Precision Clinical Medicine, 2025, **8**: pbaf014

DOI: 10.1093/pcmedi/pbaf014

Research Article

Decoding hepatobiliary-specific immune gene patterns in gastrointestinal cancers via gene ontology fingerprints, multi-omics, and experimental integration

Honglian Huang [1,2,‡] Yueping Zhan,[2,‡] Hui Zong,[1] Chenjun Huang,[2,3] Fan Yang,[4] Ziyi Wei,[1] Xin Qin,[2] M. James C. Crabbe,[5,6,7,*] Ying Wang,[2,*] Xiaoyan Zhang [1,*]

[1]Department of Bioinformatics, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China
[2]Department of Clinical Laboratory Medicine Center, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai 200437, China
[3]Department of Medicine II, University Medical Center Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim 68167, Germany
[4]State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310058, China
[5]Wolfson College, Oxford University, Oxford OX2 6UD, United Kingdom
[6]Institute of Biomedical and Environmental Science & Technology, University of Bedfordshire, Luton LU1 3JU, United Kingdom
[7]School of Life Sciences, Shanxi University, Taiyuan 030006, China
*Corresponding authors: Xiaoyan Zhang, xyzhang@tongji.edu.cn; Ying Wang, nadger_wang@139.com; M. James C. Crabbe, james.crabbe@wolfson.ox.ac.uk
‡Honglian Huang and Yueping Zhan contributed equally to this work.

Abstract

Background: Gastrointestinal (GI) cancers are characterized by high malignancy and poor prognosis. Tumors in different locations exhibit both commonalities and differences. Although immunotherapy has made progress in some GI cancers, the specific immune-related patterns in hepatobiliary tumors have not yet been fully elucidated.

Methods: Using our developed explainable gene ontology fingerprint (XGOF) method, a GI cancer GOF was established. By integrating omics data from 20 hepatocellular carcinoma (HCC) and 15 intrahepatic cholangiocarcinoma (ICC) tissues in our clinic with public databases, immune-related patterns specifically expressed in hepatobiliary tumors were identified via RNA, protein, methylation, tumor microenvironment (TME) analysis, and experimental verification.

Results: XGOF showed that GI cancers are related to diverse immune functions, especially macrophage migration. Compared to others, hepatobiliary tumors exhibit distinct patterns of gene expression, mutation, and methylation. Seven genes (APOA1, LBP, FGA, C9, APCS, ARG1, and MBL2) were identified as immune-related genes specifically decreased in hepatobiliary cancer. The impact of APOA1 on TME, prognosis, and genomic landscape in HCC was explored in prior research. In this work, the experiment confirmed the down-regulation of six genes in cancerous tissues. Moreover, LBP promoter methylation was elevated in cholangiocarcinoma. Single-cell analysis revealed downregulated immune genes in hepatocytes of HCC and cholangiocytes of ICC, enriched in humoral immunity and complement pathways. Additionally, the macrophage migration inhibitory factor (MIF) pathway was identified as a key signal in interactions between ICC tumor cells and microenvironmental cells.

Conclusion: This study identified immune-related gene patterns in hepatobiliary cancer, contributing to the discovery of novel immunotherapy targets and tumor biomarkers for future research.

Keywords: gastrointestinal cancer; hepatobiliary tumor; immune-related gene; multidimensional analysis

Introduction

Gastrointestinal (GI) cancer is a high-risk malignancy that ranks second in global tumor incidence. Its overall incidence continues to rise steadily, with a noticeable trend toward younger patients in recent years [1]. GI cancers primarily include esophageal cancer, stomach cancer, colorectal cancer, liver cancer, cholangiocarcinoma, and pancreatic cancer. Among the six most prevalent cancers in China, four are GI cancers. GI cancers have not yet been fully investigated due to their complex pathogenesis, which involves the interaction of various factors such as genetic information changes [2, 3], lifestyle factors [4], and viral infections [5]. Due to the high malignancy, poor prognosis, and elevated mortality rates, early diagnosis and treatment of GI cancers have become particularly important. While the early diagnosis of digestive tract cancers in clinical practice primarily depends on techniques like pathological biopsies, imaging scans, and serological markers, with surgery as the cornerstone of treatment supported by radiotherapy and chemotherapy [6], achieving a complete cure becomes challenging once local recurrence or malignant metastasis develops. Therefore, identifying new biomarkers and therapeutic strategies for GI cancers remains an urgent clinical challenge [7, 8].

Received 17 April 2025; revised 17 June 2025; accepted 19 June 2025. published 24 June 2025

Since Slaughter proposed the theory of field cancerization in 1953 [9], the scope of research has expanded the initial focus on head and neck squamous cell carcinoma to GI cancers [10–14]. This theory suggests that prolonged exposure to specific carcinogenic factors causes tissues in the affected area to experience genetic mutations and epigenetic alterations, thereby increasing the risk of cancer development in adjacent regions or even throughout the entire system [10–14]. A large number of research results based on this theory suggest that GI cancer may be a systemic disease resulting from an overall imbalance in biological homeostasis, where localized lesions in a single organ actually reflect the systemic pathological state of the body. For instance, Jung *et al*. [15] conducted a meta-analysis of diagnostic markers identified in colorectal cancer and found that some of these markers are also present in pancreatic cancer [16] and liver cancer [17].

On the other hand, tumor heterogeneity results in specific molecular mechanisms during the development and progression of GI cancers, whether originating from the same or different tissues. The specific expression of immune-related genes may regulate the activity of various immune cells within the tumor microenvironment (TME) [18], influence the recognition and elimination of tumor cells by immune cells, and affect interactions among immune cells. For example, in pancreatic tumors, a reduction in MHC class I molecule levels impairs antigen presentation, making it difficult for cytotoxic T cells to recognize tumor cells and allowing those cells to evade the immune system [19]. Similarly, the high expression of CXCL12 in tumor-associated fibroblasts (CAFs) within the gastric cancer TME affects the communication between CAFs and T cells, as well as between Tregs and macrophages, contributing to an immunosuppressive TME [20]. Therefore, studying and treating GI tumors from a system perspective, and investigating the common and specific biological functions and pathogenesis among different GI cancers will help to change the research paradigm and diagnostic and therapeutic strategies. Furthermore, focusing on the study of immune regulation-related genes can help enhance our understanding of the complex mechanisms of the tumor immune microenvironment in GI cancer.

In this study, we first utilized data from the Gastrointestinal Cancer Knowledge Database (GIDB) [21], previously established by our research group, to compare the commonality and specificity of molecular variation features in different types of GI cancers, which generated a landscape at three omics levels: gene mutations, expression, and methylation patterns. Furthermore, by combining omics and textual data, we analyzed the expression patterns of immune regulation-related genes across six types of GI cancers. We identified genes with specific expression in hepatobiliary malignancies and further validated the alterations in expression at the RNA and protein levels, as well as the methylation levels in the promoter regions of these key genes. Furthermore, we explored cellular composition, immune-related gene expression, and intercellular communication in hepatobiliary cancers by analyzing single-cell sequencing (scRNA-seq) data. This provides clues for future exploration of new targets for immunotherapy and tumor biomarkers.

Materials and methods

Data sources

Sequencing data of 1874 GI cancer samples were obtained from the TCGA database, including 185 cases of the TCGA-ESCA project, 633 cases of the TCGA-COADREAD project, 377 cases of the TCGA-LIHC project, 185 cases of the TCGA-PAAD project, 443 cases of the TCGA-STAD project, and 51 cases of the TCGA-CHOL project. The data type included mutational, transcriptomic, methylation and clinical data. Texts and multi-omics analysis results related to GI cancers in the GIDB database were obtained for screening GI immune-related genes. GI cancers-related literature in the PubMed database was used to construct a GI cancers explainable gene ontology fingerprint (XGOF) network [22].

XGOF network construction for GI cancers

XGOF is an interpretable gene ontology (GO) fingerprint method that quantitatively characterizes the association between genes and ontologies. XGOF uses named-entity recognition technology to extract genes and GO terms from domain literature, calculates enrichment scores by quantifying their co-occurrence frequency, and identifies significant associations based on adjusted *P*-values to construct a gene-GO knowledge network. The detailed algorithmic workflow is provided in supplementary pseudocode S1 (see online supplementary material). For GI cancers, the XGOF network was constructed with an adjusted *P*-value threshold of < 0.05. We screened GO-gene pairs related to immunity and combined them with the expression data of GI cancers from the TCGA database to calculate the correlation between immune GO and cancers. Meanwhile, using annotations of cell types and marker genes from the CellMarker [23] and Cell Taxonomy [24] databases, we analyzed the enrichment of significant genes in the XGOF network of GI cancers across immune cells, thereby calculating the association between digestive-tract tumors and immune cells. In addition, using the XGOF genes as target genes and all genes annotated in the KEGG database as background genes, we performed pathway enrichment analysis based on the Fisher test to calculate the associations between different signaling pathways and tumors.

Molecular analysis of specificity and commonality among GI cancers

We performed principal component analysis (PCA) on gene expression and methylation sequencing data of GI cancers in TCGA database. Subsequently, we integrated paired data from three omics levels (mutation, expression, methylation) and conducted clustering analysis using the R package "iClusterPlus" [25]. The mutation data were represented by a binary matrix (1: mutation, 0: no mutation). High-frequency mutations (>5%) and significantly different genes were included in the integrated cluster analysis. Parameter k was optimized through model adjustment, employing parallel computation. We explored the range of k from 1 to 10 and selected the best combination sparse model with a penalty parameter using the Bayesian information criterion (BIC). The optimal parameter k was chosen by calculating the deviation ratio, which represents the percentage of explained variation (EV) and is output by the tune.iClusterPlus function. We selected the lambda vector deviation ratio with the smallest BIC and determined the best clustering value k ($k = K + 1$) by plotting the clustering number ($K + 1$)–EV percentage curve, identifying the point where the percentage EV curve begins to decline as the optimal k value.

Based on the integrated analysis results from the GIDB database, we compared the commonality of genes of GI cancers. Then we identified common feature genes and further analyzed the molecular variations and literature counts of these genes.

Screening of immune-related genes with specific expression patterns in GI cancers

We utilized the GI cancer-associated genes from the GIDB database, integrated with GO functional enrichment analysis and manual curation, to identify immune regulation-related genes. Then, we performed clustering analysis on the RNA-seq sequencing data of these genes to obtain immune regulation-related genes exhibiting specific expression patterns.

RNA and protein level validation

Clinical samples, including 20 pairs of human hepatocellular carcinoma (HCC) and adjacent non-cancerous tissues, as well as 15 pairs of human intrahepatic cholangiocarcinoma (ICC) and adjacent non-cancerous tissues, were provided by Shanghai East Hepatobiliary Surgery Hospital. The system used for qPCR was the 7900HT Fast Real Time PCR System. Primer Premier 6.0 software was used for PCR primer design along with BLAST software for homology comparison. The primer sequences are shown in supplementary Table S1 (see online supplementary material).

Proteins were extracted using protein lysis solution (Beyotime Biotechnology Co., Ltd., Beijing), quantified by bicinchoninic acid (BCA) assay (Thermo Fisher, USA), and separated by 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). After transfer to PVDF membranes (Millipore, USA), blots were probed with primary antibodies (ab52945, Abcam, UK) overnight at 4°C, followed by horseradish peroxidase (HRP)-conjugated secondary antibodies. Signals were detected by enhanced chemiluminescence (ECL) (Thermo Fisher, USA) and analyzed using ImageJ. The relative expression level of genes was normalized by the level of glyceraldehyde-3-phosphate dehydrogenase (GAPDH).

Bisulfite amplicon sequencing

DNA methylation sequencing data for CHOL and LIHC projects were obtained from TCGA as the basis for designing methylation probes. According to the predictive results of gene expression and methylation correlation analysis, specific probes were selected for the promoter regions of the LBP gene at positions 68, 236, 271, and 315; the APCS gene at positions 33, 39, 157, and 247; and the ARG1 gene at positions 38, 53, 73, and 116. The methylation levels of the genes were validated by bisulfite amplicon sequencing (BSAS) in 9 pairs of ICC samples and 20 pairs of HCC samples. The genes to be validated in the ICC samples were APCS and LBP while APCS and ARG1 were used in the HCC samples. The raw gene data were quality controlled by FastQC and aligned to the reference sequence using Bowtie2. The methylation of each CpG cytosine was extracted by R's package methylKit, using default parameters, and the cytosine methylation level was further calculated. The methylation ratios of all CpG cytosines were calculated in the destination fragment of each sample.

Analysis of single-cell sequencing data

The ICC and HCC single-cell data were obtained from the GSE189903 dataset. We first selected samples from core tumor and normal tissues in this dataset and performed cell type annotation. Based on the annotation results, differential gene expression analysis was conducted to identify the expression patterns of six immune-related genes across different cell types using the FindMarkers function. Subsequently, functional enrichment analysis was performed on these differentially expressed genes to explore their potential immune-related biological functions using Metascape [26]. To further investigate the cellular interactions within the TME, we analyzed the interaction and communication networks between different cell types using the CellChat package [27]. We systematically compared differences in signaling patterns between tumor and adjacent non-tumor samples using the RankNet approach.

Immunofluorescence

Tissue microarrays (TMAs) for HCC and ICC (Shanghai Outdo Biotech, HLivH150CS06, HIBDA160PG01) were dewaxed in xylene and rehydrated through a graded series of decreasing alcohol concentrations. Antigen retrieval was performed using EDTA-based heat-induced epitope retrieval. After incubating the sections in 3% H_2O_2 for 20 min, they were blocked with Beyotime QuickBlock™ blocking buffer (P0260) for 30 min. Subsequently, the TMA was incubated with the primary antibody (MIF, STARTER, S0B0858) overnight at 4°C, followed by incubation with a Texas Red-X conjugated secondary antibody (Invitrogen, T-2767) for 1 h. Each step was followed by three 5-min washes with PBS. Imaging was performed using a Zeiss Tissue Gnostics tissue microscope. After excluding samples with tissue detachment, this study successfully obtained 119 valid immunofluorescence datasets. The intensity differences between tumor and adjacent normal tissues were analyzed using the Wilcoxon rank-sum test.

Results

Overview of GI cancer GOF

There were 117 878 unique gene-GO terms pairs in the GI cancers XGOF network, in which the maximum pair number was 42 531 (LIHC's GOF) (Fig. 1A). Macrophage migration (GO:1 905517) was strongly associated with most GI tumors. Additionally, complement activation was highly associated with LIHC and CHOL (Fig. 1B). According to the enrichment analysis based on the Cell Taxonomy database, DC cells, NK cells, and helper T cells show strong associations in all GI cancers (Fig. 1C). According to the CellMarker database, T cells were significantly correlated immune cell types in liver cancer, colorectal cancer, pancreatic cancer, and gastric cancer (Fig. 1D). Pathways including the TNF signaling pathway, PI3K-Akt signaling pathway, and focal adhesion were enriched to varying degrees in different types of GI cancers (supplementary Fig. S1, see online supplementary material).

Specific and common genes among different GI cancers

PCA analysis showed that ESCA, COADREAD, STAD, and PAAD had similar expression patterns at the genome-wide level, whereas LIHC and CHOL differed from them with specific expression patterns. PCA at the methylation level also demonstrated that some of the LIHC samples had specific methylation patterns (Fig. 2A).

The iClusterPlus analysis identified $k = 2$ as the optimal number of clusters, based on the point where the percentage of explained variation began to decline (Fig. 2B). The clustering analysis revealed that samples in Cluster 1 exhibited mixed cancer type clustering, failing to segregate distinctly by cancer type. In contrast, most of the LIHC samples were clustered in Cluster 2, indicating distinct gene mutation patterns specific to LIHC (Fig. 2C). It was also found that COADREAD had a high mutation frequency at the mutation level, while in LIHC the mutation rate of genes was generally lower than in other GI cancers. In addition, the annotation of clinical characteristics showed that the high prevalence of tumors in Asian populations was mainly concentrated in LIHC, STAD, and ESAC.

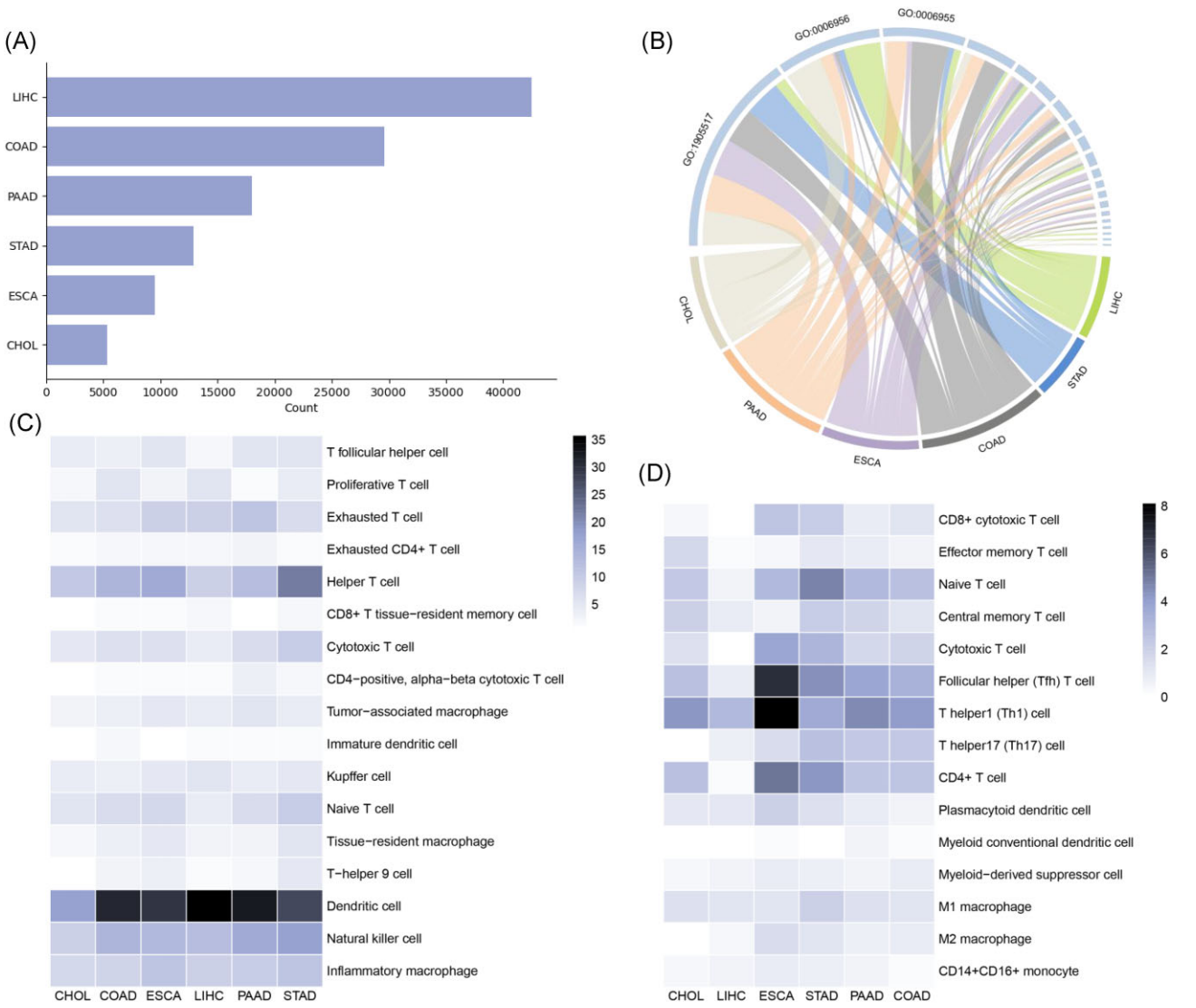

Figure 1. XGOF network construction for GI cancers. (**A**) The number of GO term pairs for six types of GI cancers. (**B**) Top 20 immune-related GO terms for GI cancers. The fan area represents the correlation between the term and cancer. GO:1 905 517, macrophage migration; GO:0006956, complement activation; GO:0006955, immune response. (**C**) Heat map of immune cell–tumor correlation based on Cell Taxonomy database annotation. (**D**) Heat map of immune cell–tumor correlation based on CellMarker database annotation.

The commonality between the molecular features of cancers may suggest a common mechanism in the tumorigenic process. A total of 122 feature genes shared by GI cancers was identified through common molecular analysis (supplementary Fig. S2, see online supplementary material). The clustering results of these genes are shown in Fig. 2D. At the mutation level, we found that some genes were mutated with high frequency (mutation frequency > 5%) in all GI cancers, e.g. MUC16 and TP53. Some of the common genes had the same molecular variants at multiple histological levels, including hypermethylated/hypo-expressed or hypomethylated/highly expressed genes, such as MMP1, SFRP1, COL14A1, CSF2, TAC1, MUC16, and FAT4. We also found several genes that were highly expressed but showed no differences at the methylation level, including AURKA, BRCA2, and DNMT3B. The ABCB1 gene, on the other hand, showed molecular variants opposite to these genes, with high-frequency mutations detected and significantly low expression level, while no difference was found at the methylation level. In addition, a number of genes were identified as high-frequency mutations in all GI cancers, but no differences were detected at the expression and methylation levels, such as in APC, ATM, CTNND1, DICER1, IGF2R, KRAS, NF1, PIK3CA, RB1CC1, ROCK1, SMAD4, and SUCO. By calculating the overall variation proportions of 122 common genes across three omics levels, we found that the top-ranked ADAM12 gene has been studied in all GI cancers, although the number of related publications remains limited. Multi-omics analysis of this gene revealed intricate molecular alterations in GI cancers, indicating that ADAM12 may play an important role in the occurrence and development of these cancers.

Identification of immune-related genes with specific expression patterns in GI cancers

Investigating the interactions between tumors and the host immune system is crucial for identifying key molecular markers of tumor prognosis, overcoming drug resistance, and developing innovative therapies. We found a total of 305 immune-related genes associated with GI cancers. Clustering analysis of these genes showed that samples in Cluster 3 included most of the LIHC and

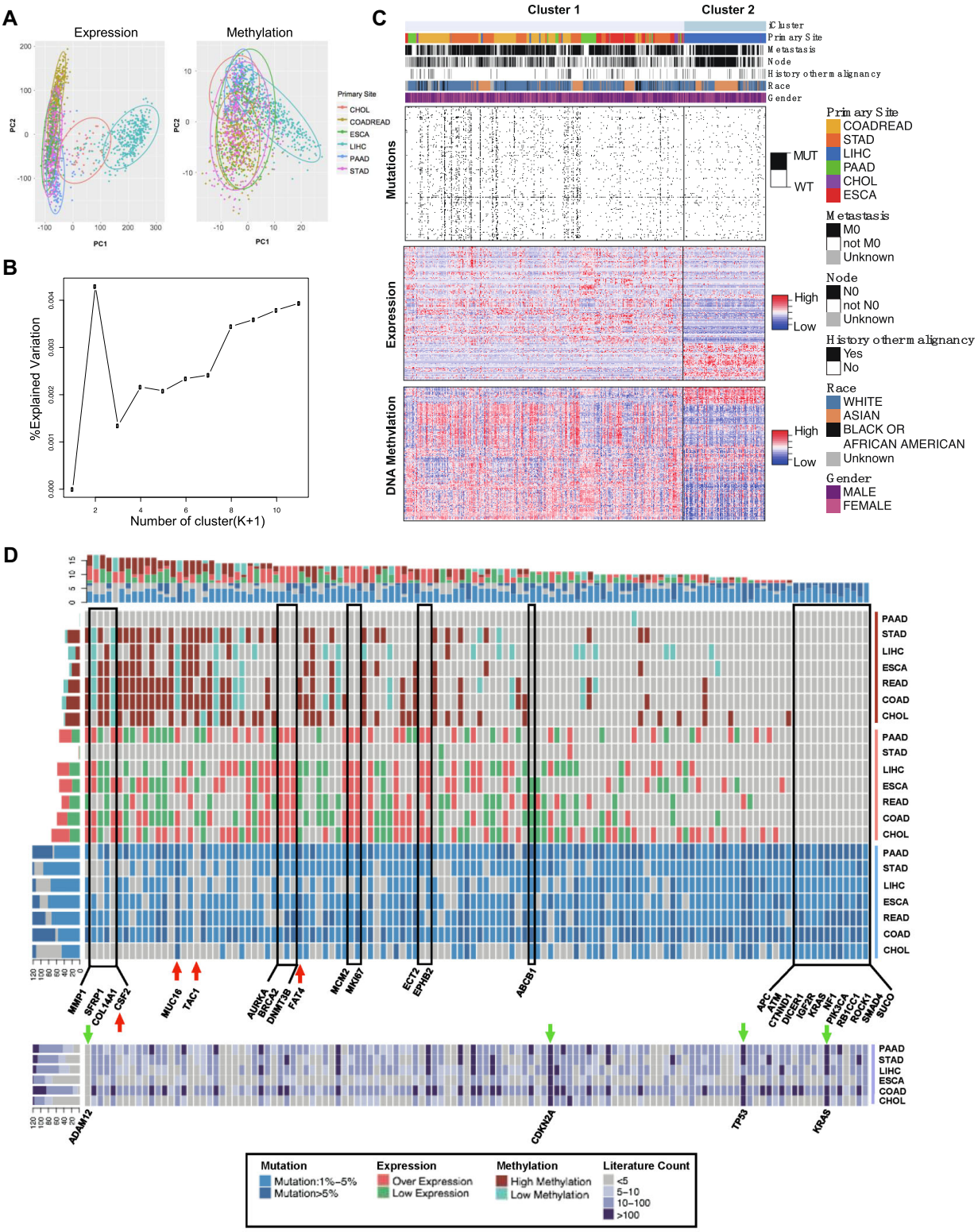

Figure 2. Molecular analysis of specificity and commonality among GI cancers. (**A**) Whole-genome expression and methylation PCA map of GI cancers. Left: PCA results at the expression level; right: PCA results at the methylation level. (**B**) Clustering number–EV percentage graph based on the iCluster method. When the % EV curve begins to decline, *k* is the optimal number of clusters. (**C**) Integrated clustering plot of GI cancers-related genes across mutation, expression, and methylation levels. The plot displays the clustering results of high-frequency mutation genes, differentially expressed genes, and differentially methylated genes associated with GI cancers. Above the clustering plot are annotations indicating sample classification, cancer information, and sample annotations under different clinical features. (**D**) Heat map of molecular features of 122 common genes in GI cancers. The upper part of the heat map indicates the features of methylation, expression, and mutation levels of these 122 genes. The lower part of the heat map indicates the corresponding number of papers. The genes labeled in the figure are the sets of genes with different features found in the papers.

CHOL cases, while samples from ESCA, STAD, COADREAD, and PAAD were dispersed between Cluster 1 and Cluster 2 (Fig. 3A). This suggests that hepatobiliary malignancies have specific gene expression patterns in immune regulation. Further analysis revealed that there were significant differences in the expression levels of seven key immune regulation genes (APOA1, FGA, C9, ARG1, MBL2, APCS, and LBP) between hepatobiliary cancers and other GI cancers. All seven genes were expressed at higher levels in hepatobiliary malignancies compared to other GI cancers. Additionally, the expression of these genes was significantly lower in LIHC and CHOL tumor tissues compared to their adjacent normal tissues (Fig. 3B–H). A detailed description of APOA1 has been provided in our previous research, therefore we will focus on the six other genes in this work [22, 28].

RNA and protein expression level analysis

In HCC samples, qPCR analysis revealed that the expression levels of LBP, FGA, MBL2, C9, ARG1, and APCS were significantly lower in tumor tissues compared to adjacent normal tissues ($P < 0.001$), confirming the downregulation of these genes (Fig. 4A). Similarly, in ICC samples, the expression of LBP, FGA, MBL2, C9, and APCS was significantly decreased ($P < 0.05$) (Fig. 4B). For the ARG1 gene in ICC samples, a similar decrease in expression was observed in some cases, but most samples had Ct values > 35 or were marked as "undetermined", suggesting unreliable detection.

Western blot analysis revealed that in HCC, protein levels of APCS, MBL2, and ARG1 were significantly reduced in tumor tissues compared to adjacent non-tumor tissues, as confirmed by ImageJ grayscale quantification ($P < 0.01$) (Fig. 4C). In ICC, only C9 exhibited significantly lower expression in tumor tissues ($P < 0.05$) (Fig. 4D).

Methylation level of gene promoter regions

DNA methylation sequencing data and RNA-seq data of HCC and ICC were obtained from TCGA. The methylation β-value and expression value of six genes were extracted for correlation calculation (Fig. 5A). We found that in gallbladder cancer, the methylation levels of LBP and APCS were negatively correlated with their expression levels, while in liver cancer, the methylation levels of ARG1 and APCS were negatively correlated with their expression levels. Subsequently, the BSAS experiment revealed that a higher methylation level was detected at position 236 of the promoter region of the LBP gene in ICC. This may suggest that the reduced expression of the LBP gene is associated with elevated methylation in its promoter region. No significant statistical differences were found in the methylation ratios of the promoter regions of the APCS gene in ICC and the two genes, ARG1 and APCS, in HCC, between tumor and adjacent non-tumor tissues ($P > 0.05$) (Fig. 5B). In nine paired cholangiocarcinoma samples, RNA expression and promoter methylation analyses revealed that LBP was significantly downregulated in tumor tissues ($P < 0.01$), with its promoter exhibiting hypermethylation (Fig. 5C). These findings suggest that LBP suppression may be associated with elevated promoter methylation.

Analysis of single-cell sequencing data

We manually annotated cell subpopulations in the GSE189903 dataset, noting that cluster 6 contained a mixture of hepatocytes and cholangiocytes. To further distinguish this subpopulation, we used specific cholangiocyte markers KRT7, KRT19, and ALB to differentiate cell types in cluster 6 (supplementary Fig. S3A-B, see online supplementary material). Cell annotation results were shown in Fig. 6A. In HCC samples, ARG1, FGA, APCS, and C9 genes were significantly downregulated in hepatocytes, while in ICC samples, these genes were downregulated in cholangiocytes (Fig. 6C, supplementary Fig. S3D) (FDR < 0.05). However, these genes exhibited no significant expression changes in other cell types within the TME. Functional enrichment analysis indicated these genes were significantly enriched in innate immune response, humoral immune response, and complement activation (Fig. 6D). Cell–cell interaction analysis further revealed significantly enhanced MIF signaling in tumor samples (supplementary Fig. S3E), with particularly strong MIF-mediated communication observed from malignant cells to multiple immune cell populations, such as macrophages, B cells, and Tregs (Fig. 6B).

Immunofluorescence

To investigate the cell–cell communication features of cholangiocarcinoma cells, we performed immunofluorescence experiments on ICC and HCC tissue microarrays. The results demonstrated that MIF expression was significantly higher in ICC tissues compared to adjacent tissues (Fig. 6E, $P = 0.043$). Figure 6F shows representative immunofluorescence staining images. However, in HCC, our experiment did not reveal an increase in MIF signaling. These findings suggest that within the ICC TME, cancer cells enhance intercellular communication signals with other cells by upregulating MIF expression.

Discussion

In this study, we compared the commonality and specificity of six GI cancers at the three histological levels of expression, mutation, and methylation. A comprehensive gene variation landscape of common gene modules in GI cancers was constructed, providing a multi-layered perspective to deepen the understanding of the molecular mechanisms underlying these tumors. The study identified and experimentally verified a set of immune-related genes with specific expression patterns in hepatobiliary tumors. In addition, we also explored the association between gene expression and immune cell infiltration in the TME.

XGOF revealed a strong association between macrophage migration and most of the GI tumors, as well as the association of GI tumors with different immune cells. Through a commonality analysis of GI tumors, we identified the expression profiles of 122 shared genes. This finding suggests that, despite differences in anatomical location and tissue origin among GI cancers, they may share certain key biological processes and exhibit convergent mechanisms in core carcinogenesis. For example, the common genes are involved in critical aspects of cancer development, including chronic inflammation (e.g. CSF225 [29] and SFRP126 [30]), cell cycle dysregulation and abnormal proliferation (e.g. AURKA27 [31]), and signaling pathway disturbances (e.g. KRAS28 [32], PIK3CA29 [33], and SMAD430 [34]). These genes may represent a “core molecular program” in GI tumors. This convergence may be linked to shared microenvironments or genetic susceptibilities within the GI system, offering valuable insights into their common etiology and potential for developing broad-spectrum therapeutic strategies.

In addition, specificity analysis revealed that hepatobiliary cancer exhibits molecular patterns distinct from other GI cancers. Further analysis focused on immune-related genes in hepatobiliary cancer, identifying a set of key genes with specific expression profiles. Our research shows that APCS, ARG1, MBL2, LBP, C9, and FGA were consistently downregulated in both liver

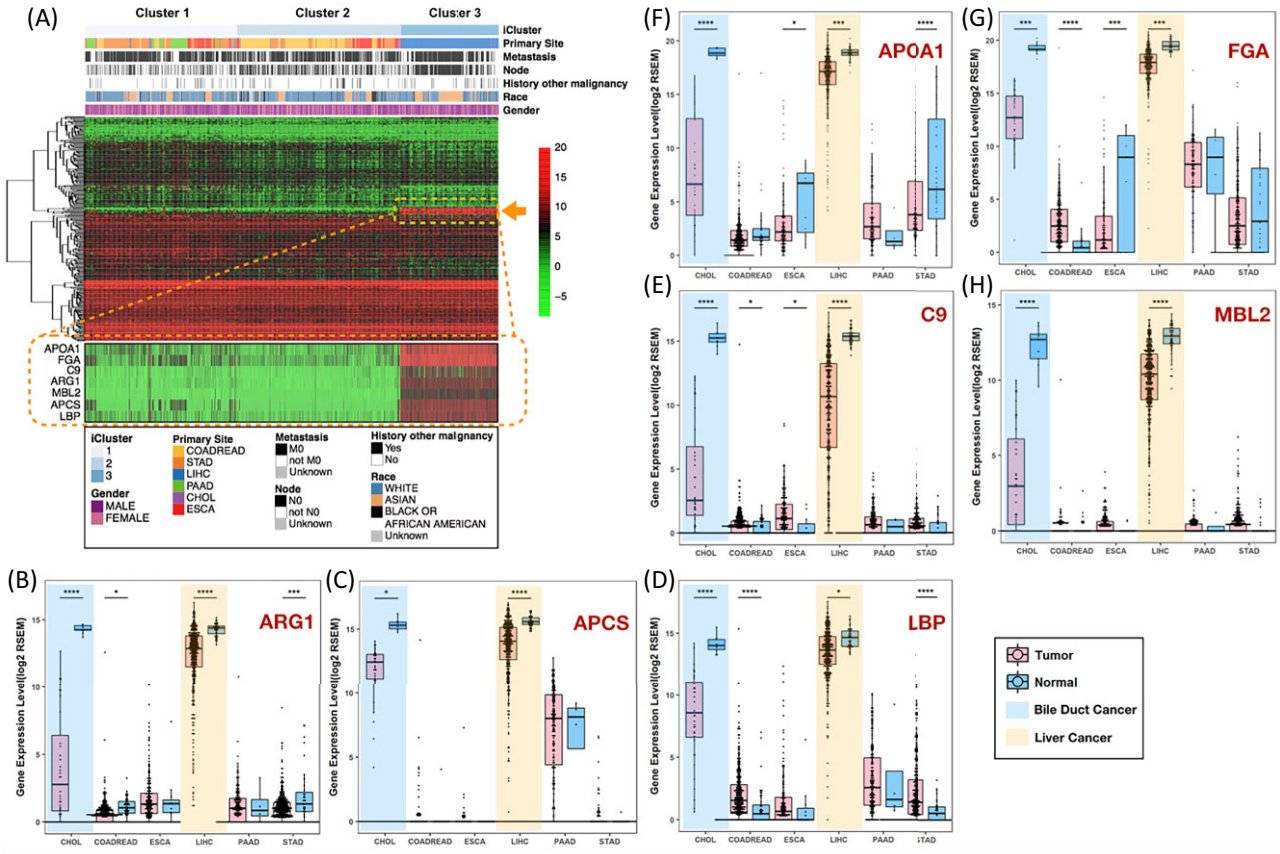

Figure 3. Immune-related genes with specific expression patterns in GI cancers. (**A**) Seven immune-related genes with specific expression in hepatobiliary malignancies. The clustering diagram shows the results of clustering the expression data of 305 immune-related genes associated with six GI cancers. The lower part of the clustering plot demonstrates the seven genes with specific expression patterns in hepatobiliary malignancies. (**B**–**H**) Boxplots of the differential expression profiles of seven genes in six GI cancers.

and gallbladder cancers, implying that these genes may act as tumor suppressors in the onset or progression of hepatobiliary cancer. Previous studies have demonstrated that ARG1 regulates metabolic reprogramming in liver cancer cells by influencing the "arginine–RBM39" signaling axis, thereby promoting the growth and proliferation of these cells [35]. FGA, as a candidate coding and non-coding driver gene, negatively regulates the YK2–STAT3–IL6 signaling pathway, thereby affecting the malignant phenotype of various liver cancer cells [36]. APCS is a potential candidate biomarker for HCC development in HCV-infected cirrhotic patients [37]. Low levels of MBL2 are associated with the prognosis of HCC [38]. LBP plays a role in regulating inflammatory responses by activating the JNK and NF-κB signaling pathways, leading to a significant increase in the expression of inflammation-related molecules [39]. Chronic inflammation can trigger hepatocyte carcinogenesis, thereby promoting cancer development and metastasis [40]. Although the mechanisms of complement C9 in hepatobiliary cancers has been less clearly described in other studies, Li *et al.* found that in the hypoxic TME of non-small cell lung cancer, the expression of C9 in tumor-associated macrophages was decreased, promoting the progression of non-small cell lung cancer [41].

The specific mechanisms underlying the downregulation of these genes remain to be clarified. In this study, we found methylation in the promoter region of the LBP gene in gallbladder cancer, which may represent a key epigenetic mechanism driving the downregulation of its expression. However, no significant differences were seen for the genes APCS and ARG1, which may suggest that the down-regulated expression of ARG1 and APCS genes is not caused by CpG methylation of the genes. We analyzed TCGA data on hepatobiliary cancers and found that APCS in cholangiocarcinoma was significantly negatively correlated with the expression of the M6A-associated gene METTL3 (supplementary Fig. S4, see online supplementary material).

Analysis of single-cell sequencing data revealed that multiple immune-related genes were significantly downregulated in hepatocytes (HCC) and cholangiocytes (ICC), with enrichment in innate immune response, humoral immune response, and complement activation. This further suggests that the suppression of these genes may weaken immune surveillance in the TME, thereby promoting immune escape. The MIF pathway serves as a key signal for intercellular communication in both HCC and ICC. In HCC, hepatocytes interact with immune cells such as macrophages and B cells via MIF, while in ICC, cholangiocytes similarly rely on MIF for communication with immune cells. As a key pro-inflammatory factor, MIF has been widely demonstrated to regulate immunosuppressive microenvironments in various cancers [42].

These findings not only reveal the potential mechanisms of dysregulated expression of six immune-related genes in hepatobiliary cancer but also provide new insights for clinical diagnosis and treatment. For diagnosis, ARG1 and APCS can be detected via ELISA, combined with AFP, to establish a highly sensitive early screening panel. Additionally, liquid biopsy-based detection of MBL2, LBP, and C9 expression further enhances the accuracy

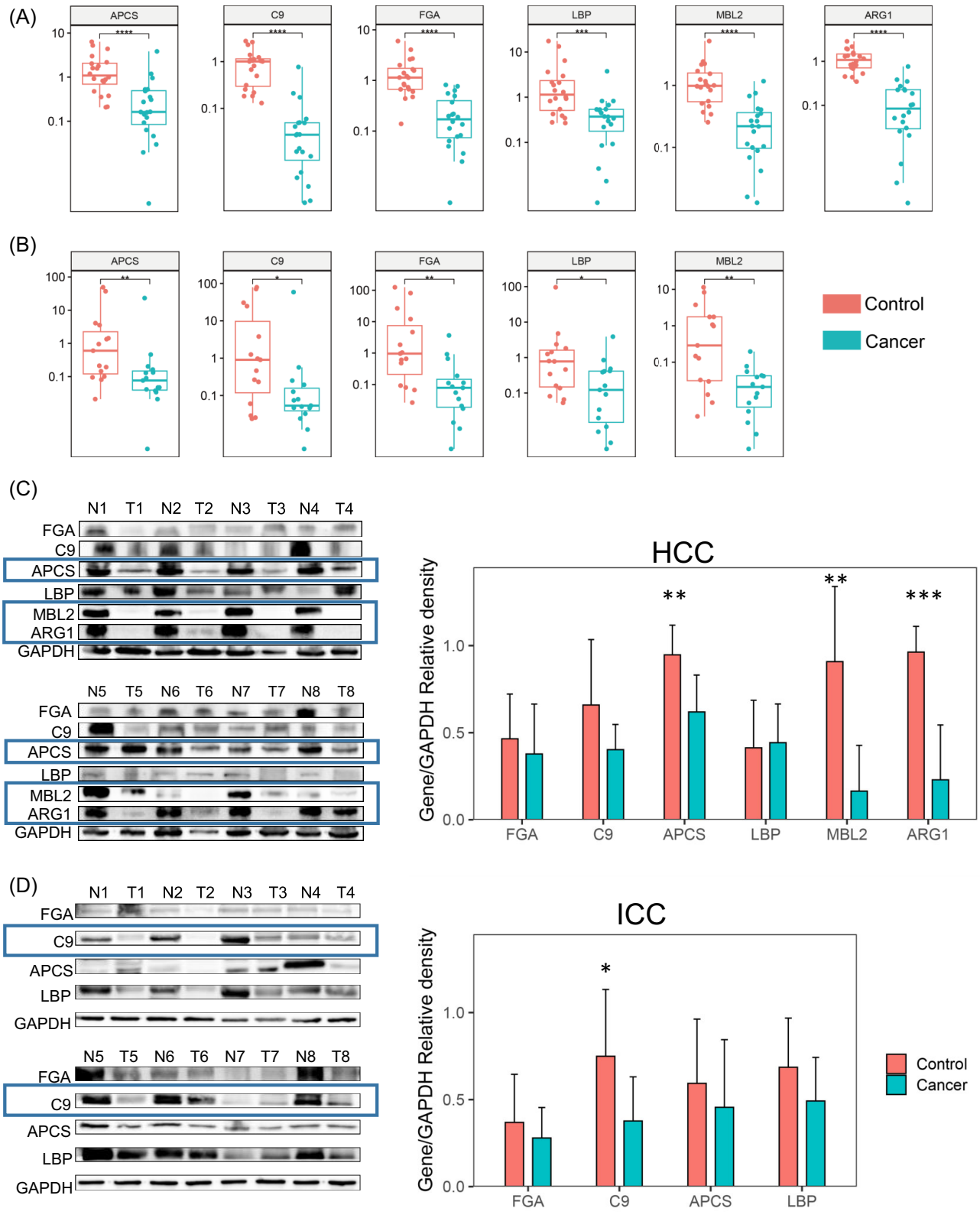

Figure 4. RNA and protein level validation. (**A**) qPCR analysis of immune-related gene expression in HCC tissues ($n = 20$). (**B**) qPCR analysis of immune-related gene expression in ICC tissues ($n = 15$). (**C**) Western blot analysis of six candidate proteins in HCC tissues, with GAPDH as the loading control. Left: representative western blot bands; right: quantification of protein density (normalized to GAPDH). (**D**) Western blot analysis of six candidate proteins in ICC tissues, as described for (C). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

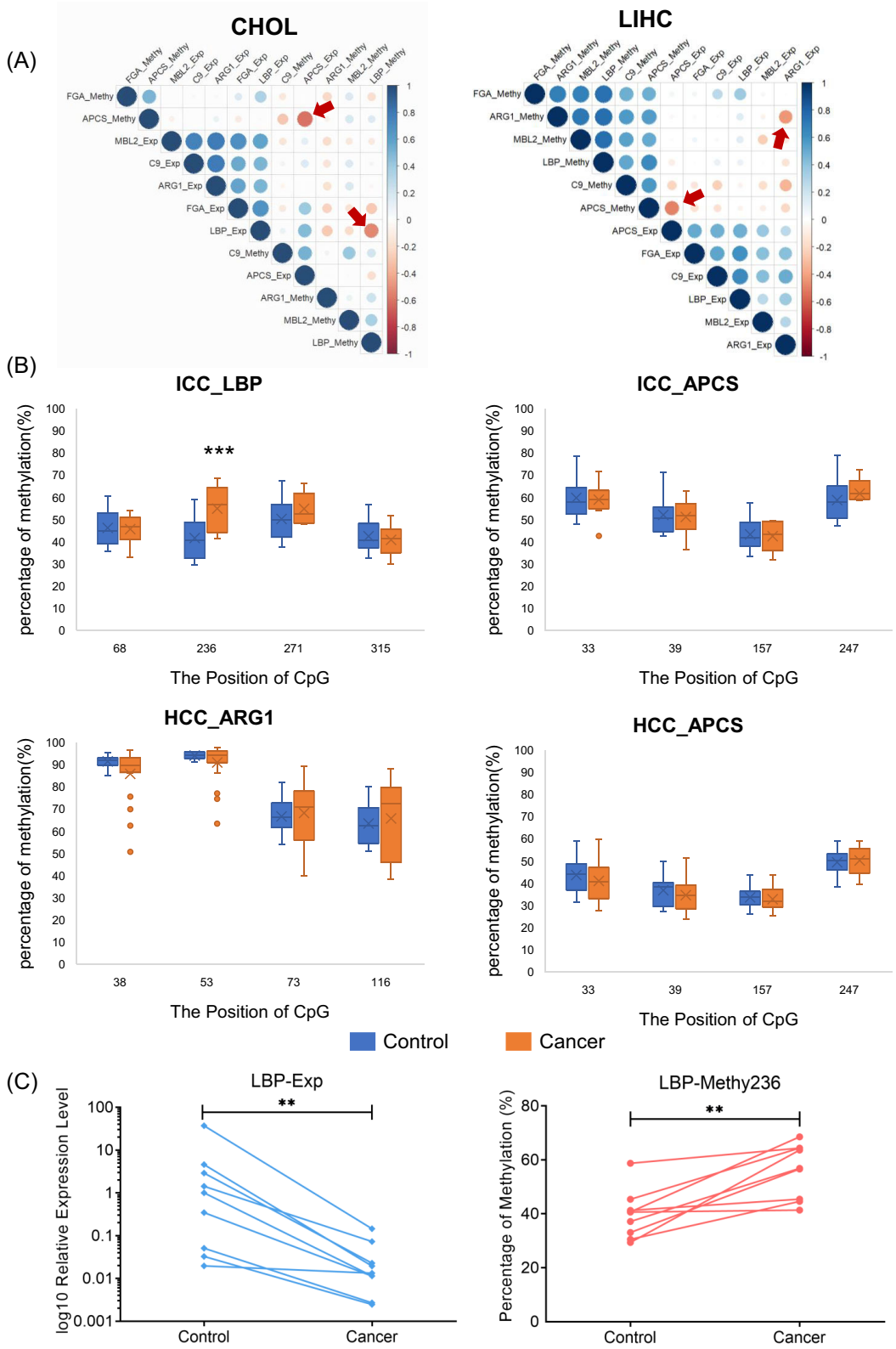

Figure 5. Methylation level of gene promoter regions. (**A**) Correlation analysis of the immune-related genes expression level and methylation level based on LIHC and CHOL samples from TCGA. Blue color indicates positive correlation, red color indicates negative correlation, and red arrows indicate genes with significant negative correlation between expression and methylation levels. (**B**) Methylation levels of APCS, LBP, and ARG1 promoter regions detected by BSAS ($n = 9$). (**C**) LBP RNA expression and methylation analysis in ICC tissues ($n = 9$). Left: RNA expression level of LBP; right: methylation level of LBP at position 236. $^{**}P < 0.01$, $^{***}P < 0.001$.

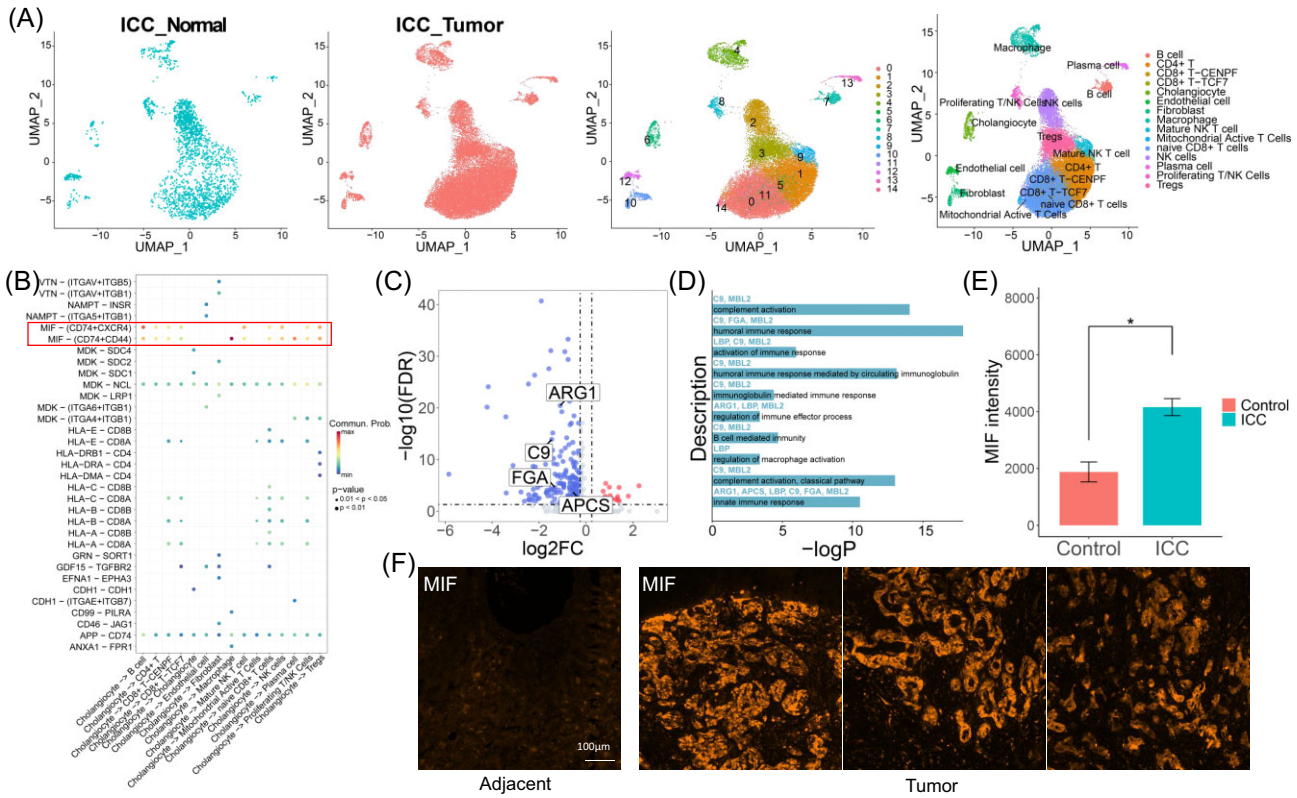

Figure 6. Single-cell analysis based on the GSE189903 dataset. (**A**) Single-cell analysis of ICC, including UMAP plots of tumor and adjacent normal tissues, unsupervised clustering cell grouping plot (Cluster 0–14), and cell-type annotation plot. (**B**) Bubble plot of cell–cell communication between cholangiocytes and other cell types in ICC. (**C**) Differentially expressed genes in cholangiocyte between normal and ICC tumor tissues (red, upregulated; blue, downregulated; |log2FC|>0.25, FDR < 0.05). (**D**) Functional enrichment analysis of differentially expressed genes. (**E**) Differences of MIF intensity between ICC cancer tissue and adjacent tissue. (**F**) Immunostaining of the ICC tissue TMA for MIF. *$P < 0.05$.

of non-invasive diagnostics. Therapeutically, small-molecule inhibitors of MIF, such as ISO-1 [43], or anti-MIF/CD74 antibodies could inhibit tumor-promoting interactions between tumor cells and macrophages, providing promising strategies for combination immunotherapy. For prognosis, low expression of MBL2 and FGA, associated with immune microenvironment dysregulation and fibrosis progression, may serve as risk-stratification markers. Future multicenter clinical studies are needed to validate the sensitivity and specificity of these potential biomarkers and to promote standardized detection methods and targeted drug development, ultimately advancing precision medicine for hepatobiliary cancers.

Conclusion

In summary, this study provides a comprehensive understanding of the specificity and commonality of six GI cancers in terms of gene expression, mutation, and methylation levels. Additionally, we identified a group of immune-related genes specifically expressed in hepatobiliary tumors, and experimentally validated their expression characteristics and methylation patterns in hepatobiliary tumor tissues. Furthermore, we explored cellular composition, immune-related gene expression, and intercellular communication in hepatobiliary cancers.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant Nos. 32370694 and 81972914), Shanghai Public Health Research Project (grant No. 2024GKM25), and the Innovation Group Project of Shanghai Municipal Health Commission (grant No. 2019CXJQ03).

Author contributions

Honglian Huang (Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing), Yueping Zhan (Validation), Hui Zong (Methodology), Chenjun Huang (Data curation), Fan Yang (Validation), Ziyi Wei (Data curation), Xin Qin (Data curation), M. James C. Crabbe (Writing – review & editing), Ying Wang (Funding acquisition, Methodology, Writing – review & editing), Xiaoyan Zhang (Funding acquisition, Writing – review & editing).

Supplementary data

Supplementary data is available at *PCMEDI Journal* online.

Conflict of interest

None declared.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Ethics Committee of the leading medical center (Shanghai Eastern Hepatobiliary Surgery Hospital, EHBHKY2020-02-012).

Availability of data and materials

Data sources and handling of the publicly available datasets used in this study are described in the Materials and Methods section. Further details and other data that supports the findings of this study are available from the corresponding authors upon request.

Reference

1 Ferlay J, Soerjomataram I, Dikshit R *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;**136**:E359–86. https://doi.org/10.1002/ijc.29210

2 Craig AJ, von Felden J, Garcia-Lezana T *et al.* Tumour evolution in hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol* 2019;**17**:139–52. https://doi.org/10.1038/s41575-019-0229-4

3 Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability—An evolving hallmark of cancer. *Nat Rev Mol Cell Biol* 2010;**11**:220–28. https://doi.org/10.1038/nrm2858

4 Tong Y, Gao H, Qi Q *et al.* High fat diet, gut microbiome and gastrointestinal cancer. *Theranostics* 2021;**11**:5889–910. https://doi.org/10.7150/thno.56157

5 Wang Y, Tong Y, Zhang Z *et al.* ViMIC: a database of human disease-related virus mutations, integration sites and cis-effects. *Nucleic Acids Res* 2022;**50**:D918–D27. https://doi.org/10.1093/nar/gkab779

6 Yang M, Lin W, Huang J *et al.* Novel immunotherapeutic approaches in gastric cancer. *Precision Clinical Medicine* 2024;**7**:pbae020. https://doi.org/10.1093/pcmedi/pbae020

7 Wooden B, Goossens N, Hoshida Y *et al.* Using big data to discover diagnostics and therapeutics for gastrointestinal and liver diseases. *Gastroenterology* 2017;**152**:53–67. https://doi.org/10.1053/j.gastro.2016.09.065

8 Zhang X, Li M, Ye S *et al.* CBD2: A functional biomarker database for colorectal cancer.. *Imeta* 2024;**3**:e155. https://doi.org/10.1002/imt2.155

9 Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* 1953;**6**:963–8.

10 Baba Y, Ishimoto T, Kurashige J *et al.* Epigenetic field cancerization in gastrointestinal cancers. *Cancer Lett* 2016;**375**:360–66. https://doi.org/10.1016/j.canlet.2016.03.009

11 Dakubo GD, Jakupciak JP, Birch-Machin MA *et al.* Clinical implications and utility of field cancerization. *Cancer Cell Int* 2007;**7**:2. https://doi.org/10.1186/1475-2867-7-2

12 Damania D, Roy HK, Subramanian H *et al.* Nanocytology of rectal colonocytes to assess risk of colon cancer based on field cancerization. *Cancer Res* 2012;**72**:2720–7. https://doi.org/10.1158/0008-5472.CAN-11-3807

13 Hyatt BJ, Paull PE, Wassef W. Gastric oncology: an update. *Curr Opin Gastroenterol* 2009;**25**:570–8. https://doi.org/10.1097/MOG.0b013e328331b5c9

14 Prevo LJ, Sanchez CA, Galipeau PC *et al.* p53-mutant clones and field effects in Barrett's esophagus. *Cancer Res* 1999;**59**:4784–7.

15 Jung Y, Lee S, Choi HS *et al.* Clinical validation of colorectal cancer biomarkers identified from bioinformatics analysis of public expression data. *Clin Cancer Res* 2011;**17**:700–9. https://doi.org/10.1158/1078-0432.CCR-10-1300

16 Bailey P, Chang DK, Nones K *et al.* Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* 2016;**531**:47–52. https://doi.org/10.1038/nature16965

17 Hoshida Y, Nijman SM, Kobayashi M *et al.* Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res* 2009;**69**:7385–92. https://doi.org/10.1158/0008-5472.CAN-09-1089

18 Jiang X, Wang J, Deng X *et al.* Role of the tumor microenvironment in PD-L1/PD-1-mediated tumor immune escape. *Mol Cancer* 2019;**18**:10. https://doi.org/10.1186/s12943-018-0928-4

19 Yamamoto K, Venida A, Yano J *et al.* Autophagy promotes immune evasion of pancreatic cancer by degrading MHC-I. *Nature* 2020;**581**:100–5. https://doi.org/10.1038/s41586-020-2229-5

20 Li X, Sun Z, Peng G *et al.* Single-cell RNA sequencing reveals a pro-invasive cancer-associated fibroblast subgroup associated with poor clinical outcomes in patients with gastric cancer. *Theranostics* 2022;**12**:620–38. https://doi.org/10.7150/thno.60540

21 Wang Y, Wang Y, Wang S *et al.* GIDB: a knowledge database for the automated curation and multidimensional analysis of molecular signatures in gastrointestinal cancer. *Database* 2019;**2019**:baz051. https://doi.org/10.1093/database/baz051

22 Wang Y, Zong H, Yang F *et al.* A knowledge empowered explainable gene ontology fingerprint approach to improve gene functional explication and prediction. *iScience* 2023;**26**:106356. https://doi.org/10.1016/j.isci.2023.106356

23 Hu C, Li T, Xu Y *et al.* CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. *Nucleic Acids Res* 2023;**51**:D870–6. https://doi.org/10.1093/nar/gkac947

24 Jiang S, Qian Q, Zhu T *et al.* Cell Taxonomy: a curated repository of cell types with multifaceted characterization. *Nucleic Acids Res* 2023;**51**:D853–D60. https://doi.org/10.1093/nar/gkac816

25 Yu G, Wang L-G, Han Y *et al.* clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 2012;**16**:284–7. https://doi.org/10.1089/omi.2011.0118

26 Zhou Y, Zhou B, Pache L *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019;**10**:1523. https://doi.org/10.1038/s41467-019-09234-6

27 Jin S, Guerrero-Juarez CF, Zhang L *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* 2021;**12**:1088. https://doi.org/10.1038/s41467-021-21246-9

28 Wang Y, Chen S, Xiao X *et al.* Impact of apolipoprotein A1 on tumor immune microenvironment, clinical prognosis and genomic landscape in hepatocellular carcinoma. *Precision Clinical Medicine* 2023;**6**:pbad021. https://doi.org/10.1093/pcmedi/pbad021

29 Wang Y, Han G, Wang K *et al.* Tumor-derived GM-CSF promotes inflammatory colon carcinogenesis via stimulating epithelial release of VEGF. *Cancer Res* 2014;**74**:716–26. https://doi.org/10.1158/0008-5472.Can-13-1459

30 Pan P, Oshima K, Huang YW *et al.* Loss of FFAR2 promotes colon cancer by epigenetic dysregulation of inflammation suppressors. *Int J Cancer* 2018;**143**:886–96. https://doi.org/10.1002/ijc.31366

31 Wang-Bishop L, Chen Z, Gomaa A *et al.* Inhibition of AURKA reduces proliferation and survival of gastrointestinal cancer cells with activated KRAS by preventing activation of RPS6KB1. *Gastroenterology* 2019;**156**:662–75. https://doi.org/10.1053/j.gastro.2018.10.030

32 Singhal A, Li BT, O'Reilly EM. Targeting KRAS in cancer. *Nat Med* 2024;**30**:969–83. https://doi.org/10.1038/s41591-024-02903-0

33 Arafeh R, Samuels Y. PIK3CA in cancer: the past 30 years. *Semin Cancer Biol* 2019;**59**:36–49. https://doi.org/10.1016/j.semcancer.2019.02.002

34 Zhao M, Mishra L, Deng C-X. The role of TGF-β/SMAD4 signaling in cancer. *International Journal of Biological Sciences* 2018;**14**:111–23. https://doi.org/10.7150/ijbs.23230

35 Mossmann D, Müller C, Park S *et al.* Arginine reprograms metabolism in liver cancer via RBM39. *Cell* 2023;**186**:5068–83. https://doi.org/10.1016/j.cell.2023.09.011

36 Chen L, Zhang C, Xue R *et al.* Deep whole-genome analysis of 494 hepatocellular carcinomas. *Nature* 2024;**627**:586–93. https://doi.org/10.1038/s41586-024-07054-3

37 Ferrín G, Ranchal I, Llamoza C *et al.* Identification of candidate biomarkers for hepatocellular carcinoma in plasma of HCV-infected cirrhotic patients by 2-D DIGE. *Liver Int* 2013;**34**:438–46. https://doi.org/10.1111/liv.12277

38 Liao H, Yang J, Xu Y *et al.* Mannose-binding lectin 2 as a potential therapeutic target for hepatocellular carcinoma: multi-omics analysis and experimental validation. *Cancers* 2023;**15**:4900. https://doi.org/10.3390/cancers15194900

39 Zhang Q, Shen X, Yuan X *et al.* Lipopolysaccharide binding protein resists hepatic oxidative stress by regulating lipid droplet homeostasis. *Nat Commun* 2024;**15**:3213. https://doi.org/10.1038/s41467-024-47553-5

40 Karin M. New insights into the pathogenesis and treatment of non-viral hepatocellular carcinoma: a balancing act between immunosuppression and immunosurveillance. *Precision Clinical Medicine* 2018;**1**:21–8. https://doi.org/10.1093/pcmedi/pby005

41 Li L, Yang H, Li Y *et al.* Hypoxia restrains the expression of complement component 9 in tumor-associated macrophages promoting non-small cell lung cancer progression. *Cell Death Discovery* 2018;**4**:63. https://doi.org/10.1038/s41420-018-0064-3

42 Noe JT, Mitchell RA. MIF-dependent control of tumor immunity. *Front Immunol* 2020;**11**:609948. https://doi.org/10.3389/fimmu.2020.609948

43 Al-Abed Y, VanPatten S. MIF as a disease target: ISO-1 as a proof-of-concept therapeutic. *Future Medicinal Chemistry* 2010;**3**:45–63. https://doi.org/10.4155/fmc.10.281

Received 17 April 2025; revised 17 June 2025; accepted 19 June 2025. published 24 June 2025