

## RESEARCH ARTICLE

# Association between Ribosomal Protein Gene Dysregulation and Tumor Biodiversity of Hepatocellular Carcinoma

Zhimin Lu<sup>1</sup> Sicong Xu<sup>2</sup> Guofeng Zhao<sup>3</sup> Ziyi Niu<sup>4</sup> Guoxin Hou<sup>5\*</sup><sup>1</sup> Department of Outpatient, Affiliated Hospital of Jiaying University, The First Hospital of Jiaying, Jiaying, Zhejiang, China<sup>2</sup> Geneplus-Beijing Institute, Peking University Medical Industrial, Beijing, China<sup>3</sup> Beijing Moyin Biotechnology, Beijing, China<sup>4</sup> The First School of Clinical Medicine, Southern Medical University, Guangzhou, China<sup>5</sup> Department of Oncology, Affiliated Hospital of Jiaying University, The First Hospital of Jiaying, Jiaying, Zhejiang, China

**Correspondence to:** Guoxin Hou, Department of Oncology, Affiliated Hospital of Jiaying University, The First Hospital of Jiaying, Jiaying, Zhejiang, China; Email: [gxhou91@163.com](mailto:gxhou91@163.com)

**Received:** November 15, 2024;**Revised:** March 28, 2025;**Accepted:** April 10, 2025;**Published:** April 15, 2025.

**Citation:** Lu Z, Xu S, Zhao G, et al. Association between Ribosomal Protein Gene Dysregulation and Tumor Biodiversity of Hepatocellular Carcinoma. *Curr Cancer Rep*, 2025, 7(1): 254-268. <https://doi.org/10.25082/CCR.2025.01.001>

**Copyright:** © 2025 Zhimin Lu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/), which permits all noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited.



**Abstract: Background:** Tumor cells are characterized by a higher production of ribosomes, which are necessary for maintaining enhanced cell growth and subsequent cell division. **Aim:** The study aimed to develop a prognostic RPL score for hepatocellular carcinoma (HCC) and explore its association with immune evasion mechanisms mediated by tumor microenvironment alterations. **Methods:** Using single-sample gene set enrichment analysis (ssGSEA), an RPL score was constructed to estimate the dysregulation of ribosomal protein large (RPL) genes. The expression of RPL genes and their association with clinical outcomes and the tumor microenvironment (TME) were systematically investigated using bulk-seq and single-cell RNA-seq (scRNA-seq). **Results:** High expression levels of RPL in HCC were associated with poorer overall survival (OS) ( $P < 0.001$ ). The RPL score evaluated the RPL gene and verified its independent prognostic value for both OS and relapse-free survival ( $P = 0.0074$  and  $P < 0.001$ , respectively). TME analysis indicated that RPL gene dysregulation was closely associated with T cell exhaustion, myeloid-derived suppressor cell (MDSC) infiltration, and vascular dysplasia may be promoted by arginine deficiency ( $P = 7.6 \times 10^{-10}$ ). The scRNA-seq data suggested that the RPL score was positively and significantly associated with the tumor biodiversity score (ITH score). **Conclusion:** The study highlights the prognostic value of the RPL score and its potential role in mediating immune evasion of HCC, which may provide an impetus for the development of new targets for the treatment of HCC.

**Keywords:** Hepatocellular carcinoma, Ribosome protein large (RPL) genes, tumor microenvironment (TME)

## 1 Introduction

The ribosome is an important organelle for intracellular protein synthesis that influences the normal functioning of cells. Studies have shown that ribosomal proteins are not only closely related to the malignant biological behavior of tumors [1] but also significantly influence the process of endogenous antigen extraction [2]. Ribosomal biosynthesis is also related to the epithelial-mesenchymal transition (EMT) of tumor cells [3]. The regulation of ribosome biogenesis is a crucial area of study in hepatocellular carcinoma (HCC), the third leading cause of cancer deaths globally. Prior research has highlighted the significance of ribosome production factors, particularly RPF2 from the BRIX family, in tumor biology [4].

The function of ribosomal proteins is also significantly related to the metastasis of circulating tumor cells [5]. Genes coding for ribosomal proteins and translation regulators are abundant in metastatic tumor cells [6]. Overexpression of RPL15, which encodes a component of the large ribosomal subunit, increases metabolic growth in multiple tissues and selectively enhances the translation of other ribosomal proteins and cell cycle regulators. This underscores the critical role of ribosomal function in regulating tumor proliferation, metastasis, and the tumor microenvironment (TME). Despite this, comprehensive landscape analyses integrating ribosomal protein function with the TME, metabolism, and genomic context of hepatocellular carcinoma

(HCC) remain underexplored. Recent evidence also supports the notion that targeting ribosome biogenesis in tumor cells represents a promising strategy for the development of specific powerful anticancer drugs [1–3,5]. However, the association between ribosome inhibition and the evolution of TME is still unclear.

In this context, the study leverages the TME deconvolution algorithm, machine learning techniques, and multi-omics data to dissect the interaction between the RPL pathway and the TME in HCC. Herein, with the application of the TME deconvolution algorithm, machine learning algorithm, and multi-omics data of HCC, the interaction between the RPL pathway and the TME of HCC was investigated. The findings show that ribosome pathway dysregulation is closely associated with the TME of HCC and metabolism. Altogether, these findings indicate that targeting the ribosomal pathway may induce normalization of the TME of HCC and facilitate immune cell function through metabolic reprogramming.

## 2 Methods

### 2.1 Data source and preprocessing

Transcriptomic and matched clinical data from patients in TCGA were downloaded from the TCGA data portal (<https://portal.gdc.cancer.gov/>) in April 2020. The RNA-seq count data were transformed into TPM [7] to calculate the RPLscore and other gene signature scores using ssGSEA methodology. Updated clinical and pathological information for TCGA samples were obtained from the Genomic Data Commons (GDC) using the TCGAbiolinks R package [8]. Transcriptomic data and associated clinical data from GSE14520 were obtained from the GEO dataset website using the GEOquery R package. Genomic data were analyzed using R (version 4.1.0) and the R Bioconductor package.

### 2.2 Consensus clustering for TME-infiltrating cells

Tumors with qualitatively differential expression patterns of RPL were grouped by hierarchical agglomerative clustering (based on Euclidean distance and Ward's linkage). Unsupervised clustering methods (K-means) [9] for dataset analysis were used to identify RPL gene patterns and classify the patients for further analysis. A consensus clustering algorithm was applied to determine the number of clusters in The Cancer Genome Atlas liver hepatocellular carcinoma (TCGA-LIHC) cohort to assess the stability of the clusters. This procedure was performed using the ConsensusClusterPlus R package [10] and was repeated 1,000 times to ensure classification stability.

### 2.3 Dimension reduction and generation of RPLscore

An unsupervised clustering method (K-means) for the analysis of RPL genes was used to classify patients into two groups for further analysis (Figure 1B). Subsequently, the random forest classification algorithm was used to perform dimension reduction to reduce noise or redundant genes [11]. Next, 62 RPL genes (Table S2) were extracted and used to calculate the RPL score using the ssGSEA method [12].

### 2.4 Inference of Immune Cell fraction and Signature Score

Several computational tools (CIBERSORT, MCP-counter, EPIC, and xCell) [1, 14–17] were integrated to estimate infiltration of immune cells in the TCGA-LIHC RNA-seq cohort. By the approach of gene set variation analysis (GSVA) algorithm [12], Gene Ontology (GO) [18], Kyoto Encyclopedia of Genes and Genomes (KEGG) [19], REACTOME [20], and HALLMARK [21] gene sets were used to estimate the pathway enrichment scores for each of the samples. Other prevalent gene signature scores concerning the TME, that is, the tumor intrinsic pathway and metabolism, were calculated for each sample using the Immuno-Oncology Biological Research (IOBR) R package.

### 2.5 Differentially expressed genes (DEGs) analysis

All differential expressed gene analyses were conducted using the DESeq2 package [22]. DEGs analysis was performed using a generalized linear model along with the Wald statistical test, assuming that the underlying gene expression count data were distributed in a negative binomial distribution using DESeq2. DEGs with p-value < 0.05 were considered for further analysis. The adjusted p for multiple testing was calculated using the Benjamini-Hochberg correction [23].

## 2.6 Functional and pathway enrichment analysis

Gene annotation enrichment analysis was performed using the clusterProfiler R package [24]. GO [18] and KEGG [19] terms were identified with a strict cutoff of  $p < 0.01$  and a false discovery rate (FDR) of less than 0.05. Pathways that were up- and downregulated among groups were identified by running GSEA [25] of the adjusted expression data for all transcripts.

## 2.7 scRNA-seq data processing

An R-based toolkit, Seurat (version 4.0.4) [26], was used to analyze the scRNA-seq data. Specifically, the raw unique molecular identifier (UMI) matrix was processed to filter out genes detected in less than 10 cells and cells with fewer than 200 genes. The number of genes and UMI counts for each cell was quantified, and high-quality cells with thresholds of 500 UMIs, 100 genes, and less than 25% mitochondrial gene counts were preserved to ensure that most of the heterogeneous cell types were included for downstream analyses. DoubletFinder [27] was then applied to each sequencing library to remove potential doublets with an expected doublet rate of 10%, and cells with double scores (DF\_pANN) larger than the 90% quantile were filtered out. The normalized expression matrix was calculated based on the raw UMI counts after normalizing the total counts per cell (library size), scaled by  $1e6$ , and logarithmically transformed. Individual sea objects were then integrated using the canonical correlation analysis (CCA) function. The top 2,000 most variable genes from each sample were combined to identify the CCA vector.

## 2.8 Cell annotation

The gene expression values of the cells from a previous single-cell study [28] with cell annotation were used to train the prediction model using the maximum attribute dependency (MDA) algorithm, following the analysis step of the scPred package [29]. The model was then used to predict specific cell types of integrated data from the 19 scRNA-seq datasets.

## 2.9 Statistical analysis

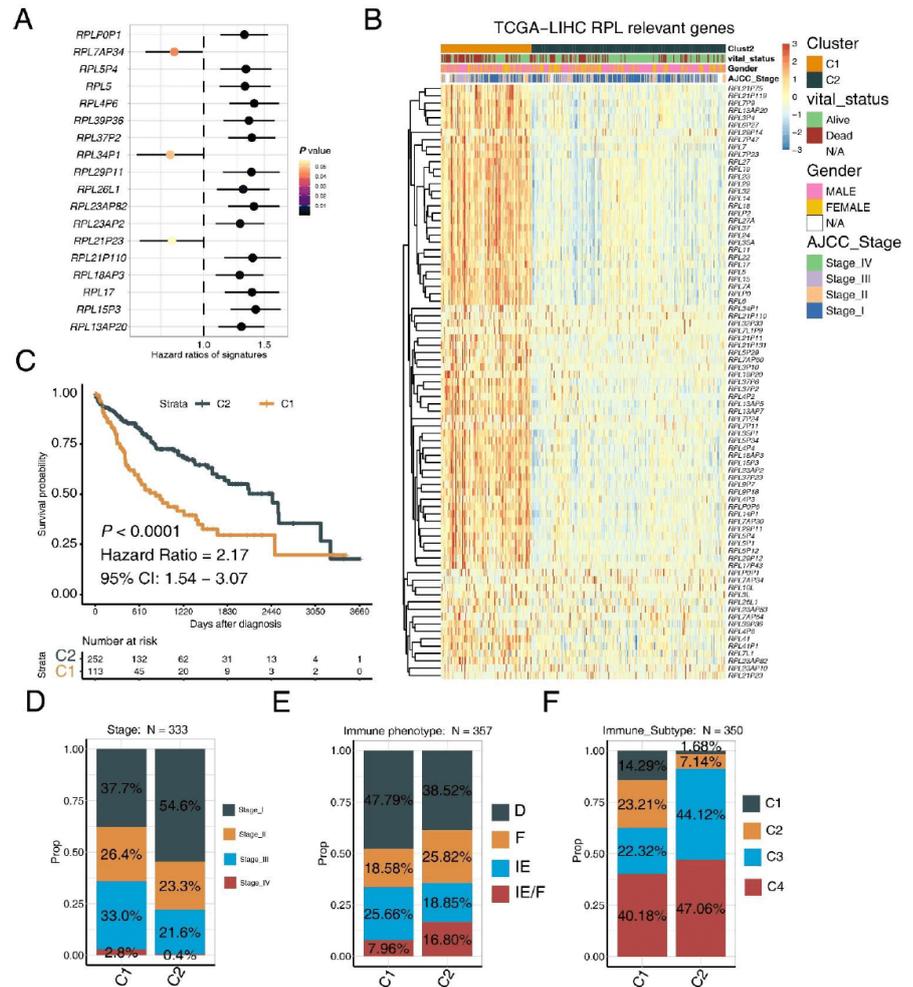
The normality of the variables was tested using the Shapiro-Wilk normality test [30]. For comparison of two groups, statistical significance for normally distributed variables was estimated using unpaired Student's t-test, and non-normally distributed variables were analyzed using the Mann-Whitney U test. For comparisons of more than two groups, Kruskal-Wallis and one-way ANOVA tests were used for non-parametric and parametric methods, respectively [31]. The correlation coefficient was calculated using Spearman and distance correlation analyses. Two-sided Fisher's exact test was used to analyze contingency tables. The best cutoff value in each dataset was evaluated based on the association between the survival outcome and RPL score in each dataset, using the survminer package. The Kaplan-Meier method was used to generate survival curves for the subgroups in each dataset, and the log-rank (Mantel-Cox) test was used to determine statistically significant differences. All statistical analyses were conducted using R (<https://www.r-project.org/>), and p values were two-sided. Furthermore, p values of less than 0.05 were considered statistically significant.

# 3 Results

## 3.1 RPL gene cluster was associated with HCC survival

Transcriptomic data of the RPL genes from TCGA-LIHC [32] were used to generate a forest plot. According to the forest plot, most RPL genes were significantly negatively correlated with poor prognosis in patients with HCC (Figure 1A). RPL genes that were statistically associated ( $P < 0.05$ ) with OS were retained for clustering. Notably, patients with RPL genes could be very stable and could be clustered into two categories: C1 and C2, respectively (Figure 1B). Survival analysis indicated a significant difference between the two clusters of LIHC. C1 patients (*i.e.*, those with high expression of the RPL gene) had significantly poorer OS than C2 patients (Figure 1C). The relationship between the RPL gene cluster and the tumor stage of patients was analyzed, and it was found that C1 tumors were significantly enriched in HCC at an advanced stage (Figure 1D,  $P = 0.008$ ). Using data concerning TME subtype [34] in a previous study, it was observed that C1 HCC was more of an immune desert tumor (47.79%), while C2 HCC was more of an inflammatory type and immune expulsion phenotype (Chi-squared test,  $P = 0.025$ ). Consistently, another type of immunophenotyping [35] also revealed a similar phenomenon (Figure 1F, chi-squared test,  $P = 8.269-10$ ). These results indicated that RPL genotyping was

not only associated with the prognosis of HCC but may also be related to the evolution of the TME in HCC.



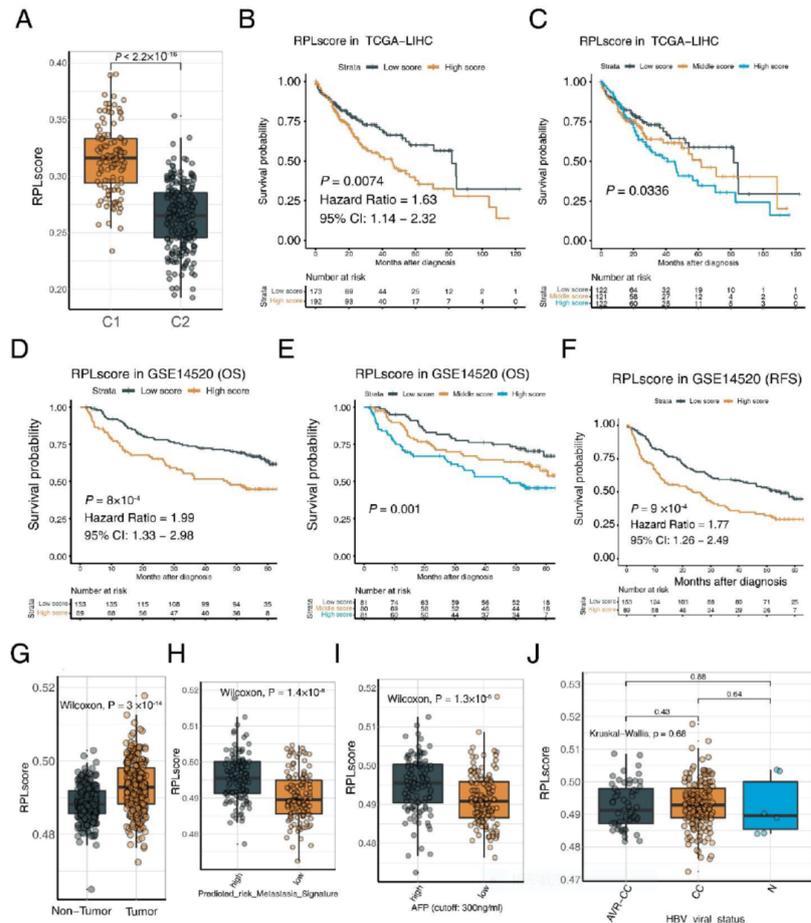
**Figure 1** High ribosome protein large gene expression associated with poor prognosis of liver cancer

**Notes:** A. Forest plot of most significant RPL genes which correlated with liver cancer overall survival. B. Clustering of RPL genes in TCGA-LIHC cohort. C. Kaplan-Meier survival analysis demonstrated that RPL gene activation (C1, cluster 1) was associated with worse overall survival than C2 (Cluster 2) tumors. ( $P < 0.0001$ , Hazard Ratio = 2.17, 95% CI: 1.54 – 3.07). D. Proportion of tumor stage in RPL gene cluster 1 and 2. E. Proportion of immune phenotypes in RPL gene clusters 1 and 2. D: immune desert, F: fibrotic, IE: immune-enriched, IE/F: immune-enriched, fibrotic. F. Proportion of immune phenotypes in RPL gene clusters 1 and 2. C1: wound healing; C2: IFN- $\gamma$  dominant; C3: inflammatory; C4: lymphocyte-depleted.

### 3.2 Construction and clinical significance of RPLscore

To better evaluate the RPL gene and its pathway in other HCC datasets, an RPL gene signature that could specifically quantify RPL-related gene clusters (Figure 1A) was constructed using a random forest algorithm, and the RPL score of each sample was obtained using the ssGSEA [32] methodology. The RPL score could significantly distinguish between patients with C1 and C2 HCC (Figure 2A) and was significantly correlated with the prognosis of HCC patients, whether taking the average value as a cutoff or three categories (Figure 2B). The same method was used for other HCC datasets (GSE14520) [36] and verified that the RPL score was significantly correlated with overall survival and relapse-free survival of HCC (Figure 2D-F). The RPL score remained significant in a multivariate model that included age, stage, and sex, which indicated that it was an independent prognostic factor (Figure 3A,  $P = 0.002$ ). Furthermore, the RPL score of the tumor tissue was significantly higher than that of the adjacent normal tissue (Figure 2G) and was significantly positively correlated with the metastasis risk score and alpha-fetoprotein (AFP) levels defined by the researcher (Figure 2H-I). However, the data did not show that the RPLscore was associated with hepatitis B virus (HBV) infection in the liver (Figure 2J). These findings suggested that higher expression of the RPL gene was a poor prognostic factor for

HCC.



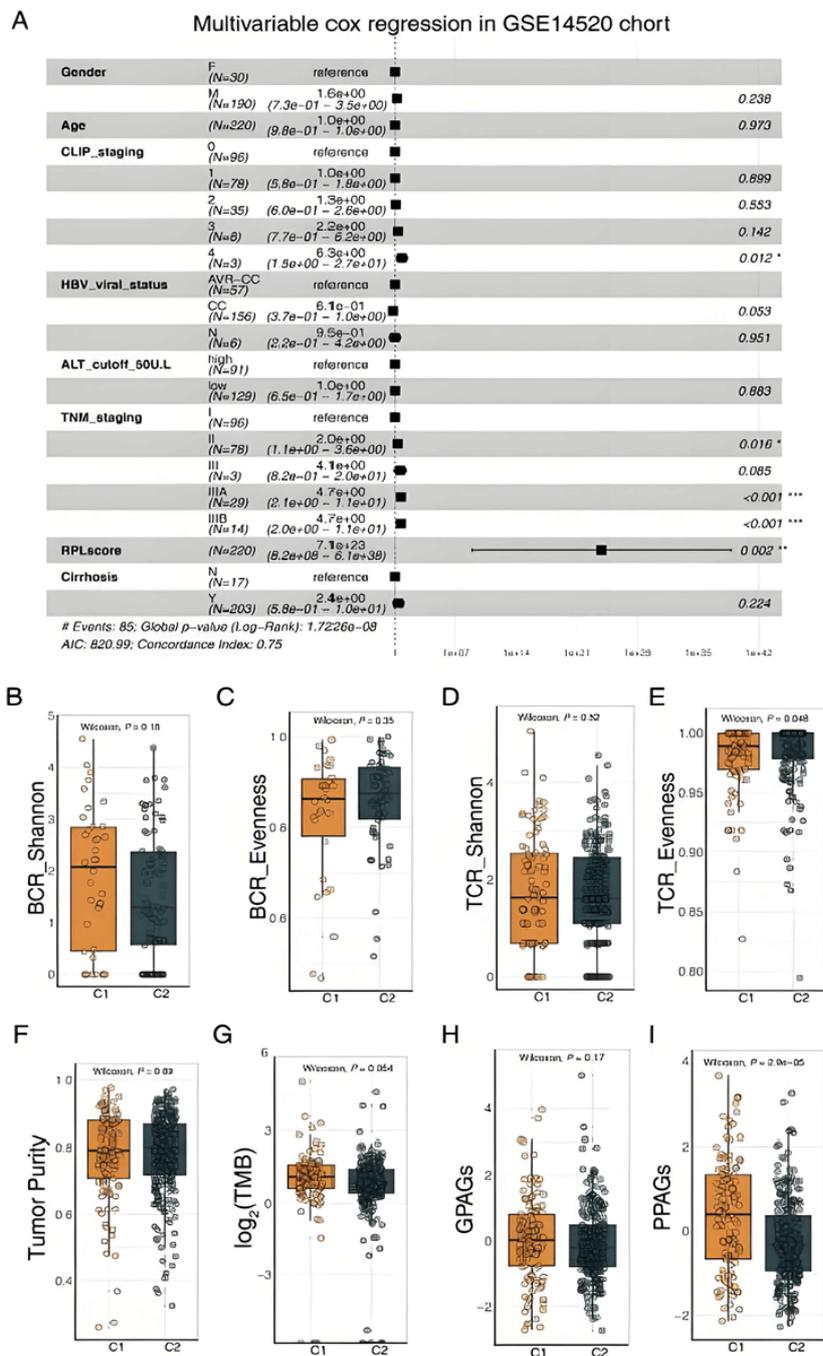
**Figure 2** RPLscore is an independent prognostic biomarker of liver cancer

**Notes:** A. RPLscore was statistically high in C1 tumors ( $P ; 2.2 \times 10^{-16}$ ). B. Kaplan-Meier survival analysis demonstrated that a lower RPL score was significantly related to more favorable overall survival in the TCGA-LIHC cohort ( $P = 0.0074$ , Hazard Ratio = 1.63, 95% CI: 1.14 - 2.32). C. Overall survival analysis showed that when we divided the RPL score into three parts, the statistical difference was still significant in the TCGA-LIHC cohort. ( $P = 0.0336$ ) D. Kaplan-Meier survival analysis demonstrated that a lower RPL score was significantly related to more favorable overall survival in the GSE14520 cohort ( $P = 8 \times 10^{-4}$ , Hazard Ratio = 1.99, 95% CI: 1.33 - 2.98). E. Overall survival analysis showed that when we divided the RPL score into three parts, the statistical difference was significant in the GSE14520 dataset. ( $P = 1 \times 10^{-3}$ ) F. Kaplan-Meier survival analysis demonstrated that a lower RPL score was significantly related to more favorable recurrence-free survival in the GSE14520 cohort ( $P = 9 \times 10^{-4}$ , Hazard Ratio = 1.77, 95% CI: 1.26 - 2.49). G. The box plot compares the RPL score of tumor and normal tissues ( $P = 3 \times 10^{-14}$ ). H. The box plot compared the RPL score of high and low metastasis score liver cancer ( $P = 1.4 \times 10^{-8}$ ). I. The box plot compared the RPL score of liver cancer with high and low AFP levels ( $P = 1.3 \times 10^{-5}$ ). J. The box plot revealed that the RPL score was not associated with HBV infection ( $P = 0.68$ ).

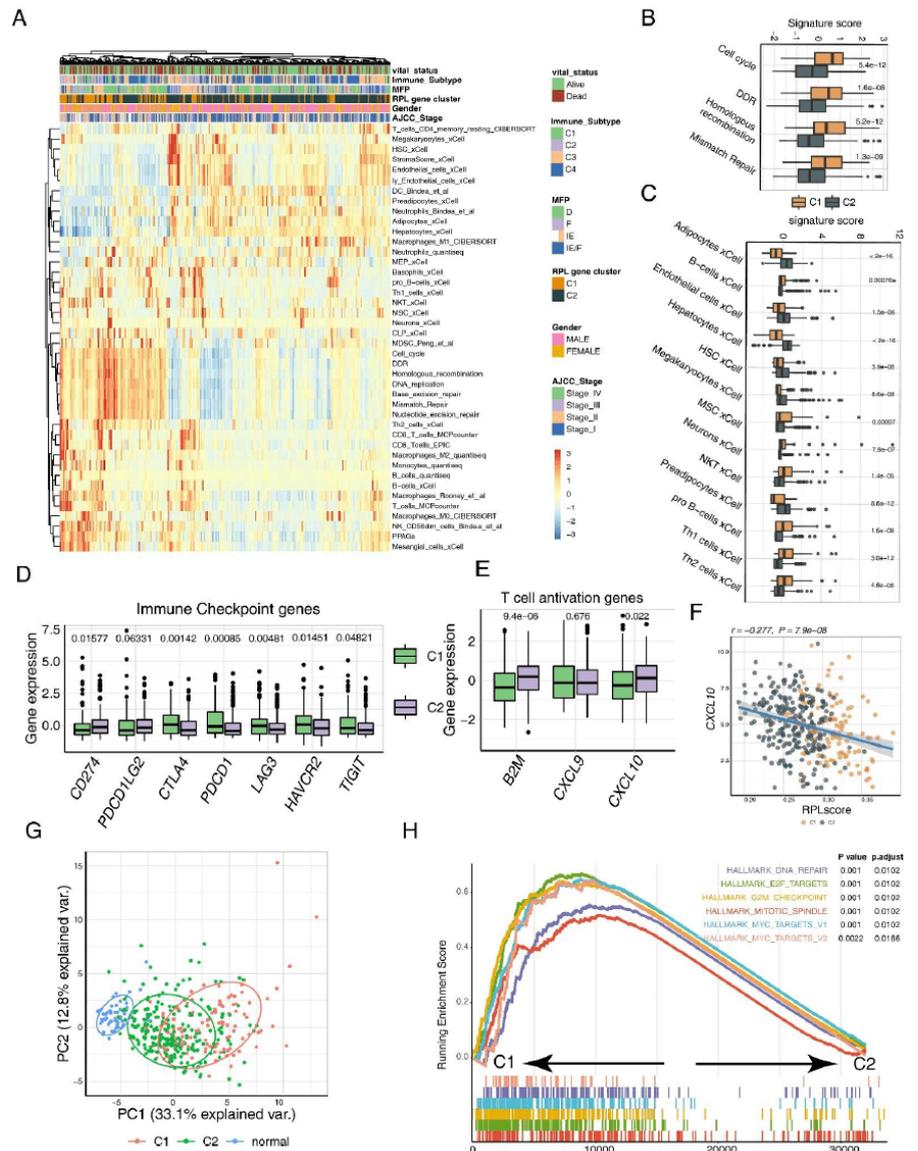
### 3.3 The immune landscape of RPL gene clusters

Various TME deconvolution algorithms [3, 14, 17, 37] and TME-related gene signatures [37] were applied to evaluate the microenvironment of each liver tumor, and correlations were estimated between these variables and RPL gene clusters in batches. It was observed that the deoxyribonucleic acid (DNA) mismatch repair pathway, cell cycle, and T cell infiltration were significantly upregulated in C1 tumors (Figure 4A-C). Although more T cells and B cells infiltrated these tumors, no significant differences were found between these two groups through the clonality and diversity analysis of B cells and T cell receptors (Figure 5B-E). It was also observed that C1 tumors had greater cell aggregation with immunosuppressive phenotypes (such as M2 macrophages and MDSCs). Furthermore, most immune checkpoint genes (CTLA4, PDCD1, LAG3, HAVCR2, and TIGIT) were upregulated in C1 tumors (Figure 4D), whereas T cell activation-related genes, *i.e.*, CXCL10 and B2M, were significantly upregulated in C2 tumors (Figure 4E-F). These results indicate that C1 tumors are immune-exhausted, which may facilitate tumor progression [38]. Consistently, the data revealed that C2 tumors exhibited significantly higher infiltration of M1 macrophages, dendritic cells, adipocytes, endothelial cells, and hepatic astrocytes (Figure 4C). This suggests that C2 tumors may exhibit a more pronounced immune activation [39]. Therefore, it was speculated whether there was a difference

in tumor purity between C1 and C2 tumors, but the data showed that there was no difference in tumor purity (*i.e.*, ABSOLUTE score) [40] between C1 and C2 tumors (Figure 3F). C1 tumors had a higher mutation rate than C2 tumors, but the difference was not statistically significant ( $P = 0.054$ , Figure 3G). These findings suggest that the microenvironment of C2 tumors may be more similar to that of a normal liver. Principal component analysis (PCA) using the microenvironment variables confirmed this speculation (Figure 4D). Relevant data also partly explain why the prognosis of C1 tumors was significantly different from that of C2 tumors (Figure 1B). Vascular dysplasia was observed in C1 tumors [41], which was previously reported to be regulated by suppressive immune cell types [42, 43] (Figure 3H-I). Analysis of differentially expressed genes [42] (Table S4) and GSEA [45] in patients with C1 and C2 tumors revealed that the DNA mismatch repair pathway and MYC pathway activation were significantly enriched in C1 tumors (Figure 4E), which was consistent with the results of the ssGSEA analysis (Figure 4A) and findings of previous studies [44].



**Figure 3** Multivariable cox regression in GSE 14520 chort



**Figure 4** The tumor microenvironment difference between two RPL gene clusters

**Notes:** A. The heatmap demonstrated clustering of tumor microenvironment cells and relevant gene signatures of hepatocellular carcinoma (TCGA-LIHC). B. The box plot indicated that the cell cycle, DNA damage response, homologous recombination, and mismatch repair pathways were upregulated dramatically in C1 tumors ( $P = 5.4 \times 10^{-12}$ ,  $1.6 \times 10^{-8}$ ,  $5.2 \times 10^{-12}$ ,  $1.3 \times 10^{-9}$ ). C. The box plot revealed that the number of adipocytes, B cells, endothelial cells, and hepatocytes was significantly higher in C2 tumors. Conversely, T and pro-B cells were enriched in C1 tumors. D. Principal component analysis indicated that the tumor microenvironment of C2 tumors was more similar to that of the normal liver. E. GSEA revealed that DNA repair and the E2F and MYC pathways were enhanced in C1 tumors.

### 3.4 RPLscore-relevant metabolism of HCC

Ribosome biogenesis is one of the most multifaceted and energy-demanding biological processes [1]. In the above-mentioned study, it was found that the TME and adipocytes of HCC were significantly correlated with the RPL gene cluster (Tables S5–6). It was speculated that there may be great differences in tumor metabolism between these two types of patients. Therefore, 144 metabolic pathways, as defined in a previous study [45], were used to evaluate their correlations (Figure 6A). The results showed a good relationship with the TME mentioned above. C2 tumors had significantly higher fatty acid catabolism, arginine biosynthesis, and steroid metabolism (Figure 6B-E). C1 tumors showed significantly activated purine metabolism and deficient arginine biosynthesis (Figure 6B). Arginine is the key raw material for the activation of T cells and the formation of immune memory. These results partly explain the immunosuppressive phenotype of C1 tumors. Using the GSE14520 data, the close relationship between these metabolic pathways and RPL score was further reproduced (Figure 5A-E).



### 3.5 Genomic variation correlates with RPLscore

To explore the malignant biological behavior of HCC mediated by the upregulation of RPL genes, the genomic data of TCGA-LIHC were further used to identify specific mutations that may be related to the RPL score. Using the data of full exon sequencing, it was noticed that tumor mutational burden (TMB) was significantly higher in the high RPL score group ( $P = 0.012$ , Figure 7A). Furthermore, ARID1A (12% vs. 3%) and TRPM6 mutations (6% vs. 1%) were significantly enriched in HCC patients with high RPLscore (Figure 7B-C). The data also showed that patients with the ARID1A mutation had a higher RPL score than those with wild-type tumors (Figure 7D). Recent studies have shown that the ARID1A mutation is closely related to the tumor mismatch repair system [46], which was confirmed by significant activation of the mismatch repair pathway in C1 tumors (Figure 7A-B). However, according to the mutation spot analysis results, no hotspots of ARID1A were enriched in tumors with higher RPLscore (Figure 7E). This was most likely caused by a lack of patients. Next, the copy number alterations of RPL-related genes in the TCGA cohort were analyzed, but no genes or sum of RPL genes with a significant trend of variance among RPL gene clusters were identified (Figure 7F-G).

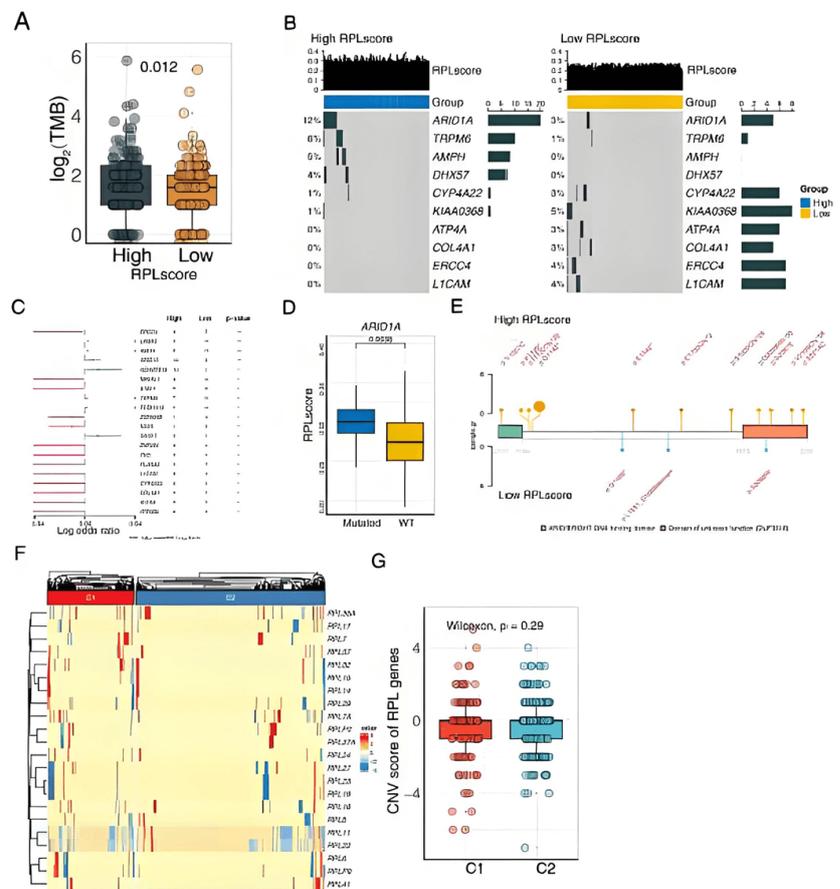


Figure 7 Malignant biological behavior of HCC mediated by the upregulation of RPL genes

### 3.6 RPLscore was associated with the biodiversity of liver tumors in single-cell resolution

To determine which cell type upregulated RPL genes, single-cell RNA-seq data from an HCC study [47] were used, which contained 19 samples derived from liver tumors. First, a serious batch effect between Datasets 1 and 2 was observed, which may contribute to misleading results of the analysis shown in Figure 8A. After cell filtration (see Methods section) and elimination of doublets [48], the CCA method [26] was used to eliminate the batch effect among samples (Figure 8B). The scPred tool [29] was used to automatically annotate the single-cell data according to previous annotations. Through this procedure, a total of 8,308 cells were obtained, including 2,802 T cells, 1,222 B cells, 819 myeloid cells, 1,490 endothelial cells, 973 fibroblasts, and 1,490 epithelial cells. Using t-distributed stochastic neighbor embedding

(t-SNE) (Figure 9A), clustering among the different cell types was observed. Marker genes of cells also provided further proof of the stability of the annotation results (Figure 8C). For example, epithelial cells had high expression of KRT19, and T cells had upregulated CD3D, CD4, CD8, etc., whereas myeloid cells exhibited significantly higher expression of CD14, CD163, and CD68. To evaluate the relationship between RPLscore and single-cell phenotype, the genes that positively correlated with RPLscore were used to estimate each cell using the AUCell algorithm [49]. The resultant data showed that malignant cells had higher up-regulation of RPLscore (Figure 9B-C). Furthermore, the average RPLscore of tumor cells was used as the RPLscore of each sample and was found to be significantly positively correlated with the tumor biodiversity score (ITH score) defined by a previous study [47], which was significantly relevant to resistance to immunotherapy. This suggests that upregulation of the RPL gene is significantly correlated with tumor heterogeneity, which may further mediate therapeutic resistance. The tumor cells (with a resolution of 0.3 of integrated data) were further classified and four subtypes were identified. Of these, Type 3 and Type 2 tumor cells had a significantly higher RPL score ( $P < 2.2 \times 10^{-16}$ ). Enrichment analysis and evaluation of metabolic pathways revealed that these cell types showed enhanced metabolic activity, with upregulated metabolic pathways related to various drugs. Conversely, antigen-presenting pathways were predominantly upregulated in Type 0 and Type 1 cells (Figure 9G-H). These results explain why B2M and CXCL10 genes were upregulated in C2 tumors (Figure 4E), which may induce anti-tumor immunity. These results also suggest that RPL gene dysregulation may facilitate tumorigenesis through metabolic reprogramming and drug metabolism, which induces tumor evolution and immune evasion.

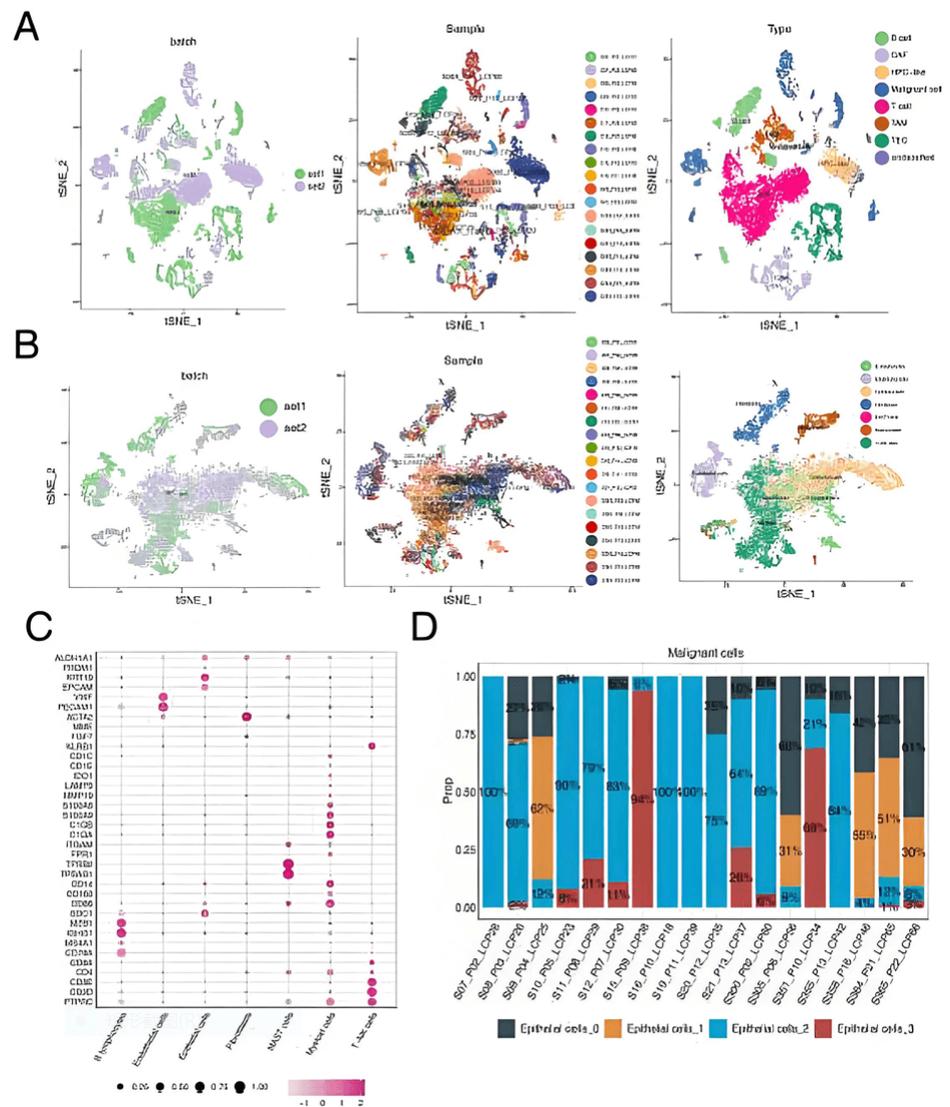
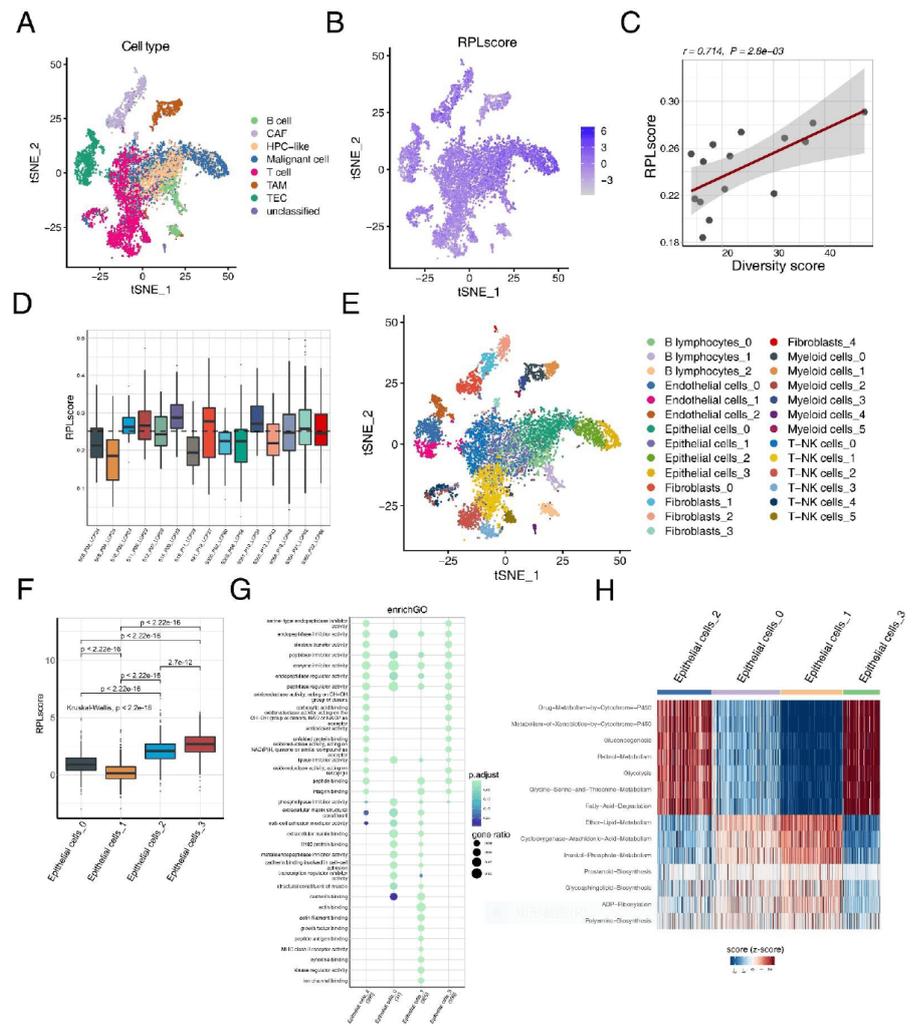


Figure 8 Batch effect between datasets 1 and 2



**Figure 9** PRLscore was associated with tumor biodiversity and immune invasion

**Notes:** A. TSNE plot shown the clustering of all cell types. B. FeaturePlot shown the RPLscore was highly upregulated in malignant cells. C. Scatter plot indicated the correlation between RPLscore and tumor biodiversity score. D. Boxplot shown the distribution of RPLscore in each sample E. Cell type re-clustering and dimension plot. F. Distribution of RPLscore between four malignant cell types. G. Enrichment analysis of differential expressed genes between four malignant cell types. H. Differential metabolism signatures between four malignant cell types.

## 4 Discussion

In this study, the relationship between RPL gene expression and the biological behavior of HCC was explored by integrating transcriptome, genomic, immunogenomic, and metabolomic data. It was found that dysregulation of RPL was associated with poorer OS and relapse-free survival in patients with HCC. Furthermore, the comprehensive analysis demonstrated that ribosome biogenesis is a promising target pathway in HCC. It was largely assumed that cancer cells become addicted to ribosomes owing to their enhanced need for protein production to maintain their unrestricted growth [1]. However, the relationship between the ribosomal pathway and TME of HCC, as well as the metabolic landscape, is largely unknown.

By employing the bootstrapping method [33], hepatocellular carcinoma (HCC) was robustly classified into two distinct clusters. The analysis of the tumor microenvironment (TME) revealed that tumors with higher RPL gene expression exhibited increased infiltration of cells with an immunosuppressive phenotype. Conversely, tumors expressing lower levels of the RPL gene displayed a TME closely resembling that of the normal liver microenvironment. This shows that the normalization of TME is a factor that should not be neglected in the treatment of HCC [50], which also suggests that RPL gene dysregulation may drive the immune suppressive phenotype of tumors.

Moreover, the key factor leading to heterogeneity in the tumor environment is probably the imbalance in the RPL pathway. It was also found that C1 tumors (high RPL score) had

significantly more vascular dysplasia, which was reported to be related to the presence of immunosuppressive cells in the tumor, such as MDSCs and macrophages. Metabolic pathway analysis explained that C1 tumors had more T cell infiltration with an immunosuppressive phenotype because of the lack of arginine biosynthesis in the TME. Arginine consumption by MDSCs, macrophages, and tumor cells [51] induces T cell exhaustion, which may promote immune evasion and proliferation of cancer cells. Furthermore, ribosomal protein synthesis requires a large amount of arginine, which exacerbates the lack of arginine in the TME [52].

To identify the intrinsic mechanism of ribosome dysregulation, genomic data from TCGA were used to identify the most related mutations or alterations in the copy number. It was observed that ARID1A mutation was mildly associated with a high RPL score. However, the mutation frequency of ARID1A is low and is not equal to the number of patients with ribosome pathway activation. By utilizing the AUCCell algorithm, the RPL score was quantified in single-cell resolution and observed to be significantly positively correlated with the tumor biodiversity score (ITH score), which was defined in a previous study [47]. This may mediate resistance to immunotherapy. Further analysis indicated that this phenotype might induce metabolic reprogramming and drug metabolism.

Furthermore, current research sheds light on additional biomarkers that might influence the TME and treatment outcomes in HCC. For instance, studies have shown the elevation of DKK-1 and sPD-L1, and specific miRNA profiles in various cohorts, suggesting a complex interplay of signaling pathways that contribute to the dynamic nature of the TME [53–55]. These external findings resonate with the observations of the RPL pathway's impact on HCC progression and immune evasion. It suggests that alongside the RPL gene dysregulation, factors like DKK-1, sPD-L1, and specific miRNAs could serve as complementary biomarkers for diagnosing and stratifying HCC, offering insights into potential therapeutic targets. This integration of diverse biomolecular pathways could pave the way for a multifaceted approach to HCC management, emphasizing the need for further research to dissect the interconnections between these pathways and their collective impact on HCC pathogenesis and therapy.

## 5 Limitations of the study

While this study provides significant insights into the relationship between RPL gene expression and tumor progression in hepatocellular carcinoma (HCC), several limitations must be noted. The findings are based on a specific set of data, which may limit generalizability across different HCC populations due to variations in genetic and environmental factors. Additionally, the cross-sectional design precludes establishing causality between RPL dysregulation and changes in the tumor microenvironment (TME). The reliance on public genomic databases also introduces potential biases related to data completeness and representativeness. Furthermore, the intricate network of interactions within the TME is not fully explored, and the causal mechanisms remain speculative without direct experimental validation. These limitations underscore the need for further longitudinal and experimental studies to confirm these findings and clarify the mechanisms involved.

## 6 Conclusion

A comprehensive analysis of the ribosomal pathway and TME was conducted, which strongly emphasized the critical role of ribosomal dysregulation in HCC. However, determining the regulatory relationship between these factors is still warranted to identify the cause of this phenomenon. Moreover, the findings indicate that targeting the ribosomal pathway may normalize the TME of HCC. Further prospective studies are required to confirm the results and hypotheses.

## Conflicts of interest

The authors declare that they have no competing interests.

## Ethics approval

This study was conducted using publicly available data obtained from GEO (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi>) and TCGA (<https://portal.gdc.cancer.gov/>), which includes data that were de-identified and contributed by research participants who provided informed consent under ethical standards. The use of this data complies with all applicable terms and conditions set. No ethical approval was required as this study did not involve any direct interaction with

human participants or the use of personal identifying information.

## Funding

This research was supported by the Key Discipline Established by Zhejiang Province and Jiaxing City Joint-Oncology Medicine (2023-SSGJ-001), National Clinical Key Specialty Construction Project-Oncology department (2023-GJZK-001), Jiaxing Key Laboratory of Clinical Laboratory Diagnosis and Transformation Research (2023-lcjyzydyzh), 2023 Jiaxing Key Discipline of Nursing (Supporting Subject) (2023-ZC-007), Science Technology Project of Jiaxing-Key Research Project (2024BZ20004), and Jiaxing Key Laboratory of Oncology Radiotherapy (2021-zlzdssys). All study sponsors had no role in the study design, collection, analysis, and interpretation of data.

## Author contribution

GX Hou and ZM Lu had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis;

GX Hou: Study concept and design;

ZM Lu: Data acquisition;

SC Xu, GF Zhao: Data analysis and interpretation;

GX Hou: Manuscript drafting;

All authors read and approved the final version of the manuscript.

## Data availability statement

The original contributions presented in this study are included in the article/supplementary material, and further inquiries can be directed to the corresponding author.

## References

- [1] Pelletier J, Thomas G, Volarević S. Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nature Reviews Cancer*. 2017, 18(1): 51-63.  
<https://doi.org/10.1038/nrc.2017.104>
- [2] Wei J, Kishton RJ, Angel M, et al. Ribosomal Proteins Regulate MHC Class I Peptide Generation for Immunosurveillance. *Molecular Cell*. 2019, 73(6): 1162-1173.e5.  
<https://doi.org/10.1016/j.molcel.2018.12.020>
- [3] Prakash V, Carson BB, Feenstra JM, et al. Ribosome biogenesis during cell cycle arrest fuels EMT in development and disease. *Nature Communications*. 2019, 10(1).  
<https://doi.org/10.1038/s41467-019-10100-8>
- [4] An Y, Xia Y, Wang Z, et al. Clinical significance of ribosome production factor 2 homolog in hepatocellular carcinoma. *Clinics and Research in Hepatology and Gastroenterology*. 2024, 48(3): 102289.  
<https://doi.org/10.1016/j.clinre.2024.102289>
- [5] Ebright RY, Lee S, Wittner BS, et al. Deregulation of ribosomal protein expression and translation promotes breast cancer metastasis. *Science*. 2020, 367(6485): 1468-1473.  
<https://doi.org/10.1126/science.aay0939>
- [6] Elhamamsy AR, Metge BJ, Alsheikh HA, et al. Ribosome Biogenesis: A Central Player in Cancer Metastasis and Therapeutic Resistance. *Cancer Research*. 2022, 82(13): 2344-2353.  
<https://doi.org/10.1158/0008-5472.can-21-4087>
- [7] Wagner GP, Kin K, Lynch VJ. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences*. 2012, 131(4): 281-285.  
<https://doi.org/10.1007/s12064-012-0162-3>
- [8] Colaprico A, Silva TC, Olsen C, et al. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Research*. 2015, 44(8): e71-e71.  
<https://doi.org/10.1093/nar/gkv1507>
- [9] Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. *Applied Statistics*. 1979, 28(1): 100.  
<https://doi.org/10.2307/2346830>
- [10] Monti S, Tamayo P, Mesirov J, et al. Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. *Machine Learning*. 2003, 52(1):91-118.
- [11] Kursu MB, Rudnicki WR. Feature Selection with the Boruta Package. *Journal of Statistical Software*. 2010, 36(11).  
<https://doi.org/10.18637/jss.v036.i11>

- [12] Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*. 2013, 14(1).  
<https://doi.org/10.1186/1471-2105-14-7>
- [13] Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods*. 2015, 12(5): 453-457.  
<https://doi.org/10.1038/nmeth.3337>
- [14] Racle J, de Jonge K, Baumgaertner P, et al. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *eLife*. 2017, 6.  
<https://doi.org/10.7554/elife.26476>
- [15] Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biology*. 2016, 17(1).  
<https://doi.org/10.1186/s13059-016-1070-5>
- [16] Li B, Severson E, Pignon JC, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biology*. 2016, 17(1).  
<https://doi.org/10.1186/s13059-016-1028-7>
- [17] Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biology*. 2017, 18(1).  
<https://doi.org/10.1186/s13059-017-1349-1>
- [18] The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*. 2018, 47(D1): D330-D338.  
<https://doi.org/10.1093/nar/gky1055>
- [19] Kanehisa M, Sato Y, Kawashima M, et al. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research*. 2015, 44(D1): D457-D462.  
<https://doi.org/10.1093/nar/gkv1070>
- [20] Croft D, O'Kelly G, Wu G, et al. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Research*. 2010, 39(Database): D691-D697.  
<https://doi.org/10.1093/nar/gkq1018>
- [21] Liberzon A, Birger C, Thorvaldsdóttir H, et al. The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Systems*. 2015, 1(6): 417-425.  
<https://doi.org/10.1016/j.cels.2015.12.004>
- [22] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Published online February 19, 2014.  
<https://doi.org/10.1101/002832>
- [23] Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 1995, 57(1): 289-300.  
<https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- [24] Yu G, Wang LG, Han Y, et al. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*. 2012, 16(5): 284-287.  
<https://doi.org/10.1089/omi.2011.0118>
- [25] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*. 2005, 102(43): 15545-15550.  
<https://doi.org/10.1073/pnas.0506580102>
- [26] Stuart T, Butler A, Hoffman P, et al. Comprehensive Integration of Single-Cell Data. *Cell*. 2019, 177(7): 1888-1902.e21.  
<https://doi.org/10.1016/j.cell.2019.05.031>
- [27] Xi NM, Li JJ. Benchmarking Computational Doublet-Detection Methods for Single-Cell RNA Sequencing Data. *Cell Systems*. 2021, 12(2): 176-194.e6.  
<https://doi.org/10.1016/j.cels.2020.11.008>
- [28] Kim N, Kim HK, Lee K, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nature Communications*. 2020, 11(1).  
<https://doi.org/10.1038/s41467-020-16164-1>
- [29] Alquicira-Hernandez J, Sathe A, Ji HP, et al. scPred: accurate supervised method for cell-type classification from single-cell RNA-seq data. *Genome Biology*. 2019, 20(1).  
<https://doi.org/10.1186/s13059-019-1862-5>
- [30] Ghasemi A, Zahediasl S. Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. *International Journal of Endocrinology and Metabolism*. 2012, 10(2): 486-489.  
<https://doi.org/10.5812/ijem.3505>
- [31] Hazra A, Gogtay N. Biostatistics series module 3: Comparing groups: Numerical variables. *Indian Journal of Dermatology*. 2016, 61(3): 251.  
<https://doi.org/10.4103/0019-5154.182416>
- [32] Cancer Genome Atlas Research Network. Electronic address wbe, Cancer Genome Atlas Research N: Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell*. 2017, 169(7):1327-1341 e1323.
- [33] Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010, 26(12): 1572-1573.  
<https://doi.org/10.1093/bioinformatics/btq170>

- [34] Bagaev A, Kotlov N, Nomie K, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell*. 2021, 39(6): 845-865.e7.  
<https://doi.org/10.1016/j.ccell.2021.04.014>
- [35] Thorsson V, Gibbs DL, Brown SD, et al. The Immune Landscape of Cancer. *Immunity*. 2018, 48(4): 812-830.
- [36] Roessler S, Jia HL, Budhu A, et al. A Unique Metastasis Gene Signature Enables Prediction of Tumor Relapse in Early-Stage Hepatocellular Carcinoma Patients. *Cancer Research*. 2010, 70(24): 10202-10212.  
<https://doi.org/10.1158/0008-5472.can-10-2607>
- [37] Zeng D, Ye Z, Shen R, et al. IOBR: Multi-Omics Immuno-Oncology Biological Research to Decode Tumor Microenvironment and Signatures. *Frontiers in Immunology*. 2021, 12.  
<https://doi.org/10.3389/fimmu.2021.687975>
- [38] Vijayan D, Young A, Teng MWL, et al. Targeting immunosuppressive adenosine in cancer. *Nature Reviews Cancer*. 2017, 17(12): 709-724.  
<https://doi.org/10.1038/nrc.2017.86>
- [39] Zeng D, Ye Z, Wu J, et al. Macrophage correlates with immunophenotype and predicts anti-PD-L1 response of urothelial cancer. *Theranostics*. 2020, 10(15): 7002-7014.  
<https://doi.org/10.7150/thno.46176>
- [40] Carter SL, Cibulskis K, Helman E, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nature Biotechnology*. 2012, 30(5): 413-421.  
<https://doi.org/10.1038/nbt.2203>
- [41] Tian L, Goldstein A, Wang H, et al. Mutual regulation of tumour vessel normalization and immunostimulatory reprogramming. *Nature*. 2017, 544(7649): 250-254.  
<https://doi.org/10.1038/nature21724>
- [42] De Palma M, Biziato D, Petrova TV. Microenvironmental regulation of tumour angiogenesis. *Nature Reviews Cancer*. 2017, 17(8): 457-474.  
<https://doi.org/10.1038/nrc.2017.51>
- [43] Koliaraki V, Prados A, Armaka M, et al. The mesenchymal context in inflammation, immunity and cancer. *Nature Immunology*. 2020, 21(9): 974-982.  
<https://doi.org/10.1038/s41590-020-0741-2>
- [44] van Riggelen J, Yetil A, Felsner DW. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nature Reviews Cancer*. 2010, 10(4): 301-309.  
<https://doi.org/10.1038/nrc2819>
- [45] Rosario SR, Long MD, Affronti HC, et al. Pan-cancer analysis of transcriptional metabolic dysregulation using The Cancer Genome Atlas. *Nature Communications*. 2018, 9(1).  
<https://doi.org/10.1038/s41467-018-07232-8>
- [46] Kelso TWR, Porter DK, Amaral ML, et al. Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in ARID1A-mutant cancers. *eLife*. 2017, 6.  
<https://doi.org/10.7554/elife.30506>
- [47] Ma L, Hernandez MO, Zhao Y, et al. Tumor Cell Biodiversity Drives Microenvironmental Reprogramming in Liver Cancer. *Cancer Cell*. 2019, 36(4): 418-430.e6.  
<https://doi.org/10.1016/j.ccell.2019.08.007>
- [48] DePasquale EAK, Schnell DJ, Van Camp PJ, et al. DoubletDecon: Deconvoluting Doublets from Single-Cell RNA-Sequencing Data. *Cell Reports*. 2019, 29(6): 1718-1727.e8.  
<https://doi.org/10.1016/j.celrep.2019.09.082>
- [49] Aibar S, González-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. *Nature Methods*. 2017, 14(11): 1083-1086.  
<https://doi.org/10.1038/nmeth.4463>
- [50] Sanmamed MF, Chen L. A Paradigm Shift in Cancer Immunotherapy: From Enhancement to Normalization. *Cell*. 2018, 175(2): 313-326.  
<https://doi.org/10.1016/j.cell.2018.09.035>
- [51] Kelly B, Pearce EL. Amino Assets: How Amino Acids Support Immunity. *Cell Metabolism*. 2020, 32(2): 154-175.  
<https://doi.org/10.1016/j.cmet.2020.06.010>
- [52] An H, Ordureau A, Körner M, et al. Systematic quantitative analysis of ribosome inventory during nutrient stress. *Nature*. 2020, 583(7815): 303-309.  
<https://doi.org/10.1038/s41586-020-2446-y>
- [53] Elhendawy M, Abdul-Baki EA, Abd-Elsalam S, et al. MicroRNA signature in hepatocellular carcinoma patients: identification of potential markers. *Molecular Biology Reports*. 2020, 47(7): 4945-4953.  
<https://doi.org/10.1007/s11033-020-05521-4>
- [54] El-Gebaly F, Abou-saif S, Elkadeem M, et al. Study of Serum Soluble Programmed Death Ligand 1 as a Prognostic Factor in Hepatocellular Carcinoma in Egyptian Patients. *Current Cancer Drug Targets*. 2019, 19(11): 896-905.  
<https://doi.org/10.2174/1568009619666190718141647>
- [55] Watany M. Study of Dickkopf-1 (DKK-1) Gene Expression in Hepatocellular Carcinoma Patients. *JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH*. Published online 2017.  
<https://doi.org/10.7860/jcdr/2017/23095.9450>

(Edited by Snowy Wang)