**SUPPLEMENTAL INFORMATION**

# Glucose deprivation induced aberrant FUT1-mediated fucosylation drives cancer stemness in hepatocellular carcinoma

Jane Ho Chun Loong, Tin-Lok Wong, Man Tong, Rakesh Sharma, Lei Zhou, Kai-Yu Ng, Hua-Jian Yu, Chi-Han Li, Kwan Man, Chung-Mau Lo, Xin-Yuan Guan, Terence K Lee, Jing-Ping Yun, Stephanie Ma

## SUPPLEMENTAL MATERIALS AND METHODS

**Reagents.** Chloroquine (CQ), 2-deoxy-D-galactose (2DGal), the PERK inhibitor GSK2656157 (PERKi), and MG-132 were purchased from Sigma-Aldrich. Sorafenib was purchased from LC Laboratories. LY294002 (AKT inhibitor) was purchased from Cell Signaling Technology.

**Clinical samples.** Formalin-fixed paraffin-embedded primary human HCC and adjacent nontumor liver tissue samples were obtained from HCC patients undergoing hepatectomy at the Sun Yat-Sen University Cancer Centre in Guangzhou, China. Samples were collected from patients who had not received any previous local or systemic treatment prior to operation.

**Limiting dilution spheroid formation assay.** Cells at limited dilutions were cultured in 100 µL of serum-free DMEM/F12 medium (Invitrogen) supplemented with 20 ng/ml human recombinant epidermal growth factor (Sigma-Aldrich), 10 ng/ml human recombinant basic fibroblast growth factor (Sigma-Aldrich), 4 mg/ml insulin (Sigma-Aldrich), B27 (1:50; Invitrogen), 500 U/ml penicillin, 500 mg/ml streptomycin (Invitrogen) and 0.25% methylcellulose (Sigma-Aldrich). Cells were cultured in suspension in poly-HEMA-coated 96-well plates. Cells were replenished with 30 µl of supplemented medium every second day. Wells were monitored for sphere formation at 7-10 days. Tumor-initiating cell frequency was calculated using extreme limiting dilution analysis.

**TCGA (The Cancer Genome Atlas) and Gene Expression Omnibus (GEO) data.** Gene expression profiling studies involving normal liver or nontumor liver and HCC tissue samples were analyzed for the expression of *ATF4* and/or *FUT1* transcripts available under Liver Hepatocellular Carcinoma (LIHC) of the TCGA Research Network and GEO (GSE14520 and GSE109211) of the National Centre for Biotechnology Information (NCBI). HCC tissue samples with *FUT1* transcripts available under TCGA LIHC were segregated into two groups: top and bottom 50%, using the mean of all *FUT1* expression levels as a cutoff point. Gene set enrichment analysis (GSEA) was then used for pathway enrichment prediction.

**RNA extraction, cDNA synthesis and quantitative real-time PCR (qPCR).** Total RNA was extracted using RNA-IsoPlus (TaKaRa), and cDNA was synthesized by PrimeScript RT Master Mix (TaKaRa). qPCR was performed with EvaGreen qPCR Master Mix (ABM) and the primers listed in **Supplemental Table 6** on a LightCycler 480 II analyzer (Roche), and the data were analyzed using LightCycler 480 II software (Roche). Relative expression differences were calculated using the $2^{-\Delta\Delta Ct}$ method.

**Protein extraction, Western blot, UEA-1 chromatography and coimmunoprecipitation.** Protein lysates were quantified and resolved on an SDS-PAGE gel, transferred onto a PVDF membrane (Millipore) and immunoblotted with primary and secondary antibodies. Antibody signals were detected using an enhanced chemiluminescence system (GE Healthcare). The following antibodies were used: GRP78 (1:1000, Abcam, ab21685), total eIF2α (1:1000, Cell Signaling Technology, 9722), p-eIF2α (1:1000, Cell Signaling Technology, 9721), total PERK (1:1000, Cell Signaling Technology, 5683), p-PERK (Thr980) (1:1000, Invitrogen, MA5-15033), ATF4 (1:1000, Santa Cruz Biotechnology, sc-200), FUT1 (1:1000, Abcam, ab198712), p-AKT (Ser473) (1:1000, Cell Signaling Technology, 9271), total AKT (1:1000, Cell Signaling Technology, 9272), p-mTOR (Ser2448) (1:1000, Cell Signaling

Technology, 2971), mTOR (1:1000, Cell Signaling Technology, 2983), p-4EBP1 (Thr7/46) (1:1000, Cell Signaling Technology, 2855S), total 4EBP1 (1:1000, Cell Signaling Technology, 9644S), LAMP2 (1:5000, Abcam, ab25631), CD147 (1:1000, Abcam, ab11572), lectin UEA-1-biotin conjugate (1:1000, Vector Laboratories, B1065), ICAM-1 (1:1000, Abcam, ab109361), EGFR (1:1000, Cell Signaling Technology, 2232S), EPHA2 (1:1000, Cell Signaling Technology, 6997S), HDGF (1:1000, Abcam, ab128921), PEBP1 (1:1000, Invitrogen, 37-2100), PARP1 (1:1000, Cell Signaling Technology, 9542S), HMGB1 (1:1000, Abcam, ab79823), CALR (1:1000, Abcam, ab92516), HMGA2 (1:1000, Abcam, ab246513), SRSF3 (1:1000, Abcam, ab73891) and β-actin (1:5000, Sigma-Aldrich, A5316). For UEA-1 chromatography, cells were lysed with ice-cold NETN buffer. A total of 500 μg of lysate was mixed with 50 μl of agarose-bound UEA-1 (AL-1063, Vector Laboratories), brought to 500 μl with PBS containing 1 mM $MnCl_2$, 1 mM $MgCl_2$ and 1 mM $CaCl_2$ and incubated with rotation at 4°C overnight. Beads were washed and subsequently extracted with 6x SDS-PAGE loading buffer. For Co-IP experiments, cells were lysed with ice-cold NETN buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA and 0.5% v/v NP40), and CD147 proteins were pulled down by anti-CD147 antibody (ab64616, Abcam). Rabbit IgG (P120-101, Bethyl Laboratories) was used as a control. CD147-binding proteins were pulled down by Protein A Sepharose beads (BioVision). Please refer to **Supplemental Table 7** for details to antibodies used in this study.

**Lentiviral production and cell transduction.** Human *FUT1*-specific shRNA expression vectors (pLKO.1) and the scrambled shRNA nontarget control (NTC) were purchased from Sigma-Aldrich. Sequences of the two shRNAs directed against human *FUT1*, *ATF4*, *CD147*, *ICAM-1*, *EGFR* and *EPHA2* are as follows:
*FUT1* clone ID NM_000148.1 (clone 544)
(CCGGCGACTGGATGTCGGAGGAGTACTCGAGTACTCCTCCGACATCCAGTCGTTTTTG)
*FUT1* clone ID NM_000148.1 (clone 565)
(CCGGCGCGGACTTGAGAGATCCTTTCTCGAGAAAGGATCTCTCAAGTCCGCGTTTTTG)
*ATF4* clone ID NM_001675 (clone 575)
(CCGGGCCAAGCACTTCAAACCTCATCTCGAGATGAGGTTTGAAGTGCTTGGCTTTTTG)
*ATF4* Clone ID NM_001675 (clone 696)
(CCGGTGGATGCCCTGTTGGGTATAGCTCGAGCTATACCCAACAGGGCATCCATTTTTG)
*CD147* clone ID NM_001728 (clone 733)
(CCGGGCTACACATTGAGAACCTGAACTCGAGTTCAGGTTCTCAATGTGTAGCTTTTTG)
*CD147* clone ID NM_001728 (clone 734)
(CCGGCCAGAATGACAAAGGCAAGAACTCGAGTTCTTGCCTTTGTCATTCTGGTTTTTG)
*ICAM-1* clone ID NM_00201 (clone 476)
(CCGGCGGCTGACGTGTGCAGTAATACTCGAGTATTACTGCACACGTCAGCCGTTTTTG)
*ICAM-1* clone ID NM_00201 (clone 629)
(CCGGCCGGTATGAGATTGTCATCATCTCGAGATGATGACAATCTCATACCGGTTTTTG)
*EGFR* clone ID NM_005228 (clone 068)
(CCGGGCCACAAAGCAGTGAATTTATCTCGAGATAAATTCACTGCTTTGTGGCTTTTTG)
*EGFR* clone ID NM_005228 (clone 634)
(CCGGGCTGGATGATAGACGCAGATACTCGAGTATCTGCGTCTATCATCCAGCTTTTTG)
*EPHA2* clone ID NM_004431 (clone 647)
(CCGGCCATCAAGATGCAGCAGTATACTCGAGTATACTGCTGCATCTTGATGGTTTTTG)
*EPHA2* clone ID NM_004431 (clone 734)
(CCGGGATAAGTTTCTATTCTGTCAGCTCGAGCTGACAGAATAGAAACTTATCTTTTTTG)

The sequence of NTC is CCGGCAACAAGATGAAGAGCACAACTCGAGTTGGTGCTCTTCATCTTGTTGTTTTT. Sequences were transfected into 293FT cells and packaged using MISSION Lentiviral Packaging Mix (Sigma-Aldrich). The open reading frame (ORF) encoding the mRNA of FUT1 was amplified from the HCC cell line Huh7 using the following primers: F (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGTGGCTCCGGAGCCATCGTCAG-3') and R (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAGGCTTAGCCAATGTCCAGAGTG-3') and cloned into the pDONRTM201 donor vector through the BP reaction. After the FUT1 ORF was found to be free of mutations, it was shuttled from pDONRTM201 to pEZ-Lv199 from GeneCopoeia through the LR reaction. pEZ-Lv199 was used as empty vector control. Sequences were transfected into 293T cells and packaged using Lenti-Pac HIV expression packaging mix (GeneCopoeia). Virus-containing supernatants were collected for subsequent transduction to establish cells with stable *ATF4/FUT1/CD147/ICAM-1/EGFR/EPHA2* repression or *FUT1* overexpression. Puromycin was used for cell selection.

**Immunohistochemistry.** Slides were heated for antigen retrieval in sodium citrate buffer, pH 6 (for CD133, ALDH1L1, Fut1, FUT1, Lewis Y, PCNA and cleaved caspase 3), and Tris-EDTA buffer, pH 9 (p-AKT). Endogenous peroxidase activity was inhibited with 3% hydrogen peroxide. Sections were subsequently incubated with CD133 (1:100, Abcam, ab19898), ALDH1L1 (1:100, Abcam, ab177463), Fut1 (1:100, Thermo Fisher Scientific, PA579287), FUT1 (1:200, Abcam, ab198712), Lewis Y (1:40, Abcam, ab3359), p-AKT (Ser473) (1:100, Cell Signaling Technology, 4060), PCNA (1:100, Abcam, ab18197) and cleaved caspase 3 (1:50, Cell Signaling Technology, 9661). The reaction was developed with the DAB+ Substrate-Chromogen System (Dako). Slides were counterstained with Mayer's hematoxylin. The expression of CD133, ALDH1L1, Fut1, FUT1, Lewis Y, PCNA and p-AKT in IHC images was quantified by NIH ImageJ v1.50 software. Five random fields were analyzed per sample. Mean OD values were taken and are expressed in fold change. The number of nuclei in the TMA images was quantified by QuPath 0.2.3 and is expressed as the mean OD per cell.

**TUNEL assay.** Detection of apoptotic cells was performed using an *in situ* cell death detection kit (Roche Diagnostics). Briefly, paraffin-embedded tissues were fixed with 15 µg/mL proteinase K in 10 mM Tris/HCl at pH 7.4. The slides were then incubated with the TUNEL reaction mixture for 1 hr at 37°C. After the slides were washed, they were incubated with a horseradish peroxidase-conjugated anti-fluorescein antibody for 30 minutes at 37°C.

**Immunofluorescence.** Cells were fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X (Sigma-Aldrich), blocked with 5% bovine serum albumin and incubated with p-AKT (Ser473) (1:400, Cell Signaling Technology, 4060) and Lewis Y (1:50, Abcam, ab3359), followed by Alexa Fluor-conjugated secondary antibody. Cells were counterstained with anti-fade DAPI (Invitrogen) and visualized by fluorescence confocal microscopy (Carl Zeiss LSM 700).

**Annexin V apoptosis assay.** Cells were treated with sorafenib for 48 hr. Following treatment, cells were harvested and stained with propidium iodide (PI) and FITC-conjugated Annexin V as provided by the Annexin V-FLUOS Staining Kit (Roche). Samples were analyzed on a BD FACSCanto II (BD Biosciences) with data analyzed by FlowJo (Tree Star, Inc.).

**Cell viability assay.** The cell viability of organoid cultures pretreated with 2DGal for 24 hr and treated with specified concentrations of sorafenib for 48 hr prior to evaluation was determined with the CellTiter-Glo Luminescent Cell Viability Assay (Promega) according to the manufacturer's protocol.

**Flow cytometry for CD133 and ALDH activity.** Flow cytometric analysis was conducted using PE-conjugated CD133 (Miltenyi Biotec) or unconjugated CD147 (Abcam) and their respective isotype controls (eBiosciences). Alexa Fluor 594 was used as a secondary antibody for CD147 (Thermo Fisher). The ALDEFLUOR kit was used for ALDH activity measurements (Stem Cell Technologies). Samples were analyzed on a BD FACSCanto II (BD Biosciences) with data analyzed by FlowJo (Tree Star, Inc.).

**Transcriptome sequencing.** Transcriptome sequencing was performed as a service at the Centre for PanorOmic Sciences – Genomics and Bioinformatics Cores of The University of Hong Kong. A poly(A)+ mRNA library was prepared using the KAPA stranded mRNA-Seq kit (Illumina). With an Illumina HiSeq1500 system, 101 bp paired-end sequencing was performed. Each sample had an average throughput of 8.9 GB and a total throughput of 17.8 An average of 94% of the bases achieved a quality score of Q30, where Q30 denotes the accuracy of a base call to be 99.9%. Alignment, expression estimation and tests for differential expression were processed by RSEM 1.2.21 and EBSeq 1.6.0. The resulting values were indicated by transcripts per million (TPM). Based on a fold-change cutoff of >5, deregulated genes were subjected to DAVID and GSEA for pathway enrichment analyses. Gene expression data were deposited in the Gene Expression Omnibus, and the accession number is GSE152927.

**Reverse phase protein array (RPPA).** Cells were lysed in RIPA buffer with protease inhibitor cocktails and PhosSTOP (Roche). Cell lysates were mixed with 4x SDS sample buffer without bromophenol blue and boiled for 5 min. Samples were sent to the RPPA Core Facility at MD Anderson Cancer Center for analysis.

**Mass spectrometry-based fucosylated peptide profiling.** An adapted protocol was used for fucosylated glycopeptide enrichment from 2 mg of total proteins/peptides using LCA and UEA-1 lectins (total peptides: lectin ratio = 1:1, in terms of wt./wt. and vol./vol.) (40). Eluted glycopeptides (with elution buffers - 200 mM α-methyl mannoside and 200 mM α-methyl glucoside in TBS for LCA; 100 mM α-methyl-fucose in TBS for UEA-I) were desalted with Sep-Pak C18 cartridges, SpeedVac dried, and reconstituted in 0.1% FA prior to nano LC-MS/MS analysis. Fucosylated peptide samples were loaded onto a Dionex UltiMate 3000 RSLC nano liquid chromatograph coupled to an Orbitrap Fusion Lumos Tribid mass spectrometer. Peptides were separated on a commercial C18 column (75 μm i.d. × 50 cm length) with 1.9 μm particle size. Separation was attained using a linear gradient of increasing buffer B (80% ACN and 0.1% formic acid) and declining buffer A (0.1% formic acid) at 300 nL/min. Buffer B was increased to 30% B in 210 min and ramped to 40% B in 10 min followed by a quick ramp to 95% B, where it was held for 5 min before a quick ramp back to 5% B. Full MS survey scan resolution was set to 120000 with an automatic gain control (AGC) target value of $2 \times 10^6$, maximum ion injection time (IT) of 30 ms, and for a scan range of 400−1500 m/z, RF lens- 50. Data were acquired in DDA mode to isolate and fragment topN multiply charged precursors according to their intensities. MS2 spectra were obtained at 30000 resolution with the following parameters: HCD triggered to obtain glycan composition, and parallel charge state enabled ETD to give glycopeptide identification - product ion enabled MS2 HCDpdEThcD method, charge state 3-8, AGC target of 1 ×

$10^5$, maximum ion injection time (IT) of 100 ms, 1.6 m/z isolation width, and normalized collisional energy of 31. Preceding precursor ions targeted for HCD were dynamically excluded after 50 s. The detection of characteristic glycan ions for HexNac, HexNac fragment, and HexNacHex with m/z 204.0867, 138.0545, and 366.1396 in the MS2 HCD spectrum triggered further fragmentation (of these product ions) via EThCD or CID activation. Additional verification (of the glycans) was gained through the detection of the subfragment ions generated. This analysis strongly enhanced the comprehensiveness of glycopeptide profiling. Data analysis was performed using Byonic version 3.3.11. Confident criteria include 1% FDR and best score >200 for fucosylated glycoprotein identification. Data are available via ProteomeXchange with identifier PXD024437.

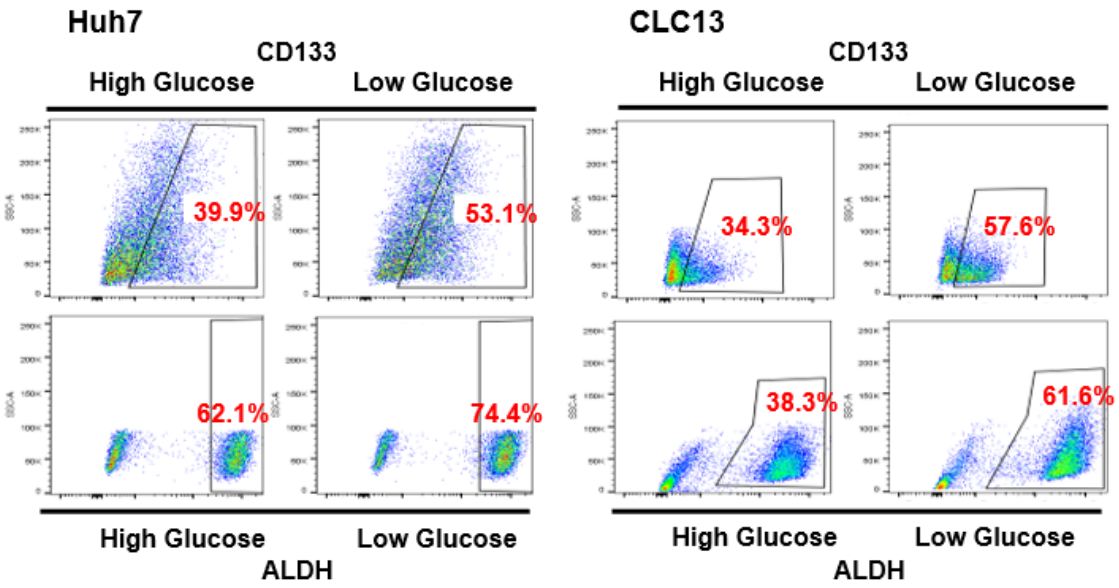
**Luciferase reporter assay.** For promoter activity, full-length *FUT1*, Δ5' *FUT1* and Δ3' *FUT1* promoter regions were amplified from the HCC cell line Huh7 and cloned into the pRL-CMV luciferase reporter vector (Promega) using the primers listed in **Supplemental Table 8**. Luciferase reporter constructs were transiently transfected into HCC cells by Lipofectamine 3000 reagent (Thermo Scientific). Luciferase activity was determined by the Dual-Glo Luciferase Assay System (Promega) and normalized based on cotransfected pRL-CMV *Renilla* luciferase control. Luciferase reporter activity represented by a ratio of firefly:*Renilla* luminescence.

**Chromatin immunoprecipitation (ChIP) assay.** ChIP was performed with the Magna ChIP G – Chromatin Immunoprecipitation Kit (Millipore) according to the manufacturer's instructions. Briefly, cells were crosslinked in the presence of 1% formaldehyde. DNA was sonicated and immunoprecipitated with anti-ATF4 (sc-200, Santa Cruz Biotechnology) or rabbit IgG control (P120-101, Bethyl Laboratories). Immunoprecipitated and eluted DNA was purified with columns and amplified by qPCR with the following primers: 5' promoter region F (5'-TCTGGTGAAAGAAACACCAGGAT-3') and R (5'- ATCTGTCCCTGCTCAGATGCT-3') and 3' promoter region F (5'- CCCTCCATTGAAAGGACCCC-3') and R (5'-CCCATTTGGCAGGAGCTACA-3').

## SUPPLEMENTAL FIGURES

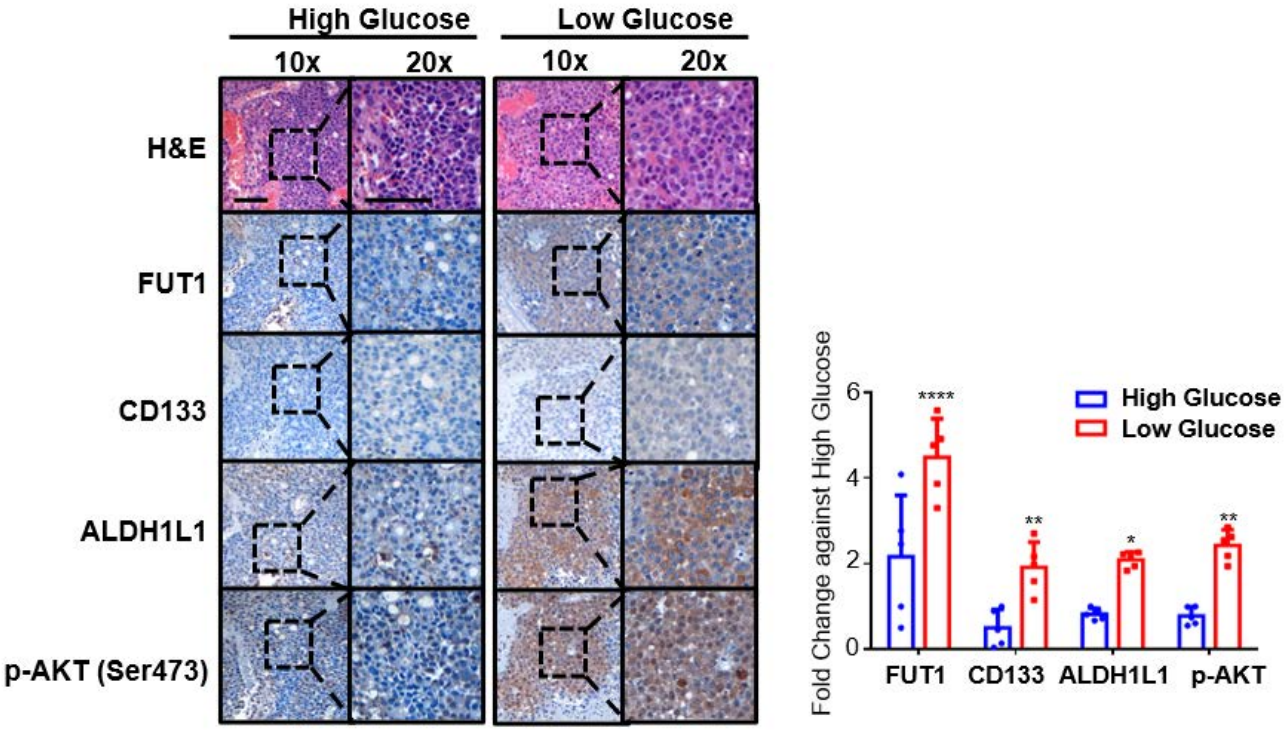
### Supplemental Figure 1. (Related to Figure 1)

**Glucose restriction promotes elevated expression of known HCC CSC markers.** Flow cytometric analysis found that CD133 expression and aldehyde dehydrogenase (ALDH) activity were enhanced when Huh7 and CLC13 HCC cells were cultured in low glucose conditions. The data shown are representative of three independent experiments.



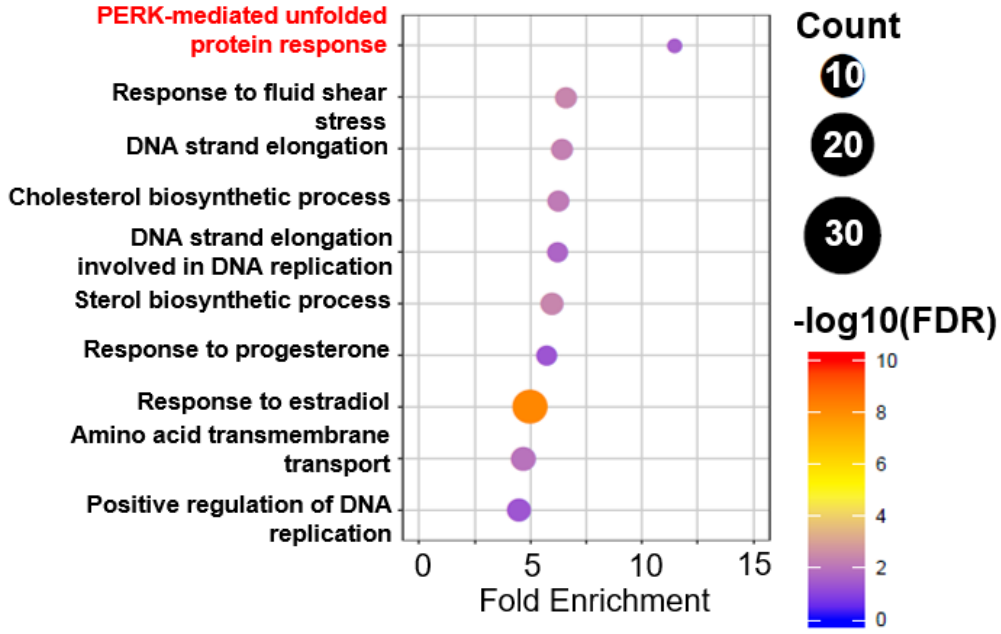
**Supplemental Figure 2. (Related to Figures 1 and 2)**

**Glucose restriction promotes elevated expression of CD133, ALDH1L1, FUT1 and p-AKT in *in vivo* HCC tumors.** H&E and immunohistochemical analysis of FUT1, CD133, ALDH1L1 and p-AKT (Ser473) expression in sections of harvested livers of the mice injected with HCC cells cultured in high or low glucose for 6 days. Data show elevated expression of CD133, ALDH1L1, FUT1 and p-AKT in tumors formed from HCC cells cultured in low glucose conditions. Scale bar = 50 µm. Scale bar in inset = 50 µm. Graph – fold change against high glucose. H&E, hematoxylin and eosin. *, ** and **** indicate $p<0.05$, $p<0.01$ and $p<0.0001$, respectively. (Unpaired Student's t-test).



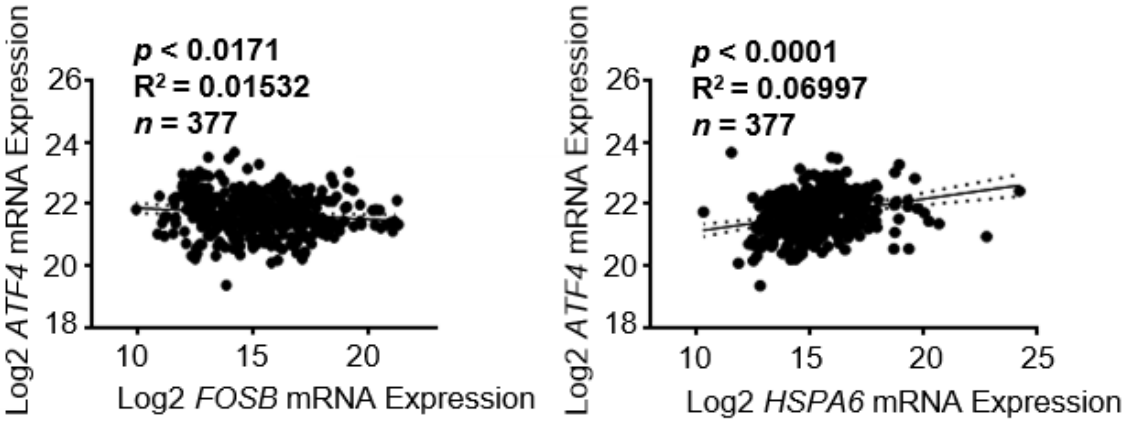
**Supplemental Figure 3. (Related to Figure 1)**

**PERK-mediated unfolded protein response was identified to be the top enriched pathway in low glucose treated HCC cells.** DAVID gene ontology analysis of differentially expressed genes identified by RNA-seq found that the PERK-mediated unfolded protein response was highly enriched in HCC cells cultured under low glucose conditions. FDR, false discovery rate.



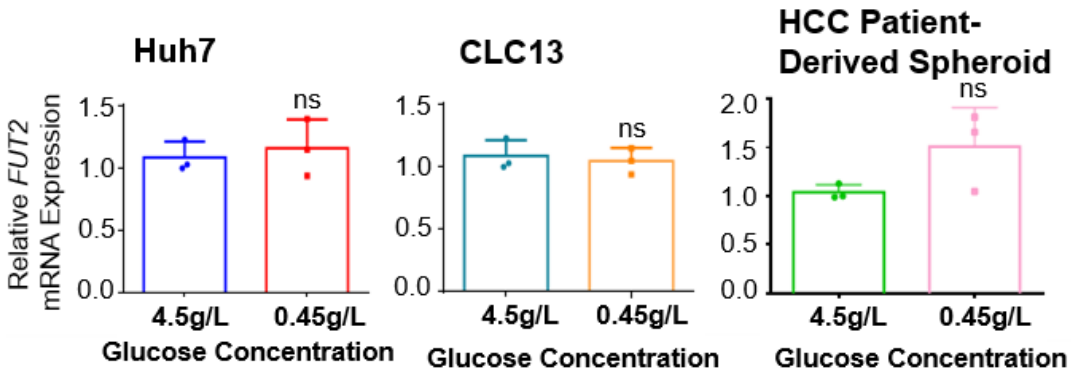
**Supplemental Figure 4. (Related to Figure 2)**

**Correlation of *ATF4* with *FOSB* and *HSPA6* in HCC clinical samples.** Analysis of *FOSB*, *HSPA6* and *ATF4* expression in The Cancer Genome Atlas (TCGA) – Liver Hepatocellular Carcinoma (LIHC) database. The expression of *ATF4* and *FOSB* was negatively correlated and thus was not explored further. The expression of *ATF4* and *HSPA6* was positively correlated, but *HSPA6* did not harbor ATF4 binding sites and thus was also not further studied. (Pearson correlation test).



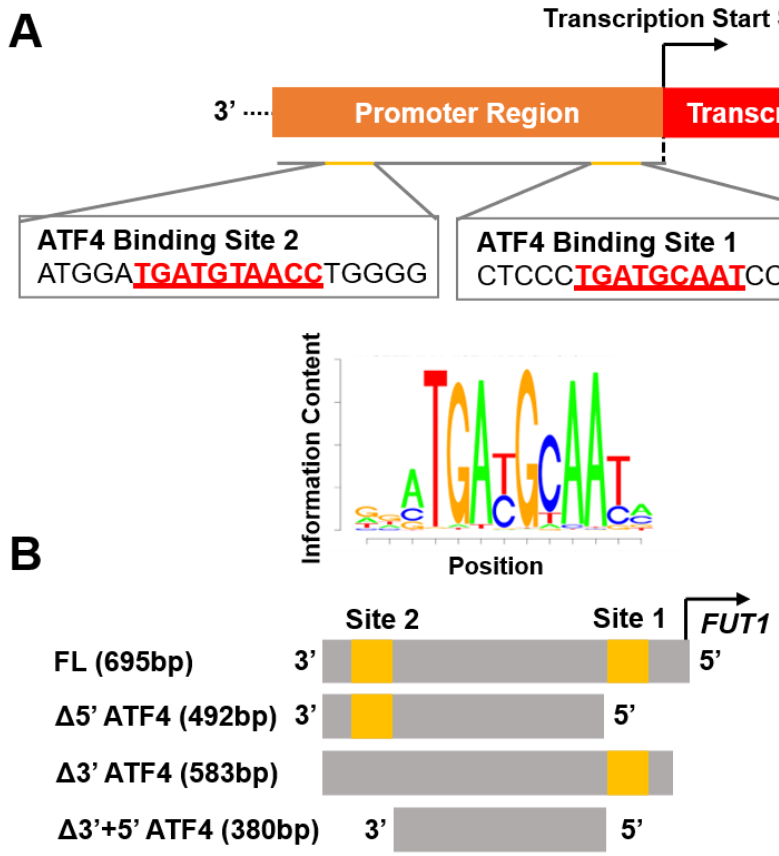
**Supplemental Figure 5. (Related to Figures 2)**

**Glucose restriction does not alter *FUT2* expression.** qPCR showed that *FUT2* expression remained unchanged after culture of Huh7, CLC13 and HCC patient-derived spheroids in low glucose culturing conditions. The data shown are representative of three independent experiments. (Unpaired Student's t-test). ns, not significant.



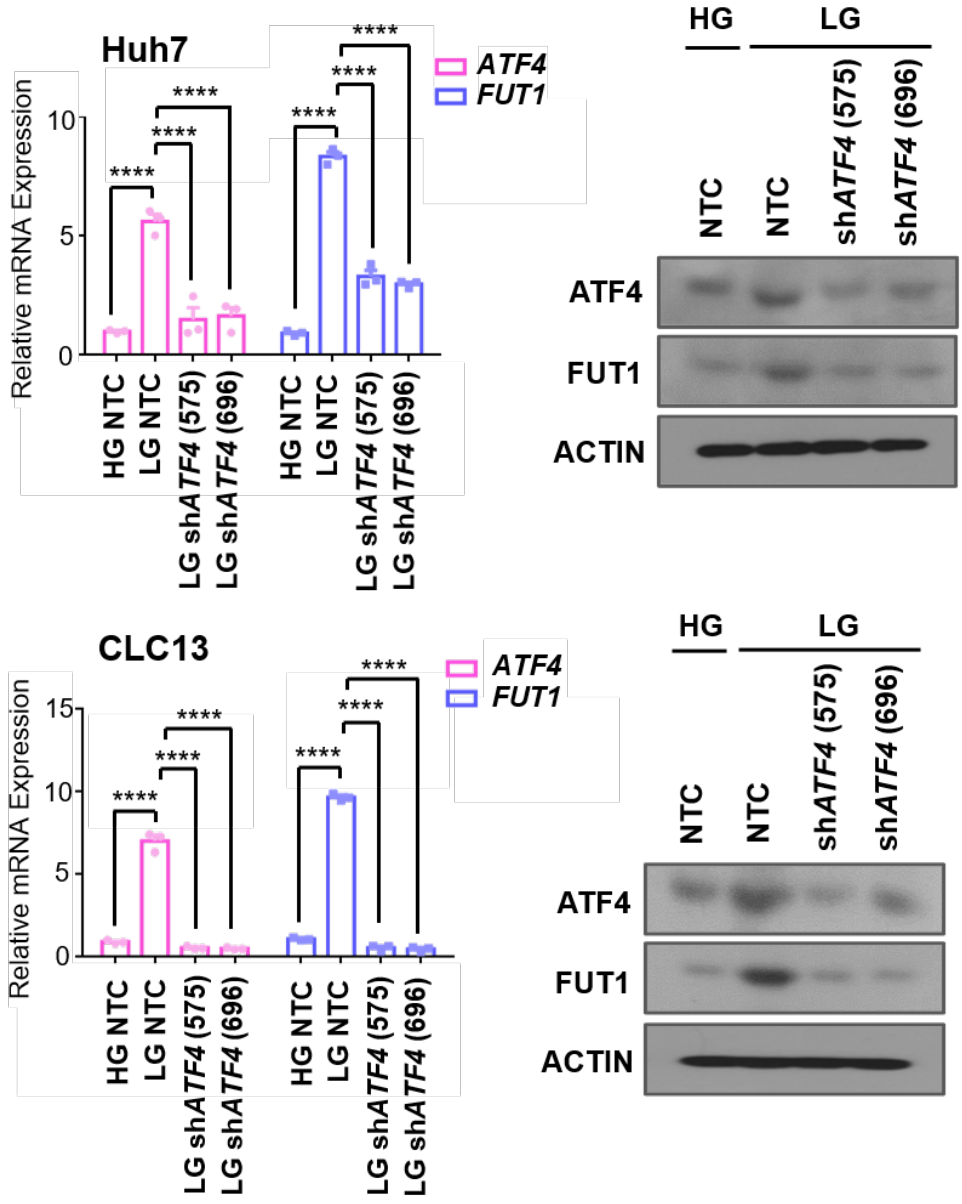
**Supplemental Figure 6. (Related to Figures 2)**

**Prediction ATF4 binding sites on *FUT1* promoter. (A)** Locations of the two ATF4 binding sites on the *FUT1* promoter region predicted by the Gene Transcription Regulation Database (GTRD: http://gtrd.biouml.org/). **(B)** Schematic diagram illustration of full-length (FL) and deletion mutants of *FUT1* used in this study. Deletion mutations designed to contain (i) the 3' binding site only (Δ5' ATF4), (ii) the 5' binding site only (Δ3' ATF4) and (iii) both the 5' and 3' binding sites deleted (Δ3' + 5' ATF4). FL, full-length.



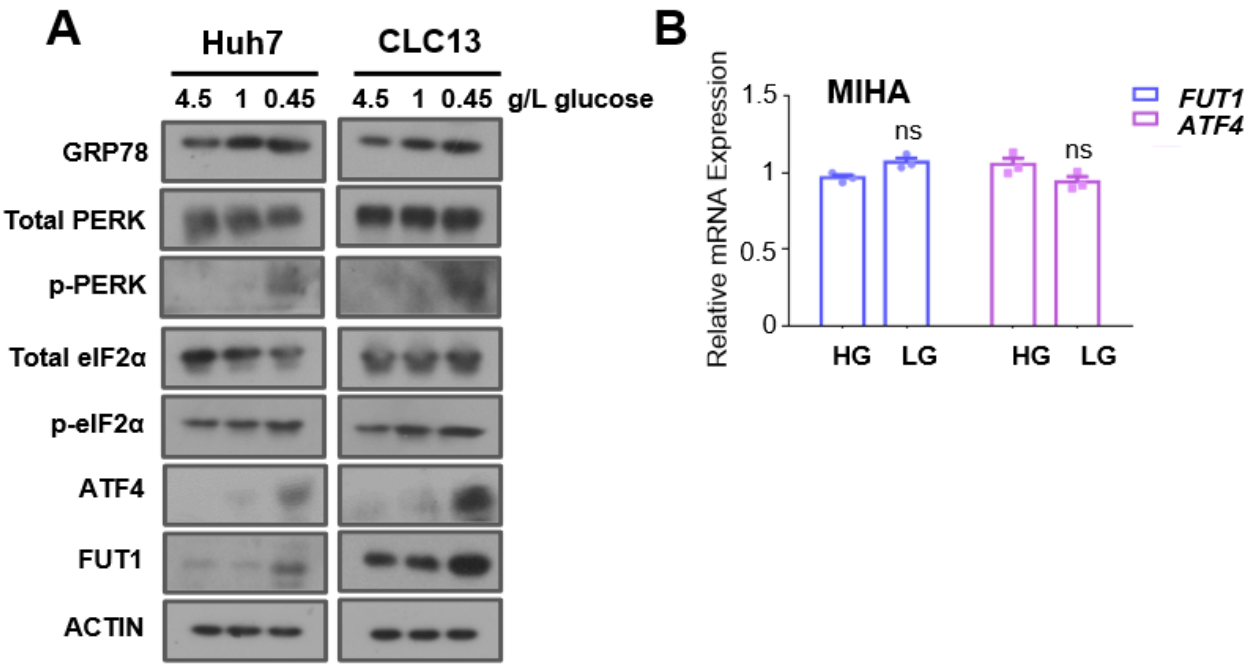
**Supplemental Figure 7. (Related to Figure 2)**

**Stable knockdown of *ATF4* expression attenuated FUT1 expression in Huh7 and CLC13 HCC cells cultured in low glucose**. The data shown are representative of three independent experiments. HG, high glucose; LG, low glucose; NTC, nontarget control; sh*ATF4* clones, 575 and 696. **** indicates $p<0.0001$. (One-way ANOVA).



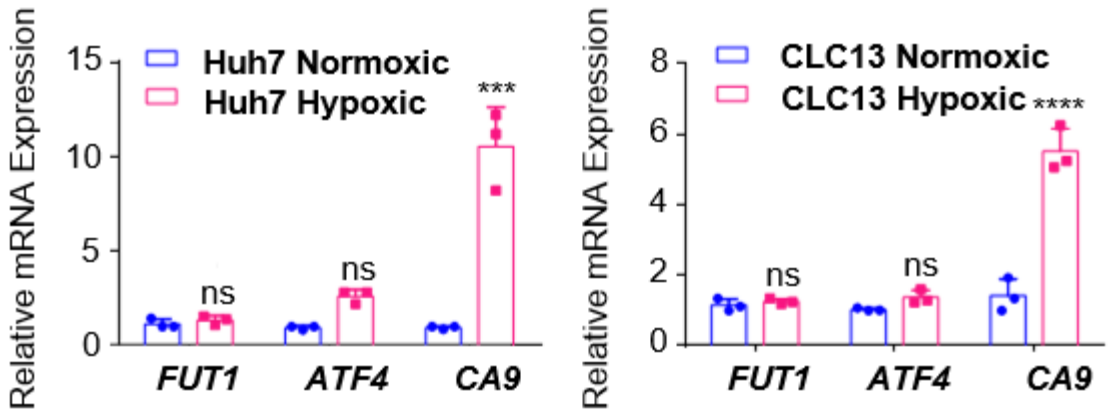
**Supplemental Figure 8. (Related to Figure 1 and 2)**

**Glucose restriction induces PERK-mediated UPR-related genes and FUT1. (A)** Huh7 and CLC13 HCC cells were cultured in medium supplemented with high (hyperglycemic; 4.5 g/L), medium (physiological; 1 g/L) and low (hypoglycemic; 0.45 g/L) glucose concentrations for 1 week. Western blot analysis found that GRP78, p-PERK, p-eIF2α, ATF4 and FUT1 were enhanced in low glucose conditions. The data shown are representative of three independent experiments. **(B)** The immortalized normal liver cell line MIHA showed no change in *ATF4* or *FUT1* expression when cultured in low glucose. The data shown are representative of three independent experiments. HG, high glucose; LG, low glucose; ns, not significant. (Unpaired Student's t-test).



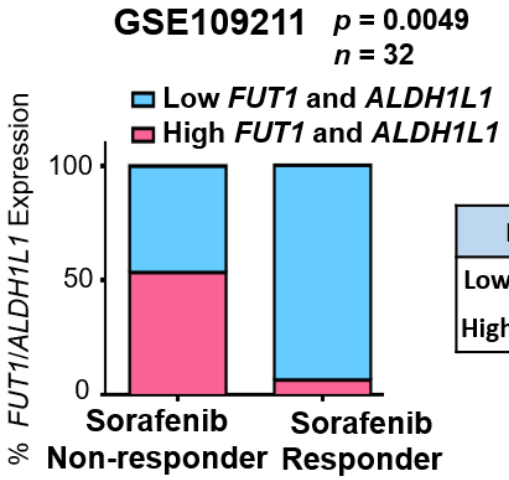
**Supplemental Figure 9. (Related to Figure 1 and 2)**

**PERK-mediated UPR activation of *FUT1* promoter activity via *ATF4* induction is not relevant to the hypoxic tumor microenvironment.** qPCR analysis of *FUT1* and *ATF4* mRNA expression showed no significant change when Huh7 and CLC13 HCC cells were cultured under normoxic (21% $O_2$) or hypoxic (1% $O_2$) conditions. *CA9*, a marker of hypoxia, was used as a positive control. The data shown are representative of three independent experiments. ns, not significant. *** and **** for $p<0.001$ and $p<0.0001$, respectively. (Unpaired Student's t-test).



**Supplemental Figure 10. (Related to Figure 3)**

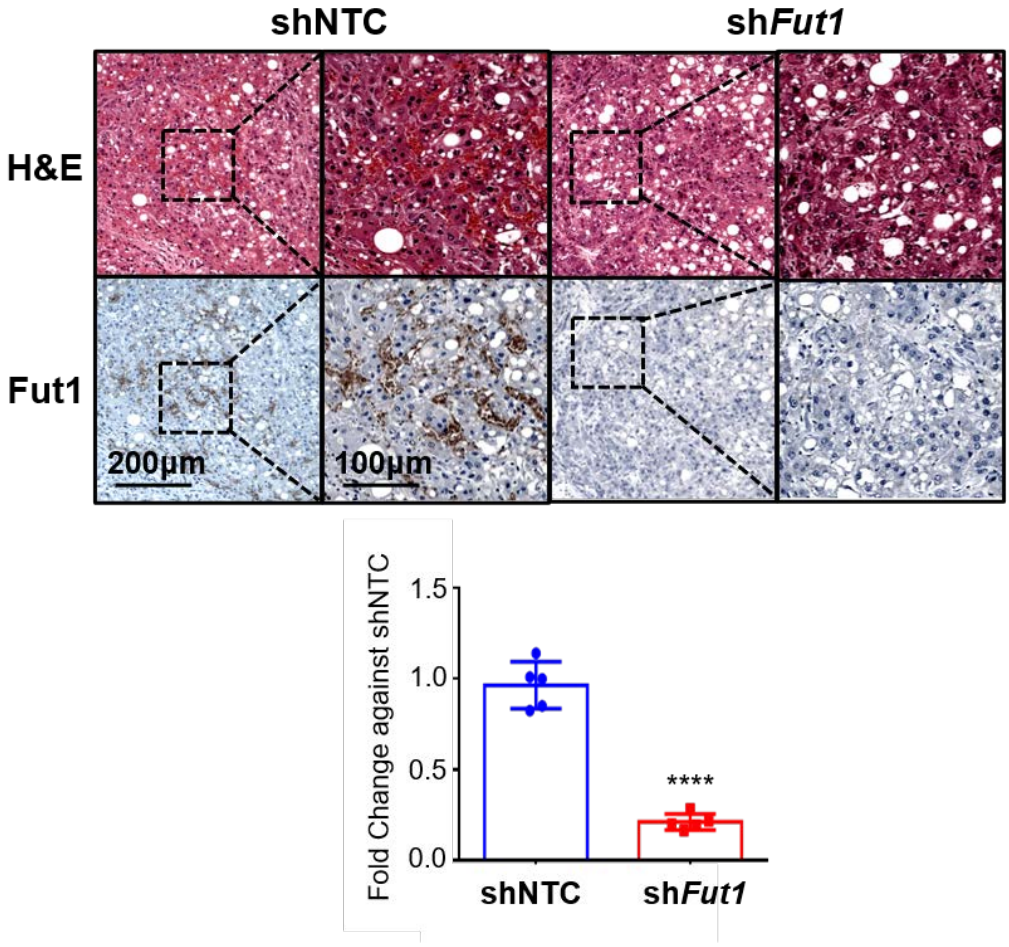
**Analysis of the GSE109211 dataset from the GEO and NCBI found that HCC patients with low *FUT1* and *ALDH1L1* levels were more responsive to sorafenib than HCC patients with high *FUT1* and *ALDH1L1* levels.** (Fisher's Exact test).




| Expression Level | Non-responder | Responder |
|---|---|---|
| Low *FUT1* and *ALDH1L1* | 7 (46.7%) | 16 (94.1%) |
| High *FUT1* and *ALDH1L1* | 8 (53.3%) | 1 (6.25%) |

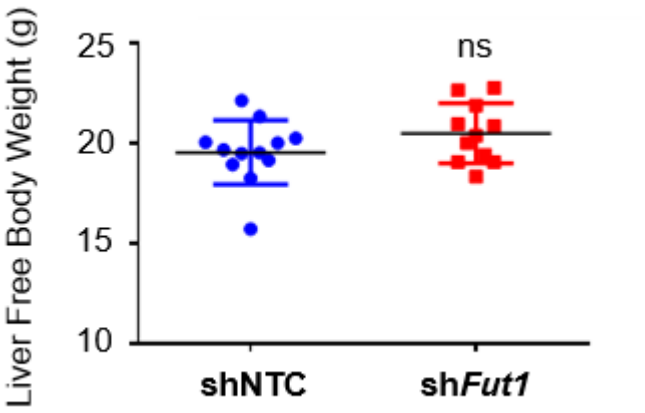
**Supplemental Figure 11. (Related to Figure 4)**

**H&E and immunohistochemical analysis of *Fut1* expression in sections of harvested livers of mice injected with shNTC or sh*Fut1* lentiviral particles.** Scale bar = 200 µm. Scale bar in inset = 100 µm. Graph – fold change against shNTC. NTC, nontarget control; sh, shRNA; H&E, hematoxylin and eosin. **** for $p<0.0001$. (Unpaired Student's t-test).



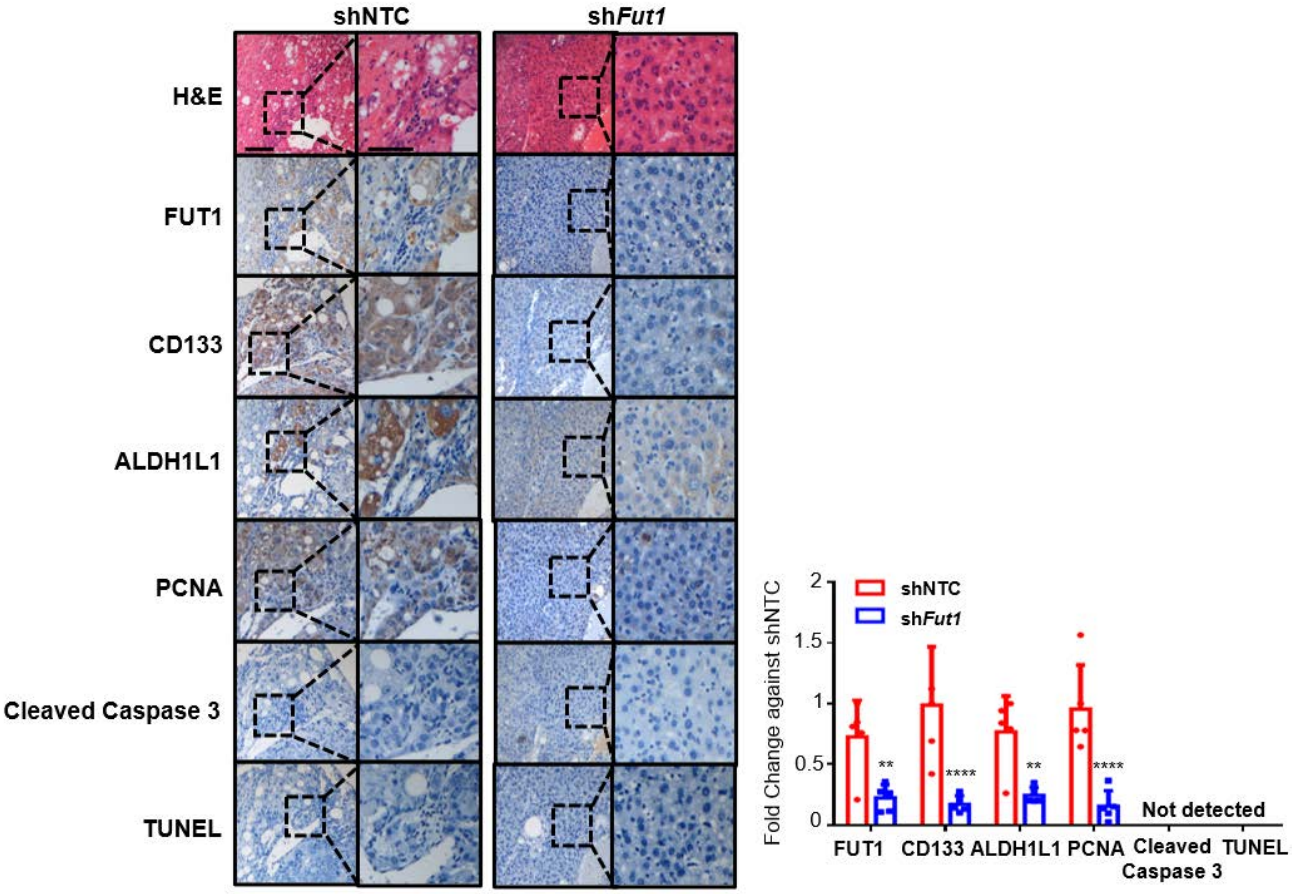
**Supplemental Figure 12. (Related to Figure 4)**

**Endogenous *Fut1* knockdown in the liver of an immunocompetent HCC mouse model did not significantly alter liver free body weight.** NTC, nontarget control; sh, shRNA; ns, not significant. (Unpaired Student's t-test).



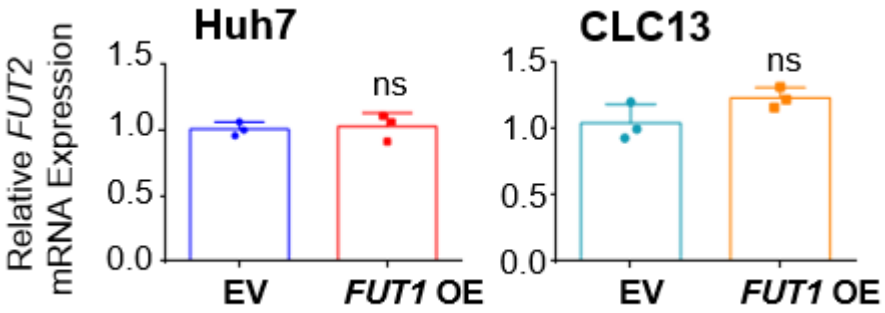
**Supplemental Figure 13. (Related to Figure 4)**

**Endogenous *Fut1* knockdown in the liver of an immunocompetent HCC mouse model attenuates hepatocarcinogenesis.** H&E and immunohistochemical analysis of the levels of FUT1, CD133, ALDH1L1, the proliferation marker PCNA and the apoptotic marker cleaved caspase 3 as well as TUNEL staining in sections of harvested livers of the mice injected with lentiviral particles of shNTC or sh*Fut1*. Data show attenuated expression of FUT1, CD133, ALDH1L1, and PCNA following sh*Fut1* knockdown. Spontaneous apoptosis was not detected in shNTC or sh*Fut1* cells. Scale bar = 100 µm. Scale bar in inset = 50 µm. Graph – fold change against shNTC. NTC, nontarget control; sh, shRNA; H&E, hematoxylin and eosin. ** and **** for $p<0.01$ and $p<0.0001$, respectively. (Unpaired Student's t-test).



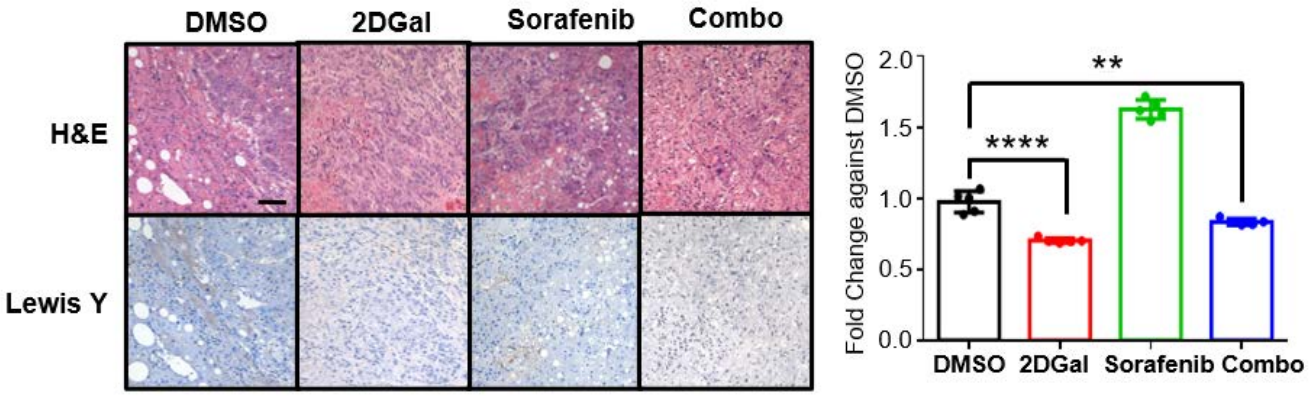
**Supplemental Figure 14. (Related to Figure 5)**

**Manipulation of *FUT1* expression did not affect *FUT2* expression.** qPCR found that *FUT2* expression remained unchanged when *FUT1* was overexpressed in Huh7 and CLC13 cells. The data shown are representative of three independent experiments. EV, empty vector control; OE, overexpression; ns, not significant. (Unpaired Student's t-test).



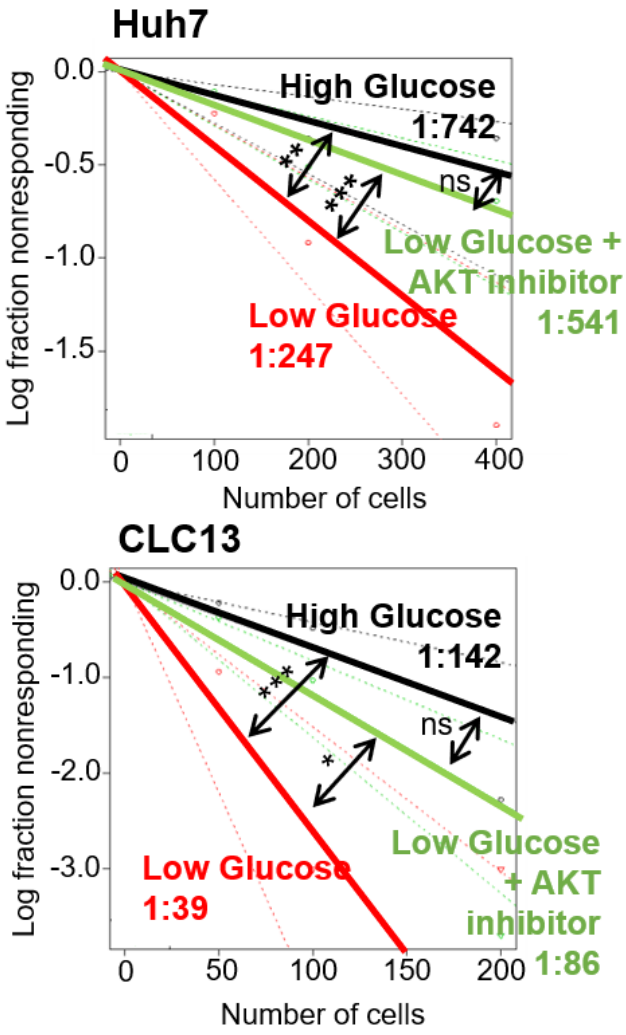
**Supplemental Figure 15. (Related to Figure 6)**

**Inhibition of α-(1,2)-fucosylation by 2DGal increases the efficacy of sorafenib and eradicates tumor-initiating cells, while 2DGal effectively impairs the formation of Lewis Y in resected livers of 2DGal and combination treatment groups.** H&E and immunohistochemical analysis of Lewis Y and p-AKT expression in resected HCC tumors of each treatment group. Scale bar = 100 µm. Graph – fold change against DMSO. H&E, hematoxylin and eosin; 2DGal, 2-deoxy-D-galactose. ** and **** for $p<0.01$ and $p<0.0001$, respectively. (One-way ANOVA).



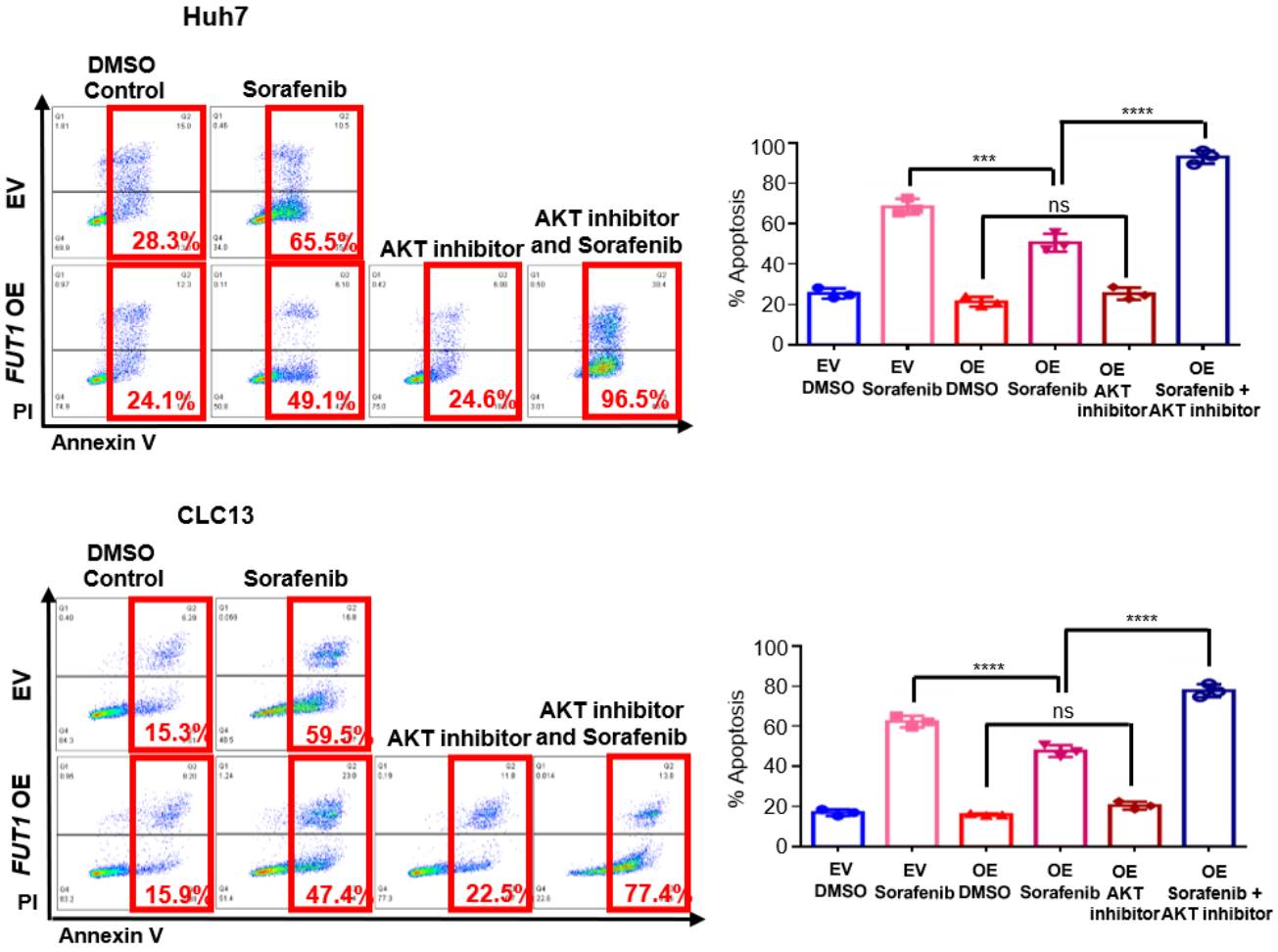
**Supplemental Figure 16. (Related to Figure 7)**

***In vitro* limiting dilution assays found that the frequency of tumor-initiating cells increased after culture in low glucose and decreased when HCC cells cultured in low glucose were treated with an AKT inhibitor (5 μM).** The data shown are representative of three independent experiments. HG, high glucose; LG, low glucose; ns, not significant. *, ** and *** indicate $p<0.05$, $p<0.01$ and $p<0.0001$, respectively. (Pairwise tests for differences in stem cell frequencies).



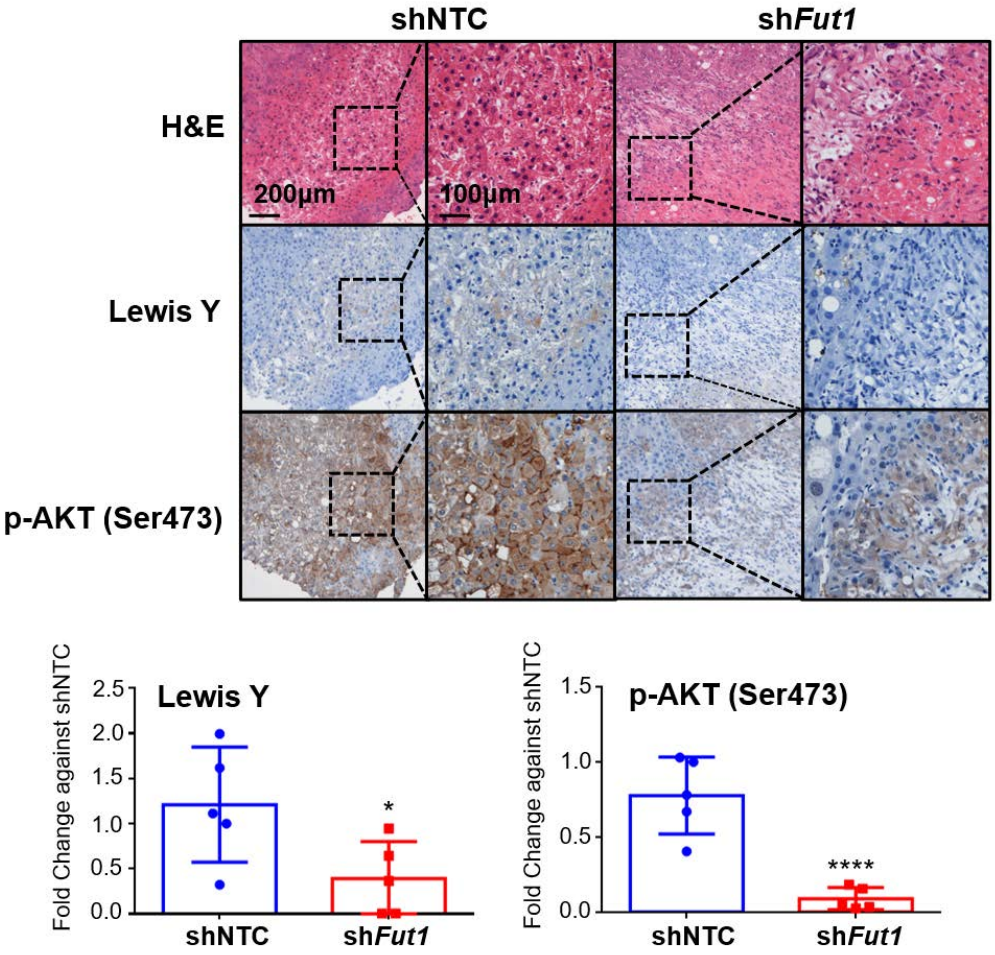
**Supplemental Figure 17. (Related to Figure 7)**

***FUT1* mediates sorafenib resistance in HCC through dysregulated AKT signaling.** Annexin V PI analysis found that *FUT1* expression augmented Huh7 cells treated with 4 μM sorafenib (72 hours) to induce apoptosis; while addition of an AKT inhibitor would reverse the effect. 2% DMSO was used as a control. The data shown are representative of three independent experiments. EV, empty vector control; OE, overexpression; PI, propidium iodide; ns, not significant. *** and **** indicate $p<0.001$ and $p<0.0001$, respectively. (One-way ANOVA).



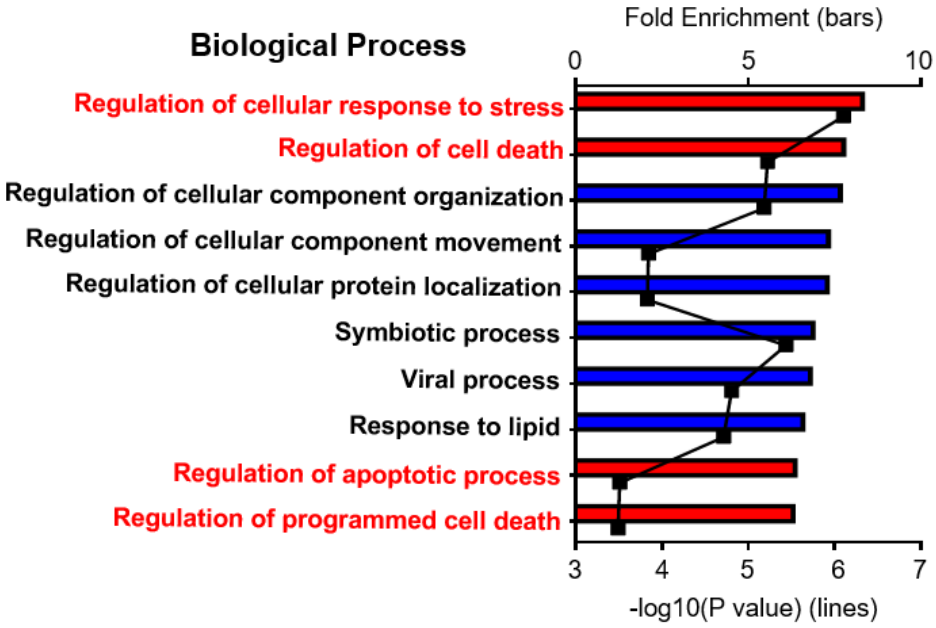
**Supplemental Figure 18. (Related to Figure 7)**

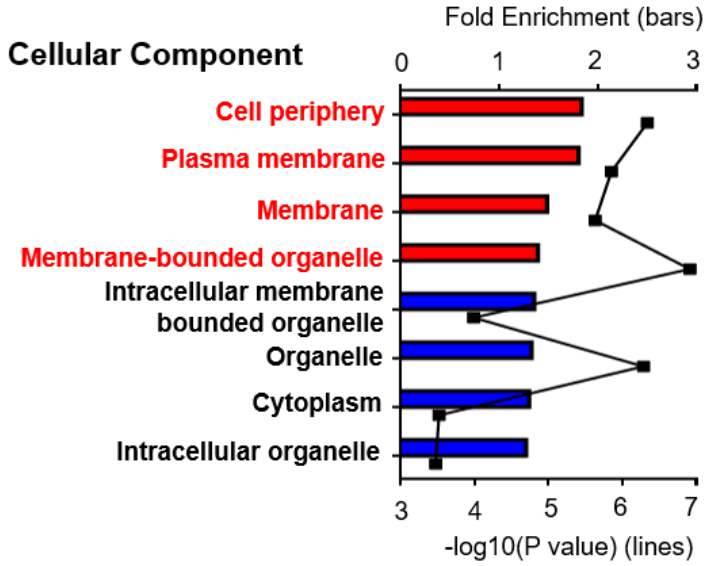
**H&E and immunohistochemical analysis of Lewis Y and p-AKT expression in sections of harvested livers of the mice injected with shNTC or sh*FUT1* lentiviral particles.** Scale bar = 200 µm. Scale bar in inset = 100 µm. Graph – fold change against shNTC. NTC, nontarget control, sh, shRNA; H&E, hematoxylin and eosin. * and **** indicate $p<0.05$ and $p<0.0001$, respectively. (Unpaired Student's t-test).



**Supplemental Figure 19. (Related to Figure 8)**

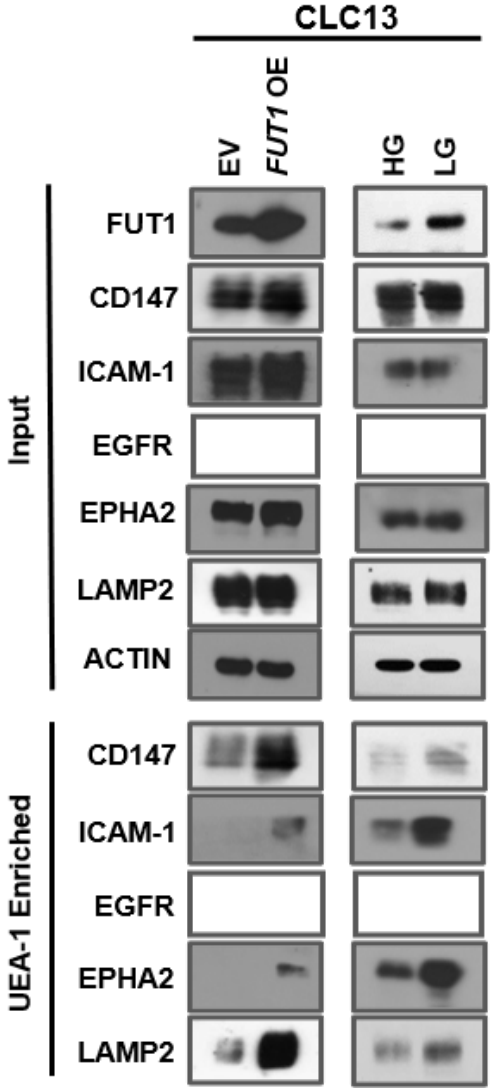
**Gene Ontology enrichment analysis (biological process and cellular component) of the 30 α-(1,2)-fucosylated proteins identified in fucosylated peptide profiling.**



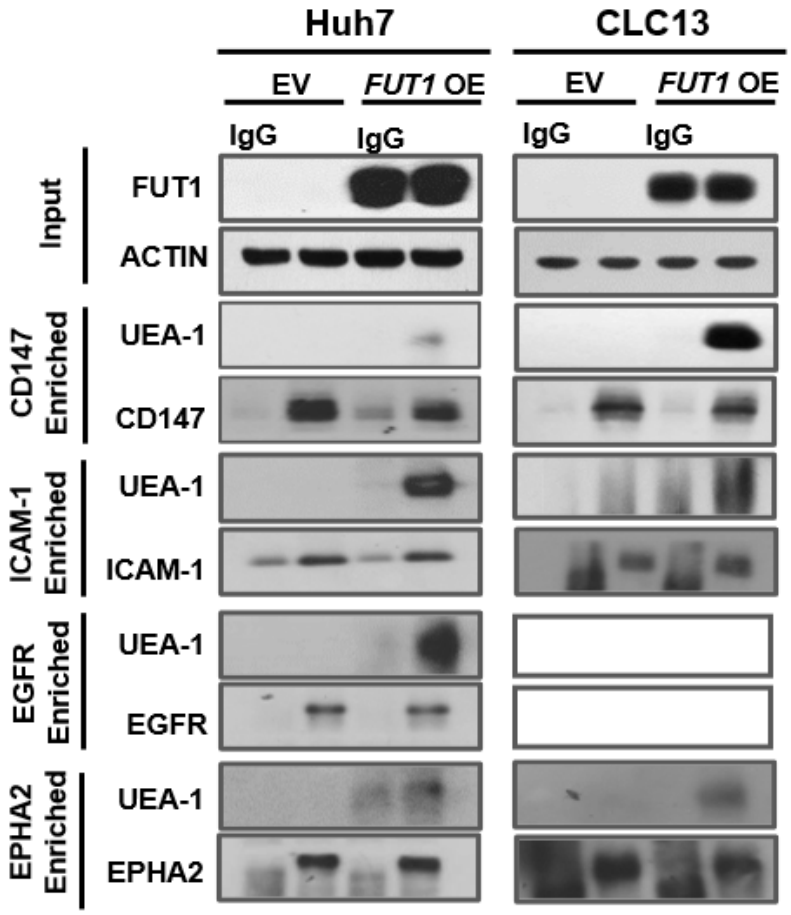
**Supplemental Figure 20. (Related to Figure 8)**

**Validation of α1,2-fucosylation on CD147, ICAM-1, EGFR and EPHA2.** UEA-1 affinity chromatography of whole cell lysates of CLC13 HCC cells transfected with empty vector (EV) or *FUT1* or CLC13 HCC cells cultured in high or low glucose, followed by Western blots with LAMP2, CD147, ICAM-1, EGFR and EPHA2 antibodies. Input shows no effect of glucose restriction and *FUT1* overexpression on protein expression. LAMP2 was used as positive control. EGFR cannot be detected in CLC13. The data shown are representative of three independent experiments. EV, empty vector control; OE, overexpression; HG, high glucose; LG, low glucose.



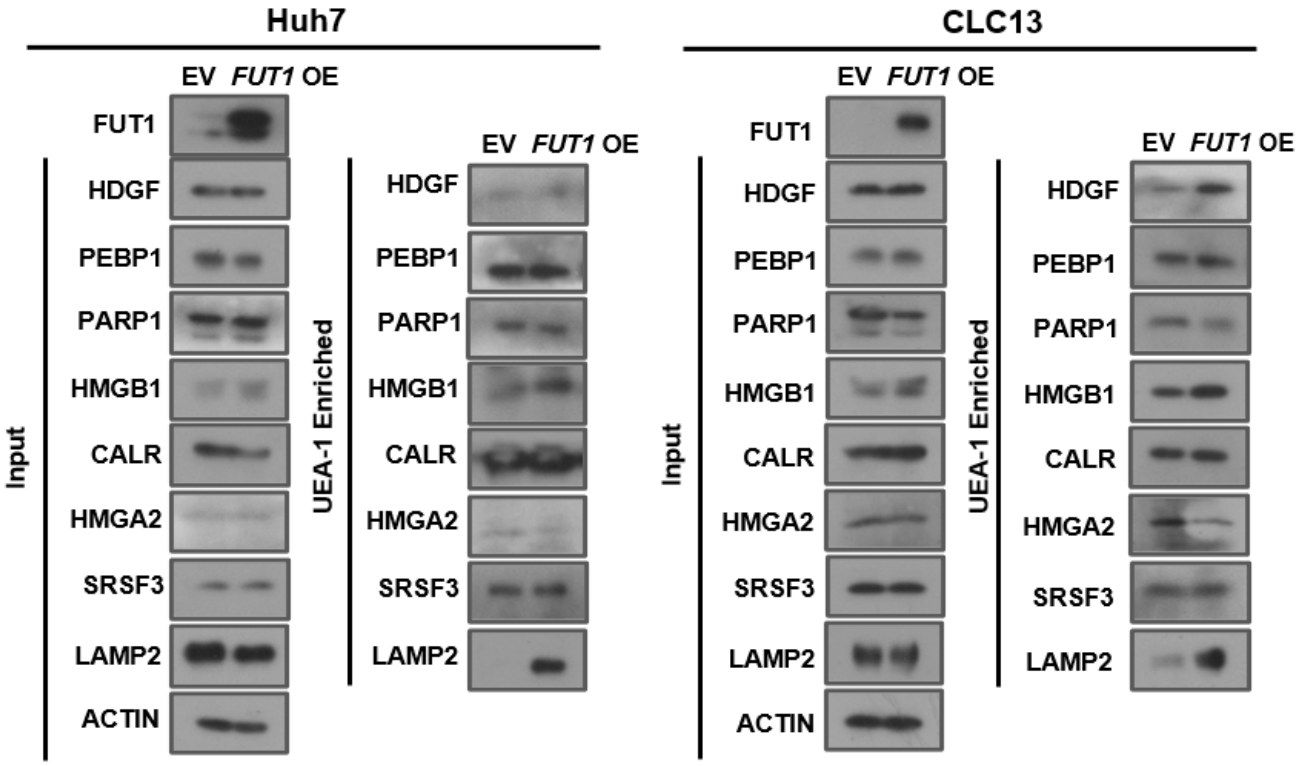
**Supplemental Figure 21. (Related to Figure 8)**

**Validation of α1,2-fucosylation on CD147, ICAM-1, EGFR and EPHA2.** Immunoprecipitation of CD147, ICAM-1, EGFR or EPHA2 from whole cell lysates of Huh7 and CLC13 HCC cells transfected with EV or with *FUT1* overexpressed followed by Western blots with UEA-1, CD147, ICAM-1, EGFR and EPHA2 antibodies. EGFR cannot be detected in CLC13. The data shown are representative of three independent experiments. EV, empty vector control; OE, overexpression.



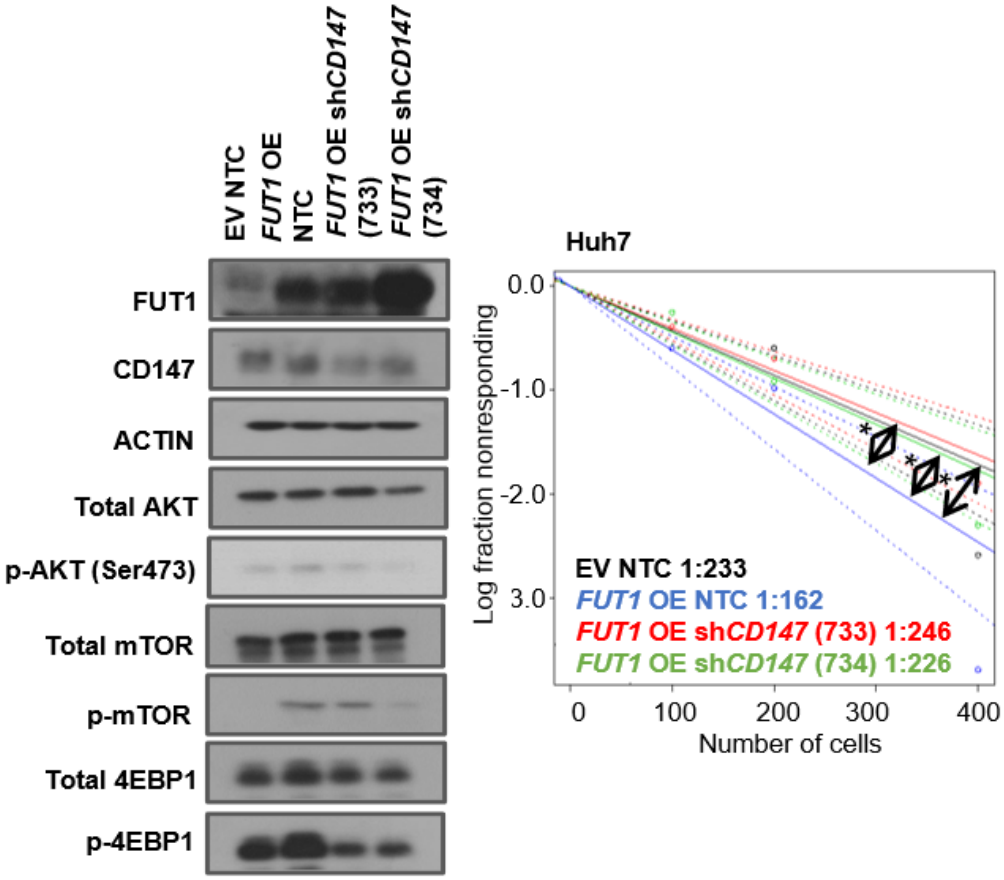
**Supplemental Figure 22. (Related to Figure 8)**

**Validation of α1,2-fucosylation on HDGF, PEBP1, PARP1, HMGB1, CALR, HMGA2 and SRSF3 proteins.** UEA-1 affinity chromatography of whole cell lysates of Huh7 and CLC13 HCC cells transfected with empty vector (EV) or with *FUT1* overexpressed, followed by Western blots with LAMP2, HDGF, PEBP1, PARP1, HMGB1, CALR, HMGA2 and SRSF3 antibodies. Input shows no effect of *FUT1* overexpression on protein expression. LAMP2 was used as positive control. The data shown are representative of three independent experiments. EV, empty vector control; OE, overexpression.



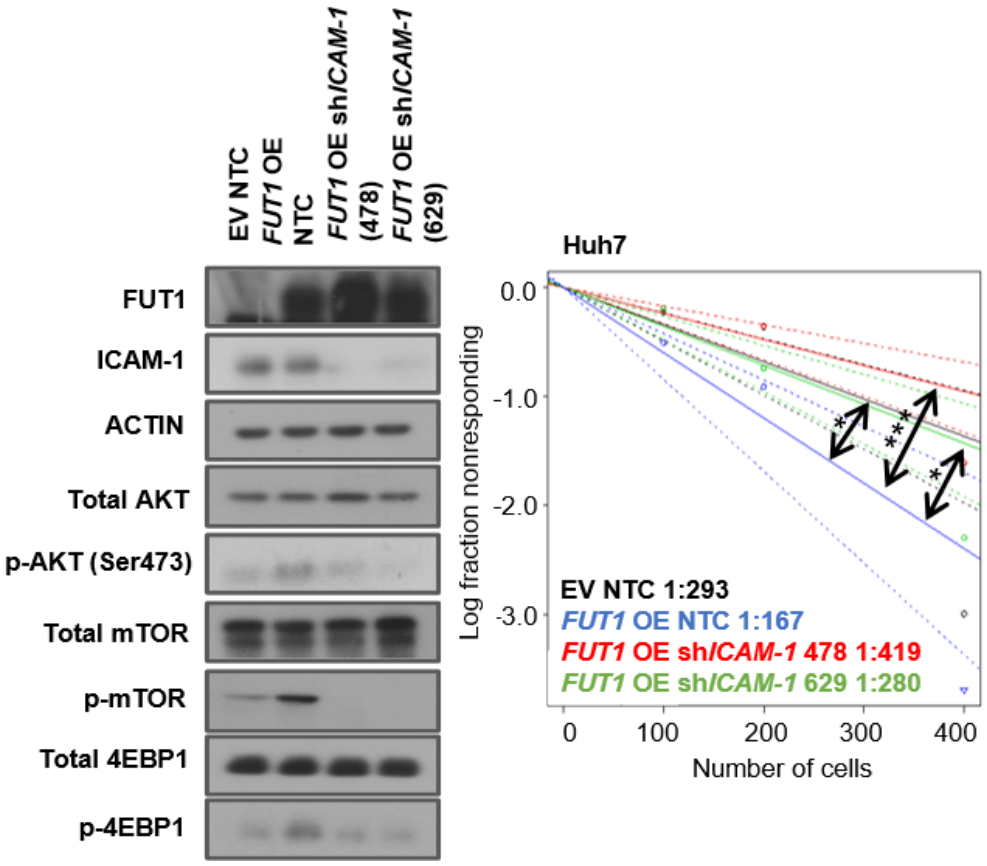
**Supplemental Figure 23. (Related to Figure 8)**

**Effects of *CD147* knockdown on AKT-mTOR-4EBP1 signaling and cancer stemness.** Western blot analysis also found that p-AKT, p-mTOR and p-4EBP1 levels were attenuated upon *CD147* knockdown in *FUT1*-overexpressing HCC Huh7 cells. *In vitro* limiting dilution assays found that the frequency of tumor-initiating cells decreased following knockdown of *CD147* in *FUT1*-overexpressing HCC Huh7 cells. All data shown are representative of three independent experiments. NTC, nontarget control; sh*CD147* clones, 733 and 734; EV, empty vector control; OE, overexpression. * for $p<0.05$. (Pairwise tests for differences in stem cell frequencies).



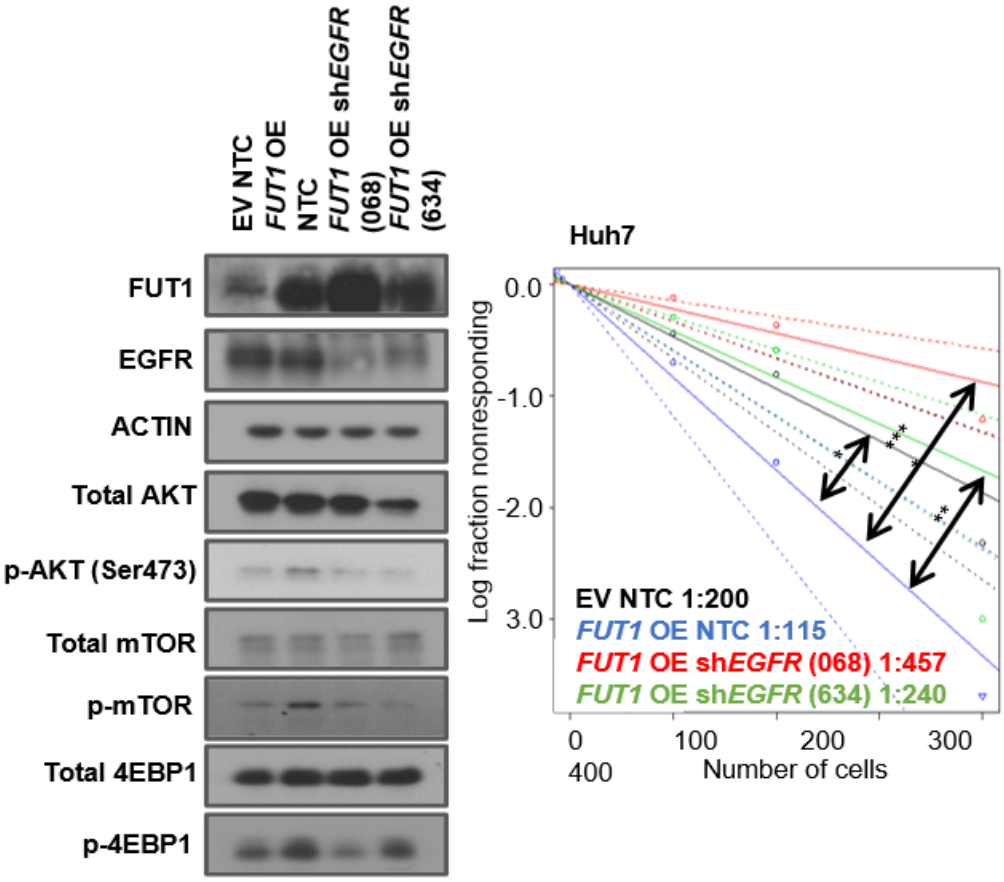
**Supplemental Figure 24. (Related to Figure 8)**

**Effects of *ICAM-1* knockdown on AKT-mTOR-4EBP1 signaling and cancer stemness.** Western blot analysis also found that p-AKT, p-mTOR and p-4EBP1 levels were attenuated upon *ICAM-1* knockdown in *FUT1*-overexpressing HCC Huh7 cells. *In vitro* limiting dilution assays found that the frequency of tumor-initiating cells decreased following knockdown of *ICAM-1* in *FUT1*-overexpressing HCC Huh7 cells. All data shown are representative of three independent experiments. NTC, nontarget control; sh*ICAM-1* clones, 478 and 629; EV, empty vector control; OE, overexpression. * and *** for $p<0.05$ and $p<0.001$, respectively. (Pairwise tests for differences in stem cell frequencies).



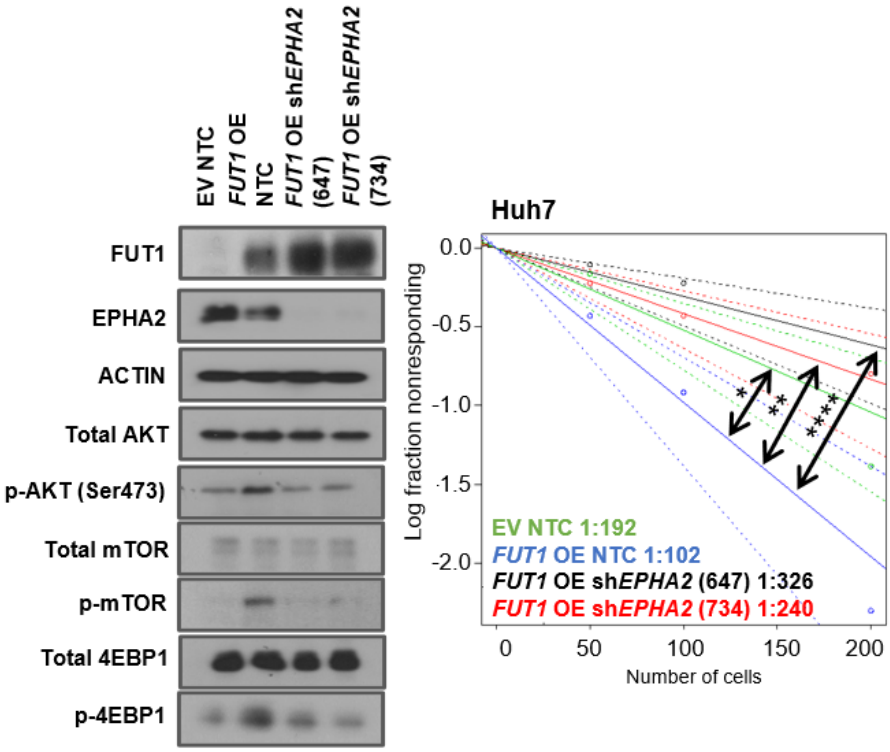
**Supplemental Figure 25. (Related to Figure 8)**

**Effects of *EGFR* knockdown on AKT-mTOR-4EBP1 signaling and cancer stemness.** Western blot analysis also found that the p-AKT, p-mTOR and p-4EBP1 levels were attenuated upon *EGFR* knockdown in the *FUT1*-overexpressing HCC Huh7 cells. *In vitro* limiting dilution assays found that the frequency of tumor-initiating cells decreased following knockdown of *EGFR* in the *FUT1*-overexpressing HCC Huh7 cells. All data shown are representative of three independent experiments. NTC, nontarget control; sh*EGFR* clones 068 and 634; EV, empty vector control; OE, overexpression. *, ** and **** indicate $p<0.05$, $p<0.01$ and $p<0.0001$, respectively. (Pairwise tests for differences in stem cell frequencies).



**Supplemental Figure 26. (Related to Figure 8)**

**Effects of *EPHA2* knockdown on AKT-mTOR-4EBP1 signaling and cancer stemness.** Western blot analysis also found that the p-AKT, p-mTOR and p-4EBP1 levels were attenuated upon *EPHA2* knockdown in the *FUT1*-overexpressing HCC Huh7 cells. *In vitro* limiting dilution assays found that the frequency of tumor-initiating cells decreased following knockdown of *EPHA2* in the *FUT1*-overexpressing HCC Huh7 cells. All data shown are representative of three independent experiments. NTC, nontarget control; sh*EPHA2* clones, 647 and 734; EV, empty vector control; OE, overexpression. *, ** and **** indicate $p<0.05$, $p<0.01$ and $p<0.0001$, respectively. (Pairwise tests for differences in stem cell frequencies).



## SUPPLEMENTAL TABLES

**Supplemental Table 1. *In vivo* limiting dilution assays comparing Huh7 HCC cells treated with high glucose (HG) or low glucose (LG).** *In vivo* limiting dilution assays found that the cells cultured in low glucose displayed an enhanced tumor incidence, expedited tumor latency and a higher frequency of tumor-initiating cells (primary implantation: *n* = 15 per group, secondary implantation: n = 5 per group).

| Cell Number | Latency (Days) | | Tumor Incidence | | Estimated TIC Frequency | |
|---|---|---|---|---|---|---|
| | HG | LG | HG | LG | HG | LG |
| **Primary Implantation** | | | | | | |
| 50,000 | 59 | 51 | 4/15 | 5/15 | 1/111043 | 1/66243 |
| 10,000 | 71 | 54 | 3/15 | 6/15 | | |
| **Secondary Implantation** | | | | | | |
| 50,000 | 52 | 49 | 1/5 | 5/5 | - | - |
| 10,000 | 67 | 49 | 1/5 | 5/5 | | |

HG for high glucose; LG for low glucose; TIC for tumor-initiating cell

**Supplemental Table 2. FUT1 expression in nontumor (NT) liver and HCC clinical samples.** Table shows the number and percentage of cases displaying no, low, medium and high staining intensity of FUT1 in paired NT and HCC samples.

| FUT1 Level | NT Liver | HCC |
|---|---|---|
| **No Signal** | 13 (23.1%) | 1 (1.6%) |
| **Low** | 26 (42.6%) | 8 (13.1%) |
| **Medium** | 20 (32.8%) | 23 (37.7%) |
| **High** | 2 (3.28%) | 29 (47.5%) |

NT for nontumor

**Supplemental Table 3. *In vivo* limiting dilution assays comparing Huh7 HCC cells with empty vector (EV) control or FUT1 overexpressed.** *In vivo* limiting dilution assays found that *FUT1* overexpression enhanced tumor incidence, expedited tumor latency and increased the frequency of tumor-initiating cells. (*n* = 6 per group).

| Cell Number | Latency (Days) | | Tumor Incidence | | Estimated TIC Frequency | |
|---|---|---|---|---|---|---|
| | EV | *FUT1* OE | EV | *FUT1* OE | EV | *FUT1* OE |
| 10,000 | 28 | 20 | 3/6 | 5/6 | 1/16652 | 1/2096 |
| 5,000 | 36 | 20 | 2/6 | 6/6 | | |
| 2,500 | - | 20 | - | 6/6 | | |

EV for empty vector control; OE for overexpression; TIC for tumor-initiating cell

**Supplemental Table 4. 30 common proteins identified by fucosylated peptide profiling.**

| Protein | # of unique peptides in *FUT1* OE | # of unique peptides in EV |
|---|---|---|
| ICAM1 | 21 | 2 |
| HMGB1 | 7 | 6 |
| BASI / CD147 | 6 | 1 |
| PARP1 | 5 | 5 |
| EGFR | 5 | 3 |
| CALR | 4 | 5 |
| LMNA | 4 | 5 |
| ALDOA | 3 | 3 |
| SRSF3 | 3 | 1 |
| ETFA | 3 | 1 |
| EPHA2 | 3 | 1 |
| HMGA2 | 2 | 3 |
| HDGF | 2 | 3 |
| MAP2 | 2 | 2 |
| PEBP1 | 2 | 1 |
| DDX5 | 1 | 2 |
| PTMA | 1 | 3 |
| HBB | 1 | 1 |
| HMGA1 | 1 | 1 |
| CLIC1 | 1 | 1 |
| FKBP4 | 1 | 1 |
| SOAT1 | 1 | 2 |
| AIFM1 | 1 | 2 |
| SRSF1 | 1 | 2 |
| GRN | 1 | 1 |
| CDC42 | 1 | 1 |
| TECR | 1 | 1 |
| SRPK1 | 1 | 1 |
| TOP2A | 1 | 1 |
| PCBP2 | 1 | 1 |

OE for overexpression

**Supplemental Table 5. PI3K-AKT cancer stemness related proteins identified by fucosylated peptide profiling.**

Marked in bold and blue are proteins that are located on the cell surface membrane.

| Protein | # of unique peptides in *FUT1* OE | # of unique peptides in EV |
|---|---|---|
| **ICAM1** | **21** | **2** |
| HMGB1 | 7 | 6 |
| **BASI / CD147** | **6** | **1** |
| PARP1 | 5 | 5 |
| **EGFR** | **5** | **3** |
| CALR | 4 | 5 |
| SRSF3 | 3 | 1 |
| **EPHA2** | **3** | **1** |
| HMGA2 | 2 | 3 |
| HDGF | 2 | 3 |
| PEBP1 | 2 | 1 |
| DDX5 | 1 | 2 |
| PTMA | 1 | 3 |
| HMGA1 | 1 | 1 |
| CLIC1 | 1 | 1 |
| FKBP4 | 1 | 1 |
| AIFM1 | 1 | 2 |
| SRSF1 | 1 | 2 |
| GRN | 1 | 1 |
| CDC42 | 1 | 1 |
| SRPK1 | 1 | 1 |
| TOP2A | 1 | 1 |
| PCBP2 | 1 | 1 |

OE for overexpression

**Supplemental Table 6. Primers used for qPCR.**

| Gene | Primer |
|---|---|
| Human *ATF4* | F (5’- GGCCAAGCACTTCAAACC-3’) |
| | R (5’-GAGAAGGCATCCTCCTTGCT-3’) |
| Human *GRP78* | F (5’-CCCGAGAACACGGTCTTTGA-3’) |
| | R (5’-TTCAACCACCTTGAACGGCA-3’) |
| Human *PERK* | F (5’-CTCAGCGACGCGAGTACC-3’) |
| | R (5’-TGATAATTACTAATGACCTGCCGC-3’) |
| Human *eIF2α* | F (5’-GGGAAGCAAGTCTGGTCTCTG-3’) |
| | R (5’-CCGGCATICTGAGGTATGTGT-3’) |
| Human *FUT1* | F (5’-AGCAACGGCATGGAGTGGTGTA-3’) |
| | R (5’-AAGCCGAAGGTGCCAATGGTCA-3’) |
| Human *FUT2* | F (5’-CTACCACCTGAACGACTGGATG-3’) |
| | R (5’- AGGGTGAACTCCTGGAGGATCT-3’) |
| Human *β-ACTIN* | F (5’- CATCCACGAAACTACCTTCAACTCC-3’) |
| | R (5’- GAGCCGCCGATCCACACG-3’) |
| Human *CA9* | F (5'- CTGGTGACTCTCGGCTACAGCT-3') |
| | R (5'-CTAGGATGTCACCAGCAGCCAGG-3') |
| Mouse *Fut1* | F (5’- GGCAGGTTTGGTAACCAGATG-3’) |
| | R (5’-GAAGGCTTGGCGACCATTG-3’) |
| Mouse *Gapdh* | F (5’AGGTCGGTGTGAACGGATTTG-3’) |
| | R (5’-GTAGACCATGTAGTTGAGGTC-3’) |

F for forward and R for reverse.

**Supplemental Table 7. Antibodies used for immunohistochemistry, western blots and flow cytometry.**

| Name | Supplier | Cat no. | Clone no. |
|---|---|---|---|
| β-actin | Sigma | A5316 | AC-74 |
| α-Tubulin | Sigma | T9026 | DM1A |
| GRP78 | Abcam | ab21685 | Polyclonal |
| Total eIF2α | Cell Signaling Technology | 9722 | Polyclonal |
| p-eIF2α | Cell Signaling Technology | 9721 | Polyclonal |
| Total PERK | Cell Signaling Technology | 5683 | D11A8 |
| p-PERK (Thr980) | Invitrogen | MA5-15033 | G.305.4 |
| ATF4 | Santa Cruz Biotechnology | sc-200 | Polyclonal |
| FUT1 | Abcam | ab198712 | Polyclonal |
| FUT1 | Thermo Fisher Scientific | PA579287 | Polyclonal |
| p-AKT (Ser473) | Cell Signaling Technology | 4060 | Polyclonal |
| p-AKT (Ser473) | Cell Signaling Technology | 9271 | Polyclonal |
| Total AKT | Cell Signaling Technology | 9272 | Polyclonal |
| p-mTOR (Ser2448) | Cell Signaling Technology | 2971 | Polyclonal |
| Total mTOR | Cell Signaling Technology | 2983 | 7C10 |
| p-4EBP1 | Cell Signaling Technology | 2855S | 236B4 |
| Total 4EBP1 | Cell Signaling Technology | 9644S | 53H11 |
| LAMP2 | Abcam | ab25631 | H4B4 |
| CD147 | Abcam | ab11572 | Polyclonal |
| CD147 | Abcam | ab666 | MEM-M6/1 |
| ICAM-1 | Abcam | ab109361 | EPR4776 |
| EGFR | Cell Signaling Technology | 2232S | Polyclonal |
| EPHA2 | Cell Signaling Technology | 6997S | D4A2 |
| HDGF | Abcam | ab128921 | EPR7898 |
| PEBP1 | Invitrogen | 37-2100 | 3E12D7 |
| PARP1 | Cell Signaling Technology | 9542S | Polyclonal |
| HMGB1 | Abcam | ab79823 | EPR3507 |
| CALR | Abcam | ab92516 | EPR3924 |
| HMGA2 | Abcam | ab246513 | EPR23215-46 |
| SRSF3 | Abcam | ab73891 | Polyclonal |
| PE-conjugated CD133 | Miltenyi Biotec | 130-113-108 | AC133 |
| CD133 | Abcam | ab19898 | Polyclonal |
| ALDH1L1 | Abcam | ab177463 | EPR12743(B) |
| Lewis Y | Abcam | ab3359 | F3 |
| PCNA | Abcam | ab18197 | Polyclonal |
| Cleaved caspase 3 | Cell Signaling Technology | 9661 | Polyclonal |

**Supplemental Table 8. Primers used for cloning of constructs used in luciferase reporter assay.**

| Construct | Primers |
|---|---|
| Full-length *FUT1* | F (5'-CCCAAGCTTGGGCCCCAGGTTATTTCAGGAGC-3') |
| | R (5'-CCGCTCGAGCGGTCAGCTGACAGAACACCCTT-3') |
| Δ5' *FUT1* | F (5'-CCCAAGCTTGGGTGAGCAGGGACAGATGAGG-3') |
| | R (5'- CCGCTCGAGCGGTCAGCTGACAGAACACCCTT-3') |
| Δ3' *FUT1* | F (5'-CCCAAGCTTGGGCCCCAGGTTATTTCAGGAGC-3') |
| | R (5'-CCGCTCGAGCGGGGGTCTAGGGGATGAGGGA-3') |

F for forward and R for reverse.