SUPPLEMENTARY MATERIALS for Keysar et al.

SUPPLEMENTARY METHODS

Generation of patient-derived xenografts (PDX)

A clinical pathologist grossed tumors after resection and non-diagnostic, non-necrotic portions were utilized. Tumors were placed in collecting medium consisting of DMEM supplemented with 10% fetal bovine serum (FBS), 200units/mL penicillin, and 200µg/mL streptomycin, and cut into 3x3x3mm pieces. Up to five mice (10 tumors) were implanted per patient case in this initial engraftment phase (F1 generation). The right and left hind flanks were sterilized and small incisions on the right and left hind flank create subcutaneous pockets. Prepared 3x3x3mm tumor pieces were dipped in Matrigel (Corning, Corning, NY) and inserted into the pocket. When tumors reach 1500mm^3 they were passed to a second colony of animals (5-10 animals per case; F2 generation) just as the fresh human tumors were.

Human papilloma virus (HPV) testing

In situ hybridization of HPV low and high-risk types was conducted using the Ventana INFORM HPV II and HPV III automated assays (Ventana Medical Systems, Oro Valley, AZ) on 4 µm-thick paraffin embedded tissue sections.

Fluorescence activated cell sorting (FACS) and flow cytometry

Tumor tissue was finely minced with a scalpel and dissociated in DMEM containing 1mg/ml collagenase IV (Worthington, Lakewood, NJ) at 37°C for 1.5h. Cells were filtered (40µm) and red blood cells were lysed in ACK buffer (Life Technologies, Carlsbad, CA). Staining with Aldefluor (Stem Cell Technologies, Vancouver, Canada) followed the manufacturer's instructions. Briefly, cells were suspended in Aldefluor staining buffer containing Aldefluor reagent (5µl/ml) and incubated at 37^{o}C for 30min. DEAB (N,N-diethylaminobenzaldehyde) was used as a negative control for setting gates. Following incubating with Aldefluor, cells were stained with the following antibodies; IgG controls, 1:10 anti-human CD44 (BD Biosciences, San Jose, CA), 1:100 anti-mouse H-2Kd (BioLegend, San Diego, CA), 1:100 anti-mouse CD31 (BioLegend) 1:100 anti-mouse CD34 (BioLegend), 1:100 anti-mouse CD45 (BioLegend) at the dilution of

100µl/1x10^6 cells. Cell sorting was performed using a MoFlo XDP (Beckman Coulter, Fort Collins, CO) flow cytometric analysis was completed on a CyAn ADP (Beckman Coulter).

Animal anesthesia and pain management

All procedures were approved by the University of Colorado Institutional Animal Care and Use Committee (IACUC). For cell injection and PDX tumor implantation animals were anesthetized with Isoflurane (Induction at 5%, maintained at 1-2%). Prior to tumor irradiation mice were anesthetized using a combination of ketamine (80mg/kg) and xylazine (16mg/kg). In preparation for surgical implantation of PDX tumor tissue, and for 48h following the procedure, animals received buprenorphine injections (1 mg/kg) every 12h.

***In vivo* treatment studies**

PX-866 (Oncothyreon, Seattle, WA) was provided under a Material Transfer Agreement. Docetaxel, LY2109761 cetuximab and ZSTK474 were purchased from Selleckchem, Houston, TX. Animals were irradiated using a RS-2000 irradiator (Rad Source Technologies, Suwanee, GA) as described (1). Mice were anesthetized using a combination of ketamine (80mg/kg) and xylazine (16mg/kg) before being placed under lead shielding leaving only the flank tumors exposed. Animals bearing tumors were irradiated using a 160KeV source, at 25 mAmp, and at a dose rate of 115 cGy/min. Treatments prior to sorting were: control, XRT (1 dose/4Gy), docetaxel (1 dose by intraperitoneal injection, 25mg/kg), cetuximab (2 doses by Intraperitoneal injection on consecutive days, 40mg/kg), ZSTK474 (2 doses by oral gavage on consecutive days, 40mg/kg), or PX-866 (2 doses by oral gavage on consecutive days, 2mg/kg). Tumors were extracted 48h following the initial dose and were immediately processed for cell sorting.

Cell lines and *in vitro* drugs

013C, 036C, and 067C cells were derived from tumor tissue using RM_K media [DMEM:F12 (3:1) with 10% FBS, Insulin (5µg/ml), EGF (10ng/ml), hydrocortisone (0.4µg/ml), transferrin (5µg/ml), penicillin (200units/mL), and streptomycin (200ug/mL)]. LY2109761, erlotinib, 4EGI-1 and AZD6244 were acquired commercially from Selleckchem.

Immunocytochemistry (ICC)

Immunofluorescence was previously described (2). Briefly, cells grown in chamber wells were fixed in 4% paraformaldehyde for 10min at RT then rinsed with PBST 3X for 5min before permeablizing in 0.2% Triton X-100 for 10 min, before blocking in 1% BSA in PBS for 30min at RT with shaking. Cells were incubated with E-cadherin (Cell Signaling, Danvers, MA) for 1.5h at RT at 1:50 dilution. Cells were rinsed in PBST 3X for 5min, once with PBS and mounted with Prolong anti-fade with DAPI (Life Technologies). Images were taken on a Zeiss Axio Observer Z1 inverted microscope using Zeiss software Rel. 4.8 (Zeiss, Oberkochen, Germany).

Tumor sphere assay

Cell lines or PDX-derived CSCs were plated in ultra-low attachment plates at a concentration of $1x10^5$ (6 well plate) or $2.5x10^4$ (24 well plate) per well. Media was supplemented 4, 7 and 10 days following cell seeding. Cells were allowed to form spheres for 10 or 14 days for cell lines or CSCs respectively. Spheres were imaged, counted and measured using a Zeiss Axio Observer Z1 inverted microscope (Zeiss software Rel. 4.8).

***In vitro* Matrigel coated invasion assay**

Matrigel-coated 8µm pore 6 well inserts were purchased from Corning, Corning, NY. Matrigel inserts were hydrated in serum DMEM media for 3h at 37^oC. Cells ($5x10^5$) were added to the insert in DMEM containing 0.5% FBS while DMEM containing 10% FBS was used as a chemoattractant. Plates were incubated for 24h before fixing with 10% formalin (15min) and staining with 0.5% crystal violet (15min). Invasion was quantified as cells/view at 5X magnification for 6 fields/insert. Experimental conditions were run in triplicate and experiments were repeated three times.

MTS terazolium assay

Cells (013C=3,000, 067C=6,000) were plated in 96-well plates and incubated overnight. Drug was added and plates were incubated for 72h. To analyze, 20µl of MTS reagent was added to each well before

plates were incubated for 30min at 37^{o}C. Absorbance was measured using a Synergy 2 microplate reader (Bio-Tek, Winooski, VT).

siRNA experiments

Cells were seeded in 6 well plates and incubated for 24h. Media was replaced with serum free DMEM for 30min prior to transfection with 1µl/ml Dharmafect1 and 50-100nM siRNA (GE Dharmacon, Lafayette, CO). Cells were incubated for 24h before DMEM containing 20% FBS was added and cells were incubated for another 48h-72h.

Gene (cDNA) overexpression

For gene overexpression experiments HEK293T cells were transfected with an empty (control) pMICH-mCherry retroviral vector, or vector containing cDNA for *SOX2,* and the pCL-Ampho packaging plasmid. 013C, 036C, 067C cells and CUHN013 CSCs were transduced with the resulting viral media and cells were selected by mCherry expression (FACS).

***ALDH1A1* promoter luciferase assay**

The *ALDH1A1* promoter luciferase reporter construct (Switchgear Genomics, Menlo Park, CA) containing a DNA fragment ranging 702bp upstream- 366bp downstream of *ALDH1A1*. Positive, negative and *GAPDH* promoter controls were used. Cells were seeded in white walled 96-well plates and transfected per the manufacturer's instructions and incubated for 48h. Luminescence was measured on a Synergy 2 micro-plate reader (Bio-Tek).

Chromatin immunoprecipitation (ChIP) analysis

The Millipore EZ-Chip system was used following the manufacturer's recommended protocol (EMD Millipore, Billerica, MA). *ALDH1A1* primers:

Forward-5'-TGCCCTAGGTGTTACAAATAAGT

Reverse 5'-GCGTGCCTGAGGATGACATTTCT

RNA Immunoprecipitation (RIP) analysis

The Millipore Magna-RIP assay was used following the manufacturer's recommended protocol (EMD Millipore). Reverse transcription was done using *SOX2* primers:

Forward-5'-CGTTCATCGACGAGGCTAAGCG

Reverse-5'-GAGCTGGTCATGGAGTTGTACTGC

RNA isolation and gene expression analysis

RNA was extracted using RNeasy purification kits (Qiagen, Germantown, MD) according to the manufacturer's instructions. RNA concentration and quality was measured using the Nanodrop ND-1000 (Thermo Scientific, Waltham, MA). RNA was reverse-transcribed to cDNA in 20µL reactions using the Verso cDNA Synthesis Kit (Thermo Scientific). Reverse transcription reactions followed the protocol recommendations and were performed using the Verti 96-Well Thermal Cycler (Applied Biosystems, Foster City, CA). TaqMan primer probes (Applied Biosystems), PCR amplification and probe detection were accomplished using the StepOnePlus Real-Time PCR System (Applied Biosystems). All data are representative of experiments performed at least three times in triplicate.

Protein isolation and western blotting

Cell pellets were lysed in 30-100µl RIPA lysis buffer containing 5µl/ml PMSF. Protein conentration was measured using the ELx800 absorbance microplate reader (BioTek) according to the manufacturer's instructions. 30ng of protein was loaded per well into NuPage Novex 4-12% Bis-Tris Midi Gel (Life Technologies), transferred using the iBlot Gel Transfer Stack System (Life Technologies). Primary antibodies and dilutions; 1:2000 Actin (pan) (Cell Signaling), 1:1000 AKT (pan) (Cell Signaling), 1: 500 pAKT (Cell Signaling), 1:1000 AKT1 (Cell Signaling), 1:500 ALDH1A1 (Sigma Aldrich, St. Louis, MO), 1:750 4EBP1 (Cell Signaling), 1:750 p4EBP1 (Cell Signaling), 1:1000 E-cadherin (Cell Signaling), 1:1000 EIF4E (Abcam, Cambridge, United Kingdom), 1:1000 pS6K (Cell Signaling), 1:1000 SMAD2/3 (Cell Signaling), 1:1000 pSMAD2 (Cell Signaling), 1:1000 pSMAD3 (Cell Signaling), 1:500 SOX2 (Cell Signaling). Secondary anti-rabbit IgG was purchased from Jackson ImmunoResearch (West Grove, PA), and used at a 1:50,000 dilution. Signal was visualized using Immobilon Western chemiluminescent HRP substrate (EMD Millipore) on x-ray

film. Quantification of relative protein levels was completed using ImageJ software version 1.5 (National Institutes of Health, imagej.nih.gov).

Immunohistochemistry (IHC)

IHC analyses were performed as previously described (3). Slides were de-paraffinized and re-hydrated in graded concentrations of alcohol by standard techniques before antigen retrieval in citrate buffer pH 6.0 (Dako, Carpinteria, CA) at 105^{o}F for 20min. All staining were done in a Dako Autostainer and slides were incubated in 3% H_2O_2 for 10min, followed by primary antibody (60min at RT). Primary antibodies and dilutions; 1:1000 ALDH1A1 (Sigma Aldrich), 1:1000 p4EBP1 (Cell Signaling), 1:100 E-cadherin (Dako), 1:100 EGFR (Cell Signaling), 1:100 pS6K (Cell Signaling), 1:200 SOX2 (Cell Signaling). Staining was developed using the following conditions: EnVision + Dual Link System HRP (Dako) for 30min and substrate-chromogen (DAB+) Solution (Dako) for 7min. Slides were then counterstain with Automated Hematoxylin (Dako) for 5min.

mRNA Sequencing and Bioinformatics

Processed RNA was sent to the UCCC Genomics and Microarray Core who performed cDNA library generation, Illumina HiSeq RNAseq (Illumina, San Diego, CA) with a read length of 100bp, and generation of FASTQ files. On average, we obtained about 65 million reads per sample. Reads were checked for quality using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), adapter sequences were removed using cutadapt (4), and finally poor quality ends were removed using trimmomatic (5). Reads had average unique mapping of 71% to the hg19 reference genome using the tuxedo workflow as previously described (6). Results were analyzed using the cummeRbund package to give gene FPKM estimates for each sample as well as pairwise significance values between sample groups. To determine differential pathway enrichment, FPKM values were analyzed by the Gene Set Enrichment Analysis (GSEA) package (7). Genes were ranked based on a measure of each gene's signal to noise ratio with respect to the two groups. Enrichment scores (ES) were calculated by increasing a running-sum statistic when a gene belongs to the set and decreasing it when the gene does not. GSEA estimates the statistical significance of the ES by a two-sided modified Kolmogorov-Smirnov permutation test. *P* less than .05 was considered significant.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Serial implantation of HNSCC CSCs and an *in vivo* re-implantation assay to measure the sensitivity of CSCs to a panel of therapies. (**A**) Hematoxylin eosin staining of tumors derived from CSCs ($ALDH^+CD44^{high}$) and single marker positive cells ($ALDH^+CD44^{low}$ and $ALDH^-CD44^{high}$) and marker negative cells ($ALDH^-CD44^{low}$). Scale bars=100 µm. (**B**) Tumor growth rates for initial and serial implantation of highly tumorigenic CSC and less tumorigenic populations. Less tumorigenic single positive populations formed tumors that grew slower than CSC derived tumors. (**C**) Sorted CSC and non-CSC population comparison and tumor recapitulation mRNA-sequencing experimental designs. (**D**) Non-CSC $ALDH^-CD44^{high}$ (CUHN014) cells are less tumorigenic and more sensitive to therapy than the $ALDH^+CD44^{hgh}$ CSC population as shown by percent survival (death=tumors>500 mm^3). $ALDH^+CD44^{high}$ (CSCs) were not implanted at 1,000 cells/injection were not included because 100 cells readily generated tumors. (**E**) Tumor growth of pretreated (control, docetaxel, cetuximab and ZSTK474) CSC populations (CUHN014=10 $ALDH^+CD44^{high}$ cells, CUHN022=10,000 $ALDH^+$ cells, CUHN013=10,000 $ALDH^+CD44^{high}$ cells). Open data points represent individual tumor measurements, while closed points and lines represent the average of all tumors for each treatment. Numbers represent tumor take rate (tumor volume>500mm^3). Tumor growth was only somewhat inhibited by docetaxel or cetuximab. ALDH=aldehyde dehydrogenase, CSC=cancer stem cell, HNSCC=head and neck squamous cell carcinoma.

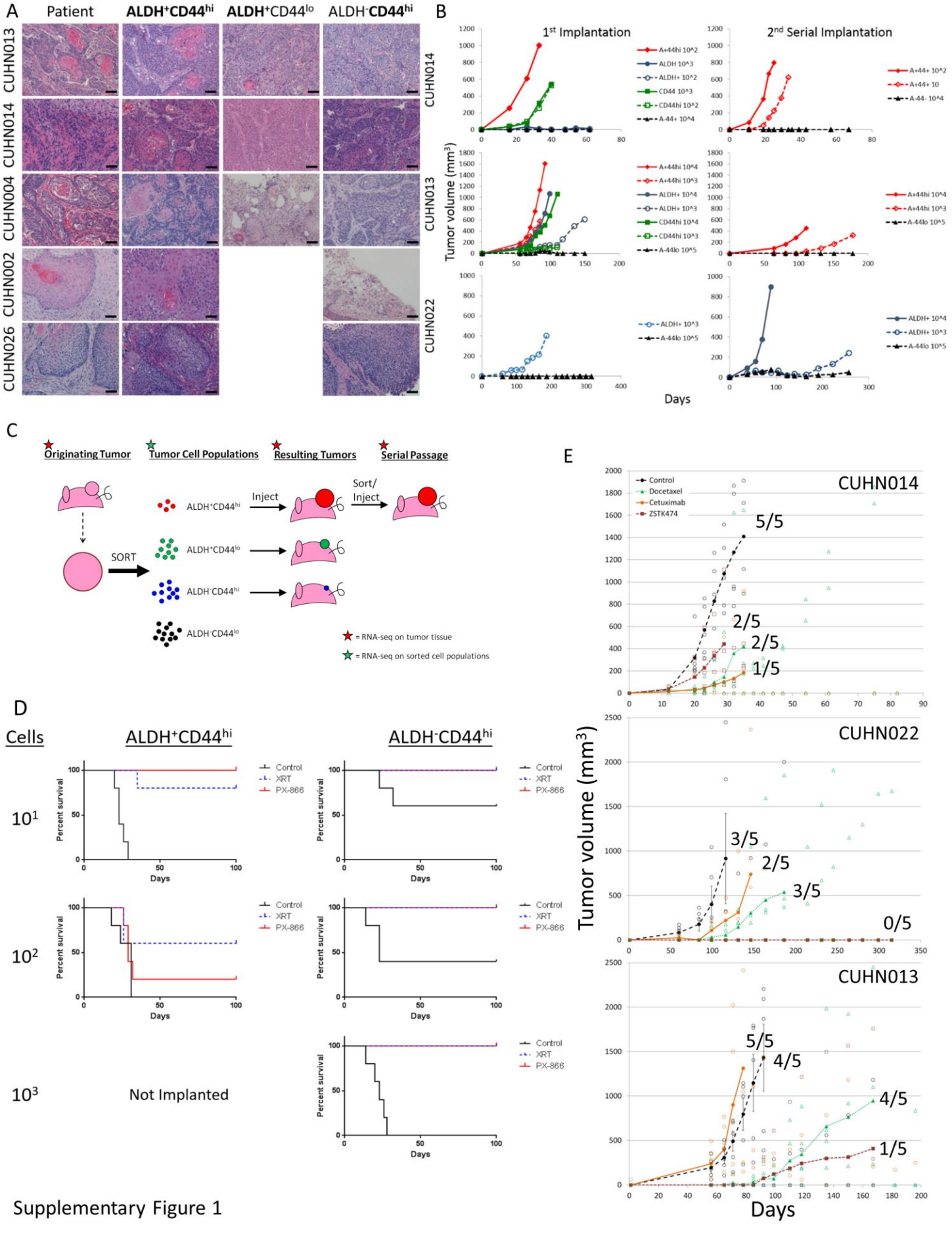
Supplementary Figure 2. Role of PI3K/mTOR signaling in the regulation of SOX2 and the ALDH+ phenotype. (**A**) CUHN013 patient tumor-derived cell line (013C) faithfully recapitulates tumor morphology when compared to the PDX from the same patient. (**B**) PI3K inhibition decreases *ALDH1A1* expression in monolayer cells. (**C**) ZSTK474 treatment decreased pAKT and SOX2 protein levels in 013C cells. (**D**) Inhibition of EGFR (erlotinib) or MEK/ERK (AZD6244) did not reduce *ALDH1A1* levels in 013C cells. (**E**) Representative diagram of the SOX2/OCT4 response elements (SORE6) reporter system [4]. (**F**) Spheroids generated from tumor derived CSCs have large populations of SORE6-$mCherry^+$ cells. (**G**) PI3K inhibition decreases the SORE6-GFP^+ population in 013C sphere cultures. (**H**) Relative mRNA levels after siRNA silencing of PI3K/mTOR signaling genes (AKT1, S6K, EIF4E) in 013C, 067C cells. (**I**) Silencing of AKT1 and EIF4E decreased SOX2 protein

levels in 013C and 067C cells. Quantification of SOX2 protein, relative to Actin loading control, was completed using ImageJ software. Results are presented as mean±SD of four (013C) and three (067C) independent experiments. Statistical significance was calculated by two-tailed student's t test (* *P*<.05, ** *P*<.01). CSC=cancer stem cell, mTOR=mechanistic target of rapamycin, PDX=patient-derived xenograft.

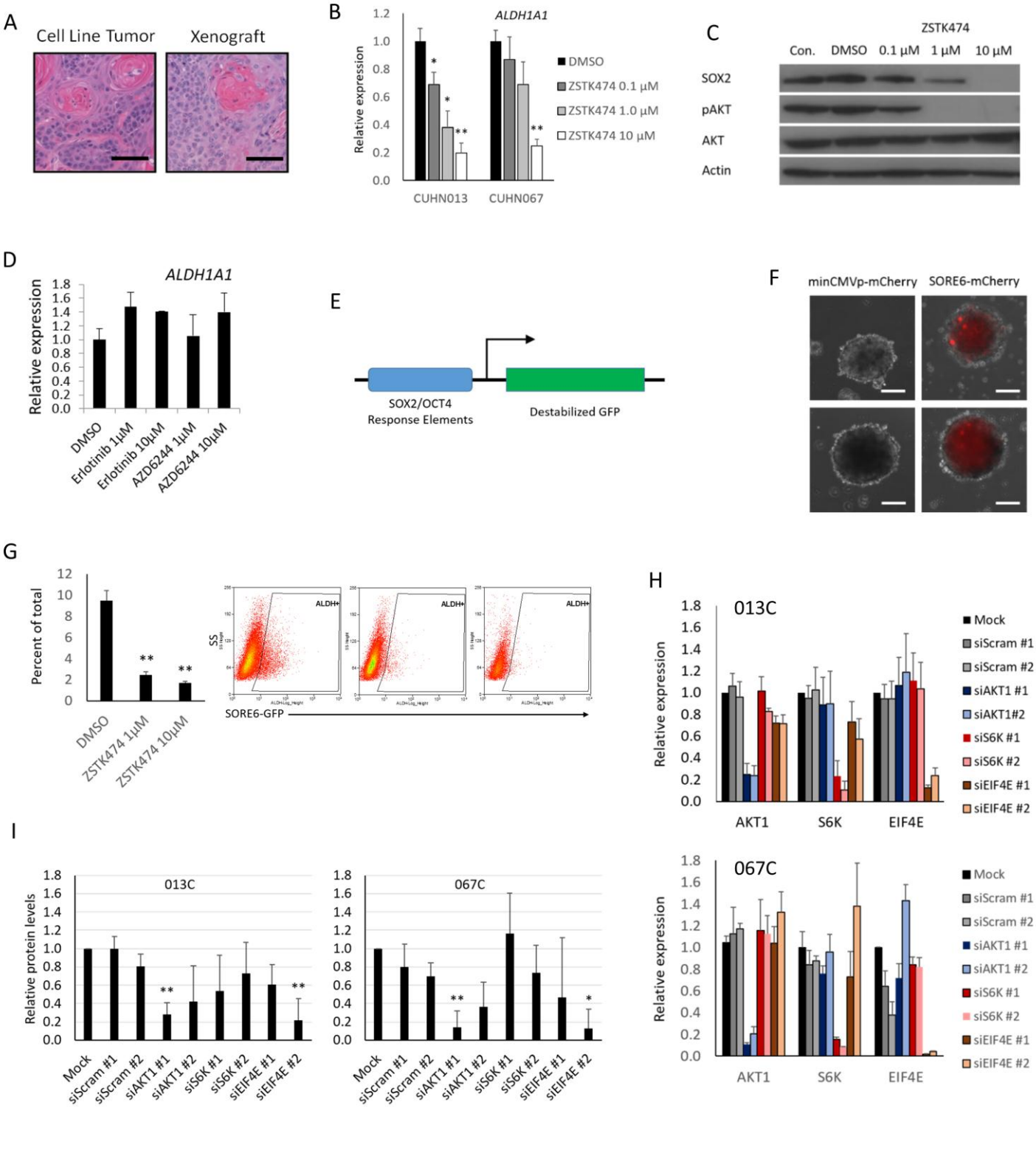
Supplementary Figure 3. Role of SOX2 in expression of *ALDH1A1* and the ALDH+ population. (**A**) Treating 013C cells with 4EGI-1 (48 hours) decreased the $ALDH^+$ population. (**B**) PI3K inhibition decreased levels of phosphorylated (Ser65) 4EBP1 in 013C and 067C cells. (**C**) Fold enrichment of SOX2 mRNA directly bound to immunoprecipitated EIF4E in 013C cells measured by real-time PCR. Cells exogenously expressing SOX2 had increased SOX2 mRNA bound to immunoprecipitated EIF4E. (**D**) Exogenous expression of *SOX2* increases *ALDH1A1* mRNA levels in 013C and 067C cell lines. (**E**) IHC staining demonstrates increased SOX2 and ALDH1A1 levels in 013C cells exogenously expressing SOX. (**F**) PI3K inhibition (ZSTK474) or exogenous SOX2 expression do not alter CD44 cell surface protein levels in 013C and 067C cells measured by flow cytometry. Results are presented as mean±SD of three independent experiments. Statistical significance was calculated by two-tailed student's t test (* *P*<.05, ** *P*<.01). ALDH=aldehyde dehydrogenase, IHC=immunohistochemistry, PI3K=phosphoinositide 3-kinase.

Supplementary Figure 4. Role of cancer stem cell pathways in cell morphology and tumor growth. (**A**) Inhibition of TGF-β with LY2109761 decreased pSMAD3, but not pSMAD2, in 013C and 067C cell lines. (**B**) LY2109761 treatment did not alter expression of *SNAI1* or *CDH1* in CUHN013 tumor derived CSCs grown as spheres. (**C**) TGF-β1 ligand stimulated expression of *SNAI1*, but did not suppress *CDH1*, in 013C and 067C cell lines but this effect was blocked by LY2109761 treatment. (**D**) Stimulation of serum starved HNSCC cells with TGF-β1 ligand or inhibition with LY2109761 did not alter the invasiveness of 013C cells. Results are presented as mean±SD of two independent experiments in triplicate. Statistical significance was calculated by two-tailed student's t test (* *P*<.05, ** *P*<.01). (**E**) CUHN013 CSCs easily form spheres under low attachment growth conditions while marker negative tumor cells do not. Results are presented as mean±SD of 2 independent experiments with four replicates each. Statistical significance was calculated by two-tailed student's t test (* *P*<.05, ** *P*<.01). (**F**) Resulting tumors from implantation of 013C-parental, 013C-empty

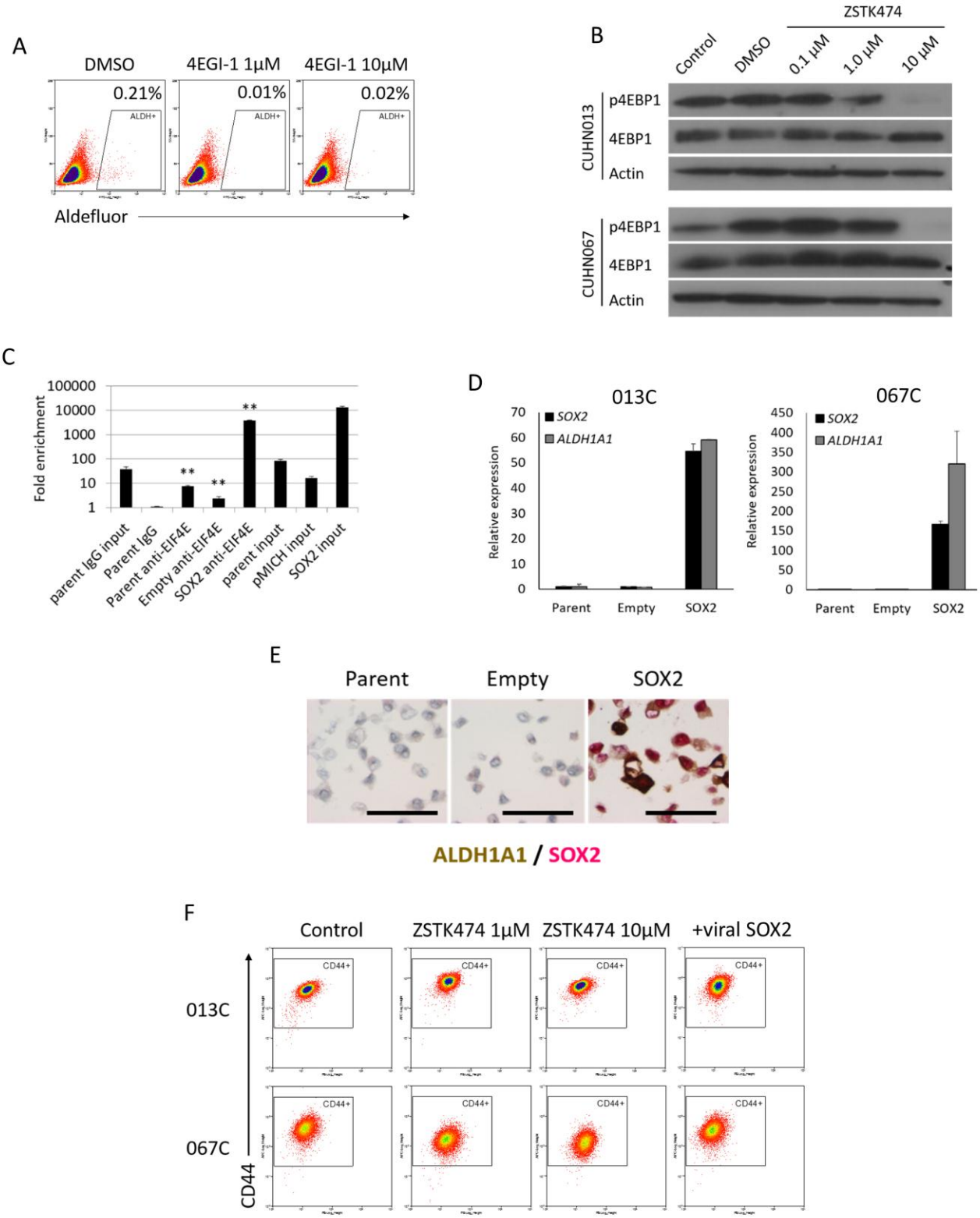
vector and 013C-SOX2 cells in nude mice. ALDH=aldehyde dehydrogenase, CSC=cancer stem cell, HNSCC=head and neck squamous cell carcinoma, IHC=immunohistochemistry, PI3K=phosphoinositide 3-kinase, TGF-β=transforming growth factor beta.



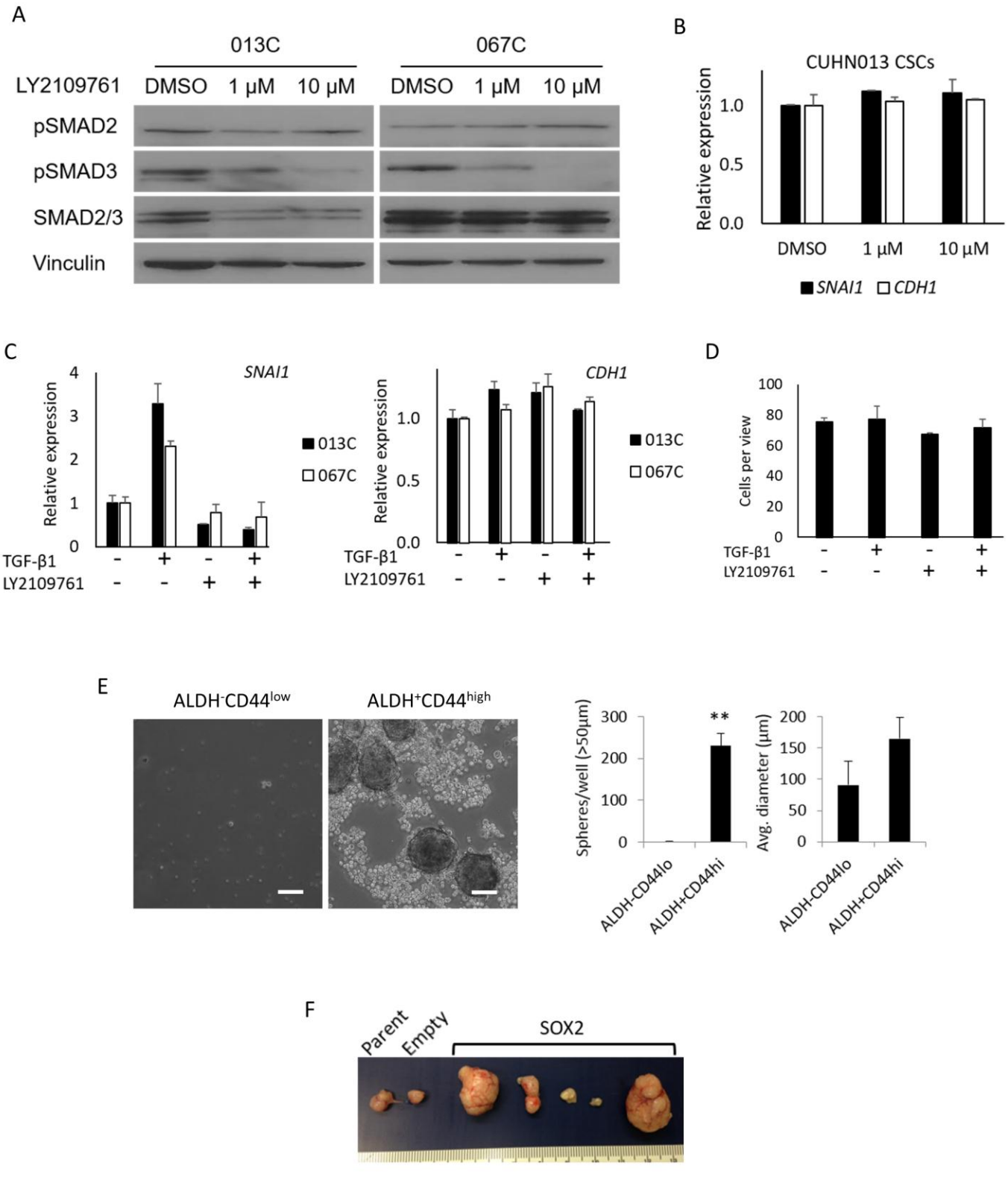
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4

Supplementary Table 1. Tumor location, HPV status, *TP53* status, *PIK3CA* status and smoking history for the 10 patients that originated the PDX models used in CSC studies and the 013C, 036C and 067C cell lines used in the *in vitro* experiments*

Case	Site	HPV	TP53	PIK3CA	Smoking
CUHN002	Tongue	-	Mutant	WT	Yes
CUHN004	Floor of Mouth	-	Mutant	WT	Yes
CUHN013	Floor of Mouth	-	Mutant	Amplified	Yes
CUHN014	Base of Tongue	+	WT	Amplified	Yes
CUHN022	Tonsil	+	WT	WT	No
CUHN026	Tongue	-	Mutant	Mutant	Yes
CUHN036	Pharynx	-	Mutant	Amplified	Yes
CUHN047	Tonsil	+	WT	Amplified	No
CUHN049	Tongue	-	Mutant	Amplified	Yes
CUHN067	Base of Tongue	-	Mutant	WT	Yes
CUHN070	Pharynx	-	Mutant	WT	Yes
CUHN098	Tongue	+	WT	WT	No

*CSC=cancer stem cell, HPV=human papilloma virus, PDX=patient-derived xenograft, WT=wild type.

Supplementary Table 2. All significant GSEA pathways from the analysis used to generate the selected genes list for Figure 2A.

NAME	*P*-value*	FDR *q*-value
HALLMARK PATHWAYS		
HALLMARK_MYC_TARGETS_V1	<0.001	<0.001
HALLMARK_G2M_CHECKPOINT	<0.001	<0.001
HALLMARK_MTORC1_SIGNALING	<0.001	0.001
HALLMARK_TNFA_SIGNALING_VIA_NFKB	<0.001	0.009
HALLMARK_E2F_TARGETS	<0.001	0.01
HALLMARK_OXIDATIVE_PHOSPHORYLATION	<0.001	0.01
HALLMARK_PI3K_AKT_MTOR_SIGNALING	<0.001	0.03
HALLMARK_TGF_BETA_SIGNALING	0.005	0.03
HALLMARK_ANDROGEN_RESPONSE	<0.001	0.03
KEGG PATHWAYS		
KEGG_SPLICEOSOME	<0.001	<0.001
KEGG_UBIQUITIN_MEDIATED_PROTEOLYSIS	<0.001	0.001
KEGG_PROTEASOME	<0.001	0.002
KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS	<0.001	0.003
KEGG_RENAL_CELL_CARCINOMA	<0.001	0.004
KEGG_REGULATION_OF_AUTOPHAGY	0.002	0.006
KEGG_RNA_DEGRADATION	<0.001	0.007
KEGG_BASAL_TRANSCRIPTION_FACTORS	0.004	0.008
KEGG_RIBOSOME	0.002	0.011
KEGG_PROTEIN_EXPORT	0.005	0.02
KEGG_PARKINSONS_DISEASE	<0.001	0.02
KEGG_ENDOCYTOSIS	<0.001	0.03
KEGG_NEUROTROPHIN_SIGNALING_PATHWAY	<0.001	0.04
KEGG_CYSTEINE_AND_METHIONINE_METABOLISM	0.032	0.05
NCI PATHWAY INTERACTION DATABASE		
PID_FOXO_PATHWAY	<0.001	<0.001
PID_P38_ALPHA_BETA_DOWNSTREAM_PATHWAY	<0.001	<0.001
PID_TELOMERASE_PATHWAY	<0.001	0.002
PID_HIF1_TFPATHWAY	<0.001	0.002
PID_ECADHERIN_NASCENT_AJ_PATHWAY	<0.001	0.003
PID_AR_TF_PATHWAY	<0.001	0.003
PID_AR_PATHWAY	0.003	0.005
PID_MYC_ACTIV_PATHWAY	<0.001	0.006
PID_P53_REGULATION_PATHWAY	<0.001	0.007
PID_CDC42_PATHWAY	0.003	0.008
PID_MTOR_4PATHWAY	<0.001	0.009
PID_NFKAPPAB_CANONICAL_PATHWAY	0.005	0.01
PID_VEGFR1_2_PATHWAY	<0.001	0.01
PID_ERBB1_DOWNSTREAM_PATHWAY	<0.001	0.01
PID_HES_HEY_PATHWAY	0.01	0.01
PID_ECADHERIN_STABILIZATION_PATHWAY	0.01	0.02
PID_PS1_PATHWAY	<0.001	0.02
PID_HDAC_CLASSI_PATHWAY	0.005	0.02
PID_SMAD2_3NUCLEAR_PATHWAY	0.008	0.02
PID_TGFBR_PATHWAY	0.01	0.03

PID_RHOA_PATHWAY	0.01	0.03
PID_BCR_5PATHWAY	0.008	0.03
PID_IFNG_PATHWAY	0.02	0.03
PID_NECTIN_PATHWAY	0.02	0.03
PID_FAK_PATHWAY	0.01	0.03
PID_PDGFRB_PATHWAY	<0.001	0.04
PID_MET_PATHWAY	0.01	0.04
PID_ILK_PATHWAY	0.03	0.05
REACTOME PATHWAYS		
REACTOME_CIRCADIAN_CLOCK	<0.001	<0.001
REACTOME_MRNA_3_END_PROCESSING	<0.001	<0.001
REACTOME_MRNA_PROCESSING	<0.001	<0.001
REACTOME_MRNA_SPLICING	<0.001	<0.001
REACTOME_PROCESSING_OF_CAPPED_INTRON_CONTAINING_PRE_MRNA	<0.001	<0.001
REACTOME_RNA_POL_II_TRANSCRIPTION	<0.001	<0.001
REACTOME_TRANSCRIPTION	<0.001	<0.001
REACTOME_APC_C_CDH1_MEDIATED_DEGRADATION_OF_CDC20_AND_OTHER_APC_C_CDH1_TARGETED_PROTEINS_IN_LATE_MITOSIS_EARLY_G1	<0.001	<0.001
REACTOME_REGULATION_OF_MRNA_STABILITY_BY_PROTEINS_THAT_BIND_AU_RICH_ELEMENTS	<0.001	<0.001
REACTOME_SCFSKP2_MEDIATED_DEGRADATION_OF_P27_P21	<0.001	<0.001
REACTOME_REGULATION_OF_MITOTIC_CELL_CYCLE	<0.001	<0.001
REACTOME_AUTODEGRADATION_OF_CDH1_BY_CDH1_APC_C	<0.001	<0.001
REACTOME_CLEAVAGE_OF_GROWING_TRANSCRIPT_IN_THE_TERMINATION_REGION_	<0.001	<0.001
REACTOME_CYCLIN_E_ASSOCIATED_EVENTS_DURING_G1_S_TRANSITION_	<0.001	<0.001
REACTOME_BMAL1_CLOCK_NPAS2_ACTIVATES_CIRCADIAN_EXPRESSION	<0.001	<0.001
REACTOME_APC_C_CDC20_MEDIATED_DEGRADATION_OF_MITOTIC_PROTEINS	<0.001	<0.001
REACTOME_ANTIGEN_PROCESSING_UBIQUITINATION_PROTEASOME_DEGRADATION	<0.001	<0.001
REACTOME_ACTIVATION_OF_NF_KAPPAB_IN_B_CELLS	<0.001	<0.001
REACTOME_MEIOSIS	<0.001	<0.001
REACTOME_HIV_LIFE_CYCLE	<0.001	<0.001
REACTOME_REGULATION_OF_ORNITHINE_DECARBOXYLASE_ODC	<0.001	<0.001
REACTOME_P53_DEPENDENT_G1_DNA_DAMAGE_RESPONSE	<0.001	<0.001
REACTOME_CYTOSOLIC_TRNA_AMINOACYLATION	<0.001	<0.001
REACTOME_TRANSLATION	<0.001	<0.001
REACTOME_LATE_PHASE_OF_HIV_LIFE_CYCLE	<0.001	<0.001
REACTOME_PACKAGING_OF_TELOMERE_ENDS	<0.001	<0.001
REACTOME_RNA_POL_I_PROMOTER_OPENING	<0.001	<0.001
REACTOME_AUTODEGRADATION_OF_THE_E3_UBIQUITIN_LIGASE_COP1	<0.001	<0.001
REACTOME_SIGNALING_BY_WNT	<0.001	<0.001
REACTOME_TRANSPORT_OF_MATURE_TRANSCRIPT_TO_CYTOPLASM	<0.001	<0.001
REACTOME_MEIOTIC_RECOMBINATION	<0.001	<0.001
REACTOME_DESTABILIZATION_OF_MRNA_BY_AUF1_HNRNP_D0	<0.001	<0.001
REACTOME_INFLUENZA_LIFE_CYCLE	<0.001	<0.001
REACTOME_METABOLISM_OF_MRNA	<0.001	<0.001
REACTOME_METABOLISM_OF_RNA	<0.001	<0.001
REACTOME_HIV_INFECTION	<0.001	<0.001
REACTOME_UNFOLDED_PROTEIN_RESPONSE	<0.001	<0.001
REACTOME_RNA_POL_I_TRANSCRIPTION	<0.001	<0.001
REACTOME_SIGNALING_BY_HIPPO	<0.001	<0.001
REACTOME_SRP_DEPENDENT_COTRANSLATIONAL_PROTEIN_TARGETING_TO_MEMBRANE	<0.001	<0.001

REACTOME_SCF_BETA_TRCP_MEDIATED_DEGRADATION_OF_EMI1	<0.001	<0.001
REACTOME_RNA_POL_II_PRE_TRANSCRIPTION_EVENTS	<0.001	<0.001
REACTOME_MEIOTIC_SYNAPSIS	<0.001	<0.001
REACTOME_CLASS_I_MHC_MEDIATED_ANTIGEN_PROCESSING_PRESENTATION	<0.001	<0.001
REACTOME_ER_PHAGOSOME_PATHWAY	<0.001	0.001
REACTOME_MRNA_SPLICING_MINOR_PATHWAY	<0.001	0.001
REACTOME_RORA_ACTIVATES_CIRCADIAN_EXPRESSION	<0.001	0.001
REACTOME_PERK_REGULATED_GENE_EXPRESSION	<0.001	0.001
REACTOME_CROSS_PRESENTATION_OF_SOLUBLE_EXOGENOUS_ANTIGENS_ENDOSOMES	<0.001	0.001
REACTOME_DIABETES_PATHWAYS	<0.001	0.002
REACTOME_VIF_MEDIATED_DEGRADATION_OF_APOBEC3G	0.002	0.002
REACTOME_FORMATION_OF_RNA_POL_II_ELONGATION_COMPLEX_	<0.001	0.002
REACTOME_CIRCADIAN_REPRESSION_OF_EXPRESSION_BY_REV_ERBA	<0.001	0.002
REACTOME_INFLUENZA_VIRAL_RNA_TRANSCRIPTION_AND_REPLICATION	<0.001	0.002
REACTOME_TRANSCRIPTIONAL_ACTIVITY_OF_SMAD2_SMAD3_SMAD4_HETEROTRIMER	<0.001	0.002
REACTOME_ACTIVATION_OF_GENES_BY_ATF4	0.005	0.003
REACTOME_NONSENSE_MEDIATED_DECAY_ENHANCED_BY_THE_EXON_JUNCTION_COMPLEX	<0.001	0.003
REACTOME_RNA_POL_I_RNA_POL_III_AND_MITOCHONDRIAL_TRANSCRIPTION	<0.001	0.003
REACTOME_MRNA_CAPPING	<0.001	0.003
REACTOME_3_UTR_MEDIATED_TRANSLATIONAL_REGULATION	<0.001	0.003
REACTOME_P53_INDEPENDENT_G1_S_DNA_DAMAGE_CHECKPOINT	<0.001	0.033
REACTOME_SIGNALING_BY_TGF_BETA_RECEPTOR_COMPLEX	<0.001	0.003
REACTOME_RNA_POL_II_TRANSCRIPTION_PRE_INITIATION_AND_PROMOTER_OPENING	<0.001	0.004
REACTOME_AMYLOIDS	<0.001	0.005
REACTOME_SMAD2_SMAD3_SMAD4_HETEROTRIMER_REGULATES_TRANSCRIPTION	0.002	0.005
REACTOME_EGFR_DOWNREGULATION	<0.001	0.005
REACTOME_ELONGATION_ARREST_AND_RECOVERY	0.005	0.007
REACTOME_CDK_MEDIATED_PHOSPHORYLATION_AND_REMOVAL_OF_CDC6	<0.001	0.007
REACTOME_FORMATION_OF_THE_TERNARY_COMPLEX_AND_SUBSEQUENTLY_THE_43S_COMPLEX	<0.001	0.007
REACTOME_FORMATION_OF_THE_HIV1_EARLY_ELONGATION_COMPLEX	<0.001	0.007
REACTOME_CDT1_ASSOCIATION_WITH_THE_CDC6_ORC_ORIGIN_COMPLEX	<0.001	0.008
REACTOME_ACTIVATION_OF_THE_MRNA_UPON_BINDING_OF_THE_CAP_BINDING_COMPLEX_AND _EIFS_AND_SUBSEQUENT_BINDING_TO_43S	0.003	0.008
REACTOME_PEPTIDE_CHAIN_ELONGATION	0.003	0.009
REACTOME_DOWNSTREAM_SIGNALING_EVENTS_OF_B_CELL_RECEPTOR_BCR	<0.001	0.01
REACTOME_SIGNALLING_TO_ERKS	0.007	0.01
REACTOME_PROCESSING_OF_CAPPED_INTRONLESS_PRE_MRNA	0.005	0.01
REACTOME_G1_PHASE	0.01	0.02
REACTOME_ABORTIVE_ELONGATION_OF_HIV1_TRANSCRIPT_IN_THE_ABSENCE_OF_TAT	0.02	0.02
REACTOME_ANTIVIRAL_MECHANISM_BY_IFN_STIMULATED_GENES	0.003	0.02
REACTOME_DOWNREGULATION_OF_SMAD2_3_SMAD4_TRANSCRIPTIONAL_ACTIVITY	0.02	0.02
REACTOME_HOST_INTERACTIONS_OF_HIV_FACTORS	<0.001	0.02
REACTOME_APC_C_CDC20_MEDIATED_DEGRADATION_OF_CYCLIN_B	0.02	0.02
REACTOME_MITOTIC_G2_G2_M_PHASES	0.003	0.02
REACTOME_SIGNALING_BY_FGFR_MUTANTS	0.02	0.02
REACTOME_DOWNREGULATION_OF_TGF_BETA_RECEPTOR_SIGNALING	0.01	0.02
REACTOME_APC_CDC20_MEDIATED_DEGRADATION_OF_NEK2A	0.02	0.03
REACTOME_SIGNALING_BY_THE_B_CELL_RECEPTOR_BCR	0.003	0.03
REACTOME_PPARA_ACTIVATES_GENE_EXPRESSION	<0.001	0.03
REACTOME_INHIBITION_OF_THE_PROTEOLYTIC_ACTIVITY_OF_APC_C_REQUIRED_FOR_THE_ONSET _OF_ANAPHASE_BY_MITOTIC_SPINDLE_CHECKPOINT_COMPONENTS	0.03	0.03

REACTOME_CELL_CYCLE	<0.001	0.03
REACTOME_APOPTOSIS	<0.001	0.03
REACTOME_REGULATORY_RNA_PATHWAYS	0.02	0.03
REACTOME_TRANSPORT_OF_MATURE_MRNA_DERIVED_FROM_AN_INTRONLESS_TRANSCRIPT	0.02	0.03
REACTOME_REGULATION_OF_APOPTOSIS	0.008	0.04
REACTOME_DEADENYLATION_DEPENDENT_MRNA_DECAY	0.01	0.04
REACTOME_PRE_NOTCH_TRANSCRIPTION_AND_TRANSLATION	0.03	0.04
REACTOME_DEPOSITION_OF_NEW_CENPA_CONTAINING_NUCLEOSOMES_AT_THE_CENTROMERE	0.02	0.04
REACTOME_SPHINGOLIPID_DE_NOVO_BIOSYNTHESIS	0.02	0.04
REACTOME_APOPTOTIC_EXECUTION_PHASE	0.016	0.04
REACTOME_ACTIVATION_OF_CHAPERONE_GENES_BY_XBP1S	0.02	0.05
REACTOME_ORC1_REMOVAL_FROM_CHROMATIN	0.003	0.05
REACTOME_TGF_BETA_RECEPTOR_SIGNALING_ACTIVATES_SMADS	0.04	0.05
REACTOME_YAP1_AND_WWTR1_TAZ_STIMULATED_GENE_EXPRESSION	0.04	0.05
REACTOME_GOLGI_ASSOCIATED_VESICLE_BIOGENESIS	0.02	0.05
REACTOME_PRE_NOTCH_EXPRESSION_AND_PROCESSING	0.03	0.05
REACTOME_SULFUR_AMINO_ACID_METABOLISM	0.05	0.05

*Statistical significance of GSEA data was calculated by a two-sided modified Kolmogorov-Smirnov permutation test.

Supplementary Table 3. Significant Hallmark Pathways determined by Gene Set Enrichment Analysis (GSEA) comparing HPV-negative (CUHN004, CUHN013) and HPV-positive (CUHN014, CUHN022) CSC populations

HALLMARK PATHWAYS	***P*-value***	**q-value**
TNF-alpha signaling via NF-κβ	<0.001	<0.001
Epithelial to mesenchymal transition	<0.001	<0.001
Hypoxia	<0.001	<0.001
P53 Pathway	<0.001	<0.001
TGF-beta signaling	<0.001	<0.001
DNA Repair	<0.001	0.002
UV Response	<0.001	0.002
Coagulation	<0.001	0.002
Apoptosis	<0.001	0.002
Apical junction	<0.001	0.002
Glycolysis	<0.001	0.003
Oxidative phosphorylation	<0.001	0.006
PI3K/AKT/mTOR signaling	<0.001	0.008
IL2/STAT5 signaling	<0.001	0.02
Angiogenesis	0.03	0.03
Adipogenesis	<0.001	0.03
Unfolded protein response	<0.001	0.03
Myogenesis	0.003	0.03
Wnt/Beta-catenin signaling	0.03	0.04
mTORC1 signaling	0.006	0.04
Notch signaling	0.04	0.04
Complement	0.003	0.04
Androgen response	0.01	0.05
Kras signaling up	0.009	0.05

*Statistical significance of GSEA data was calculated by a two-sided modified Kolmogorov-Smirnov permutation test.

Supplementary Table 4. Tumor take rate of CUHN013, CUHN014 and CUHN022 CSC and non-CSC populations that were pretreated *in vivo* (cetuximab, docetaxel, XRT, PX-866 or ZSTK474) before resecting tumors and cell sorting*

Population	Control	Cetuximab	Docetaxel	XRT	PX-866	ZSTK474
CUHN013						
Negative (10^5)	0/10 (0.0%)	0/10 (0.0%)	0/10 (0.0%)	0/10 (0.0%)	0/10 (0.0%)	0/10 (0.0%)
$CD44^{high}$ (10^3)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)
$CD44^{high}$ (10^4)	3/5 (60.0%)	3/5 (60.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)
$ALDH^+$ (10^3)	1/5 (20.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)
$ALDH^+$ (10^4)	4/5 (80.0%)	0/5 (0.0%)	1/5 (20.0%)	0/5 (0.0%)	2/5 (40.0%)	2/5 (40.0%)
$ALDH^+$ $CD44^{high}$ (10^3)	1/5 (20.0%)	1/5 (20.0%)	1/5 (20.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)
$ALDH^+$ $CD44^{high}$ (10^4)	5/5 (100.0%)	4/5 (80.0%)	4/5 (80.0%)	4/5 (80.0%)	3/5 (60.0%)	1/5 (20.0%)
CUHN022						
Negative (10^5)	0/10 (0.0%)	0/10 (0.0%)	0/10 (0.0%)	0/10 (0.0%)	0/10 (0.0%)	0/10 (0.0%)
$ALDH^+$ (103)	1/5 (20.0%)	2/5 (40.0%)	1/5 (20.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)
$ALDH^+$ (104)	3/5 (60.0%)	2/5 (40.0%)	3/5 (60.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)
CUHN014						
Negative (10^5)	2/5 (40.0%)	1/5 (20.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)
$CD44^{high}$ (10^1)	2/5 (40.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)	0/5 (0.0%)
$CD44^{high}$ (10^2)	3/5 (60.0%)	0/5 (0.0%)	1/5 (20.0%)	0/5 (0.0%)	0/5 (0.0%)	3/5 (60.0%)
$CD44^{high}$ (10^3)	5/5 (100.0%)	2/5 (40.0%)	1/5 (20.0%)	0/5 (0.0%)	0/5 (0.0%)	3/5 (60.0%)
$ALDH^+$ $CD44^{high}$ (10^1)	5/5 (100.0%)	2/5 (40.0%)	5/5 (100.0%)	1/5 (20.0%)	0/5 (0.0%)	1/5 (20.0%)
$ALDH^+$ $CD44^{high}$ (10^2)	5/5 (100.0%)	2/5 (40.0%)	5/5 (100.0%)	2/5 (40.0%)	4/5 (80.0%)	4/5 (80.0%)

* Values represent the number of resulting tumors (>500 mm^3) over the total number of implantations. CSCs consistently generated more tumors than non-CSC populations while XRT or PI3K inhibitors efficiently blocked CSC initiated tumor growth across PDX cases. CSC=cancer stem cell. PI3K=phosphoinositide 3-kinase. PDX=patient-derived xenograft. XRT=radiation therapy.

REFERENCES

1. Gan GN, Eagles J, Keysar SB*, et al.* Hedgehog signaling drives radioresistance and stroma-driven tumor repopulation in head and neck squamous cancers. Cancer Res 2014;74(23):7024-36.

2. Keysar SB, Le PN, Anderson RT*, et al.* Hedgehog Signaling Alters Reliance on EGF Receptor Signaling and Mediates Anti-EGFR Therapeutic Resistance in Head and Neck Cancer. Cancer Res 2013;73(11):3381-92.

3. Keysar SB, Astling DP, Anderson RT*, et al.* A patient tumor transplant model of squamous cell cancer identifies PI3K inhibitors as candidate therapeutics in defined molecular bins. Mol Oncol 2013;7(4):776-90.

4. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 2011;17(1).

5. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30(15):2114-20.

6. Trapnell C, Roberts A, Goff L*, et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 2012;7(3):562-78.

7. Subramanian A, Tamayo P, Mootha VK*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102(43):15545-50.