Extracellular heparan 6-O-endosulfatases SULF1 and SULF2 in HNSC and other malignancies

Yang Yang[1], Jaeil Ahn[2], Nathan Edwards[1,3], Julius Benicky[3,4], Aaron M. Rozeboom[4], Bruce Davidson[5], Christina Karamboulas[6], Kevin C.J. Nixon[6], Laurie Ailles[6,7], Radoslav Goldman[1,3,4*]

[1]Department of Biochemistry and Molecular & Cell Biology, Georgetown University, Washington, D.C., USA

[2]Department of Biostatistics, Bioinformatics and Biomathematics, Georgetown University, Washington, D.C., USA

[3]Clinical and Translational Glycoscience Research Center, Georgetown University, Washington, D.C., USA

[4]Department of Oncology, Lombardi Comprehensive Cancer Center, Georgetown University, Washington, D.C., USA

[5]Department of Otolaryngology-Head and Neck Surgery, MedStar Georgetown University Hospital, Washington, D.C. USA

[6]Princess Margaret Cancer Centre, University Health Network, Toronto, ON Canada

[7]Department of Medical Biophysics, University of Toronto, Toronto, ON Canada

*Address correspondence to: Radoslav Goldman, Department of Oncology, NRB E207, 3970 Reservoir Rd NW, Washington, DC 20057, USA. Phone: +1 202-687-9868; Email: rg26@georgetown.edu;

The authors declare no potential conflicts of interest.

ABSTRACT

SULF1 and SULF2 are oncogenic in a number of human malignancies, including head and neck squamous cell carcinoma (HNSC). The function of these two heparan sulfate editing enzymes was previously considered largely redundant but the biology of cancer suggests differences that we explore in our RNAseq and RNAScope studies of HNSC and in a pan cancer analysis using the TCGA and CPTAC (proteomics) data. Our studies document a consistent upregulation of SULF1 and SULF2 in HNSC which is associated with poor survival outcomes. SULF2 expression increases in multiple malignancies but less consistently than SULF1, which uniformly increases in the tumor tissues and negatively impacts survival in several types of cancer. Meanwhile, SULF1 showed low expression in cancer cell lines and a scRNAseq study of HNSC shows that SULF1 is not supplied by epithelial tumor cells, like SULF2, but is secreted by cancer associated fibroblasts. Our RNAScope and PDX analysis of the HNSC tissues fully confirm the stromal source of SULF1 and explain the uniform impact of this enzyme on the biology of multiple malignancies. In summary, the SULF1 enzyme, supplied by a subset of cancer associated fibroblasts, is upregulated and negatively impacts HNSC survival at an early stage of the disease progression while the SULF2 enzyme, supplied by tumor cells, impacts survival at later stages of HNSC. This paradigm is common to multiple malignancies and suggests a potential for diagnostic and therapeutic targeting of the heparan sulfatases in cancer diseases.

INTRODUCTION

Eukaryotic sulfatases are primarily lysosomal enzymes that hydrolyze sulfate esters during the degradation of macromolecules, often in conjunction with glycosidases[1]. The sulfatase family of proteins shares a common catalytic mechanism using a formyl glycine residue in the active site for catalysis[2]. However, human 6-*O*-endosulfatases SULF1 and SULF2 are distinct from all other sulfatases in that they are neutral pH extracellular enzymes that edit the sulfation of heparan sulfate proteoglycans (HSPG) instead of degrading them[3]. The HSPG family of core proteins carries one or several serine/threonine residues covalently attached to a heparan sulfate glycosaminoglycan chain further modified by *N*-deacetylation and sulfation, epimerization, and variable *O*-sulfation[4]. Four sulfation sites at the *N*-, 3-*O*-, and 6-*O*-positions of the glucosamine and at the 2-*O*-position of the glucuronic acid[4] regulate protein interactions and the 6-*O*-sulfate-dependent interactions are critical regulators of the pathophysiology of multicellular organisms[5–8]. The SULF1 and SULF2 enzymes are the only post-synthetic editors of the 6-*O*-sulfation at the internal glucosamines of highly sulfated HSPG domains and their activity defines many critical interactions at the cell-surface and in the extracellular matrix (ECM).

HSPG exquisitely regulate embryogenesis, organogenesis, and physiology of nearly all organs by adjusting gradients of at least 600 proteins including growth factors, cytokines, chemokines, proteases, or collagens[6,9]. The highly specific SULF activities liberate sequestered HS-binding proteins which regulate matrix remodeling, immune infiltration, or signaling of the respective cognate receptors[7,10,11]. The determinants of ligand binding are under intense investigation[12–14] and systemic rules need to be further elucidated. We know, however, that heparan 6-*O*-sulfation is essential for binding of many ligands including VEGF, FGF-1, FGF-10, IL8, HGF, Wnt ligands or L- and P- selectins[4,10]. SULFs represent, therefore, an essential regulatory element that controls the HS-dependent developmental and pathophysiological processes including cancer progression[15–17].

Human SULF1 and SULF2 are 65% identical and there is no clear difference in the substrate-specificity of the two enzymes[18]. However, the impact of the two enzymes on cancer diseases is distinct. SULF2 is upregulated and oncogenic in various cancers[7] and a recent study documented that anti-SULF2 antibodies prevent tumor growth in a mouse model of cholangiocarcinoma[19]. The reported impact of SULF1 on cancer progression is less consistent. Widespread low expression of the SULF1 transcript is observed in cancer cell lines [20,21] and prior studies suggested a tumor suppressor function of SULF1 in ovarian, breast, and liver cancers[16]. In contrast, increased SULF1 expression is observed in a wide range of human tumors[7,16] and high SULF1 expression is associated with advanced primary tumor status, higher histological grade, and worse survival in urothelial carcinoma[22]. We have shown that SULF1 and SULF2 enzymes increase in tumor tissues of patients with HNSC and that the increase is associated with poor survival[23]. In this study, we therefore examined available datasets to find which cancers are affected by the 6-O-endosulfatases and we carried out experiments that verify the unifying concepts in their impact on HNSC and other cancers.

METHODS

Differential expression of SULF1 and SULF2 mRNAs in 32 TCGA studies

RNA-seq data and clinical information of 9,160 patients enrolled in 32 cancer studies conducted by the Cancer Genome Atlas (TCGA) consortium (portal.gdc.cancer.gov) and corresponding non-disease tissues from the Genotype-Tissue Expression (GTEx) project (gtexportal.org/home) were downloaded from UCSC-Xena on 02-26-2021 (xenabrowser.net/datapages); SULF1 and SULF2 mRNA was quantified as $\log_2$(RSEM counts+1) in both datasets. For the differential expression analysis, we selected 14 TCGA cancer studies with n>10 of paired tumor and normal tissues (**Table 1**) and we used a Wilcoxon rank sum test to compare paired tumor and non-tumor tissues, where the $\log_2$-fold change ($|\text{Log}_2\text{FC}|$) >1 and the false discovery rate (FDR) <0.05 across studies were considered as statistically significant. SULF1 and SULF2 mRNA expression was further compared by Wilcoxon rank sum test

between unpaired tumor tissues of 24 cancer studies and corresponding non-disease tissues of the same organs reported in GTEx (**Supplemental Table 1A**).

Differential expression of SULF proteins in 10 CPTAC studies

Proteomics data and clinical information of 1,247 patients enrolled in 10 cancer studies with matched tumor and adjacent non-tumor tissue-pairs conducted by the Clinical Proteomic Tumor Analysis Consortium (CPTAC) were downloaded from the Proteomics Data Commons (proteomic.datacommons.cancer.gov/pdc) and the CPTAC Data Portal[24] on 05-18-2020. Protein abundance, as determined by the CPTAC Common Data Analysis Pipeline[25] quantified the $\log_2$ ratio of individual proteins to an internal control of each study, using only peptides not shared between quantified proteins. We analyzed the differential expression between tumor and paired normal tissues by paired t-test; we compared SULF expression at different cancer stages by one-way ANOVA and computed the corresponding FDRs (**Table 2**).

Pan-cancer survival analysis based on SULF1 and SULF2 mRNA

The impact of SULF mRNA expression on time to event endpoints as cancer driven progression-free interval (PFI), defined by the TCGA research network[26], was evaluated by the Kaplan-Meier method in 20 TCGA cancer studies with >100 patients and >40 PFI events (**Supplemental Table 1B**). We identified optimal SULF cutoff values yielding the smallest p-value from the log-rank test when both SULF low and high groups, in each study, have at least 25% patients. We used the multivariable Cox-proportional hazard model (MCPH) adjusting for age and gender and summarized hazard ratios (HR) with 95% confidence intervals (CI). We used FDR<0.05 to call statistical significance. We further evaluated the HNSC TCGA dataset for an association between SULF1 or SULF2 mRNA expression and PFI at different pathological stages of the disease. We analyzed separately the early stage (stage I and II, n=96) and the late stage (stage III and IV, n=346) tumors. In each subset, we used the log-rank test as above.

A confirmatory survival analysis was completed on HNSC patients (n=88) enrolled at the Princess Margaret Cancer Centre, University of Toronto in line with IRB approved protocols. Majority of the patients have stage IV oral cancer (**Supplemental Table 2**). RNA was extracted from the snap frozen tumor tissues using a Qiagen RNeasy mini-kit and sample library preparation was done using the Illumina TruSeq stranded total RNA sample preparation kit. Sequencing used a 100-cycle paired read protocol and multiplexing to obtain ~75 million reads/sample on a Novaseq S4 flow cell using XP mode. Transcript abundances in transcripts per million (TPM) were generated from trimmed reads using Kallisto (v. 0.46.1) and the gencodev33 human transcriptome or a combined human-mouse transcriptome with reads aligning to mouse removed prior to analysis. Gene-level abundances (TPM) were calculated using the Bioconductor package tximport (v. 1.24.0). Survival analyses for disease-free interval were performed using methods described above.

Cell-specific expression of SULF1 and SULF2 in HNSC

We evaluated a published HNSC single cell RNA-seq dataset[27] that profiled transcriptomes of 5,578 cells from tumor tissues of 18 HNSC patients and identified the type of individual cells based on copy-number variations, karyotypes and expression signatures. We downloaded the data from the UCSC Cell Browser (cells.ucsc.edu/?ds=head-neck) and we quantified the SULF1 and SULF2 mRNAs as $\log_2$(TPM+1). Each cell-type expressing SULF1 or SULF2 was defined as a percentage of cells with non-zero expression values (**Supplemental Table 3**). Differential expression of SULF1 and SULF2 between the cell types was calculated by non-paired t-test.

Correlation studies of the SULF enzymes with CAF and other proteins

We computed Pearson's correlations between SULF1 and all proteins/corresponding mRNAs in the CPTAC study of HNSC[28]. To explore magnitude of the correlations between SULF1 and cancer associated fibroblasts (CAF), the averaged Pearson's correlation between CAF1 proteins (n=206) and SULF1 was compared with the averaged SULF1 correlation with all proteins (n=10,073) where p-value

was obtained by performing 10,000 permutations using a randomly selected subset of proteins (n=206). We adopted the CAF1 and CAF2 definitions by Puram et al[27]. In addition, we adopted the CAF definition of a COL11A1-related CAF subset associated with pre-metastatic locally invasive tumors proposed by Anastassiou[29,30] (**Supplemental Table 4**).

RNAscope analysis of SULF1 and SULF2 in OSCC tissues

We selected tumor tissues of patients (n=20) with carcinoma of the oral cavity (OSCC) for the analysis of SULF1 and SULF2 expression based on *in situ* hybridization. The patients were either node positive (n=10) or node negative (n=10) and the node positive group had, in general, poor survival outcomes (**Supplemental Table 5**). FFPE sections (5 µm) of the patient's tumors were baked at 60°C, deparaffinized, and dehydrated. The RNAScope assay (RNAscope Multiplex Fluorescent Reagent Kit v2 #323100) was done according to the manufacturers protocol with probes for Sulf1 and Sulf2 (ACD 403581-C3 and ACD 502241) paired with OPALs 650 and 570 (Akoya FP1496001KT and FP1488001KT, respectively). After the final wash, slides were prepped for IHC and incubated for 60min with anti-panCK antibody (M3515 DAKO), anti-mouse HRP secondary (DAKO K400111-2), OPAL TSA 520 (Akoya # FP1487001KT), and DAPI (Akoya # FP1490).

Slides were scanned at 10× magnification using the Vectra 3.0 Automated Quantitative Pathology Imaging System (Akoya). Whole slide scans were viewed with Phenochart (Akoya) and high-powered images at 20× (resolution of 0.5 µm per pixel) were selected for multispectral image capture. Three to 20 multispectral image regions of interest (ROIs; 669 µm × 500 µm) were captured in the tumor and normal adjacent regions on each slide. A selection of 10-15 representative multispectral images spanning all tissue sections was used to train the inForm software (tissue/cell segmentation and phenotyping tools). All the settings applied to the training images were saved within an algorithm for batch analysis of all the multispectral images for the project. The analysts were blinded to the patient status and all the raw data were consolidated in PhenoptrReports (Akoya). We quantified the

total number of cells in the panCK+ tumor area and the adjacent panCK-area (stroma); these compartments were analyzed for the SULF1+ and SULF2+ cells (**Supplemental Table 5**).

Expression of SULF1 and SULF2 in a PDX model of OSCC

PDX models were generated as described[31]. RNAseq was carried out on snap-frozen tissues from early passage (passage 1 to 3) PDX models, as described above. Expression of SULF1 and SULF2 was compared between patient and xenograft samples using a paired Wilcoxon rank sum test and visualized using the R package ggplot2 (v. 3.3.6).

Study Approval

The studies involving human participants were reviewed and approved by the Georgetown University-MedStar Health Institutional Review Board and the University Health Network Research Ethics Board. The patients/participants provided their written informed consent to participate in this study.

RESULTS

SULF1 and SULF2 mRNA expression in different cancer types

We extended our analysis of HNSC[23] by retrieving the RNA-seq data of 9,160 patients in 32 cancer studies from the TCGA database to develop a systematic evaluation of SULF1 and SULF2 expression. We compared tumor and paired normal tissues in 14 cancer studies with >10 available paired samples (**Table 1**). SULF1 is significantly upregulated in tumor tissues of 10 of the 14 studies of which 9 show >2-fold upregulation. The highest SULF1 log2FC is observed in LUAD ($log_2FC = 2.78$, FDR<0.001). SULF2 is overexpressed in 8 cancer types of which 5 increase >2-fold. The highest SULF2 FC is observed in ESCA ($log_2FC = 2.76$, FDR<0.001). Besides the wide-scale overexpression in tumor tissues, SULF1 is significantly downregulated in KICH and THCA and SULF2 in PRAD but none of the studies reaches a 2-fold decrease.

Our comparison of SULF mRNA in 32 TCGA tumors with the corresponding non-disease tissues from GTEx (**Supplemental Table 1A**) shows that SULF1 is significantly increased in 18 cancers of which 16 show >2-fold upregulation. SULF2 is significantly higher in 16 cancer studies, of which 11 show >2-fold increase. All the TCGA studies with significant SULF1 or SULF2 upregulation in paired tissues retain the trend in our GTEx analysis. However, the slight decreases observed for SULF1 in KICH and THCA or for SULF2 in PRAD (**Table 1**) are not confirmed in the GTEX comparison. The decrease in SULF1 is not significant for any of the GTEX comparisons; SULF2 expression is significantly lower in OV (log_2FC=-2.27, FDR<0.001) and UCEC (log_2FC=-1.76, FDR<0.001).

SULF1 and SULF2 proteins in 10 CPTAC studies

We analyzed differential expression of SULF proteins in paired tumor and normal tissues of 10 proteomics studies from the CPTAC consortium (**Table 2**). Seven of the cancer types (LUSC, LUAD, HNSC, KIRC, BRCA, COAD, and LIHC/HCC-HBV) overlap with the TCGA datasets, which enables comparison of the expression at the transcriptional and translational levels. Similar to the pervasive upregulation of SULF1 mRNA, SULF1 protein is significantly upregulated in tumor compared to paired normal tissues in 9 of the 10 studies. Four studies showed >2-fold increase of SULF1 protein in tumor tissues (**Table 2**) with the highest fold-change observed in HNSC (log_2FC=1.59, FDR<0.001). Only the smaller size (n=12) study of OSC did not show any difference in SULF1 protein. SULF2 protein is significantly upregulated in the tumors of 6 studies but a >2-fold increase is only observed in PDAC (log_2FC=1.14, FDR<0.001). In addition, we saw a significant downregulation of SULF2 protein in tumor tissues of HBV-related HCC (log_2FC=-0.33, FDR<0.001) and UCEC (log_2FC =-0.43, FDR=0.01). This is consistent with the reduced expression of SULF2 mRNA in tumor tissues of LIHC (log_2FC=-0.481, FDR=0.116, **Table 1**) and UCEC (log_2FC=-1.76, FDR<0.001, **Supplemental Table 1A**). Based on the SULF expression in the two independent datasets (TCGA and CPTAC), we conclude that SULF1 is commonly upregulated across different cancer types while SULF2 overexpression is more restricted to certain cancer pathologies.

We observed >2-fold upregulation of both SULF1 and SULF2 mRNA in four cancer studies from the TCGA (HNSC, ESCA, LUSC, and STAD) and all these cancers remain significantly upregulated compared to the GTEX normal tissues (**Figure 1A, 1B**). At the same time, SULF1 and SULF2 proteins are significantly elevated in HNSC and LUSC (**Table 1**) while the STAD and ESCA studies were not reported at the time of our analysis. Pancreatic cancer (PDAC) is the only CPTAC study with both SULF1 and SULF2 protein significantly upregulated >2-fold in tumor compared to paired normal tissues (**Figure 1C**). Limited size (n=4 paired samples) of the PDAC study prevented our analysis of paired mRNA expression in the TCGA dataset. However, SULF1 and SULF2 mRNAs are >20-fold higher in the PDAC tumor tissues than in the non-disease pancreatic tissue from the GTEx (SULF1 $\log_2$FC=6.81, SULF2 $\log_2$FC=4.85, both p<0.001, **Figure 1D**), which is consistent with the large difference in the protein expression (**Table 2**). The survival analyses presented below further support the impact of SULF enzymes on HNSC and PDAC and warrant additional study.

Association of SULF1 and SULF2 mRNA expression with survival outcomes

The association of SULF1 or SULF2 mRNA expression and PFI was analyzed by univariate log-rank tests and compared to published studies (**Supplemental Table 1B**). Our literature search found five studies (bladder[22], breast[32], lung[16], gastric[33], and liver[11]) showing that high SULF1 expression is associated with poor survival outcomes. We observed an adverse prognostic trend for these cancers in the TCGA studies but the associations did not reach significance (**Supplemental Table 1B**). In addition, high SULF1 is a significant (FDR<0.05) prognostic factor in KIRP (HR=2.693), PAAD (HR=2.365), CESC (HR=2.227), COAD (HR=1.867), and LGG (HR=1.479); we are not aware of studies reporting these associations. We note that SULF1 is significantly increased at the mRNA (**Table 1**) and protein (**Table 2**) levels in COAD and negatively impacts survival; such associations deserve further attention. The association of high SULF1 with poor survival (**Figure 2**) in many cancers is quite remarkable especially in view of the fact that the SULF1 transcript is low in most cancer cell lines (**Supplemental Figure 1**) [20,21].

We found studies of 8 cancers (bladder[22], esophagus[34], head and neck[23], kidney[35], liver[36], lung[37], and pancreatic[38]) showing significant association of SULF2 with survival but the impact is less uniform. Our analysis confirms the reported associations (all FDR<0.05) of high SULF2 expression with poor PFI of patients with HNSC (HR=1.687), LIHC (HR =1.587), and PAAD (HR =1.724) (**Figure 2**). An adverse association of SULF2 with PFI in ESCA (HR=1.352, p=0.158) did not reach statistical significance as reported[34]. In addition, high SULF2 expression is significantly (FDR<0.05) associated with better PFI in LGG (HR=0.354, p<0.001) and UCEC (HR=0.415, p=0.004) (**Supplemental Table 1B**) and a favorable prognostic impact of high SULF2 was reported in clear cell renal carcinoma[35] (HR=0.07, p=0.015, n=49) and lung squamous cell carcinoma[37] (HR=0.11, p=0.02, n=51). These observations were, however, not corroborated in the larger TCGA studies (**Supplemental Table 1B**).

Survival impact of SULF1 and SULF2 in HNSC differs by pathological tumor stage

We have shown that high SULF1 or SULF2 expression in HNSC is associated with poor survival outcomes[23]. This association was further confirmed in our study of 88 HNSC patients enrolled at the University of Toronto; we used an optimized cutoff of SULF1 (52 high and 36 low expressors) or SULF2 (24 high and 64 low expressors) to show that high expression of either gene is associated with poor disease-free interval (p<0.001) **(Supplemental Figure 2)**. The study has limited size but provides an important independent verification of the results.

SULF1 is associated with poor survival in univariate analysis[23] but, contrary to SULF2, loses a significant impact when analyzed in a multivariable model together with SULF2, age, gender, smoking history, tumor stage, and radiation therapy. However, the TCGA sample-set used in the analysis is dominated by tumors of stage 3 and 4. To further evaluate the impact of SULF1, we analyzed separately HNSC patients with early (stage 1 and 2) or late (stage 3 and 4) tumors. We observe that the survival outcomes differ by stage even though SULF1 and SULF2 expression does not differ between early- and late-stage tumors[23]. High SULF1 mRNA expression in tumor is associated with poor PFI in early-stage patients (HR=2.327, p=0.023, **Figure 3A**) but not in late-stage patients

(HR=1.034, p=0.842). On the contrary, SULF2 mRNA overexpression is significantly associated with poor PFI outcomes in late-stage patients (HR=1.794, p<0.001, **Figure 3B**) but not in early-stage patients (HR=1.457, p=0.866). The high impact of SULF1 in the early tumors is even more pronounced in a multivariable model[23] that includes SULF1, SULF2, age, gender, smoking, and radiation therapy. While SULF2 is insignificant in the 97 stage 1 and 2 patients (HR=1.32 (95% CI, 0.58-3.0), p=0.59), high SULF1 remains an independent predictor of poor PFI (HR=4.61 (1.88-11.3) p<0.001). We speculate that this pattern of adverse prognostic impact is associated with a SULF1 function in the local spread of the disease at an early stage which is complemented by SULF2 activity at later stages of the HNSC progression.

Cell-specific expression of SULF enzymes in HNSC

Analysis of a single cell RNA-seq dataset of HNSC[27] showed that the percentage of SULF1 and SULF2 positive cells varies substantially across 9 cell types (**Supplemental Table 3**). SULF2 is expressed in 63% of all tumor cells (n=1,389 of 2215) which is the highest representation among all the cell types. The positivity of SULF1 is the highest in fibroblasts (48%, n=691 of 1,440) compared to <20% in any other cell type (**Supplemental Table 3**, **Figure 4A**). SULF1 is expressed in only 14% of tumor cells (n=309 of 2,215).

SULF1 expression is significantly higher in fibroblasts compared to tumor epithelial cells in terms of both percent positivity and expression; in contrast, SULF2 expression has the opposite trend (**Figure 4**). The mean SULF1 mRNA, represented as $\log_2$(TPM+1), is 0.172 in tumor cells compared to 1.099 in fibroblasts (p<0.001); the mean value of SULF2 mRNA is 1.145 in tumor cells compared to 0.386 in fibroblasts (p<0.001) (**Figure 4A**). The percentage of positive cells among individual patients ranges from 2.4-54.5% for SULF1 and 43.1-91.7% for SULF2 in tumor cells, and from 31.3-71.0% for SULF1 and 8.7-43.8% for SULF2 in fibroblasts (**Figure 4B**). Interestingly, the mean expression of SULF1 in SULF1-positive cells remains significantly higher in fibroblasts than in tumor cells (2.291 vs 1.232, p<0.001); however, the mean expression of SULF2 in SULF2-positive fibroblasts and tumor cells

is the same (1.826 vs 1.812, p=0.8, **Figure 4C**). These results suggest that SULF1 is expressed to a high degree by a large sub-population of fibroblasts; SULF2 is expressed to the same degree in the fibroblasts and tumor cells but the population of tumor cells expressing SULF2 is much bigger than that of the fibroblasts. We conclude that SULF1 and SULF2 in HNSC derive primarily from the fibroblasts and tumor cells, respectively. The results show that the expression of SULF1 and SULF2 in HNSC is regulated by different mechanisms which leads to an independent regulation of the HS-dependent signaling activities by the two enzymes.

To further strengthen the observation that SULF1 is expressed in fibroblasts and SULF2 in tumor cells, we analyzed the RNA-seq data from the Cancer Cell Line Encyclopedia (CCLE). SULF1 expression in fibroblast cells is distinctly higher compared to all the cancer cell lines (**Supplemental Figure 1A**) but SULF2 expression is the highest in cell lines from neuroblastoma, HNSC, and other cancers (**Supplemental Figure 1B**). This suggests that SULF1 expression in fibroblasts of the tumor tissues is not unique for HNSC but more likely a pan-cancer event.

A final demonstration of the expression of SULF1 in fibroblasts comes from our PDX studies of 42 HNSC patients (**Figure 5**) showing that in all but one case the expression of SULF1 decreases in the PDX compared to the primary tumor (median primary tumor TPM=52.6, median PDX TPM = 1.7, $p<0.001$; Wilcoxon rank-sum test). This is in line with the expansion of tumor cells and loss of the transplanted stroma commonly observed in the PDX models. In contrast, SULF2 expression in the PDX increases (median primary tumor TPM = 50, median PDX TPM = 103, $p<0.01$; Wilcoxon rank-sum test) which confirms that the tumor cell is the major source of this enzyme.

SULF1 expression in cancer-associated fibroblasts

Dominant expression of SULF1 in fibroblasts and its increase in HNSC tissues, in spite of low expression in the HNSC cell lines, prompted us to analyze its connection with cancer-associated fibroblasts (CAF). We examined the correlation of genes (n=206) of CAF1, a HNSC CAF defined previously[27], with SULF1 in the CPTAC HNSC study. Analysis of the Pearson's correlation coefficients

shows that 122 proteins in the CPTAC dataset are correlated with SULF1 with r>0.55, of which 44 belong to the CAF1 cluster (**Supplemental Table 4**). The distribution of the correlation coefficients of the CAF1 genes is shifted to significantly higher values compared to other proteins (n=10,073, p<0.001) (**Figure 4D**). The correlations of SULF2 protein with the CAF1 genes is weaker, as expected, and we observed a similar trend in the RNAseq data from the CPTAC HNSC study (**Supplemental Table 4**). The expression signature of the HNSC CAF1[30] overlaps substantially with a COL11A1-expressing CAF which is defined by an invasive pre-metastatic phenotype[29,30]. This subset of CAF was observed in multiple cancers (HNSC, ovarian, pancreatic, colorectal) which strongly suggests that SULF1 is impactful in multiple cancers, in addition to HNSC.

The association of SULF1 with CAF is further supported by our RNAscope analysis of SULF1 and SULF2 in 20 OSCC patients (**Figure 6**). The *in-situ* hybridization clearly shows that SULF1 expressing cells are more common in the stroma (mean 24.9% stroma vs 8.4% tumor, p<0.001) while SULF2 expression is higher in the cancer cells (mean 22.7% stroma vs 52.5% tumor, p<0.001) (**Supplemental Table 5**). In an exploratory analysis, we separated the OSCC patients into node positive (n=10) and node-negative (n=10) groups and we observe a trend for higher SULF1 and SULF2 expression in the node-positive cases (**Figure 6**). The results support that SULF1-positive CAF accumulate at the invasive front of the HNSC at a pre-metastatic stage and facilitate local invasion.

DISCUSSION

Previous studies showed that SULF1 and SULF2 are upregulated in several cancers[4,16,22,23,34,37,38]. SULF2 is considered oncogenic[7], the SULF1 impact is more controversial[7,39,40]. Are there unifying trends in the cancer biology of the heparan 6-O-endosulfatases supporting their impact on HNSC and other cancers? Proteogenomic analysis of the TCGA and CPTAC datasets conclusively documents a significant elevation of SULF1 and SULF2 in multiple cancer tissues compared to adjacent (**Table 1**) or normal (**Supplemental Table 1**) counterparts. This is uniformly corroborated by protein increases

(**Table 2**) and at least six cancers (BRCA, HNSC, KIRC, LUSC, LUAD, PAAD) consistently upregulate both SULF1 and SULF2. PDAC mRNA was not reported in the TCGA adjacent normal but both mRNA (compared to normal) and protein show some of the highest increases overall. Other cancers upregulate one of the SULFs (e.g. SULF1 in COAD) or the SULFs remain unchanged, but decreases, like SULF2 in UCEC, are rare.

At the same time, high expression of SULF1 or SULF2 is typically associated with poor survival (**Supplemental Table 1B**). Among all the cancers examined, we observe most consistent impact of SULF1 and SULF2 in PAAD and HNSC (**Figure 1 and 2**) but other cancers are affected as well. HNSC is an interesting case because SULF2 negatively affects survival of stage III and IV patients while the impact of SULF1 is more prominent in stage I and II cancer patients (**Figure 3**). Associations of high expression with improved outcomes are restricted to SULF2 (e.g. in UCEC) but remain exceptional. High SULF1 predicts significantly lower PFI in at least 5 cancers and other malignancies, reported previously, follow a similar trend (**Supplemental Table 1B**). The consistent increases of SULF1 in cancer tissues and their uniform association with poor survival outcomes are quite remarkable because SULF1 expression in cancer cell lines is typically low[20,21] (**Supplemental Figure 1**). However, our analyses of HNSC (**Figure 6**) and analyses of other cancers[41] show that SULF1 is high in cancer tissues and supplied by the CAF, a cell type associated with cancer invasion, metastasis, and the escape from immune surveillance[30,42,43] and so far overlooked in the cancer biology of the SULF enzymes.

Our PDX study of HNSC shows that SULF1, contrary to SULF2, disappears in the transplanted tumors, as expected for a gene expressed by the stromal cells (**Figure 5**). The scRNAseq data[27] show a strong correlation of SULF1 with CAF1 genes (**Figure 4**) and the genes of the COL11A1-expressing CAF associated with locally invasive pre-metastatic cancer disease[29] (**Supplemental Table 4**). SULF1 is a gene typical of this subtype of CAF in several cancers[30] which supports that the CAF supply SULF1 not only in HNSC but in general; the function of SULF1 in the biology of the CAF deserves further attention. Finally, our RNAScope study shows that SULF1+ cells localize to the stroma while SULF2+

cells overlap to a large degree with the cytokeratin+ tumor cells (**Figure 3**). Our results also suggest that SULF1+ cells in the stroma are more abundant in node positive tumors which is in line with recent papers associating low tumor/stroma ratio[44] or the presence of CAF[45] with poor HNSC survival outcomes.

In conclusion, our study confirms overexpression of SULF1 and SULF2 in various cancers which is commonly associated with poor survival outcomes; the 6-O-endosulfatases emerge as interesting targets for cancer monitoring and therapeutic intervention. It is expected that the enzymes determine survival of HNSCC patients by adjusting gradients of heparan sulfate binding proteins in the microenvironment of tumors. We have strong evidence that SULF1 is supplied by CAF while SULF2 is provided primarily by the cancer cells which has important consequences because the secreted SULF enzymes act locally due to strong non-covalent interactions with cell surfaces[7]. In addition, recent data suggest that their activity is regulated by cell-specific posttranslational modifications[46]. We know that SULFs regulate oncogenic pathways but they also adjust matrix structure, angiogenesis, or immune responses[3,4,7,8,10,15–17,47,48] and their function at the tumor/stroma interface in different cancers needs further study.

Acknowledgments

The authors want to thank Dr. Dimitris Anastassiou for his insightful analyses of the CAF.

REFERENCE

1. Morimoto-Tomita M, Uchimura K, Werb Z, Hemmerich S, Rosen SD. Cloning and characterization of two extracellular heparin-degrading endosulfatases in mice and humans. *J Biol Chem*. 2002;277(51):49175-49185. doi:10.1074/jbc.M205131200

2. Appel MJ, Bertozzi CR. Formylglycine, a Post-Translationally Generated Residue with Unique Catalytic Capabilities and Biotechnology Applications. *ACS Chem Biol*. 2015;10(1):72-84. doi:10.1021/cb500897w

3. Hanson SR, Best MD, Wong CH. Sulfatases: structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew Chem Int Ed Engl*. 2004;43(43):5736-5763. doi:10.1002/anie.200300632

4. Morimoto-Tomita M, Uchimura K, Bistrup A, et al. Sulf-2, a proangiogenic heparan sulfate endosulfatase, is upregulated in breast cancer. *Neoplasia*. 2005;7(11):1001-1010. doi:10.1593/neo.05496

5. Habuchi H, Habuchi O, Kimata K. Sulfation pattern in glycosaminoglycan: does it have a code? *Glycoconj J*. 2004;21(1-2):47-52. doi:10.1023/B:GLYC.0000043747.87325.5e

6. Bishop JR, Schuksz M, Esko JD. Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature*. 2007;446(7139):1030-1037. doi:10.1038/nature05817

7. Rosen SD, Lemjabbar-Alaoui H. Sulf-2: an extracellular modulator of cell signaling and a cancer target candidate. *Expert Opin Ther Targets*. 2010;14(9):935-949. doi:10.1517/14728222.2010.504718

8. Ai X, Do AT, Lozynska O, Kusche-Gullberg M, Lindahl U, Emerson CPJ. QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. *J Cell Biol*. 2003;162(2):341-351. doi:10.1083/jcb.200212083

9. Sarrazin S, Lamanna WC, Esko JD. Heparan sulfate proteoglycans. *Cold Spring Harb Perspect Biol*. 2011;3(7). doi:10.1101/cshperspect.a004952

10. Uchimura K, Morimoto-Tomita M, Bistrup A, et al. HSulf-2, an extracellular endoglucosamine-6-sulfatase, selectively mobilizes heparin-bound growth factors and chemokines: effects on VEGF, FGF-1, and SDF-1. *BMC Biochem*. 2006;7:2. doi:10.1186/1471-2091-7-2

11. Dhanasekaran R, Nakamura I, Hu C, et al. Activation of the transforming growth factor-β/SMAD transcriptional pathway underlies a novel tumor-promoting role of sulfatase 1 in hepatocellular carcinoma. *Hepatology*. 2015;61(4):1269-1283. doi:10.1002/hep.27658

12. Chopra P, Joshi A, Wu J, et al. The 3-O-sulfation of heparan sulfate modulates protein binding and lyase degradation. *Proc Natl Acad Sci U S A*. 2021;118(3). doi:10.1073/pnas.2012935118

13. Nagarajan B, Holmes SG, Sankaranarayanan NV, Desai UR. Molecular dynamics simulations to understand glycosaminoglycan interactions in the free- and protein-bound states. *Curr Opin Struct Biol*. 2022;74:102356. doi:10.1016/j.sbi.2022.102356

14. Qiu H, Shi S, Yue J, et al. A mutant-cell library for systematic analysis of heparan sulfate structure-function relationships. *Nat Methods*. 2018;15(11):889-899. doi:10.1038/s41592-018-0189-6

15. El Masri R, Seffouh A, Lortat-Jacob H, Vivès RR. The “in and out” of glucosamine 6-O-sulfation: the 6th sense of heparan sulfate. *Glycoconj J*. 2017;34(3):285-298. doi:10.1007/s10719-016-9736-5

16. Bret C, Moreaux J, Schved JF, Hose D, Klein B. SULFs in human neoplasia: implication as progression and prognosis factors. *J Transl Med*. 2011;9:72. doi:10.1186/1479-5876-9-72

17. Marques C, Reis CA, Vivès RR, Magalhães A. Heparan Sulfate Biosynthesis and Sulfation Profiles as Modulators of Cancer Signalling and Progression. *Front Oncol*. 2021;11:778752. doi:10.3389/fonc.2021.778752

18. Holmes R. Comparative and evolutionary studies of vertebrate extracellular sulfatase genes and proteins: SULF1 and SULF2. *J Proteomics Bioinform*. 2017;10(2).

19. Luo X, Campbell NA, He L, et al. SULF2 monoclonal antibody 5D5 suppresses human cholangiocarcinoma xenograft growth via regulation of a SULF2-PDGFRβ-YAP signaling axis. *Hepatology*. Published online March 18, 2021. doi:10.1002/hep.31817

20. Yang YW, Phillips JJ, Jablons DM, Lemjabbar-Alaoui H. Development of novel monoclonal antibodies and immunoassays for sensitive and specific detection of SULF1 endosulfatase. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 2021;1865(2):129802. doi:10.1016/j.bbagen.2020.129802

21. Lai J, Chien J, Staub J, et al. Loss of HSulf-1 up-regulates heparin-binding growth factor signaling in cancer. *J Biol Chem*. 2003;278(25):23107-23117. doi:10.1074/jbc.M302203200

22. Lee HY, Yeh BW, Chan TC, et al. Sulfatase-1 overexpression indicates poor prognosis in urothelial carcinoma of the urinary bladder and upper tract. *Oncotarget*. 2017;8(29):47216-47229. doi:10.18632/oncotarget.17590

23. Yang Y, Ahn J, Raghunathan R, et al. Expression of the Extracellular Sulfatase SULF2 Affects Survival of Head and Neck Squamous Cell Carcinoma Patients. *Front Oncol*. 2020;10:582827. doi:10.3389/fonc.2020.582827

24. Edwards NJ, Oberti M, Thangudu RR, et al. The CPTAC Data Portal: A Resource for Cancer Proteomics Research. *J Proteome Res*. 2015;14(6):2707-2713. doi:10.1021/pr501254j

25. Rudnick PA, Markey SP, Roth J, et al. A Description of the Clinical Proteomic Tumor Analysis Consortium (CPTAC) Common Data Analysis Pipeline. *J Proteome Res*. 2016;15(3):1023-1032. doi:10.1021/acs.jproteome.5b01091

26. Liu J, Lichtenberg T, Hoadley KA, et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell*. 2018;173(2):400-416.e11. doi:10.1016/j.cell.2018.02.052

27. Puram SV, Tirosh I, Parikh AS, et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell*. 2017;171(7):1611-1624.e24. doi:10.1016/j.cell.2017.10.044

28. Huang C, Chen L, Savage SR, et al. Proteogenomic insights into the biology and treatment of HPV-negative head and neck squamous cell carcinoma. *Cancer Cell*. 2021;39(3):361-379.e16. doi:10.1016/j.ccell.2020.12.007

29. Kim H, Watkinson J, Varadan V, Anastassiou D. Multi-cancer computational analysis reveals invasion-associated variant of desmoplastic reaction involving INHBA, THBS2 and COL11A1. *BMC Med Genomics*. 2010;3:51. doi:10.1186/1755-8794-3-51

30. Zhu K, Cai L, Cui C, de Los Toyos JR, Anastassiou D. Single-cell analysis reveals the pan-cancer invasiveness-associated transition of adipose-derived stromal cells into COL11A1-expressing cancer-associated fibroblasts. *PLoS Comput Biol*. 2021;17(7):e1009228. doi:10.1371/journal.pcbi.1009228

31. Karamboulas C, Bruce JP, Hope AJ, et al. Patient-Derived Xenografts for Prognostication and Personalized Treatment for Head and Neck Squamous Cell Carcinoma. *Cell Rep*. 2018;25(5):1318-1331.e4. doi:10.1016/j.celrep.2018.10.004

32. Chen F, Zhang Z, Yu Y, Liu Q, Pu F. HSulf‐1 and palbociclib exert synergistic antitumor effects on RB‐positive triple‐negative breast cancer. *Int J Oncol*. 2020;57(1):223-236. doi:10.3892/ijo.2020.5057

33. Hu Y, Li J, Luo H, Song W, Yang J. Differential Expression of COL1A1, COL1A2, COL6A3, and SULF1 as Prognostic Biomarkers in Gastric Cancer. *Int J Gen Med*. 2021;14:5835-5843. doi:10.2147/IJGM.S321265

34. Lui NS, van Zante A, Rosen SD, Jablons DM, Lemjabbar-Alaoui H. SULF2 expression by immunohistochemistry and overall survival in oesophageal cancer: a cohort study. *BMJ Open*. 2012;2(6):e001624. doi:10.1136/bmjopen-2012-001624

35. Kumagai S, Ishibashi K, Kataoka M, et al. Impact of Sulfatase-2 on cancer progression and prognosis in patients with renal cell carcinoma. *Cancer Sci*. 2016;107(11):1632-1641. doi:10.1111/cas.13074

36. Lai JP, Sandhu DS, Yu C, et al. Sulfatase 2 up-regulates glypican 3, promotes fibroblast growth factor signaling, and decreases survival in hepatocellular carcinoma. *Hepatology*. 2008;47(4):1211-1222. doi:10.1002/hep.22202

37. Lui NS, Yang YW, van Zante A, Buchanan P, Jablons DM, Lemjabbar-Alaoui H. SULF2 Expression Is a Potential Diagnostic and Prognostic Marker in Lung Cancer. *PLoS One*. 2016;11(2):e0148911. doi:10.1371/journal.pone.0148911

513 38. Alhasan SF, Haugk B, Ogle LF, et al. Sulfatase-2: a prognostic biomarker and candidate
514 therapeutic target in patients with pancreatic ductal adenocarcinoma. *Br J Cancer*.
515 2016;115(7):797-804. doi:10.1038/bjc.2016.264

516 39. Lai JP, Sandhu DS, Shire AM, Roberts LR. The tumor suppressor function of human sulfatase 1
517 (SULF1) in carcinogenesis. *J Gastrointest Cancer*. 2008;39(1-4):149-158. doi:10.1007/s12029-
518 009-9058-y

519 40. Hammond E, Khurana A, Shridhar V, Dredge K. The Role of Heparanase and Sulfatases in the
520 Modification of Heparan Sulfate Proteoglycans within the Tumor Microenvironment and
521 Opportunities for Novel Cancer Therapeutics. *Front Oncol*. 2014;4:195.
522 doi:10.3389/fonc.2014.00195

523 41. Hussain A, Voisin V, Poon S, et al. Distinct fibroblast functional states drive clinical outcomes in
524 ovarian cancer and are regulated by TCF21. *J Exp Med*. 2020;217(8). doi:10.1084/jem.20191094

525 42. Chakravarthy A, Khan L, Bensler NP, Bose P, De Carvalho DD. TGF-β-associated extracellular
526 matrix genes link cancer-associated fibroblasts to immune evasion and immunotherapy failure.
527 *Nat Commun*. 2018;9(1):4692. doi:10.1038/s41467-018-06654-8

528 43. Chen Y, McAndrews KM, Kalluri R. Clinical and therapeutic relevance of cancer-associated
529 fibroblasts. *Nat Rev Clin Oncol*. 2021;18(12):792-804. doi:10.1038/s41571-021-00546-5

530 44. Almangush A, Alabi RO, Troiano G, et al. Clinical significance of tumor-stroma ratio in head and
531 neck cancer: a systematic review and meta-analysis. *BMC Cancer*. 2021;21(1):480.
532 doi:10.1186/s12885-021-08222-8

533 45. Marsh D, Suchak K, Moutasim KA, et al. Stromal features are predictive of disease mortality in
534 oral cancer patients. *J Pathol*. 2011;223(4):470-481. doi:10.1002/path.2830

535 46. El Masri R, Seffouh A, Roelants C, et al. Extracellular endosulfatase Sulf-2 harbors a
536 chondroitin/dermatan sulfate chain that modulates its enzyme activity. *Cell Rep*.
537 2022;38(11):110516. doi:10.1016/j.celrep.2022.110516

538 47. Lemjabbar-Alaoui H, van Zante A, Singer MS, et al. Sulf-2, a heparan sulfate endosulfatase,
539 promotes human lung carcinogenesis. *Oncogene*. 2010;29(5):635-646. doi:10.1038/onc.2009.365

540 48. Nawroth R, van Zante A, Cervantes S, McManus M, Hebrok M, Rosen SD. Extracellular
541 sulfatases, elements of the Wnt signaling pathway, positively regulate growth and tumorigenicity of
542 human pancreatic cancer cells. *PLoS One*. 2007;2(4):e392. doi:10.1371/journal.pone.0000392

543

544

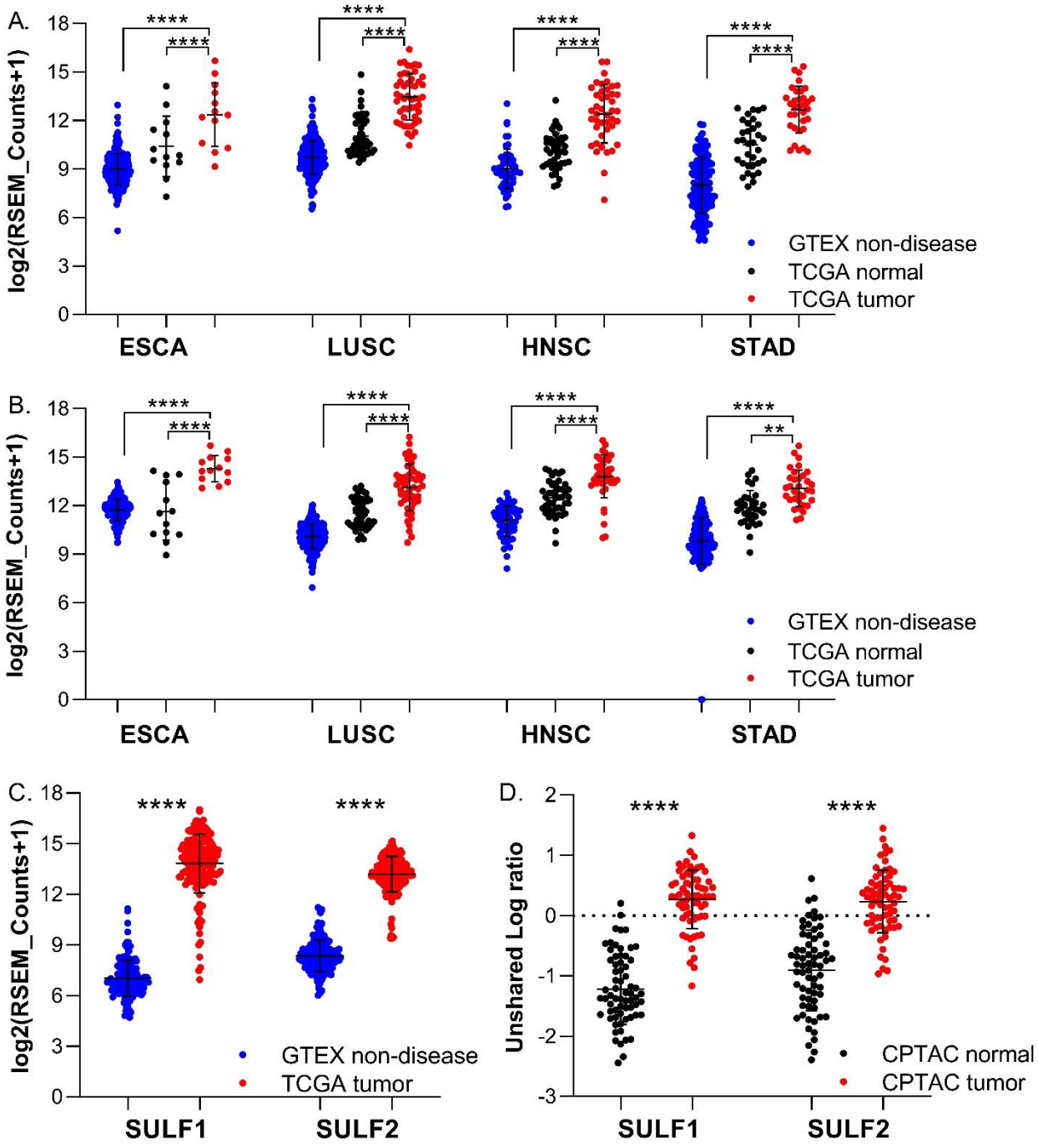
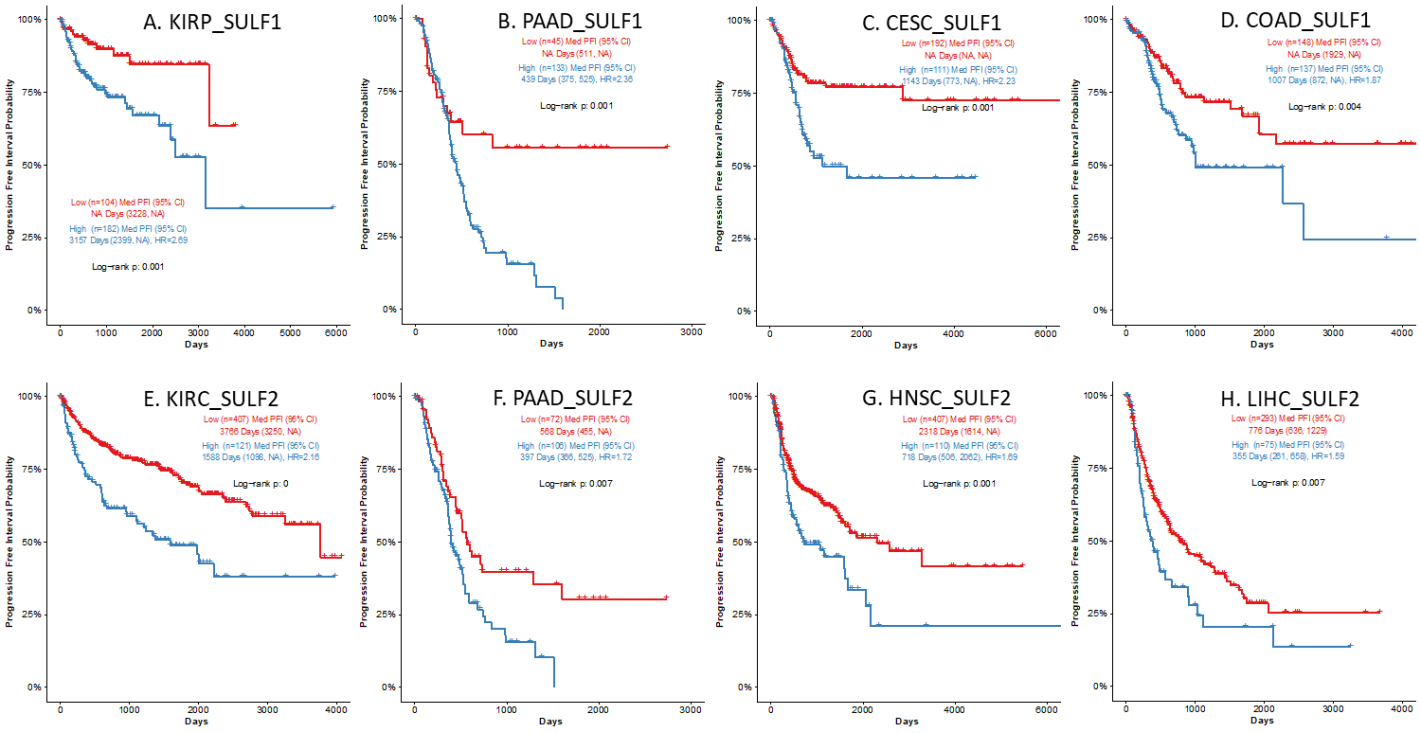


Figure 1. Expression of SULF1 and SULF2 in five cancer types. (A) SULF1 and (B) SULF2 mRNA in tumor compared to paired normal tissue of cancer patients in TCGA datasets and to non-disease tissues of healthy donors from GTEx; (C) SULF1 and SULF2 mRNA in the PAAD tumor tissues from TCGA compared to non-cancerous pancreatic tissue from GTEX. (D) SULF1 and SULF2 protein in the PDAC tumor and adjacent non-cancer tissues from CPTAC; p<0.0001(****), p<0.01(**).



552

553 **Figure 2. SULF1 and SULF2 expression in tumor tissues is associated with poor survival.** High

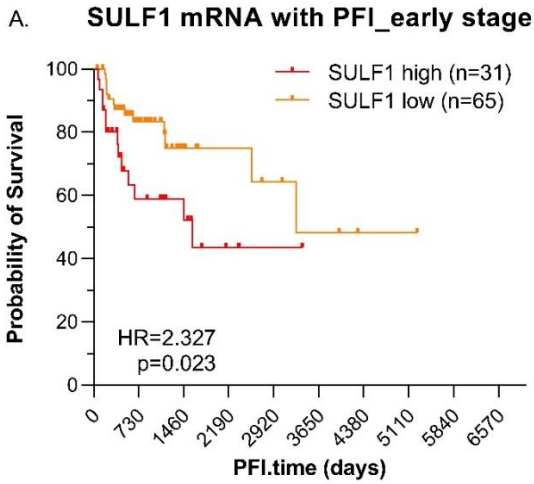
554 SULF1 (a-d) or high SULF2 (e-h) expression is significantly (HR>1.5, FDR<0.05) associated with poor

555 progression free interval (PFI) in the following TCGA cancer studies: a. KIRP, b. PAAD, c. CESC, d.

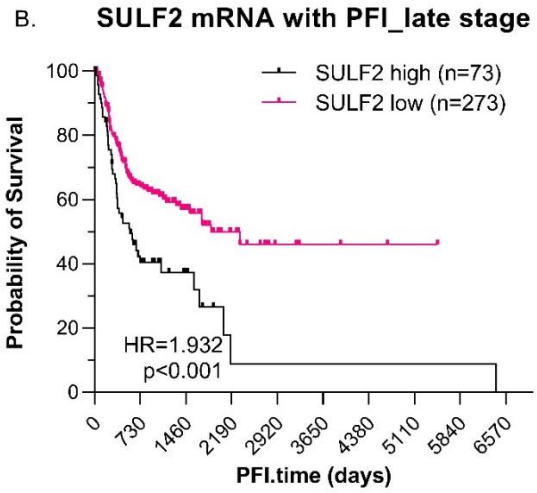
556 COAD for SULF1; and e. KIRC, f. PAAD, g. HNSC, and h. LIHC for SULF2.

557

A.



B.



C.

	Early stage (I&II)		Late stage (III&IV)	
	HR	P-val	HR	P-val
SULF1	2.327	0.023	1.034	0.842
SULF2	1.457	0.866	1.932	<0.001

Figure 3. Impact of SULF1 and SULF2 expression on PFI of HNSC patients differs between early and late stage tumors. (A) SULF1 is associated with PFI in early stage HNSC; (B) SULF2 is associated with PFI in late stage HNSC. (C) Summary statistics of the PFI in the early and late stage tumors

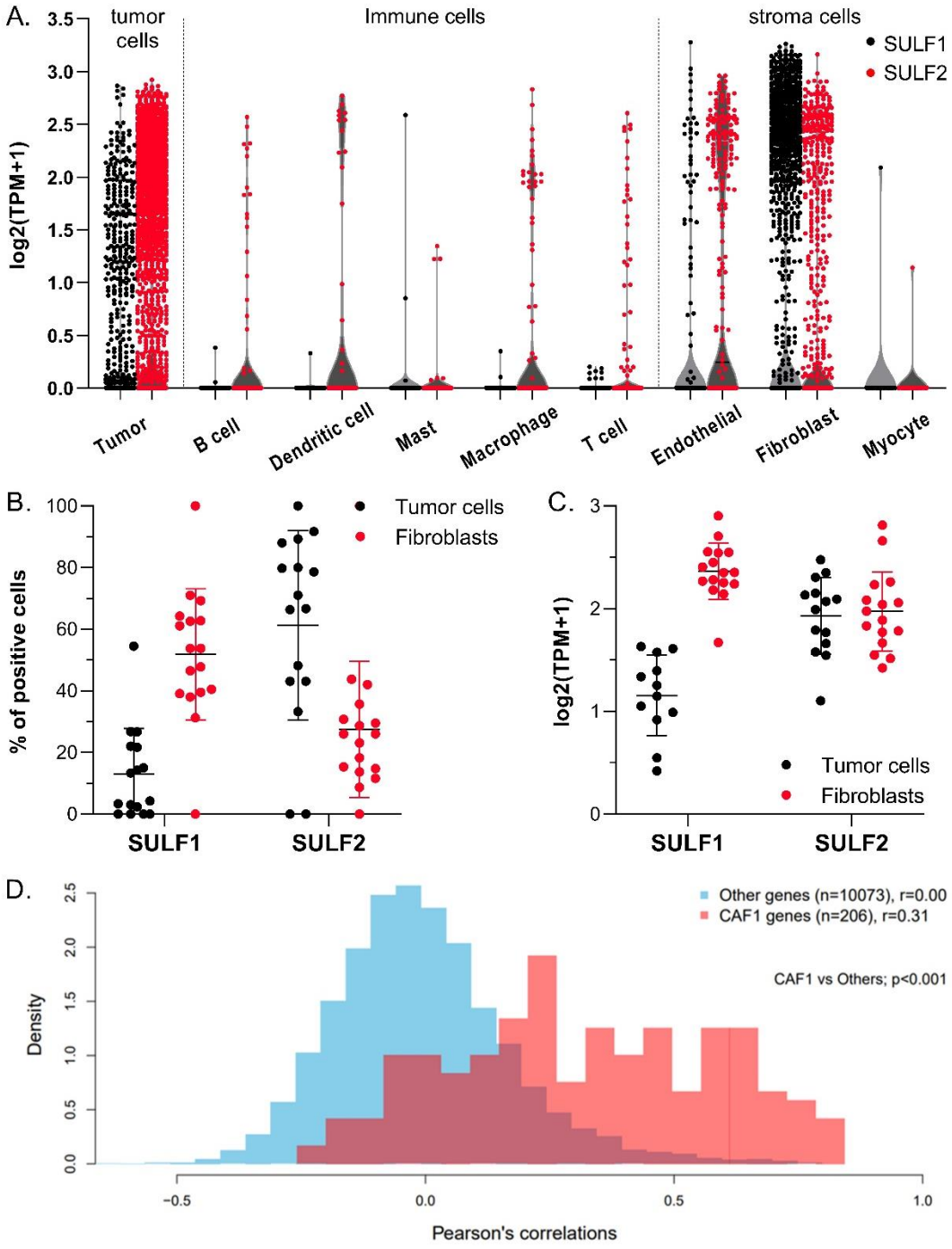
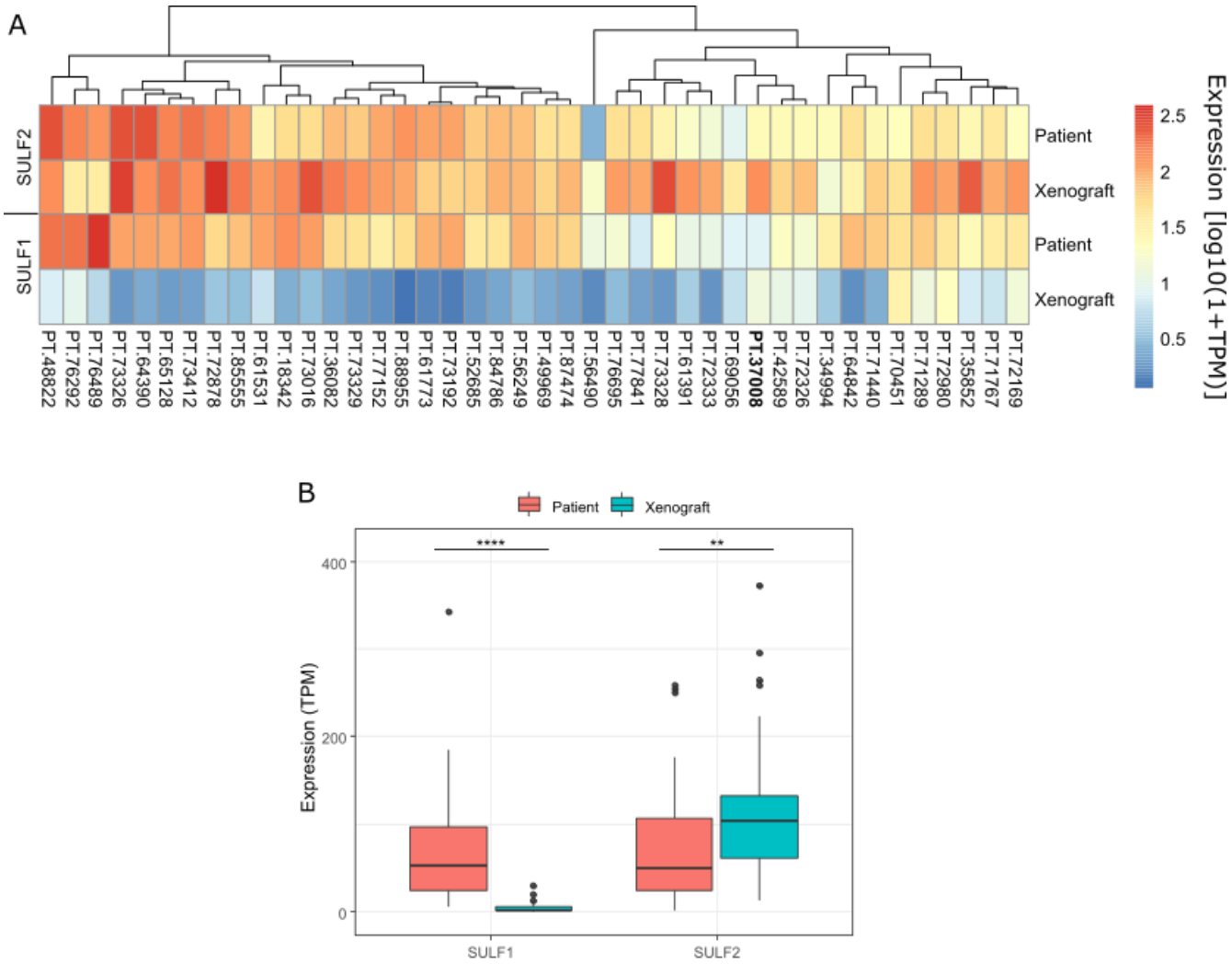


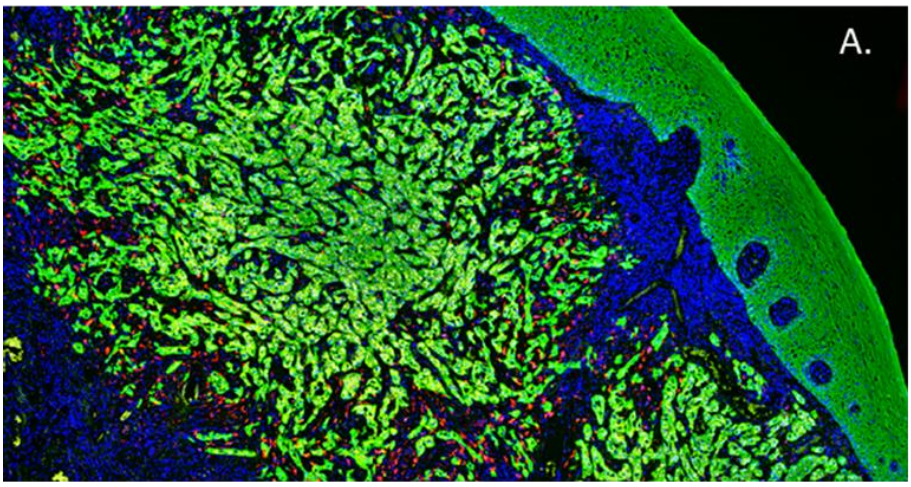
Figure 4. SULF1 and SULF2 are differentially expressed in HNSC epithelial and fibroblast cells. (A) SULF1 and SULF2 mRNA expression in different cell types detected in HNSC tumors; (B) Percent of cells expressing SULF1 or SULF2 differ between the tumor epithelial and fibroblast cells ($p<0.0001$, two-way ANOVA); (C) Expression of SULF1 mRNA in SULF-positive cells differs between tumor epithelial cells and fibroblast ($p<0.001$) but SULF2 does not ($p=0.800$); and (D) Distribution of Pearson's correlation coefficients of SULF1 with CAF1 (n=206) and other (n=10,073) proteins measured in the CPTAC HNSC study[28]. Correlation coefficients of SULF1 protein (red bars) with the CAF1 proteins are significantly ($p<0.001$) higher than the correlations with other proteins. Analyses and the definition of CAF1 are based on a single cell RNA-seq study of 5,578 cells in tumor tissues of 18 HNSC patients[27].



574

575 **Figure 5: Comparison of RNA expression of SULF1 and SULF2 in primary HNSC tumors (n=42)**
576 **and their PDX in immunocompromised mice**: A. Heatmap of log10(1+TPM) expression of SULF1 and
577 SULF2. Patient ID in bold represents the one case in which SULF1 expression does not decrease from
578 patient to PDX. B. Boxplots of SULF1 and SULF2 expression (TPM) in Patient and PDX samples
579 (**p<0.01, ****p<1e-4; Wilcoxon rank-sum test).

580



B.	SULF1+ cells, stroma		SULF2+ cells, tumor	
	Node-negative group	Node-positive group	Node-negative group	Node-positive group
Mean	18%	34%	43%	62%
Median	15%	33%	52%	67%
St. Dev.	15%	12%	31%	25%
p-value	0.02		0.13	

Figure 6. RNAScope of OSCC tumors: A. a tongue cancer stained for DAPI (blue), cytokeratin (green), and with *in situ* probes for SULF1 (red) and SULF2 (yellow). SULF1 localizes to the stroma, SULF2 is expressed mainly in the tumor epithelial cells. Adjacent normal tissue is mostly negative for both SULF1 and SULF2. **B.** comparison of SULF1+ and SULF2+ cell counts in the cancer cell (cytokeratin+) and adjacent stroma (cytokeratin-) of OSCC patients (n=20). Percentage of SULF1+ cells in the stroma is significantly higher (p=0.023) in node positive patients (n=10) with poor survival than in node negative (n=10) patients with good survival. SULF2* cells in the tumor follow a similar trend but the difference is not significant.

TCGA study			SULF1		SULF2	
Project	**Primary name**	**No. Pairs**	**log_2FC**	**FDR**	**log_2FC**	**FDR**
LUAD	lung adenocarcinoma	58	2.78	5.73E-12	0.74	8.52E-04
ESCA	esophageal carcinoma	13	2.76	6.50E-03	2.65	8.52E-04
LUSC	lung squamous cell carcinoma	50	2.62	8.70E-13	1.58	1.31E-08
COAD	colon adenocarcinoma	26	2.55	7.78E-05	0.5	2.12E-02
HNSC	head and neck squamous cell carcinoma	43	2.52	5.20E-07	1.32	3.44E-05
STAD	stomach adenocarcinoma	33	2.46	1.48E-05	1.22	8.52E-04
BLCA	bladder urothelial carcinoma	19	2.14	2.80E-04	0.56	2.75E-01
BRCA	breast invasive carcinoma	112	2.03	1.08E-25	0.82	1.31E-08
KIRC	kidney renal clear cell carcinoma	72	1.31	1.85E-06	0.98	4.47E-07
LIHC	liver hepatocellular carcinoma	50	0.9	1.07E-01	-0.48	1.16E-01
KIRP	kidney renal papillary cell carcinoma	32	0.23	8.75E-01	1.31	4.45E-05
PRAD	prostate adenocarcinoma	51	-0.15	1.07E-01	-0.75	4.45E-05
THCA	thyroid carcinoma	59	-0.51	4.98E-03	0.26	4.01E-02
KICH	kidney chromophobe	25	-0.6	1.23E-02	-0.77	9.36E-02

Table 1. Differential expression of SULF1 and SULF2 mRNA between paired tumor and normal tissues in 14 TCGA studies. Entries with fold-change>2 and FDR<0.05 are in bold.

CPTAC study			SULF1		SULF2	
Project	**Primary name**	**No. Pairs**	**log_2FC**	**FDR**	**log_2FC**	**FDR**
HNSC	head and neck squamous cell carcinoma	68	1.59	1.90E-15	0.5	5.37E-07
BRCA	breast invasive carcinoma	17	1.51	2.80E-05	0.87	2.14E-03
PDAC	pancreatic ductal adenocarcinoma	66	1.49	5.47E-18	1.14	1.45E-16
LUSC	lung squamous cell carcinoma	102	1.35	3.82E-28	0.45	1.62E-11
LUAD	lung adenocarcinoma	100	0.95	5.47E-18	0.19	5.37E-03
COAD	colon adenocarcinoma	96	0.73	2.16E-14	ND	ND
UCEC	uterine corpus endometrial carcinoma	30	0.5	1.69E-03	-0.43	1.09E-02
KIRC	clear cell renal cell carcinoma	84	0.46	4.16E-04	0.26	5.22E-03
HBV-HCC	HBV-related hepatocellular carcinoma	160	0.23	5.69E-02	-0.33	6.83E-07
OSC_JHU	ovarian serous cystadenocarcinoma	12	0.43	7.04E-02	0.44	3.80E-02
OSC_PNNL	ovarian serous cystadenocarcinoma	10	0.07	4.32E-01	-0.58	3.13E-01

Table 2. Differential expression of SULF1 and SULF2 proteins between tumor and adjacent normal tissues in 10 CPTAC studies. Entries with $|(log_2FC)|>1$ and FDR<0.05 are in bold. ND, not detected.

Supplemental Table 1. A. SULF1 and SULF2 mRNA expression across 32 TCGA studies compared to non-disease tissues from GTEx. SULF1 and SULF2 mRNAs are quantified as $\log_2$ (RSEM_counts+1). Differential expressions, represented by $\log_2$FC and p-value, were calculated by non-paired t-test. **B. Association between SULF mRNA expression and PFI outcomes of patients in 20 TCGA studies.** Differential expression of SULFs is represented as $\log_2$FC between the paired tumor and normal tissue of patients in each cancer study. The HR and p-values were calculated by the Kaplan-Meier method and the log-rank test using optimal cutoff of SULF expressions in each cancer type (>100 patients, >40 events for PFI).

Supplemental Table 2: Characteristics of the HNSC patients enrolled at the Princess Margaret Cancer Center, University of Toronto.

Supplemental Table 3. Distribution of SULF1- or SULF2- expressing cells among cell-types analyzed in a scRNAseq study of HNSC tumor tissues. The table summarizes single-cell RNA-seq data of 5578 cells from tumor tissues of 18 HNSC patients[27]. SULF1+ and SULF2+ cells were defined as individual cells with non-zero value of $\log_2$(TPM+1). Number and percentage of SULF1+ and SULF2+ cells in each cell type was calculated based on the 5,578 cells.

Supplemental Table 4. Genes (n=100) with the highest SULF1 Pearson's correlations (FDR<0.05) at the protein level in the CPTAC study[28]. ns, non-significant

Supplemental Table 5: Percentage of SULF1 and SULF2 positive cells in the tumor and stroma of the 20 OSCC patients (A) and corresponding clinical characteristics of the patients (B).