The Prognostic Potential and Pro-tumorigenicity of Serglycin was Associated with Tumor Microenvironment in Liver Cancer

Chengzeng Zou[1], Yue Li[1], Wenhai Guo[1], Jiongshan Zhang[1]☆, Yongwei Li[1]☆

[1] Department of Traditional Chinese Medicine, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China

☆ Corresponding author: Dr. Yongwei Li, E-Mail: liyongw@mail.sysu.edu.cn, Jiongshan Zhang [1]☆, E-Mail:zhjshan2@mail.sysu.edu.cn

Chengzeng Zou, Yue Li, and Wenhai Guo are co-authors

600 Tianhe Road, Guangzhou, Guangdong, China. Tel: +0086 020 85253497

Supported by Science and Technology Planning Project of Guangdong Province, China, no. 2017A030313738, 2022A1515011689

The authors declare no potential conflicts of interest.

Abstract

Background Serglycin (SRGN) is a prominent hematopoietic proteoglycan and regulates tumorigenesis in malignancies. However, the role of serglycin is unclear in pan-cancer, especially liver hepatocellular cancer (LIHC).

Materials and methods The expression and prognostic potential of SRGN in LIHC and other pan-cancer was investigated using the bioinformatics databases PrognoScan, GEPIA, Kaplan-Meier Plotter, and TIMER. HepG2 cells were transfected with a SRGN overexpressing vector. Proliferation, invasion, sorafenib resistance, and vasculature were examined *in vitro*. A subcutaneous xenograft tumor was constructed in nude mice.

Results The prognostic potential of SRGN was inconsistent in pan-cancer. The Kaplan-Meier Plotter indicated that SRGN expression was a favorable prognostic factor correcting for stage, grade, AJCC_T stage, sex, race, alcohol consumption, hepatitis virus infection, and sorafenib treatment in liver cancer. TIMER 2.0 showed that T cells CD8+, macrophage M1, macrophage M2, and endothelial cells(ECs) were strongly correlated with SRGN mRNA expression ($r=0.552$, $P=5.79e^{-29}$; $r=0.517$, $P=5.84e^{-25}$; $r=0.696$, $P=3.26e^{51}$; and $r=0.522$, $P=1.67e^{-25}$, respectively), and had prognostic potential in LIHC in the cohort with low or high levels of SRGN (HR=0.53, P=0.048; HR=3.09, P=0.00633; HR=1.99, P=0.0366; and HR=0.364, P=0.00638, respectively). SRGN promoted HepG2 cell *in vitro* and *vivo* proliferation, weak sorafenib resistance, invasion, and vasculature. CD206 and CD80 were up-regulated and down-regulated, respectively in subcutaneous tumor tissues.

Conclusions The prognostic potential of serglycin is limited in LIHC. However, the pro-tumorigenic properties of SRGN and its closely association with tumor microenvironment cells may contribute to the development of new treatment strategies.

Key words: Prognosis; Liver hepatocellular carcinoma; Serglycin; Pro-tumorigenicity

Introduction

More than 257 million people are infected with hepatitis B virus (HBV) worldwide [1]. HBV has a broad spectrum of clinical manifestations ranging from a healthy carrier state to acute and chronic hepatitis [chronic hepatitis B (CHB)], liver cirrhosis (LC), primary liver cancer [mainly liver hepatocellular carcinoma (LIHC)], and end-stage liver disease. Most patients recover from acute HBV infection, and 1–5% of patients develop LIHC [2]. HBV-specific B-cell antibodies and functional HBV-specific T cells terminate HBV infections [3]. HBV infection-based microenvironment consists of immunosuppressive and inflammatory cytokines and immune cells (macrophages, neutrophils, lymphocytes, etc.) [4-6]. It promotes more complication in LIHC tumor microenvironment (TME). Immunotherapy PD-1/PD-L1 blockade was regarded as a hotspot and promising therapy in HCC [4]. However, irAE (immune-related adverse event) caused by PD-1/PD-L1 inhibitor has unbearable side-effects. To clarify the mechanism linking immune and malignancy is important to investigate new treatment to LIHC. Serglycin protein (SRGN gene) with features of inflammatory and malignancy is a prominent hematopoietic proteoglycan. It interacts with proteases, chemokines, cytokines, and is required for the formation of secretory granules [7-9]. The HepG2.2.15 cell line is derived from HepG2 human hepatoma cells chronically infected with HBV [10]. In our previous work, we analyzed the Human Genome U133 Plus 2.0 Array (Platform GPL570; Affymetrix, Santa Clara, CA) and showed the presence of extensive changes in global gene transcription in HepG2.215 compared with HepG2 cells. SRGN was the most significantly differentially expressed gene [11]. Further study showed that the levels of serglycin were increased in peripheral

monocytes, neutrophils, and lymphocytes in HBV-infected LIHC patients compared with healthy adults. Hematopoietic SRGN was associated with unfavorable prognosis regarding over survival (OS) in LIHC [12]. Based on these results, we investigated comprehensively the prognostic potential of SRGN in LIHC by bioinformatics databases with reference to the study of serglycin in pan-cancer.

Materials and methods

PrognoScan database analysis

The prognostic potential of SRGN expression in cancers was analyzed using the PrognoScan database from microarray datasets (http://www.abren.net/ PrognoScan/) [13].

Gene expression profiling interactive analysis

The Gene Expression Profiling Interactive Analysis (GEPIA) database (http://gepia. cancer-pku.cn/index.html, http://gepia2021.cancer-pku.cn/) was used to analyze the data from The Cancer Genome Atlas (TCGA) and the GTEx projects [14]. The survival curves of OS and relapse-free survival (RFS) were generated with the Cox model and log-rank test in pan-cancer. SRGN expression and immune cells were analyzed in LIHC.

Kaplan-Meier plotter database analysis

The association of SRGN expression with survival was evaluated using the Kaplan-Meier (*KM*) plotter database in breast, liver, ovarian, lung, and gastric cancer (http://kmplot.com/analysis/(registration-free*KM-plotter*)) [15]. Differential expression of

SRGN was studied in pan-cancer, and analyzed in tumor, normal, and metastatic tissues of LIHC.

TIMER database analysis

The Tumor Immune Estimation Resource (TIMER) was used to analyze tumor-infiltrating cell types in the tumor microenvironment (TME) using data from TCGA (https:// cistrome.shinyapps.io/timer/, http:// timer.comp-genomics.org/)[16]. The prognostic potential of SRGN expression was analyzed with or without immune infiltrates including B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, dendritic cells, and subsets by TIMER2.0 Gene-Outcome, TIMER Survival, and TIMER2.0 Immune Outcome, respectively.

Correlation module were used to investigate correlations between SRGN and gene markers of CD8+ T cells, monocytes, TAMs, M1 macrophages, M2 macrophages, or six serglycin-associated genes by the Spearman's correlation analysis. The results were represented as heatmap in pan-cancer and scatter plots between SRGN and gene markers in LIHC, respectivelt. SRGN expression level was used for the Y-axis, and related genes for the X-axis.

Overexpression of SRGN in HepG2 cell line

A detailed description is provided in the Supplementary Methods including lentiviral transduction,quantitative PCR and Western blotting, CCK8 assay of cell viability and Transwell invasion assay, angiogenesis and vasculogenic mimicry, and immuno-fluorescence and HE staining.

Animal experiments

Male BALB/c nude mice (5-weeks-old) were purchased from Guangzhou University of Chinese Medicine Laboratory Animal Center (Guangzhou, China). Tumor cells of 5×10^6/0.1 mL PBS were randomly injected into the right axillary region. The mice were monitored twice per week for 4 weeks for tumor formation. The animal experiments were approved by the Laboratory Animal Center of Third Affiliated Hospital of Sun Yat-Sen University.

Statistical analysis

Survival curves were generated using PrognoScan, GEPIA, *KM-plotter*, and the TIMER database. The results are expressed as Cox P-values or log-rank P-value, and hazard ratio (HR) with or without 95% confidence interval (CI). The value of Spearman's correlations was determined using the following criteria: 0.1–0.3: weak correlation, 0.3–0.5: moderate correlation, and 0.5–1.0: strong correlation (Cohen, J., Statistical power analysis for the behavioral sciences (2nd ed.); Hillsdale, NJ: Lawrence Erlbaum Associates (1988); page 83). Data were analyzed using SPSS 22.0 (IBM Corp. Armonk, NY, USA). Measurement data were compared by *t* test or ANOVA. Count data were compared with the Chi-square test. Differences between data means with skewed distribution and variance were analyzed with the rank sum test. $P < 0.05$ was considered statistically significant.

Results

Expression levels of SRGN mRNA in pan-caner

In a previous study, we showed that SRGN was the most downregulated gene in

HepG2.215 cells compared with HepG2 cells. The 30 most differentially expressed genes are shown in Figure 1 and Supplementary Excel.

The differential expression of SRGN mRNA between tumor and adjacent normal tissues was analyzed by TIMER and the results are shown in Figure 1. SRGN levels were significantly higher in head and neck cancer (HNSC) and kidney renal clear cell carcinoma (KIRC) (red) than in adjacent normal tissues (blue). SRGN expression was higher in HPV-positive HNSC than in HPV-negative ones. SRGN levels were significantly lower in bladder urothelial carcinoma(BLCA), breast invasive carcinoma (BRCA), colon adenocarcinoma, kidney chromophobe, kidney renal papillary cell carcinoma, LIHC, lung squamous cell carcinoma, lung adenocarcinoma (LUAD), prostate adenocarcinoma, rectum adenocarcinoma, and uterine corpus endometrial carcinoma (red) than blue adjacent normal tissues(blue). SRGN was lower in skin cutaneous melanoma (SKCM) tumor than SKCM metastasis (violet). However, GEPIA showed that SRGN levels did not differ between LIHC tissues and normal tissues (Figure 2A, Supplementary 2). The *KM-plotter* analysis showed that the levels of SRGN mRNA had significant difference between tumor and adjacent normal tissues in 19 cancers except for adrenal, liver and renal_PA cancers (Supplementary 3). The *KM-plotter* analysis of Genechip data showed that SRGN levels were significantly higher in adjacent normal tissues than LIHC tumor tissues (non-paired tumor vs. adjacent normal tissues, $P = 5.85e^{-11}$; paired tumor *vs.* adjacent normal tissues, $P = 9.29e^{-10}$). This was consistent with the comparison between tumor, normal, and metastasis of LIHC ($P = 7.92e^{-11}$, Figure 2B). SRGN levels from RNA-Seq data showed no differences.

Prognostic potential of SRGN in pan-Cancer according to PrognoScan, GEPIA, Kaplan-Meier Plotter Database, and TIMER

The relationship between SRGN expression and prognosis in pan-cancer is shown in Supplementary 4-7 and Table 1. By PrognoScan analysis, SRGN expression had a significant impact on prognosis in five cancer types, including acute myeloid leukemia (AML), diffuse large B-cell lymphoma (DLBCL), BRCA, lung cancer and ovarian cancer (OV). In AML and DLBCL, high SRGN expression was associated with poor prognosis. However, SRGN showed favorable prognostic potential in three solid tumors: BRCA for distant metastasis-free survival (DMFS) in 86/29 (patients with low/high expression of SRGN) and 91/195 patients, respectively; lung cancer for RFS in 182/22 patients; and OV for OS in 87/46 patients. According to GEPIA, SRGN had prognostic value only in testicular germ cell tumors (TGCT), glioblastoma multiforme(GBM), and SKCM. According to the gene_outcome module of TIMER, SRGN was a significant prognostic factor for OS in sarcoma, SKCM, and SKCM metastasis correcting for age, race, sex, and tumor purity (Supplementary Figure 6 and Table 1).

The *KM-plotter* was used to examine the prognostic potential of SRGN mRNA in pan-cancer based on genechip and RNAseq. High SRGN levels were associated with poor prognosis for OS, PFS and PPS in OV (Supplementary 7A). In BRCA, SRGN was a risk factor for PPS (post progression survival) and RFS, and a protective factor for OS (Supplementary 7B). In gastric cancer, high SRGN expression was associated with a favorable prognosis regarding OS, first progression (FP), and PPS (Supplementary 7C). In lung cancer, high SRGN levels were significantly associated with favorable or poor

prognosis for OS (201859_at and 201858_s_at, respectively), and favorable prognosis for FP (Supplementary 7D).

SRGN is a favorable prognostic factor in liver cancer according to the Kaplan-Meier Plotter Database.

The relationship between SRGN expression and the survival of liver cancer patients was investigated using the *KM-plotter*. High expression of SRGN was associated with better OS, RFS, PFS, and DSS after correcting for stages, grades, AJCC_T stages, vascular invasion, sex, race, sorafenib treatment, hepatitis virus, and alcohol consumption (Supplementary 7E; RNA_SeqID 5552; PFS HR = 0.67, 95% CI = 0.5–0.91, P = 0.0095, median survival 16.73 *vs.* 30.4 months; OS HR = 0.68, 95% CI = 0.47–0.97, P = 0.034, median survival 52 *vs.* 70.5 months; RFS HR = 0.65, 95% CI = 0.46–0.92, P = 0.015, median survival 21.87 *vs.* 40.97 months; and DSS HR = 0.58, 95% CI = 0.35–0.97, P = 0.034). However, the median survival for DSS was 104.17 *vs.* 84.4 months in the low *vs.* high expression cohort. The survival prognostic potential of serglycin may change with time. It was a time-dependent covariate.

SRGN was further analyzed in association with clinicopathological factors in liver cancer (Table 2). High SRGN mRNA expression was correlated with favorable OS and PFS in stage 2–3+4, RFS in stages 2–3, and DSS in stages 2+3–3+4. In AJCC_T 2 and 3 , higher SRGN expression levels had better OS, PFS, and DSS. In AJCC_T 2, high SRGN expression was associated with favorable RFS. High SRGN expression was associated with better OS and DSS in Grade 2 cancer, better RFS in grades 2–3 cancer, and better PFS in grades 1–3 tumors. High SRGN expression was associated with favorable OS, RFS, PFS,

and DSS in men, Asians, or patients with alcohol consumption history, and with favorable OS in women or white patients. Overexpression of SRGN in patients with hepatitis virus infection was associated with favorable prognosis for RFS and PFS, whereas in those without virus infection, SRGN was associated with favorable OS and DSS.

The prognostic potential of SRGN expression associated with immune cells in pan-cancer

SRGN expression associated with six immune infiltrates was analyzed by TIMER 2.0. In cholangiocarcinoma (CHOL), LIHC, LUAD, SKCM-metastasis, and thymoma (THYM), SRGN expression was a prognostic factor correcting for six immune infiltrates and clinical factors (age, sex, stage, and tumor purity). High SRGN expression was a risk factor for OS in CHOL, LUAD, and THYM, whereas it was a favorable prognostic factor in LIHC and SKCM metastasis (Supplementary 8). Only the log-rank test in SKCM metastasis was significantly different between the groups with low and high SRGN levels (P=0, Supplementary 9).

The correlation and prognostic potential of SRGN with six immune cells in LIHC

SRGN expression was significantly positively correlated with CD4+ T cells, B cells, macrophages, CD8+ T cells, neutrophils, and dendritic cells (Supplementary10). The correlation was moderate to strong (r = 0.34, 0.483, 0.579, 0.616, 0.631, and 0.714, respectively; P = $0.40e^{-11}$, $1.73e^{-21}$, $7.02e^{-32}$, $4.16e^{-37}$, $9.47e^{-40}$, and $2.94e^{54}$, respectively). Cancer types (LIHC), clinical characteristics (age, stage, tumor purity, sex, and race), six immune infiltrates, and gene symbols (SRGN) were analyzed using a multivariable Cox proportional hazard model. Of 305 patients enrolled, 101 died. The results showed that race

(black), stage 4, macrophages, and dendritic cells were risk factors for OS, whereas stage 3, CD4+ T cells, CD8+ T cells, and SRGN were favorable factors (Supplementary 8). However, the log-rank test of six immune cells and SRGN showed no significant survival differences (Supplementary 9).

The correlation and prognostic potential of SRGN with TME cells in LIHC

Each immune cell subset and other TME cell types were subjected to analysis of correlation and outcome using TIMER 2.0. SRGN expression was significantly negatively correlated with tumor purity (Rho = -0.53, P = $1.67e^{-26}$). Eleven cell subsets were negatively correlated with SRGN expression *vs.* 32 positively correlated cell subsets (Table 3, Supplementary 11). Immune cells included ten T cell CD4+ subsets (T cell CD4+, T cell CD4+ naïve (by two algorithms CIBERSORT and XCELL), T cell CD4+ central memory, T cell CD4+ effector memory, T cell CD4+ memory activated, T cell CD4+ memory resting, T cell CD4+ memory, T cell CD4+ (non-regulatory), and T cell CD4+ Th1, T cell CD4+ Th2), two T cell CD8+ subsets (T cell CD8+ central memory and T cell CD8+ effector memory), three B cell subsets (B cell naïve, B cell plasma_(CIBERSORT-ABS and XCELL), and B cell memory), T cell regulatory, T cell follicular helper, monocytes, four macrophage subsets (macrophage_(EPIC and TIMER), macrophage M0, macrophage M1, macrophage M2_(CIBERSORT-ABS and TIDE), common lymphoid progenitors, neutrophils, three NK cell subsets (NK cell, NK cell resting, and NK cell activated), three myeloid dendritic cell subsets (myeloid dendritic cell**,** myeloid dendritic cell activated, and myeloid dendritic cell resting), and T cell NK. T cell CD4+ naïve, B cell plasma, macrophage, and macrophage M2 had negative or positive correlation with SRGN expression as determined by different

algorithm, *e g*., macrophage M2 (CIBERSORT-ABS: r = 0.696, P = $3.26e^{51}$; TIDE: r = -0.516, P = $0.66e^{-25}$). T cell CD8+, macrophage M1, myeloid dendritic cells, NK cell activated, and ECs had a strong and positive correlation with SRGN expression levels (r = 0.552, P = $5.79e^{-29}$; r = 0.517, P = $5.84e^{-25}$; r = 0.627, P = $4.97e^{-39}$; r = 0.524, P = $1.11e^{-25}$; r = 0.522, P = $1.67e^{-25}$). The analysis *via* GEPIA2021 database showed that the proportion of macrophages M2 was the highest in LIHC, LIHC normal, and liver tissues, respectively; T cell CD8+, macrophages M1 and M2 were higher in LIHC and LIHC normal tissues than in non-cancerous liver tissue, whereas monocytes were the highest in non-cancerous liver tissue than in LIHC and LIHC normal tissues; SRGN levels are consistent with immune cell distribution. The levels of SRGN in immune cells of LIHC and LIHC normal tissues were distributed as M2>M1> T cell CD8+> monocytes> M0 (Figure C-F). The remaining six cell types [such as hematopoietic stem cells (HSCs) and myeloid-derived suppressor cells (MDSCs)] showed no correlation with SRGN expression.

We further analyzed the correlations between SRGN expression and the biomarkers of T cell CD8+, macrophages M1 and M2, TAMs, monocytes in pan-cancer, especially LIHC (Supplementary 12 and 13). All of them and CTLA4, PDCD1, CD274 , CD80, CD44, VEGFC, and TGFβ2 were positively correlated with SRGN in LIHC.

TME infiltrates (each of 21 cell types were analyzed), clinical characteristics (age, stage, purity, sex, and race), and SRGN were analyzed for OS of LIHC using the Coxph Model. The cell types and subsets are shown in Table 4 and Figure 3 with SRGN as a significant prognostic factor in the Cox model. By log-rank test, high infiltration of MDSCs,

macrophage M1, or monocytes_MCPCOUNTER) were associated with unfavorable OS in the cohort with high SRGN expression (MDSCs HR = 2.78, P = 0.00668; M1 HR = 3.09, P = 0.00633; monocytes HR = 2.75, P = 0.025), whereas high infiltration of T cell CD8+, T cell CD4+ memory resting, ECs, or HSCs was associated with favorable prognosis for OS (T cell CD8+ HR = 0.53, P = 0.048; T cell CD4+ memory resting HR = 0.501, P = 0.0383; ECs HR = 0.364, P = 0.00638; HSCs HR = 0.379, P = 0.00868). In the cohort with low SRGN expression, high infiltration of macrophage_TIMER, macrophage M0_CIBERSORT, and macrophage M2_CIBERSORT-ABS were unfavorable factors for OS (macrophage HR = 1.82, P = 0.0432; M0 HR = 1.86, P = 0.0356; M2 HR = 1.99, P = 0.0366), all of them and T cell CD8+ showed time-dependent survival curves. The log-rank test of SRGN expression with the other six cell types showed no significant difference in LIHC (T cell CD4+ central memory, T cell CD4+ Th2, Tregs, monocytes (XCELL), myeloid dendritic cells, and myeloid dendritic cell resting).

The *in vitro* and *in vivo* protumorigenicity of SRGN

SRGN promoted the proliferation of SRGN-overexpressing HepG2SG cells compared with the control HepG2SG-NC cells ($^{*}P < 0.05$, $^{**}P < 0.01$, Figure 4A-C). Sorafenib treatment showed weak resistance in HepG2SG cells (1 < RI < 2, RI: resistance index, Figure 4D). The weight of subcutaneous xenograft tumors of HepG2SG cells was higher than that of HepG2SG-NC cells ($^{*}P < 0.05$, Figure 4E). SRGN and CD206 were up-regulated, while CD80 down-regulated significantly in subcutaneous tumor tissues in HepG2SG xenograft mice compared with HepG2 SG-NC mice ($^{*}P < 0.05$, $^{**}P < 0.01$,

Figure 4F). HE staining showed that the cells were distributed in clumps or strands in subcutaneous tumor tissues. The mean number of tubular structures was higher in HUVECs cultured with HepG2SG supernatants than those with HepG2SG-NC supernatants ($^{***}P \leq 0.001$, Figure 4I). SRGN promoted a higher number of VM tubes in HepG2SG cells than that in HepG2SG-NC cells ($^{**}P < 0.01$, Figure4J).

Discussion

SRGN is a hematopoietic intracellular proteoglycan expressed in secretory granules and vesicles and secreted to the extracellular medium [17]. The core protein of SRGN consists of eight serine/glycine repeats attached by variable glycosaminoglycan (GAG) bonds depending on cell type and status [18-19]. SRGN has a pro-tumorigenic effect to benefit tumor cells in multiple ways [20]. Several malignant cells depend on SRGN to promote aggregation [21]. However, it remains controversial.

By TIMER, GEPIA, and *KM-Plotter* database analysis, the SRGN levels were the highest in AML than other cancers. SRGN expression was a prognostic risk factor for survival in four AML groups and two DLBCL groups by PrognoScan analysis. The prognostic potential of SRGN mRNA expression were consistent with those of SRGN protein expression examined by immunohistochemistry (IHC) and ELISA [22], and the pro-tumorigenic effect of SRGN was recently identified again in AML [23]. However, the consistency could not be shown in solid tumors.

IHC studies in lung, breast, prostate, and colon cancer cells showed that SRGN expression was up-regulated in advanced tumors and the activated TME [24]. In the present

study, the levels of SRGN mRNA were decreased in both BRCA and lung cancer compared with those in normal adjacent tissues according to TIMER, *KM-plotter*, and GEPIA as well as reported by Oncomine [25]. By PrognoScan analysis, high levels of SRGN mRNA were associated with a favorable prognosis in OV, lung cancer, and BRCA in a small sample size. However, the expression of SRGN mRNA was also a prognostic risk factor in a large sample size in the three cancers *via KM-plotter*. Further study identified the pro-tumorigenic effect of SRGN in lung cancer and BRCA. SRGN was correlated with drug resistance and stemness in BRCA, promoting invasion, metastasis and inducing BRCA cell aggregation [25-29]. High levels of SRGN predict a poor prognosis in LUAD, promoting NSCLC cell migration [30-31]. The result about SKCM was consistent with the report that serglycin was a favorable factor in SKCM and SKCM-metastasis tissues via GEPIA and TIMER 2.0[25], which indicated that our study was reproducible and reliable. The prognostic potential and tumorigenicity of serglycin has rarely studied in sarcoma, GBM, TGCT, gastric cancer, CHOL and THYM.

The controversy is more evident in LIHC. The levels of SRGN mRNA were lower in LIHC than in adjacent normal tissues via TIMER, whereas they were higher in liver cancer *vs.* normal tissues via Oncomine [25]. There was no difference via GEPIA, and inconsistency between Genechip data and RNA-seq data via *KM-plotter*. IHC staining showed that SRGN levels were higher in 56.7% of HCC tissues *vs.* 3.1% of adjacent, non-tumor tissues [9]. In our prior study, the levels of hematopoietic SRGN examined by flow cytometry were higher in CHB, LC, and LIHC than in healthy adults; whereas the levels of SRGN mRNA were decreased in HepG2.215 compared with HepG2 cells [11-12].

Transfected with SRGN-overexpressing vector, HepG2 with low *in vivo* tumorigenicity showed significant proliferation *in vitro* and *in vivo*. LIHC generally develops from hepatitis or cirrhosis induced by HBV; SRGN expression may be affected by inflammation and malignant transformation, which leads to more complexity in follow survival analysis.

The expression of SRGN mRNA showed no prognostic potential in LIHC *via* PrognoScan, GEPIA or TIMER 2.0 (analyzed by TIMER2.0 Gene_Outcome module (with no TME cells analyzed)). According to the *KM-plotter* database, SRGN mRNA was a favorable prognostic factor after adjusting for stage, grade, AJCC_T, sex, race, alcohol consumption, hepatitis virus infection, sorafenib treatment, and vascular invasion. However, the expression of SRGN protein was an unfavorable factor for OS or time to recurrence [9, 12]. The prognostic potential of SRGN transcriptome and protein was contradictory as lung cancer and BRCA. As a prominent hematopoietic proteoglycan, SRGN should be analyzed together with TME cells in the multivariate Cox regression model.

The liver is the major hematopoietic organ during the fetal period. Hepatic hematopoiesis may generate the lymphoid, myeloid, and erythroid lineages, and is involved in tumor surveillance and rejection [32-33]. SRGN maintains homeostasis of the body's immune cell populations, controlling the magnitude and durability of the immune response [34]. In the present study, six prominent immune cells showed strong correlations with SRGN. SRGN was a favorable prognostic factor for OS associated with macrophages, dendritic cells, T cell CD4+, and T cell CD8+. However, there are abundant immune cell subsets, the Cox regression were further analyzed.

SRGN has been studied in neutrophils, cytotoxic T lymphocytes, macrophages, T cell CD8+, platelets, and NK cells [8, 35-39]. The negative correlation between SRGN and tumor purity suggested that SRGN may be prominently expressed in TME infiltrating cells, and the correlation between SRGN and TME cells was reliable. The results showed that monocytes were the most prevalent in normal liver tissues than in LIHC and LIHC normal tissues. Macrophages. M2, acquire pro-tumorigenic properties by responding to CCL2 and IL-10, which triggers the release of cytokines, chemokines, and growth factors to promote malignancy [40]. But M2 cells with immunosuppressive properties are the highest among the five immune cells in healthy liver tissues as in LIHC and LIHC normal tissues. Maybe for the liver is an organ with natural immune tolerance [4]. M1 and T cell CD8+ were differentiated for antagonizing pro-tumorigenisis. Both levels of them were second to M2 in LIHC and LIHC normal tissues. The results suggested that the differentiation or activation of macrophages/ monocytes may vary in the physiological and pathological liver tissues. SRGN can be induced in macrophages [35]. The GAG of SRGN varies according to the differentiation of monocytes to macrophages or the activation of macrophages [18]. The present result showed that the levels of SRGN mRNA in T cell CD8+ and macrophage subsets/monocytes were consistent with the proportional distribution of these cells in LIHC, LIHC normal and healthy liver tissues. The Spearman's correlation analysis indicated the significant correlations between SRGN and the biomarkers or these cells in LIHC. The *in vivo* results identified CD206 increased, while CD80 down-regulated in HepG2SG mice. It was contradictory to the positive correlation between SRGN and CD80 *via* TIMER2.0.

SRGN might promote macrophage/monocyte differentiation, or migration to LIHC as reported in myeloma [41]. The TME was immunosuppressive in HepG2SG mice.

The survival analysis showed that macrophage (TIMER), M0, and pro-tumorigenic M2 associated with low SRGN levels, while monocytes (MCPCOUNTER) and “antitumorigenic” M1 with high SRGN levels, were unfavorable prognostic factors. The prior associated with low SRGN might maintain the immunosuppressive status. The latter with high SRGN might promote excessive proinflammatory status into immune exhaustion or tolerance status. It’s similar as that M1 cultured with GM-CSF converted “antitumorigenic” to promote tumor metastasis and invasion (Zhupei Guan, Jinbo Yang. M1 macrophages combine with platelets and tumor cells to promote tumor metastasis. Master's Thesis, Lanzhou University, 2020). The prognostic potential of SRGN was closely related to macrophage/monocyte subsets. The positive correlations between SRGN and CTLA4, PD-1 and PD-L1 also suggested that SRGN was related to the immune escape. The immune therapy, *e g*., PD-1 inhibitor, might have effect on SRGN-positive LIHC.

T cell CD4 subsets had extensive, weak to moderate correlations with SRGN (seven positive and three negative). However, only high T cell CD4+ memory resting and T cell CD8 in the cohort with high SRGN expression showed protective effects against LIHC by the log-rank test. Maybe high levels of SRGN play a role in the cytotoxicity of T cells to LIHC as it has been reported in virus infection [37, 39]. MDSCs have immunosuppressive activity against T-cell responses. High SRGN and MDSCs indicated an unfavorable

prognosis. Further study is warranted to investigate the interaction between SRGN and these immune cells.

The prognostic potential of SRGN mRNA had no difference between patients with vascular invasion or not via *KM-plotter*. However, It has been reported that high SRGN may up-regulate the proliferation of HUVECs [42]. The SRGN protein was positively associated with vascular invasion in HCC patients [9]. The present study showed that ECs and VEGFC were strongly correlated with SRGN, and SRGN induced vasculature of both HUVECs and HepG2. High levels of the circulating HSCs and endothelial progenitor cells were associated with worse OS in HCC patients [43]. However, high SRGN mRNA levels with high HSCs or ECs were associated with favorable OS by TIMER 2.0 analysis. Both high levels of SRGN and ECs might promote tumor vascular normalization？Further study is to investigate the controversy.

The prognostic capability may be variable with different covariates in Cox regression model. Without regard to algorithms, the expression pattern of SRGN (mRNA and protein), in addition to clinicopathological factors, observation endpoint, sample size, etc., may result in contradictory survival difference in pan-cancer except for blood cancers. One of two SRGN transcript variants is protein-coding [GEPIA, provided by RefSeq, Jul 2010]. The present study constructed SRGN-overexpressing vector (based on the same sequence of ID5552 analyzed in *KM-plotter*) and identified pro-tumorigenicity and metastasis in LIHC. Otnd ti Menyhaame, et al also noted that accession for survival by protein was not consistent with that by the transcriptome [44]. They reviewed 318 genes related to HCC

survival; even none had equivalent prognostic value as stage. TME cells may have very important effect on the multivariate Cox regression model. Serglycin is a limited prognostic factor, but the study revealed relationships between SRGN and monocyte/macrophage subsets, ECs, HSCs, MDSC, T cell CD8+, or T cell CD4+ memory resting, which may contribute to new treatment strategies.

Acknowledgement We thank International Science Editing (http://www. International scienceediting.com) for editing this manuscript.

References

1. Ortega-Prieto AM, Cherry C, Gunn H, Dorner M. In vivo model systems for hepatitis B virus research. ACS Infect Dis. 2019 10; 5(5):688-702.

2. Gurtsevitch VE. Human oncogenic viruses: Hepatitis B and hepatitis C viruses and their role in hepatocarcinogenesis. Biochemistry (Moscow), 2008; 73(5): 504-513.

3. Bertoletti A, Ferrari C. Adaptive immunity in HBV infection. J Hepatol. 2016; 64(1 Suppl):S71-S83.

4. Li B, Yan C, Zhu J, Chen X, Fu Q, et al. Anti-PD-1/PD-L1 Blockade Immunotherapy Employed in Treating Hepatitis B Virus Infection-Related Advanced Hepatocellular Carcinoma: A Literature Review. Front Immunol. 2020; 11: 1037.

5. Guidotti LG, Inverso D, Sironi L, et al. Immunosurveillance of the liver by intravascular effector CD8(+) T cells. Cell 2015; 161:486–500.

6. Douglas Hanahan, Robert A Weinberg. Hallmarks of cancer: The next generation. Cell. 2011; 144: 646-674.

7. Korpetinou A, Skandalis SS, Labropoulou VT, Smirlaki G, Noulas A, Karamanos NK, Theocharis AD. Serglycin: At the crossroad of inflammation and malignancy. Front Oncol. 2014; 3: 327.

8. Xin-Jian Li and Chao-Nan Qian. Serglycin in human cancers. Chin J Cancer. 2011; 30 (9):585-9.

9. He L, Zhou X, Qu C, Tang Y, Zhang Q, Hong J. Serglycin (SRGN) overexpression predicts poor prognosis in hepatocellular carcinoma patients. Med Oncol. 2013; 30: 707.

10. Li Y, Zhu M, Guo Y, Chen W, Li G. Full-length hepatitis B virus sequences from naive patients with fluctuation of viral load during ADV monotherapy. Virus Genes 2010; 40:155-162.

11. Yongwei Li, Mingfen Zhu, Gang Li. Expression of serglycin in HepG2 cells with the different HBV infection status. Journal of Gastroenterology and Hepatology. 2013; 28(Supplement 3): 913.

12. Li Y, Chen H, Lu H, Zou Z, Li Y. Prognostic significance of hematopoietic-cell serglycin for the survival of hepatocellular carcinoma: A single-center retrospective study. Comb Chem High Throughput Screen. 2021; 24(7):986-995.

13 . Mizuno, H., Kitada, K., Nakai, K. et al. PrognoScan: a new database for meta-analysis of the prognostic value of genes. BMC Med Genomics. 2009; 2:18.

14. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017; 45(W1):W98-W102.

15. Nagy A, Munkárcsy G, Győrffy B. Pancancer survival analysis of cancer hallmark genes, Scientific Reports, 2021; 11(1): 6047.

16. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: a web server for comprehensive analysis of tumor- infiltrating immune cells. Cancer Res. 2017; 77: e108-e110.

17. Dimitra Manou, Nikos K. Karamanos, Achilleas D. Theocharis. Tumorigenic functions of serglycin: Regulatory roles in epithelial to mesenchymal transition and oncogenic signaling. Seminars in Cancer Biology.2020；62:108-115.

18. Kolset SO, Kjellen L, Seljelid R, Lindah U. Changes in glycosaminoglycan biosynthesis during differentiation in vitro of human monocytes. Biochem J .1983; 210 (3):661–7.

19. Uhlin-Hansen L, Eskeland T, Kolset, S O. Modulation of the expression of chondroitin sulfate proteoglycan in stimulated human monocytes. J Biol Chem. 1989; 264(25):14916–22

20. Angeliki Korpetinou, Spyros S Skandalis, Vassiliki T Labropoulou, Gianna Smirlaki, Argyrios Noulas, Nikos K Karamanos, Achilleas D Theocharis. Serglycin: At the crossroad of inflammation and malignancy. Front Oncol. 2014; 3: 327.

21. Manou D, Bouris P, Kletsas D, Götte M, Greve B, Moustakas A, Karamanos NK, Theocharis AD. Serglycin activates pro-tumorigenic signaling and controls glioblastoma cell stemness, differentiation and invasive potential. Matrix Biol Plus. 2020; 6-7: 100033.

22. Niemann CU, Kjeldsen L, Ralfkiaer E, Jensen MK, Borregaard N. Serglycin proteoglycan in hematologic malignancies: A marker of acute myeloid leukemia. Leukemia. 2007; 21(12):2406-10.

23. Peng L, Zhang Y, Xin H. lncRNA SNHG3 facilitates acute myeloid leukemia cell growth via the regulation of miR-758-3p/SRGN axis. J Cell Biochem. 2020; 121(2): 1023-1031.

24. Korpetinou A, Papachristou DJ, Lampropoulou A, Bouris P, Labropoulou VT, Noulas A, Karamanos NK, Theocharis AD. Increased expression of serglycin in specific carcinomas and aggressive cancer cell lines. Biomed Res Int. 2015; 2015: 690721.

25. Wang X, Xiong H, Liang D, Chen Z, Li X, Zhang K. The role of SRGN in the survival and immune infiltrates of skin cutaneous melanoma (SKCM) and SKCM-metastasis patients. BMC Cancer. 2020; 20(1): 378.

26. Yang T, Fu Z, Zhang Y, Wang M, Mao C, Ge W. Serum proteomics analysis of candidate predictive biomarker panel for the diagnosis of trastuzumab-based therapy resistant breast cancer. Biomed Pharmacother. 2020; 129: 110465.

27. Zhang Z, Qiu N, Yin J, Zhang J, Liu H, Guo W, Liu M, Liu T, Chen D, Luo K, Li H, He Z, Liu J, Zheng G. SRGN crosstalk with YAP to maintain chemoresistance and stemness in breast cancer cells by modulating HDAC2 expression. Theranostics. 2020 4;

10(10):4290-4307.

28. Bouris P, Manou D, Sopaki-Valalaki A, Kolokotroni A, Moustakas A, Kapoor A, Iozzo RV, Karamanos NK, Theocharis AD. Serglycin promotes breast cancer cell aggressiveness: Induction of epithelial to mesenchymal transition, proteolytic activity and IL-8 signaling. Matrix Biol. 2018; 74: 35-51.

29. Zhang Z, Deng Y, Zheng G, Jia X, Xiong Y, Luo K, Qiu Q, Qiu N, Yin J, Lu M, Liu H, Gu Y, He Z.SRGN- TGFbeta2 regulatory loop confers invasion and metastasis in triple-negative breast cancer. Oncogenesis. 2017; 6(7):e360.

30. Guo JY, Chiu CH, Wang MJ, Li FA, Chen JY.J Proteoglycan serglycin promotes non-small cell lung cancer cell migration through the interaction of its glycosaminoglycans with CD44. Biomed Sci. 2020; 27(1):2.

31. Guo JY, Hsu HS, Tyan SW, Li FY, Shew JY, Lee WH, Chen JY. Serglycin in tumor microenvironment promotes non-small cell lung cancer aggressiveness in a CD44-dependent manner. Oncogene. 2017 ; 36(17): 2457-2471.

32. Kotton D. N., Fabian A. J., Mulligan R. C. A novel stem-cell population in adult liver with potent hematopoietic-reconstitution activity. Blood. 2005; 106(5):1574–1580.

33. Golden-Mason L., Curry M. P., Nolan N., et al. Differential expression of lymphoid and myeloid markers on differentiating hematopoietic stem cells in normal and tumor-bearing adult human liver. Hepatology. 2000; 31(6):1251–1256.

34. Kolset SO, Pejler G. Serglycin: a structural and functional chameleon with wide impact on immune cells. J Immunol. 2011;187(10):4927–33.

35. Zernichow L, Abrink M, Hallgren J, Grujic M, Pejler G, Kolset SO. Serglycin is the major secreted proteoglycan in macrophages and has a role in the regulation of macrophage tumor necrosis factor-alpha secretion in response to lipopolysaccharide. J Biol Chem. 2006; 281(37):26792–801.

36. Niemann CU, Abrink M, Pejler G, Fischer RL, Christensen EI, Knight SD, et al. Neutrophil elastase depends on serglycin proteoglycan for localization in granules. Blood. 2007; 109(10): 4478– 86.

37. Grujic M, Christensen JP, Sorensen MR, Abrink M, Pejler G, Thomsen AR. Delayed contraction of the CD8+ T cell response toward lymphocytic choriomeningitis virus infection in mice lacking serglycin. J Immunol. 2008; 181(2):1043–51.

38. Chanzu H, Lykins J, Wigna-Kumar S, Joshi S, Pokrovskaya I, Storrie B, Pejler G, Wood JP, Whiteheart SW. Platelet alpha-granule cargo packaging and release are affected by the luminal proteoglycan, serglycin. J Thromb Haemost. 2021; 19(4):1082-1095.

39. Sutton VR, Brennan AJ, Ellis S, Danne J, Thia K, Jenkins MR, Voskoboinik I, Pejler G, Johnstone RW, Andrews DM, Trapani JA. Serglycin determines secretory granule repertoire and regulates natural killer cell and cytotoxic T lymphocyte cytotoxicity. FEBS J. 2016 ; 283(5):947-61.

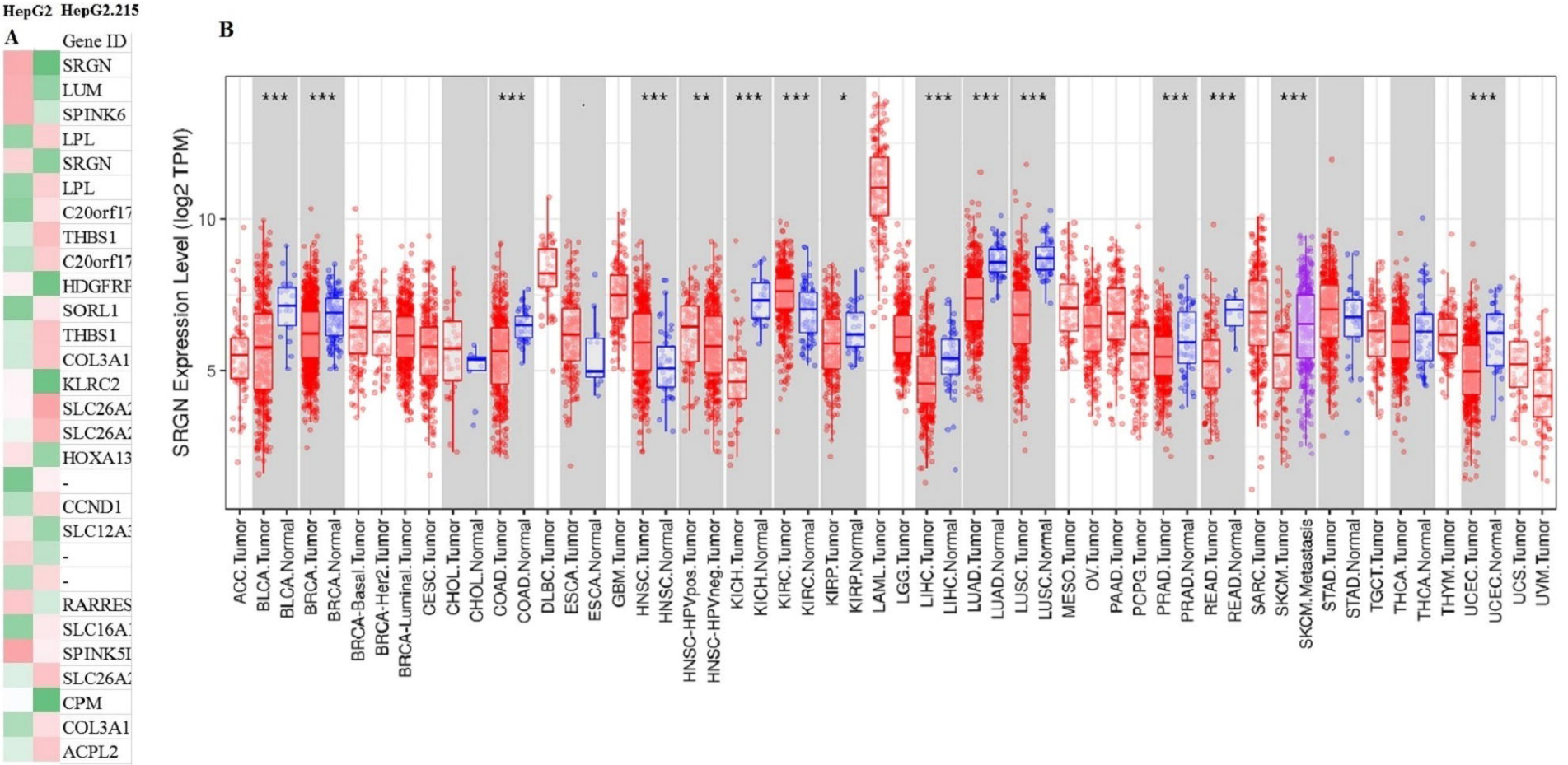
40. EvansA, CostelloE. The role of inflammatory cells in fostering pancreatic cancer cell growth and invasion. Front Physiol. 2012; 3:270.

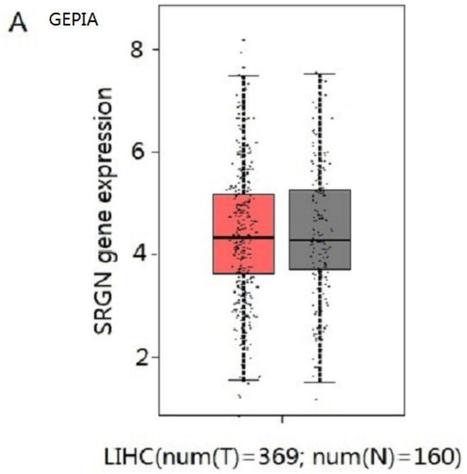
41. Purushothaman A, Bandari SK, Chandrashekar DS, Jones RJ, Lee HC, Weber DM, Orlowski RZ.Chondroitin sulfate proteoglycan serglycin influences protein cargo loading and functions of tumor-derived exosomes. Oncotarget. 2017; 8(43):73723- 73732.

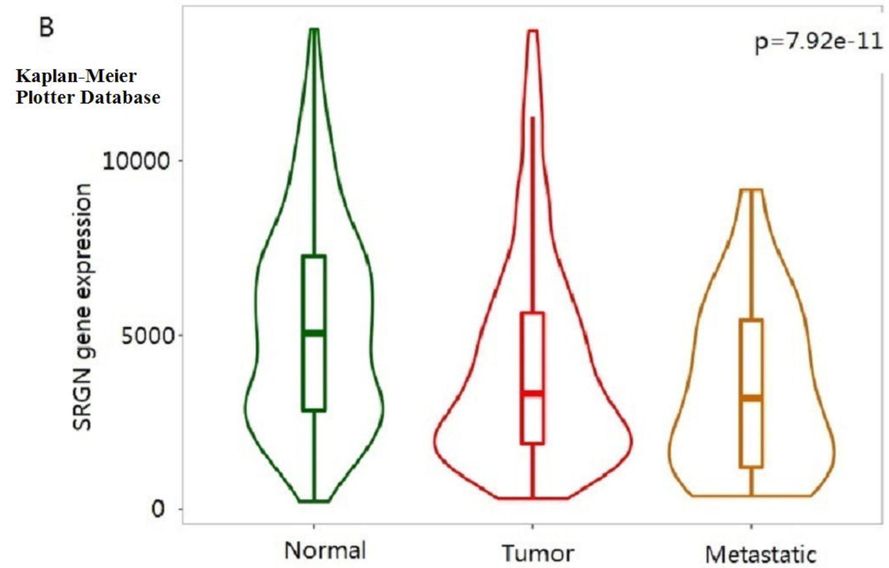
42. Ma Q, Gu W, Li T, Zhang K, Cui Y, Qu K, Wang N, Humphry R, Durkan C, Qiu J, Wang G. SRGN, a new identified shear-stress- responsive gene in endothelial cells. Mol Cell Biochem. 2020; 474(1-2): 15-26.

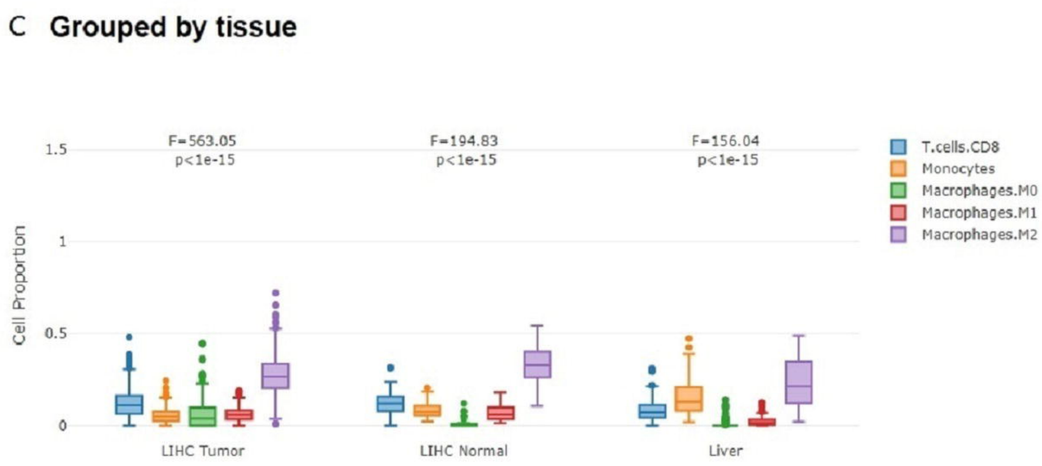
43. Zahran AM, Abdel-Rahim MH, Refaat A, Sayed M, Othman MM, Khalak LMR, Hetta HF.Circulating hematopoietic stem cells,endothelialprogenitor cells and cancer stem cells in hepatocellular carcinoma patients: contribution to diagnosis and prognosis. Acta Oncol. 2020; 59(1):33-39.

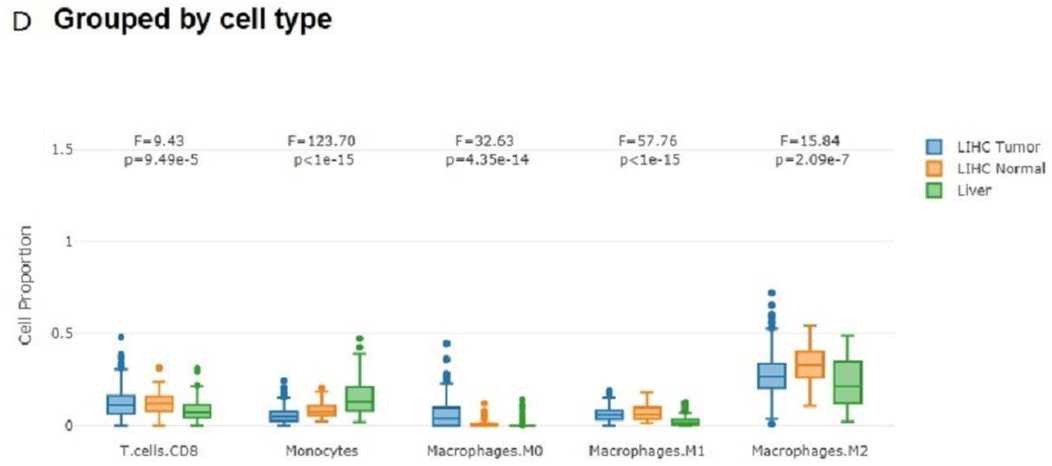
44 Menyhart O, Nagy A, Gyorffy B. Determining consistent prognostic biomarkers of overall survival and vascular invasion in hepatocellular carcinoma. R Soc Open Sci**.** 2018; 5(122):181006.

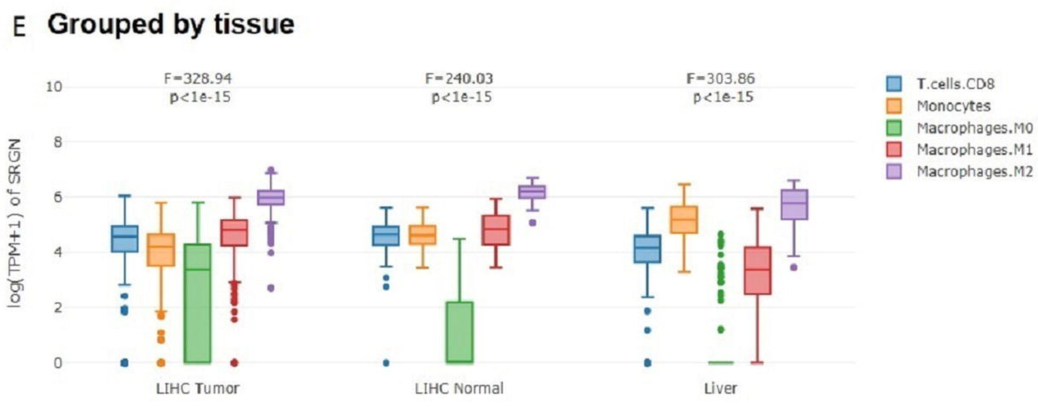

HepG2
HepG2.215
A
Gene ID
SRGN
LUM
SPINK6
LPL
SRGN
LPL
C20orf17
THBS1
C20orf17
HDGFRF
SORL1
THBS1
COL3A1
KLRC2
SLC26A2
SLC26A2
HOXA13
-
CCND1
SLC12A2
-
-
RARRES
SLC16A1
SPINK5I
SLC26A2
CPM
COL3A1
ACPL2
B
SRGN Expression Level (log2 TPM)
10
5
ACC.Tumor
BLCA.Tumor
BLCA.Normal
BRCA.Tumor
BRCA.Normal
BRCA-Basal.Tumor
BRCA-Her2.Tumor
BRCA-Luminal.Tumor
CESC.Tumor
CHOL.Tumor
CHOL.Normal
COAD.Tumor
COAD.Normal
DLBC.Tumor
ESCA.Tumor
ESCA.Normal
GBM.Tumor
HNSC.Tumor
HNSC.Normal
HNSC-HPVpos.Tumor
HNSC-HPVneg.Tumor
KICH.Tumor
KICH.Normal
KIRC.Tumor
KIRC.Normal
KIRP.Tumor
KIRP.Normal
LAML.Tumor
LGG.Tumor
LIHC.Tumor
LIHC.Normal
LUAD.Tumor
LUAD.Normal
LUSC.Tumor
LUSC.Normal
MESO.Tumor
OV.Tumor
PAAD.Tumor
PCPG.Tumor
PRAD.Tumor
PRAD.Normal
READ.Tumor
READ.Normal
SARC.Tumor
SKCM.Tumor
SKCM.Metastasis
STAD.Tumor
STAD.Normal
TGCT.Tumor
THCA.Tumor
THCA.Normal
THYM.Tumor
UCEC.Tumor
UCEC.Normal
UCS.Tumor
UVM.Tumor

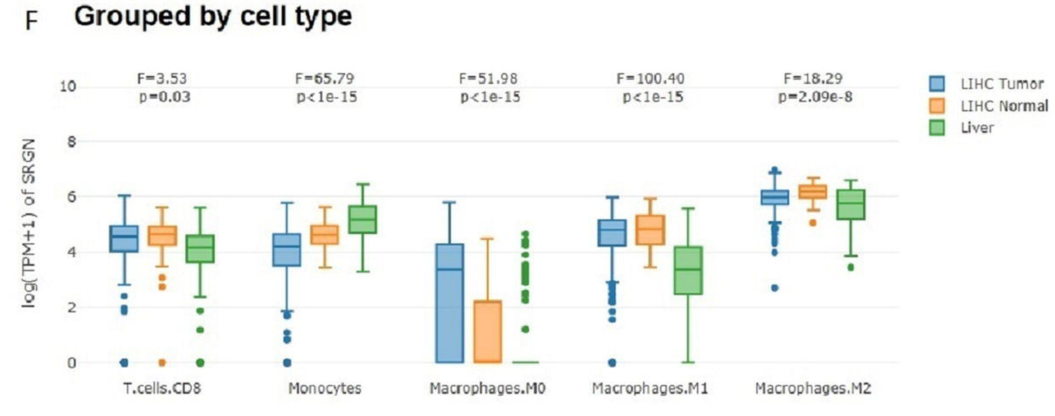

A
GEPIA
SRGN gene expression
8
6
4
2
LIHC(num(T)=369; num(N)=160)

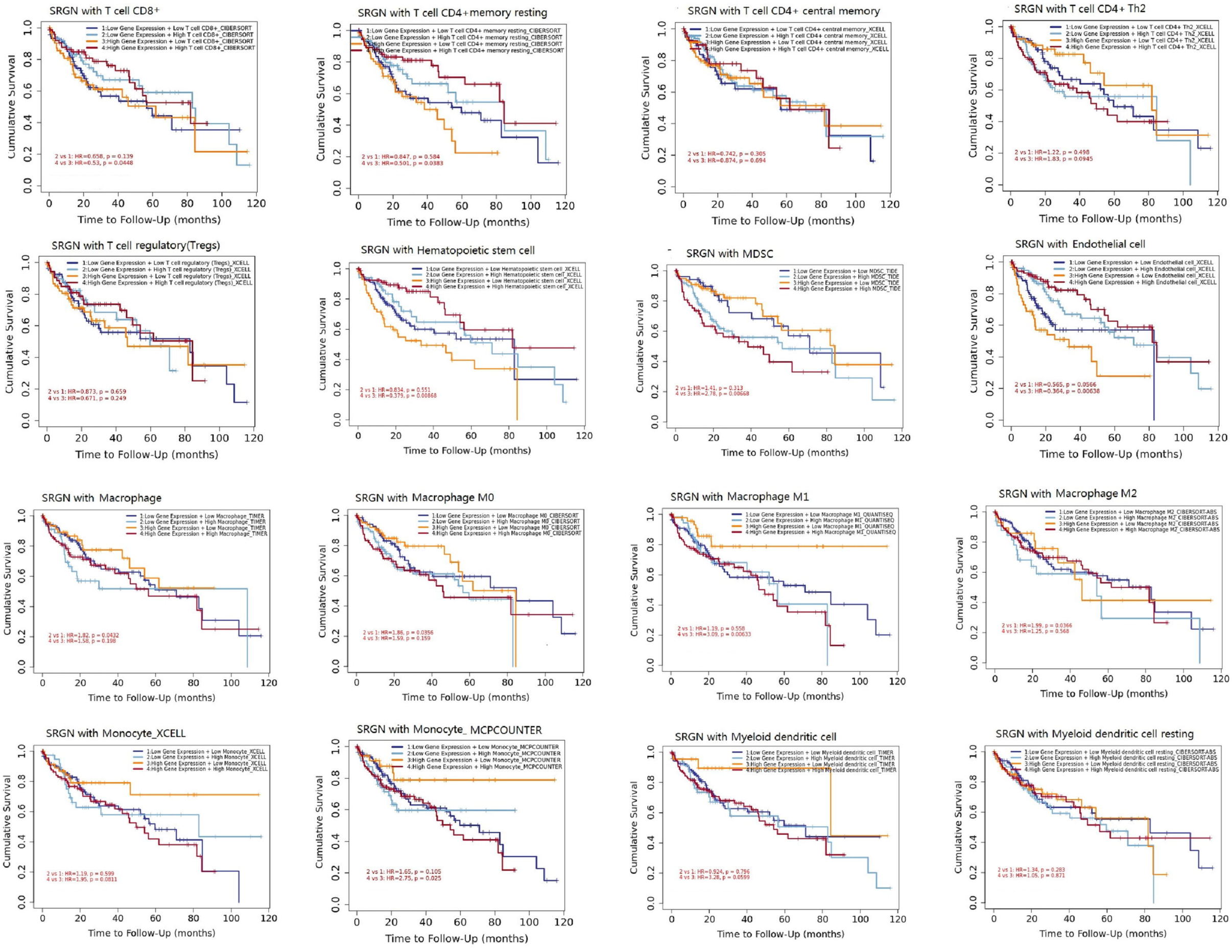

B
Kaplan-Meier
Plotter Database
p=7.92e-11
SRGN gene expression
10000
5000
0
Normal
Tumor
Metastatic

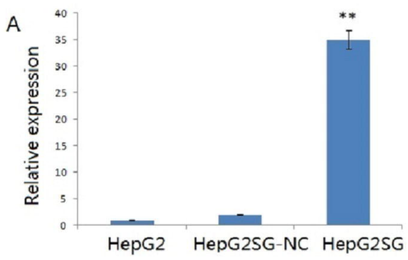

C Grouped by tissue
F=563.05
p<1e-15
F=194.83
p<1e-15
F=156.04
p<1e-15
T.cells.CD8
Monocytes
Macrophages.M0
Macrophages.M1
Macrophages.M2
Cell Proportion
1.5
1
0.5
0
LIHC Tumor
LIHC Normal
Liver

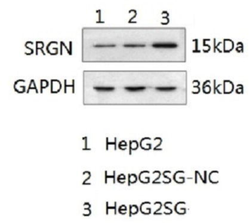

D Grouped by cell type
F=9.43
p=9.49e-5
F=123.70
p<1e-15
F=32.63
p=4.35e-14
F=57.76
p<1e-15
F=15.84
p=2.09e-7
LIHC Tumor
LIHC Normal
Liver
Cell Proportion
1.5
1
0.5
0
T.cells.CD8
Monocytes
Macrophages.M0
Macrophages.M1
Macrophages.M2

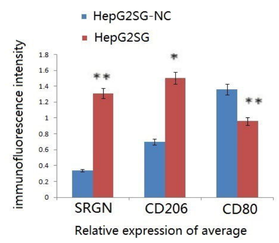

E Grouped by tissue
F=328.94
p<1e-15
F=240.03
p<1e-15
F=303.86
p<1e-15
T.cells.CD8
Monocytes
Macrophages.M0
Macrophages.M1
Macrophages.M2
log(TPM+1) of SRGN
10
8
6
4
2
0
LIHC Tumor
LIHC Normal
Liver

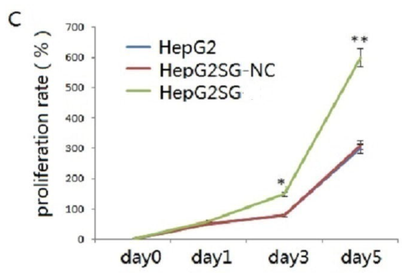

F Grouped by cell type
F=3.53
p=0.03
F=65.79
p<1e-15
F=51.98
p<1e-15
F=100.40
p<1e-15
F=18.29
p=2.09e-8
LIHC Tumor
LIHC Normal
Liver
log(TPM+1) of SRGN
10
8
6
4
2
0
T.cells.CD8
Monocytes
Macrophages.M0
Macrophages.M1
Macrophages.M2

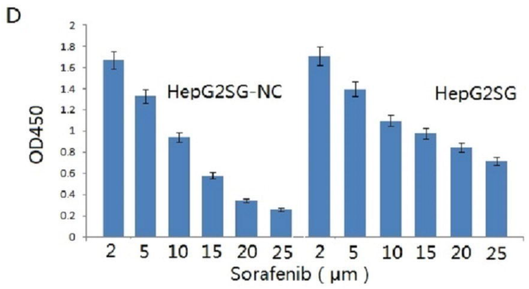

SRGN with T cell CD8+
1:Low Gene Expression + Low T cell CD8+_CIBERSORT
2:Low Gene Expression + High T cell CD8+_CIBERSORT
3:High Gene Expression + Low T cell CD8+_CIBERSORT
4:High Gene Expression + High T cell CD8+_CIBERSORT
2 vs 1: HR=0.658, p = 0.139
4 vs 3: HR=0.53, p = 0.0448
SRGN with T cell CD4+memory resting
1:Low Gene Expression + Low T cell CD4+ memory resting_CIBERSORT
2:Low Gene Expression + High T cell CD4+ memory resting_CIBERSORT
3:High Gene Expression + Low T cell CD4+ memory resting_CIBERSORT
4:High Gene Expression + High T cell CD4+ memory resting_CIBERSORT
2 vs 1: HR=0.847, p = 0.584
4 vs 3: HR=0.501, p = 0.0383
SRGN with T cell CD4+ central memory
1:Low Gene Expression + Low T cell CD4+ central memory_XCELL
2:Low Gene Expression + High T cell CD4+ central memory_XCELL
3:High Gene Expression + Low T cell CD4+ central memory_XCELL
4:High Gene Expression + High T cell CD4+ central memory_XCELL
2 vs 1: HR=0.742, p = 0.305
4 vs 3: HR=0.874, p = 0.694
SRGN with T cell CD4+ Th2
1:Low Gene Expression + Low T cell CD4+ Th2_XCELL
2:Low Gene Expression + High T cell CD4+ Th2_XCELL
3:High Gene Expression + Low T cell CD4+ Th2_XCELL
4:High Gene Expression + High T cell CD4+ Th2_XCELL
2 vs 1: HR=1.22, p = 0.498
4 vs 3: HR=1.83, p = 0.0945
SRGN with T cell regulatory(Tregs)
1:Low Gene Expression + Low T cell regulatory (Tregs)_XCELL
2:Low Gene Expression + High T cell regulatory (Tregs)_XCELL
3:High Gene Expression + Low T cell regulatory (Tregs)_XCELL
4:High Gene Expression + High T cell regulatory (Tregs)_XCELL
2 vs 1: HR=0.873, p = 0.659
4 vs 3: HR=0.671, p = 0.249
SRGN with Hematopoietic stem cell
1:Low Gene Expression + Low Hematopoietic stem cell_XCELL
2:Low Gene Expression + High Hematopoietic stem cell_XCELL
3:High Gene Expression + Low Hematopoietic stem cell_XCELL
4:High Gene Expression + High Hematopoietic stem cell_XCELL
2 vs 1: HR=0.834, p = 0.551
4 vs 3: HR=0.379, p = 0.00868
SRGN with MDSC
1:Low Gene Expression + Low MDSC_TIDE
2:Low Gene Expression + High MDSC_TIDE
3:High Gene Expression + Low MDSC_TIDE
4:High Gene Expression + High MDSC_TIDE
2 vs 1: HR=1.41, p = 0.313
4 vs 3: HR=2.78, p = 0.00668
SRGN with Endothelial cell
1:Low Gene Expression + Low Endothelial cell_XCELL
2:Low Gene Expression + High Endothelial cell_XCELL
3:High Gene Expression + Low Endothelial cell_XCELL
4:High Gene Expression + High Endothelial cell_XCELL
2 vs 1: HR=0.565, p = 0.0566
4 vs 3: HR=0.364, p = 0.00638
SRGN with Macrophage
1:Low Gene Expression + Low Macrophage_TIMER
2:Low Gene Expression + High Macrophage_TIMER
3:High Gene Expression + Low Macrophage_TIMER
4:High Gene Expression + High Macrophage_TIMER
2 vs 1: HR=1.82, p = 0.0432
4 vs 3: HR=1.58, p = 0.198
SRGN with Macrophage M0
1:Low Gene Expression + Low Macrophage M0_CIBERSORT
2:Low Gene Expression + High Macrophage M0_CIBERSORT
3:High Gene Expression + Low Macrophage M0_CIBERSORT
4:High Gene Expression + High Macrophage M0_CIBERSORT
2 vs 1: HR=1.86, p = 0.0356
4 vs 3: HR=1.59, p = 0.159
SRGN with Macrophage M1
1:Low Gene Expression + Low Macrophage M1_QUANTISEQ
2:Low Gene Expression + High Macrophage M1_QUANTISEQ
3:High Gene Expression + Low Macrophage M1_QUANTISEQ
4:High Gene Expression + High Macrophage M1_QUANTISEQ
2 vs 1: HR=1.19, p = 0.558
4 vs 3: HR=3.09, p = 0.00633
SRGN with Macrophage M2
1:Low Gene Expression + Low Macrophage M2_CIBERSORT-ABS
2:Low Gene Expression + High Macrophage M2_CIBERSORT-ABS
3:High Gene Expression + Low Macrophage M2_CIBERSORT-ABS
4:High Gene Expression + High Macrophage M2_CIBERSORT-ABS
2 vs 1: HR=1.99, p = 0.0366
4 vs 3: HR=1.25, p = 0.568
SRGN with Monocyte_XCELL
1:Low Gene Expression + Low Monocyte_XCELL
2:Low Gene Expression + High Monocyte_XCELL
3:High Gene Expression + Low Monocyte_XCELL
4:High Gene Expression + High Monocyte_XCELL
2 vs 1: HR=1.19, p = 0.599
4 vs 3: HR=1.95, p = 0.0811
SRGN with Monocyte_ MCPCOUNTER
1:Low Gene Expression + Low Monocyte_MCPCOUNTER
2:Low Gene Expression + High Monocyte_MCPCOUNTER
3:High Gene Expression + Low Monocyte_MCPCOUNTER
4:High Gene Expression + High Monocyte_MCPCOUNTER
2 vs 1: HR=1.65, p = 0.105
4 vs 3: HR=2.75, p = 0.025
SRGN with Myeloid dendritic cell
1:Low Gene Expression + Low Myeloid dendritic cell_TIMER
2:Low Gene Expression + High Myeloid dendritic cell_TIMER
3:High Gene Expression + Low Myeloid dendritic cell_TIMER
4:High Gene Expression + High Myeloid dendritic cell_TIMER
2 vs 1: HR=0.924, p = 0.796
4 vs 3: HR=3.28, p = 0.0599
SRGN with Myeloid dendritic cell resting
1:Low Gene Expression + Low Myeloid dendritic cell resting_CIBERSORT-ABS
2:Low Gene Expression + High Myeloid dendritic cell resting_CIBERSORT-ABS
3:High Gene Expression + Low Myeloid dendritic cell resting_CIBERSORT-ABS
4:High Gene Expression + High Myeloid dendritic cell resting_CIBERSORT-ABS
2 vs 1: HR=1.34, p = 0.283
4 vs 3: HR=1.05, p = 0.871
Cumulative Survival
Time to Follow-Up (months)

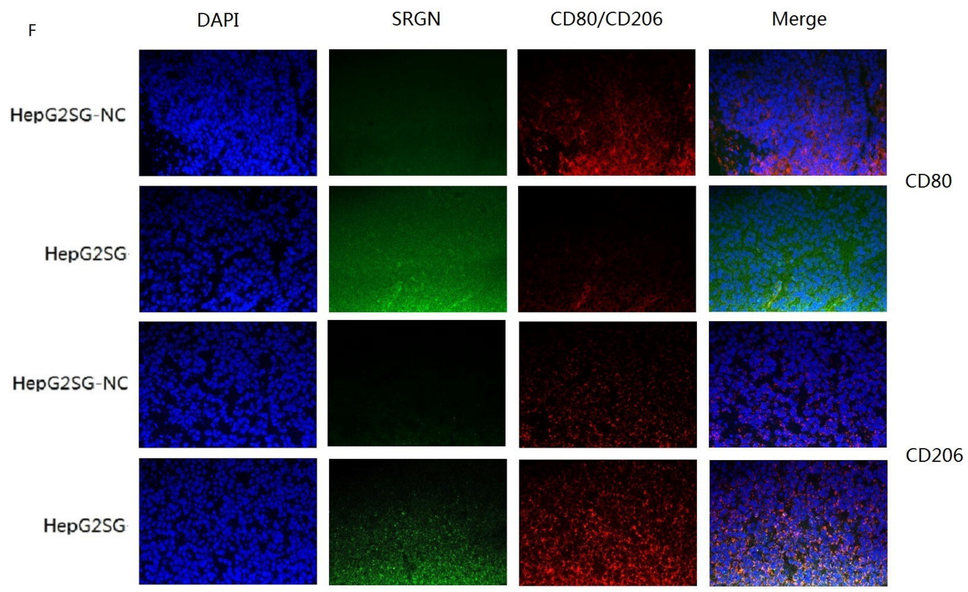

A
Relative expression
HepG2
HepG2SG-NC
HepG2SG

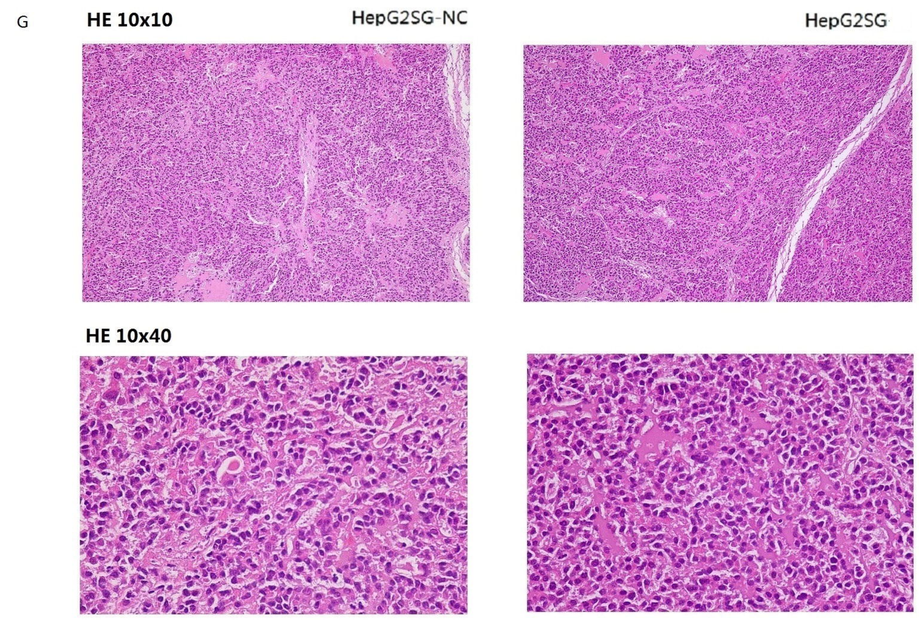

B
1 2 3
SRGN
15kDa
GAPDH
36kDa
1 HepG2
2 HepG2SG-NC
3 HepG2SG

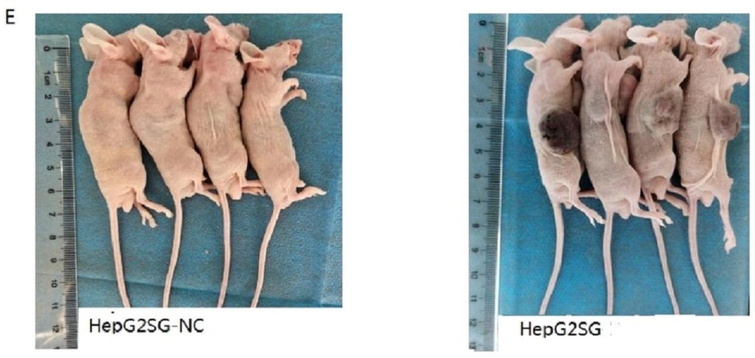

HepG2SG-NC
HepG2SG
immunofluorescence intensity
SRGN
CD206
CD80
Relative expression of average

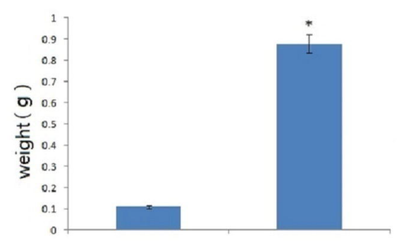

C
HepG2
HepG2SG-NC
HepG2SG
proliferation rate (%)
day0
day1
day3
day5

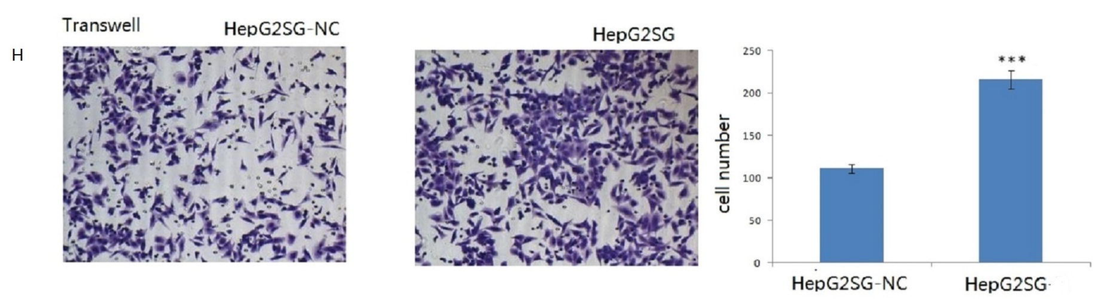

D
HepG2SG-NC
HepG2SG
OD450
Sorafenib (μm)

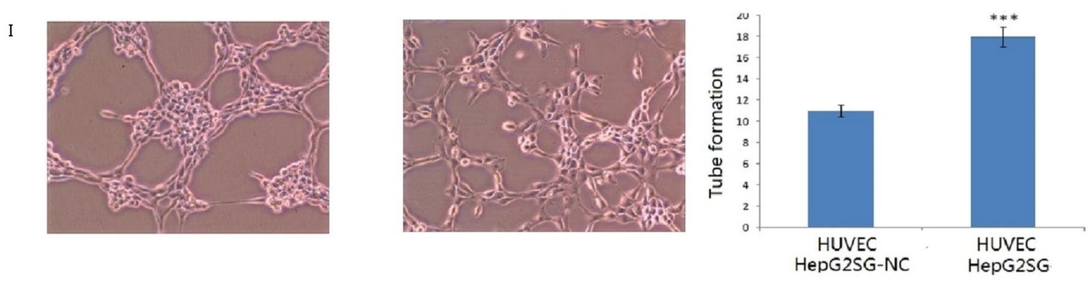

F
DAPI
SRGN
CD80/CD206
Merge
HepG2SG-NC
HepG2SG
CD80
HepG2SG-NC
HepG2SG
CD206


G
HE 10x10
HepG2SG-NC
HepG2SG
HE 10x40


E
HepG2SG-NC
HepG2SG


weight (g)


H
Transwell
HepG2SG-NC
HepG2SG
cell number
HepG2SG-NC
HepG2SG


I
Tube formation
HUVEC
HepG2SG-NC
HUVEC
HepG2SG

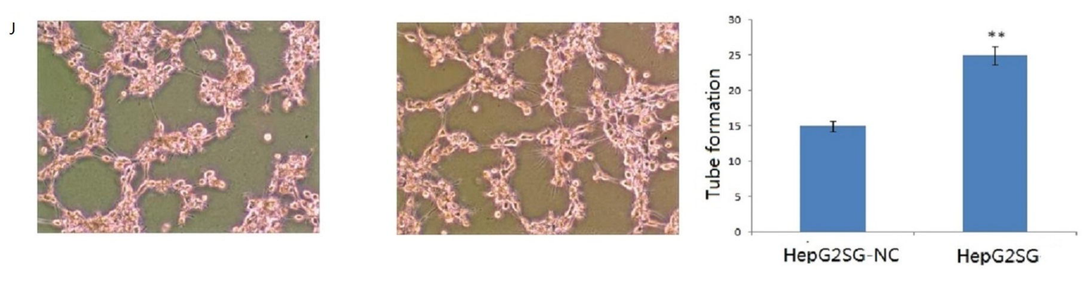

J
Tube formation
HepG2SG-NC
HepG2SG

Figure legend

Figure1(A) The prior study of 30 differential genes between HepG2.215 and HepG2 using Human Genome U133 Plus 2.0 Array; (B) DiffExp of SRGN in tumor (red column) was compared with adjacent normal tissues (Blue column) across all TCGA tumors by the Wilcoxon test via TIMER(Diff Exp module) (**P*< 0.05, ***P*<0.01, ****P*< 0.001).

Figure2 (A) The expression of SRGN mRNA in LIHC and normal tissues via GEPIA; (B) SRGN gene expression in tumor, normal and metastatic liver cancer tissues via the Kaplan-Meier plotter database; (C, D) grouped by tissue and cell type, the cell proportion of T cell CD8+, monocytes, macrophages.M0, M2 and M1 in non-cancerous liver, LIHC tumor and LIHC normal tissues via GEPIA2021; (E, F) grouped by tissue and cell type, the levels of SRGN in T cell CD8+, monocyte, macrophages. M0, M1 and M2 in LIHC tumor, LIHC normal and liver via GEPIA 2021.

Figure 3 Kaplan-Meier survival curves comparisons between the cohorts with the high and low levels of SRGN expression associated with high and low levels of infiltrating cells in LIHC via TIMER2.0.

Figure 4 In *vitro* and *in vivo* experiment of the pro-tumorigenicity of SRGN (A) Relative expression of SRGN in HepG2SG (overexpressing SRGN) and HepG2SG-NC (transfected with control blank vector) examined by quantitative RT-PCR; (B) The expression of serglycin protein in HepG2SG and HepG2SG-NC examined by Western-blotting; (C) the proliferation rate of HepG2SG and HepG2SG-NC by CCK8 assay (*p<0.05, **p<0.01); (D) The proliferation rate of

HepG2SG and HepG2SG-NC cells treated with sorafenib, 2> resistance index>1; (E) Subcutaneous xenograft tumor with HepG2SG and HepG2SG-NC cells in nude mice (*$p<0.05$) ; (F) Immunofluorescence of subcutaneous tumor tissues. Both SRGN and CD206 were up-regulated, while CD80 down-regulated in HepG2SG mice compared with HepG2SG-NC ones (*$p<0.05$, **$p<0.01$). Blue light for DAPI staining, green light for FITC marked SRGN staining, red light for CY3 marked CD80 or CD206 staining; (G) HE staining for subcutaneous tumor tissues. (H) Invasion of HepG2SG and HepG2SG-NC examined by Transwell assay (***$p<0.001$); (I) Angiogenesis of umbilical vein endothelial cells (HUVEC) cultured with supernatant of HepG2SG and HepG2SG-NC, respectively (***$p\leq0.001$); (J) Vasculogenic mimicry of HepG2SG-R and HepG2SG-NC (**$p<0.01$).

.

Table1 The correlation of serglycin expression and prognosis in cancers via four bioinformatics databases

Database and cancer	**ID**	**Patients**	**Survival**	**Hazard ratio(HR)**	**p-value**
PrognoScan	ID	Sample size	Survival	HR(95%CI)	Cox P
AML	1554676_at	35/44	OS	1.35(1.03-1.76)	0.03
	201859_at	69/10	OS	2.59(1.01-6.63	0.05
	201858_s_at	9/49	OS	1.99(1.31-3.0)	0.001
	201859_at	36/22	OS	2.67(1.48-4.82)	0.001
DLBCL	201858_s_at	22/31	EFS	2.49(1.03-6.02)	0.04
	201858_s_at	22/31	OS	2.59(1.02-6.56)	0.05
Breast cancer	201859_at	86/29	DMSF	0.31(0.12-0.83)	0.02
	201858_s_at	91/195	DMSF	0.77(0.62-0.96)	0.02
Lung cancer	1554676_at	182/22	RFS	0.46(0.29-0.72)	0.001
Ovarian cancer	201859_at	87/46	OS	0.82(0.70-0.97)	0.03
GEPIA[#]	ID	Sample size	Survival	HR(high)	p-value
TGCT	NA	68/68	OS	9.1e8	0.022
GBM	NA	80/81	RFS	1.7	0.013
SKCM	NA	228/229	OS	0.51	5.2e-07
SKCM	NA	228/229	RFS	0.74	0.014
TIMER2.0	ID	Sample size	Survival	HR	p-value
SARC	NA	260	OS	0.762	0.0213
SKCM	NA	471	OS	0.746	0.000873
SKCM metastasis	NA	368	OS	0.734	0.00184
ACC	NA	79	OS	1.49	0.0659
LGG	NA	516	OS	1.19	0.0691
Kaplan-Meier plotter[*]	ID	Sample size	Survival	HR(95%CI)	p-value
Breast cancer	201859 _at	458	PPS	1.39(1.09-1.76)	0.0066
	201859_at	4929	RFS	1.15(1.03-1.27)	0.012
	201858_s_at	1879	OS	0.74(0.61-0.91)	0.0038
	201858_s_at	4929	RFS	1.13(1.01-1.27)	0.034
Ovarian cancer	201858_s_at	1435	PFS	1.26,(1.11-1.45)	0.00054
	201859_at	1435	PFS	1.3(1.12-1.5)	0.00038
	201859_at	1656	OS	1.15(1-1.33)	0.051
	201859_at	782	PPS	1.21(1-1.46)	0.05
Lung cancer	201859_at	982	FP	0.58(0.48- 0.71)	9.9e-08
	201859_at	1925	OS	0.78 (0.69 -0.89)	0.00022
	201858_s_at	1925	OS	1.25(1.1-1.42)	0.00081
Gastric cancer	201858_s_at	875	OS	0.55(0.46-0.66)	6.3e-12
	201858_s_at	498	PPS	0.49(0.38-0.62)	9.7e-10
	201858_s_at	640	FP	0.57(0.47-0.7)	4.8e-08
	201859_at	640	FP		1e-07

	201859_at	875	OS	0.57(0.46-0.7)	2.1e-10
	201859_at	498	PPS	0.58(0.49-0.69)	3.8e-07
				0.56(0.44-0.7)	
Liver cancer	5552	366	PFS	0.67(0.5-0.91)	0.0095
	5552	364	OS	0.68(0.47-0.97)	0.034
	5552	313	RFS	0.65((0.46-0.92)	0.015
	5552	357	DSS	0.58(0.35-0.97)	0.034

The survival analysis of *SRGN* expression in pan-cancer. NA: the data was not shown; *Split patients by auto select best cutoff, other factors such as pathology, risk factors, patient, and follow up threshold, etc., were chosen all; [#]group cutoff: median, cutoff -high/low: 50%. AML, acute myeloid leukemia ; DLBCL, diffuse large B-cell lymphoma ; LGG, brain lower grade glioma; ACC, adrenocortical carcinoma; SKCM, skin cutaneous melanoma; TGCT, testicular germ cell tumor; SARC, sarcoma; GBM, glioblastoma multiforme; OS, overall survival; PFS, progression free survival; RFS, relapse-free survival; PPS, post progression survival; FP, first progression; DSS, disease-specific survival; DMFS, distant metastasis-free survival; EFS, event free survival.

Table 2 Correlation of *SRGN* mRNA expression and prognosis in liver cancer with different clinical factors by Kaplan-Meier plotter.

Clinical characteristics*	**OS (n= 364)**			**RFS(n=313)**		
	N	**Hazard ratio**	**p-value**	**N**	**Hazard ratio**	**p-value**
Male	246	0.56(0.36-0.88)	**0.01**	208	0.56(0.36-0.88)	**0.0051**
Female	118	0.49(0.24-0.99)	**0.044**	105	0.56(0.28-1.1)	0.085
White	181	0.58(0.36-0.93)	**0.022**	147	0.75(0.48-1.18)	0.21
Asian	155	0.51(0.28-0.94)	**0.028**	143	0.55(0.33-0.92)	**0.022**
Hepatitis virus(yes)	150	0.72(0.38-1.37)	0.31	138	0.53(0.31-0.88)	**0.014**
(no)	167	0.39(0.23-0.64)	**0.00014**	142	0.65(0.38-1.12)	0.12
Alcohol consumption(yes)	115	0.35(0.18-0.7)	**0.0017**	98	0.52(0.28-0.97)	**0.037**
(no)	202	0.65(0.38-1.1)	0.1	182	0.63(0.38-1.02)	0.06
Sorafenib treatment	29	0.51(0.17-1.54)	0.23	22	0.41(0.16-1.07)	0.06
Vascular invasion(none)	203	0.65(0.38-1.12)	0.12	175	0.74(0.45-1.21)	0.23
(micro)	90	2.14(0.74-6.22)	0.15	81	0.59(0.27-1.3)	0.18
AJCC_T(1)	180	0.76(0.42-1.39)	0.37	160	1.27(0.72-2.24)	0.4
(2)	90	0.44(0.21-0.95)	**0.031**	79	0.46(0.21-1.01)	**0.046**
(3)	78	0.46(0.23-0.9)	**0.021**	65	0.56(0.27-1.15)	0.11
Grade (1)	55	0.44(0.17-1.14)	0.084	45	1.74(0.64-4.71)	0.27
2	174	0.55(0.33-0.94)	**0.026**	147	0.53(0.31-0.93)	**0.024**
3	118	1.3(0.69-2.46)	0.42	106	0.54(0.31-0.93)	**0.025**
Stage 1	170	0.73(0.39-1.38)	0.34	153	1.55(0.87-2.76)	0.14
1+2	253	0.67(0.42-1.09)	0.11	227	0.74(0.48-1.13)	0.16
2	83	0.45(0.2-1.01)	**0.046**	74	0.51(0.25-1.05)	0.063
2+3	166	0.49(0.3-0.79)	**0.0028**	142	0.52(0.33-0.81)	**0.0033**
3	83	0.41(0.21-0.8)	**0.0071**	68	0.54(0.29-1.01)	0.051
3+4	87	0.45(0.24-0.85)	**0.011**	68	0.54(0.29-1.01)	0.051

Clinical characteristics		PFS (n=366)			DSS(n=357)	
	N	Hazard ratio	p-value	N	Hazard ratio	p-value
Male	246	0.54 (0.37-0.77)	**0.00073**	241	0.53 (0.29-0.97)	**0.037**
Female	120	0.56 (0.31-1.01)	0.051	116	0.41 (0.16-1.08)	0.062
White	183	0.73 (0.49-1.07)	0.11	177	0.62 (0.35-1.11)	0.1
Asian	155	0.58 (0.36-0.94)	**0.025**	152	0.31 (0.14-0.68)	**0.002**
Hepatitis virus(yes)	152	0.55 (0.34-0.88)	**0.011**	149	1.54 (0.67-3.55)	0.31
(no)	167	0.64 (0.39-1.03)	0.062	163	0.37 (0.2-0.69)	**0.0013**
Alcohol consumption(yes)	115	0.43 (0.25-0.74)	**0.002**	115	0.29 (0.13-0.66)	**0.0017**
(no)	204	0.68 (0.43-1.08)	0.098	197	0.63 (0.31-1.29)	0.2
Sorafenib treatment	30	0.37 (0.15-0.9)	**0.023**	29	0.51 (0.17-1.54)	0.23
Vascular invasion(none)	204	0.84 (0.54-1.3)	0.43	200	1.46 (0.7-3.04)	0.31
(micro)	91	0.61 (0.34-1.08)	0.084	88	4.62(0.6-35.68)	0.11
AJCC_T(1)	180	1.58 (0.95-2.63)	0.073	177	3.82(0.9-16.26)	0.051
(2)	92	0.45 (0.25-0.8)	**0.0049**	89	0.36(0.13-1.04)	**0.049**
(3)	78	0.46 (0.24-0.88)	**0.016**	75	0.4(0.18-0.89)	**0.021**
Grade (1)	55	0.4 (0.18-0.9)	**0.023**	55	0.32(0.09-1.17)	0.072
2	175	0.6 (0.37-0.96)	**0.031**	169	0.48(0.24-0.95)	**0.032**
3	119	0.54 (0.32-0.93)	**0.024**	116	1.31(0.61-2.83)	0.49
Stage 1	170	1.62 (0.96-2.74)	0.071	167	3.18(0.74-13.74)	0.1
1+2	254		0.088	249	0.67(0.34-1.34)	0.26
2	84	0.69 (0.45-1.06)	**0.0083**	82	0.33(0.1-1.08)	0.054
2+3	167		**0.00043**	163	0.38(0.21-0.7)	**0.0013**
3	83	0.44 (0.24-0.82)	**0.0074**	81	0.36(0.17-0.78)	**0.007**
3+4	88	0.49 (0.33-0.74)	**0.0066**	84	0.37(0.18-0.79)	**0.0075**

		0.45（0.24-0.82）			
		0.45（0.25-0.81）			

* The sample numbers with some clinical characteristics were too low for meaningful analysis, such as Black or Africa American, Grade 4, macro invasion, etc. N: sample size.

Table3. The correlation between *SRGN* and infiltrating cells in LIHC by TIMER2.0

TME Cells	Rho	p-value
T cell CD4+	0.368	1.76e-12
T cell CD4+ naïve_CIBERSORT)	-0.152	4.69e-03
T cell CD4+ naïve_XCELL)	0.171	1.44e-03
T cell CD4+ central memory	-0.262	7.79e-07
T cell CD4+ effector memory	0.107	4.8e-02
T cell CD4+ memory activated	0.12	2.64e-02
T cell CD4+ memory resting	0.412	1.55e-15
T cell CD4+ memory	0.305	7.2e-09
T cell CD4+(non-regulatory)	0.175	1.1e-03
T cell CD4+ Th1	-0.234	1.09e05
T cell CD4+ Th2	0.206	1.19e-04
T cell CD8+	**0.552**	**5.79e-29**
T cell CD8+ central memory	0.336	1.49e-10
T cell CD8+ effector memory	0.119	2.65e-02
B cell	0.308	5.19e-09
B cell naïve	0.145	7.06e-03
B cell plasma_CIBERSORT-ABS	0.121	2.4e-02
B cell plasma_XCELL	-0.11	4.12e-02
B cell memory	0.108	4.49e-02
T cell regulatory	0.399	1.34e-14
T cell follicular helper	0.252	2.22e-06
Monocyte	0.481	2.21e-21
Macrophage/monocyte	0.481	2.21e-21
Macrophage_EPIC	-0.382	1.9e-13
Macrophage_TIMER	0.481	2.31e-21
Macrophage M0	-0.123	2.23e-02
Macrophage M1	**0.517**	**5.84e-25**
Macrophage M2_CIBERSORT-ABS	**0.696**	**3.26e-51**
Macrophage M2_TIDE	**-0.516**	**6.66e-25**
Myloid dendritic cell	**0.627**	**4.97e-39**
Myloid dendritic cell activated	0.475	7.62e-21
Myloid dendritic cell resting	0.142	8.13e-03
Common lymphoid progenitor	0.211	8.13e-05
Neutrophil	0.426	1.11e-16

NK cell	-0.115	3.26e-02
NK cell resting	-0.164	2.18e-03
NK cell activated	**0.524**	**1.11e-25**
T cell NK	-0.147	6.11e-03
Mast cell	0.109	4.27e-02
Mast cell resting	0.192	3.23e-04
Mast cell activated	-0.166	2.03e03
Endothelial cell	**0.522**	**1.67e-25**
Cancer associated fibroblast	0.456	4.09e-19

Table4. The survival analysis of *SRGN* expression associated with infiltrating cells in LIHC by TIMER2.0

Cells in TME	**Low SRGN expression Hazard ratio**	**p-value**	**High SRGN expression Hazard ratio**	**p-value**
T cell CD8+	0.658	0.139	0.53	**0.0448**
T cell CD4+ memory resting	0.847	0.584	0.501	**0.0383**
Hematopoietic stem cell	0.834	0.551	0.379	**0.00868**
MDSC	1.41	0.313	2.78	**0.00668**
Endothelial cell	0.565	0.0566	0.364	**0.00638**
Monocyte(MCPcounter)	1.65	0.105	2.75	**0.025**
Macrophage	1.82	**0.0432**	1.58	0.198
Macrophage M1	1.19	0.558	3.09	**0.00633**
Macrophage 0	1.86	**0.0356**	1.59	0.159
Macrophage 2	1.99	**0.0366**	1.25	0.568
T cell CD4+ central memory	0.742	0.305	0.874	0.694
T cell CD4+ Th2	1.22	0.498	1.83	0.0945
Tregs	0.873	0.659	0.671	0.249
Monocyte(XCELL)	1.19	0.599	1.95	0.0811
Myloid dendrtic cell	0.924	0.796	3.28	0.0599
Myloid dendrtic cell resting	1.34	0.283	1.05	0.871

TME: tumor environment