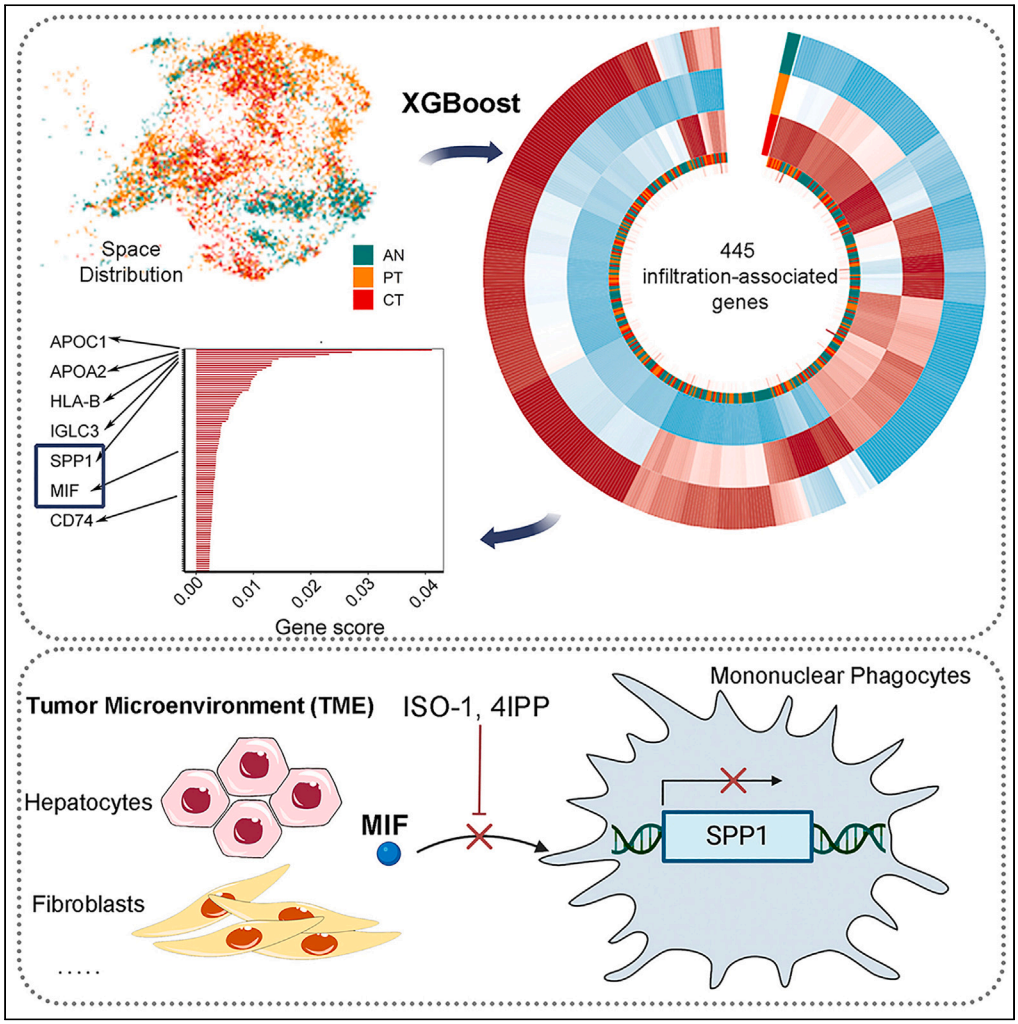
iScience

CellPress
OPEN ACCESS

Article

MIF is a critical regulator of mononuclear phagocytic infiltration in hepatocellular carcinoma



Yunxi Liao, Chenyang Wu, Yang Li, Jinhua Wen, Dongyu Zhao

jhwen@bjmu.edu.cn (J.W.)
zhaodongyu@bjmu.edu.cn (D.Z.)

Highlights

The integrated scRNA-seq data portrait the mononuclear phagocytes in HCC

XGBoost model identified 445 infiltration-associated genes in tumors

MIF is one of the upstream regulators of SPP1 to promote macrophage migration

MIF and SPP1 promote tumor metastasis and invasion

Liao et al., iScience 26, 107273
August 18, 2023
https://doi.org/10.1016/j.isci.2023.107273

iScience

CellPress
OPEN ACCESS

Article

MIF is a critical regulator of mononuclear phagocytic infiltration in hepatocellular carcinoma

Yunxi Liao,[1,2] Chenyang Wu,[1,2] Yang Li,[3] Jinhua Wen,[3,*] and Dongyu Zhao[1,2,4,*]

SUMMARY

Immunotherapy targeting tumor-associated macrophages (TAMs) is a promising approach to treating cancer. However, the limited drug targets and ambiguous mechanisms impede the development of clinical immunotherapy strategies. To elucidate the underlying processes involved in mononuclear phagocyte (MNP) infiltration and phenotypic changes in hepatocellular carcinoma (HCC), we integrated single-cell RNA-sequencing data from 100,030 cells derived from patients with HCC and healthy individuals and compared the phenotypes and origins of the MNPs in the tumor core, tumor periphery, adjacent normal tissue, and healthy liver samples. Using machine learning and multi-omics analyses, we identified 445 infiltration-associated genes and potential drug targets affecting this process. Through *in vitro* experiments, we found that the expression of macrophage migration inhibitory factor (MIF) is the upstream regulator of secreted phosphoprotein 1 (SPP1) and promote migration in TAMs. Our findings also indicate that MIF promotes tumor metastasis and invasion and is a promising potential target for treating HCC.

INTRODUCTION

Hepatocellular carcinoma (HCC) is an aggressive malignancy accounting for >80% of primary liver cancers. HCC is the fourth most common cause of cancer-related death worldwide, posing a serious threat to people's health and lives.[1] In addition to tumor cells, fibroblasts, endothelial cells, and immune cells in the tumor comprise the tumor microenvironment (TME). Tumor-associated macrophages are considered a major component of the TME and are related to patient prognosis and tumor resistance to therapy. Macrophages in the liver are categorized into tissue-resident macrophages (Kupffer cells) and bone-marrow-derived macrophages according to their origin.[2] The TME can induce these macrophages to undergo a series of phenotypic and functional alterations, yielding tumor-associated macrophages (TAMs). Monocytes extravasate into the tissue and interact with the TME during infiltration and differentiation, becoming TAMs that can promote tumor progression.

Mononuclear phagocytes (MNPs) include macrophages and monocytes, and their precursors are dispersed across various organs and tissues throughout the body. It has previously been demonstrated that macrophages can be divided into two groups, M1 and M2.[3] M1 macrophages are considered "classically activated" macrophages that can be activated with lipopolysaccharide and IFN-γ. M1 macrophages express TNF-α, iNOS, IL-1β, and CXCL9/10, recruit T cells, and promote inflammation. M2 macrophages are "alternatively activated" by IL-4 and IL-13. M2 macrophages express CD206, CD163, TREM2, and other anti-inflammatory markers and do not play a pro-inflammatory role.[3,4] However, these two distinct types do not comprehensively represent the highly heterogeneous nature of macrophages[5] such as TAMs, a heterogeneous group of cells expressing both M1 and M2 markers. Many TAMs with immunosuppressive effects exhibit an M2-like phenotype and do not inhibit tumors.[6] These TAMs secrete IL-10, TGF-β, VEGFA, and other tumor-promoting cytokines, hampering T cell cytotoxicity and promoting tumor metastasis.[7] Cancer patient single-cell RNA-seq (scRNA-seq) data have revealed previously unrecognized dynamics and heterogeneity among TAMs.[8–10]

Due to cellular plasticity, numerous MNPs exhibit distinct phenotypes, implying a diverse role in disease. Many studies have now demonstrated that macrophages in the TME promote tumor development[11,12] and targeting TAMs has emerged as a promising direction for drug development and therapeutic strategies.[13] There are two main TAM-targeting approaches: changing the function of the TAMs or inhibiting TAMs infiltration.[14] Several cytokines, including IL-6, IL-1β, CSF-1, and VEGFA, and chemokines CCL2, CCL3, CCL4,

[1]Department of Biomedical Informatics, School of Basic Medical Sciences, Peking University, Beijing 100191, China

[2]State Key Laboratory of Vascular Homeostasis and Remodeling, Peking University, Beijing 100191, China

[3]Department of Cell Biology, School of Basic Medical Sciences, Peking University, Beijing 100191, China

[4]Lead contact

*Correspondence: jhwen@bjmu.edu.cn (J.W.), zhaodongyu@bjmu.edu.cn (D.Z.)

https://doi.org/10.1016/j.isci.2023.107273

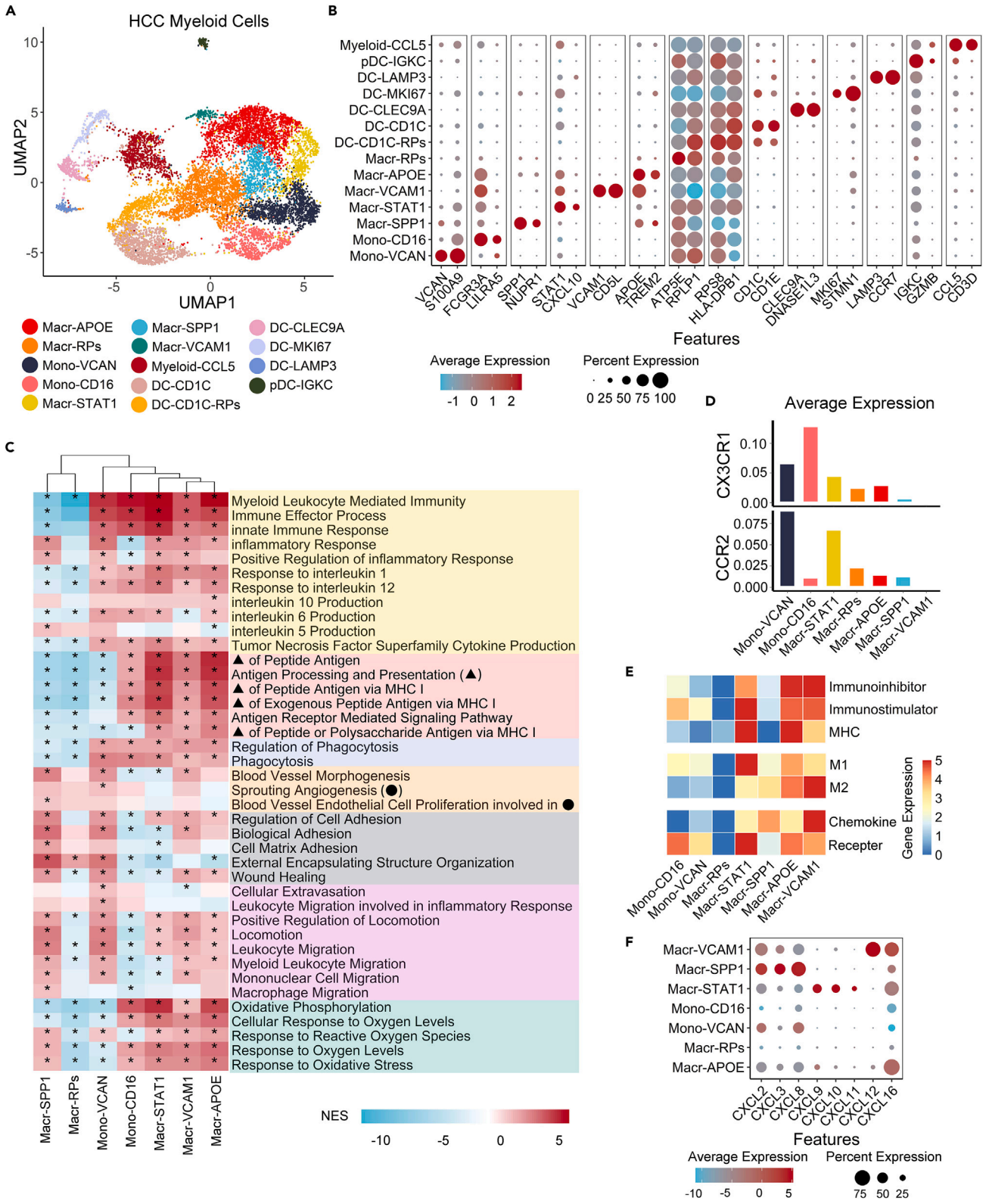


Figure 1. Identification of the MNPs in HCC

(A) UMAP plot showing the clusters of the 10,381 HCC myeloid cells.

(B) Dim plot showing the marker genes expression of the HCC myeloid cells.

Figure 1. *Continued*

(C) Heatmap showing ssGSEA results of the 6,939 HCC MNPs in the BP gene set in the GO term enrichment. Heatmap colors show the normalized gene enrichment scores (NES), and label "*" represents the adjusted p value <0.05. The different color blocks of row names distinguish the pathway types, from top to bottom, the pathway types are "inflammation & immune", "antigen presentation", "phagocytosis", "angiogenesis", "remodeling cell adhesion and matrix", "cell migration", and "oxidative stress". "▲" represents "Antigen Processing and Presentation", "●" represents "Sprouting Angiogenesis".

(D) Bar plot showing the average expression of CCR2 and CX3CR1 in each type of MNPs.

(E) Heatmaps showing the expression levels of cell polarization, chemotaxis, and immune-related genes in each MNP subtypes.

(F) Dim plot showing the expression of CXCLs in each type of MNP.

and CXCL12, have been demonstrated to promote monocyte recruitment in animal models of lung, breast, and pancreatic cancer.[15] This finding has led to the development of many drugs targeting MNP infiltration, predominately targeting the CCR2–CCL2 and CXCR4–CXCL12 pathways.[16] Strategies targeting these pathways have shown efficacy in HCC mouse models.[13]

However, due to the complexity of the TME, additional signaling pathways that mediate intercellular interactions within the TME need to be illustrated. Limited targets and unclear mechanisms impede the development of clinical treatment strategies. To further understand the role of MNPs in HCC and identify potential drug targets influencing MNP infiltration, we integrated single-cell transcriptomic data from patients with HCC and healthy individuals. We used a gradient-boosted decision tree (GBDT) approach to distinguish different spatial locations of MNPs. A multi-classifier was trained to identify genes affecting the spatial distribution of cells, and the effects of these genes were further confirmed with microarray and RNA-seq data. Through data mining and experimental validation, we demonstrated that the migration inhibitory factor (MIF) gene could influence TAM infiltration, providing a potential target for treating HCC.

RESULTS

MNPs show phenotype and function heterogeneity in HCC

To better understand the heterogeneity of the immune system in liver cancer, we evaluated data from 75,696 immune cells from two published HCC single-cell datasets, Zhang Q et al. and Sharma A et al. The cells were obtained from distinct spatial regions: adjacent normal (AN) tissue, the tumor periphery (TP), and the tumor core (TC) (Figures S1A–S1C). We divided the cells into 12 groups based on the expression of classical immune cell marker genes, such as CD8A, IL-7R, TCLA4, IFG, and CD79A (Figures S1D, S1E, and Table S1). We further focused on myeloid cells, comprising dendritic cells (DCs), monocytes, and macrophages, and classified the cells into 13 subtypes, including seven MNP subtypes, five conventional dendritic cells subtypes, and one plasmacytoid dendritic cells subtype after re-clustering (Figure 1A and 1B). To exclude the effect of ribosomal proteins (RPs), we compared cell clusters before and after ribosome removal and found that DC-CD1C-RPs belonged to the DC-CD1C subtype and Macr-RPs represented a separate subtype (Figures S1F and S1G).

Gene Ontology (GO) enrichment analysis[17] was performed for the 7 MNP subtypes (Figure 1C). The Mono-VCAN subtype has high motility and is enriched in the cell extravasation pathway; thus, these cells might migrate during the inflammatory response. In contrast, the Mono-CD16 subtype with low motility exhibited a more mature phenotype (Figure 1C). The Macr-STAT1 and Macr-APOE subtypes have a strong antigen-presenting functions and are significantly enriched in myeloid-mediated immunity and immune effects (Figure 1C). The Macr-APOE subtype expressed M2 markers, such as CD163, MRC1, and IL-10, and classical complement C1q (C1QA/B/C) (Figure S1H). It has been reported that tumors with a high number of macrophages and a high level of C1q expression establish an immunosuppressive environment.[18] Another type of immunosuppressive macrophage was the Macr-secreted phosphoprotein 1 (SPP1) subtype. These cells strongly expressed SPP1 and TREM2 (Figures 1B and S1H). The Macr-SPP1 subtype demonstrated overexpression of pathways involved in cell adhesion, extracellular matrix remodeling, and angiogenesis (Figure 1C).

To identify direct factors affecting MNP infiltration, we evaluated the expression of chemokines and their ligands in each cell subtype. The inflammatory Mono-VCAN subtype showed the highest level of CCR2 expression, indicating the initial event of $CX3CR1^{Lo}CCR2^{Hi}$ monocytes recruitment from the vasculature to the tumor; infiltrating monocytes are then further transformed into $CX3CR1^{Hi}CCR2^{Lo}$ monocytes (the Mono-CD16 subtype) (Figure 1D). The $CCR2^{+}$ Macr-STAT1 subtype expressed many chemokine receptors and M1 marker genes, suggesting that the Macr-STAT1 subtype is present in the early stages of monocyte differentiation into macrophages (Figures 1D and 1E).

The Macr-SPP1 subtype was also highly enriched in cell migration-related pathways and had low expression of chemokine receptors but higher chemokine levels. These cells primarily secrete CXCL1/2/3/8 (Figure 1F), which are CXCR1/2 ligands. The aforementioned chemokines primarily recruit myeloid cells, such as granulocytes, promote the development of immunosuppressive cells, and inhibit T and NK cell cytotoxicity.[19,20] In addition, secretion of IL-5 by Macr-SPP1 cells may also inhibit NK cell function by recruiting eosinophils.[21] Due to the low enrichment scores in the immune-related pathways (Figure 1C), Macr-SPP1 cells may not secrete inflammatory factors such as IL-6 and TNF-α. Overall, we distinguished immunosuppressive and activated MNP subtypes in the tumor microenvironment using their distinct gene expression characteristics.

MNPs show spatial distribution preferences

We investigated the cells' origin by using scRNA-seq of MNPs from healthy liver datasets (MacParland et al.[22], Tamburini et al.[23], and Ramachandran et al.[24]) (Figure S2A). In contrast to tissue-resident Kupffer cells, monocytes and mononuclear-derived macrophages are collectively referred to as "MDM" in healthy livers (Figure S2B). VCAM1, MARCO, VSIG4, CD163, and other markers of Kupffer cells were highly expressed in the KC-VCAM1 subtype. A small fraction of cells, MDM-CD14 cells, expressed CD14, ITGAM (CD11b), and FCN1 and may be located in the peripheral blood. In addition to classification by tissue origin, MDMs can be classified as inflammatory or non-inflammatory. Compared to MDM-CD14 cells, MDM-VCAN cells express higher levels of CCR2 and inflammatory factors (e.g., S100A8 and S100A9) (Figure S2C). MDM-VCAN cells account for the vast majority of MNPs in the liver.

Subsequently, we examined the similarities among cellular identities in HCC and healthy liver samples by directly calculating correlation coefficients and using a logistic regression model with elastic net regularization for each cellular identity defined by the clusters identified in each sample type.[25] In this study, Macr-SPP1 and Macr-STAT1 were two novel mononuclear-derived macrophage subtypes that evolved in HCC with little connection to a specific cell subtype found in healthy livers (Figure 2A). Additionally, we found that Macr-VCAM1 and KC-VCAM1 cells are quite similar (Figures 2A and 2B). A small cluster of Macr-VCAM1 cells prominently expressing Kupffer cell markers was found at the margin of the Macr-APOE cluster. The distribution of most Macr-APOE cells overlapped with KC-VCAM1 cells in the UMAP descending analysis, indicating that Macr-APOE cells are derived from Kupffer cells, while Macr-APOE cells were derived from the mononuclear-derived Macr-STAT1 subtype (Figure 2C).

The proportion of non-inflammatory Macr-APOE, Macr-RP, and Macr-SPP1 cells is significantly enhanced in HCC samples compared to in healthy liver samples, increasing from 35% in the AN tissue to 70% in the TP and 71% in the TC, respectively (Figures 2C–2E). Under the influence of the TME, the ratio of Mono-CD16/Mono-VCAN cells increased from the AN site to the TC, and infiltrating monocytes lost their inflammation-related phenotype. Macr-STAT1 cells with an M1 phenotype were mainly found at the TP, with low numbers in the TC (Figure 2C), supporting the idea that macrophages display more M2 characteristics during tumor infiltration.[6] Ro/e, the ratio of observed cell number over the expected cell number of a given combination of MNPs cell cluster and spatial distribution, was calculated to determine changes in cellular distribution among spatial locations (Figure 2F). These results indicated that infiltrating MNPs are influenced by the TME in different locations and show spatial distribution preferences.

MIF is the main secretory signal for MNPs interacting with the microenvironment

To investigate the environmental signals affecting MNP infiltration, we reintegrated non-immune cells (hepatocytes, endothelial cells, and fibroblasts) from Sharma A et al. (Figure 3A). Cell-cell commutation with CellChat revealed that non-immune cells primarily produce MIF, annexin, visfatin, and $C3^{+}$ complement to interact with MNPs (Figures 3B and S3). It has been reported that C3 complement can promote M2 polarization.[6,26] Tumor hepatocytes may generate C3 complement, attracting $C3AR^{+}$ TAMs (Macr-APOE) (Figure S3). Notably, MIF-related pathways were present in all three types of samples and varied across adjacent normal and malignant tissues (Figure S3). The primary source of MIF in the TME was fibroblast cells, along with autocrine signaling from MNPs, whereas in AN sites, hepatocytes are the primary non-immune cells that secrete MIF (Figures 3C, and S3). Given the widespread expression of MIF receptors CD74 and CXCR4 in MNPs in HCC, the MIF/CXCR4 axis may be the primary signaling pathway promoting macrophage infiltration (Figure 3C), whereas the other ligand of CXCR4, CXCL12, was detected only in endothelial cells in AN sites with low intensity (Figure S4). The expression levels of MIF were higher in Macr-SPP1 cells than in the other MNPs (Figures 3C and 3D). Both signals reached maximum intensity (Figure S4), and their signaling network structures were similar in the tumor sites (Figure 3E), suggesting these two genes have potential correlations. In summary, the

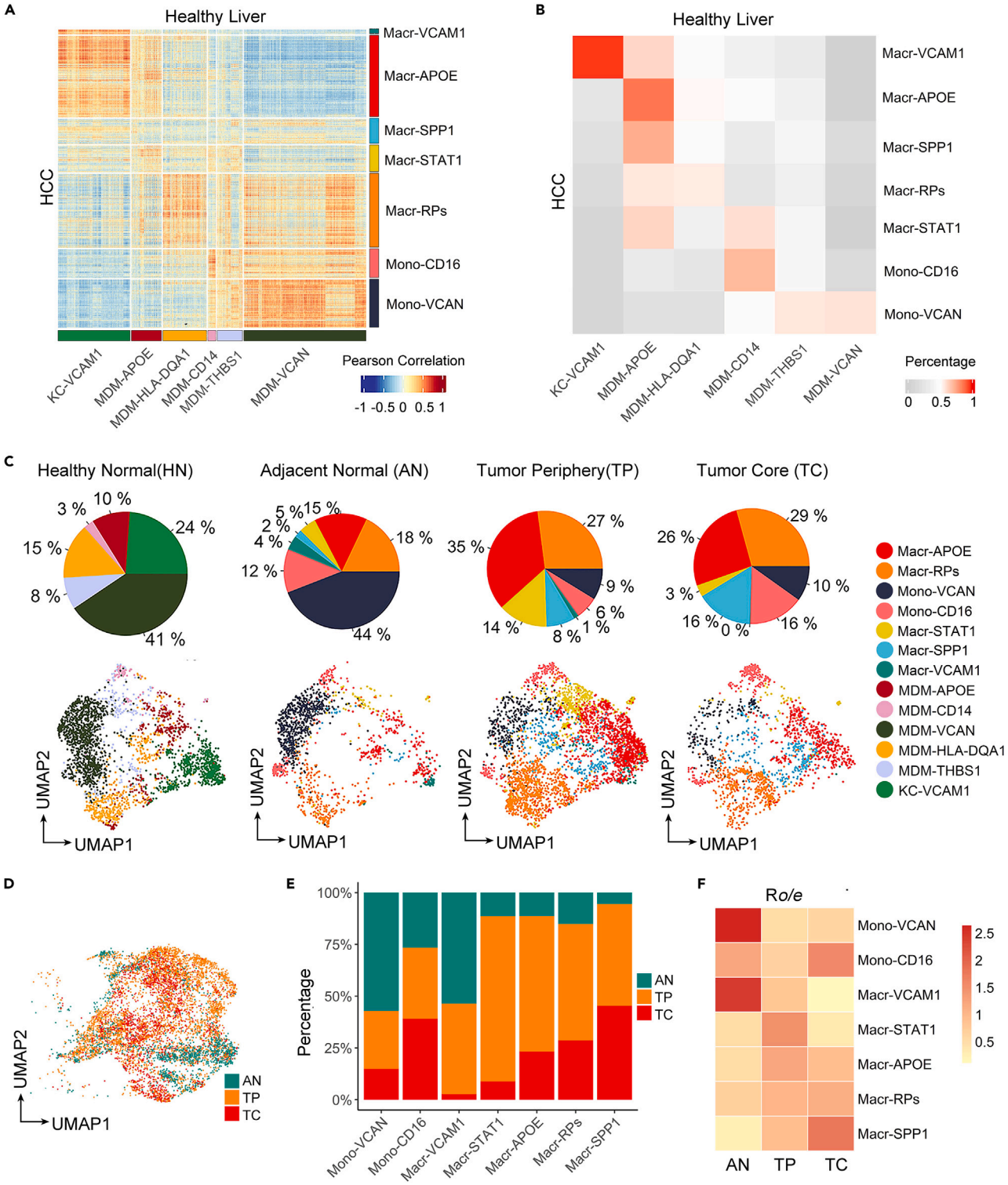


Figure 2. Differences in the MNPs in the different spatial locations

(A) Heatmap showing the correlation of MNP types in healthy livers (columns) and HCC (rows).

(B) Heatmap showing the similarity between different MNP types in healthy livers (columns; 3,039 cells) and HCC (rows; 6,939 cells).

(C) Pie charts and UMAP plots showing the spatial distribution and proportion of each MNP subtype.

(D–E) UMAP plot and bar plot showing the spatial distribution of the MNPs in HCC.

(F) Heatmap showing the spatial preference of each cell type estimated using the *Ro/e* method.

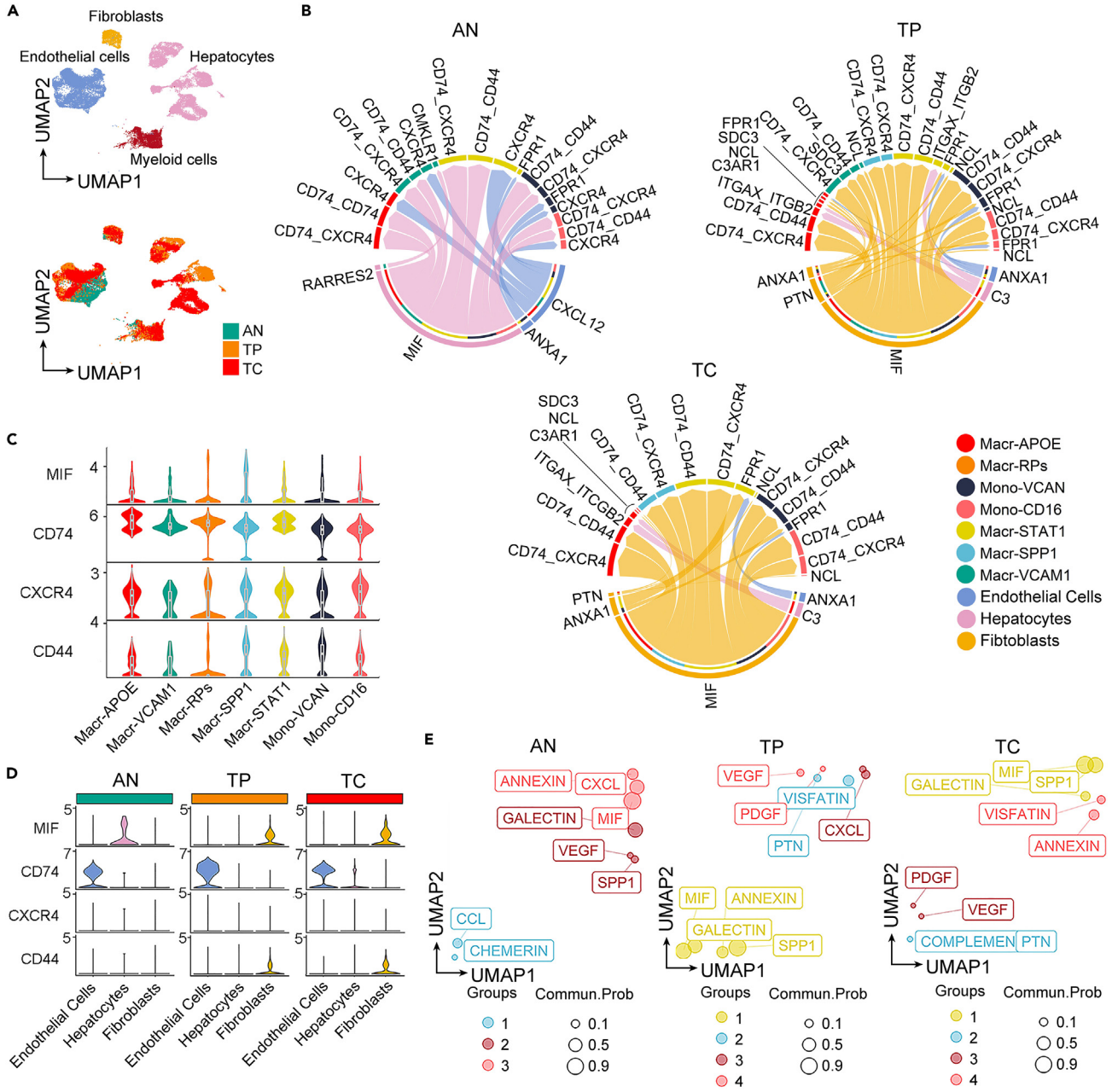


Figure 3. Analysis of cell communication through secretory signals by CellChat

(A) UMAP plots showing the 21,295 non-immune cells and 6,939 MNPs and their spatial location used to run CellChat.

(B) Chord diagrams showing the secretory signals predicted by CellChat for the communication between non-immune cells and MNPs.

(C) Violin plots showing the RNA expression levels of ligand-receptor pairs of MIF-related pathways in MNPs.

(D) Violin plots showing the RNA expression levels of ligand-receptor pairs of MIF-related pathways in the non-immune cells of different locations.

(E) UMAP plots showing the structure similarity of the secretory signals in the different spatial locations.

disparities in cellular spatial distribution and communication signals provide a snapshot of different macroenvironments, and the TC possesses more immunosuppressive factors than AN sites.

The cellular trajectory in tumor sites reveals MNP infiltration processes

To explore how monocytes transform during infiltration into tumors, we deduced the cellular infiltration and differentiation process using analysis of pseudotime and RNA velocity. To prevent batch effects arising

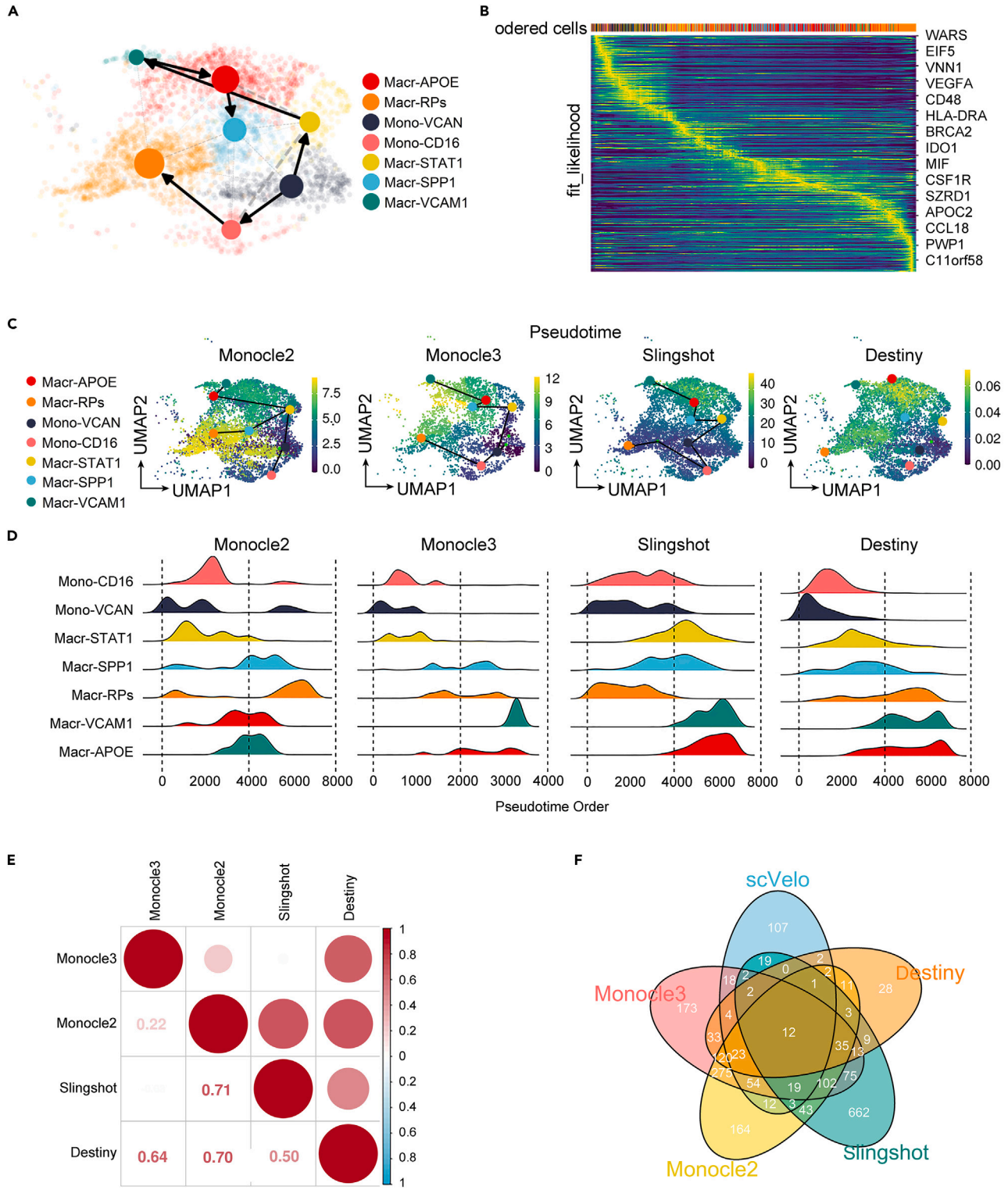


Figure 4. Cell trajectory analysis of MNPs during infiltration in HCC

(A) PAGA plot showing the predicted trajectory by scVelo.

(B) Heatmap showing the top 100 genes in the molecular dynamic model most likely to affect differentiation based on latent time for RNA splicing.

Figure 4. *Continued*

(C) UMAP plots showing the pseudotime by four TI methods. The trajectory is simplified according to the predicted trajectory inferred by the TI methods. The colorful dots in the trajectory represent the cells where the median pseudotime of each MNP subtype and the small green dots represent the starting cells with pseudotime 0.

(D) Density plots showing the cell pseudotime order inferred by the TI methods.

(E) Heatmap showing the Pearson correlation coefficient when compared to the pseudo-time order of cells estimated by the TI methods.

(F) Venn diagram showing the driver genes found in the five-cell trajectory.

from using multiple HCC datasets, we selected one dataset (Sharma A et al.) for RNA velocity analysis[27] and discovered that the cells were divided into two branches derived from the Mono-VCAN subtype, monocyte-like and macrophage-like progressing from M1 (Macr-STAT1) to M2 (Macr-APOE) (Figure 4A and 4B). Additionally, the RNA velocity between the Macr-VCAM1 and Macr-APOE subtype flow confirmed that Macr-APOE cells are partially derived from Kupffer cells (Figure 4A).

Within the trajectory inference, we set the CX3CR1^{Lo}CCR2^{Hi} Mono-VCAN subtype as the root cells (Figure 4C). Although Macr-VCAM1 cells are consistent with the tissue-resident Kupffer cells in HCC and are not involved in monocyte infiltration, we included these cells in our analysis to obtain a full overview of the MNPs. Monocle2 (Figures S5A–S5C), Monocle3 (Figure S5D),[28] Slingshot,[29] and Destiny (Figures S5E–S5H)[30] produced similar pseudotime orders (Figures 4C–4E), in agreement with the RNA velocity results. Overall, the trajectory illuminates the progression of inflammatory monocytes (abundant in the AN sites) to M1 macrophages (accumulating at the TP) and immunoinhibitory TAMs, consistent with MNPs' spatial preference (Figure 2C), indicating a dynamic process of MNP infiltration in the tumor.

Machine learning determined the 445 infiltration-associated genes in MNPs

We used a machine learning approach to identify the main factors influencing MNP infiltration. Because cellular trajectories reveal this dynamic process of infiltration, we used 2026 essential genes obtained from RNA velocity analysis and pseudotime inference as model features (Figures 4F, and S5). The expression matrix of these 2026 genes was used as the input for the model, and three spatial sites were used as the output. This approach was used to identify the critical genes influencing the MNP cells' distribution. We employed three models based on GBDT, Catboost,[31] LightGBM,[32] and XGBoost[33] (Figure S6A). Because single-cell data are high-dimensional sparse matrices, we first employed Boruta software to perform feature selection by importance rating, lowering the gene number from 2026 to 446 (Figures S6B and S6C).

The performance of the models was similar with this dataset using the selected hyperparameters, and the built ternary classifiers had an accuracy of 85% on the same test set after hyperparameter tuning (Figure S7A). We selected the XGBoost model (Figures S7B and S7C) as it demonstrated the highest accuracy (85.7%), and high AUC_{ROC} and AUC_{PR} values (Figures 5A and 5B, Table 1). The accuracy reached 94.6% when the normal and malignant tissues were separated (Figure S7A). Intriguingly, both SPP1 and MIF were listed at the top when ranking the importance of the model's features (Figure 5C). The model was externally verified using scRNA-seq datasets (Cheng S et al.) containing other HCC and alternate tumor types.[10] Due to the lack of distinction between the TP and TC in the other datasets, we could only use the model to produce binary (normal vs. malignant) predictions, and the model performance was good (Figures 5D and 5E). GO term enrichment analysis of the selected genes revealed that MNP chemotaxis and migration were involved in MNP infiltration (Figures 5F and 5G).

The infiltration-associated genes enable the classification of TME types

To determine whether the infiltration-associated genes identified through machine learning can be used to characterize the MNP-related microenvironment in tissues, we calculated the highest expressed gene in each tissue and divided the 445 genes into three corresponding categories (AN, TP, and TC) (Figure S7B, Table S5). In addition, we integrated two sets of bulk transcriptome data from patients with HCC, including microarray data[34–41] and data from the TCGA[42] and ICGA[43] (Figures S8A–S8F). The tumor tissues were grouped into three groups (AN, TP, and TC) by comparing the weighted average of gene expression. When combined with patient survival data, we discovered that AN-type patients have a better prognosis than TP- or TC-type patients (Figures 6A–6C), as the aforementioned spatial distribution preferences of MNPs (Figures 2C–2F) in AN imply an immune microenvironment with a high proportion of monocytes (especially inflammatory monocytes) and a low proportion of macrophages. Regarding the TNM stage and Edmondson–Steiner grade, most patients with stage I and G1-2 had AN-type HCC (Figure 6D). The AN-type microenvironments are low-risk factors.

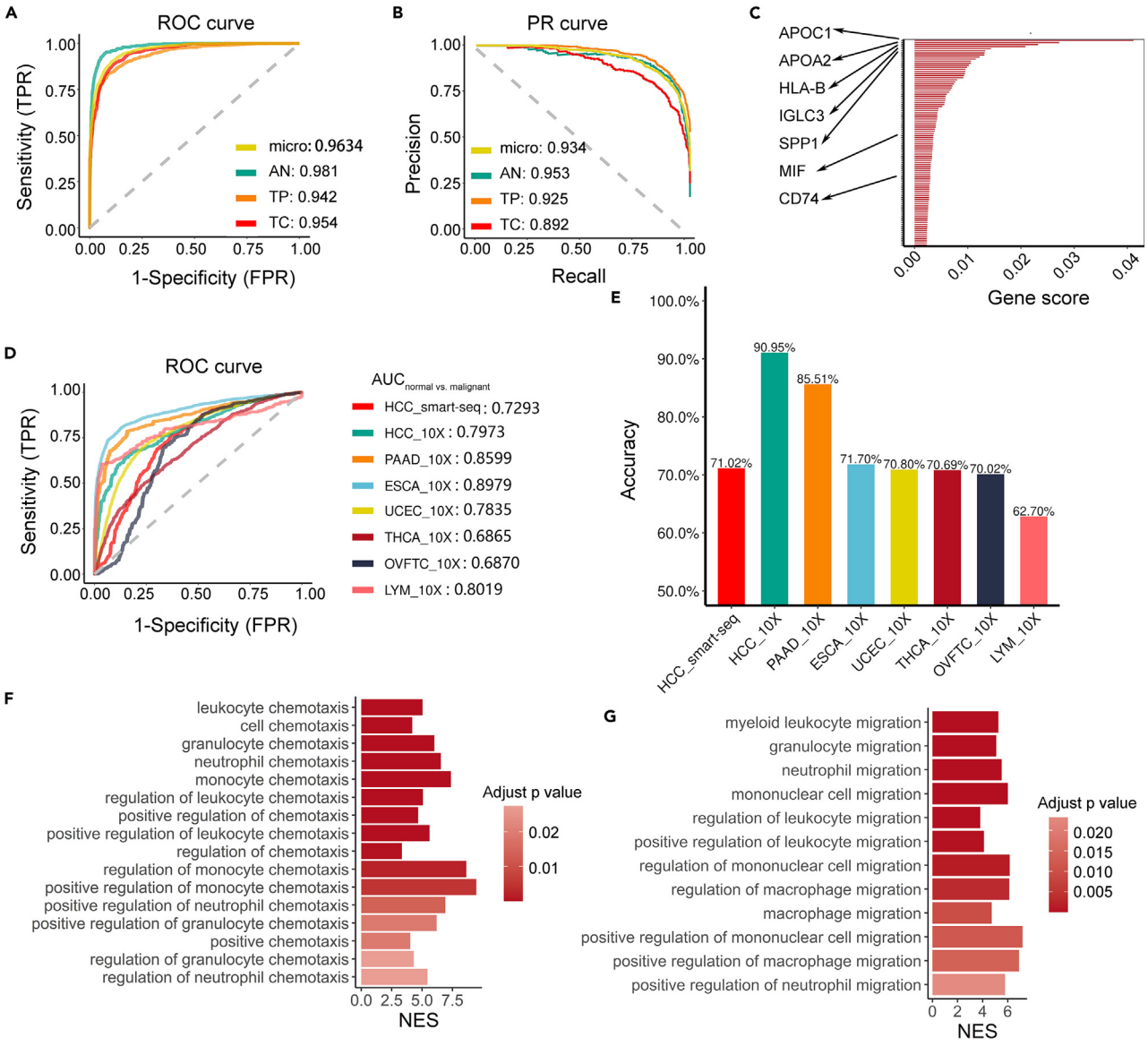


Figure 5. Predicting the spatial location of the HCC MNPs using the XGBoost model

(A) ROC curves of the selected XGBoost models.

(B) PR curves (precision-recall curve) of the selected XGBoost model.

(C) Bar plot shows the importance scores ranking of the feature genes in the XGBoost model.

(D) ROC curves of the selected XGBoost models on the external validation datasets.

(E) Bar plot representing the accuracy of the XGBoost model as a binary classifier in external validation.

(F and G) Bar plot representing the pathways associated with cell chemotaxis (F) and migration (G) enriched in GOBP for the 446 feature genes selected by Boruta.

Next, we compared the three categories of genes with differential genes between tumor and normal tissue (Figure 6E). AN genes and low-expressed genes in the tumor were regarded as low-risk genes, and TP or TC genes and high-expressed genes in the tumor were considered high-risk genes. There was a significant overlap within the high-risk genes (Figure 6F), considered tumor-related genes affecting the immune microenvironment. The top-ranked genes in terms of feature importance, SPP1 and MIF, overlapped with high-risk genes and were highly expressed in tumor tissues, indicating poor prognosis (Figures S9A and S9B). These two genes were negatively correlated with $CD8^+$ T cell infiltration in tumors and positively correlated with the degree of monocyte and macrophage infiltration (Figure 6G). In summary, genes that affect MNP infiltration, especially SPP1 and MIF, influence the tumor microenvironment and cancer prognosis.

Table 1. XGBoost model evaluation metrics

	Triple classifier		Binary classifier		
	Macro average	Micro average	AN	TP	TC
AUC_{ROC}	0.959	0.9634	0.981	0.942	0.954
F1 score	0.849	0.857	0.866	0.879	0.802
AUC_{PR} (mAP)	0.923	0.934	0.953	0.925	0.892

MIF is a critical factor that regulates SPP1 expression in TAMs

To further explore the relationship between MIF and SPP1, we explored the correlation of MIF and SPP1 in TCGA, ICGC, and tissue microarray datasets and found that MIF and SPP1 demonstrate high correlation (Figure S9C). We selected 50 samples with the highest MIF expression and 50 with the lowest MIF expression as the MIF-high-variable samples from the TCGA and ICGC datasets, respectively, and performed a differential gene (DEG) analysis. We found that in the MIF-high-variable datasets, SPP1's logFC was similar to that of MIF; however, in the SPP1-high-variable datasets, MIF was not significantly overexpressed (Figure S9D). In addition, the common DEGs between MIF- and SPP1-high-variable samples substantially overlapped (Figure S9E). These data indicate that these two genes might share a common regulatory pathway and that MIF is a possible upstream regulator of SPP1.

Then, we developed an *in vitro* model of TAMs induced by HCC tumor-conditioned media (TCM-TAMs) (Figure S10A) to further elucidate the possible regulatory relationship between MIF and SPP1 within MNPs in the TME. The TCM-TAMs express M1 and M2 marker genes (CD86, CD80, CD206, and CD163) and associated cytokines (IL-1β, IL-10, TNF-α, and TGF-β) in a dose-dependent manner as the TCM concentration increases (Figure S10B). To test the role of MIF in the induced macrophages, we treated the TCM-TAMs with MIF inhibitors 4-IPP and ISO-1 and found that suppression of MIF decreased SPP1 expression in both mRNA and protein levels (Figures 7A and 7B). 4-IPP and ISO-1 help inhibit the migration of TAMs. Conversely, inhibition of MIF does not impair cell proliferation (Figure S10C). These data suggest that inhibition of MIF is a potential approach to reducing TAM migration.

We analyzed the effects of 20 μM 4IPP treatment and SPP1 knockdown on TAMs using RNA-seq, which resulted in reduced levels of SPP1 mRNA expressed by TAMs (Figure 7E). The corresponding differential genes were identified after inhibiting the MIF pathway activation and SPP1 expression, respectively (Figure S10D), and were found to show a large overlap (Figure 7F). Notably, the knockdown of SPP1 did not result in decreased MIF expression (Figure 7G). Moreover, MIF-activated genes were found to be enriched in SPP1-activated genes, while MIF-repressed genes were enriched in SPP1-repressed genes, indicating that SPP1 is downstream of MIF (Figure 7H). Gene set enrichment analysis in KEGG revealed that the DEGs obtained from the MIF or SPP1-high-variable datasets are enriched in similar pathways (Figure S11A). MIF and SPP1 regulate multiple metabolic pathways, including fatty acid metabolism, and enhance extracellular matrix receptor interactions and basal cell carcinoma-related pathways (Figures S7A and S7B). The DEGs between 4IPP-treated or SPP1-KD TAMs with respective control were enriched in many pathways associated with cell migration (Figure 7I). Therefore, MIF and SPP1 share common downstream targets and may regulate macrophage migration through some common pathways.

We also constructed a gene regulatory network of cancer-associated transcription factors and shared DEGs in MIF- or SPP1-high-variable datasets in the TCGA or ICGC. The most central target gene in the regulatory network was PRSS22 (brain-specific serine protease 4), which has been reported to promote liver cancer cell migration and invasion *in vitro* and *in vivo*[44]. Other genes associated with tumor metastasis and invasion, such as PROM1,[45] B3GALT5,[46] RHOV,[47] WNT7B,[48] and EPN3,[49] were also identified (Figure S11B). All nodal genes are enriched in extracellular matrix regulation, cell adhesion and migration, and other pathways related to tumor invasion (Figure S11C). The degradation of the extracellular matrix is the first step of tumor invasion; tumor cells subsequently invade the basement membrane and surrounding stroma. Since the expression level of matrix metalloproteinases was significantly reduced in 4IPP-treated or SPP1-KD TAMs (Figure S10D), we performed tumor invasion experiments on HepG2 cells using SPP1 knockdown or MIF-suppressed TAMs as variables and found that MIF and SPP1 indeed inhibit tumor invasion (Figure S11D). Overall, MIF and SPP1 may enhance the aggressiveness of tumors by regulating the aforementioned cancer-related genes in HCC, promoting tumor migration and invasion.

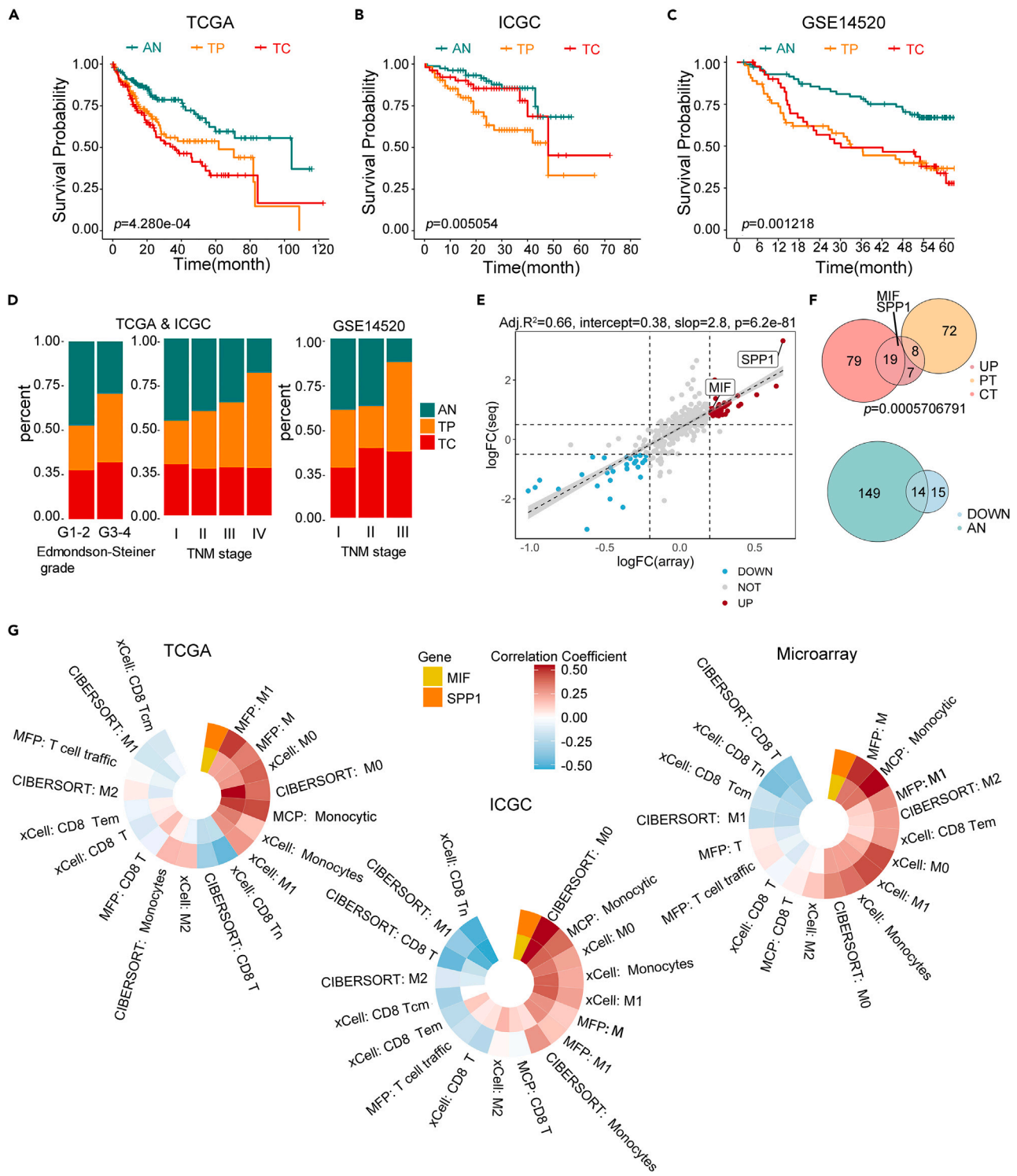


Figure 6. The infiltration-associated genes from the model training can predict HCC patient prognosis

(A–C) Overall survival rates of patients with HCC with different microenvironment types divided by the infiltration-associated genes. (A) AN = 142, TP = 113, TC = 109; (B) AN = 99, TP = 78, TC = 65; (C) AN = 90, TP = 67, TC = 52.

D) Stacked bar plot of the TNM stage (I, II, III, IV) and Edmondson–Steiner grade (G1-G2 and G3-G4) by microenvironment types.

(E) Scatterplot showing gene expression fold changes (tumor vs. normal) of integrated microarray data and RNA-seq data.

Figure 6. *Continued*

(F) Venn diagrams showing the overlap between the genes upregulated in tumors and the TP or TC genes. The p value of the hypergeometric test is shown.
(G) Heatmaps showing the Pearson correlation coefficient between MIF or SPP1 and the gene signature scores related to $CD8^+$ T cells and MNPs calculated by the MCP-counter algorithm, CIBERSORT algorithm, xCell method, and MFP method.

DISCUSSION

In the study, we obtained important genes involved in the infiltration process of MNPs after model training. The highly ranked genes demonstrated high expression in the TC MNPs, and many are induced by hypoxia, including various heat shock proteins, such as FTL,[50] NUPR1,[51,52] TXNIP,[53] SOD1,[54] HMOX1,[55] and PRDX1,[56] indicating that MNPs are under oxidative stress (Figure 2A). Additionally, several genes reported to promote tumor development, such as ORM1,[57] S100A10,[58,59] AIF1,[60] TGFBI,[61] and TREM2,[62] were included in the feature genes highly expressed in the MNPs of the TC.

At the tissue level, the infiltration-associated genes identified both represent MNPs and correlate with $CD8^+$ T cell infiltration. Their expression reflects, to some extent, the whole TME due to the variable interactions, correlating with prognosis. This finding further illustrates that these 445 important genes play critical roles in tumor development, particularly MIF and SPP1. Some of these genes are less reported, and their functions in tumor immunity require further exploration.

MIF is expressed in various cell types[63] and was initially thought to inhibit macrophage migration.[64] Subsequent studies have revealed that MIF recruits and promotes macrophage infiltration in tumors[65] and induces macrophage M2-type polarization.[66] In contrast to earlier studies, our findings suggest that the fibroblast cells, together with autocrine signaling, are the primary source of MIF in the TME. We identified hepatocytes in AN sites as the primary non-immune cells that secrete MIF (Figures 4D and 4E). Our *in vitro* experiment confirmed that SPP1 expression is influenced by the MIF pathway and that MIF promotes TAM migration. Collectively, these results suggest that MIF is a promising drug target and diagnostic indicator in HCC. 4-IPP is a specific suicide substrate of MIF that can irreversibly bind to MIF covalently and inhibit its biological activity.[67] ISO-1 is a competitive inhibitor of MIF1's tautomerase activity, and both ISO-1 and hydroxyphenylpyruvate, a substrate of MIF, can bind to the same site of the MIF tautomerase active site.[68] We simulated the blockade of MIF activation on corresponding receptors on macrophages in the tumor microenvironment by adding MIF inhibitors to the cell culture medium. For the first time, we identified an association between MIF and SPP1 expression in both MNPs and tumor tissue. Furthermore, we determined that MIF is a possible regulator of SPP1, promoting macrophage infiltration and tumor invasion.

Osteopontin (OPN) is the final product of the SPP1 gene. Recently, many studies have focused on the role of OPN in the tumor microenvironment because it is considered a key molecule in promoting tumor development. OPN has secreted (sOPN) and intracellular (iOPN) forms with different functions depending on their location inside or outside the cell; they are translation isoforms and are thus difficult to isolate.[69] Our results demonstrate that MIF affects the level of SPP1 mRNA from the transcriptional level, and MIF inhibitors can reduce the total OPN protein and sOPN, but it is challenging to precisely determine whether iOPN is affected.

A recent study revealed that MIF knockdown in lung tissue in rats slowed the progression of bleomycin-induced pulmonary fibrosis and decreased SPP1 expression.[70] However, the underlying mechanism by which MIF regulates SPP1 remains unknown and requires further exploration. It has been observed that 4-IPP and ISO-1 inhibit the p65 pathway, and a review summarized the upstream regulatory genes of SPP1, revealing that both RUNX2 and STAT3 can bind to the SPP1 promoter and upregulate transcription.[71] MIF may regulate OPN expression in macrophages through the p65/RUNX2 or STAT3 pathways. MIF originates from various cells of the tumor microenvironment, especially fibroblasts, and serves as a signal for environmental stimulation of mononuclear macrophages, whereas SPP1 is mainly expressed by macrophages and is a more direct regulator of migratory capacity and pro-tumor-invasive ability. Additionally, CD44 participation is required for MIF activation of the CD74 receptor, and an sOPN domain exists that can bind to CD44.[69] Matrix metalloproteinases, another important CD44 ligand, are commonly upregulated in many cancers and play an important role in promoting tumor angiogenesis, progression, invasion, and metastasis.[72] Binding of CD44 on macrophage membranes induced by sOPN and THP-1 activates p65-related pathways and promotes cell migration.[73] Thus, sOPN may regulate macrophage phenotype in conjunction with MIF by activating CD44.

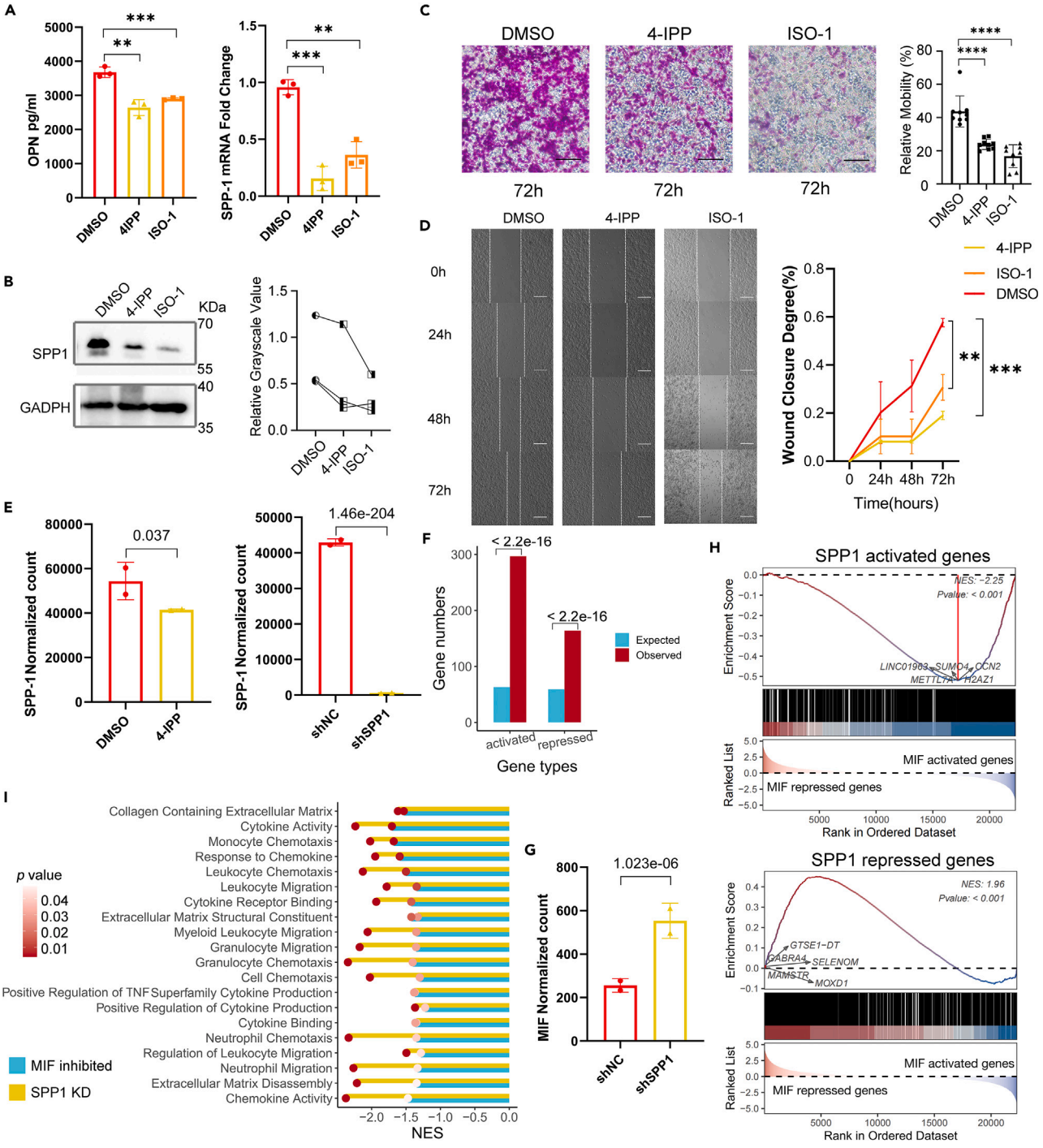


Figure 7. MIF can regulate the expression of SPP1 in the TAMs and promote cell migration

(A) Bar plots showing osteopontin protein levels in TAMs supernatants (left) and SPP1 mRNA levels (RT–qPCR, right) in the TAMs treated with DMSO, 20 μM 4-IPP, or 5 μM ISO-1. The data are represented as mean ± SD. The p values of the unpaired Student's t-test are shown, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$.

(B) Western blotting results showing the expression of SPP1 and MIF in TAMs treated with DMSO (control), 20 μM 4-IPP, or 5 μM ISO-1. The line chart shows the relative gray value calculated by ImageJ.

(C) Transwell assays for the migration of TAMs treated with DMSO, 20 μM 4-IPP, or 5 μM ISO-1. The migration time is 72 h. The length of the black lines in the figure represents 50 μm. The data are represented as mean ± SD. The p values of the unpaired Student's *t* test are shown, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$.

Figure 7. *Continued*

(D) Wound healing assays for the migration of TAMs treated with DMSO, 20 μM 4-IPP, or 5 μM ISO-1. The pictures showing the wound closure after scratching the wounds for 0, 24, 48, and 72 h. The length of the solid white lines in the figure represents 100 μm. The data are represented as mean ± SD. The p values of the paired Student's *t* test are shown, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$.

(E) Bar plots showing the mRNA expression levels of SPP1 measured by RNA-seq. The data are represented as mean ± SD. The p values of differential gene analysis by edgeR are shown.

(F) Bar plot showing the results of the Fisher test for MIF- and SPP1-activated/repressed genes.

(G) Bar plots showing the mRNA expression levels of MIF measured by RNA-seq. The data are represented as mean ± SD. The p values of differential gene analysis by edgeR are shown.

(H) GSEA showing enrichment of the SPP1-activated and -repressed genes in genes regulated by MIF in TAMs.

(I) GSEA showing the enrichment of the DEGs between 4IPP-treated or SPP1-KD TAMs and control on migration-associated pathways.

Sorafenib and lenvatinib are established as first-line drugs for the treatment of hepatocellular carcinoma. Recent studies have demonstrated that cancer-associated fibroblast-derived SPP1 promotes resistance to sorafenib and lenvatinib in HCC through the activation of oncogenic signals and the promotion of epithelial-to-mesenchymal transition.[74] While rarely reported in HCC, tyrosine kinase inhibitors can also impact the MIF signaling pathway. For example, studies have shown that the tyrosine kinase inhibitor genistein can inhibit the upregulation of MIF mRNA in primary synovial fibroblasts.[75] Moreover, in non-small-cell lung cancer, MIF expression can influence the therapeutic efficacy of epidermal growth factor receptor-tyrosine kinase inhibitors like osimertinib and gefitinib. Targeting MIF, a widely expressed secreted protein in various cells, presents challenges in achieving specificity. Various MIF inhibitors have been developed, including ISO-1, 4-IPP, and IPG1094. IPG1094 has undergone evaluation in clinical trials and received approval from the US FDA for the treatment of solid tumors, hematological tumors, psoriasis, and multiple sclerosis (ClinicalTrials.gov Identifier: NCT05112159). The clinical application of MIF inhibitors is still in its early stages, and more clinical trials are needed to evaluate their safety and efficacy.

In conclusion, our study analyzed the relationship between the origin of macrophage cells in healthy livers and HCC in detail, revealed the dynamics of MNPs in the tumor environment through cell trajectory analysis, identified genes affecting MNP infiltration and differentiation, and screened the genes for prognostic applications. We revealed a previously unknown link between MIF and SPP1 in HCC, providing new potential targets for tumor immunotherapy targeting macrophages.

Limitations of the study

Our study provides new targets and directions for immunotherapy of hepatocellular carcinoma, with MIF and SPP1 being identified as key factors in promoting macrophage infiltration and tumor invasion. We agree that our study fails to demonstrate a specific molecular mechanism for this regulatory relationship. Furthermore, our data suggest that MIF affects the overall protein levels and sOPN levels of OPN *in vitro*, as shown by our western blot and ELISA experiments. However, it is still not possible to know exactly if iOPN is affected. Here, further studies need to be done to explore the regulatory mechanisms between MIF and SPP1.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell culture and treatment
- METHOD DETAILS
  - Plotting cell growth curves
  - Reverse transcription qPCR analysis
  - Western blot analysis
  - ELISA
  - Macrophage migration assays
  - Tumor invasion assay
  - Lentiviral transfection experiments

- Library construction, quality control and sequencing for RNA-seq
- Integration and processing of the 10x myeloid single cell dataset
- Similarity analysis of cell types between health livers and HCC tissues
- Determination of the spatial preference of MNPs
- Analysis of cell-cell communication
- Distinguishing polarization, chemotaxis, and immune-related phenotypes
- Inference of cell trajectories
- Data preprocessing for training machine learning models
- Model construction
- Model evaluation
- Integration of the bulk RNA-seq and microarray datasets
- Feature gene classification and sample typing
- Estimation of immune cell infiltration
- Analysis for RNA-seq data
- Differential expression analysis for tumor and normal samples
- Construction of gene regulatory networks
- Survival analysis
- Functional enrichment analysis

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107273.

ACKNOWLEDGMENTS

We thank Huanran Tan for providing the THP-1 cells. We thank Xianwen Ren and Ankur Sharma for providing the raw scRNA-seq data. We thank Hongkui Deng for providing experimental equipment and instruments. We thank LetPub (www.letpub.com) for linguistic assistance and pre-submission expert review. This work was supported in part by the National Key R&D Program of China (2021YFF1201100 and 2017YFA0103002), the National Natural Science Foundation of China (NSFC, 32270603) and the Fundamental Research Funds for the Central Universities (BMU2021YJ057).

AUTHOR CONTRIBUTIONS

D.Z., J.W., Ya.L., and Yu.L. conceived the project and designed the experiments. Yu.L. performed experiments and data analyses, and C.W. assisted with the data analysis. Yu.L. wrote the manuscript, with all authors contributing to providing feedback.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 28, 2022
Revised: May 3, 2023
Accepted: June 28, 2023
Published: July 2, 2023

REFERENCES

1. Yang, J.D., Hainaut, P., Gores, G.J., Amadou, A., Plymoth, A., and Roberts, L.R. (2019). A global view of hepatocellular carcinoma: trends, risk, prevention and management. Nat. Rev. Gastroenterol. Hepatol. *16*, 589–604. https://doi.org/10.1038/s41575-019-0186-y.

2. Tacke, F. (2017). Targeting hepatic macrophages to treat liver diseases. J. Hepatol. *66*, 1300–1312. https://doi.org/10.1016/j.jhep.2017.02.026.

3. Duluc, D., Delneste, Y., Tan, F., Moles, M.-P., Grimaud, L., Lenoir, J., Preisser, L., Anegon, I., Catala, L., Ifrah, N., et al. (2007). Tumor-associated leukemia inhibitory factor and IL-6 skew monocyte differentiation into tumor-associated macrophage-like cells. Blood *110*, 4319–4330. https://doi.org/10.1182/blood-2007-02-072587.

4. Rőszer, T. (2015). Understanding the Mysterious M2 Macrophage through Activation Markers and Effector Mechanisms. Mediat. Inflamm. *2015*, 816460. https://doi.org/10.1155/2015/816460.

5. Orecchioni, M., Ghosheh, Y., Pramod, A.B., and Ley, K. (2019). Macrophage Polarization: Different Gene Signatures in M1(LPS+) vs. Classically and M2(LPS–) vs. Alternatively Activated Macrophages. Front. Immunol. *10*, 1084. https://doi.org/10.3389/fimmu.2019.01084.

6. Locati, M., Curtale, G., and Mantovani, A. (2020). Diversity, Mechanisms, and Significance of Macrophage Plasticity. Annu. Rev. Pathol. *15*, 123–147. https://doi.org/10.1146/annurev-pathmechdis-012418-012718.

7. DeNardo, D.G., and Ruffell, B. (2019). Macrophages as regulators of tumour immunity and immunotherapy. Nat. Rev. Immunol. *19*, 369–382. https://doi.org/10.1038/s41577-019-0127-6.

8. Zilionis, R., Engblom, C., Pfirschke, C., Savova, V., Zemmour, D., Saatcioglu, H.D., Krishnan, I., Maroni, G., Meyerovitz, C.V., Kerwin, C.M., et al. (2019). Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. Immunity *50*, 1317–1334.e10. https://doi.org/10.1016/j.immuni.2019.03.009.

9. Zhang, L., Li, Z., Skrzypczynska, K.M., Fang, Q., Zhang, W., O'Brien, S.A., He, Y., Wang, L., Zhang, Q., Kim, A., et al. (2020). Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer. Cell *181*, 442–459.e29. https://doi.org/10.1016/j.cell.2020.03.048.

10. Cheng, S., Li, Z., Gao, R., Xing, B., Gao, Y., Yang, Y., Qin, S., Zhang, L., Ouyang, H., Du, P., et al. (2021). A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. Cell *184*, 792–809.e23. https://doi.org/10.1016/j.cell.2021.01.010.

11. Cheng, Y., Ma, X.-l., Wei, Y.-q., and Wei, X.-W. (2019). Potential roles and targeted therapy of the CXCLs/CXCR2 axis in cancer and inflammatory diseases. Biochim. Biophys. Acta Rev. Canc *1871*, 289–312. https://doi.org/10.1016/j.bbcan.2019.01.005.

12. Sun, Y., Wu, L., Zhong, Y., Zhou, K., Hou, Y., Wang, Z., Zhang, Z., Xie, J., Wang, C., Chen, D., et al. (2021). Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. Cell *184*, 404–421.e16. https://doi.org/10.1016/j.cell.2020.11.041.

13. Cassetta, L., and Pollard, J.W. (2018). Targeting macrophages: therapeutic approaches in cancer. Nat. Rev. Drug Discov. *17*, 887–904. https://doi.org/10.1038/nrd.2018.169.

14. Mantovani, A., Marchesi, F., Malesci, A., Laghi, L., and Allavena, P. (2017). Tumour-associated macrophages as treatment targets in oncology. Nat. Rev. Clin. Oncol. *14*, 399–416. https://doi.org/10.1038/nrclinonc.2016.217.

15. Schmid, M.C., Franco, I., Kang, S.W., Hirsch, E., Quilliam, L.A., and Varner, J.A. (2013). PI3-kinase γ promotes Rap1a-mediated activation of myeloid cell integrin α4β1, leading to tumor inflammation and growth. PLoS One *8*, e60226. https://doi.org/10.1371/journal.pone.0060226.

16. Ngambenjawong, C., Gustafson, H.H., and Pun, S.H. (2017). Progress in tumor-associated macrophage (TAM)-targeted therapeutics. Adv. Drug Deliv. Rev. *114*, 206–221. https://doi.org/10.1016/j.addr.2017.04.010.

17. The Gene Ontology Consortium (2019). The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. *47*, D330–D338. https://doi.org/10.1093/nar/gky1055.

18. Roumenina, L.T., Daugan, M.V., Noé, R., Petitprez, F., Vano, Y.A., Sanchez-Salas, R., Becht, E., Meilleroux, J., Clec'h, B.L., Giraldo, N.A., et al. (2019). Tumor cells hijack macrophage-produced complement C1q to promote tumor growth. Cancer Immunol. Res. *7*, 1091–1105.

19. Greene, S., Robbins, Y., Mydlarz, W.K., Huynh, A.P., Schmitt, N.C., Friedman, J., Horn, L.A., Palena, C., Schlom, J., Maeda, D.Y., et al. (2020). Inhibition of MDSC Trafficking with SX-682, a CXCR1/2 Inhibitor, Enhances NK-Cell Immunotherapy in Head and Neck Cancer Models. Clin. Cancer Res. *26*, 1420–1431. https://doi.org/10.1158/1078-0432.CCR-19-2625.

20. Teijeira, Á., Garasa, S., Gato, M., Alfaro, C., Migueliz, I., Cirella, A., de Andrea, C., Ochoa, M.C., Otano, I., Etxeberria, I., et al. (2020). CXCR1 and CXCR2 chemokine receptor agonists produced by tumors induce neutrophil extracellular traps that interfere with immune cytotoxicity. Immunity *52*, 856–871.e8.

21. Schuijs, M.J., Png, S., Richard, A.C., Tsyben, A., Hamm, G., Stockis, J., Garcia, C., Pinaud, S., Nicholls, A., Ros, X.R., et al. (2020). ILC2-driven innate immune checkpoint mechanism antagonizes NK cell antimetastatic function in the lung. Nat. Immunol. *21*, 998–1009. https://doi.org/10.1038/s41590-020-0745-y.

22. MacParland, S.A., Liu, J.C., Ma, X.-Z., Innes, B.T., Bartczak, A.M., Gage, B.K., Manuel, J., Khuu, N., Echeverri, J., Linares, I., et al. (2018). Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat. Commun. *9*, 4383. https://doi.org/10.1038/s41467-018-06318-7.

23. Tamburini, B.A.J., Finlon, J.M., Gillen, A.E., Kriss, M.S., Riemondy, K.A., Fu, R., Schuyler, R.P., Hesselberth, J.R., Rosen, H.R., and Burchill, M.A. (2019). Chronic liver disease in humans causes expansion and differentiation of liver lymphatic endothelial cells, *10*. https://doi.org/10.3389/fimmu.2019.01036.

24. Ramachandran, P., Dobie, R., Wilson-Kanamori, J.R., Dora, E.F., Henderson, B.E.P., Luu, N.T., Portman, J.R., Matchett, K.P., Brice, M., Marwick, J.A., et al. (2019). Resolving the fibrotic niche of human liver cirrhosis at single-cell level. Nature *575*, 512–518. https://doi.org/10.1038/s41586-019-1631-3.

25. Young, M.D., Mitchell, T.J., Vieira Braga, F.A., Tran, M.G.B., Stewart, B.J., Ferdinand, J.R., Collord, G., Botting, R.A., Popescu, D.-M., Loudon, K.W., et al. (2018). Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. Science *361*, 594–599. https://doi.org/10.1126/science.aat1699.

26. Chaumonnot, K., Masson, S., Sikner, H., Bouchard, A., Baverel, V., Bellaye, P.-S., Collin, B., Garrido, C., and Kohli, E. (2021). The HSP GRP94 interacts with macrophage intracellular complement C3 and impacts M2 profile during ER stress. Cell Death Dis. *12*, 114. https://doi.org/10.1038/s41419-020-03288-x.

27. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol. *38*, 1408–1414. https://doi.org/10.1038/s41587-020-0591-3.

28. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502. https://doi.org/10.1038/s41586-019-0969-x.

29. Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genom. *19*, 477. https://doi.org/10.1186/s12864-018-4772-0.

30. Angerer, P., Haghverdi, L., Büttner, M., Theis, F.J., Marr, C., and Buettner, F. (2016). destiny: diffusion maps for large-scale single-cell data in R. Bioinformatics *32*, 1241–1243. https://doi.org/10.1093/bioinformatics/btv715.

31. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., and Gulin, A.J.a.p.a. (2017). CatBoost: unbiased boosting with categorical features.

32. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y (2017). Lightgbm: A highly efficient gradient boosting decision tree. 30, 3146-3154.

33. Chen, T., and Guestrin, C. (2016). Xgboost: A Scalable Tree Boosting System. pp. 785-794.

34. Wang, S.M., Ooi, L.L.P.J., and Hui, K.M. (2007). Identification and validation of a novel gene signature associated with the recurrence of human hepatocellular carcinoma. Clin. Cancer Res. *13*, 6275–6283. https://doi.org/10.1158/1078-0432.CCR-06-2236.

35. Chiang, D.Y., Villanueva, A., Hoshida, Y., Peix, J., Newell, P., Minguez, B., LeBlanc, A.C., Donovan, D.J., Thung, S.N., Solé, M., et al. (2008). Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. Cancer Res. *68*, 6779–6788. https://doi.org/10.1158/0008-5472.CAN-08-0742.

36. Roessler, S., Jia, H.-L., Budhu, A., Forgues, M., Ye, Q.-H., Lee, J.-S., Thorgeirsson, S.S., Sun, Z., Tang, Z.-Y., Qin, L.-X., and Wang, X.W. (2010). A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. Cancer Res. *70*, 10202–10212. https://doi.org/10.1158/0008-5472.CAN-10-2607.

37. Hodo, Y., Honda, M., Tanaka, A., Nomura, Y., Arai, K., Yamashita, T., Sakai, Y., Yamashita, T., Mizukoshi, E., Sakai, A., et al. (2013). Association of interleukin-28B genotype and hepatocellular carcinoma recurrence in patients with chronic hepatitis C. Clin. Cancer Res. *19*, 1827–1837. https://doi.org/10.1158/1078-0432.CCR-12-1641.

38. Lim, H.-Y., Sohn, I., Deng, S., Lee, J., Jung, S.H., Mao, M., Xu, J., Wang, K., Shi, S., Joh, J.W., et al. (2013). Prediction of Disease-free Survival in Hepatocellular Carcinoma by Gene Expression Profiling. Ann. Surg Oncol.

20, 3747–3753. https://doi.org/10.1245/s10434-013-3070-y.

39. Dong, B., Lee, J.-S., Park, Y.-Y., Yang, F., Xu, G., Huang, W., Finegold, M.J., and Moore, D.D. (2015). Activating CAR and β-catenin induces uncontrolled liver growth and tumorigenesis. Nat. Commun. *6*, 5944. https://doi.org/10.1038/ncomms6944.

40. Villanueva, A., Portela, A., Sayols, S., Battiston, C., Hoshida, Y., Méndez-González, J., Imbeaud, S., Letouzé, E., Hernandez-Gea, V., Cornella, H., et al. (2015). DNA methylation-based prognosis and epidrivers in hepatocellular carcinoma. Multicenter Study. Hepatology *61*, 1945–1956. https://doi.org/10.1002/hep.27732.

41. Grinchuk, O.V., Yenamandra, S.P., Iyer, R., Singh, M., Lee, H.K., Lim, K.H., Chow, P.K.-H., and Kuznetsov, V.A. (2018). Tumor-adjacent tissue co-expression profile analysis reveals pro-oncogenic ribosomal gene signature for prognosis of resectable hepatocellular carcinoma. Mol. Oncol. *12*, 89–113. https://doi.org/10.1002/1878-0261.12153.

42. Hutter, C., and Zenklusen, J.C. (2018). The Cancer Genome Atlas: Creating Lasting Value beyond Its Data. Cell *173*, 283–285. https://doi.org/10.1016/j.cell.2018.03.042.

43. The International Cancer Genome Consortium. International network of cancer genome projects, (2010). 464, 993-998.

44. Chen, C.-Y., Chung, I.H., Tsai, M.-M., Tseng, Y.-H., Chi, H.-C., Tsai, C.-Y., Lin, Y.-H., Wang, Y.-C., Chen, C.-P., Wu, T.-I., et al. (2014). Thyroid hormone enhanced human hepatoma cell motility involves brain-specific serine protease 4 activation via ERK signaling. Mol. Cancer *13*, 162. https://doi.org/10.1186/1476-4598-13-162.

45. Zhou, L., Yu, K.H., Wong, T.L., Zhang, Z., Chan, C.H., Loong, J.H., Che, N., Yu, H.J., Tan, K.V., Tong, M., et al. (2022). Lineage tracing and single-cell analysis reveal proliferative Prom1+ tumour-propagating cells and their dynamic cellular transition during liver cancer progression. Gut *71*, 1656–1668. https://doi.org/10.1136/gutjnl-2021-324321.

46. Liao, Y.-M., Wang, Y.-H., Hung, J.-T., Lin, Y.-J., Huang, Y.-L., Liao, G.-S., Hsu, Y.-L., Wu, J.-C., and Yu, A.L. (2021). High B3GALT5 expression confers poor clinical outcome and contributes to tumor progression and metastasis in breast cancer. Breast Cancer Res. *23*, 5. https://doi.org/10.1186/s13058-020-01381-9.

47. Zhang, D., Jiang, Q., Ge, X., Shi, Y., Ye, T., Mi, Y., Xie, T., Li, Q., and Ye, Q. (2021). RHOV promotes lung adenocarcinoma cell growth and metastasis through JNK/c-Jun pathway. Int. J. Biol. Sci. *17*, 2622–2632. https://doi.org/10.7150/ijbs.59939.

48. Gao, Q., Yang, L., Shen, A., Li, Y., Li, Y., Hu, S., Yang, R., Wang, X., Yao, X., and Shen, G. (2021). A WNT7B-m6A-TCF7L2 positive feedback loop promotes gastric cancer progression and metastasis. Signal Transduct. Targeted Ther. *6*, 43. https://doi.org/10.1038/s41392-020-00397-z.

49. Schiano Lomoriello, I., Giangreco, G., Iavarone, C., Tordonato, C., Caldieri, G., Serio, G., Confalonieri, S., Freddi, S., Bianchi, F., Pirroni, S., et al. (2020). A self-sustaining endocytic-based loop promotes breast cancer plasticity leading to aggressiveness and pro-metastatic behavior. Nat. Commun. *11*, 3020. https://doi.org/10.1038/s41467-020-16836-y.

50. Liu, J., Gao, L., Zhan, N., Xu, P., Yang, J., Yuan, F., Xu, Y., Cai, Q., Geng, R., and Chen, Q. (2020). Hypoxia induced ferritin light chain (FTL) promoted epithelia mesenchymal transition and chemoresistance of glioma. J. Exp. Clin. Cancer Res. *39*, 137. https://doi.org/10.1186/s13046-020-01641-8.

51. Liu, J., Song, X., Kuang, F., Zhang, Q., Xie, Y., Kang, R., Kroemer, G., and Tang, D. (2021). NUPR1 is a critical repressor of ferroptosis. Nat. Commun. *12*, 647. https://doi.org/10.1038/s41467-021-20904-2.

52. Huang, C., Santofimia-Castaño, P., and Iovanna, J. (2021). NUPR1: A Critical Regulator of the Antioxidant System. Cancers *13*, 3670. https://doi.org/10.3390/cancers13153670.

53. Ye, Z., Zhuo, Q., Hu, Q., Xu, X., Mengqi, l., Zhang, Z., Xu, W., Liu, W., Fan, G., Qin, Y., et al. (2021). FBW7-NRA41-SCD1 axis synchronously regulates apoptosis and ferroptosis in pancreatic cancer cells. Redox Biol. *38*, 101807. https://doi.org/10.1016/j.redox.2020.101807.

54. Banks, C.J., and Andersen, J.L. (2019). Mechanisms of SOD1 regulation by post-translational modifications. Redox Biol. *26*, 101270. https://doi.org/10.1016/j.redox.2019.101270.

55. Fang, X., Wang, H., Han, D., Xie, E., Yang, X., Wei, J., Gu, S., Gao, F., Zhu, N., Yin, X., et al. (2019). Ferroptosis as a target for protection against cardiomyopathy. Proc. Natl. Acad. Sci. USA *116*, 2672–2680. https://doi.org/10.1073/pnas.1821022116.

56. Jiang, Y., Cao, W., Wu, K., Qin, X., Wang, X., Li, Y., Yu, B., Zhang, Z., Wang, X., Yan, M., et al. (2019). LncRNA LINC00460 promotes EMT in head and neck squamous cell carcinoma by facilitating peroxiredoxin-1 into the nucleus. J. Exp. Clin. Cancer Res. *38*, 365. https://doi.org/10.1186/s13046-019-1364-z.

57. Matsusaka, K., Fujiwara, Y., Pan, C., Esumi, S., Saito, Y., Bi, J., Nakamura, Y., Mukunoki, A., Takeo, T., Nakagata, N., et al. (2021). α1-Acid Glycoprotein Enhances the Immunosuppressive and Protumor Functions of Tumor-Associated Macrophages. Cancer Res. *81*, 4545–4559. https://doi.org/10.1158/0008-5472.CAN-20-3471.

58. Lou, Y., Han, M., Liu, H., Niu, Y., Liang, Y., Guo, J., Zhang, W., and Wang, H. (2020). Essential roles of S100A10 in Toll-like receptor signaling and immunity to infection. Cell. Mol. Immunol. *17*, 1053–1062. https://doi.org/10.1038/s41423-019-0278-1.

59. O'Connell, P.A., Surette, A.P., Liwski, R.S., Svenningsson, P., and Waisman, D.M. (2010). S100A10 regulates plasminogen-dependent macrophage invasion. Blood *116*, 1136–1146. https://doi.org/10.1182/blood-2010-01-264754 %J Blood.

60. Cai, H., Zhu, X.-D., Ao, J.-Y., Ye, B.-G., Zhang, Y.-Y., Chai, Z.-T., Wang, C.-H., Shi, W.-K., Cao, M.-Q., Li, X.-L., and Sun, H.-C. (2017). Colony-stimulating factor-1-induced AIF1 expression in tumor-associated macrophages enhances the progression of hepatocellular carcinoma. OncoImmunology *6*, e1333213. https://doi.org/10.1080/2162402X.2017.1333213.

61. Lecker, L.S.M., Berlato, C., Maniati, E., Delaine-Smith, R., Pearce, O.M.T., Heath, O., Nichols, S.J., Trevisan, C., Novak, M., McDermott, J., et al. (2021). TGFBI Production by Macrophages Contributes to an Immunosuppressive Microenvironment in Ovarian Cancer. Cancer Res. *81*, 5706–5719. https://doi.org/10.1158/0008-5472.CAN-21-0536.

62. Molgora, M., Esaulova, E., Vermi, W., Hou, J., Chen, Y., Luo, J., Brioschi, S., Bugatti, M., Omodei, A.S., Ricci, B., et al. (2020). TREM2 Modulation Remodels the Tumor Myeloid Landscape Enhancing Anti-PD-1 Immunotherapy. Cell *182*, 886–900.e17. https://doi.org/10.1016/j.cell.2020.07.013.

63. Sumaiya, K., Langford, D., Natarajaseenivasan, K., and Shanmughapriya, S. (2022). Macrophage migration inhibitory factor (MIF): A multifaceted cytokine regulated by genetic and physiological strategies. Pharmacol. Ther. *233*, 108024. https://doi.org/10.1016/j.pharmthera.2021.108024.

64. Bloom, B.R., and Bennett, B. (1966). Mechanism of a reaction in vitro associated with delayed-type hypersensitivity. Science *153*, 80–82. https://doi.org/10.1126/science.153.3731.80.

65. Yang, T., Wang, Y., Dai, W., Zheng, X., Wang, J., Song, S., Fang, L., Zhou, J., Wu, W., and Gu, J. (2018). Increased B3GALNT2 in hepatocellular carcinoma promotes macrophage recruitment via reducing acetoacetate secretion and elevating MIF activity. J. Hematol. Oncol. *11*, 50. https://doi.org/10.1186/s13045-018-0595-3.

66. Zhao, J., Li, H., Zhao, S., Wang, E., Zhu, J., Feng, D., Zhu, Y., Dou, W., Fan, Q., Hu, J., et al. (2021). Epigenetic silencing of miR-144/451a cluster contributes to HCC progression via paracrine HGF/MIF-mediated TAM remodeling. Mol. Cancer *20*, 46. https://doi.org/10.1186/s12943-021-01343-5.

67. Winner, M., Meier, J., Zierow, S., Rendon, B.E., Crichlow, G.V., Riggs, R., Bucala, R., Leng, L., Smith, N., Lolis, E., et al. (2008). A novel, macrophage migration inhibitory factor suicide substrate inhibits motility and growth of lung cancer cells. Cancer Res. *68*, 7253–7257. https://doi.org/10.1158/0008-5472.Can-07-6227.

68. Al-Abed, Y., Dabideen, D., Aljabari, B., Valster, A., Messmer, D., Ochani, M., Tanovic, M., Ochani, K., Bacher, M., Nicoletti, F., et al. (2005). ISO-1 binding to the tautomerase active site of MIF inhibits its pro-inflammatory activity and increases survival in severe sepsis.

J. Biol. Chem. *280*, 36541–36544. https://doi.org/10.1074/jbc.C500243200.

69. Aggarwal, N., Deerhake, M.E., DiPalma, D., Shahi, S.K., Gaggioli, M.R., Mangalam, A.K., and Shinohara, M.L. (2021). Secreted osteopontin from $CD4^+$ T cells limits acute graft-versus-host disease. Cell Rep. *37*. https://doi.org/10.1016/j.celrep.2021.110170.

70. Luo, Y., Yi, H., Huang, X., Lin, G., Kuang, Y., Guo, Y., and Xie, C. (2021). Inhibition of macrophage migration inhibitory factor (MIF) as a therapeutic target in bleomycin-induced pulmonary fibrosis rats. Am. J. Physiol. Lung Cell Mol. Physiol. *321*, L6–L16. https://doi.org/10.1152/ajplung.00288.2020.

71. Song, Z., Chen, W., Athavale, D., Ge, X., Desert, R., Das, S., Han, H., and Nieto, N. (2021). Osteopontin takes center stage in chronic liver disease. Hepatology *73*, 1594–1608.

72. Hassn Mesrati, M., Syafruddin, S.E., Mohtar, M.A., and Syahir, A. (2021). CD44: A multifunctional mediator of cancer progression. Biomolecules *11*, 1850.

73. Jiang, X., Zhang, F., Ji, X., Dong, F., Yu, H., Xue, M., Qiu, Y., Yang, F., Hu, X., and Bao, Z. (2021). Lipid-injured hepatocytes release sOPN to improve macrophage migration via CD44 engagement and pFak-NFκB signaling. Cytokine *142*, 155474. https://doi.org/10.1016/j.cyto.2021.155474.

74. Eun, J.W., Yoon, J.H., Ahn, H.R., Kim, S., Kim, Y.B., Lim, S.B., Park, W., Kang, T.W., Baek, G.O., Yoon, M.G., et al. (2023). Cancer-associated fibroblast-derived secreted phosphoprotein 1 contributes to resistance of hepatocellular carcinoma to sorafenib and lenvatinib. Cancer Commun. *43*, 455–479. https://doi.org/10.1002/cac2.12414.

75. Onodera, S., Kaneda, K., Mizue, Y., Koyama, Y., Fujinaga, M., and Nishihira, J. (2000). Macrophage migration inhibitory factor up-regulates expression of matrix metalloproteinases in synovial fibroblasts of rheumatoid arthritis. J. Biol. Chem. *275*, 444–450. https://doi.org/10.1074/jbc.275.1.444.

76. Zhang, Q., He, Y., Luo, N., Patel, S.J., Han, Y., Gao, R., Modak, M., Carotta, S., Haslinger, C., Kind, D., et al. (2019). Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma. Cell *179*, 829–845.e20. https://doi.org/10.1016/j.cell.2019.10.003.

77. Sharma, A., Seow, J.J.W., Dutertre, C.-A., Pai, R., Blériot, C., Mishra, A., Wong, R.M.M., Singh, G.S.N., Sudhagar, S., Khalilnezhad, S., et al. (2020). Onco-fetal Reprogramming of Endothelial Cells Drives Immunosuppressive Macrophages in Hepatocellular Carcinoma. Cell *183*, 377–394.e21. https://doi.org/10.1016/j.cell.2020.08.040.

78. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36. https://doi.org/10.1186/gb-2013-14-4-r36.

79. Chen, K., Xi, Y., Pan, X., Li, Z., Kaestner, K., Tyler, J., Dent, S., He, X., and Li, W. (2013). DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing. Genome Res. *23*, 341–351. https://doi.org/10.1101/gr.142067.112.

80. McGinnis, C.S., Murrow, L.M., and Gartner, Z.J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. Cell Syst. *8*, 329–337.e4. https://doi.org/10.1016/j.cels.2019.03.003.

81. Manimaran, S., Selby, H.M., Okrah, K., Ruberman, C., Leek, J.T., Quackenbush, J., Haibe-Kains, B., Bravo, H.C., and Johnson, W.E. (2016). BatchQC: interactive software for evaluating sample and batch effects in genomic data. Bioinformatics *32*, 3836–3838.

82. Kursa, M.B., and Rudnicki, W.R. (2010). Feature selection with the Boruta package. J. Stat. Software *36*, 1–13.

83. Jin, S., Guerrero-Juarez, C.F., Zhang, L., Chang, I., Ramos, R., Kuan, C.-H., Myung, P., Plikus, M.V., and Nie, Q. (2021). Inference and analysis of cell-cell communication using CellChat. Nat. Commun. *12*, 1088. https://doi.org/10.1038/s41467-021-21246-9.

84. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.

85. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast gene set enrichment analysis. Preprint at bioRxiv. https://doi.org/10.1101/060012.

86. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-r., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296. https://doi.org/10.1038/s41592-019-0619-0.

87. Becht, E., Giraldo, N.A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., Selves, J., Laurent-Puig, P., Sautès-Fridman, C., Fridman, W.H., and de Reyniès, A. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. *17*, 218–220.

88. Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nat. Methods *14*, 979–982. https://doi.org/10.1038/nmeth.4402.

89. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., and Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinf. *12*, 77. https://doi.org/10.1186/1471-2105-12-77.

90. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

91. Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E., and Storey, J.D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics *28*, 882–883. https://doi.org/10.1093/bioinformatics/bts034 %J Bioinformatics.

92. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498. https://doi.org/10.1038/s41586-018-0414-6.

93. Aran, D., Hu, Z., and Butte, A.J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. *18*, 220–314.

94. Zhang, L., Yu, X., Zheng, L., Zhang, Y., Li, Y., Fang, Q., Gao, R., Kang, B., Zhang, Q., Huang, J.Y., et al. (2018). Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature *564*, 268–272. https://doi.org/10.1038/s41586-018-0694-x.

95. Zhang, X., Lan, Y., Xu, J., Quan, F., Zhao, E., Deng, C., Luo, T., Xu, L., Liao, G., Yan, M., et al. (2018). CellMarker: a manually curated resource of cell markers in human and mouse. Nucleic Acids Res. *47*, D721–D728. https://doi.org/10.1093/nar/gky900.

96. Wu, X.-Z., and Zhou, Z.-H. (2017). A Unified View of Multi-Label Performance Measures. In Proceedings of the 34th International Conference on Machine Learning, P. Doina and T. Yee Whye, eds. (PMLR).

97. International Cancer Genome Consortium, Hudson, T.J., Anderson, W., Artez, A., Barker, A.D., Bernabé, R.R., Bernabé, R.R., Bhan, M.K., Calvo, F., Eerola, I., et al. (2010). International network of cancer genome projects. Nature *464*, 993–998. https://doi.org/10.1038/nature08987.

98. Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics *8*, 118–127. https://doi.org/10.1093/biostatistics/kxj037 %J Biostatistics.

99. Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., and Alizadeh, A.A. (2015). Robust enumeration of cell subsets from tissue expression profiles. Nat. Methods *12*, 453–457.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-SPP1	Abcam	Cat# ab214050
Anti-MIF	Abcam	Cat# ab187064
Anti-GADPH	Abcam	Cat# ab8245
HRP-labeled Goat Anti-Rabbit IgG (H+L)	Beyotime	Cat# A0208
Bacterial and virus strains		
MIF shRNA lentiviral	Genechem	
Chemicals, peptides, and recombinant proteins		
PMA	Santa Cruz	CAS 16561-29-8; Cat# sc-3576
ISO-1	Selleck	CAS 478336-92-4; Cat# I287808
4-IPP	Aladdin	CAS 41270-96-6; Cat# S7732
2-mercaptoethanol	Aladdin	CAS 60-24-2; Cat# M301573
Tween 20	Ameresco	CAS 9005-64-5
Methanol	Tongguang Fine	CAS 67-56-1
Chloroform	Tongguang Fine	CAS 67-66-3
Anhydrous ethanol	Tongguang Fine	CAS 563-64-4
Isopropanol	Tongguang Fine	CAS 67-63-0
Critical commercial assays		
Cell Counting Kit-8	MCE	Cat# HY-K0301
TB Green PCR Mix	Takara	Cat# RR420A
cDNA Reverse Transcriptase	Thermo Fisher Scientific	Cat# EP0751
BCA Protein Assay Kit	Applygen	Cat# P1511
Enhanced Chemiluminescence Kit	Advansta	Cat# K-12049-D50
Human osteopontin ELISA Kit	Multisciences	Cat# 70-EK1135-48
Protease Inhibitor	Applygen	Cat# C1055
TRIzol® Reagent	Applygen	Cat# R1030
Trypsin-EDTA (0.25%)	Gibco	Cat# 25200056
Matrigel	Solarbio	Cat# M8370
Skimmed milk powder	Wondersun	Cat# 100012705274
Deposited data		
10× Genomics scRNA-seq data	Zhang et al.[76]	GEO: GSE140228
10× Genomics scRNA-seq data	Sharma et al.[77]	GEO: GSE156625
10× Genomics scRNA-seq data	MacParland et al.[22]	GEO: GSE115469
10× Genomics scRNA-seq data	Tamburini BAJ et al.[23]	GEO: GSE129933
10× Genomics scRNA-seq data	Ramachandran P et al.[24]	GEO: GSE136103
10× Genomics scRNA-seq data	Sun Y et al.[12]	CNGBdb: CNP0000650
scRNA Smart-seq data	Zhang et al.[76]	GEO: GSE140228
10× Genomics scRNA-seq data	Cheng et al.[10]	GEO: GSE154763
RNA-Seq data	This paper	GEO: GSE230666
RNA-Seq data	TCGA	TCGA-LIHC
RNA-Seq data	ICGC	LIRI-JP

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
microarray data	Wang et al.[34]	GEO: GSE121248
microarray data	Hodo et al.[37]	GEO: GSE41804
microarray data	Chiang et al.[35]	GEO: GSE9843
microarray data	Lim et al.[38]	GEO: GSE36376
microarray data	Dong et al.[39]	GEO: GSE43619
microarray data	Grinchuk et al.[41]	GEO: GSE76427
microarray data	Roessler et al.[36]	GEO: GSE14520
microarray data	Villanueva et al.[40]	GEO: GSE63898
processed data	This paper	https://figshare.com/account/home#/projects/164332
codes	This paper	https://github.com/LiaoYunxi/HCC_MNPs_MIF
Experimental models: Cell lines		
Human: THP-1 cells	Dr. Huanran Tan's laboratory, School of Basic Medicine at Peking University	N/A
Human: HepG2 cells	ATCC	Cat# HB-8065
Oligonucleotides		
Oligonucleotides/primers are listed in Table S4.	RuiBiotech	N/A
Software and algorithms		
TopHat	Kim et al.[78]	https://ccb.jhu.edu/software/tophat/index.shtml
HTSeq	N/A	https://htseq.readthedocs.io/en/master/
BEDTools	N/A	https://bedtools.readthedocs.io/en/latest/
bedGraphToBigWig	N/A	https://www.encodeproject.org/software/bedgraphtobigwig/
DANPOS	Chen et al.[79]	https://sites.google.com/site/danposdoc/
DoubletFinder	McGinnis et al.[80]	https://github.com/chris-mcginnis-ucsf/DoubletFinder
LightGBM	Ke et al.[32]	https://github.com/Microsoft/LightGBM
scvelo	Bergen et al.[27]	https://scvelo.readthedocs.io/
BatchQC	Manimaran et al.[81]	https://github.com/mani2012/BatchQC
Boruta	Kursa and Rudnicki[82]	https://cran.r-project.org/web/packages/Boruta/index.html
caret	N/A	https://cran.r-project.org/web/packages/caret/index.html
Catboost	Prokhorenkova et al.[31]	https://catboost.ai/
CellChat	Jin et al.[83]	http://www.cellchat.org/
edgeR	N/A	https://bioconductor.org/packages/release/bioc/html/edgeR.html
clusterProfiler	Wu et al.[84]	http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html
destiny	Angerer et al.[30]	http://www.bioconductor.org/packages/release/bioc/html/destiny.html
Fgsea	Korotkevich et al.[85]	http://www.bioconductor.org/packages/release/bioc/html/fgsea.html
gglmannotate	N/A	https://github.com/wilkox/gglmannotate
glmnet	N/A	https://cran.r-project.org/web/packages/glmnet/index.html

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Harmony	Korsunsky et al.[86]	https://github.com/immunogenomics/harmony
MCPcounter	Becht et al.[87]	https://github.com/ebecht/MCPcounter
modEvA	N/A	https://cran.r-project.org/web/packages/modEvA/index.html
monocle	Qiu et al.[88]	http://cole-trapnell-lab.github.io/monocle-release/
monocle3	Cao et al.[28]	http://cole-trapnell-lab.github.io/monocle-release/
msigdbr	N/A	http://www.bioconductor.org/packages/release/data/experiment/html/msigdb.html
Presto	N/A	https://cran.r-project.org/web/packages/presto/index.html
pROC	Robin et al.[89]	https://cran.r-project.org/web/packages/pROC/index.html
randomForest	N/A	https://cran.r-project.org/web/packages/randomForest/index.html
Seurat	Hao et al.[90]	https://satijalab.org/Seurat/index.html
slingshot	Street et al.[29]	http://www.bioconductor.org/packages/release/bioc/html/slingshot.html
survival	N/A	https://cran.r-project.org/web/packages/survival/index.html
survminer	N/A	https://cran.r-project.org/web/packages/survminer/index.html
Sva	Leek et al.[91]	http://www.bioconductor.org/packages/release/bioc/html/sva.html
velocyto.R	La Manno et al.[92]	http://velocyto.org/
xCell	Aran et al.[93]	https://github.com/dviraran/xCell
XGBoost	Chen and Guestrin[33]	https://cran.r-project.org/web/packages/xgboost/index.html
Ggraph	N/A	https://cran.r-project.org/web/packages/ggraph/index.html
Other		
CFX384 Real-Time system	BIO-RAD	
6.5 mm Transwell® with 8.0 µm Pore Polycarbonate Membrane Insert	Corning	Cat# 3422
polyvinylidene difluoride	Millipore	Cat# ISEQ00010
RPMI 1640 medium	Gibco	Cat# C11875500BT
DMEM medium	Gibco	Cat# C11995500BT
FBS	Gibco	Cat# A3161001C, 10099-141

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dongyu Zhao (zhaodongyu@bjmu.edu.cn).

Materials availability

Reagents and methods in this study will be made available by the lead contact upon request.

Data and code availability

- The publiced scRNA-seq and microarray data used for analysis of MNPs and the data generated for this manuscript are stored in Gene Expression Omnibus. Accession numbers are listed in the key resources table (KRT). The tumor microenvironment RNA-seq data for TME classification were obtained from the TCGA-LIHC and LIRI-JP project.
- All original code has been deposited at github, and the URL is listed in the key resources table.
- The processed data can be downloaded from figshare, and the URL is also listed in the key resources table. Other additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture and treatment

HepG2 (ATCC) cells were grown in DMEM medium supplemented with 10% fetal bovine serum (FBS). THP-1 (ATCC) cells were donated by Professor Tan Huanran (School of Basic Medicine at Peking University) and grown in RPMI 1640 medium supplemented with 2-mercaptoethanol to a final concentration of 0.05 mM and 10% FBS. Cells were incubated at 37°C in a 95% humidified atmosphere containing 5% CO_2. Both cell lines have not been authenticated or tested for mycoplasma contamination.

A total of 1×10^6/ml THP-1 cells were differentiated into M0 macrophages by 24 h incubation with 150 mM phorbol-12-myristate 13-acetate.

Culture supernatant was harvested 36 h after HepG2 cell starvation treatment to use as TCM for developing HCC tumor-associated macrophages (TAMs). M0 macrophages were cultured with different concentration gradients of TCM for 24 h for HCC-TAM induction. For the MIF inhibitor treatment experiments, 5 μM ISO-1 and 20 μM 4-IPP were added to the TCM at the step of TAM induction for 24-h induction to investigate the effect of MIF on HCC-TAMs.

METHOD DETAILS

Plotting cell growth curves

THP-1 cells were treated with MIF inhibitors for cell growth curves. The standard curve was plotted by seeding 0, 5×10^3, 1×10^4, 2×10^4, and 4×10^4 THP-1 cells/well into a 96-well plate, adding 10 μL of CCK-8 reagent, incubating for 2 h, and then measuring the absorbance at 450 nm. For cell growth curves, the cells were seeded into 96-well plates (8×10^3 cells/well) with different doses of drug (ISO-1: 1.25, 2.5, 5, 10 μM; 4-IPP: 2.5, 5, 10, 20 μM). Cell viability was determined by Cell Counting Kit-8 every 2 days. ISO-1 was obtained from Selleck, and 4-IPP was obtained from Aladdin.

Reverse transcription qPCR analysis

Quantitative real-time PCR (qRT–PCR) was performed to measure the mRNA levels CD86, CD80, CD206, CD163, SPP1, IL-1β, IL-10, TNF-α, and TGF-β in cells. The qPCR primers used in this experiment are shown in Table S3. Total RNA was extracted from the cells using TRIzol according to the manufacturer's instructions. After reverse transcription using cDNA reverse transcriptase and Oligo (dT) primer, qRT–PCR was performed using TB Green PCR Mix and the CFX384 Real-Time system. mRNA expression levels were normalized to β-actin. Data are presented as expression quantification relative to the M0 macrophage control determined by the $2^{-\Delta\Delta Ct}$ method.

Western blot analysis

Whole cells were lysed in RIPA buffer containing a protease inhibitor. The protein concentration was measured with a BCA Protein Assay Kit. Protein samples were separated using sodium dodecylsulfate-polyacrylamide gel electrophoresis and transferred onto polyvinylidene difluoride. The membranes were blocked with 5% nonfat milk for 2 h and incubated overnight with the indicated primary antibody. A secondary antibody was added, and incubation was conducted for 1.5 h at room temperature. Finally, the blots were visualized with an enhanced chemiluminescence kit and the gray value was analyzed using ImageJ software.

ELISA

After the induction of THP-1 cells into M0 cells, HCC-TCM was used to further induce HCC-TAMs; 5 μM ISO-1 and 20 μM 4-IPP were added to TCM, and after 24 hours, the supernatants were collected in 1mL after centrifuged at 1000 rpm at 4°C for 5 min to remove cells and diluted 10-fold. The OPN content was measured using the Human osteopontin ELISA Kit, according to the manufacturer's protocol.

Macrophage migration assays

Fresh 1×10^5 TCM-TAMs suspended in 150 μL serum-free 1640 were added to the upper chamber of 8 μm pore transwell inserts (Corning, 3422). The inserts were then moved to 24-well plates containing 800 μL RPMI 1640 medium with 20% FBS and incubated at 37°C for 2 days. The migrated cells were fixed in methanol and stained with 0.5% Crystal Violet. Migratory cells in the lower chamber were quantified using five random fields. The area of stained cells was calculated using ImageJ software.

For the wound-healing assay, THP-1 cells (1×10^6/ml) were seeded in 24-well plates and inducted into TCM-TAMs. The monolayer was then disrupted by scraping with 10 μL pipette tips, and photographs were taken at 0 and 72 h using a phase-contrast microscope. Wound areas were calculated by manually tracing the cell-free area in the images using ImageJ software.

Tumor invasion assay

Matrigel was diluted with cold serum-free DMEM medium at a 1:40 ratio to obtain a working concentration of 200 μg/mL after melting at 4°C. A 100 μL aliquot of the diluted gel was added to each Transwell insert and incubated at 37°C for 4 hours. To prepare for cell seeding, 200 μL of serum-free medium was added to each Transwell insert, followed by aspiration after 30 minutes. HepG2 cells were digested with trypsin, re-suspended in a serum-free suspension at 5×10^5 cells/mL, and added to the Transwell inserts. The inserts were then placed into 24-well plates containing HCC-TAM, 5 μM ISO-1, and 20 μM 4IPP-treated HCC-TAM for the experimental groups, respectively. Next, 800 μL of DMEM medium containing 20% FBS was added to each well of the 24-well plates. After 48 hours of incubation at 37°C, staining and counting were performed using the same method as that used for the macrophage migration assays.

Lentiviral transfection experiments

The lentiviral vector containing SPP1 shRNA was GV115 with the sequence of components hU6-MCS-CMV-EGFP.

Two hundred thousand THP-1 cells were inoculated in a 24-well plate, and the cells were transfected with 3×10^7 TU/ml virus and A solution, centrifuged at 1000 rpm for 1 hour at room temperature, and placed in a cell culture incubator. The solution was changed after 13 hours. After one week of culture, cells with strong positive expression of GFP were sorted by flow cytometry.

Library construction, quality control and sequencing for RNA-seq

We induced THP-1 cells with lentivirus carrying shRNA and negative control virus into HCC-TAMs, and extracted total RNA for RNA-seq. The NEB Next Ultra RNA Library Prep Kit for Illumina was used to generate sequencing libraries, and index codes were added to distinguish between samples. After mRNA purification and cDNA synthesis, cDNA fragments of preferred length were selected, and PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers, and Index Primer. The libraries were then sequenced on Illumina platforms. The same RNA sequencing process was used for 4IPP-treated HCC-TAMs and their control.

Integration and processing of the 10x myeloid single cell dataset

The 10x scRNA-seq datasets from HCC immune cells were obtained from the GEO database (Gene Expression Omnibus). A total of 75,696 immune cells from AN, TP, and TC sites in the data from Zhang Q et al.[76] and Sharma A et al.[77] were selected for integration based on the original annotations in the two datasets. First, we used the 'Seurat' R package to normalize the two datasets separately. We obtained 4000 variable genes, scaled them, and ran a PCA. The data were quality controlled by removing cells with mitochondrial gene percentages greater than 10% and nFeature > 200. Doublets in both datasets were found and removed separately using the DoubletFinder R package[80] (Figure S1A) with the following setting parameters: PCs = 30, pN = 0.25, and nExp = 5% of the total number of cells. We first used Seurat v4 methods[90] to

match shared cell populations across datasets and remove batch effects. These methods first identified cross-dataset pairs of cells in a matched biological state ('anchors') using the FindIntegrationAnchors function. These anchors were used to integrate the two datasets with IntegrateData. Then, we ran an integrated analysis on all cells, calculated the k-nearest neighbors, and identified clusters of cells using a shared nearest neighbor modularity optimization-based clustering algorithm.

We isolated myeloid cells for reintegration using the Harmony algorithm.[86] After creating the two datasets as Seruat objects, a PCA matrix with 50 components was computed for the Harmony algorithm. Harmony embeds gene expression data into a low-dimensional space via PCA. Then, it uses an iterative process to remove data-specific effects.

Similarity analysis of cell types between health livers and HCC tissues

We calculated the Pearson's correlation coefficient of each cell's gene expression levels and performed hierarchical clustering of cells based on the correlation coefficients to examine cell type similarity. In addition, we trained a logistic regression model using elastic net regularization on the cellular identities[25] defined by the cell types in HCC and healthy liver using the "glmnet" R package. In training this model, we used 10-fold cross-validation and set alpha = 0.99 to produce strong regularization but to prevent the exclusion of strongly co-linear genes.

Determination of the spatial preference of MNPs

To measure the enrichment of MNPs across different spatial distributions, we calculated *Ro/e*, the ratio of observed cell number over the expected cell number of a given combination of MNP types and spatial location as follows: ($Ro/e = \frac{observed}{expected}$).[94] The expected cell number for each combination of MNP types and tissues was obtained from the chi-squared test and defined as: $\frac{(observed-expecyed)^2}{expected}$. This method only indicates the divergence of observations from random expectations. $Ro/e > 1$ indicates that cells of the given MNP type are more frequently observed expected at random in the specific microenvironment. $Ro/e < 1$ indicates that cells of the given MNP type are observed less frequently than expected at random in the specific microenvironment.

Analysis of cell-cell communication

Secreted signals were ascertained by evaluating the expression of pairs of ligand-receptor pairs within cell types in different tissues according to the CellChat R toolkit.[83] We set min. cells = 10 for scanning signals and chose a K parameter from 2–5 in the identifyCommunicationPatterns function based on NMF (non-negative matrix factorization) during the analysis of MNPs and non-immune cells from various spatial locations.

Distinguishing polarization, chemotaxis, and immune-related phenotypes

We collected the reported gene signatures of the M1/M2 phenotype,[9,95] calculated their average expression values in different MNPs, and normalized them to 0–5. The known chemotaxis or immune-related genes were obtained from ImmPort (https://www.immport.org/). The three gene signature types are shown in Table S4.

Inference of cell trajectories

To characterize the potential process of cell functional changes and determine the potential lineage differentiation between diverse MNPs, we first applied Velocyto[92] to obtain the loom files of the dataset from Sharma A et al using the GRCh38 database (version 1.2.0) as the human genome reference from the bam files provided by the authors. Subsequently, we used scVelo to calculate and obtain the PAGA (Partition-based graph abstraction) graph. scVelo[27] is a scalable python toolkit that generalizes RNA velocity through dynamical modeling, recovering the latent time of the underlying cellular processes and systematically detecting driver genes via their characterization by high likelihoods. The parameter settings for the Sharma A et al. dataset were min_shared_counts = 5, n_top_genes = 2000, n_pcs = 23, n_neighbors = 30.

We applied four methods to infer cell trajectories, all using Mono-VCAN as root cells. To prevent batch effects, the built-in dimensionality reduction calculation was skipped, and the calculation was performed directly using the data stored in the Seurat object after removing batch effects with Harmony. When running Monocle 3,[28] we randomly sampled 50% of the cells. We found the driver genes that vary between

types of MNPs in UMAP space using the function graph_test function. In Slingshot,[29] we used the associationTest function to invert the variance-covariance matrix of the consecutive nodes along a lineage as contrasts to find driver genes. In Monocle 2,[88] we used marker genes obtained from the function FindAllMarkers function in the Seurat package for ordering cells and the differentialGeneTest function to model a gene's expression level as a smooth, nonlinear function of pseudotime and obtain p values after the likelihood ratio tests. We used the Destiny algorithm based on diffusion maps and found driver genes by calculating gene relevance.

Finally, we compared the order of cells' pseudotime generalized by Monocle3, Monocle2, slingshot, and the Destiny algorithm through Pearson correlation coefficients.

Data preprocessing for training machine learning models

To identify the key genes affecting MNP infiltration, we trained machine learning models using GBDT. Each cell was considered a sample, and the labels corresponded to spatial sites (AN, TP, and TC). We took and merged the driver genes obtained from the four trajectory inference methods as model features to construct a cell expression matrix with 2026 gene features (Slingshot: 1000 genes, Monocle 3: 931 genes, Monocle 2: 879 genes, Destiny: 298 genes) and scVelo (281 genes). The criteria for selecting the 2026 genes were based on the following

- Genes with a probability greater than 0 as driver genes in scVelo
- Genes with Moran's I greater than 0.1 in Monocle 3
- Genes identified as driver genes by QR decomposition in slingshot
- Genes with p-values less than 0.05 from the differentialGeneTest function in Monocle2
- Genes with high gene relevance identified by the Destiny algorithm on the basis of diffusion maps

To improve the effectiveness and quality of machine learning, we first preprocessed the cellular transcriptome expression matrix as the input data. scRNA-seq data represent a high-dimensional sparse matrix. Considering the difficulty of missing value completion, we used Seurat to normalize the data, keeping the original sequencing count of zeroes unchanged. We used the quantile method to eliminate data noise by deleting cell samples containing over 200 extreme expression values, and we used the Boruta algorithm[82] for feature engineering. The Boruta R package is a wrapper built around the random forest classification algorithm. We ran the Boruta function with the default settings and removed the rejected features for the next training.

Model construction

We used XGBoost (Extreme Gradient Boosting),[33] LightGBM (Light Gradient Boosting Machine)[32] and Catboost[31] to train ML models. We divided the training and test set into 7:3 and trained the model using 10-fold cross-validation, using the same training and test sets for all three methods. Other scRNA-seq datasets (Cheng S et al) from HCC and alternate tumor types were used to externally verify the model.[10] We used the caret R package to implement the XGBoost and Catboost algorithms, with the following parameter settings:

XGBoost: objective = "multi:softprob", num_class = 3, eta=0.1, max_depth = 7, nrounds = 10000, booster = "gbtree", eval_metric = "error", subsample = 0.5, colsample_bytree = 1, gamma = 0.5, min_child_weight = 1.

Catboost: loss_function = 'MultiClass', classes_count = 3, learning_rate = 0.2, depth = 7, iterations = 10000, border_count = 64, rsm = 0.95, l2_leaf_reg = 0.01, metric_period = 50.

We implemented LightGBM algorithm in python 3.7 with the parameter settings: objective= 'multiclass', num_class = 3, max_depth = 7, num_iterations = 10000, learning_rate= 0.1, lambda_l1= 0.1, lambda_l2 = 0.2.

Model evaluation

We evaluate the model by comparing the accuracy, AUC (Area under the ROC Curve), mAP (mean Average Precision), and F1-score using the test set and external validation datasets. We calculated the AUC of the binary classifier using the roc function of the pROC R package[89] and AP by the AUC function of the modEvA R package when the set curve = "PR". The performance of the multi-label classifier was measured as macro-F1, macro-AUC, macro-AP, micro-F1, micro-AUC and micro-AP.[96]

Integration of the bulk RNA-seq and microarray datasets

To integrate two sets of tissue transcriptome data from HCC patients, we downloaded the raw count data from the TCGA,[42] ICGC[97] and Gene Expression Omnibus (GEO) databases (https://www.ncbi.nlm.nih.gov/) and normalized the gene expression as transcripts per million (TPM). Given the presence of multiple sources of variation, we applied two parametric ComBat[98] implemented with the sva R package[91] to remove batch effects. Regarding the microarray data, ComBat was first applied to remove batch variability using the common prob, and then a second ComBat step was run to remove prob variability. We also used the BatchQC R package[81] to detect batch effects.

Feature gene classification and sample typing

We first compared the spatial site of highest expression for each gene, dividing the 445 feature genes into the corresponding three categories (AN, TP, and TC). We normalized the gene expression among the model-selected features for bulk-tissue RNA data by calculating the Z-sore. The genes of each sample were divided into three groups according to the gene categories, and the weighted average of each gene category of each sample was calculated as a spatial $score_{MNP}$, and the weight was the importance value obtained by the XGBoost model. Then, each sample was typed according to the highest spatial $score_{MNP}$ of the corresponding spatial type.

$$\text{spatial score}_{MNP} = \frac{\sum_n gene\ expression \times weight}{n}$$

Estimation of immune cell infiltration

We applied four immune infiltration estimation methods to quantify the abundance of tumor infiltration immune cells in HCC specimens, focusing on the elements associated with MNPs and $CD8^+$T cells. We normalized the data obtained from these methods by Z-score and calculated their correlation coefficients with gene expression.

CIBERSORT[99] estimates the proportions of 22 immune cell types in a mixed cell population using normalized data. The CIBERSORT R source code and the immune signature matrix file called "LM22.txt" were obtained online (https://cibersort.stanford.edu/). xCell[93] estimates the abundance scores of 64 immune and stromal cell types. The Microenvironment Cell Populations-counter (MCP-counter) method[87] quantifies the absolute abundance of eight immune and two stromal cell populations in heterogeneous tissues. We ran xCell and MCP-counter on the corresponding R packages. The MF (Molecular Functional) Portrait can yield qualitative and quantitative descriptions of modules built based on the expression of 29 gene expression signatures, with the size of each module corresponding to the intensity of the normalized ssGSEA score. The MFP python source code was obtained online (https://github.com/BostonGene/MFP).

Analysis for RNA-seq data

RNA-seq raw reads were mapped to the rat genome version rn7 using TopHat[78], and expression value (number of raw reads) for each gene was determined by the software HTSeq. Normalized (using the upper-quartile method) expression values and differentially expressed genes were determined by edgeR. We used the function genomeCoverageBed in BEDTools, along with the nor2total function in DANPOS[79], and the tool bedGraphToBigWig to generate a BigWig file that contains RNA-seq signal (read density) at each base pair across the genome.

Differential expression analysis for tumor and normal samples

We calculated the log2 fold change (log2FC) of the expression (TPM) for each feature gene by subtracting the log2 transformed median count in each group and used the Wilcoxon Signed Rank Test to evaluate the significance of each feature gene, with multiple hypothesis correction performed using the Benjamini–Hochberg procedure. Genes with |logFC| >0.5 in the integrated RNA-seq dataset, |logFC| >0.1 in the

integrated microarray dataset, and adjusted p-value < 0.05 in both, were considered DEGs. In addition, the linear correlation between feature genes' expression levels in the microarray and RNA-seq datasets was calculated using the geom_lmannotate function, and the significance of the overlap between DEGs and feature gene classification was calculated using the hypergeometric test.

We ranked patient samples by MIF expression levels and constructed MIF-high-variable datasets from the top 50 and tail 50 samples in the TCGA and ICGC datasets. Genes with |logFC| >0 and a p-value < 0.05 in both datasets were considered DEGs. We constructed the SPP1-high-variable datasets using the same method. The upregulated and downregulated genes were separated in the high-variable datasets of the two genes. The hypergeometric distribution test was used to calculate the significance of the degree of overlap of the upregulated and downregulated genes.

Construction of gene regulatory networks

Pearson correlation coefficients were calculated for cancer-associated TFs and shared DEGs in MIF- or SPP1-high-variable datasets in the TCGA or ICGC. Shared interactions with a p-value < 0.05, and a |correlation coefficient| > 0.4 in both the TCGA and ICGC datasets were selected. The degree of centrality was calculated with Pearson's correlation coefficient as the weight. Cancer-related TFs were obtained through CistromeCancer (https://cistrome.org/CistromeCancer/CancerTarget/), and gene regulatory networks were mapped using the 'ggraph' R package.

Survival analysis

Clinical data were obtained from the TCGA-LIHC, ICGC-LIHC-JP, and GEO datasets. The Cox proportional-hazards regression models were analyzed using the survival R package to test the correlation between selected genes or microenvironment types and patients' survival. The Kaplan–Meier survival curves were plotted and compared using the log-rank test. The ggsurvplot function in the survminer R package was used for data visualization.

Functional enrichment analysis

The pathway enrichment analysis of feature genes was performed by the 'clusterProfiler' R package.[84] P-values were obtained using an empirical phenotype-based permutation test, and pathways with p-values < 0.05 were selected. The p-values were corrected for multiple hypotheses using the Benjamini–Hochberg method. GO biological processes (GOBP) with an adjusted p-value <0.05 were employed. The GOBP gene sets were obtained using the msigdbr R package.

The gene set enrichment analysis (GSEA) of the MNPs was performed using the fgsea R package.[85] The DEGs in each cell subtype were calculated using the Wilcox function in the R package presto, and the logFC was replaced by $AUC_{ROC.}$ Genes with |logFC| >0 and p < 0.05 in MIF- or SPP1-high-variable datasets were subjected to KEGG pathway enrichment analysis by the same method described above.

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistical details of experiments can be found in the figure legends, and the statistical details of bioinformatics analysis can be found in each related part of the method details section.