CANCER GENOMICS & PROTEOMICS *23*: 281-299 (2026)
doi: 10.21873/cgp.20577

Decoding the Mechanisms of Hepatocellular Carcinoma Cancer Stem Cells and Identifying Potential Therapeutic Strategies Based on Single-cell Omics

XIAOTIAN TAN[1*], QIAOXIAN LUO[1], YING GE[1], NING DENG[1], PING JIN[1], MENGHUAN SONG[1], CAROLINA OI LAM UNG[1,2,3], MEIWAN CHEN[1*] and HAO HU[1,2,3*]

[1]State Key Laboratory of Mechanism and Quality Research of Chinese Medicine, Institute of Chinese Medical Sciences, University of Macau, Taipa, Macau;
[2]Centre for Pharmaceutical Regulatory Sciences, University of Macau, Taipa, Macau;
[3]Department of Public Health and Medical Administration, Faculty of Health Sciences, University of Macau, Taipa, Macau

Abstract

Background/Aim: Cancer stem cells (CSCs) play key roles in hepatocellular carcinoma (HCC) initiation, progression, therapeutic resistance, and recurrence, yet their cellular and spatial heterogeneity remains poorly understood. This study aimed to systematically characterize HCC-associated CSCs and identify prognostic biomarkers and potential therapeutic strategies using single-cell omics.

Materials and Methods: Single-cell RNA sequencing and spatial transcriptomics data were obtained from HCCDB v2.0. Malignant cells were re-clustered using Harmony-based batch correction, followed by uniform manifold approximation and projection (UMAP) and Louvain clustering. Copy number variation analysis validated malignant identities. CSC-associated molecular features were characterized using differential expression, gene regulatory network analysis (SCENIC), pathway enrichment (GSVA), pseudotime trajectory inference (Monocle, CytoTRACE2), and cell-cell communication analysis (CellChat). CSC-specific genes were integrated with GEO survival datasets (GSE76427, GSE14520) to construct a prognostic model, and potential CSC-targeting compounds were predicted using Connectivity Map.

Results: Six malignant subpopulations were identified, including a progenitor-like CSC subset expressing EPCAM, SOX9, and SOX4. Spatial transcriptomics revealed CSC enrichment at the tumor-stroma interface. CSCs exhibited strong stemness, metabolic plasticity, and invasive potential, with activation of WNT/β-catenin, TGF-β, Notch, EMT, MYC, and mTORC1 pathways. Key transcription factors (TEAD2, SOX4, HNF1B, KLF7) were identified. A 12-gene CSC-derived

continued

[*]These Authors contributed equally to this work.

Hao Hu, State Key Laboratory of Mechanism and Quality Research of Chinese Medicine, Institute of Chinese Medical Sciences, University of Macau, Taipa, Macau. Tel: +853 88228538, e-mail: haohu@um.edu.mo

Received December 14, 2025 | Revised January 18, 2026 | Accepted February 3, 2026

signature stratified patients into distinct risk groups with significantly different overall survival. Several candidate compounds, including fluspirilene, genistein, and daunorubicin, showed potential CSC-suppressive activity.
Conclusion: This study provides a comprehensive single-cell-based atlas of CSCs in HCC, highlighting their spatial niches, regulatory programs, and clinical relevance. The identified prognostic signature and candidate drugs offer promising avenues for CSC-targeted therapies.

Keywords: Single-cell RNA sequencing, hepatocellular carcinoma, liver cancer, stem cell.

Introduction

Liver cancer is among the most prevalent and lethal malignancies worldwide. In 2022, an estimated 865,269 new cases were reported, ranking sixth in global cancer incidence, while 757,948 deaths ranked third in cancer-related mortality (1). At the current trajectory, the global incidence of liver cancer is projected to nearly double by 2050 (2). Primary liver cancer, which arises from liver parenchyma, is dominated by hepatocellular carcinoma (HCC) (3), whereas metastatic liver cancer results from secondary tumors disseminating from organs such as the colon, stomach, and pancreas (4-7). Chronic viral hepatitis (HBV and HCV), cirrhosis, long-term alcohol exposure, and metabolic disorders remain the major risk factors for HCC development (8). Although surgical resection, local ablative therapies, and systemic treatments have improved clinical outcomes to some extent, HCC remains highly heterogeneous, relapse-prone, and intrinsically drug-resistant, resulting in persistently poor overall survival (9, 10). Consequently, the discovery of novel therapeutic strategies or effective drug-repurposing approaches for HCC remains an urgent clinical priority (11).

Accumulating evidence suggests that cancer stem cells (CSCs) play a pivotal role in HCC initiation, progression, metastasis, and therapeutic resistance (12). CSCs possess stem-like properties–including self-renewal, multipotent differentiation, and pronounced drug resistance–that enable them to drive tumor growth and fuel recurrence even after clinically effective treatments (12-14). Elucidating the molecular signatures of HCC-associated CSCs, defining their prognostic relevance, and identifying CSC-specific therapeutic vulnerabilities are therefore critical steps toward improving clinical management and designing precision therapies (15).

Compared to traditional bulk RNA sequencing, single-cell RNA sequencing (scRNA-seq) provides high-resolution profiling of individual cells, enabling the dissection of cellular heterogeneity, identification of rare cell populations, and reconstruction of gene regulatory networks and lineage trajectories (16, 17). Recent advances integrating single-cell and spatial transcriptomics further allow the localization of distinct cell states within the tumor microenvironment (TME), offering deeper insights into tumor architecture, tumor-immune interactions, and spatially constrained regulatory programs (18-20). These approaches hold promise for discovering clinically relevant biomarkers and therapeutic targets in HCC.

Despite growing recognition of the importance of CSCs in HCC progression and treatment failure, current studies largely rely on single-modality datasets–most commonly scRNA-seq–and lack a comprehensive multi-omics integration. Crucial aspects such as the spatial niches of CSCs, their intercellular communication networks, transcriptional regulators, developmental hierarchies, and clinically actionable molecular determinants remain insufficiently defined. To address these gaps, the present study constructs a multimodal atlas of HCC CSCs, integrating single-cell and spatial transcriptomics, regulatory network inference, developmental trajectory reconstruction, clinical prognostic modeling, and drug-response prediction. This integrative framework provides a systematic characterization of CSC malignancy states, spatial organization, stemness-associated regulatory programs, prognostic biomarkers, and potential

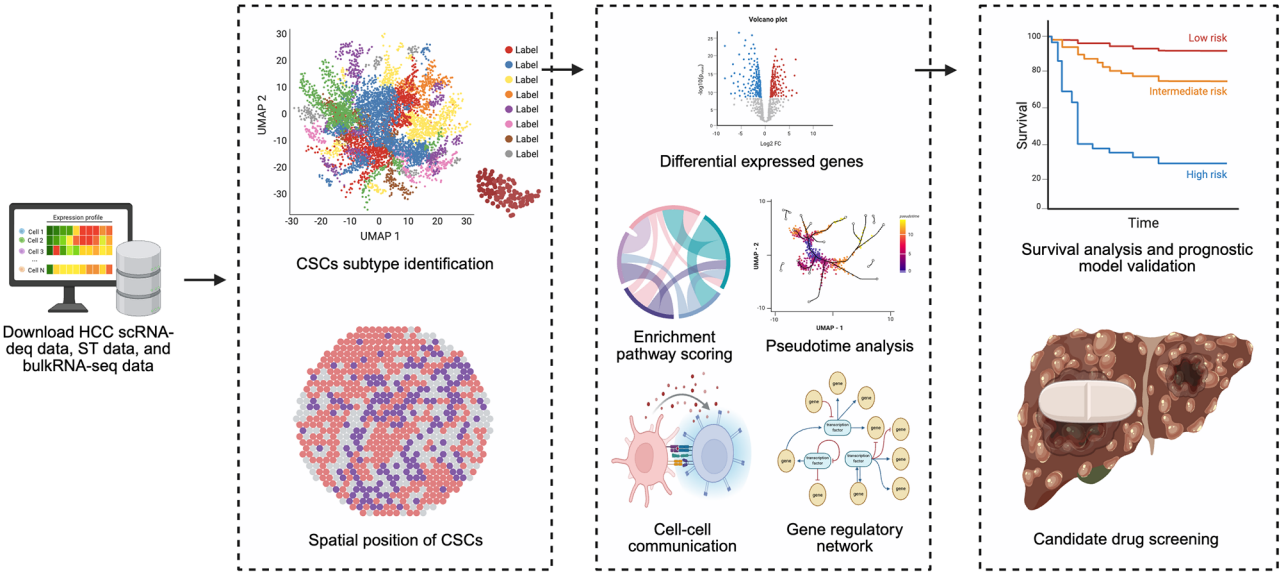

Figure 1. *Overview of the integrated multi-omics workflow, encompassing single-cell transcriptomic profiling, malignant cell subpopulation delineation, spatial mapping of CSC niches, regulatory network and developmental trajectory inference, construction of a CSC-derived prognostic model, and CMap-based therapeutic compound screening. Created with BioRender.com. CSCs: Cancer stem cells; CMap: connectivity map.*

therapeutic compounds with relevance to CSC-targeted treatment (Figure 1).

Materials and Methods

Single-cell sequencing data sources and preprocessing. Single-cell RNA sequencing (scRNA-seq) data were obtained from the HCCDBsc3 dataset in the Single-cell RNA-seq Datasets module of the HCCDB v2.0 database (21). The dataset comprises 61,428 cells derived from seven hepatocellular carcinoma (HCC) patients. Raw data underwent standardized quality control, during which cells were retained according to the following criteria: mitochondrial gene percentage ≤20%, number of detected features (nFeature) between 200 and 10,000, and total UMI counts (nCount) between 500 and 30,000. Data preprocessing was conducted using the Seurat (v4.3.0) package (22). Gene expression matrices were normalized, and highly variable genes were identified. Principal component analysis (PCA) was applied, and the top 30 principal components were selected for downstream dimensionality reduction (23). To mitigate inter-sample batch effects, canonical correlation analysis (CCA) was applied for data integration (24). Clustering was performed using a shared nearest neighbor (SNN) graph-based approach, and dimensionality reduction for visualization was achieved using the uniform manifold approximation and projection (UMAP) algorithm (25). Cell-type annotation was carried out by integrating canonical marker genes reported in the literature, resulting in the identification of 7 major cell types and 17 secondary subpopulations.

Extraction and re-clustering of malignant cell subsets. To investigate intratumoral heterogeneity in greater detail, we focused our analysis on tumor-derived malignant cell populations. Cells annotated as “Malignant cells” in the original dataset were extracted for downstream analyses. These cells were subsequently re-normalized and scaled using the Seurat (v5.3.0) pipeline (22), and principal component analysis (PCA) was re-performed based on highly variable genes. To further reduce inter-sample

variability, batch correction was applied in the principal component space using the Harmony algorithm (26). Nonlinear dimensionality reduction was then conducted using UMAP (25), followed by clustering with a shared nearest neighbor graph and Louvain community detection (27), which partitioned malignant cells into 7 transcriptionally distinct subpopulations. Marker genes for each malignant subpopulation were identified through differential expression analysis. Functional interpretation and biological annotation were performed by integrating evidence from the literature and curated databases, including CellMarker and PanglaoDB (28, 29). We found two clusters exhibiting highly similar marker gene expression profiles and consistent functional characteristics. We merged these two clusters for subsequent analyses to enhance biological interpretability and analytical robustness. Ultimately, six subpopulations were identified.

Cell proportion analysis. To assess tumor heterogeneity and inter-patient variability in CSC distribution, both relative and absolute proportions of the six malignant cell subpopulations were calculated for each sample. Relative proportions were defined as the fraction of cells in each subpopulation over the total number of cells per sample. These proportions were visualized using bar plots and stacked charts to compare the enrichment patterns of CSC subpopulations across patients.

Spatial transcriptomics (ST) analysis. ST data were obtained from HCC-1, HCC-3, and HCC-4 samples within the Spatial Transcriptomics module of the HCCDB v2.0 database (21). Spots corresponding to tumor tissue regions were selected for downstream analysis. The AddModuleScore algorithm (30) was applied to compute the average expression levels of each malignant cell cluster at the single-cell level, and these expression profiles were mapped onto the corresponding H&E-stained tissue images. By integrating the marker gene signatures of CSC subpopulations, their spatial localization and enrichment regions within tumor tissues were delineated. Spatial distribution patterns and the associated tumor microenvironmental features of CSCs were further analyzed in the context of histological morphology.

Copy number variation (CNV) analysis. To validate the classification of malignant cell subsets and distinguish true tumor cells from potential immune cell contamination, CNV analysis was performed using the inferCNV (v1.24.0) package (https://github.com/broadinstitute/inferCNV). Immune cells were used as a reference population, and gene copy number signals were estimated through a moving window smoothing algorithm. CNV heatmaps were generated using the ward.D2 hierarchical clustering method to visualize genomic copy number aberrations across subpopulations, facilitating the accurate identification of malignant cells.

Differentially expressed genes (DEGs) analysis. Differential gene expression among the six malignant cell subpopulations, as well as between CSC and non-CSC populations, was analyzed using the FindMarkers and FindAllMarkers functions from the Seurat package. Genes were considered differentially expressed if they met the following criteria: adjusted p-value <0.05, mean $\log_2$ fold change ($|\log_2 FC|$) >1.5, expression rate in the CSC population (pct.1) >0.4, and expression rate in the non-CSC population (pct.2) <0.2.

Enrichment pathway scoring. Gene set variation analysis (GSVA) was performed using the GSVA (v2.2.0) package in combination with GSEABase (v1.70.0) and msigdbr (v24.1.0) (https://www.gsea-msigdb.org/gsea/msigdb). The analysis was conducted on the six malignant cell subpopulations using the Hallmark gene sets (31). Enrichment scores were computed for 50 canonical HALLMARK pathways to reveal functional heterogeneity among CSC subpopulations, encompassing processes such as metabolism, cell cycle regulation, stemness maintenance, and immune evasion.

Pseudotime trajectory analysis. To investigate the differentiation potential and developmental trajectories of CSCs, pseudotime analysis was performed using Monocle2 (v2.36.0) and Monocle3 (v1.4.26) (32, 33). Gene dynamics models were constructed to visualize the expression patterns of key marker genes along the trajectories. Additionally, CytoTRACE 2 (v1.1.0) scoring was integrated

to evaluate the stemness and differentiation states of individual cells (34).

Cell-cell communication. Ligand-receptor interaction networks among the six malignant cell subpopulations were constructed using the CellChat package (v1.6.1) (35). Communication probability matrices were calculated to identify significant signaling pathways, revealing intercellular interaction patterns between CSCs and other tumor subpopulations. Particular attention was given to signaling axes associated with immune regulation, angiogenesis, and tumor progression.

Gene regulatory networks. Gene regulatory networks for each malignant cell subpopulation were inferred using the SCENIC framework (v1.3.1), incorporating GENIE3 and AUCell modules (36). Core transcriptional regulators and their downstream target gene sets were identified for each subpopulation, and regulatory activity scores (AUC values) were computed to assess the potential roles of CSC-specific transcription factors in maintaining stemness and driving tumor progression.

Building a prognostic model using GEO data. Differentially expressed genes between CSC and non-CSC subpopulations were integrated with transcriptomic and survival data from the GSE76427 and GSE14520 cohorts, encompassing 336 tumor samples (37, 38). Genes significantly associated with overall survival were first identified via univariate Cox regression. A multigene prognostic signature was then constructed using LASSO-Cox regression. Patients were stratified into high-risk and low-risk groups based on their calculated risk scores, and the predictive performance of the model was assessed using Kaplan-Meier survival analysis and time-dependent ROC curves (39, 40).

Drug screening by Connectivity Map (CMap). To identify potential therapeutics capable of reversing the molecular characteristics of CSCs, we leveraged the CMap database (41). Differentially expressed genes between CSC and non-CSC populations ($|\log_2 \text{FoldChange}| > 1.5$, adjusted p-value <0.05) were classified as upregulated or downregulated and submitted to the L1000 platform (42). The CMap algorithm computed a Connectivity Score (NCS) for each compound, reflecting the inverse correlation between the drug-induced transcriptional profile and the CSC molecular signature. Compounds with negative NCS values were considered as potential inhibitors or phenotype-reversing agents for CSCs. Mechanism-of-action (MOA) annotations were retrieved from the CMap database, and MOA enrichment analysis was performed to identify recurrent functional classes among the candidate drugs.

Ethics statement. The single-cell RNA sequencing and bulk RNA sequencing data used in this study were all from publicly available sources, therefore no ethical review was required.

Results

Metadata cell classification. Following data integration, a total of 61,428 cells were retained for downstream analysis. Using UMAP coordinates derived from the metadata, we constructed a two-dimensional visualization depicting seven major cell types: B cells, endothelial cells, malignant cells, myeloid cells, NK/T cells, plasma B cells, and stromal cells (Figure 2A). A higher-resolution UMAP further resolved 17 distinct cellular subtypes, including B cells, basophils, $CD4^+$ T cells, $CD8^+$ T cells, cDC1s, cDC2s, circulating NK/NKT cells, endothelial cells, macrophages, malignant cells, monocytes and monocyte-derived cells, neutrophils, plasma B cells, resident NK cells, stromal cells, and Tregs (Figure 2B).

Extraction of malignant cell subpopulations, followed by dimensionality reduction, clustering, and cell annotation. A total of 21,447 cells annotated as "Malignant cells" were extracted for in-depth analysis. Seurat objects were reconstructed for this subset, followed by matrix normalization, principal component analysis (PCA), and UMAP-based dimensionality reduction. After rigorous quality control, these malignant cells were clustered into six distinct subpopulations, which were manually annotated

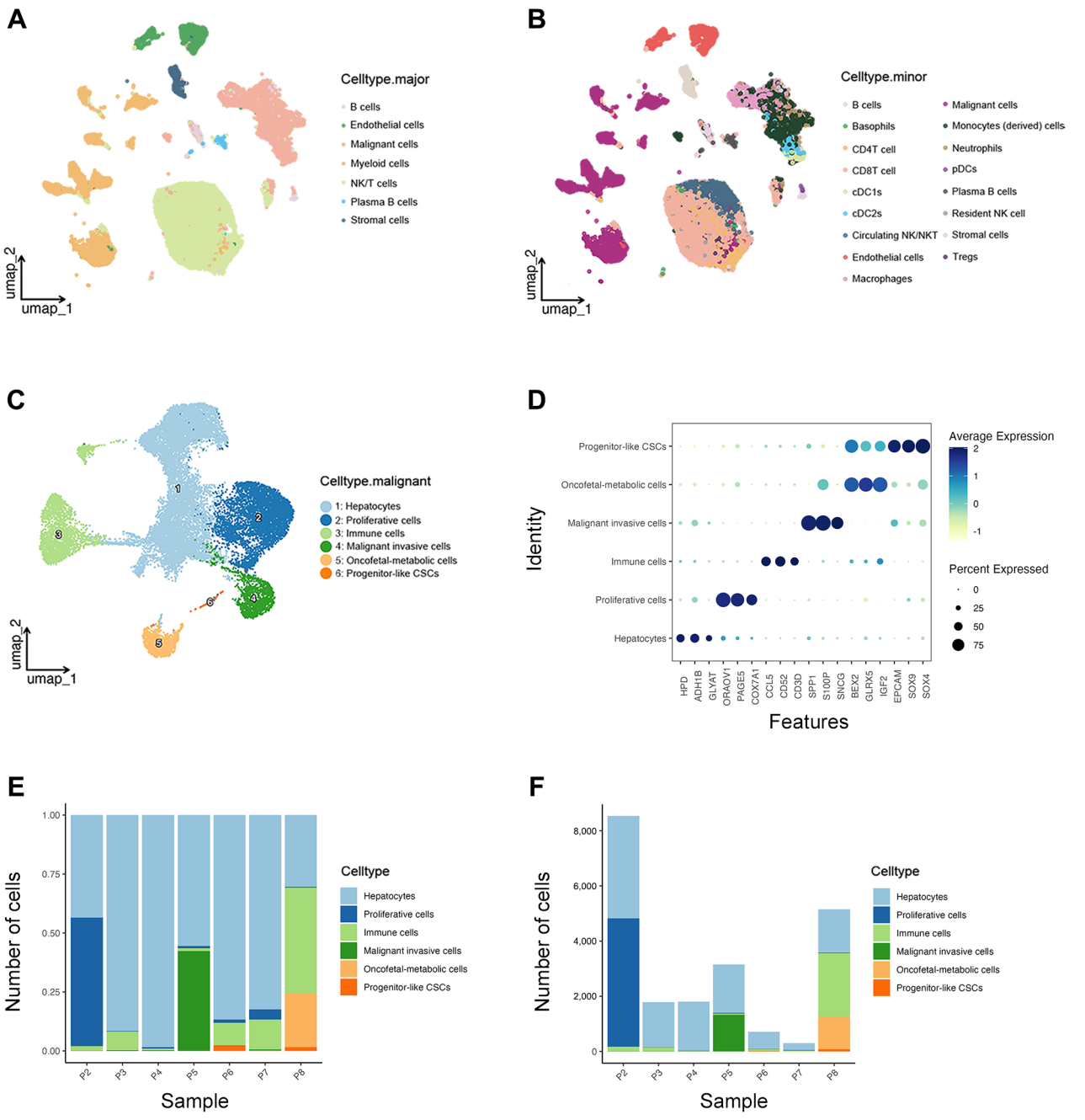

Figure 2. *Extraction of malignant cells and identification of CSC subsets. (A–B) UMAP visualization of 61,428 cells from seven HCC patients, classified into seven major lineages (A) and 17 finer subtypes (B) after stringent quality control (mitochondrial content <20%; 200–10,000 detected genes; 500–30,000 UMIs). Data were normalized using SCTransform, integrated via canonical correlation analysis (CCA, 30 PCs), and visualized with UMAP. (C) Re-clustering of 21,447 malignant cells following Harmony batch correction, using 15 principal components. (D) Marker-based annotation of six malignant states: hepatocytes, proliferative cells, immune-interacting cells, invasive malignant cells, oncofetal–metabolic cells, and progenitor-like CSCs. (E–F) Inter-patient heterogeneity of malignant cell states, showing both relative proportions (E) and absolute cell counts (F). Notably, CSCs were selectively enriched in specific patients, underscoring divergent malignant programs across HCC tumors. CSCs: Cancer stem cells; UMAP: uniform manifold approximation and projection; HCC: hepatocellular carcinoma.*

using differential gene expression profiles and literature references. The identified subpopulations include: Hepatocytes (HPD, ADH1B, GLYAT) (43-45), proliferative cells (ORAOV1, PAGE5, COX7A1) (46-48), Immune cells (CCL5, CD52, CD3D) (49-51), malignant invasive cells (SPP1, S100P, SNCG) (52-54), oncofetal-metabolic cells (BEX2, GLRX5, IGF2) (55-57), and progenitor-like CSCs (EPCAM, SOX9, SOX4) (Figure 2C-D) (58-60). Cell proportion analysis revealed that mature hepatocytes consistently comprised the largest fraction across samples, whereas proliferative cells were enriched in patient 2, and CSCs were particularly prominent in patients 6 and 8 (Figure 2E-F), highlighting the pronounced inter-patient heterogeneity of malignant cell subpopulations.

Spatial position of CSCs. Spatial transcriptomics analysis of tumor tissues from three HCC patients revealed pronounced spatial heterogeneity of CSC subsets within the tumor microenvironment (Figure 3A). CSCs were predominantly enriched at the tumor margins, whereas their presence was relatively sparse within the tumor core. This pattern was consistently observed across all three samples, suggesting a preferential localization of CSCs at the tumor–stromal interface, where microenvironmental cues–such as hypoxia, inflammatory signals, and immune-suppressive niches-may support stemness maintenance. Furthermore, spatial mapping of CSC marker genes (*e.g.*, SOX9, HES4, WNK2, ITGA2, DKK1, PLEKHB1) demonstrated markedly elevated expression at the tumor periphery, mirroring the distribution patterns identified in single-cell clustering. These findings indicate that CSCs exhibit defined histological niches and likely contribute to tumor invasion and recurrence.

Copy number variation. CNV analysis revealed pronounced genomic alterations within the malignant cell population (Figure 3B). Except for the infiltrating immune cells, which displayed relatively stable genomes, other malignant subpopulations exhibited heterogeneous CNV patterns with substantial intercellular variability. These findings corroborate and further validate the previously defined malignant subtypes, highlighting the genomic heterogeneity underlying HCC tumor cell populations.

DEGs and gene regulatory network. We performed differential gene expression analysis between CSC and non-CSC populations (Figure 3C). Notably, the oncofetal-metabolic and CSC populations exhibited overlapping transcriptional profiles, both expressing stemness- and metabolism-associated marker genes such as ACSL3, COX7B2, and CTAG2 (61-63). However, the CSC subset was distinguished by elevated expression of epithelial and stemness/migratory markers, including SOX4, SOX9, and EPCAM (58-60). In addition, the CSC population showed high expression of DKK1, WNK2, ITGA2, and HES4, indicative of oncofetal-like developmental programs, active Notch/Wnt signaling, and enhanced capabilities in microenvironmental interactions, ion homeostasis, and extracellular matrix remodeling (64-67). Collectively, these features define a CSC-like phenotype characterized by "undifferentiated, plastic, and migration/colonization potential."

To further delineate regulatory mechanisms, we constructed a CSC-specific gene regulatory network and identified key transcription factors (TFs) (Figure 3D). Several TFs were significantly upregulated or exhibited elevated regulatory activity in progenitor-like CSCs, including HNF1B, TEAD2, SOX4, and KLF7, which emerged as central regulatory nodes with high AUC scores. These TFs likely play pivotal roles in maintaining CSC stemness and self-renewal capacity. For instance, TEAD2 may drive CSC proliferation and resistance to apoptosis via the Hippo-YAP/TAZ pathway (68), and its elevated expression correlates with tumor progression, poor prognosis, enhanced epithelial-mesenchymal transition, and angiogenesis in HCC (69). SOX4 promotes oncogenic transformation and stem cell-like self-renewal by reprogramming transcriptional circuits and modulating tumor suppressor pathways such as p53 (70). KLF7 regulates proliferation, differentiation, and survival, with aberrant expression linked to tumor growth and metastasis in HCC (71). HNF1B controls intrahepatic bile duct and gallbladder differentiation, mediates liver metabolic

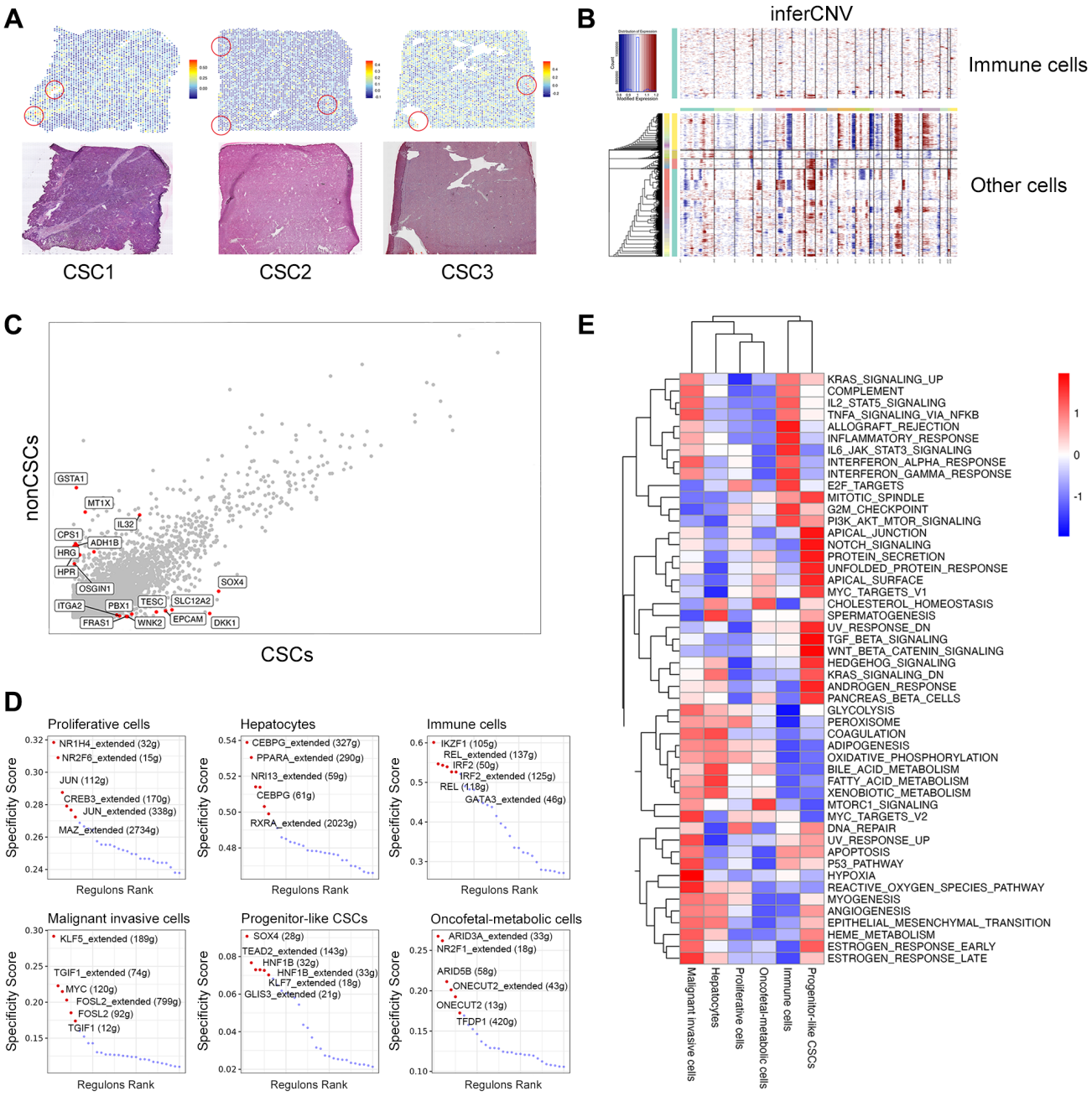

Figure 3. *Spatial location, differentially expressed genes, transcriptional identity, and pathway activity of CSCs. (A) Spatial transcriptomics of three HCC specimens demonstrates consistent enrichment of CSC signatures at the tumor–stroma interface. CSC module scores (AddModuleScore) were projected onto histological sections, highlighting preferential localization at invasive tumor margins. (B) inferCNV analysis, using immune cells as reference, reveals widespread chromosomal gains and losses across malignant populations, confirming their malignant identity and validating cluster assignments. (C) Differential gene expression analysis distinguishes CSCs (e.g., SOX4, EPCAM, DKK1) from non-CSC populations (e.g., HPR, ADH1B, IL32). (D) SCENIC-based regulatory network analysis identifies a core CSC transcriptional circuitry dominated by TEAD2, SOX4, HNF1B, and KLF7, each exhibiting elevated AUC activity specifically in progenitor-like CSCs. (E) GSVA Hallmark scoring reveals coordinated activation of WNT–β-catenin, TGF-β, Hedgehog, Notch, EMT, MYC, and mTORC1 pathways in CSCs, defining a transcriptionally active, epigenetically plastic, high-stemness malignant state. CSCs: Cancer stem cells; HCC: hepatocellular carcinoma; CNV: copy number variation; AUC: area under the curve; GSVA: gene set variation analysis.*

functions, and may contribute to maintaining CSC metabolic homeostasis and hepatocyte lineage characteristics (72).

Enrichment pathway scoring. Hallmark pathway-based enrichment analysis (Figure 3E) revealed that the progenitor-like CSC subpopulation displayed pronounced activation of multiple developmental and stemness-associated signaling pathways, including WNT-β-catenin, TGF-β, Hedgehog, Notch, Mitotic spindle, Apical junction, Protein secretion, EMT, Androgen response, mTORC1, MYC targets, and Cholesterol homeostasis. The concurrent activation of these pathways underscores the strong self-renewal, stemness maintenance, and proliferative capacity of this subpopulation, as well as its enhanced migratory and invasive potential through EMT-related gene programs. Additionally, upregulation of the Apical junction and Apical surface pathways reflects an active state in preserving epithelial polarity and cell-cell adhesion, which may facilitate interactions with the tumor microenvironment and the sensing of extracellular cues. The observed downregulation of the UV response pathway suggests enhanced stress tolerance, survival advantages, and resilience to genotoxic stress. Meanwhile, the relative suppression of KRAS signaling down indicates a phenotype characterized by low KRAS activity, high plasticity, and preserved stemness, potentially contributing to tumor recurrence and therapy resistance. Interestingly, the enrichment of pathways related to pancreatic β-cell function implies that CSCs may adopt metabolic, secretory, and energy-sensing features reminiscent of endocrine cells, further supporting their functional versatility. Collectively, these findings portray CSCs as a transcriptionally active, highly plastic population, with coordinated regulation of stemness, metabolic adaptation, and microenvironmental responsiveness, providing a mechanistic basis for their pivotal role in HCC progression, relapse, and treatment refractoriness.

Cell development trajectory. To explore the developmental hierarchy of malignant cells, we reconstructed single-cell trajectories using Monocle2. Genes associated with cell progression were selected, and dimensionality reduction was performed using the DDRTree algorithm, allowing cells to be ordered along a pseudo-temporal trajectory (Figure 4A-B). Our analysis revealed that subsets of tumor-infiltrating immune cells, CSCs, and oncofetal-metabolic cells may represent the earliest states in malignant cell development. We further performed gene clustering based on pseudo-temporal expression dynamics to identify groups of genes with similar temporal trends (Figure 4C). Notably, marker genes of early-stage immune infiltrating cells, CSCs, and oncofetal-metabolic cells exhibited high expression at the beginning of the trajectory, supporting their role as progenitors in tumor progression. Additionally, differentiation state maps generated using CytoTRACE 2 (Figure 4D) and Monocle3 (Figure 4E) corroborated these findings, confirming the hierarchical organization and stemness potential of these cell populations.

Cell chat. To investigate intercellular signaling, we calculated ligand-receptor interaction probabilities among six malignant cell subtypes and constructed a communication network. The heatmap (Figure 5A) illustrates the maximum efferent and afferent signals for each population. Notably, CSCs exhibit strong outgoing and incoming signals within the MDK, AGT, and IGF pathways. NMF-based pattern clustering (73) (Figure 5B) confirmed that CSCs act as the dominant signal emitters in these pathways, motivating a focused analysis of their ligand–receptor pairs. Using CSCs as signal emitters and all six subtypes (including CSCs) as receivers, ligand–receptor analysis revealed significant self-renewal and proliferation-related signaling (Figure 5C). For example, MDK promotes tumor cell proliferation, survival, invasion, and angiogenesis (Figure 5D), and its interaction with NCL (Figure 5G) may directly drive CSC self-renewal and proliferation. Activation of the MDK-SDC2 axis (Figure 5H) may enhance CSC self-renewal, proliferation, and migration via PI3K/AKT, FAK-SRC, and MAPK pathways. As a core extracellular matrix signaling integrator, SDC2 may further maintain CSC stemness and malignant potential by mediating MDK-dependent angiogenesis and

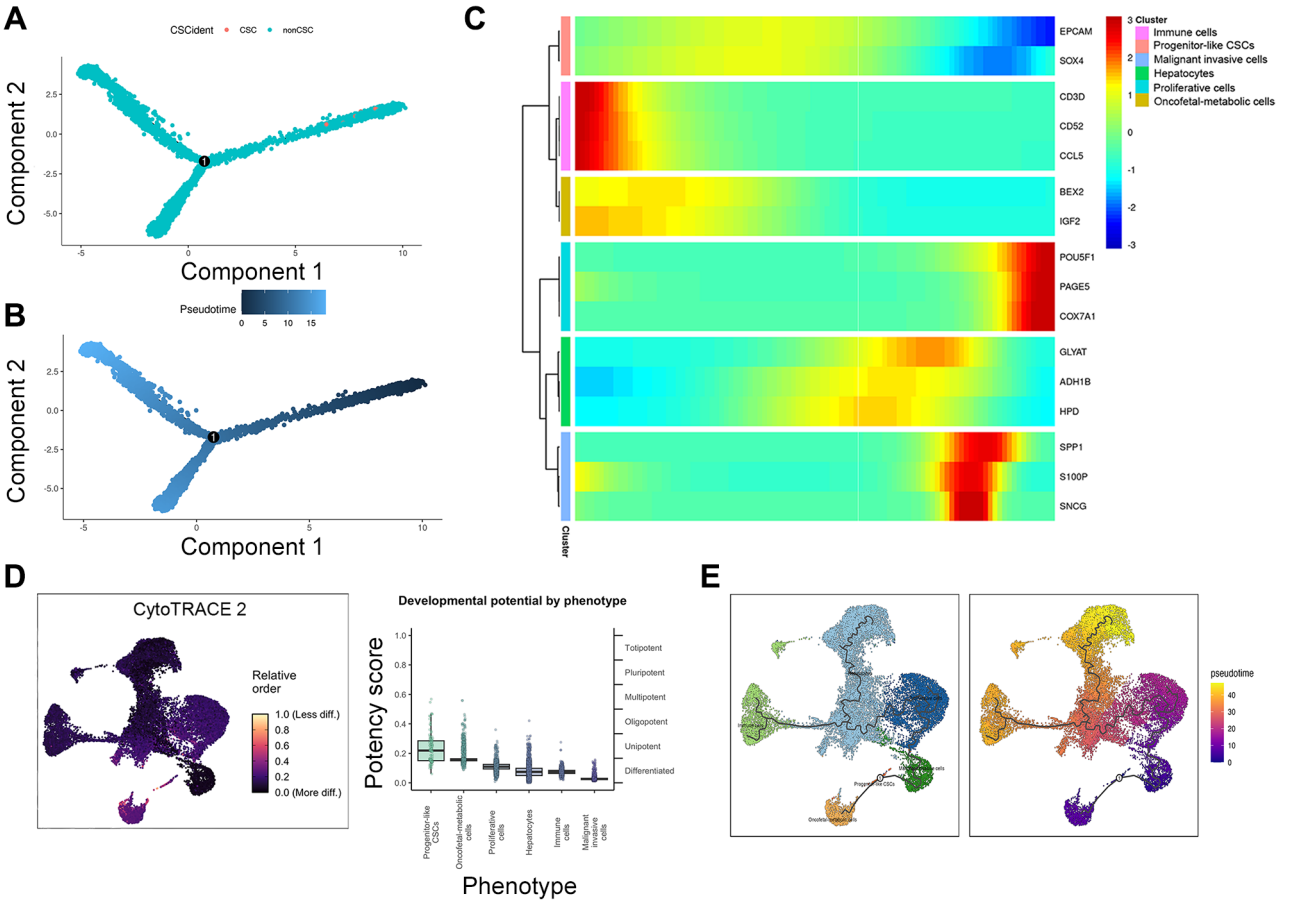

Figure 4. *Developmental trajectory of CSCs in malignant cells. (A–B) Monocle2 DDRTree reconstruction reveals a branched developmental continuum among malignant cells, with CSCs predominantly occupying early pseudotime states, indicative of potential origins or transitional intermediates. (C) Pseudotime-resolved heatmap illustrates coordinated early expression of CSC and oncofetal–metabolic gene signatures, supporting their inferred developmental positioning. (D) CytoTRACE2 scoring independently confirms high developmental potential in CSCs and oncofetal–metabolic cells, reinforcing their placement at the apex of stem/progenitor hierarchy. (E) Monocle3 pseudotime analysis of CSC-related developmental genes visualized on UMAP highlights dynamic gene expression along the trajectory. CSCs: Cancer stem cells; HCC: hepatocellular carcinoma; UMAP: uniform manifold approximation and projection.*

immune microenvironment remodeling (74). AGT (Figure 5E and I) plays a central role in the renin-angiotensin system (RAS), suggesting that CSCs may exploit RAS signaling to regulate oxidative stress and energy supply, thereby maintaining stemness (75). IGF2R mutations or loss of heterozygosity weaken its negative regulatory effect on IGF signaling, resulting in sustained IGF pathway activation (Figure 5F) (76). This provides a molecular basis for the activation of the IGF2-IGF2R axis within CSCs. Our analysis revealed enhanced IGF2-IGF2R signaling between progenitor-like CSCs and oncofetal-metabolic cells: IGF2 secreted by CSCs activates IGF2R in neighboring metabolically active cells, promoting their metabolic activity and niche remodeling; conversely, oncofetal-metabolic cells feed back to CSCs through IGF2-IGF2R, reinforcing PI3K/AKT/mTOR signaling and stemness gene expression78. This bidirectional communication establishes a stable "stemness-metabolic interaction module" (Figure 5J). Overall, these signaling axes reveal key intercellular communication networks within the CSC population in the tumor microenvironment, providing a mechanistic basis for potential targeted interventions.

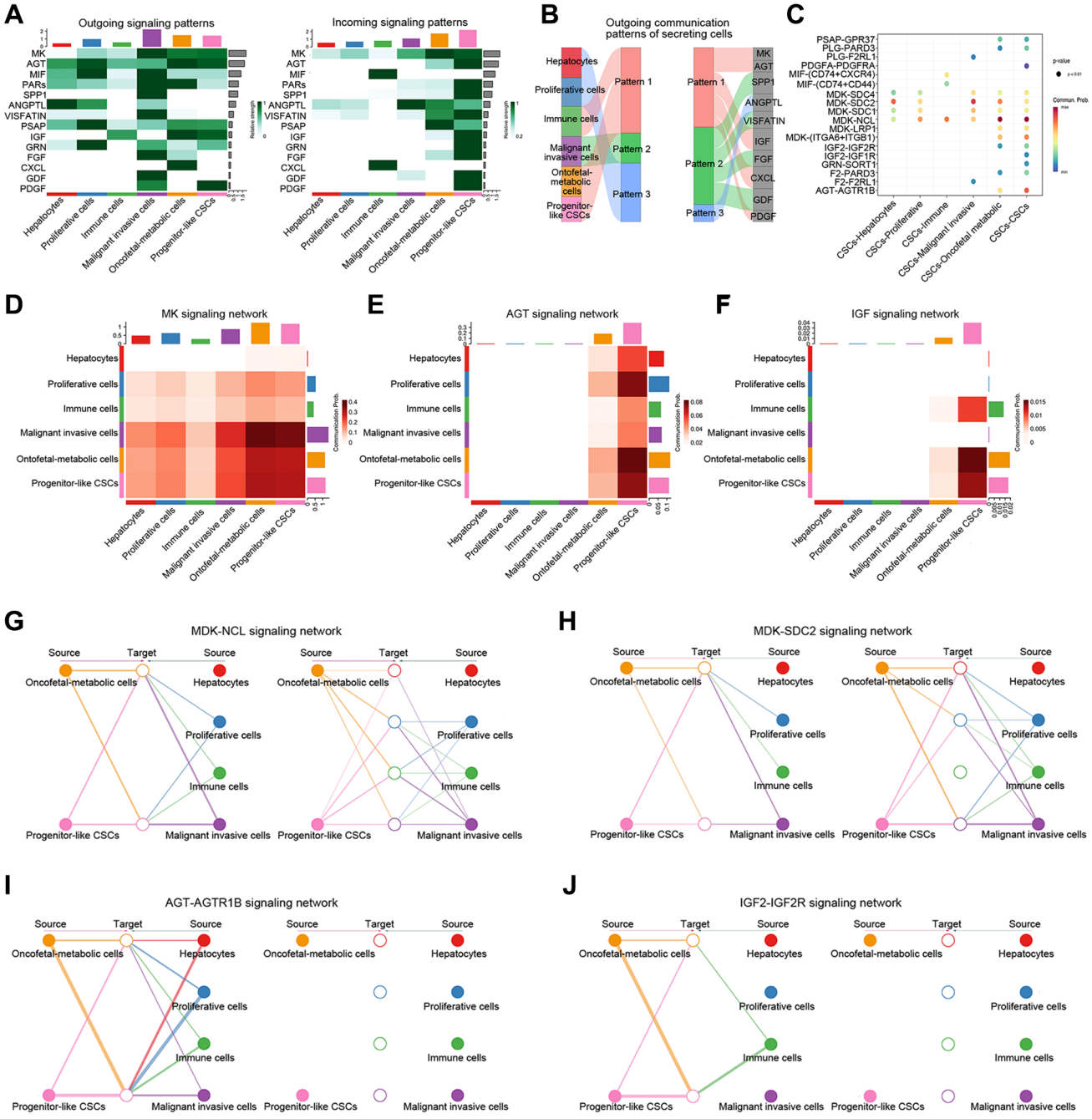

Figure 5. *CSC-centered communication networks reveal dominant MDK, AGT, and IGF signaling axes. (A) Network centrality analysis identifies the signaling pathways contributing most to efferent and afferent communication across six cell subtypes. (B) NMF reveals distinct cell communication patterns, with decomposition into three major modules using the selectK approach. (C) CSCs act as ligand-emitting populations, while the remaining subtypes serve as signal receivers, illustrating the differential ligand–receptor axis intensities. (D–F) Heatmaps display the MK, AGT, and IGF signaling pathways, with rows representing signal-emitting cells and columns representing signal-receiving cells. Color intensity reflects signaling strength, while the side and top bars indicate cumulative efferent and afferent signal intensity, respectively. (G–J) Visualization of individual ligand–receptor pairs mediating intercellular communication. The extractEnrichedLR function was used to identify key ligand–receptor interactions and associated signaling genes for each pathway, highlighting the pairs contributing most to signal transduction. CSCs: Cancer stem cells; NMF: nonnegative matrix factorization.*

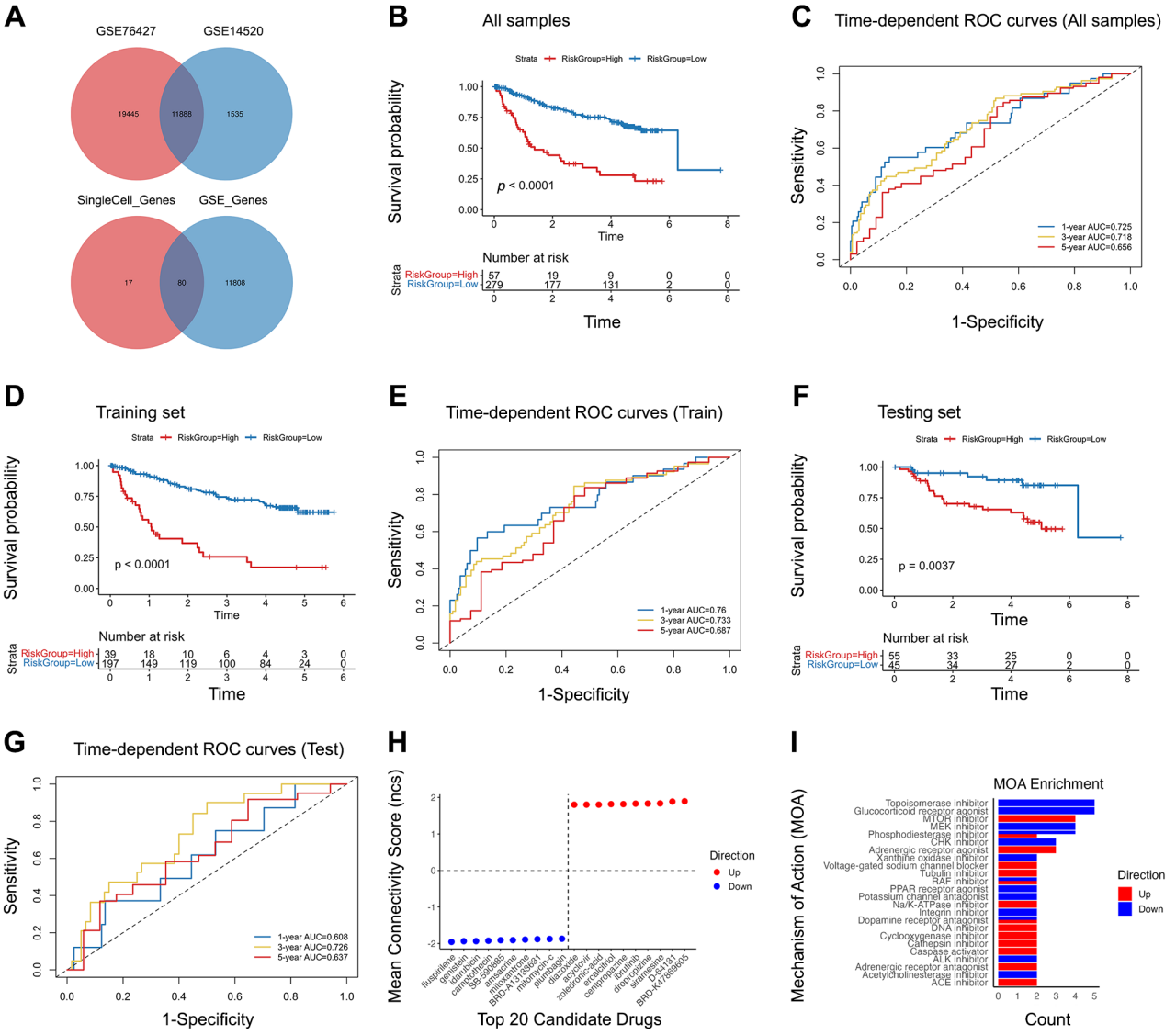

Figure 6. *CSC-derived prognostic signature and CMap-based identification of CSC-targeting compounds. (A) Integrative workflow (Venn diagram) combining CSC-specific differentially expressed genes with GEO survival datasets, yielding 80 overlapping genes across 336 evaluable HCC patients. (B–G) A 12-gene CSC-derived signature, constructed via LASSO–Cox regression, robustly stratifies patients into high- and low-risk groups across training, test, and full cohorts. Time-dependent ROC curves confirm strong predictive performance. (H) CMap-based reversal analysis identifies compounds exhibiting strong negative connectivity to CSC transcriptional programs–including fluspirilene, genistein, daunorubicin, and camptothecin–suggesting potential CSC-suppressive activity. (I) MOA enrichment reveals convergence on pathways critical to CSC biology, including topoisomerase inhibition, mTOR blockade, MEK inhibition, CHK inhibition, and phosphodiesterase inhibition. CSCs: Cancer stem cells; CMap: connectivity map; HCC: hepatocellular carcinoma; MOA: mechanism-of-action.*

Clinical prognostic model construction and drug screening. To evaluate the clinical relevance of CSC transcriptional signatures, we integrated differentially expressed genes from single-cell analyses with survival and clinical data from the GSE14520 and GSE76427 cohorts, encompassing 336 HCC samples and 80 candidate genes (Figure 6A). The cohort was randomly split into training and test sets at a 70:30 ratio. LASSO-Cox regression analysis on the training

set identified a 12-gene signature (ALCAM, C8orf4, DKK1, GSTP1, IVNS1ABP, MARCKSL1, NR2F1, PLEKHA1, RALGPS2, SOX4, TLE4, UTRN) significantly associated with overall survival. These genes are implicated in CSC-related biological processes, including cell adhesion, metabolic reprogramming, stemness maintenance, and signaling regulation. Kaplan-Meier survival analyses across all samples, training, and test sets demonstrated that patients stratified into the high-risk group exhibited significantly poorer overall survival than those in the low-risk group (Figure 6B-G). Time-dependent ROC curves further confirmed the stability and predictive performance of the model, indicating that this 12-gene signature effectively captures CSC transcriptional features in clinical HCC samples, reflecting stem-like and invasive malignant states.

To identify candidate therapeutics capable of targeting CSC phenotypes, up- and downregulated CSC gene sets were queried against the CMap database. Compounds exhibiting a significant negative correlation with the CSC transcriptional signature, including fluspirilene, genistein, daunorubicin, camptothecin, and amastatin, were predicted to potentially suppress CSC activity, whereas positively correlated drugs (*e.g.*, ibuprofen, diazoxide) may reinforce CSC-like states (Figure 6H). Mechanism-of-action (MOA) enrichment analysis revealed that candidate compounds predominantly act through topoisomerase inhibition, mTOR inhibition, MEK inhibition, phosphodiesterase inhibition, and CHK inhibition (Figure 6I). These pathways converge on core CSC signaling networks, including PI3K/AKT/mTOR, MAPK, DNA damage repair, and cell cycle regulation, suggesting that the identified drugs may exert anti-CSC effects by disrupting metabolic activity, signaling maintenance, and genome stability. Collectively, the CMap-based analysis highlights a repertoire of candidate compounds for targeting CSC-driven malignancy in HCC, offering a framework for potential therapeutic intervention.

Discussion

This study provides a systematic characterization of the molecular features, spatial distribution, and potential roles of CSCs in HCC within the tumor microenvironment, offering new insights into HCC initiation, progression, and therapy resistance. By integrating single-cell transcriptomics, spatial transcriptomics, and GEO clinical data, we delineated the tissue localization and stemness-maintaining mechanisms of CSC subsets and proposed potential targeted therapeutic strategies.

First, the marked enrichment of CSCs at the tumor periphery suggests a preference for specific microenvironmental niches, including hypoxia, inflammatory cues, and immunosuppressive signals. This spatial heterogeneity may underlie the high recurrence and metastasis rates observed in HCC. CSCs concurrently express epithelial markers (EPCAM, SOX9, SOX4) and stemness-associated genes (DKK1, HES4, WNK2), along with enhanced metabolic and secretory functions, indicating a highly plastic and migratory phenotype. These observations are consistent with previous descriptions of tumor stem cells as "undifferentiated, plastic, and invasive" (12, 13).

Second, gene regulatory network analysis identified key transcription factors, including TEAD2, SOX4, HNF1B, and KLF7, with elevated regulatory activity in CSCs, underscoring their roles in sustaining stemness, proliferation, and signaling plasticity. These factors engage in tumor-promoting and anti-apoptotic programs through Hippo-YAP/TAZ, Wnt/β-catenin, and p53-associated pathways, further highlighting CSCs as central drivers of malignant progression in HCC (77-79).

Third, pseudotime trajectory analysis and CytoTRACE scoring positioned CSCs at the apex of malignant cell differentiation, reflecting high stemness and differentiation potential. Intercellular communication analysis further revealed that CSCs and oncofetal-metabolic cells form a homeostatic interaction module via the IGF2-IGF2R axis, MDK, and AGT signaling, coordinating stemness maintenance and metabolic reprogramming. These findings emphasize the critical regulatory influence of the tumor microenvironment on CSC survival and function.

Moreover, a 12-gene prognostic model derived from CSC-specific differentially expressed genes and GEO

survival data effectively stratified patients into high- and low-risk groups, with high-risk patients exhibiting significantly reduced overall survival. This model validates the clinical relevance of CSC transcriptional features and may serve as a tool for risk stratification and personalized treatment in HCC.

Finally, CMap-based drug screening identified several small molecules, including fluspirilene, genistein, and daunorubicin, with the potential to reverse the CSC transcriptional program by targeting PI3K/AKT/mTOR, MAPK, and DNA repair pathways. These results provide a rational starting point for the development of CSC-directed therapeutic strategies.

Despite these advances, this study has limitations. First, our findings are primarily computational and derived from public datasets, lacking experimental validation in vitro or in vivo. Second, the coverage of spatial transcriptomics data remains limited, restricting a full understanding of dynamic interactions between CSCs and immune or stromal compartments. Future work integrating multi-omics approaches, such as proteomics (80), in vivo models, and longitudinal clinical samples, will be essential to validate CSC functions and assess the therapeutic efficacy of CSC-targeting agents.

Conclusion

In summary, we constructed a comprehensive multi-modal atlas of CSCs in hepatocellular carcinoma by integrating single-cell and spatial transcriptomics, regulatory network inference, developmental trajectory analysis, clinical prognostic modeling, and drug-response prediction. Our findings reveal that CSCs reside in a spatially conserved niche at the tumor–stroma interface, exhibit high transcriptional plasticity, and engage in self-reinforcing signaling circuits–particularly via MDK, AGT, and IGF axes–that collectively sustain stemness, metabolic reprogramming, and malignant progression. The CSC-derived 12-gene prognostic signature effectively stratifies patient survival, highlighting the translational relevance of CSC molecular features. Moreover, CMap analysis identified several candidate compounds capable of reversing CSC-associated transcriptional programs, providing a rational framework for the development of CSC-targeted therapies. Overall, this study advances our understanding of CSC-driven heterogeneity, recurrence, and treatment resistance in HCC, and offers actionable molecular insights for precision oncology.

Data Availability

The raw data of scRNA-seq and ST results can be downloaded from HCCDB v2.0 (http://lifeome.net:809/#/home). The bulkRNA-seq data can be downloaded from https://www.ncbi.nlm.nih.gov/geo/

Conflicts of Interest

The Authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' Contributions

Hao Hu: Conceptualization, Resources, Supervision, Project administration, Funding acquisition, Writing – Review & Editing. Meiwan Chen: Conceptualization, Resources, Supervision, Project administration, Funding acquisition, Writing – Review & Editing. Xiaotian Tan: Investigation, Formal analysis, Methodology, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project administration. Qiaoxian Luo: Writing – Review & Editing, Project administration, Data Curation, Formal analysis, Validation. Ying Ge: Writing – Review & Editing, Methodology, Software. Ping Jin: Project administration, Writing – Review & Editing, Methodology. Carolina Oi Lam Ung: Conceptualization, Resources, Supervision, Project administration, Writing – Review & Editing. Menghuan Song: Supervision, Project administration, Writing – Review & Editing, Validation. Ning Deng: Visualization, Methodology, Writing – Review & Editing.

Funding

This research is supported by the funding of the University of Macau (MYRGGRG2023-00059-ICMS) and the Science and Technology Development Fund, Macau SAR (File No. 0014/2023/AKP).

Acknowledgements

We are grateful for the kind support from the Pizza Group at the University of Macau. We thank the developers and maintainers of BioRender for providing the illustration platform used to create the figures in this study. We also acknowledge the Seurat R package and its contributors for offering powerful tools for single-cell RNA-seq data processing and analysis. We would also like to thank Peking Union Medical College Hospital and Tsinghua University for building the HCCDB v2.0 dataset. Their efforts greatly facilitated the completion of this work.

Artificial Intelligence (AI) Disclosure

During the preparation of this manuscript, a large language model (ChatGPT-5, OpenAI) was used solely for language editing and stylistic improvements in select paragraphs. No sections involving the generation, analysis, or interpretation of research data were produced by generative AI. All scientific content was created and verified by the authors. Furthermore, no figures or visual data were generated or modified using generative AI or machine learning-based image enhancement tools.

References

1 Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A: Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 74(3): 229-263, 2024. DOI: 10.3322/caac.21834

2 Chan SL, Sun HC, Xu Y, Zeng H, El-Serag HB, Lee JM, Schwartz ME, Finn RS, Seong J, Wang XW, Paradis V, Abou-Alfa GK, Rimassa L, Kao JH, Zhang BH, Llovet JM, Bruix J, Yip TC, Wong VW, Wong GL, Chan LL, Liu MQ, Gao Q, Shen F, Kelley RK, Cheng AL, Kurosaki M, Toyoda H, Chen WX, Murakami T, Liang P, Zucman-Rossi J, Minami Y, Miyayama S, Wang K, Man K, Hasegawa K, Li Q, Tsuchiya K, Xu L, Chew V, Chow P, Hoshida Y, Lujambio A, Ng IO, Sakamoto M, Park YN, Yau T, Kudo M, Fan J, Zhou J: The Lancet Commission on addressing the global hepatocellular carcinoma burden: comprehensive strategies from prevention to treatment. Lancet 406(10504): 731-778, 2025. DOI: 10.1016/s0140-6736(25)01042-6

3 Sia D, Villanueva A, Friedman SL, Llovet JM: Liver cancer cell of origin, molecular class, and effects on patient prognosis. Gastroenterology 152(4): 745-761, 2017. DOI: 10.1053/j.gastro.2016.11.048

4 Gupta GP, Massagué J: Cancer metastasis: building a framework. Cell 127(4): 679-695, 2006. DOI: 10.1016/j.cell.2006.11.001

5 Bertocchi A, Carloni S, Ravenda PS, Bertalot G, Spadoni I, Lo Cascio A, Gandini S, Lizier M, Braga D, Asnicar F, Segata N, Klaver C, Brescia P, Rossi E, Anselmo A, Guglietta S, Maroli A, Spaggiari P, Tarazona N, Cervantes A, Marsoni S, Lazzari L, Jodice MG, Luise C, Erreni M, Pece S, Di Fiore PP, Viale G, Spinelli A, Pozzi C, Penna G, Rescigno M: Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver. Cancer Cell 39(5): 708-724.e11, 2021. DOI: 10.1016/j.ccell.2021.03.004

6 Riihimäki M, Hemminki A, Sundquist K, Sundquist J, Hemminki K: Metastatic spread in patients with gastric cancer. Oncotarget 7(32): 52307-52316, 2016. DOI: 10.18632/oncotarget.10740

7 Zambirinis CP, Midya A, Chakraborty J, Chou JF, Zheng J, McIntyre CA, Koszalka MA, Wang T, Do RK, Balachandran VP, Drebin JA, Kingham TP, D'Angelica MI, Allen PJ, Gönen M, Simpson AL, Jarnagin WR: Recurrence after resection of pancreatic cancer: can radiomics predict patients at greatest risk of liver metastasis? Ann Surg Oncol 29(8): 4962-4974, 2022. DOI: 10.1245/s10434-022-11579-0

8 Chidambaranathan-Reghupaty S, Fisher PB, Sarkar D: Hepatocellular carcinoma (HCC): Epidemiology, etiology and molecular classification. Adv Cancer Res 149: 1-61, 2021. DOI: 10.1016/bs.acr.2020.10.001

9 Labrie M, Brugge JS, Mills GB, Zervantonakis IK: Therapy resistance: opportunities created by adaptive responses to targeted therapies in cancer. Nat Rev Cancer 22(6): 323-339, 2022. DOI: 10.1038/s41568-022-00454-5

10 Chen C, Wang Z, Ding Y, Qin Y: Tumor microenvironment-mediated immune evasion in hepatocellular carcinoma. Front Immunol 14: 1133308, 2023. DOI: 10.3389/fimmu.2023.1133308

11 Chen X, Shi J, Xue Y, Lai Y, Song M, Ung COL, Hu H: Critical analysis of the liver cancer policies and programs in China: implications for international liver cancer control. Glob Health Res Policy 10(1): 50, 2025. DOI: 10.1186/s41256-025-00450-w

12 Quiroz Reyes AG, Lozano Sepulveda SA, Martinez-Acuña N, Islas JF, Gonzalez PD, Heredia Torres TG, Perez JR, Garza Treviño EN:

Cancer stem cell and hepatic stellate cells in hepatocellular carcinoma. Technol Cancer Res Treat 22: 15330338231163677, 2023. DOI: 10.1177/15330338231163677

13 Zhou L, He L, Liu C, Qiu H, Zheng L, Sample KM, Wu Q, Li J, Xie K, Ampuero J, Li Z, Lv D, Liu M, Romero-Gómez M, Hu Y, Tang H: Liver cancer stem cell dissemination and metastasis: uncovering the role of NRCAM in hepatocellular carcinoma. J Exp Clin Cancer Res 42(1): 311, 2023. DOI: 10.1186/s13046-023-02893-w

14 Wang Z, Yan N, Sheng H, Xiao Y, Sun J, Cao C: Single-cell transcriptomic analysis reveals an immunosuppressive network between POSTN CAFs and ACKR1 ECs in TKI-resistant lung cancer. Cancer Genomics Proteomics 21(1): 65-78, 2024. DOI: 10.21873/cgp.20430

15 Balaji N, Kukal S, Bhat A, Pradhan N, Minocha S, Kumar S: A quartet of cancer stem cell niches in hepatocellular carcinoma. Cytokine Growth Factor Rev 79: 39-51, 2024. DOI: 10.1016/j.cytogfr.2024.08.007

16 Hwang B, Lee JH, Bang D: Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med 50(8): 1-14, 2018. DOI: 10.1038/s12276-018-0071-8

17 Molla Desta G, Birhanu AG: Advancements in single-cell RNA sequencing and spatial transcriptomics: transforming biomedical research. Acta Biochim Pol 72: 13922, 2025. DOI: 10.3389/abp.2025.13922

18 Tang X, Huang Y, Lei J, Luo H, Zhu X: The single-cell sequencing: new developments and medical applications. Cell Biosci 9: 53, 2019. DOI: 10.1186/s13578-019-0314-y

19 Lu Y, Yang A, Quan C, Pan Y, Zhang H, Li Y, Gao C, Lu H, Wang X, Cao P, Chen H, Lu S, Zhou G: A single-cell atlas of the multicellular ecosystem of primary and metastatic hepatocellular carcinoma. Nat Commun 13(1): 4594, 2022. DOI: 10.1038/s41467-022-32283-3

20 Sun Y, Wu L, Zhong Y, Zhou K, Hou Y, Wang Z, Zhang Z, Xie J, Wang C, Chen D, Huang Y, Wei X, Shi Y, Zhao Z, Li Y, Guo Z, Yu Q, Xu L, Volpe G, Qiu S, Zhou J, Ward C, Sun H, Yin Y, Xu X, Wang X, Esteban MA, Yang H, Wang J, Dean M, Zhang Y, Liu S, Yang X, Fan J: Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. Cell 184(2): 404-421.e16, 2021. DOI: 10.1016/j.cell.2020.11.041

21 Jiang Z, Wu Y, Miao Y, Deng K, Yang F, Xu S, Wang Y, You R, Zhang L, Fan Y, Guo W, Lian Q, Chen L, Zhang X, Zheng Y, Gu J: HCCDB v2.0: Decompose expression variations by single-cell RNA-seq and spatial transcriptomics in HCC. Genomics Proteomics Bioinformatics 22(1): qzae011, 2024. DOI: 10.1093/gpbjnl/qzae011

22 Butler A, Hoffman P, Smibert P, Papalexi E, Satija R: Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36(5): 411-420, 2018. DOI: 10.1038/nbt.4096

23 Maćkiewicz A, Ratajczak W: Principal components analysis (PCA). Computers & Geosciences 19(3): 303-342, 1993. DOI: 10.1016/0098-3004(93)90090-R

24 Knapp TR: Canonical correlation analysis: A general parametric significance-testing system. Psychological Bulletin 85(2): 410-416, 1978. DOI: 10.1037/0033-2909.85.2.410

25 McInnes L, Healy J, Saul N, Großberger L: UMAP: Uniform manifold approximation and projection. Journal of Open Source Software 3(29): 861, 2018. DOI: 10.21105/joss.00861

26 Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh PR, Raychaudhuri S: Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16(12): 1289-1296, 2019. DOI: 10.1038/s41592-019-0619-0

27 Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E: Fast unfolding of communities in large networks. J Stat Mech 2008(10): P10008, 2008. DOI: 10.1088/1742-5468/2008/10/P10008

28 Hu C, Li T, Xu Y, Zhang X, Li F, Bai J, Chen J, Jiang W, Yang K, Ou Q, Li X, Wang P, Zhang Y: 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 51(D1): D870-D876, 2023. DOI: 10.1093/nar/gkac947

29 Franzén O, Gan LM, Björkegren JLM: PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. Database (Oxford) 2019: baz046, 2019. DOI: 10.1093/database/baz046

30 Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, Fallahi-Sichani M, Dutton-Regester K, Lin JR, Cohen O, Shah P, Lu D, Genshaft AS, Hughes TK, Ziegler CG, Kazer SW, Gaillard A, Kolb KE, Villani AC, Johannessen CM, Andreev AY, Van Allen EM, Bertagnolli M, Sorger PK, Sullivan RJ, Flaherty KT, Frederick DT, Jané-Valbuena J, Yoon CH, Rozenblatt-Rosen O, Shalek AK, Regev A, Garraway LA: Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 352(6282): 189-196, 2016. DOI: 10.1126/science.aad0501

31 Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P: The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 1(6): 417-425, 2015. DOI: 10.1016/j.cels.2015.12.004

32 Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL: The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 32(4): 381-386, 2014. DOI: 10.1038/nbt.2859

33 Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C: Reversed graph embedding resolves complex single-cell trajectories. Nat Methods 14(10): 979-982, 2017. DOI: 10.1038/nmeth.4402

34 Kang M, Gulati GS, Brown EL, Qi Z, Avagyan S, Armenteros JJA, Gleyzer R, Zhang W, Steen CB, D'Silva JP, Schwab J, Clarke MF, Chaudhuri AA, Newman AM: Improved reconstruction of single-cell developmental potential with CytoTRACE 2. Nat Methods 22(11): 2258-2263, 2025. DOI: 10.1038/s41592-025-02857-2

35 Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, Myung P, Plikus MV, Nie Q: Inference and analysis of cell-cell communication using CellChat. Nat Commun 12(1): 1088, 2021. DOI: 10.1038/s41467-021-21246-9
36 Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J, van den Oord J, Atak ZK, Wouters J, Aerts S: SCENIC: single-cell regulatory network inference and clustering. Nat Methods 14(11): 1083-1086, 2017. DOI: 10.1038/nmeth.4463
37 Roessler S, Jia HL, Budhu A, Forgues M, Ye QH, Lee JS, Thorgeirsson SS, Sun Z, Tang ZY, Qin LX, Wang XW: A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. Cancer Res 70(24): 10202-10212, 2010. DOI: 10.1158/0008-5472.Can-10-2607
38 Grinchuk OV, Yenamandra SP, Iyer R, Singh M, Lee HK, Lim KH, Chow PK, Kuznetsov VA: Tumor-adjacent tissue co-expression profile analysis reveals pro-oncogenic ribosomal gene signature for prognosis of resectable hepatocellular carcinoma. Mol Oncol 12(1): 89-113, 2018. DOI: 10.1002/1878-0261.12153
39 Friedman JH, Hastie T, Tibshirani R: Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software 33(1): 1-22, 2010. DOI: 10.18637/jss.v033.i01
40 Blanche P, Dartigues JF, Jacqmin-Gadda H: Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med 32(30): 5381-5397, 2013. DOI: 10.1002/sim.5958
41 Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR: The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. Science 313(5795): 1929-1935, 2006. DOI: 10.1126/science.1132939
42 Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, Donahue M, Julian B, Khan M, Wadden D, Smith IC, Lam D, Liberzon A, Toder C, Bagul M, Orzechowski M, Enache OM, Piccioni F, Johnson SA, Lyons NJ, Berger AH, Shamji AF, Brooks AN, Vrcic A, Flynn C, Rosains J, Takeda DY, Hu R, Davison D, Lamb J, Ardlie K, Hogstrom L, Greenside P, Gray NS, Clemons PA, Silver S, Wu X, Zhao WN, Read-Button W, Wu X, Haggarty SJ, Ronco LV, Boehm JS, Schreiber SL, Doench JG, Bittker JA, Root DE, Wong B, Golub TR: A Next generation connectivity map: L1000 platform and the first 1,000,000 profiles. Cell 171(6): 1437-1452.e17, 2017. DOI: 10.1016/j.cell.2017.10.049
43 Gu P, Yang Q, Chen B, Bie YN, Liu W, Tian Y, Luo H, Xu T, Liang C, Ye X, Liu Y, Tang X, Gu W: Genetically blocking HPD via CRISPR-Cas9 protects against lethal liver injury in a pig model of tyrosinemia type I. Mol Ther Methods Clin Dev 21: 530-547, 2021. DOI: 10.1016/j.omtm.2021.04.002
44 Jiang F, Zhou S, Xia C, Lu J, Wang B, Wang X, Shen J, Ding W, Yin M, Dai F, Fu S: Downregulation of GLYAT correlates with tumour progression and poor prognosis in hepatocellular carcinoma. J Cell Mol Med 28(21): e70197, 2024. DOI: 10.1111/jcmm.70197
45 Watson BR, Paul B, Rahman RU, Amir-Zilberstein L, Segerstolpe Å, Epstein ET, Murphy S, Geistlinger L, Lee T, Shih A, Deguine J, Xavier RJ, Moffitt JR, Mullen AC: Spatial transcriptomics of healthy and fibrotic human liver at single-cell resolution. Nat Commun 16(1): 319, 2025. DOI: 10.1038/s41467-024-55325-4
46 Ha SY, Yeo SY, Lee KW, Kim SH: Validation of ORAOV1 as a new treatment target in hepatocellular carcinoma. J Cancer Res Clin Oncol 147(2): 423-433, 2021. DOI: 10.1007/s00432-020-03437-x
47 Nylund C, Rappu P, Pakula E, Heino A, Laato L, Elo LL, Vihinen P, Pyrhönen S, Owen GR, Larjava H, Kallajoki M, Heino J: Melanoma-associated cancer-testis antigen 16 (CT16) regulates the expression of apoptotic and antiapoptotic genes and promotes cell survival. PLoS One 7(9): e45382, 2012. DOI: 10.1371/journal.pone.0045382
48 West MD, Labat I, Sternberg H, Larocca D, Nasonkin I, Chapman KB, Singh R, Makarev E, Aliper A, Kazennov A, Alekseenko A, Shuvalov N, Cheskidova E, Alekseev A, Artemov A, Putin E, Mamoshina P, Pryanichnikov N, Larocca J, Copeland K, Izumchenko E, Korzinkin M, Zhavoronkov A: Use of deep neural network ensembles to identify embryonic-fetal transition markers: repression of COX7A1 in embryonic and cancer cells. Oncotarget 9(8): 7796-7811, 2017. DOI: 10.18632/oncotarget.23748
49 Tan W, Zhang W, Strasner A, Grivennikov S, Cheng JQ, Hoffman RM, Karin M: Tumour-infiltrating regulatory T cells stimulate mammary cancer metastasis through RANKL-RANK signalling. Nature 470(7335): 548-553, 2011. DOI: 10.1038/nature09707
50 Watanabe T, Masuyama J, Sohma Y, Inazawa H, Horie K, Kojima K, Uemura Y, Aoki Y, Kaga S, Minota S, Tanaka T, Yamaguchi Y, Kobayashi T, Serizawa I: CD52 is a novel costimulatory molecule for induction of CD4+ regulatory T cells. Clin Immunol 120(3): 247-259, 2006. DOI: 10.1016/j.clim.2006.05.006
51 Wei Z, Shen Y, Zhou C, Cao Y, Deng H, Shen Z: CD3D: a prognostic biomarker associated with immune infiltration and immunotherapeutic response in head and neck squamous cell carcinoma. Bioengineered 13(5): 13784-13800, 2022. DOI: 10.1080/21655979.2022.2084254
52 Wang Y, Liu X, Zhang H, Sun L, Zhou Y, Jin H, Zhang H, Zhang H, Liu J, Guo H, Nie Y, Wu K, Fan D, Zhang H, Liu L: Hypoxia-inducible lncRNA-AK058003 promotes gastric cancer metastasis by targeting γ-synuclein. Neoplasia 16(12): 1094-1106, 2014. DOI: 10.1016/j.neo.2014.10.008
53 Aishima S, Fujita N, Mano Y, Kubo Y, Tanaka Y, Taketomi A, Shirabe K, Maehara Y, Oda Y: Different roles of S100P overexpression in intrahepatic cholangiocarcinoma: carcinogenesis of perihilar type and aggressive behavior of

peripheral type. Am J Surg Pathol 35(4): 590-598, 2011. DOI: 10.1097/PAS.0b013e31820ffdf1
54 Shang S, Plymoth A, Ge S, Feng Z, Rosen HR, Sangrajrang S, Hainaut P, Marrero JA, Beretta L: Identification of osteopontin as a novel marker for early hepatocellular carcinoma. Hepatology 55(2): 483-490, 2012. DOI: 10.1002/hep.24703
55 Fukushi D, Shibuya-Takahashi R, Mochizuki M, Fujimori H, Kogure T, Sugai T, Iwai W, Wakui Y, Abue M, Murakami K, Nakamura Y, Yasuda J, Yamaguchi K, Sugamura K, Shibata C, Katayose Y, Satoh K, Tamai K: BEX2 is required for maintaining dormant cancer stem cell in hepatocellular carcinoma. Cancer Sci 112(11): 4580-4592, 2021. DOI: 10.1111/cas.15115
56 Mollbrink A, Jawad R, Vlamis-Gardikas A, Edenvik P, Isaksson B, Danielsson O, Stål P, Fernandes A: Expression of thioredoxins and glutaredoxins in human hepatocellular carcinoma: correlation to cell proliferation, tumor size and metabolic syndrome. Int J Immunopathol Pharmacol 27(2): 169-183, 2014. DOI: 10.1177/039463201402700204
57 Martinez-Quetglas I, Pinyol R, Dauch D, Torrecilla S, Tovar V, Moeini A, Alsinet C, Portela A, Rodriguez-Carunchio L, Solé M, Lujambio A, Villanueva A, Thung S, Esteller M, Zender L, Llovet JM: IGF2 is up-regulated by epigenetic mechanisms in hepatocellular carcinomas and is an actionable oncogene product in experimental models. Gastroenterology 151(6): 1192-1205, 2016. DOI: 10.1053/j.gastro.2016.09.001
58 Yang Q, Liang Y, Shi Y, Shang J, Huang X: The ALKBH5/SOX4 axis promotes liver cancer stem cell properties via activating the SHH signaling pathway. J Cancer Res Clin Oncol 149(17): 15499-15510, 2023. DOI: 10.1007/s00432-023-05309-6
59 Liu C, Liu L, Chen X, Cheng J, Zhang H, Shen J, Shan J, Xu Y, Yang Z, Lai M, Qian C: Sox9 regulates self-renewal and tumorigenicity by promoting symmetrical cell division of cancer stem cells in hepatocellular carcinoma. Hepatology 64(1): 117-129, 2016. DOI: 10.1002/hep.28509
60 Yamashita T, Ji J, Budhu A, Forgues M, Yang W, Wang HY, Jia H, Ye Q, Qin LX, Wauthier E, Reid LM, Minato H, Honda M, Kaneko S, Tang ZY, Wang XW: EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features. Gastroenterology 136(3): 1012-1024, 2009. DOI: 10.1053/j.gastro.2008.12.004
61 Wang Z, Embaye KS, Yang Q, Qin L, Zhang C, Liu L, Zhan X, Zhang F, Wang X, Qin S: A novel metabolism-related signature as a candidate prognostic biomarker for hepatocellular carcinoma. J Hepatocell Carcinoma Volume 8: 119-132, 2021. DOI: 10.2147/jhc.S294108
62 Wang X, Zhou X, Liu J, Liu Z, Zhang L, Gong Y, Huang J, Yu L, Wang Q, Yang C, Liao X, Yu T, Han C, Zhu G, Ye X, Peng T: Genome wide investigation of the clinical implications and molecular mechanism of long noncoding RNA LINC00668 and protein coding genes in hepatocellular carcinoma. Int J Oncol 55(4): 860-878, 2019. DOI: 10.3892/ijo.2019.4858
63 Huang L, Xu R, Chen S, Lin C, Li W, Li S, Saw PE, Zhang L, Xu X: Modulating lipid metabolism by nanoparticles (NPs)-mediated ACSL3 silencing to inhibit hepatocellular carcinoma growth and metastasis. Mol Cancer 24(1): 73, 2025. DOI: 10.1186/s12943-025-02274-1
64 Shen Q, Fan J, Yang XR, Tan Y, Zhao W, Xu Y, Wang N, Niu Y, Wu Z, Zhou J, Qiu SJ, Shi YH, Yu B, Tang N, Chu W, Wang M, Wu J, Zhang Z, Yang S, Gu J, Wang H, Qin W: Serum DKK1 as a protein biomarker for the diagnosis of hepatocellular carcinoma: a large-scale, multicentre study. Lancet Oncol 13(8): 817-826, 2012. DOI: 10.1016/s1470-2045(12)70233-4
65 Serysheva E, Berhane H, Grumolato L, Demir K, Balmer S, Bodak M, Boutros M, Aaronson S, Mlodzik M, Jenny A: Wnk kinases are positive regulators of canonical Wnt/β-catenin signalling. EMBO Rep 14(8): 718-725, 2013. DOI: 10.1038/embor.2013.88
66 Förster C, Mäkela S, Wärri A, Kietz S, Becker D, Hultenby K, Warner M, Gustafsson JA: Involvement of estrogen receptor beta in terminal differentiation of mammary gland epithelium. Proc Natl Acad Sci U S A 99(24): 15578-15583, 2002. DOI: 10.1073/pnas.192561299
67 De Decker M, Lavaert M, Roels J, Tilleman L, Vandekerckhove B, Leclercq G, Van Nieuwerburgh F, Van Vlierberghe P, Taghon T: HES1 and HES4 have non-redundant roles downstream of Notch during early human T-cell development. Haematologica 106(1): 130-141, 2021. DOI: 10.3324/haematol.2019.226126
68 Tian W, Yu J, Tomchick DR, Pan D, Luo X: Structural and functional analysis of the YAP-binding domain of human TEAD2. Proc Natl Acad Sci U S A 107(16): 7293-7298, 2010. DOI: 10.1073/pnas.1000293107
69 Joo JS, Cho SY, Rou WS, Kim JS, Kang SH, Lee ES, Moon HS, Kim SH, Sung JK, Kwon IS, Eun HS, Lee BS: TEAD2 as a novel prognostic factor for hepatocellular carcinoma. Oncol Rep 43(6): 1785-1796, 2020. DOI: 10.3892/or.2020.7578
70 Chen ZZ, Huang L, Wu YH, Zhai WJ, Zhu PP, Gao YF: LncSox4 promotes the self-renewal of liver tumour-initiating cells through Stat3-mediated Sox4 expression. Nat Commun 7: 12598, 2016. DOI: 10.1038/ncomms12598
71 Gupta R, Malvi P, Parajuli KR, Janostiak R, Bugide S, Cai G, Zhu LJ, Green MR, Wajapeyee N: KLF7 promotes pancreatic cancer growth and metastasis by up-regulating ISG expression and maintaining Golgi complex integrity. Proc Natl Acad Sci U S A 117(22): 12341-12351, 2020. DOI: 10.1073/pnas.2005156117
72 Coffinier C, Gresh L, Fiette L, Tronche F, Schütz G, Babinet C, Pontoglio M, Yaniv M, Barra J: Bile system morphogenesis defects and liver dysfunction upon targeted deletion of HNF1β. Development 129(8): 1829-1838, 2002. DOI: 10.1242/dev.129.8.1829
73 Lopes N, Ribeiro B: Non-negative matrix factorization (NMF). In: Machine learning for adaptive many-core machines - a practical approach. Cham, Switzerland, Springer International Publishing, pp. 127-154, 2015.
74 Noguer O, Villena J, Lorita J, Vilaró S, Reina M: Syndecan-2 downregulation impairs angiogenesis in human microvascular

endothelial cells. Exp Cell Res 315(5): 795-808, 2009. DOI: 10.1016/j.yexcr.2008.11.016
75 Nakajima T, Jorde LB, Ishigami T, Umemura S, Emi M, Lalouel JM, Inoue I: Nucleotide diversity and haplotype structure of the human angiotensinogen gene in two populations. Am J Hum Genet 70(1): 108-123, 2002. DOI: 10.1086/338454
76 Tovar V, Alsinet C, Villanueva A, Hoshida Y, Chiang DY, Solé M, Thung S, Moyano S, Toffanin S, Mínguez B, Cabellos L, Peix J, Schwartz M, Mazzaferro V, Bruix J, Llovet JM: IGF activation in a molecular subclass of hepatocellular carcinoma and pre-clinical efficacy of IGF-1R blockage. J Hepatol 52(4): 550-559, 2010. DOI: 10.1016/j.jhep.2010.01.015
77 Heyliger S, Turley TN, Roach T, Saulsbury MD, Taka E, Reynolds JP, Copland JA 3rd, Kase AM, Reams RR: Analysis of peroxisomal ABCD3 transporter as a prognostic factor in clear cell renal cell carcinoma. Cancer Genomics Proteomics 22(5): 698-715, 2025. DOI: 10.21873/cgp.20530
78 Yin D, Li X, Yang X, Shang X, Li Z, Geng J, Xu Y, Xu Z, Wang Z, Shang Z, Yang Z, Hu L, Li Q, Wang J, Song X, Li X, Wang X: Huqi formula suppresses hepatocellular carcinoma growth by modulating the PI3K/AKT/mTOR pathway and promoting T cell infiltration. Chin Med 20(1): 25, 2025. DOI: 10.1186/s13020-025-01061-w
79 Yang A, Huang H, Xie J, Tian Y, Wang L, Liu D, Wei X, Tan P, Chai X, Zha X, Tu P, Hu Z: Interfering with the AKT/mTOR/STAT3/ID1 signaling axis with usenamine A restrains the proliferative and invasive potential of human hepatocellular carcinoma cells. Chin Med 19(1): 4, 2024. DOI: 10.1186/s13020-023-00875-w
80 Ge Y, Li J, Ming C, Ung COL, Lai Y, Hu H: Proteomic and transcriptomic analyses define molecular subtypes, identify biomarkers, and suggest potential therapeutic agent for early-stage HBV-related hepatocellular carcinoma. J Proteome Res 24(12): 6238-6251, 2025. DOI: 10.1021/acs.jproteome.5c00828