The transcription factor SALL4 regulates stemness of EpCAM-positive hepatocellular carcinoma

The Transcription Factor SALL4 Regulates Stemness of EpCAM-positive Hepatocellular Carcinoma

Sha Sha Zeng[1], Taro Yamashita[1, 2], Mitsumasa Kondo[1], Kouki Nio[1], Takehiro Hayashi[1], Yasumasa Hara[1], Yoshimoto Nomura[1], Mariko Yoshida[1], Tomoyuki Hayashi[1], Naoki Oishi[1], Hiroko Ikeda[3], Masao Honda[1], and Shuichi Kaneko[1]

Departments of [1]Gastroenterology, [2]General Medicine, and [3]Pathology, Kanazawa University Hospital, Kanazawa, Ishikawa, Japan

Corresponding Author: Taro Yamashita M.D., Ph.D.

Department of General Medicine/Gastroenterology, Kanazawa University Hospital,

13-1 Takara-Machi, Kanazawa, Ishikawa 920-8641, Japan

Tel.: +81-76-265-2042; Fax: +81-76-234-4281

E-mail: taroy@m-kanazawa.jp

Word count: 4,966

Number of figures: 4

Number of tables: 1

Abbreviations: CSC, cancer stem cell; HCC, hepatocellular carcinoma; EpCAM, epithelial cell adhesion molecule; AFP, alpha fetoprotein; HpSC-HCC, hepatic stem cell-like HCC; MH-HCC, mature hepatocyte-like HCC; SALL4, Sal-like 4 (Drosophila); qRT-PCR, quantitative reverse transcription-polymerase chain reaction ;

HDAC, histone deacetylase; SAHA, suberoylanilide hydroxamic acid; SBHA, suberic bis-hydroxamic acid; NuRD, nucleosome remodeling and deacetylase.

Conflicts of interest: All authors declare that they have no conflicts of interest.

Financial support: This study was supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology, Japan (23590967), a grant from the Japanese Society of Gastroenterology, a grant from the Ministry of Health, Labour and Welfare, and a grant from the National Cancer Center Research and Development Fund (23-B-5), Japan.

Abstract

Background/Aims: Recent evidence suggests that hepatocellular carcinoma can be classified into certain molecular subtypes with distinct prognoses based on the stem/maturational status of the tumor. We investigated the transcription program deregulated in hepatocellular carcinomas with stem cell features.

Methods: Gene and protein expression profiles were obtained from 238 (analyzed by microarray), 144 (analyzed by immunohistochemistry), and 61 (analyzed by qRT-PCR) hepatocellular carcinoma cases. Activation/suppression of an identified transcription factor was used to evaluate its role in cell lines. The relationship of the transcription factor and prognosis was statistically examined.

Results: The transcription factor SALL4, known to regulate stemness in embryonic and hematopoietic stem cells, was found to be activated in a hepatocellular carcinoma subtype with stem cell features. SALL4-positive hepatocellular carcinoma patients were associated with high values of serum alpha fetoprotein, high frequency of hepatitis B virus infection, and poor prognosis after surgery compared with SALL4-negative patients. Activation of SALL4 enhanced spheroid formation and invasion capacities, key characteristics of cancer stem cells, and up-regulated the hepatic stem cell markers *KRT19*, *EPCAM*, and *CD44* in cell lines. Knockdown of SALL4 resulted in the down-regulation of these stem cell markers, together with attenuation of the invasion capacity. The SALL4 expression status was associated with histone deacetylase activity in cell lines, and the histone deacetylase inhibitor successfully suppressed proliferation of SALL4-positive hepatocellular carcinoma cells.

Conclusions: SALL4 is a valuable biomarker and therapeutic target for the diagnosis and treatment of hepatocellular carcinoma with stem cell features.

Introduction

Cancer is a heterogeneous disease in terms of morphology and clinical behavior. This heterogeneity has traditionally been explained by the clonal evolution of cancer cells and the accumulation of serial stochastic genetic/epigenetic changes [1]. The alteration of the microenvironment by tumor stromal cells is also considered to contribute to the development of the heterogeneous nature of the tumor through the activation of various signaling pathways in cancer cells, including epithelial mesenchymal transition programs [2].

Recent evidence suggests that a subset of tumor cells with stem cell features, known as cancer stem cells (CSCs), are capable of self-renewal and can give rise to relatively differentiated cells, thereby forming heterogeneous tumor cell populations [3]. CSCs were also found to generate tumors more efficiently in immunodeficient mice than non-cancer stem cells in various solid tumors as well as hematological malignancies [4]. CSCs are also more metastatic and chemo/radiation-resistant than non-CSCs and are therefore considered to be a pivotal target for tumor eradication [5, 6].

Hepatocellular carcinoma (HCC) is a leading cause of cancer death worldwide [7]. Recently, we proposed a novel HCC classification system based on the expression status of the hepatic stem/progenitor markers epithelial cell adhesion molecule (EpCAM) and alpha fetoprotein (AFP) [8]. EpCAM-positive ($^{+}$) AFP^{+} HCC (hepatic stem cell-like HCC; HpSC-HCC) is characterized by an onset of disease at younger ages, activation of Wnt/β-catenin signaling, a high frequency of portal vein invasion and poor prognosis after radical resection, compared with $EpCAM^{-}$ AFP^{-} HCC (mature hepatocyte-like HCC; MH-HCC) [9]. *EPCAM* is a target gene of Wnt/β-catenin

signaling, and EpCAM$^+$ HCC cells isolated from primary HCC and cell lines show CSC features including tumorigenicity, invasiveness, and resistance to fluorouracil [9, 10]. Thus, EpCAM appears to be a potentially useful marker for the isolation of liver CSCs in HpSC-HCC. However, key transcriptional programs responsible for the maintenance of EpCAM$^+$ CSCs are still unclear.

In this study, we aimed to clarify the transcriptional programs deregulated in HpSC-HCC using a gene expression profiling approach. We found that the *SALL4* gene encoding Sal-like 4 (*Drosophila*) (SALL4), a zinc finger transcriptional activator and vertebrate orthologue of the *Drosophila* gene spalt (*sal*) [11], was up-regulated in HpSC-HCC. In adults, SALL4 is known to be expressed in hematopoietic stem cells and their malignancies, but its role in HCC has not yet been fully elucidated [12-14]. We therefore investigated the role of SALL4 in the regulation and maintenance of EpCAM$^+$ HCC.

Materials and methods

Clinical HCC specimens

A total of 144 HCC tissues and adjacent non-cancerous liver tissues were obtained from patients who underwent hepatectomy for HCC treatment from 2002 to 2010 at Kanazawa University Hospital, Kanazawa, Japan. These samples were formalin-fixed and paraffin-embedded, and used for immunohistochemistry (IHC). A further 61 HCC samples were obtained from patients who underwent hepatectomy from 2008 to 2011; these were freshly snap-frozen in liquid nitrogen and used for RNA analysis. Of these 61 HCCs, eight and 36 cases were defined as HpSC-HCC and MH-HCC, respectively, according to previously described criteria [8].

Twenty-seven HCC cases were included in both the IHC cohort (n =144) and quantitative reverse transcription-polymerase chain reaction (qRT-PCR) cohort (n = 61), and SALL4 gene and protein expression were compared between these cases. An additional fresh HpSC-HCC sample was obtained from a surgically resected specimen and immediately used for preparation of a single-cell suspension. All experimental and tissue acquisition procedures were approved by the Ethics Committee and the Institutional Review Board of Kanazawa University Hospital. All patients provided written informed consent.

Microarray analysis

Detailed information on microarray analysis is available in the supplemental text.

Cell culture and reagents

Human liver cancer cell lines HuH1, HuH7, HLE, and HLF were obtained from the Japanese Collection of Research Bioresources (JCRB), and Hep3B and SK-Hep-1 were obtained from the American Type Culture Collection (ATCC). Single-cell suspensions of a primary HCC tissue were prepared as described previously [15]. Detailed information is available in the supplemental text. The histone deacetylase (HDAC) inhibitor suberic bis-hydroxamic acid (SBHA) and suberoylanilide hydroxamic acid (SAHA) were obtained from Cayman Chemical (Ann Arbor, MI). Plasmid constructs pCMV6-SALL4 (encoding *SALL4A*), pCMV6-SALL4-GFP, and 29mer shRNA constructs against human SALL4 (No. 7412) were obtained from OriGene Technologies, Inc. (Rockville, MD). These constructs were transfected using Lipofectamine 2000 (Life Technologies, Carlsbad, CA) according to the manufacturer's protocol.

Western blotting

Whole cell lysates were prepared using RIPA lysis buffer. Nuclear and cytoplasmic proteins were extracted using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Pierce Biotechnology Inc., Rockford, IL). Mouse monoclonal antibody to human Sall4 clone 6E3 (Abnova, Walnut, CA), rabbit polyclonal antibodies to human Lamin B1 (Cell Signaling Technology Inc., Danvers, MA), and mouse monoclonal anti-β-actin antibody (Sigma-Aldrich, St. Louis, MO) were used. Immune complexes were visualized by enhanced chemiluminescence (Amersham Biosciences Corp., Piscataway, NJ) as described previously [15, 16].

Quantitative reverse transcription-polymerase chain reaction (qRT-PCR)

Detailed information on qRT-PCR is available in the supplemental text.

IHC and immunofluorescence (IF) analyses

IHC was performed using an Envision+ kit (Dako, Carpinteria, CA) according to the manufacturer's instructions. Anti-SALL4 monoclonal antibody 6E3 (Abnova, Walnut, CA), anti-EpCAM monoclonal antibody VU-1D9 (Oncogene Research Products, San Diego, CA), and anti-CK19 monoclonal antibody RCK108 (Dako Japan, Tokyo, Japan) were used for detecting SALL4, EpCAM, and CK19, respectively. Anti-Sall4 rabbit polyclonal antibodies (ab29112) (Abnova) and vector red (Vector Laboratories Inc., Burlingame, CA) were used for double color IHC analysis. Samples with >5% positive staining in a given area were considered to be positive for a particular antibody. For IF analyses, Alexa 488 fluorescein isothiocyanate (FITC)-conjugated anti-mouse immunoglobulin G (IgG) (Life Technologies) was used as a secondary antibody.

Cell proliferation, spheroid formation, invasion, and HDAC activity assay

Detailed information on this topic is available in the supplemental text.

Statistical analyses

Student *t*-tests were performed with GraphPad Prism software 5.0 (GraphPad Software, San Diego, CA) to compare various test groups assayed by cell proliferation assays and qRT-PCR analysis. Spearman's correlation analysis and Kaplan-Meier survival analysis were also performed with GraphPad Prism software 5.0 (GraphPad Software).

Results

Activation of SALL4 in HpSC-HCC. To elucidate the transcriptional programs deregulated in HpSC-HCC, we performed class-comparison analyses and identified 793 genes showing significant differences in differential expression between HpSC-HCC (n = 60) and MH-HCC (n = 96) ($P < 0.001$), as previously described [9]. Of them, 455 genes were specifically up-regulated in HpSC-HCC, and we performed transcription factor analysis using this gene set to identify their transcriptional regulators by MetaCore software. We identified four transcription factor genes, *SALL4*, *NFYA*, *TP53*, and *SP1*, that were potentially activated in HpSC-HCC (Fig. 1A). Involvement of *TP53* and *SP1* in the stemness of HCC has previously been described [17, 18], but the roles of *SALL4* and *NFYA* were unclear.

We investigated the interaction networks affected by *SALL4* and *NFYA* using the MetaCore dataset. We showed that *SALL4* might be a regulator of Akt signaling (*SP1*), Wnt signaling (*TCF7L2*), and epigenetic modification (*JARID2*, *DMRT1*, *DNMT3B*) [19], and could potentially regulate two other transcriptional regulators, *SP1* and *NFYA*, through Akt and Myb signaling pathways (Fig. 1B). As a recent study indicated that *SALL4* is a direct target of the Wnt signaling pathway [20], which is dominantly activated in HpSC-HCC [9], we focused on the expression of *SALL4* in HpSC-HCC, and confirmed its up-regulation in HpSC-HCC compared with MH-HCC as evaluated by microarray data (Fig. 1C). We validated this using an independent HCC cohort evaluated by qRT-PCR (Fig. 1D). We further examined the expression of *SALL4*, *EPCAM*, and *AFP* using microarray data of 238 HCC cases (Fig. 1E) and qRT-PCR data of 61 HCC cases (Fig. 1F). For the tumor/non-tumor ratios, we identified a weak positive correlation between *SALL4* and *EPCAM* ($r = 0.31$, $P < 0.0001$) and between

SALL4 and *AFP* (r = 0.31, $P = 0.0003$) in the microarray cohort. We further evaluated expression of these genes in HCC tissues by qRT-PCR, and we validated the strong positive correlation between *SALL4* and *EPCAM* (r = 0.70, $P < 0.0001$) and between *SALL4* and AFP (r = 0.66, $P < 0.0001$) in the independent cohort.

Next we performed IHC analysis of 144 HCC cases surgically resected at Kanazawa University Hospital. We first confirmed the nuclear accumulation of SALL4 stained by an anti-human SALL4 antibody (Fig. 2A). We further confirmed the concordance of SALL4 protein expression evaluated by IHC, and *SALL4* gene expression evaluated by qRT-PCR using the same samples (Fig. 2B). We detected the nuclear expression of SALL4 in 43 of 144 HCC cases (Table 1). After evaluating the clinicopathological characteristics of SALL4-positive and -negative HCC cases, we identified that SALL4-positive HCCs were associated with a significantly high frequency of hepatitis B virus (HBV) infection and significantly high serum AFP values. We further identified that SALL4-positive HCCs were associated with expression of the hepatic stem cell markers EpCAM and CK19. Co-expression of SALL4, EpCAM, and CK19 was confirmed by double color IHC analysis (Fig. 2C). Evaluation of the survival outcome of these surgically resected HCC cases by Kaplan-Meier survival analysis indicated that SALL4-positive HCCs were associated with significantly lower recurrence-free survival outcomes within one year compared with SALL4-negative HCCs ($P = 0.0049$) (Fig. 2D).

Because SALL4 expression was positively correlated with EpCAM and AFP expression in primary HCC cases, we evaluated the expression of SALL4 in $EpCAM^+$ AFP^+ and $EpCAM^-$ AFP^- HCC cell lines. Consistent with the primary HCC data, two of three $EpCAM^+$ AFP^+ HCC cell lines (Hep3B and HuH7) abundantly expressed SALL4,

as shown by qRT-PCR (Fig. 2E) and Western blotting (Fig. 2F). We identified the expression of two isoforms of SALL4 proteins with molecular weights of 165 kD (SALL4A) and 115 kD (SALL4B), and SALL4B was found to be the dominant endogenous isoform in HCC cell lines. All $EpCAM^-$ AFP^- HCC cell lines (SK-Hep-1, HLE, and HLF) and one $EpCAM^+$ AFP^+ cell line (HuH1) did not express SALL4. Nuclear accumulation of SALL4 in Hep3B and HuH7 cells was confirmed by IHC using subcutaneous tumors developed in xenotransplanted NOD/SCID mice (Fig. 2G). We further evaluated the expression of *EPCAM* and *SALL4* using single cell suspensions derived from a surgically resected primary HCC. $EpCAM^+$ and $EpCAM^-$ cells were separated by magnetic beads, and we revealed a strong spheroid formation capacity of sorted $EpCAM^+$ cells compared with $EpCAM^-$ cells (Fig. 2H, left panel). Interestingly, when comparing the expression of *SALL4* in these sorted cells, we identified a high expression of *SALL4* in sorted $EpCAM^+$ cells compared with $EpCAM^-$ cells (Fig. 2H, right panel), indicating that SALL4 is activated in $EpCAM^+$ liver CSCs.

SALL4 regulates stemness of HpSC-HCC. To explore the role of SALL4 in HpSC-HCC, we evaluated the effect of its overexpression in HuH1 cells which showed little expression of SALL4 irrespective of $EpCAM^+$ and AFP^+ HpSC-HCC phenotype. We transfected plasmid constructs encoding SALL4 (pCMV6-SALL4) or control (pCMV7), and we similarly identified the expression of two isoforms by using this construct (Fig. 3A). Evaluation of the subcellular localization of GFP-tagged SALL4 (pCMV6-SALL4-GFP) showed that it could be detected in both the cytoplasm and nucleus (Fig. 3B). We observed strong up-regulation of the hepatic stem cell marker *KRT19*, modest up-regulation of *EPCAM* and *CD44*, and down-regulation of the mature hepatocyte marker *ALB* in HuH1 cells transfected with pCMV6-SALL4 compared with

control (Fig. 3C). Up-regulation of CK19 by SALL4 overexpression was also confirmed at the protein level by IF analysis (Fig. 3D). Phenotypically, SALL4 overexpression in HuH1 cells resulted in the significant activation of spheroid formation and invasion capacities with activation of *SNAI1*, which induces epithelial-mesenchymal transition, compared with the control (Fig. 3E & 3F, Supplemental Fig. 1A).

We further investigated the effect of SALL4 knockdown in HuH7 cells which intrinsically expressed high levels of SALL4. Expression of *SALL4* was decreased to 50% in HuH7 cells transfected with SALL4 sh-RNA compared with control when evaluated by qRT-PCR (Fig.4A). However, the reduction of SALL4 protein was more evident when evaluated by Western blotting, suggesting that this sh-RNA construct might work at the translational as well as the transcriptional level (Fig. 4B). Knock down of *SALL4* resulted in a compromised invasion capacity and spheroid formation capacity with decreased expression of *EPCAM* and *CD44* in HuH7 cells (Fig. 4C & 4D, Supplemental Fig. 1B and 1C).

SALL4 and HDAC activity in HpSC-HCC. The above data suggested that SALL4 is a good target and biomarker for the diagnosis and treatment of HpSC-HCCs. However, it is difficult to directly target SALL4 as no studies have investigated the inhibition of its transcription using chemical or other approaches [21]. We therefore re-investigated the interaction networks associated with *SALL4*, and found that SALL4 activation appeared to induce epigenetic modification (Fig. 1B). In particular, a recent study suggested that SALL4 forms a nucleosome remodeling and deacetylase (NuRD) complex with HDACs and potentially regulates HDAC activity [22]. We therefore confirmed that *SALL4* knock down resulted in the reduced activity of total HDAC in HuH7 cells (Fig. 4E). We also evaluated the effect of the overexpression of SALL4 in

HuH1 and HLE cells, which do not express SALL4 endogenously, and SALL4 overexpression was found to result in a modest increase of HDAC activity and mild enhancement of chemosensitivity to an HDAC inhibitor SBHA in both cell lines (Supplemental Fig. 2A and 2B). We further investigated HDAC activity in two SALL4-positive (Hep3B, HuH7) and two SALL4-negative (HLE, HLF) HCC cell lines. Interestingly, high HDAC activities were detected in SALL4-positive compared with SALL4-negative HCC cell lines (Fig. 4F). The HDAC inhibitor SBHA was found to inhibit proliferation of SALL4-positive HCC cell lines at 10 μM concentration. By contrast, SBHA had little effect on the proliferation of SALL4-negative HCC cell lines at this concentration (Fig. 4G). SBHA treatment suppressed the expression of SALL4 gene/protein expression in SALL4-positive HuH7 and Hep3B cell lines (Supplemental Fig. 3A and 3B). We further investigated the effect of SAHA, an additional HDAC inhibitor, in these HCC cell lines, and SAHA was found to more efficiently suppress the cell proliferation of SALL4-positive cell lines compared with SALL4-negative cell lines (Supplemental Fig. 3C).

Taken together, our data suggest a pivotal role for the transcription factor SALL4 in regulating the stemness of HpSC-HCC. SALL4 was detected in HpSC-HCCs with poor prognosis, and inactivation of SALL4 resulted in a reduced invasion/spheroid formation capacity and decreased expression of hepatic stem cell markers. The HDAC inhibitors inhibited proliferation of SALL4-positive HCC cell lines with a reduction of SALL4 gene/protein expression, suggesting their potential in the treatment of SALL4-positive HpSC-HCC.

Discussion

Stemness traits in cancer cells are currently of great interest because they may explain the clinical outcome of patients according to the malignant nature of their tumor. Recently, we proposed an HCC classification system based on the stem/maturation status of the tumor by EpCAM and AFP expression status [8]. These HCC subtypes showed distinct gene expression patterns with features resembling particular stages of liver lineages. Among them, HpSC-HCC was characterized by a highly invasive nature, chemoresistance to fluorouracil, and poor prognosis after radical resection, warranting the development of a novel therapeutic approach against this HCC subtype [9].

In this study, we showed that the transcription factor SALL4 was activated in HpSC-HCC and that SALL4 might regulate HCC stemness, as characterized by the activation of EpCAM, CK19, and CD44 with highly tumorigenic and invasive natures. Furthermore, we identified that SALL4-positive HCC cell lines tended to show high HDAC activity and chemosensitivity to the HDAC inhibitors SBHA and SAHA. This study reveals for the first time the utility of SBHA for the treatment of HCC with stem cell features.

SALL4 is a zinc finger transcription factor originally cloned based on sequence homology to *Drosophila sal* [11]. *SALL4* mutations are associated with Okihiro syndrome, a human disease involving multiple organ defects [23, 24]. SALL4 plays a fundamental role in the maintenance of embryonic stem cells, potentially through interaction with Oct4, Sox2, and Nanog [25-30]. Furthermore, knockdown of SALL4 significantly reduces the efficiency of induced pluripotent stem cell generation [31]. *SALL4* is also expressed in hematopoietic stem cells and leukemia cells, where it regulates their maintenance [14, 32]. SALL4 is known to encode two isoforms

(SALL4A and SALL4B), and a recent study suggested the important role of SALL4B on maintaining the stemness of embryonic stem cells [25]. Interestingly, our data indicated that SALL4B is also a dominant form in HpSC-HCC cell lines. It is unclear how SALL4 isoform expression is regulated in cancer, and future studies are required to explore the mechanisms of SALL4 isoform regulation.

In the liver, SALL4 is expressed in fetal hepatic stem/progenitors but not in adult hepatocytes, and a mouse study demonstrated that inhibition of SALL4 in hepatic stem/progenitors contributes to their differentiation [33]. Interestingly, recent studies indicated that AFP-producing gastric cancer expresses SALL4, suggesting that SALL4 might play a role in the hepatoid differentiation of gastric cancer [34]. Consistently, our data indicated a positive correlation between *SALL4*, *AFP*, and *EPCAM* expression in two independent HCC cohorts. Strikingly, SALL4 was recently shown to be expressed in a subset of human liver cancers with poor prognoses, while modification of SALL4 expression resulted in the alteration of cell proliferation *in vitro* and tumor growth *in vivo*, consistent with our current study [35]. A recent study reported the expression of SALL4 in 46% of HCC cases, which is almost comparable to our present study [36]. Furthermore, a very recent study of two independent large cohorts demonstrated that SALL4 is a marker for a progenitor subclass of HCC with an aggressive phenotype [37]. It is still unclear how SALL4 expression is regulated and which target genes are directly activated by SALL4 binding. Future studies using next generation sequencing are required to fully understand the mechanisms of SALL4 regulation of HCC stemness.

In this study, we demonstrated that SALL4-positive HCC cell lines have high HDAC activity and chemosensitivity against the HDAC inhibitors SBHA and SAHA compared with SALL4-negative HCC cell lines. SALL4 was recently found to directly

connect with the epigenetic modulator NuRD complex [22], thereby possible affecting the histone modification associated with stemness. The NuRD complex is a multiunit chromatin remodeling complex containing chromodomain-helicase-DNA-binding proteins and HDACs that regulate histone deacetylation [38]. Its role in cancer is still controversial, while its function in HCC has not yet been determined.

Our data suggest that SALL4 plays a role in controlling HDAC activity and contributing to the maintenance of HCC with stem cell features. Consistently, HDAC inhibitors might be useful for the eradication of SALL4-positive HCC cells through their inhibitory effects on histone deacetylation by NuRD [39]. Encouragingly, a recent study demonstrated the utility of a SALL4-binding peptide to inhibit its binding to phosphatase and tensin homolog deleted on chromosome 10 (PTEN) through interaction with HDAC, thereby targeting leukemia cells [21]. Further studies are required to understand the relationship between SALL4, the NuRD complex, and the maintenance of stemness in HCC.

Acknowledgments

We thank Ms. Masayo Baba and Ms. Nami Nishiyama for excellent technical assistance.

References

[1] Nowell PC. The clonal evolution of tumor cell populations. Science 1976;194:23-28.

[2] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144:646-674.

[3] Jordan CT, Guzman ML, Noble M. Cancer stem cells. The New England journal of medicine 2006;355:1253-1261.

[4] Clarke MF, Dick JE, Dirks PB, Eaves CJ, Jamieson CH, Jones DL, et al. Cancer stem cells--perspectives on current status and future directions: AACR Workshop on cancer stem cells. Cancer Res 2006;66:9339-9344.

[5] Dean M, Fojo T, Bates S. Tumour stem cells and drug resistance. Nat Rev Cancer 2005;5:275-284.

[6] Visvader JE, Lindeman GJ. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. Nat Rev Cancer 2008;8:755-768.

[7] Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin 2011;61:69-90.

[8] Yamashita T, Forgues M, Wang W, Kim JW, Ye Q, Jia H, et al. EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. Cancer Res 2008;68:1451-1461.

[9] Yamashita T, Ji J, Budhu A, Forgues M, Yang W, Wang HY, et al. EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features. Gastroenterology 2009;136:1012-1024.

[10] Yamashita T, Budhu A, Forgues M, Wang XW. Activation of hepatic stem cell marker EpCAM by Wnt-beta-catenin signaling in hepatocellular carcinoma. Cancer Res 2007;67:10831-10839.

[11] de Celis JF, Barrio R. Regulation and function of Spalt proteins during animal development. The International journal of developmental biology 2009;53:1385-1398.

[12] Aguila JR, Liao W, Yang J, Avila C, Hagag N, Senzel L, et al. SALL4 is a robust stimulator for the expansion of hematopoietic stem cells. Blood 2011;118:576-585.

[13] Yang J, Chai L, Gao C, Fowles TC, Alipio Z, Dang H, et al. SALL4 is a key regulator of survival and apoptosis in human leukemic cells. Blood 2008;112:805-813.

[14] Yang J, Chai L, Liu F, Fink LM, Lin P, Silberstein LE, et al. Bmi-1 is a target gene for SALL4 in hematopoietic and leukemic cells. Proceedings of the National Academy of Sciences of the United States of America 2007;104:10494-10499.
[15] Yamashita T, Honda M, Nio K, Nakamoto Y, Takamura H, Tani T, et al. Oncostatin m renders epithelial cell adhesion molecule-positive liver cancer stem cells sensitive to 5-Fluorouracil by inducing hepatocytic differentiation. Cancer Res 2010;70:4687-4697.
[16] Yamashita T, Honda M, Takatori H, Nishino R, Minato H, Takamura H, et al. Activation of lipogenic pathway correlates with cell proliferation and poor prognosis in hepatocellular carcinoma. J Hepatol 2009;50:100-110.
[17] Woo HG, Wang XW, Budhu A, Kim YH, Kwon SM, Tang ZY, et al. Association of TP53 mutations with stem cell-like gene expression and survival of patients with hepatocellular carcinoma. Gastroenterology 2011;140:1063-1070.
[18] Ji J, Wang XW. Clinical implications of cancer stem cell biology in hepatocellular carcinoma. Seminars in oncology 2012;39:461-472.
[19] Yang J, Corsello TR, Ma Y. Stem cell gene SALL4 suppresses transcription through recruitment of DNA methyltransferases. The Journal of biological chemistry 2012;287:1996-2005.
[20] Bohm J, Sustmann C, Wilhelm C, Kohlhase J. SALL4 is directly activated by TCF/LEF in the canonical Wnt signaling pathway. Biochemical and biophysical research communications 2006;348:898-907.
[21] Gao C, Dimitrov T, Yong KJ, Tatetsu H, Jeong HW, Luo HR, et al. Targeting transcription factor SALL4 in acute myeloid leukemia by interrupting its interaction with an epigenetic complex. Blood 2013;121:1413-1421.
[22] Lu J, Jeong HW, Kong N, Yang Y, Carroll J, Luo HR, et al. Stem cell factor SALL4 represses the transcriptions of PTEN and SALL1 through an epigenetic repressor complex. PLoS One 2009;4:e5577.
[23] Al-Baradie R, Yamada K, St Hilaire C, Chan WM, Andrews C, McIntosh N, et al. Duane radial ray syndrome (Okihiro syndrome) maps to 20q13 and results from mutations in SALL4, a new member of the SAL family. American journal of human genetics 2002;71:1195-1199.
[24] Kohlhase J, Heinrich M, Schubert L, Liebers M, Kispert A, Laccone F, et al. Okihiro syndrome is caused by SALL4 mutations. Human molecular genetics 2002;11:2979-2987.

[25] Rao S, Zhen S, Roumiantsev S, McDonald LT, Yuan GC, Orkin SH. Differential roles of Sall4 isoforms in embryonic stem cell pluripotency. Molecular and cellular biology 2010;30:5364-5380.
[26] Tanimura N, Saito M, Ebisuya M, Nishida E, Ishikawa F. Stemness-related Factor Sall4 Interacts with Transcription Factors Oct-3/4 and Sox2 and Occupies Oct-Sox Elements in Mouse Embryonic Stem Cells. The Journal of biological chemistry 2013;288:5027-5038.
[27] Wu Q, Chen X, Zhang J, Loh YH, Low TY, Zhang W, et al. Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. The Journal of biological chemistry 2006;281:24090-24094.
[28] Yang J, Chai L, Fowles TC, Alipio Z, Xu D, Fink LM, et al. Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells. Proceedings of the National Academy of Sciences of the United States of America 2008;105:19756-19761.
[29] Yang J, Gao C, Chai L, Ma Y. A novel SALL4/OCT4 transcriptional feedback network for pluripotency of embryonic stem cells. PLoS One 2010;5:e10766.
[30] Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY, Soh BS, et al. Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nature cell biology 2006;8:1114-1123.
[31] Tsubooka N, Ichisaka T, Okita K, Takahashi K, Nakagawa M, Yamanaka S. Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. Genes to cells : devoted to molecular & cellular mechanisms 2009;14:683-694.
[32] Yang J, Liao W, Ma Y. Role of SALL4 in hematopoiesis. Current opinion in hematology 2012;19:287-291.
[33] Oikawa T, Kamiya A, Kakinuma S, Zeniya M, Nishinakamura R, Tajiri H, et al. Sall4 regulates cell fate decision in fetal hepatic stem/progenitor cells. Gastroenterology 2009;136:1000-1011.
[34] Ikeda H, Sato Y, Yoneda N, Harada K, Sasaki M, Kitamura S, et al. alpha-Fetoprotein-producing gastric carcinoma and combined hepatocellular and cholangiocarcinoma show similar morphology but different histogenesis with respect to SALL4 expression. Human pathology 2012;43:1955-1963.
[35] Oikawa T, Kamiya A, Zeniya M, Chikada H, Hyuck AD, Yamazaki Y, et al. SALL4, a stem cell biomarker in liver cancers. Hepatology 2012.

[36] Gonzalez-Roibon N, Katz B, Chaux A, Sharma R, Munari E, Faraj SF, et al. Immunohistochemical expression of SALL4 in hepatocellular carcinoma, a potential pitfall in the differential diagnosis of yolk sac tumors. Human pathology 2013.
[37] Yong KJ, Gao C, Lim JS, Yan B, Yang H, Dimitrov T, et al. Oncofetal gene SALL4 in aggressive hepatocellular carcinoma. The New England journal of medicine 2013;368:2266-2276.
[38] Lai AY, Wade PA. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. Nat Rev Cancer 2011;11:588-596.
[39] Marquardt JU, Thorgeirsson SS. Sall4 in "stemness"-driven hepatocarcinogenesis. The New England journal of medicine 2013;368:2316-2318.

Table 1. Clinicopathological characteristics of SALL4-positive and -negative HCC cases used for IHC analyses

Parameters	SALL4-positive (n = 43)	SALL4-negative (n = 101)	*P*-value[*]
Age (years, mean ± SE)	60.8 ± 1.8	64.6 ± 1.0	0.13
Sex (Male/Female)	27/16	70/18	0.06
Etiology (HBV/ HCV/ B+C/Other)	21/14/0/8	20/63/3/15	0.0014
Liver cirrhosis (Yes/No)	21/22	61/40	0.27
AFP (ng/ml, mean ± SE)	13,701 ± 9,292	175.5 ± 55.0	<0.0001
Histological grade[**]			
I-II	3	18	
II-III	33	68	
III-IV	7	15	0.24
Tumor size (<3cm/>3cm)	17/26	57/44	0.071
EpCAM (Positive/Negative)	27/16	29/72	0.0002
CK19 (Positive/Negative)	12/31	12/89	0.027

[*]Mann-Whitney U-test or $\chi 2$ test [**]Edmondson-Steiner

Figure legends

Figure 1 Transcription factors potentially activated in HpSC-HCC. **A.** Transcription Factor Analysis. Transcription factors regulating genes up-regulated in HpSC-HCC are listed with their *P*-values and Z-scores as calculated by MetaCore software. **B.** Interaction Network Analysis. Seven genes (*ABL1*, *DMRT1*, *DNMT3B*, *JARID2*, *NFYA*, *SP1*, and *TCF7L2*, indicated in orange) shown to be up-regulated in HpSC-HCC were identified as potential target genes regulated by *SALL4* (indicated in red). **C.** SALL4 gene expression evaluated by microarray analysis. Tumor/non-tumor (T/N) ratios of microarray data in HpSC-HCC (n = 60, indicated in red) and MH-HCC (n = 96, indicated in green). **D.** SALL4 gene expression evaluated by qRT-PCR. Gene expression of *SALL4* in HpSC-HCC (n = 8, indicated in red) and MH-HCC (indicated in green) samples. **E.** Scatter plot analysis. Gene expression levels of *EPCAM* (upper panel) and *AFP* (lower panel) were positively correlated with those of *SALL4* in microarray data (n = 238, T/N ratios), as shown by Spearman's correlation coefficients. **F.** Scatter plot analysis. Gene expression levels of *EPCAM* (upper panel) and *AFP* (lower panel) were positively correlated with those of *SALL4* in qRT-PCR data (n = 61), as shown by Spearman's correlation coefficients.

Figure 2 SALL4 expression in human primary HCCs and cell lines. **A.** Representative images of SALL4-positive (left) and -negative (right) HCC immunostaining (scale bar, 100 μm). **B.** Gene expression of *SALL4* in SALL4-positive (n = 13) and -negative HCCs (n = 14) as shown by IHC (mean ± SD). **C.** Double color IHC analysis of HCC stained with anti-SALL4 and anti-EpCAM (left panel) or anti-CK19 (right panel) antibodies (scale bar, 100 μm). **D.** Kaplan-Meier survival analysis with log-rank. Recurrence-free survival of SALL4-positive (n = 43) and -negative (n = 101) HCCs was analyzed. **E.**

SALL4 expression in $EpCAM^+$ (Hep3B, HuH7, and HuH1) and $EpCAM^-$ (SK-Hep-1, HLE, and HLF) HCC cell lines evaluated by qRT-PCR. **F.** SALL4 expression in $EpCAM^+$ and $EpCAM^-$ HCC cell lines evaluated by Western blotting. **G.** IHC analysis of SALL4 expression in subcutaneous tumors obtained from $EpCAM^+$ (HuH7 and Hep3B) HCC cell lines xenografted in NOD/SCID mice. **H.** Spheroid formation capacity of sorted $EpCAM^+$ and $EpCAM^-$ cells obtained from a primary HCC (left panel). Number of spheroids obtained from 2,000 sorted cells is indicated (n = 3, mean ± SD). Gene expression of *SALL4* in sorted $EpCAM^+$ and $EpCAM^-$ cells obtained from a primary HCC (n = 3, mean ± SD) (right panel).

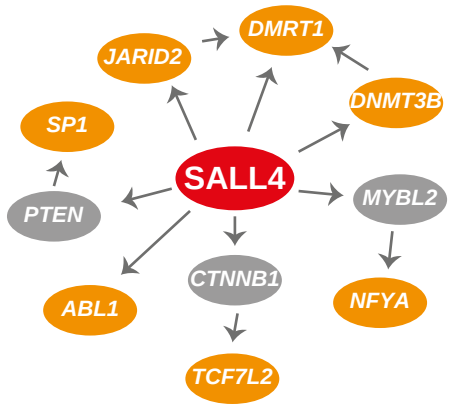
Figure 3 Effect of SALL4 overexpression. **A.** Western blots of cell lysates with anti-SALL4 antibodies. HuH1 cells were transfected with pCMV7 or pCMV6-SALL4 and incubated for 72 h. **B.** IF analysis of HuH1 cells transfected with pCMV7 or pCMV6-SALL4 and incubated for 72 h. **C.** qRT-PCR analysis of *KRT19*, *EPCAM*, *CD44*, and *ALB* in HuH1 cells transfected with pCMV7 or pCMV6-SALL4 and incubated for 48 h. **D.** IF analysis of HuH1 cells transfected with pCMV7 or pCMV6-SALL4, incubated for 72 h and stained with anti-CK19 antibodies, evaluated by the confocal laser scanning microscopy. **E.** Spheroid formation assay of HuH1 cells transfected with pCMV7 or pCMV6-SALL4. Number of spheroids obtained from 2,000 cells is indicated (n = 3, mean ± SD). **F.** Invasion assay of HuH1 cells transfected with pCMV7 or pCMV6-SALL4 (n = 3, mean ± SD).

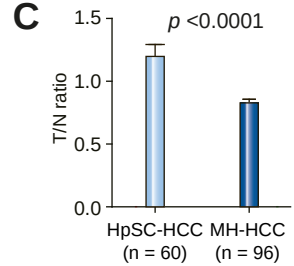
Figure 4 Effect of SALL4 knockdown and HDAC activity. **A.** qRT-PCR analysis of *SALL4* in HuH7 cells transfected with control or SALL4 sh-RNAs (n = 3, mean ± SD). **B.** Western blots of lysates obtained from HuH7 cells transfected with control or SALL4 sh-RNAs with anti-SALL4 antibodies. **C.** qRT-PCR analysis of *EPCAM* and

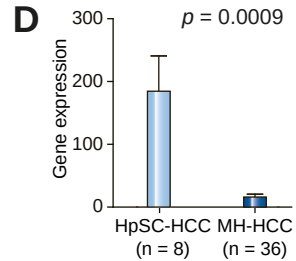
CD44 in HuH7 cells transfected with control or SALL4 sh-RNAs (n = 3, mean ± SD). **D.** Invasion assay of HuH7 cells transfected with control or SALL4 sh-RNAs (n = 3, mean ± SD). **E.** HDAC activity of nuclear extracts obtained from HuH7 cells transfected with control or SALL4 sh-RNAs. **F.** HDAC activity of nuclear extracts obtained from each cell line. HDAC activity was measured in duplicate and average amounts of deacetylated products are indicated (upper panel). Lamin B included in the nuclear extracts loaded for HDAC activity assays was measured by Western blotting (lower panel). **G.** Cell proliferation assay of HCC cell lines. Each cell line was treated with control DMSO (white bar) or 10 μM SBHA (black bar) and cultured for 72 h (n = 4, mean ± SD).

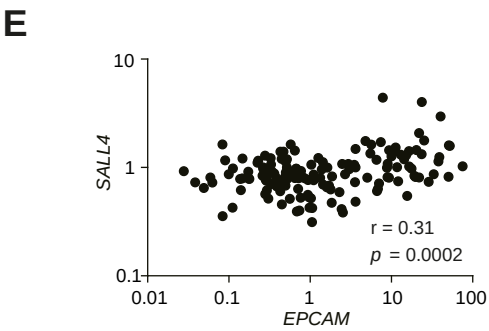
A

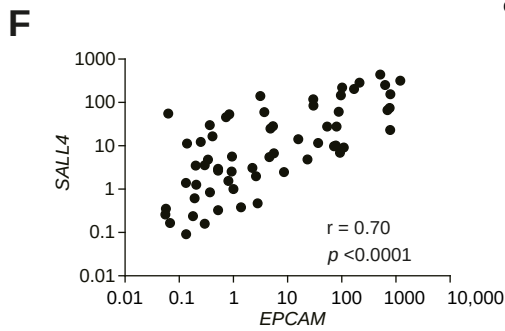
Transcription factors	*p* value	z-score
SALL4	0.0005708	3.958
NFYA	0.002483	3.267
TP53	3.923E-05	4.25
SP1	2.917E-11	6.942



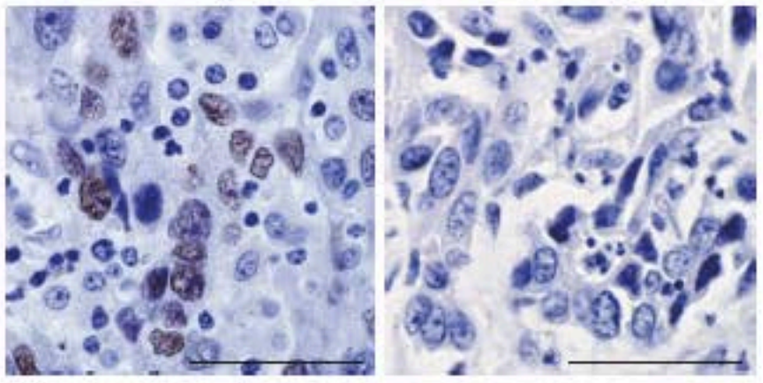




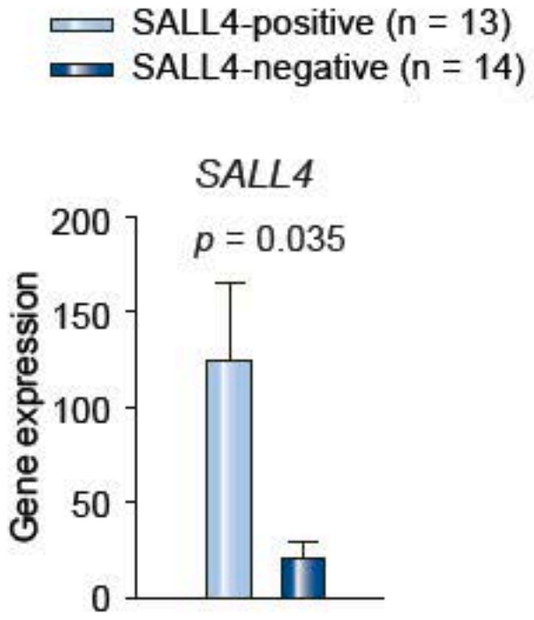




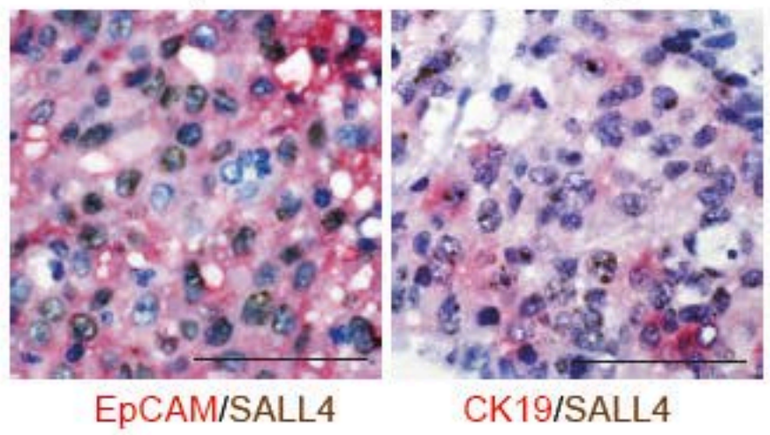
A


SALL4-positive
SALL4-negative

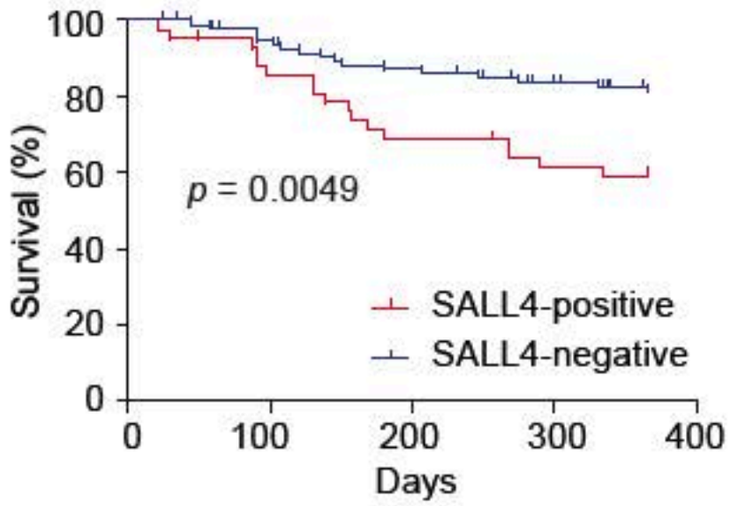
B


SALL4-positive (n = 13)
SALL4-negative (n = 14)
SALL4
p = 0.035
Gene expression
200
150
100
50
0

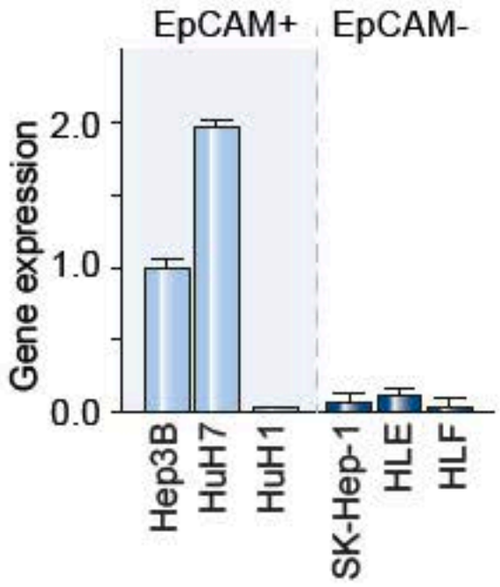
C


SALL4-positive
SALL4-negative
EpCAM/SALL4
CK19/SALL4

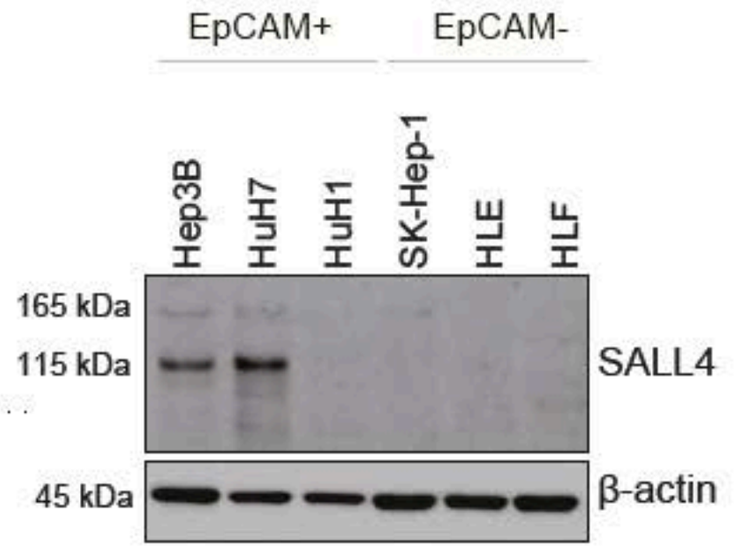
D


Survival (%)
100
80
60
40
20
0
p = 0.0049
SALL4-positive
SALL4-negative
0
100
200
300
400
Days

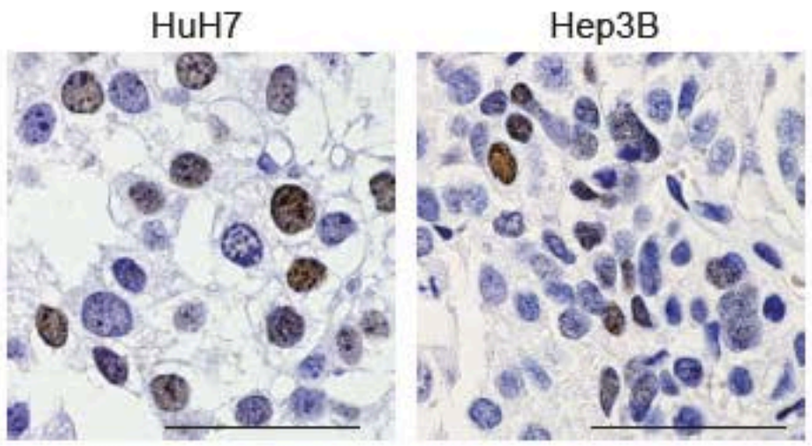
E


EpCAM+
EpCAM-
Gene expression
2.0
1.0
0.0
Hep3B
HuH7
HuH1
SK-Hep-1
HLE
HLF

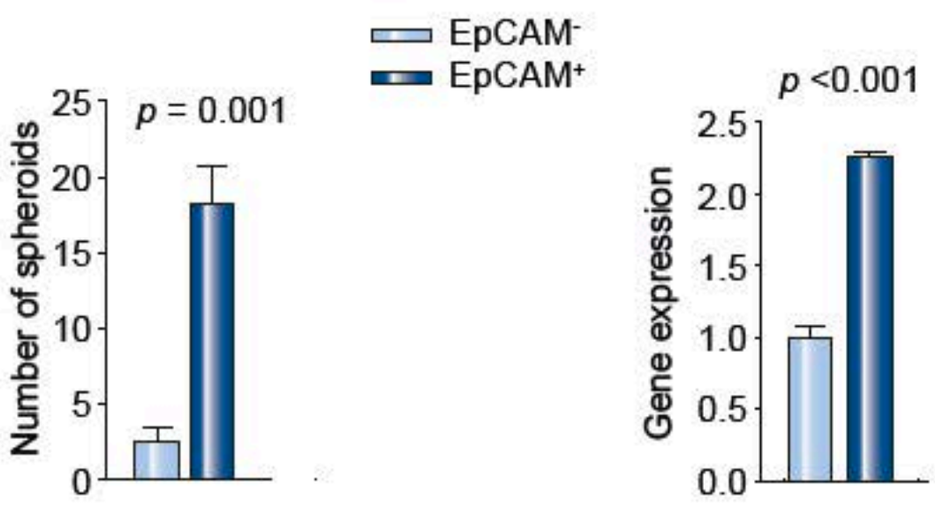
F

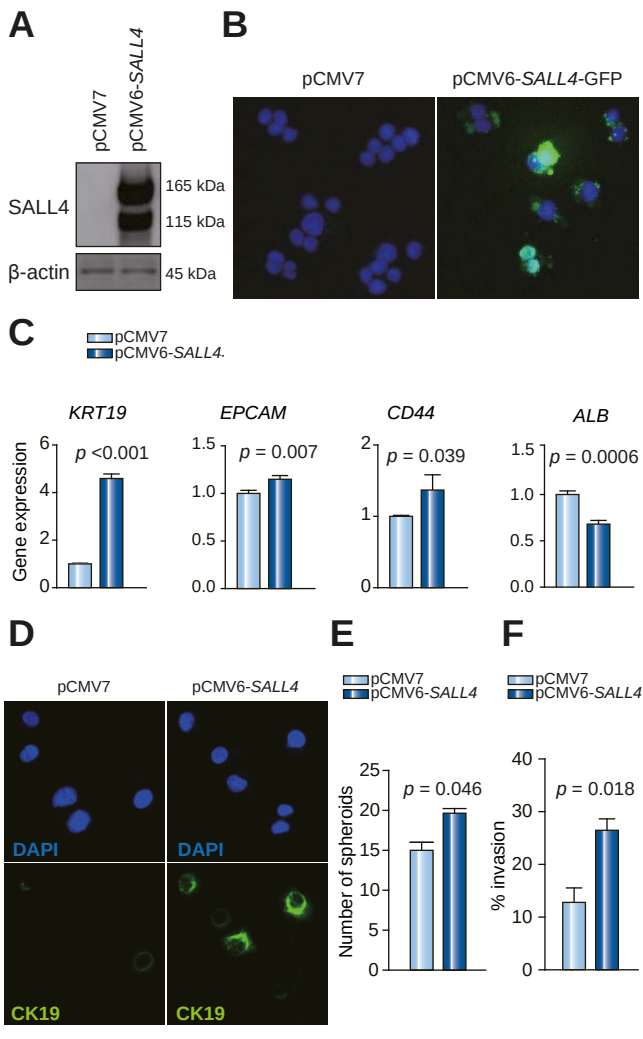

EpCAM+
EpCAM-
Hep3B
HuH7
HuH1
SK-Hep-1
HLE
HLF
165 kDa
115 kDa
SALL4
45 kDa
β-actin

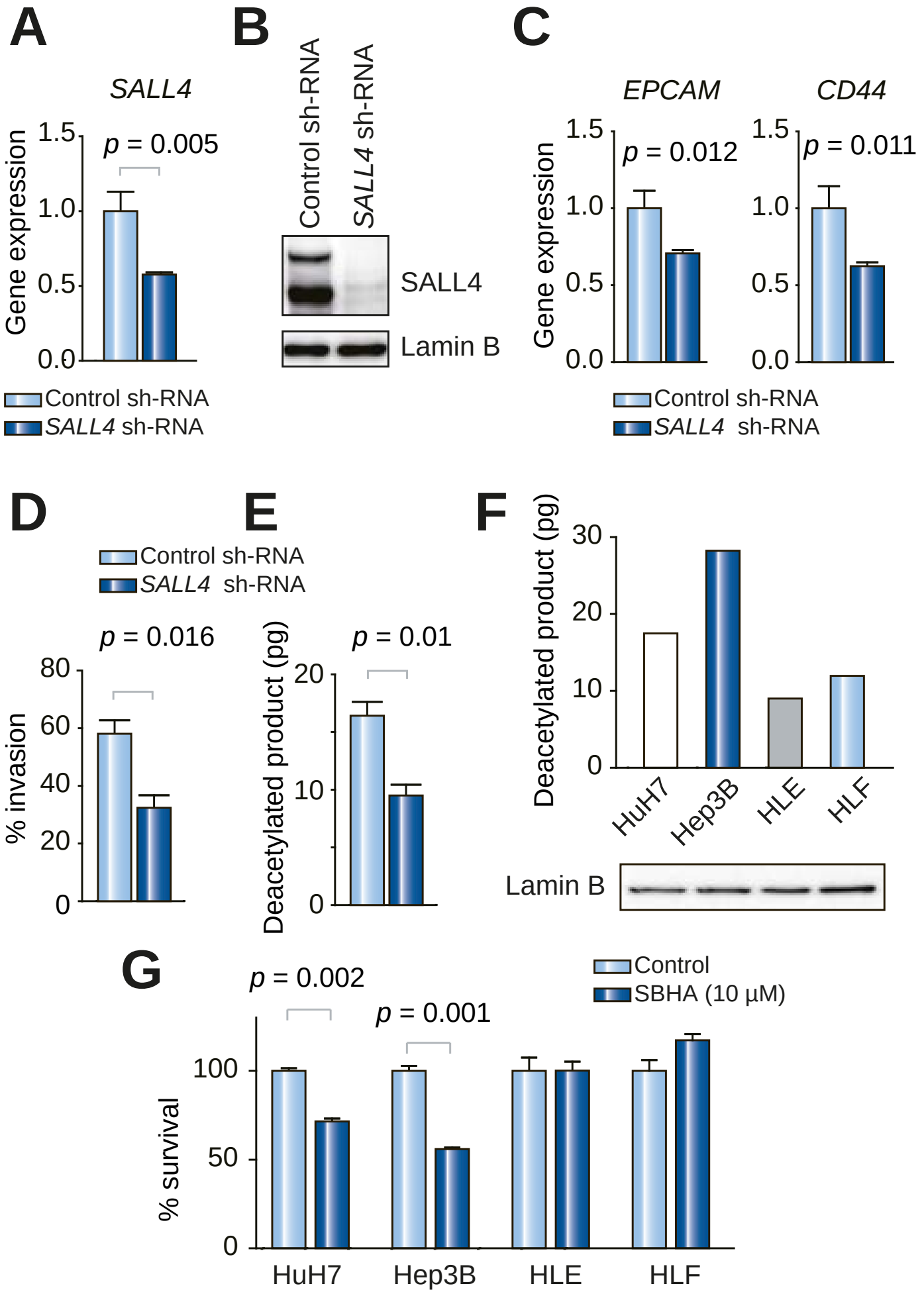
G


HuH7
Hep3B

H


EpCAM-
EpCAM+
Number of spheroids
25
20
15
10
5
0
p = 0.001
p <0.001
Gene expression
2.5
2.0
1.5
1.0
0.5
0.0


A
pCMV7
pCMV6-SALL4
SALL4
165 kDa
115 kDa
β-actin
45 kDa
B
pCMV7
pCMV6-SALL4-GFP
C
pCMV7
pCMV6-SALL4.
KRT19
EPCAM
CD44
ALB
Gene expression
p <0.001
p = 0.007
p = 0.039
p = 0.0006
D
pCMV7
pCMV6-SALL4
DAPI
DAPI
CK19
CK19
E
pCMV7
pCMV6-SALL4
Number of spheroids
p = 0.046
F
pCMV7
pCMV6-SALL4
% invasion
p = 0.018


A
SALL4
Gene expression
1.5
1.0
0.5
0.0
p = 0.005
Control sh-RNA
SALL4 sh-RNA
B
Control sh-RNA
SALL4 sh-RNA
SALL4
Lamin B
C
EPCAM
CD44
Gene expression
p = 0.012
p = 0.011
Control sh-RNA
SALL4 sh-RNA
D
Control sh-RNA
SALL4 sh-RNA
p = 0.016
% invasion
80
60
40
20
0
E
p = 0.01
Deacetylated product (pg)
20
10
0
F
Deacetylated product (pg)
30
20
10
0
HuH7
Hep3B
HLE
HLF
Lamin B
G
p = 0.002
p = 0.001
Control
SBHA (10 μM)
% survival
100
50
0
HuH7
Hep3B
HLE
HLF

Parameters	SALL4-positive (n = 43)	SALL4-negative (n = 101)	*p* value*
Age (yr, mean ± SE)	60.8 ± 1.8	64.6 ± 1.0	0.13
Sex (male/female)	27/16	70/18	0.06
Etiology (HBV/HCV/B + C/other)	21/14/0/8	20/63/3/15	0.0014
Liver cirrhosis (yes/no)	21/22	61/40	0.27
AFP (ng/ml, mean ± SE)	13,701 ± 9292	175.5 ± 55.0	<0.0001
Histological grade**			
I-II	3	18	
II-III	33	68	
III-IV	7	15	0.24
Tumor size (<3 cm/>3 cm)	17/26	57/44	0.071
EpCAM (positive/negative)	27/16	29/72	0.0002
CK19 (positive/negative)	12/31	12/89	0.027