Schraml BU, et al. Dendritic cells, monocytes and macrophages: a unified nomenclature based on ontogeny. Nat Rev Immunol. 2014;14(8):571–8. https://doi.org/10.1038/nri3712.
28. Yang J, Zhang L, Yu C, Yang XF, Wang H. Monocyte and macrophage differentiation: circulation inflammatory monocyte as biomarker for inflammatory diseases. Biomark Res. 2014;2(1):1. https://doi.org/10.1186/2050-7771-2-1.
29. Jakubzick CV, Randolph GJ, Henson PM. Monocyte differentiation and antigen-presenting functions. Nat Rev Immunol. 2017;17(6):349–62. https://doi.org/10.1038/nri.2017.28.
30. Zhang L, Li Z, Skrzypczynska KM, Fang Q, Zhang W, O'Brien SA, et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. Cell. 2020;181(2):442-459.e29. https://doi.org/10.1016/j.cell.2020.03.048.
31. Heras-Murillo I, Adán-Barrientos I, Galán M, Wculek SK, Sancho D. Dendritic cells as orchestrators of anticancer immunity and immunotherapy. Nat Rev Clin Oncol. 2024;21(4):257–77. https://doi.org/10.1038/s41571-024-00859-1.
32. Zhang Y, Chen H, Mo H, Hu X, Gao R, Zhao Y, et al. Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. Cancer Cell. 2021;39(12):1578-1593.e8. https://doi.org/10.1016/j.ccell.2021.09.010.
33. Wang G, Li J, Zhu L, Zhou Z, Ma Z, Zhang H, et al. Identification of hepatocellular carcinoma-related subtypes and development of a prognostic model: a study based on ferritinophagy-related genes. Discov Oncol. 2023;14(1):147. https://doi.org/10.1007/s12672-023-00756-6.
34. An LN, Du L, Wang LL, Chen J, Wang XR, Duan JP. Comprehensive analysis of an autophagy-related prognostic model for predicting survival based on TCGA and ICGC database in hepatocellular carcinoma patients. J Gastrointest Oncol. 2022;13(6):3154–68. https://doi.org/10.21037/jgo-22-1130.

35. Wang W, Wang L, Xie X, Yan Y, Li Y, Lu Q. A gene-based risk score model for predicting recurrence-free survival in patients with hepatocellular carcinoma. BMC Cancer. 2021;21(1):6. https://doi.org/10.1186/s12885-020-07692-6.
36. Tan Y, Wu D, Liu ZY, Yu HQ, Zheng XR, Lin XT, et al. Degradation of helicase-like transcription factor (HLTF) by β-TrCP promotes hepatocarcinogenesis via activation of the p62/mTOR axis. J Mol Cell Biol. 2023. https://doi.org/10.1093/jmcb/mjad012.
37. Cen X, Lu Y, Lu J, Zhan P, Cheng Y, Luo C, et al. Upregulation of helicase-like transcription factor predicts poor prognosis and facilitates hepatocellular carcinoma progression. Hum Cell. 2023;36(4):1477–84. https://doi.org/10.1007/s13577-023-00917-3.
38. Qiu L, Chao W, Zhong S, Ren AJ. Eukaryotic ribosomal protein S5 of the 40s subunit: structure and function. Int J Mol Sci. 2023. https://doi.org/10.3390/ijms24043386.
39. Hanley KL, Liang Y, Wang G, Lin X, Yang M, Karin M, et al. Concurrent disruption of the Ras/MAPK and NF-κB pathways induces circadian deregulation and hepatocarcinogenesis. Mol Cancer Res. 2022;20(3):337–49. https://doi.org/10.1158/1541-7786.MCR-21-0479.
40. Duan Z. MS4A3 promotes the chemosensitivity of lung cancer via THAP1/EGFR pathways. Crit Rev Eukaryot Gene Expr. 2024;34(8):1–11. https://doi.org/10.1615/CritRevEukaryotGeneExpr.2024053662.
41. Marengo A, Rosso C, Bugianesi E. Liver cancer: connections with obesity, fatty liver, and cirrhosis. Annu Rev Med. 2016;67:103–17. https://doi.org/10.1146/annurev-med-090514-013832.
42. Thomas JA, Kendall BJ, El-Serag HB, Thrift AP, Macdonald GA. Hepatocellular and extrahepatic cancer risk in people with non-alcoholic fatty liver disease. Lancet Gastroenterol Hepatol. 2024;9(2):159–69. https://doi.org/10.1016/S2468-1253(23)00275-3.
43. Polyzos SA, Chrysavgis L, Vachliotis ID, Chartampilas E, Cholongitas E. Nonalcoholic fatty liver disease and hepatocellular carcinoma: Insights in epidemiology, pathogenesis, imaging, prevention and therapy. Semin Cancer Biol. 2023;93:20–35. https://doi.org/10.1016/j.semcancer.2023.04.010.
44. Scrivens PJ, Noueihed B, Shahrzad N, Hul S, Brunet S, Sacher M. C4orf41 and TTC-15 are mammalian TRAPP components with a role at an early stage in ER-to-Golgi trafficking. Mol Biol Cell. 2011;22(12):2083–93. https://doi.org/10.1091/mbc.E10-11-0873.
45. De Visser KE, Joyce JA. The evolving tumor microenvironment: from cancer initiation to metastatic outgrowth. Cancer Cell. 2023;41(3):374–403. https://doi.org/10.1016/j.ccell.2023.02.016.
46. Ge P, Wang W, Li L, Zhang G, Gao Z, Tang Z, et al. Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of colorectal cancer. Biomed Pharmacother. 2019;118: 109228. https://doi.org/10.1016/j.biopha.2019.109228.
47. Deng L, Lu D, Bai Y, Wang Y, Bu H, Zheng H. Immune profiles of tumor microenvironment and clinical prognosis among women with triple-negative breast cancer. Cancer Epidemiol Biomarkers Prev. 2019;28(12):1977–85. https://doi.org/10.1158/1055-9965.Epi-19-0469.
48. Ning ZK, Hu CG, Huang C, Liu J, Zhou TC, Zong Z. Molecular subtypes and CD4(+) memory T cell-based signature associated with clinical outcomes in gastric cancer. Front Oncol. 2020;10: 626912. https://doi.org/10.3389/fonc.2020.626912.
49. Cui Y, Miao Y, Cao L, Guo L, Cui Y, Yan C, et al. Activation of melanocortin-1 receptor signaling in melanoma cells impairs T cell infiltration to dampen antitumor immunity. Nat Commun. 2023;14(1):5740. https://doi.org/10.1038/s41467-023-41101-3.
50. Shi J, Li G, Liu L, Yuan X, Wang Y, Gong M, et al. Establishment and validation of exhausted CD8+ T cell feature as a prognostic model of HCC. Front Immunol. 2023;14:1166052. https://doi.org/10.3389/fimmu.2023.1166052.
51. Ma RY, Black A, Qian BZ. Macrophage diversity in cancer revisited in the era of single-cell omics. Trends Immunol. 2022;43(7):546–63. https://doi.org/10.1016/j.it.2022.04.008.