

Single-cell transcriptomic inference of lactylation-related programs identifies PRDX1 as a candidate metabolic regulator in hepatocellular carcinoma

Hongmei Niu

Northwest Minzu University

Xin Zheng

Northwest Minzu University

Weili Li

Weifang People's Hospital

Yuan Niu

Gansu Provincial Second People's Hospital

Jin Liang

Gansu Provincial Hospital

Wanting Cao

Northwest Minzu University

Shuya Ma

Northwest Minzu University

Yifan Gao

Northwest Minzu University

Xingting Bao

Northwest Minzu University

Jun Luo

19151921236@163.com

Gansu University of Traditional Chinese Medicine

Research Article

Keywords: hepatocellular carcinoma, lactylation, single-cell RNA sequencing, PRDX1, tumor metabolism

Posted Date: October 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10404023/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Lactate-derived lysine lactylation has emerged as a link between tumor metabolism and gene regulation. In hepatocellular carcinoma (HCC), however, the cellular distribution of lactylation-related transcriptional programs and their relationship to malignant phenotypes remain incompletely defined.

Methods

We integrated 24 single-cell RNA-sequencing samples (16 HCC and 8 normal liver samples) with TCGA-LIHC bulk RNA-sequencing data. Lactylation-related gene-set activity was estimated using AUCell, UCell, singscore, ssGSEA, and AddModuleScore. Malignant hepatocytes were inferred by copy-number variation analysis, and high-score programs were examined by differential expression, hdWGCNA, CellChat, and CytoTRACE. Bulk transcriptomic feature selection was performed with Boruta, random forest, LASSO, and XGBoost. PRDX1 was then tested in HepG2 and Huh-7 cells by siRNA knockdown assays.

Results

Lactylation-related scores were higher in HCC than in normal tissue and were associated with poorer overall survival in TCGA-LIHC. At single-cell resolution, the highest scores were observed in hepatocytes and myeloid dendritic cells, whereas B cells and plasma cells showed lower scores. High-score malignant hepatocytes displayed increased CytoTRACE scores and distinct ligand-receptor communication patterns. Intersecting high-score differentially expressed genes with hdWGCNA modules yielded 178 candidate genes. Machine-learning prioritization nominated BSG, ERGIC3, PPM1G, and PRDX1, whose higher expression was associated with poorer survival. PRDX1 knockdown reduced HCC-cell proliferation, colony formation, migration, invasion, and intracellular lactate production.

Conclusions

This study defines a transcriptomic lactylation-related program in HCC and nominates PRDX1 as a candidate link between malignant behavior and lactate metabolism. Direct proteomic and clinical validation will be required to establish PRDX1-dependent lactylation mechanisms.

Introduction

Hepatocellular carcinoma (HCC) is the predominant histological subtype of liver cancer and remains a major cause of cancer-related mortality worldwide [1]. Despite progress in surgery, locoregional therapy, molecular targeted treatment, and immunotherapy, clinical outcomes remain limited by recurrence,

therapeutic resistance, and marked intratumoral heterogeneity [2–4]. A clearer understanding of the cellular programs that distinguish malignant hepatocytes from other components of the tumor microenvironment is therefore needed to refine prognostic assessment and identify tractable therapeutic vulnerabilities.

Metabolic reprogramming is a central feature of HCC. Cancer cells frequently increase glycolytic flux and lactate production even under oxygen-replete conditions, a phenotype commonly described as the Warburg effect [5]. Lactate is not only a metabolic by-product; it also shapes extracellular acidity, immune-cell behavior, and tumor progression [6]. More recently, lactate has been recognized as a donor for lysine lactylation, a post-translational modification that can regulate histone and non-histone proteins and may contribute to tumor growth, angiogenesis, metastasis, and therapy resistance [7].

Several studies have connected lactylation-related mechanisms with HCC biology [2, 8]. Histone lactylation marks such as H3K9la and H3K14la have been implicated in proliferation, metastasis, and cancer-stem-cell phenotypes, whereas lactylome studies suggest that non-histone lactylation may affect metabolic enzymes and adaptive tumor metabolism [8]. Nevertheless, most evidence has been generated from bulk tissue or selected experimental models, leaving the cell-type specificity of lactylation-related programs in HCC incompletely resolved.

Single-cell RNA sequencing (scRNA-seq) provides a framework for studying tumor heterogeneity and cell-cell communication at cellular resolution. Here, we integrated public scRNA-seq and bulk RNA-seq datasets to infer lactylation-related transcriptional activity across HCC cell populations, identify malignant hepatocyte programs associated with high lactylation-related scores, and prioritize candidate genes using co-expression and machine-learning approaches. We further performed in vitro knockdown experiments to test PRDX1, the most experimentally tractable candidate emerging from this analysis, in HCC-cell proliferation, migration, invasion, and lactate production.

Methods

Data Acquisition

We analyzed 24 scRNA-seq samples, comprising 16 HCC and 8 normal liver samples, from the GEO database (GSE149614) [9] and Mendeley Data (<https://data.mendeley.com/datasets/skxr2fz79n/1>). Bulk RNA-seq and clinical data were obtained from the TCGA-LIHC cohort, including 365 HCC and 50 normal samples. Detailed information on the included datasets is provided in Table S1 [10–12].

scRNA-Seq Data Processing

For scRNA-seq quality control, cells with fewer than 200 detected genes, more than 7,000 detected genes, or mitochondrial gene content greater than 20% were removed, and genes expressed in fewer than three cells were excluded. A total of 146,515 cells remained for downstream analysis. Data were

normalized and integrated using the Seurat workflow [13, 14]. The top 3,000 highly variable genes were selected with FindVariableFeatures, principal-component analysis was used for dimension reduction, and Harmony was applied to mitigate batch effects [15]. Cells were clustered with FindClusters at a resolution of 0.8 and annotated according to canonical marker genes and cluster-specific expression patterns.

Evaluation of Lactylation Activity

A set of 332 lactylation-related genes was obtained from published literature [16] (Table S2). Because transcriptomic data do not directly measure protein lactylation, we refer to the resulting readout as a lactylation-related activity score. AUCell, UCell, singscore, ssGSEA, and AddModuleScore were used to estimate lactylation-related activity for each cell [17, 18]. Cells below the 25th percentile were assigned to the low-score group, cells between the 25th and 75th percentiles to the intermediate-score group, and cells above the 75th percentile to the high-score group. Differentially expressed genes between high- and low-score cells were identified using FindMarkers. In parallel, hdWGCNA was used to identify co-expression modules associated with lactylation-related activity [19]. Genes shared by high-score differentially expressed genes and lactylation-associated hdWGCNA modules were retained as candidate genes.

Functional Enrichment Analysis

Gene Ontology (GO) and Disease Ontology (DO) enrichment analyses were performed using the clusterProfiler package in R [20]. GO terms were examined across biological process, cellular component, and molecular function categories, and DO enrichment was used to evaluate disease associations among candidate genes.

Identification of HCC Feature Genes

To prioritize robust HCC feature genes from the candidate gene set, we applied four complementary feature-selection methods: LASSO regression, random forest (RF), XGBoost, and Boruta [21–24]. LASSO used 10-fold cross-validation to select genes with non-zero coefficients. RF feature importance was estimated after evaluating model stability across 1-500 trees. XGBoost and Boruta were used to provide additional rankings of informative genes. Genes shared by all four approaches were retained as optimal feature genes.

Expression and Diagnostic Relevance of Optimal Feature Genes

Expression differences in optimal feature genes between HCC and control samples were assessed using the Wilcoxon rank-sum test. Prognostic relevance was evaluated by Kaplan-Meier survival analysis after

stratifying TCGA-LIHC samples by median gene expression. Receiver operating characteristic curves were used when diagnostic discrimination was assessed.

Pseudotime analysis

Monocle2 was used to infer single-cell trajectories [25]. After dimensionality reduction and cell ordering, cell-state transitions were visualized in low-dimensional space. CytoTRACE was then applied with default parameters as an independent estimate of transcriptional diversity and differentiation state [26, 27].

Cell–cell interaction analysis

Cell-cell communication was inferred using CellChat [28]. Ligand-receptor interactions among annotated cell types and lactylation-score-defined malignant hepatocyte groups were summarized by interaction number, interaction strength, signaling roles, and selected ligand-receptor pairs.

Cell culture and transfection

Human HCC cell lines HepG2 and Huh-7 were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin under 37°C in a humidified atmosphere with 5% CO₂. PRDX1 expression was knocked down using two specific small interfering RNAs (siRNAs), labeled Si1 and Si2, targeting PRDX1 mRNA. A non-targeting siRNA (NC) was used as a control. Transfections were performed using Lipofectamine 3000 (Thermo Fisher Scientific, USA) according to the manufacturer's protocol. Cells were harvested 48 hours post-transfection for further analysis.

CCK-8 proliferation assay

Cell proliferation was assessed using the Cell Counting Kit-8 (CCK-8, Dojindo Laboratories, Japan). HepG2 and Huh-7 cells were seeded in 96-well plates at a density of 2×10^3 cells per well. After transfection with PRDX1-specific siRNAs or NC, the cells were incubated with CCK-8 solution, and optical density (OD) at 450 nm was measured daily over five days using a microplate reader.

Colony formation assay

Following transfection, HepG2 and Huh-7 cells were seeded in 6-well plates at a density of 500 cells per well and allowed to grow for two weeks. Colonies were then fixed with 4% paraformaldehyde, stained with 0.1% crystal violet, and imaged. Colonies containing more than 50 cells were counted under a microscope. Quantitative analysis was performed to compare colony numbers across different groups.

Transwell migration and invasion assays

Transwell chambers (8 µm pore size; Corning, USA) were used to assess cell migration and invasion. For the migration assay, 5×10^4 transfected cells in serum-free medium were added to the upper chamber, while medium with 10% FBS was added to the lower chamber as a chemoattractant. After 24 hours, cells that migrated to the bottom of the membrane were fixed, stained with 0.1% crystal violet, and counted. For the invasion assay, the procedure was similar, except that the upper chamber was pre-coated with Matrigel (BD Biosciences, USA).

Wound healing assay

The wound healing assay was conducted to evaluate cell migration. Transfected HepG2 and Huh-7 cells were seeded in 6-well plates and grown to 90% confluency. A uniform scratch was made in the cell monolayer using a sterile 200 µL pipette tip. The cells were then washed with PBS to remove debris and incubated in serum-free DMEM. Images of the wound area were captured at 0 and 24 hours to assess wound closure, and the relative migration distance was calculated using ImageJ software. The experiment was performed in triplicate for statistical reliability.

Lactate assay

Intracellular lactate levels were measured using a lactate assay kit (Sigma-Aldrich, USA) following the manufacturer's instructions. HepG2 and Huh-7 cells transfected with PRDX1 siRNAs or NC were harvested, lysed, and the lactate concentration was quantified by measuring the absorbance at 570 nm using a microplate reader. The relative lactate levels were calculated and compared across groups. Each assay was performed in triplicate.

Statistical analysis

All data processing, statistical analysis, and visualization were performed in R (version 4.1.3). For two-group comparisons, Wilcoxon rank-sum tests or Student's t-tests were used as appropriate. Pearson and Spearman correlation tests were used to evaluate associations between variables. All P values were two-sided, and $P < 0.05$ was considered statistically significant.

Results

Single-cell profiling defines major cell populations in HCC

After quality control, 146,515 cells from 16 HCC and 8 normal liver samples were retained for analysis. Harmony integration produced broadly mixed sample distributions, indicating that major batch effects were reduced (Fig. 1A). Unsupervised clustering identified 30 clusters (Fig. 1B), which were annotated as macrophages, fibroblasts, T cells, B cells, NK cells, monocytes, endothelial cells, myeloid dendritic cells (MDCs), and hepatocytes according to canonical marker genes (Fig. 1C-E). These annotations provided the basis for evaluating lactylation-related transcriptional programs across the HCC ecosystem.

Lactylation-related activity scores are heterogeneous across HCC cell types

In TCGA-LIHC bulk RNA-seq data, the lactylation-related gene-set score was higher in HCC than in normal samples by ssGSEA analysis (Fig. 2A, $P < 0.001$). Patients with high scores had poorer overall survival than those with low scores when stratified by the median value (Fig. 2B, $P = 0.032$). At single-cell resolution, AUCell, UCell, singscore, ssGSEA, and AddModuleScore consistently showed marked heterogeneity across cell types (Fig. 2C-E). Hepatocytes and MDCs displayed relatively high scores, whereas B cells and plasma cells showed lower scores (Fig. 2F-H). High-score cells were more frequent in HCC samples than in normal samples (Fig. 2I). Because malignant hepatocytes drive much of the HCC transcriptomic program, inferCNV was used to distinguish malignant from non-malignant hepatocytes. K-means clustering of copy-number profiles identified 12,932 malignant hepatocytes for downstream analysis (Fig. 2J; Table S3).

High-score malignant hepatocytes show increased stemness-associated transcriptional features

To focus on malignant hepatocyte heterogeneity, malignant cells were stratified into low-, intermediate-, and high-score groups using quartile thresholds (low, < 1.649 ; high, > 2.481 ; Fig. 3A, B). Among inferred malignant hepatocytes, high-, intermediate-, and low-score cells accounted for 63.51%, 31.56%, and 4.93% of cells, respectively (Fig. 3C, D). Reclustering showed that lactylation-related scores were not evenly distributed but instead formed distinct high-, intermediate-, and low-score malignant-cell states (Fig. 3E, F). CytoTRACE analysis indicated that high-score malignant cells had higher transcriptional diversity scores and spatially overlapped with regions of high lactylation-related activity (Fig. 3G-K), suggesting an association between the high-score program and stemness-like or less differentiated malignant-cell states.

High-score malignant hepatocytes show altered cell-cell communication patterns

We next examined whether high-score malignant hepatocytes differed in inferred intercellular communication. CellChat analysis indicated that high-score malignant cells occupied prominent signaling roles and showed stronger inferred communication with other cell populations than low-score malignant cells (Fig. 4A, B). Signaling-role analysis highlighted outgoing pathways including MIF, MK, ANGPTL, GDF, PROS, BAG, and FGF, and incoming pathways including MK, ANGPTL, PTN, EGF, TWEAK, and PERIOSTIN (Fig. 4C). Ligand-receptor bubble plots further suggested distinct communication patterns between lactylation-score-defined malignant cells and other HCC cell types (Fig. 4D-F). These data support an association between the high-score malignant-cell state and altered tumor-microenvironmental signaling.

hdWGCNA identifies gene modules associated with high lactylation-related scores

To identify transcriptional modules associated with high lactylation-related activity, hdWGCNA was applied to malignant hepatocytes. A soft-thresholding power of 7 was selected to approximate scale-free topology while maintaining network connectivity (Fig. 5A). Seven modules were identified (Fig. 5B), and hub genes within each module were ranked by eigengene-based connectivity (Fig. 5C, D). Several modules, including black, red, blue, turquoise, green, and yellow, showed specificity for high-score cells (Fig. 5E). The black module showed the strongest positive association with the high-score malignant-cell group (Fig. 5F). Based on top hub genes from these modules, 589 genes were retained as high-score-associated hub genes (Table S4).

Candidate genes associated with high lactylation-related malignant-cell programs

Differential expression analysis between high- and low-score malignant hepatocytes identified 577 upregulated genes in the high-score group ($\text{avg_log2FC} > 0.25$ and adjusted $P < 0.05$; Fig. 6A; Table S5). Intersecting these genes with hdWGCNA hub genes yielded 178 candidate genes associated with high lactylation-related malignant-cell programs (Fig. 6B; Table S6). DO enrichment linked these genes to liver-related and inflammation-related disease categories, including hepatitis-associated terms relevant to HCC risk (Fig. 6C; Table S7). GO enrichment highlighted neutrophil activation, antigen processing and presentation, vacuolar and lytic vacuole membranes, oxidoreductase activity, and antioxidant activity (Fig. 6D; Table S8), suggesting that immune regulation, vesicular processes, and redox metabolism may contribute to the high-score malignant-cell state.

Machine learning prioritizes four candidate feature genes

The 178 candidate genes were further analyzed using Boruta, LASSO, XGBoost, and RF to prioritize stable feature genes. Boruta identified 12 confirmed feature genes (Fig. 7A, B; Table S9). LASSO retained 12 genes after 10-fold cross-validation (Fig. 7C, D), XGBoost identified 20 genes with importance > 0.02 (Fig. 7E; Table S10), and RF ranked 26 genes with importance greater than zero (Fig. 7F, G). The intersection of the four approaches nominated four optimal feature genes: BSG, ERGIC3, PPM1G, and PRDX1 (Fig. 7H, I).

Feature genes are associated with prognosis and are enriched in malignant hepatocytes

TCGA-LIHC patients were stratified into high- and low-expression groups according to the median expression of each feature gene. Higher expression of BSG, ERGIC3, PPM1G, and PRDX1 was associated with poorer overall survival (Fig. 8A-D). Single-cell analysis showed that these genes were preferentially

expressed in malignant hepatocytes relative to several non-malignant cell types (Fig. 8E-G). Within malignant hepatocytes, expression of the four genes was most evident in cells with high or intermediate lactylation-related scores (Fig. 8H), supporting their relevance to the high-score malignant-cell program.

PRDX1 supports malignant phenotypes and lactate production in HCC cells

Because PRDX1 was consistently prioritized and has reported links to redox and metabolic regulation, we tested its functional relevance in HepG2 and Huh-7 cells using two independent siRNAs. PRDX1 knockdown decreased cell viability over five days in CCK-8 assays (Fig. 9A, B) and reduced colony formation (Fig. 9C). Transwell and wound-healing assays showed that PRDX1 silencing impaired migration and invasion in both cell lines (Fig. 9D-G). Finally, intracellular lactate levels were lower after PRDX1 knockdown (Fig. 9H, I). Together, these findings support a role for PRDX1 in maintaining malignant HCC-cell phenotypes and lactate production, although direct PRDX1-dependent protein lactylation remains to be demonstrated.

Discussion

HCC is characterized by metabolic reprogramming, cellular heterogeneity, and extensive interaction between malignant hepatocytes and the tumor microenvironment [29–32]. By integrating scRNA-seq, bulk RNA-seq, co-expression analysis, machine learning, and in vitro perturbation, this study maps lactylation-related transcriptional programs in HCC and identifies PRDX1 as a candidate metabolic regulator linked to malignant phenotypes.

Lactate accumulation is a hallmark of glycolytic tumor metabolism and has regulatory roles beyond energy production [33–37]. Lactylation provides a mechanistic framework by which lactate abundance may influence chromatin state, protein function, and tumor-microenvironmental behavior [38–42]. Consistent with these concepts, our bulk analysis found that a lactylation-related gene-set score was elevated in HCC and associated with poorer survival. Single-cell analysis further indicated that this program is not uniformly distributed but enriched in hepatocytes and MDCs, with relatively low scores in B cells and plasma cells.

Within malignant hepatocytes, high lactylation-related scores were associated with increased CytoTRACE scores and distinct inferred communication patterns. These findings align with studies linking lactate and lactylation to cancer stem-cell properties, immune remodeling, and malignant progression [43–45]. However, transcriptomic scoring should be interpreted as an indirect readout. It captures coordinated expression of lactylation-related genes rather than the absolute abundance of lactylated proteins or specific lactylation sites.

The candidate genes prioritized here overlap conceptually with prior studies that used lactylation-related gene signatures to stratify HCC prognosis and treatment response [49, 50]. Our analysis adds cellular

resolution by linking candidate genes to malignant hepatocyte states and then testing PRDX1 experimentally. The four nominated genes have biological plausibility in HCC. ERGIC3 participates in ER-Golgi trafficking and has been linked to epithelial-mesenchymal transition [51–53]. PPM1G regulates phosphorylation-dependent nuclear processes and has been associated with HCC progression and stemness [54–57]. BSG/CD147 is a well-established regulator of tumor metabolism, invasion, immune interactions, and HCC therapeutic targeting [62–66].

PRDX1 emerged as a particularly relevant candidate because of its roles in redox homeostasis, metabolism, and tumor biology [58–61, 67, 68]. Our siRNA experiments showed that PRDX1 knockdown suppressed proliferation, colony formation, migration, invasion, and lactate production in two HCC cell lines. These data support PRDX1 as a functional contributor to malignant HCC-cell behavior and lactate-linked metabolic activity. They do not yet prove that PRDX1 directly controls protein lactylation; rather, they nominate PRDX1 as a candidate upstream regulator that should be tested by lactylation-specific assays.

The study has several limitations. First, lactylation-related scores were inferred from RNA expression and should be validated using direct lactylation measurements, such as pan-lysine lactylation, histone lactylation markers, or mass spectrometry-based lactylomics. Second, the four-gene signature requires validation in independent external cohorts and, ideally, in clinical HCC tissue by immunohistochemistry or spatial profiling. Third, PRDX1 knockdown was tested in vitro only; rescue, overexpression, lactate-supplementation, and in vivo experiments would strengthen causal inference. Finally, cell-cell communication and pseudotime analyses are computational inferences and require experimental validation.

In conclusion, this study identifies a lactylation-related transcriptomic program enriched in malignant HCC hepatocytes and linked to stemness-associated features, altered cell-cell communication, and poor prognosis. Integrated feature selection nominates BSG, ERGIC3, PPM1G, and PRDX1, and functional assays support PRDX1 as a candidate regulator of HCC-cell malignant phenotypes and lactate production. These findings provide a focused basis for future work directly testing PRDX1-dependent lactylation mechanisms in HCC.

Declarations

Declaration of Competing 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.

Ethics Statement

The study was conducted in accordance with the Declaration of Helsinki. Ethical review and approval were waived for this study as it exclusively utilized publicly available, de-identified data from the GEO and TCGA databases.

Funding

This work was supported by the National Natural Science Foundation of China (Regional Program; Grant No. 82360783); the Fundamental Research Funds for the Central Universities (Grant No. 31920250007); the Gansu Provincial Science and Technology Program for Basic Research—Natural Science Foundation of Gansu Province (Grant No. 25JRRA040); the National Natural Science Foundation of China (Young Scientists Fund; Grant No. 12405395); the Youth Science and Technology Foundation of Gansu Province (Grant No. 24JRRA998); and the Undergraduate Innovation and Entrepreneurship Training Program of Northwest Minzu University (Grant No. 202510742011).

Author Contribution

Hongmei Niu: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. Xin Zheng: Data curation, Validation, Investigation. Weili Li: Validation, Investigation, Visualization. Yuan Niu: Software, Formal analysis. Jin Liang: Resources, Data curation. Wanting Cao: Software, Visualization. Shuya Ma: Investigation, Validation. Yifan Gao: Resources, Data curation. Xingting Bao: Conceptualization, Supervision, Writing – review & editing, Project administration. Jun Luo: Conceptualization, Supervision, Funding acquisition, Writing – review & editing. All authors have read and agreed to the published version of the manuscript.

Data Availability

Publicly available datasets were analyzed in this study. The scRNA-seq data are available from the Gene Expression Omnibus under accession GSE149614 and from Mendeley Data (<https://data.mendeley.com/datasets/skrx2fz79n/1>). Bulk RNA-seq data are available from The Cancer Genome Atlas (TCGA-LIHC project). Additional processed data supporting the findings are provided in the supplementary tables or are available from the corresponding authors upon reasonable request.

References

1. Bray F, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424.
2. Xu H, et al. Royal jelly acid suppresses hepatocellular carcinoma tumorigenicity by inhibiting H3 histone lactylation at H3K9la and H3K14la sites. *Phytomedicine*. 2023;118:154940.
3. Lin S, et al. Single-cell RNA-Seq Elucidates the Crosstalk Between Cancer Stem Cells and the Tumor Microenvironment in Hepatocellular Carcinoma. *J Cancer*. 2024;15(4):1093–109.

4. Yu Y, et al. Hepatic stellate cells promote hepatocellular carcinoma development by regulating histone lactylation: Novel insights from single-cell RNA sequencing and spatial transcriptomics analyses. *Cancer Lett.* 2024;604:217243.
5. Pascale RM et al. The Warburg Effect 97 Years after Its Discovery. *Cancers (Basel)*, 2020. 12(10).
6. Zhang Y, et al. The function and mechanism of lactate and lactylation in tumor metabolism and microenvironment. *Genes Dis.* 2023;10(5):2029–37.
7. Lv X, Lv Y, Dai X. Lactate, histone lactylation and cancer hallmarks. *Expert Rev Mol Med.* 2023;25:e7.
8. Yang Z, et al. Lactylome analysis suggests lactylation-dependent mechanisms of metabolic adaptation in hepatocellular carcinoma. *Nat Metab.* 2023;5(1):61–79.
9. Lu Y, et al. A single-cell atlas of the multicellular ecosystem of primary and metastatic hepatocellular carcinoma. *Nat Commun.* 2022;13(1):4594.
10. Roessler S, et al. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. *Cancer Res.* 2010;70(24):10202–12.
11. Grinchuk OV, et al. Tumor-adjacent tissue co-expression profile analysis reveals pro-oncogenic ribosomal gene signature for prognosis of resectable hepatocellular carcinoma. *Mol Oncol.* 2018;12(1):89–113.
12. Ning J, et al. Macrophage-coated tumor cluster aggravates hepatoma invasion and immunotherapy resistance via generating local immune deprivation. *Cell Rep Med.* 2024;5(5):101505.
13. Butler A, et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol.* 2018;36(5):411–20.
14. Sun X, et al. Architecting the metabolic reprogramming survival risk framework in LUAD through single-cell landscape analysis: three-stage ensemble learning with genetic algorithm optimization. *J Transl Med.* 2024;22(1):353.
15. Korsunsky I, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods.* 2019;16(12):1289–96.
16. Huang H, et al. A multi-dimensional approach to unravel the intricacies of lactylation related signature for prognostic and therapeutic insight in colorectal cancer. *J Transl Med.* 2024;22(1):211.
17. Du C, et al. Machine learning algorithms integrate bulk and single-cell RNA data to unveil oxidative stress following intracerebral hemorrhage. *Int Immunopharmacol.* 2024;137:112449.
18. Aibar S, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods.* 2017;14(11):1083–6.
19. Han J, et al. Intratumoral immune heterogeneity of prostate cancer characterized by typing and hub genes. *J Cell Mol Med.* 2023;27(1):101–12.
20. Yu G, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics.* 2012;16(5):284–7.
21. Zhang S, et al. Identifying DNA-binding proteins based on multi-features and LASSO feature selection. *Biopolymers.* 2021;112(2):e23419.

22. Walker AM, et al. Evaluating the performance of random forest and iterative random forest based methods when applied to gene expression data. *Comput Struct Biotechnol J*. 2022;20:3372–86.
23. Fang ZG, et al. Application of a data-driven XGBoost model for the prediction of COVID-19 in the USA: a time-series study. *BMJ Open*. 2022;12(7):e056685.
24. Huang RH, et al. A machine learning framework develops a DNA replication stress model for predicting clinical outcomes and therapeutic vulnerability in primary prostate cancer. *J Transl Med*. 2023;21(1):20.
25. Qiu X, et al. Single-cell mRNA quantification and differential analysis with Censur. *Nat Methods*. 2017;14(3):309–15.
26. Cheng S, et al. A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell*. 2021;184(3):792–e80923.
27. Gulati GS, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science*. 2020;367(6476):405–11.
28. Jin S, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. 2021;12(1):1088.
29. Vogel A, et al. Hepatocellular carcinoma. *Lancet*. 2022;400(10360):1345–62.
30. Yang JD, et al. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol*. 2019;16(10):589–604.
31. Brown ZJ, et al. Management of Hepatocellular Carcinoma: A Review. *JAMA Surg*. 2023;158(4):410–20.
32. Wang Z, et al. Precision diagnosis of hepatocellular carcinoma. *Chin Med J (Engl)*. 2023;136(10):1155–65.
33. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell*, 2017. 169(7): pp. 1327–e134123.
34. Du D, et al. Metabolic dysregulation and emerging therapeutical targets for hepatocellular carcinoma. *Acta Pharm Sin B*. 2022;12(2):558–80.
35. Li X, et al. Lactate metabolism in human health and disease. *Signal Transduct Target Ther*. 2022;7(1):305.
36. Rabinowitz JD, Enerbäck S. Lactate: the ugly duckling of energy metabolism. *Nat Metab*. 2020;2(7):566–71.
37. Ye L, Jiang Y, Zhang M. Crosstalk between glucose metabolism, lactate production and immune response modulation. *Cytokine Growth Factor Rev*. 2022;68:81–92.
38. Wang T, et al. Lactate-induced protein lactylation: A bridge between epigenetics and metabolic reprogramming in cancer. *Cell Prolif*. 2023;56(10):e13478.
39. Chen AN, et al. Lactylation, a Novel Metabolic Reprogramming Code: Current Status and Prospects. *Front Immunol*. 2021;12:688910.

40. Yang Z, Zheng Y, Gao Q. Lysine lactylation in the regulation of tumor biology. *Trends Endocrinol Metab.* 2024;35(8):720–31.
41. Qu J, Li P, Sun Z. Histone lactylation regulates cancer progression by reshaping the tumor microenvironment. *Front Immunol.* 2023;14:1284344.
42. Zhang D, et al. Metabolic regulation of gene expression by histone lactylation. *Nature.* 2019;574(7779):575–80.
43. Cai J et al. Targeting SRSF10 might inhibit M2 macrophage polarization and potentiate anti-PD-1 therapy in hepatocellular carcinoma. *Cancer Commun (Lond)*, 2024.
44. Jin L, et al. FGF21-Sirtuin 3 Axis Confers the Protective Effects of Exercise Against Diabetic Cardiomyopathy by Governing Mitochondrial Integrity. *Circulation.* 2022;146(20):1537–57.
45. Pan L, et al. Demethylzeylasteral targets lactate by inhibiting histone lactylation to suppress the tumorigenicity of liver cancer stem cells. *Pharmacol Res.* 2022;181:106270.
46. Narita T, Weinert BT, Choudhary C. Functions and mechanisms of non-histone protein acetylation. *Nat Rev Mol Cell Biol.* 2019;20(3):156–74.
47. Shvedunova M, Akhtar A. Modulation of cellular processes by histone and non-histone protein acetylation. *Nat Rev Mol Cell Biol.* 2022;23(5):329–49.
48. Gu J, et al. Tumor metabolite lactate promotes tumorigenesis by modulating MOESIN lactylation and enhancing TGF- β signaling in regulatory T cells. *Cell Rep.* 2022;39(12):110986.
49. Cheng Z et al. Lactylation-Related Gene Signature Effectively Predicts Prognosis and Treatment Responsiveness in Hepatocellular Carcinoma. *Pharmaceuticals (Basel)*, 2023. 16(5).
50. Wu Q, et al. Integrated analysis of histone lysine lactylation (Kla)-specific genes suggests that NR6A1, OSBP2 and UNC119B are novel therapeutic targets for hepatocellular carcinoma. *Sci Rep.* 2023;13(1):18642.
51. Yoo W, et al. The E3 ubiquitin ligase MARCH2 regulates ERGIC3-dependent trafficking of secretory proteins. *J Biol Chem.* 2019;294(28):10900–12.
52. Guan L, et al. ERGIC2 and ERGIC3 regulate the ER-to-Golgi transport of gap junction proteins in metazoans. *Traffic.* 2022;23(3):140–57.
53. Zhang LY, et al. miR-490-3p modulates cell growth and epithelial to mesenchymal transition of hepatocellular carcinoma cells by targeting endoplasmic reticulum-Golgi intermediate compartment protein 3 (ERGIC3). *J Biol Chem.* 2013;288(6):4035–47.
54. Kamada R, et al. Metal-dependent Ser/Thr protein phosphatase PPM family: Evolution, structures, diseases and inhibitors. *Pharmacol Ther.* 2020;215:107622.
55. Chen D, et al. PPM1G promotes the progression of hepatocellular carcinoma via phosphorylation regulation of alternative splicing protein SRSF3. *Cell Death Dis.* 2021;12(8):722.
56. Hu W, et al. PPM1G promotes cell proliferation via modulating mutant GOF p53 protein expression in hepatocellular carcinoma. *iScience.* 2024;27(3):109116.

57. Zhang X et al. PPM1G and its diagnostic, prognostic and therapeutic potential in HCC (Review). *Int J Oncol*. 2024. 65(5).
58. Guan X, et al. Dual role of PRDX1 in redox-regulation and tumorigenesis: Past and future. *Free Radic Biol Med*. 2024;210:120–9.
59. Xu M, et al. Expression and prognostic roles of PRDXs gene family in hepatocellular carcinoma. *J Transl Med*. 2021;19(1):126.
60. Zhang Y, et al. UHRF2 promotes the malignancy of hepatocellular carcinoma by PARP1 mediated autophagy. *Cell Signal*. 2023;109:110782.
61. An Y, et al. Peroxiredoxin 1 is essential for natamycin-triggered apoptosis and protective autophagy in hepatocellular carcinoma. *Cancer Lett*. 2021;521:210–23.
62. Asgari R, et al. CD147 and MMPs as key factors in physiological and pathological processes. *Biomed Pharmacother*. 2023;157:113983.
63. Kanekura T. CD147/Basigin Is Involved in the Development of Malignant Tumors and T-Cell-Mediated Immunological Disorders via Regulation of Glycolysis. *Int J Mol Sci*. 2023. 24(24).
64. Li J, et al. CD147 reprograms fatty acid metabolism in hepatocellular carcinoma cells through Akt/mTOR/SREBP1c and P38/PPAR α pathways. *J Hepatol*. 2015;63(6):1378–89.
65. Huang Q, et al. CD147 promotes reprogramming of glucose metabolism and cell proliferation in HCC cells by inhibiting the p53-dependent signaling pathway. *J Hepatol*. 2014;61(4):859–66.
66. Tseng HC, et al. Efficacy of anti-CD147 chimeric antigen receptors targeting hepatocellular carcinoma. *Nat Commun*. 2020;11(1):4810.
67. Sun HH, et al. PRDX1 Influences The Occurrence and Progression of Liver Cancer by Inhibiting Mitochondrial Apoptosis Pathway. *Cell J*. 2022;24(11):657–64.
68. Moretton A, et al. A metabolic map of the DNA damage response identifies PRDX1 in the control of nuclear ROS scavenging and aspartate availability. *Mol Syst Biol*. 2023;19(7):e11267.
69. Wang K, et al. CD147-K148me2-Driven Tumor Cell-Macrophage Crosstalk Provokes NSCLC Immunosuppression via the CCL5/CCR5 Axis. *Adv Sci (Weinh)*. 2024;11(29):e2400611.
70. Dalton KM et al. Catastrophic ATP loss underlies a metabolic combination therapy tailored for MYCN-amplified neuroblastoma. *Proc Natl Acad Sci U S A*. 2021. 118(13).
71. Su J, Chen X, Kanekura T. A CD147-targeting siRNA inhibits the proliferation, invasiveness, and VEGF production of human malignant melanoma cells by down-regulating glycolysis. *Cancer Lett*. 2009;273(1):140–7.
72. Wang K, et al. Di-methylation of CD147-K234 Promotes the Progression of NSCLC by Enhancing Lactate Export. *Cell Metab*. 2021;33(1):160–e1736.

Figures

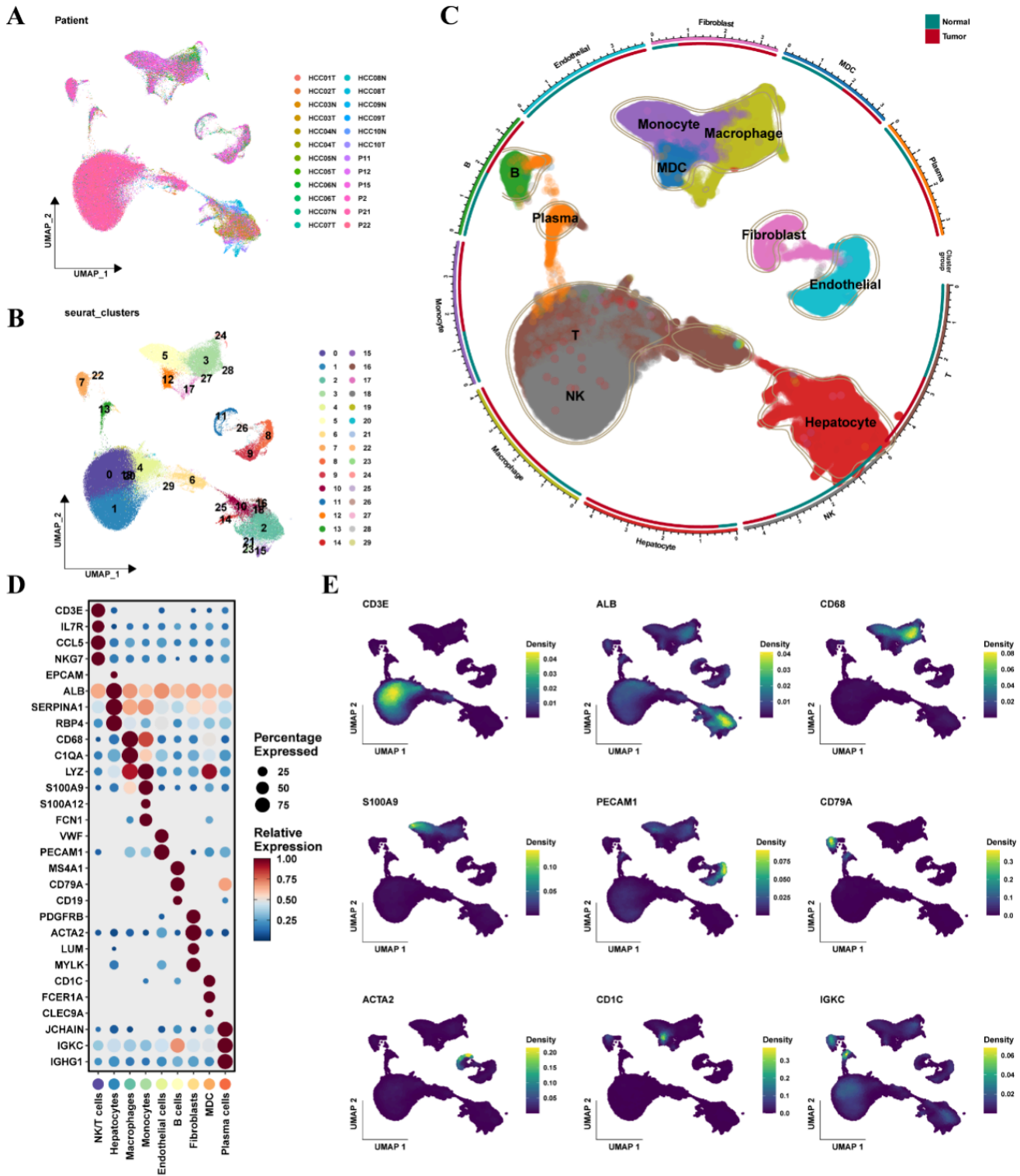


Figure 1

Overview of scRNA-seq cell populations in HCC. (A) Sample distribution after integration. (B) UMAP plot showing 30 cell clusters. (C) UMAP plot colored by major cell-type annotation. (D) Dot plot of representative marker genes for each annotated cell type. (E) UMAP feature plots for selected marker genes.

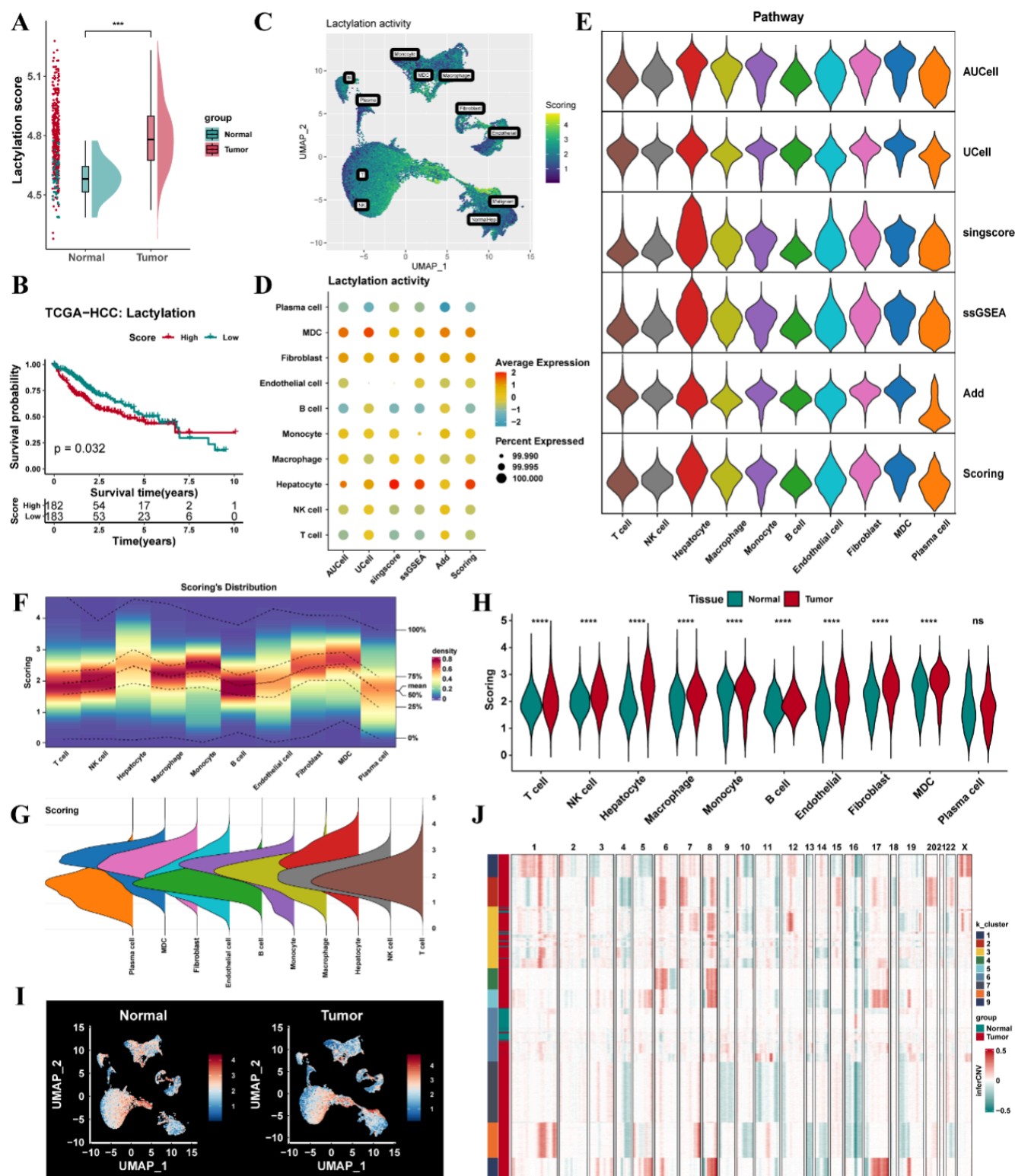


Figure 2

Lactylation-related activity scores across HCC datasets and cell types. (A) ssGSEA lactylation-related score in TCGA-LIHC HCC and normal samples. (B) Overall survival of patients with high versus low lactylation-related scores. (C) UMAP visualization of lactylation-related score distribution. (D, E) Dot plots and violin plots showing scores estimated by AUCell, UCell, singscore, ssGSEA, and AddModuleScore. (F, G) Heatmap and ridge plot of score distributions across 10 cell types. (H) Cell-type-specific score

differences between normal and HCC samples. (I) UMAP distribution of high-score cells in normal and HCC samples. (J) inferCNV analysis of hepatocyte populations.

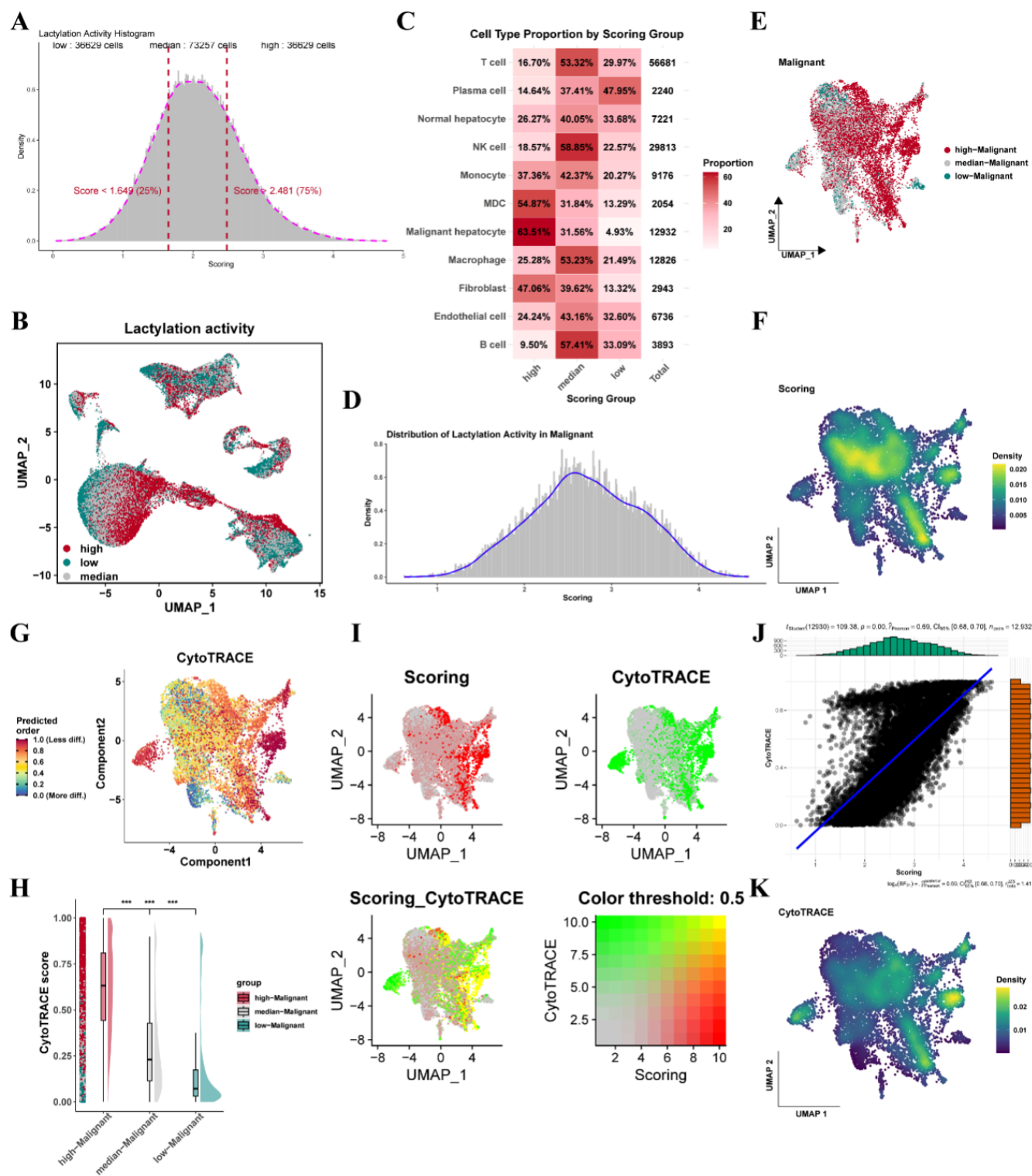


Figure 3

Lactylation-related score states in malignant hepatocytes. (A, B) Distribution of total HCC cells and quartile-based stratification into high-, intermediate-, and low-score groups. (C) Cell-type proportions

across score groups. (D) Distribution of lactylation-related scores in malignant hepatocytes. (E) UMAP plot of malignant hepatocyte score states. (F) Density of lactylation-related scores in malignant hepatocytes. (G) CytoTRACE analysis. (H) CytoTRACE scores across malignant hepatocyte score states. (I-K) Colocalization, correlation, and density analyses of CytoTRACE and lactylation-related scores.

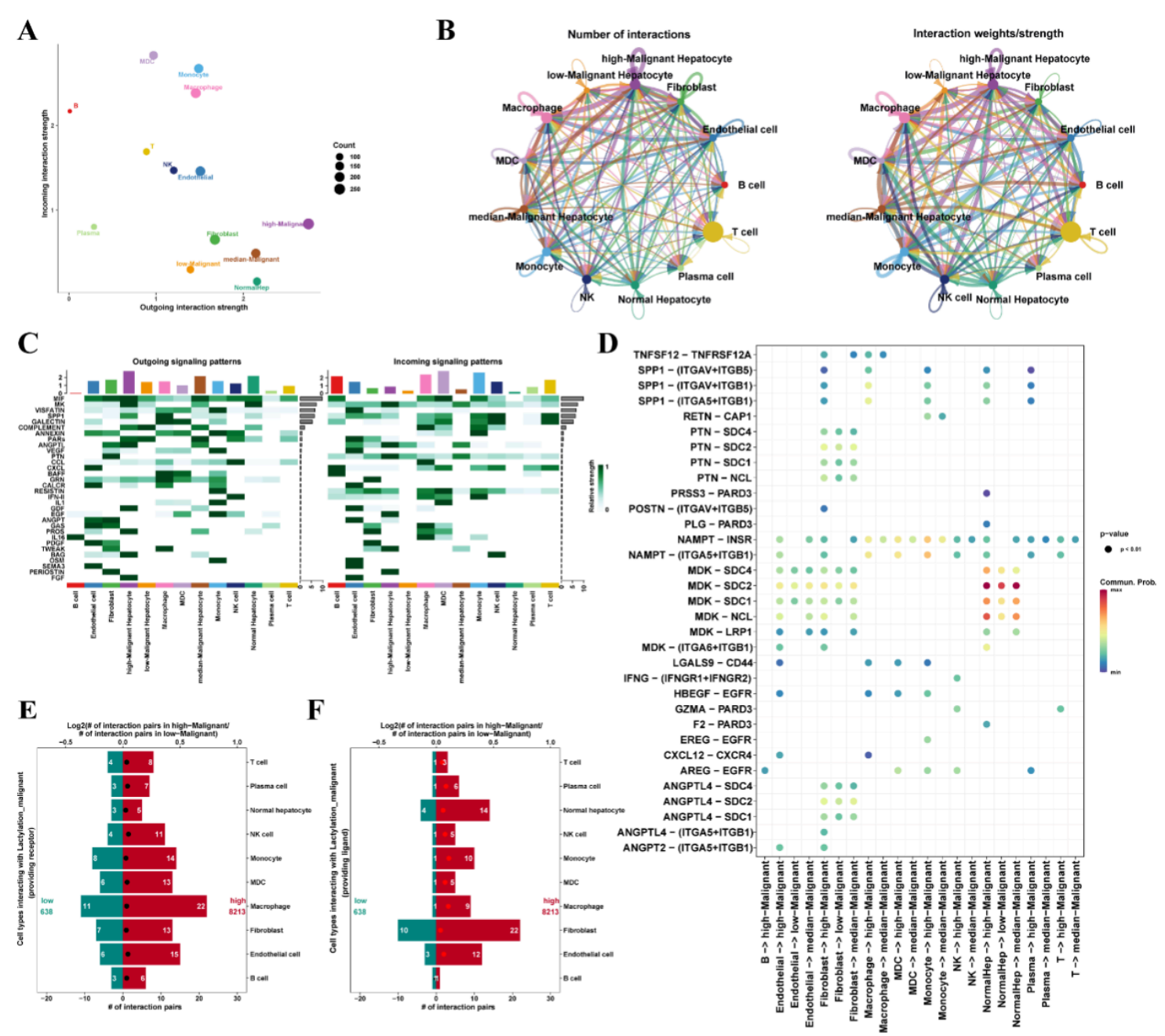


Figure 4

Cell-cell communication associated with lactylation-related malignant hepatocyte states. (A) Inferred incoming and outgoing signaling roles of cell types. (B) Number and strength of inferred interactions in the HCC cell-cell communication network. (C) Signaling-role analysis across pathways. (D) Ligand-receptor bubble plot of selected interactions. (E, F) Cell types interacting with lactylation-score-defined malignant cells through receptor- and ligand-providing interactions.

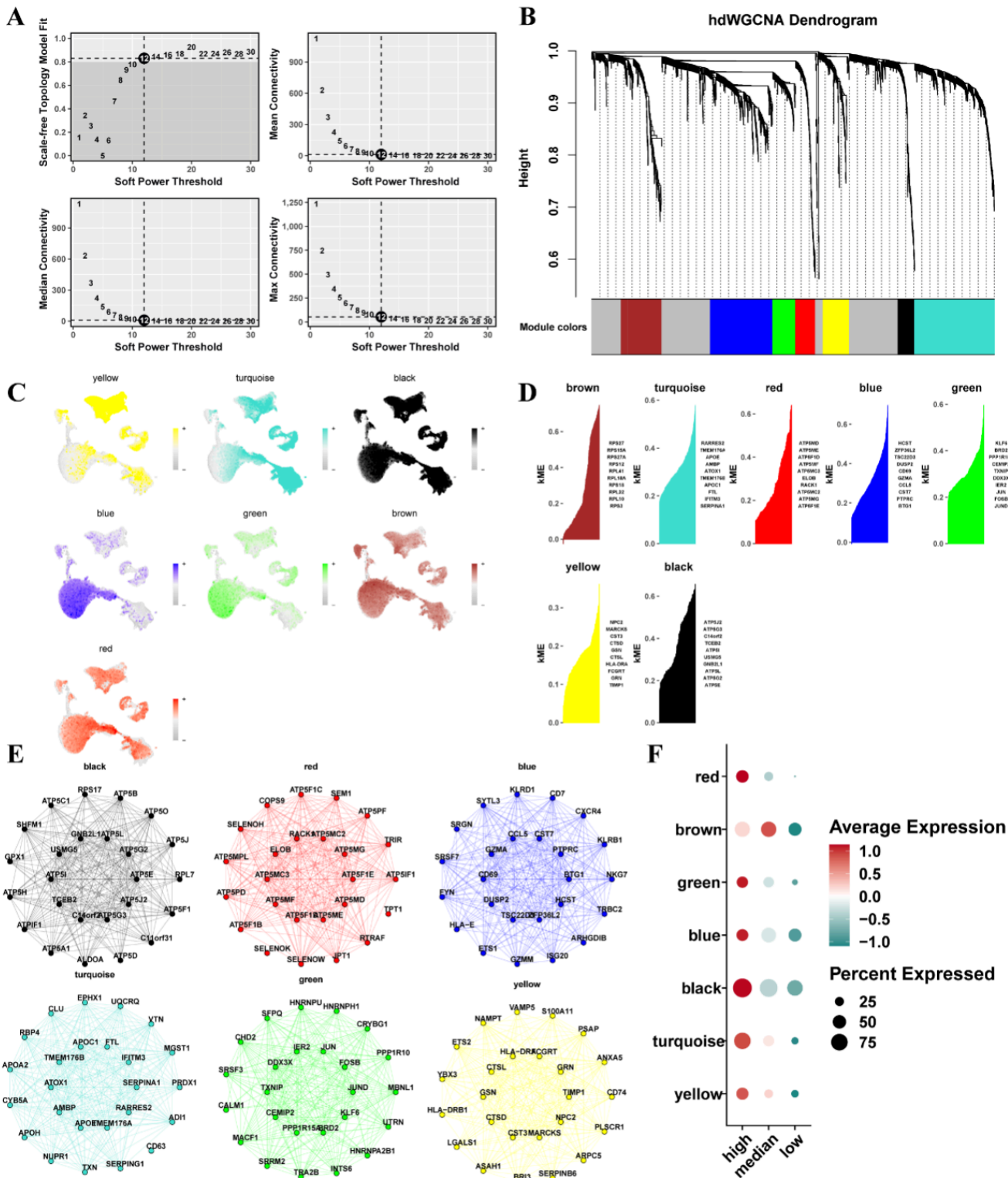


Figure 5

hdWGCNA analysis of malignant hepatocytes. (A) Selection of soft-thresholding power and network connectivity metrics. (B) Module dendrogram identifying seven co-expression modules. (C) Module scores in malignant hepatocytes. (D) Hub genes ranked by module eigengene-based connectivity. (E) Module networks for selected high-score-associated modules. (F) Module scores across lactylation-related score groups.

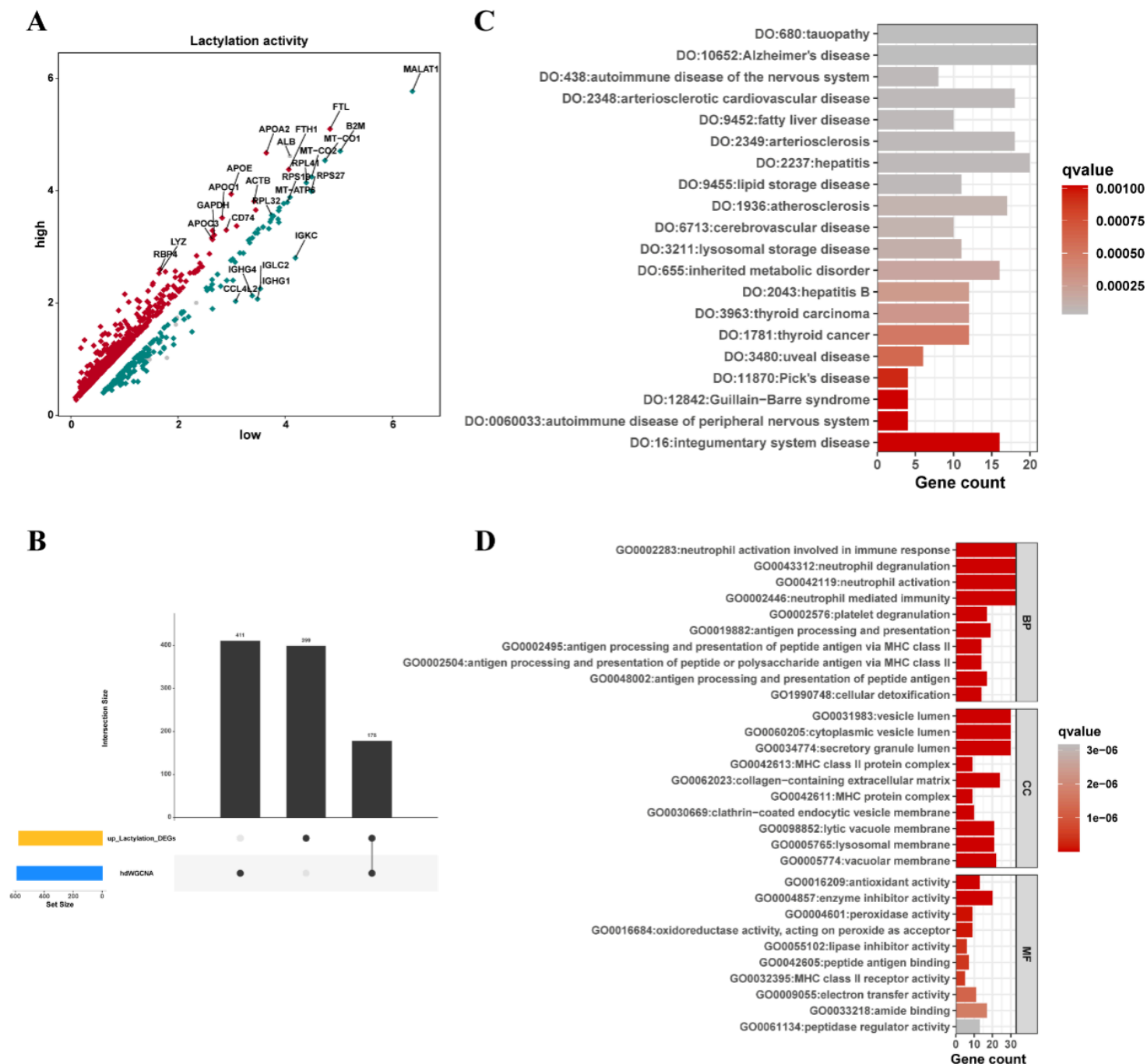


Figure 6

Candidate genes associated with high lactylation-related malignant hepatocytes. (A) Differential expression analysis comparing high- and low-score malignant hepatocytes. (B) Venn diagram showing the intersection between high-score upregulated genes and hdWGCNA hub genes. (C) Disease Ontology enrichment analysis. (D) Gene Ontology enrichment analysis.

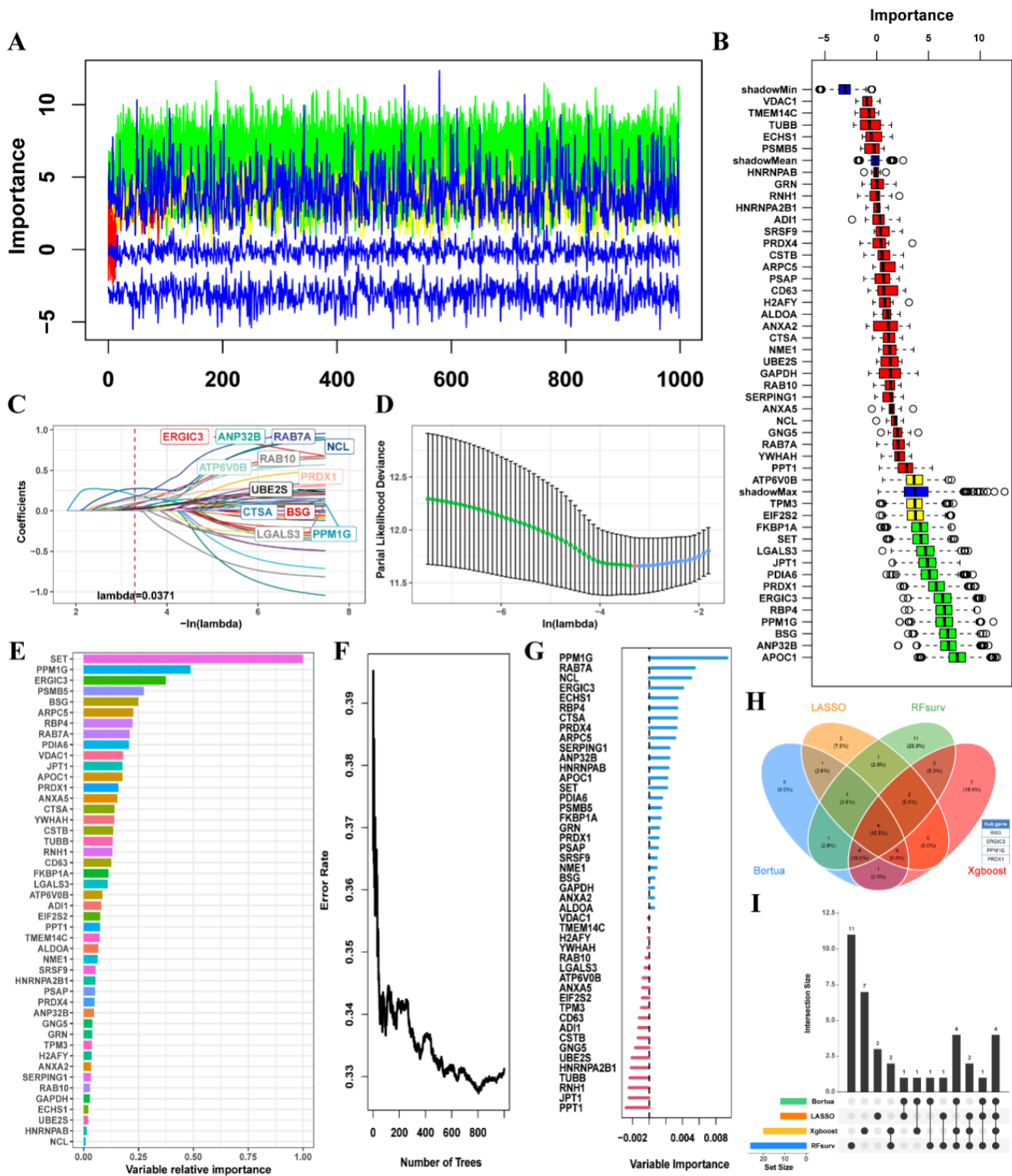


Figure 7

Machine-learning prioritization of candidate feature genes. (A) Boruta feature-selection history. (B) Boruta-confirmed features; yellow denotes confirmed features and other colors denote shadow attributes. (C) LASSO coefficient profiles. (D) Cross-validation curve for LASSO. (E) XGBoost feature-importance ranking. (F) RF error rate across tree numbers. (G) RF feature-importance ranking. (H, I) Venn and UpSet plots showing the four shared feature genes identified by all algorithms.

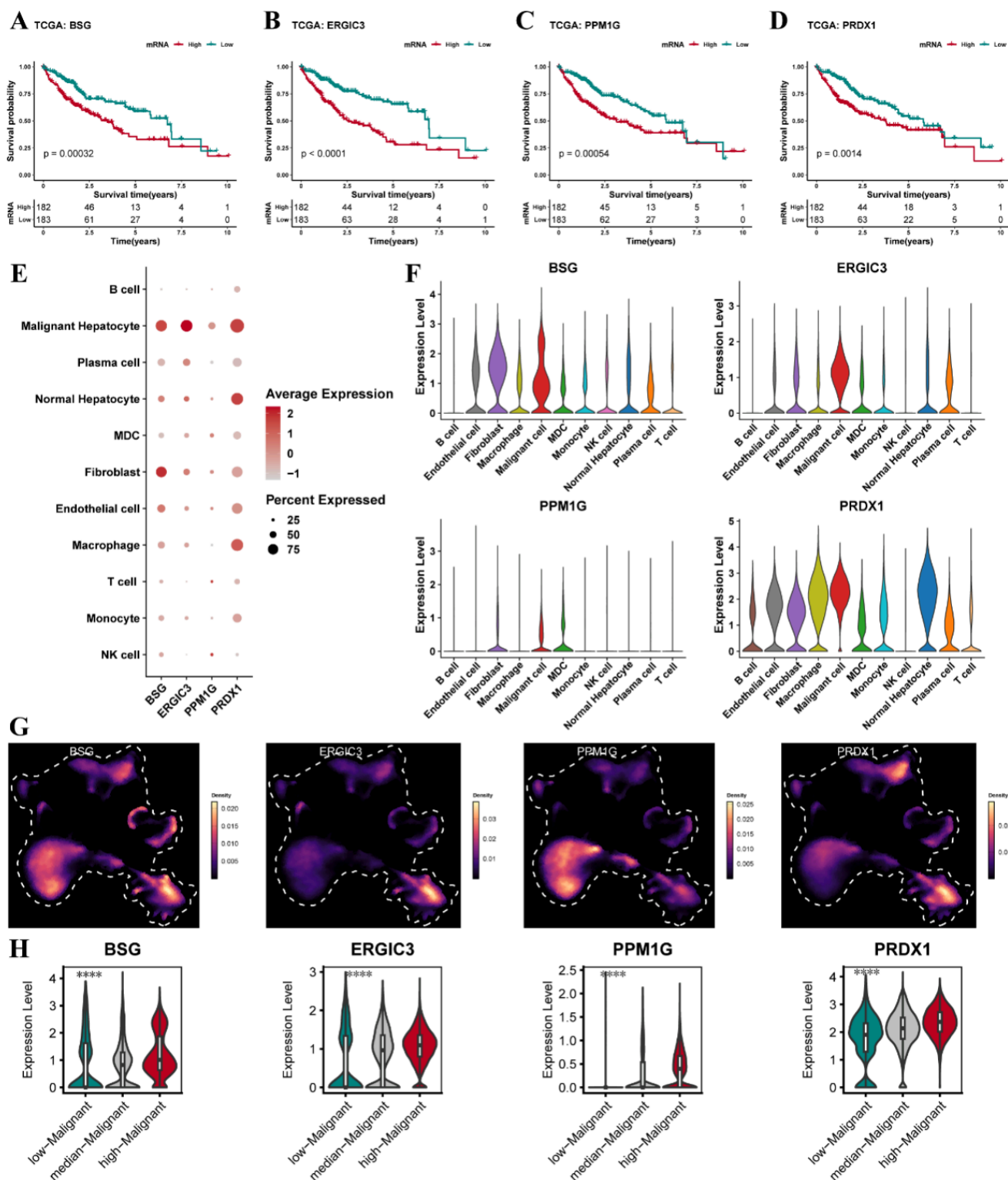


Figure 8

Prognostic and single-cell expression patterns of feature genes. (A-D) Kaplan-Meier overall survival analyses for BSG, ERGIC3, PPM1G, and PRDX1. (E) Dot plot showing feature-gene expression across scRNA-seq cell types. (F) Violin plots of feature-gene expression at single-cell resolution. (G) UMAP density plots for BSG, ERGIC3, PPM1G, and PRDX1. (H) Violin plots showing feature-gene expression across malignant hepatocyte score states.

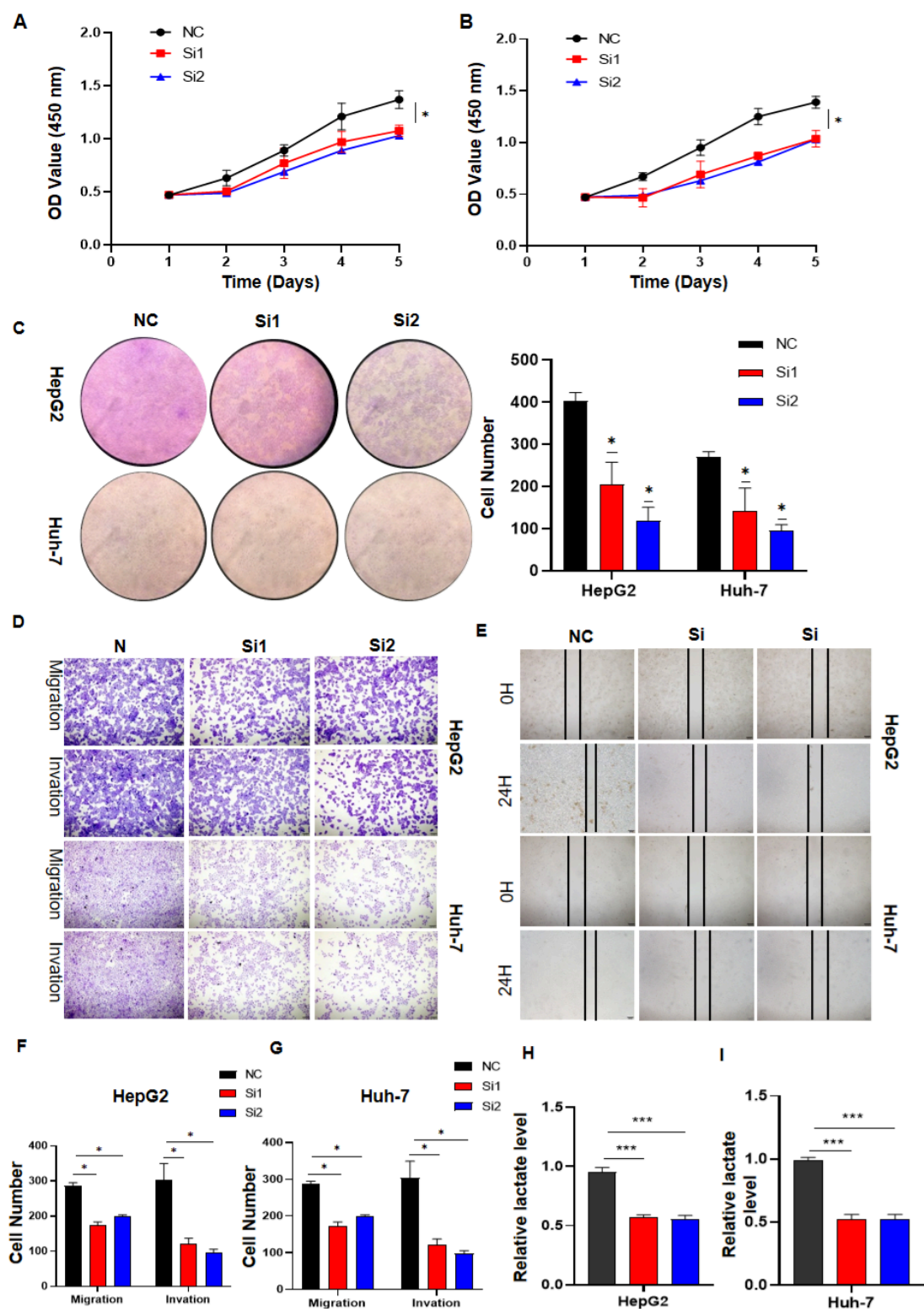


Figure 9

Functional validation of PRDX1 in HCC cells. (A, B) CCK-8 assays showing the effect of PRDX1 knockdown on HepG2 and Huh-7 cell proliferation over five days. NC, non-targeting control; Si1 and Si2, independent siRNAs targeting PRDX1. (C) Colony-formation assay and quantification. (D) Transwell migration and invasion images. (E) Wound-healing assay images at 0 and 24 h. (F, G) Quantification of

migration and invasion in HepG2 and Huh-7 cells. (H, I) Relative intracellular lactate levels after PRDX1 knockdown. *P < 0.05, **P < 0.01, ***P < 0.001.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable1.xlsx](#)