Published online 27 May 2022

NAR Cancer, 2022, Vol. 4, No. 2 **1**
https://doi.org/10.1093/narcan/zcac017

EHF is a novel regulator of cellular redox metabolism and predicts patient prognosis in HNSCC

Akinsola Oyelakin [1], Kasturi Bala Nayak [2], Alexandra Ruth Glathar [2], Christian Gluck [2], Theresa Wrynn [2], Antonio Tugores [3], Rose-Anne Romano [1,*] and Satrajit Sinha [2,*]

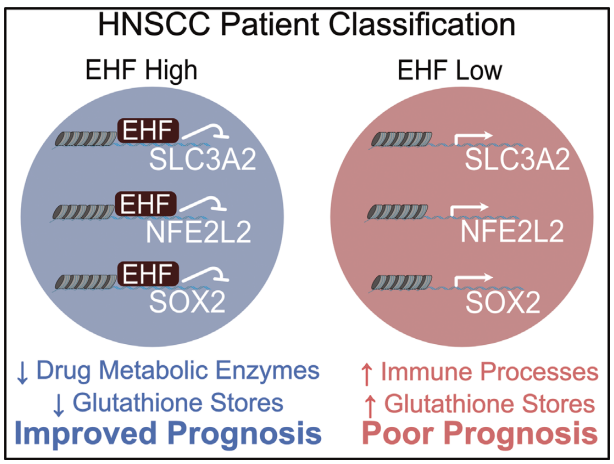
[1]Department of Oral Biology, School of Dental Medicine, State University of New York at Buffalo, Buffalo, NY, USA, [2]Department of Biochemistry, Jacobs School of Medicine and Biomedical Sciences, State University of New York at Buffalo, Buffalo, NY, USA and [3]Unidad de Investigación, Complejo Hospitalario Universitario Insular Materno Infantil Avda Maritima del Sur, Las Palmas de Gran Canaria, Spain

Received May 26, 2021; Revised April 28, 2022; Editorial Decision May 04, 2022; Accepted May 18, 2022

ABSTRACT

Head and Neck Squamous Cell Carcinoma (HNSCC) is a heterogeneous disease with relatively high morbidity and mortality rates. The lack of effective therapies, high recurrence rates and drug resistance driven in part, by tumor heterogeneity, contribute to the poor prognosis for patients diagnosed with this cancer. This problem is further exacerbated by the fact that key regulatory factors contributing to the disease diversity remains largely elusive. Here, we have identified EHF as an important member of the ETS family of transcription factors that is highly expressed in normal oral tissues, but lost during HNSCC progression. Interestingly, HNSCC tumors and cell lines exhibited a dichotomy of high and low EHF expression, and patients whose tumors retained EHF expression showed significantly better prognosis, suggesting a potential tumor suppressive role for EHF. To address this, we have performed gain and loss of function studies and leveraged bulk and single-cell cancer genomic datasets to identify global EHF targets by RNA-sequencing (RNA-seq) and Chromatin Immunoprecipitation and next generation sequencing (ChIP-seq) experiments of HNSCC cell lines. These mechanistic studies have revealed that EHF, acts as a regulator of a broad spectrum of metabolic processes, specifically targeting regulators of redox homeostasis such as NRF2 and SOX2. Our immunostaining results confirm the mutually exclusive expression patterns of EHF and SOX2 in HNSCC tumors and suggest a possible role for these two factors in establishing discrete metabolic states within the tumor microenvironment. Taken together, EHF may serve as a novel prognostic marker for classifying HNSCC patients for actionable and targeted therapeutic intervention.

GRAPHICAL ABSTRACT



INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) cancer incidences have continued their upward trend in recent years (1,2). Patient outlook for this disease is influenced by a number of factors, including stage of presentation, extent of loco-regional involvement, metastasis and HPV status, with the latter portending a better prognosis (2–5). Treatment options for HNSCC remain rather limited and blunt-forced, due to the heterogeneous nature of the disease and the lack of a deep understanding of the tumor biology. Indeed, despite intensive research efforts and the advent of antibody-based therapies, there has only been a marginal improvement in patient outcome (6–9). Hence, there is a dire need to better understand the molecular attributes of HNSCC and to identify patients that are likely to benefit from targeted therapeutic interventions.

*To whom correspondence should be addressed. Tel: +1 716 881 7994; Email: ssinha2@buffalo.edu
Correspondence may also be address to Rose-Anne Romano. Email: rromano2@buffalo.edu

One of the prominent molecular characteristics of tumors of all origin, including HNSCC, is the dysregulated expression and activity of Transcription Factors (TFs). Such TFs are often important for lineage specification, orchestrating faithful developmental programs and maintaining tissue homeostasis (10). However, by co-opting and repurposing these crucial TFs, tumor cells feed their oncogene addiction and direct the aberrant gene expression programs that underlie the myriad hallmark properties of cancers (11–13). This is very well typified by the E26 transformation specific (ETS) family of TFs which regulate a wide spectrum of cellular and molecular processes, including proliferation, differentiation, tissue remodeling and cancer-specific metabolic states, to name a few (14,15). Not surprisingly, altered expression and/or activity of ETS factors is quite commonly and intimately associated with various stages of tumor initiation, progression and metastasis (16,17). In this regard, we and others have shown that ETS1, the founding and prototypical member of the ETS family, plays a key oncogenic role in HNSCC where it directs a mesenchymal-enriched gene expression program (18,19). Similarly, a crucial role for ETS2 and ELK3 has been demonstrated in skin squamous cell carcinoma (SCC), where these two ETS family members appear to be critical in regulating cancer-specific enhancers and in orchestrating the expression of a cohort of oncogenic, growth-promoting, and epi-immune genes (20). Although such studies have highlighted the relevance of these oncogenic ETS factors, it is likely that additional family members also play important roles in HNSCC.

Here, we have utilized a comprehensive repertoire of genomic, epigenomic and genetic tools, to identify the ETS transcription factor EHF, as a novel regulator of HNSCC. We show that expression of EHF is lost in HNSCC compared to normal tissues across multiple large-scale transcriptomic datasets. We also demonstrate that patients with higher levels of *EHF* (*EHF*^{high}) expression have a significantly better prognosis compared to their *EHF*^{low} counterparts. To investigate the molecular function of EHF and identify its transcriptional targets we utilized several preclinical cancer cell line models of HNSCC. Our RNA-seq and ChIP-seq based studies have allowed us to identify a well-defined EHF-driven gene signature that matches well with the gene expression profiles derived from a recently reported scRNA-seq dataset of *EHF*^{high} and *EHF*^{low} HNSCC tumors (21). Notably, we demonstrate EHF acts as a repressor of the cellular redox master regulator *NFE2L2* (NRF2), and the stem cell factor, SOX2 and regulates broad cellular metabolic processes that sustain oncogenic states. Finally, we show that EHF and SOX2 protein expression is heterogenous in nature and localized to mutually exclusive regions in HNSCC tumors. Our studies reaffirm the notion that intratumoral heterogeneity is associated with distinct transcriptional and metabolic states and highlights EHF as a valuable candidate for functional tumor stratification.

MATERIALS AND METHODS

Cell culture

Cell culture conditions for HNSCC cell lines, SCC25 and CAL-27 have been previously described (18), A253 cells were obtained as a gift from Dr. Jill Kramer (School of Dental Medicine, SUNY Buffalo) and were grown in McCoy's 5a Modified Medium supplemented with 5% fetal bovine serum and 1% penicillin/streptomycin. All cell lines used in the study were periodically authenticated by short tandem repeat profiling and tested for mycoplasma contamination.

Plasmids for lentiviral constructs of gene expression and shRNA knockdown

Plasmids. A full-length EHF cDNA encoding the 300 amino acids of the human EHF with an in-frame 3' FLAG tag in the pcDNA3.1^{+}/C- (K)-DYK vector was obtained from Genescript. This plasmid was used as a template to generate a mutant version (Mut) of EHF containing an alteration of a highly-conserved Lysine (K) to Alanine (A) at position 262 in the DNA-binding domain of EHF. The counterpart of this conserved K262 residue in other ETS proteins, such as ELF5 and PDEF has been shown to be critical in mediating DNA binding (22–24). The primers used for 2-step PCR based site directed mutagenesis of EHF are Forward primer 5′-AAC AGC ATG ACC TAT GAA GCG CTC AGC CGA GCT ATG AGA-3′ and Reverse primer 5′-TCT CAT AGC TCG GCT GAG CGC TTC ATA GGT CAT GCT GCT GTT-3′. Wildtype and mutant EHF cDNA sequences were sub-cloned into the BamH1-Xho1 sites in the pLEX (Open Biosystems) or pInducer20 lentiviral plasmids (25) to generate stable or Tet-inducible over-expression systems. Both wildtype and mutant EHF plasmids were confirmed by sequence analysis. Lentivirus-mediated depletion of EHF in CAL-27 cells was performed using the pGIPZ system. GIPZ Lentiviral shRNAs (Clone IDs 3336663, 64 and 65) targeting EHF were purchased from Horizon.

Lentiviral production. Viral packaging into HEK-293T cells was done using the Trans-Lentiviral shRNA packaging system. Forty-eight hours after packaging, lentiviral particles were harvested from cell supernatant, and transduced into target cells. 4μg/ml polybrene was added to the viral supernatant before transduction to prevent the clumping of viral particles and aid transduction efficiency. Stably transduced target cells were selected using 2 μg/ml puromycin (GIPZ and pLEX systems) or 400 μg/ml Geneticin (pInducer20). pInducer20 transduced A253 and SCC25 cells were treated with 0.01, 0.05, 0.1, 0.5 and 1.0 μg/ml doxycycline for 72hrs to induce dose-dependent EHF overexpression.

Clonogenic assay

HNSCC cells were seeded into 6-well plates at a density of 500 cells/well. The cells were maintained in culture for 9 days at 37°C and 5% CO_2 after which the cells were fixed for 2hrs with a 1:3 (v/v) Acetic Acid-Methanol solution, washed twice with phosphate buffered saline (PBS) at pH 7.0, and stained with 2% (w/v) Methylene Blue in 50% Ethanol for 20 minutes. Excess stain was rinsed off with tap water and the plates were air-dried before imaging and counting.

RNA-sequencing analysis

Reads were mapped to the reference human genome (GRCh38/hg19 build) using HISAT2 v2.1.0 (26–28). Reads aligning to the reference genome were quantified with featureCounts v1.5.3 (29) to generate a matrix of raw counts which was then further processed in the R statistical software, to generate normalized expression values in Transcripts per million (TPM) values according to published methods (30). Differential gene expression analysis comparing control to test condition was carried out using DESeq2 v1.24.0 (31). Differentially expressed genes were called at a false discovery rate (FDR) value of 0.1 or lower.

TCGA analysis

TPM normalized values of the TCGA dataset (GSE62944) (32) for both normal and tumor samples were obtained from the GEO website. The data was imported into R and filtered to retain the HNSC dataset. The 502 tumor and 44 matched-normal data matrix files were further filtered to retain only the expression values of the 28 members of the ETS family of transcription factors ordered in decrease average expression across the tumors. The resulting matrix was then used to generate a heatmap.

HNSCC patient survival analysis

The R package, RTCGA v1.16.0 was used to obtain the clinical data for HNSC patients. The clinical data was merged with the expression data, and ranked by decreasing EHF expression. The package survminer v0.4.8 was used for Kaplan Meier analysis and to generate the survival plot. Optimal cut point between EHF^{high} and EHF^{low} patients was determined using the R package, Evaluate Cutpoint (33).

ChIP-sequencing analysis

CAL-27 cells were grown to 80% confluency in 150 mm culture dish and cross-linked with 1% formaldehyde for 10 min. The cross-linking reaction was quenched with Glycine to a final concentration of 125 mM. Fixed CAL-27 cells were prepared for chromatin shearing using Diagenode low sodium dodecyl sulphate (SDS) shearing kit for Transcription Factors (Diagenode) and sonicated using a Diagenode Bioruptor to obtain sheared chromatin of between 200 and 400 bp. EHF immunoprecipitation was done using 2 μg of anti-EHF antibody (Diagenode). Following immunoprecipitation, DNA–protein complex was then subjected to cross-link reversal and proteinase K treatment. Input DNA and DNA from immunoprecipitation experiments were purified and concentrated using Qiagen MinElute kit. DNA libraries were prepared from concentrated DNA using ThruPLEX DNA-seq kit (Rubicon Genomics), and single-end sequencing was done on an Illumina NovaSeq 6000. The sequencing reads from all experiments were mapped to the Homo sapiens genome (hg19 build) using Bowtie v1.1.1 (34) with the parameter m = 1 to remove all reads mapping to multiple genomic loci. Peak calling was then performed using MACS2 v2.0.10 (35,36) with a minimum FDR cutoff of 0.05 and resultant peaks were matched to the nearest gene using Genomic Region Enrichment Annotation Tool (GREAT) v3.0.0 (37) with default settings (basal extension). For visualization of ChIP peaks, deeptools v3.3.2 (38) was used to generate bigwig files which were then uploaded into the Integrative Genomics Viewer (IGV, Broad Institute).

Motif analysis

De novo motif analysis on the 20,273 EHF genomic binding sites was preformed using MEME-ChIP (39). The resulting enriched motifs were compared against the Homo sapiens Comprehensive Model Collection (HOCOMOCO) database (40) to identify matching transcription factor motifs using TomTom (41). The Input sequences for MEME-ChIP analysis consisting of 500 bases of genomic sequences centered on the EHF ChIP-seq binding summits were generated using BEDtools (42).

Microarray analysis

Microarray datasets GSE31056 (43) and GSE30784 (44) were downloaded from the Gene Expression Omnibus (GEO) using the R Bioconductor package affy v1.62.0. Normalization was performed on each dataset before the generation of graphs.

Analysis of single-cell RNA-seq dataset of HNSCC

TPM normalized data from GSE103322 (21) was reanalyzed using Seurat v3.2.2 (45). Briefly, the pre-annotated files were uploaded into R, filtered to retain only the tumor (or malignant) cells and converted into a Seurat object for downstream clustering analysis. We retained 15 of the patient samples that were sequenced, discarding samples with <200 cells sequenced per patient. We evaluated the distribution of *EHF* expression and performed differential gene expression analysis based on EHF^{high} and EHF^{low} clusters.

Immunohistochemistry

Tumor microarray slides were deparaffinized and sequentially rehydrated in reducing concentrations of ethanol in water. After heat-induced antigen retrieval in sodium citrate, tissue slides were incubated overnight according to the standard IHC protocol with primary antibodies specific to: EHF (5A5.5) (46) and SOX2 (D1C7J, Cell Signaling Technologies) and subsequently stained with the Impact DAB substrate kit (SK-4105, Vector labs). Counterstaining was done using hematoxylin (H-304, Vector labs) after which the slides were rinsed in tap water, air dried and the coverslip was mounted using permount (Fisher).

Western blot analysis

Whole cell lysates were prepared from cells grown to 80-90% confluency by adding Laemmli Sample Buffer (BioRad) and collecting the lysates. The proteins in the lysates were then denatured at ∼100°C for 10 minutes. Equal amounts of protein extracts were loaded in SDS-polyacrylamide gels and subjected to electrophoresis, followed by transfer to Immunoblot polyvinylidene fluoride (PVDF) membranes (BioRad). Immunoblotting was performed using

the following antibodies: EHF (5A5.5); SOX2 (D6D9, Cell signaling technologies); GAPDH (MAB374, EMD Millipore); β-Tubulin (MAB3408, EMD Millipore); AKR1C1 (CPTC-AKR1C1-1, DHSB); AKR1C2 (CPTC-AKR1C2-1, DHSB); NRF2 (D1Z9C, cell signaling technologies); ELF3 (MAB5787, R&D), SLC3A2 (15193-1-AP, Proteintech) and GPX2 (MAB5470-SP, R&D). Membranes were washed off in 0.05% Tween-20 in Tris Buffered Saline (TBS-T). The membranes were next incubated with species-specific secondary antibodies, and washed in TBS-T. The Enhanced Chemi-Luminescence (ECL) method was used to detect fractionated proteins, and the membranes were either electronically imaged using BioRad ChemiDoc imaging system (BioRad), or bioblot radiograph films (LPS).

Cellular glutathione quantification

GSH/GSSH ratio was determined using the GSH/GSSG-Glo assay kit (Promega, WI).

Briefly, 5000 cells were seeded in triplicates into 96-well plates and incubated for 24–48 h. At the time of the assay, growth media was aspirated and replaced with either total glutathione lysis buffer or oxidized glutathione lysis buffer. To achieve efficient lysis, the plate was placed on a plate shaker for 5 min at room temperature. Next, Luciferin generation reagent was added to each well, and the plate was incubated for 30 minutes at room temperature. Finally, Luciferin detection reagent was added to each well and the resulting luminescence was quantified using a Cytation 1 Imager (BioTek, VT). Graphs were plotted using Microsoft Excel (Microsoft, WA).

Statistical analysis

Statistical and bioinformatics analyses were performed on the R statistical platform. Data shown within this manuscript were derived from representative experiments. Three or more experimental replicates were performed. Error bars in bar graphs in Figure 2C denote standard deviations (SD) for four biological replicates for each experiment. Student's t-test for unpaired data was performed for Figure 1C.

RESULTS

EHF expression is downregulated in a subset of HNSCC tumors and confers better clinical outcome

To determine the expression pattern of the 28 members of the human ETS family of genes, we queried the HNSC dataset of The Cancer Genome Atlas (TCGA) consisting of 502 cancer and 44 matched-normal tissues. Expression-based ranking of the ETS family members revealed *ETS2*, *ELF3* and *EHF* to be the most highly expressed in cancer tissues as well as in normal HNSCC samples (Figure 1A). However, while the expression of *ETS2* was generally uniform across all cancer samples analyzed, the expression of *ELF3* and *EHF* was highly variable (Figure 1A). We modeled the density distribution of *ELF3* and *EHF* due to their variable expression, comparing their cancer to normal expression patterns and determined that EHF expression, in contrast to *ELF3* expression, was attenuated in a larger cohort of the cancer samples while normal tissues expressed very high levels of EHF (Figure 1B). This differential expression of *EHF* was easily discernible when 44 normal samples were compared to their corresponding matched cancer tissues (Supplementary Figure S1A). In agreement, the loss of *EHF* expression has also been shown to be highly correlated with malignancy in other epithelial-rich tumors (46–48). We next examined additional microarray-based datasets of HNSCC tumors (43,44) which confirmed that *EHF* expression was progressively lost in dysplasia and cancer, validating our observations from the TCGA datasets (Figure 1C). To determine the clinical significance of dysregulated *EHF* expression, we performed Kaplan-Meier survival analysis by comparing the outcome of patients with high *EHF* expressing tumors (EHF^{high}) to those with low *EHF* expressing tumors (EHF^{low}). This analysis revealed that patients with EHF^{high} tumors were associated with better overall survival suggesting a tumor suppressor role for *EHF* (Supplementary Figure S1B).

Enforced EHF expression in HNSCC cells blocks clonogenicity

The variable expression of *EHF* in the tumor samples prompted us to examine whether HNSCC cell lines can also segregate into similar EHF^{high} and EHF^{low} groups and thus serve as appropriate models for loss-of-function and gain-of-function studies. To address this, we performed western blot experiments on several well-established HNSCC cell lines representing different molecular subtypes and anatomical origins. We found EHF expression to be quite variable across the HNSCC cell lines with some cell lines such as CAL-27 expressing robust levels of EHF and others either modest or low levels (Figure 2A). As a control, we also performed immunoblot for a closely related ETS family member, ELF3, which showed rather uniform expression across these HNSCC cell lines. Armed with this information, we decided to first perform gain-of-function studies of EHF in two cell lines, A253 and SCC25 which respectively expressed modest or low levels of endogenous EHF, while expressing high levels of ELF3. Towards this end, we established two lentiviral based systems, a doxycycline inducible (pInducer) and stable (pLEX) system to drive the expression of a FLAG-epitope tagged version of EHF in both A253 and SCC25 cell lines. To determine whether the possible cellular and molecular outcomes of EHF overexpression are dependent on the direct engagement of EHF with its genomic targets, we also generated a DNA-binding mutant (EHF-Mut) version of FLAG-EHF construct containing an Alanine substitution for a key Lysine residue in the ETS DNA-binding domain.

Robust and comparable expression of EHF and EHF-Mut proteins were detected in the stably infected A253 and SCC25 cell lines (Figure 2B). We also evaluated the toxic effect of the overexpressed proteins in both cell lines by presence of cleaved caspase 3. Our analysis showed that overexpression of either EHF or EHF-Mut was not toxic to the cells (Figure 2B). A previous study showed that forced expression of EHF leads to a block in clonogenicity of oral cancer cells (49). To confirm and validate this finding, we

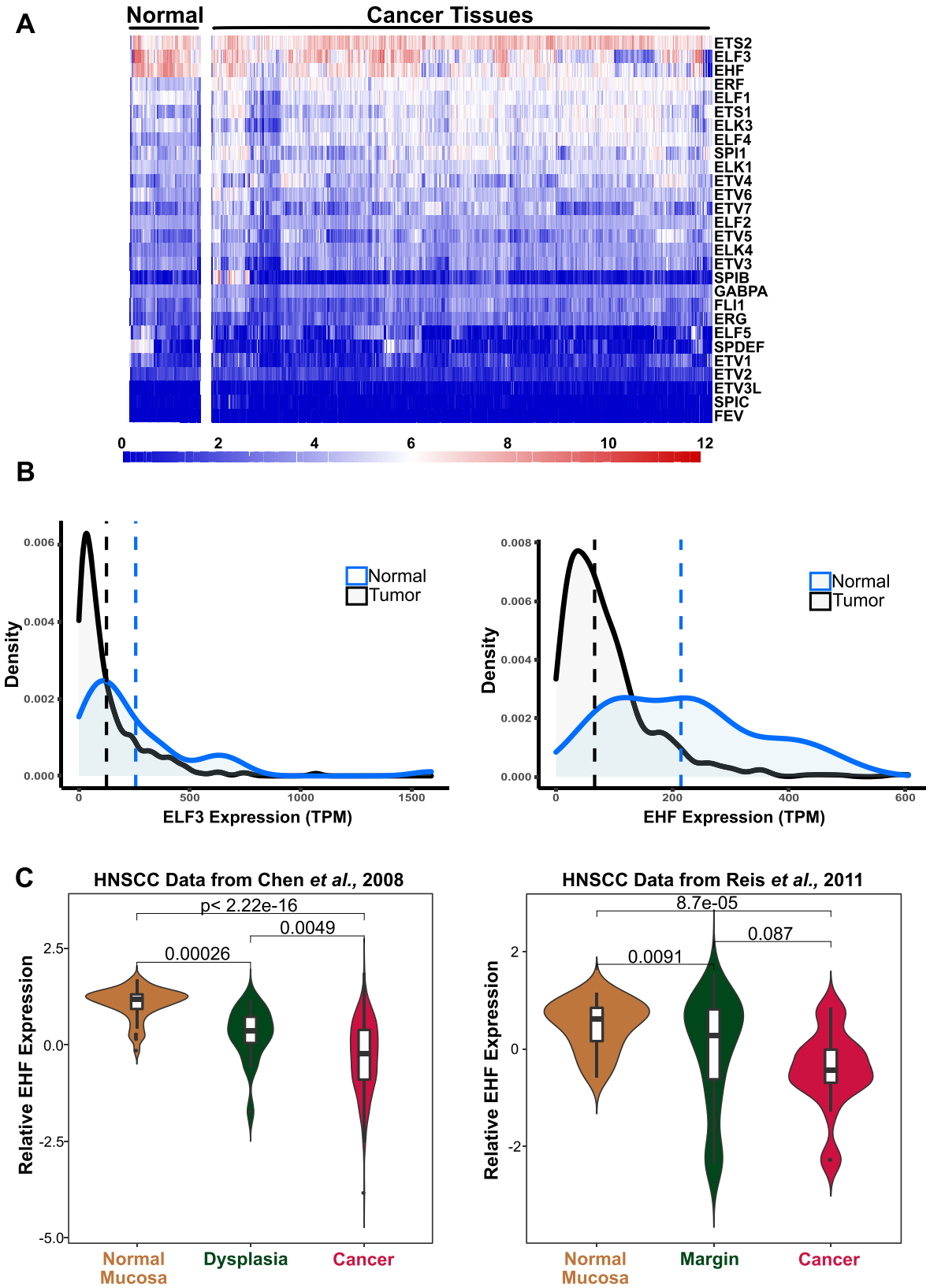


Figure 1. Expression of ETS Family in HNSCC. (**A**) Heatmap showing the expression of 28 ETS family members in the TCGA-HNSC cancer and matched normal dataset. (**B**) Density distribution plot comparing the range of expression of the two ETS family members (*ELF3* and *EHF*) in cancer and matched normal TCGA-HNSC tissues. Dotted lines mark the median expression values for both cancer (black) and normal (blue) tissues. (**C**) Violin plots showing the loss of *EHF* expression during HNSCC progression from two independent datasets. Statistical significance was determined using unpaired t-test.

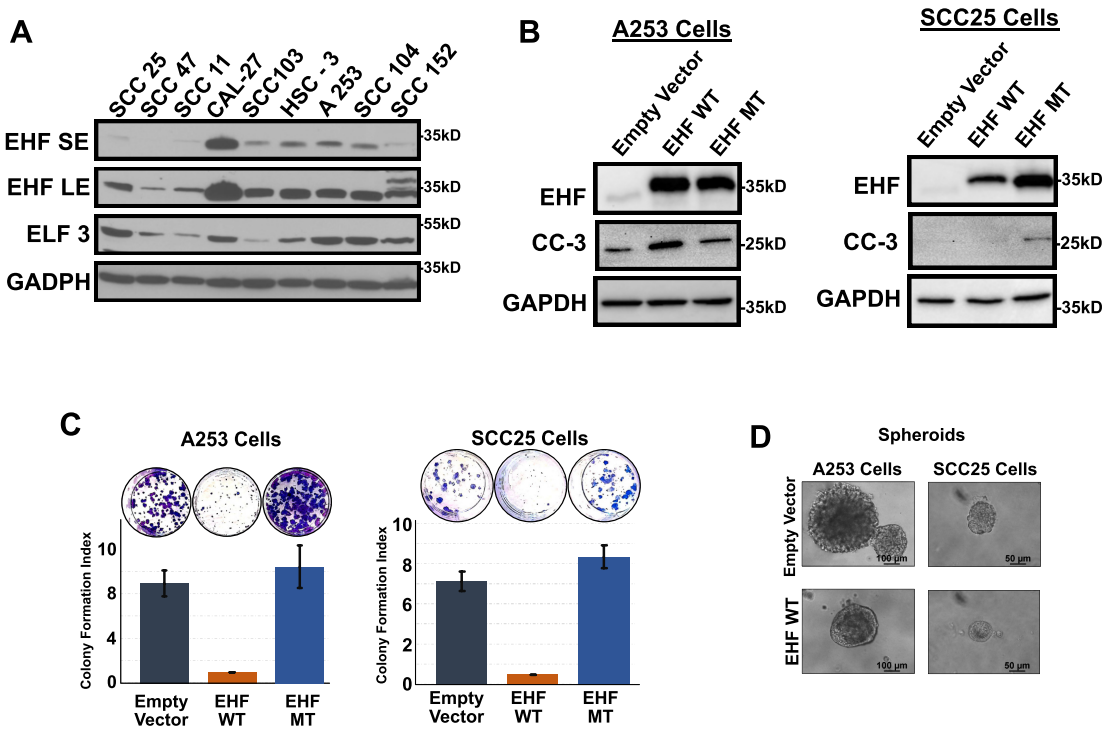


Figure 2. EHF Expression Patterns and Functional Analysis of EHF Overexpression in HNSCC Cell Lines. (**A**) Western blot showing EHF and ELF3 expression across a panel of model HNSCC cell lines. LE and SE indicate long and short exposures respectively. (**B**) Western blot showing the induced expression of wildtype and DNA-binding deficient EHF in two HNSCC cell line models. (**C**) Wildtype EHF but not EHF-Mut blocks clonogenic potential in HNSCC cells ($n = 4$ replicates) (**D**) Tumor spheroids obtained from cells grown on Matrigel. Overexpression of EHF blocks the ability of both A253 and SCC25 cells to form spheroids. Bar graph data is presented as mean ± standard deviation (error bars).

performed clonogenic assays on stable EHF expressing cell lines. As shown in Figure 2C, induction of EHF greatly impaired the ability of these cells to form colonies. The ability of cancer cells to form tumors *in-vivo* can be determined from their ability to form spheroids *in-vitro*. We evaluated the effect of EHF and EHF-Mut expression on tumor spheroid formation. Our analysis showed that the expression of wildtype EHF impaired tumor formation in both A253 and SCC25 cells (Figure 2D).

EHF regulates genes involved in cellular metabolism and immune response

To investigate the functional relevance of EHF overexpression, we next turned to our inducible EHF expression system. First, we confirmed that EHF over-expression could be modulated in a dose-dependent fashion upon administration of doxycycline to the cell culture media (Figure 3A). We next carried out RNA-sequencing based studies using the SCC25 inducible system in which EHF and EHF-Mut were expressed for a period of 48 hours. We reasoned that the short-term induction of EHF would better capture the direct transcriptomic changes rather than those that are due to secondary effects. We observed robust changes in the gene expression profile with 999 upregulated and 978 downregulated genes identified as differentially expressed at a false discovery rate (FDR) of 0.1 upon inducible expression of EHF (Figure 3B & Supplementary Table S1). Notably, under similar conditions, inducing EHF-Mut expression (Supplementary Figure S2A) led to minimal changes in gene expression when similar statistical cut-offs were applied for the differential gene expression analysis, further confirming that the EHF-Mut was transcriptionally deficient (Supplementary Figure S2B). To better understand the functional consequences of EHF overexpression, we performed gene set enrichment analysis on the EHF overexpression dataset using the hallmark gene set from the molecular signature database, MSigDB (50). Our analysis revealed the enrichment of glycolysis and other metabolic processes, cell cycle, DNA repair, oxidative stress response, drug metabolism and epithelial-to-mesenchymal transition (EMT) among the top cellular processes that are downregulated by EHF expression (Figure 3C). We also observed the positive enrichment of genes involved in immune response and inflammatory processes in response to EHF overexpression (Figure 3D). This finding was interesting, given the reported role of EHF in modulating the response of immune checkpoint inhibitors in pancreatic cancer (51) and suggested that EHF might also influence the interactions between the tumor epithelial cells and the immune microenvironment in HNSCC.

To account for any effects resulting from intrinsic molecular and functional attributes that are inherent to a specific HNSCC cell line, and to establish a broader consensus list of potential EHF targets, we repeated the RNA-seq experiments using A253 cells in which EHF was induced after doxycycline treatment for a period of 48 hours (Figure 3E). In this case, we observed 623 upregulated genes and 372 downregulated genes (FDR of 0.1), resulting from

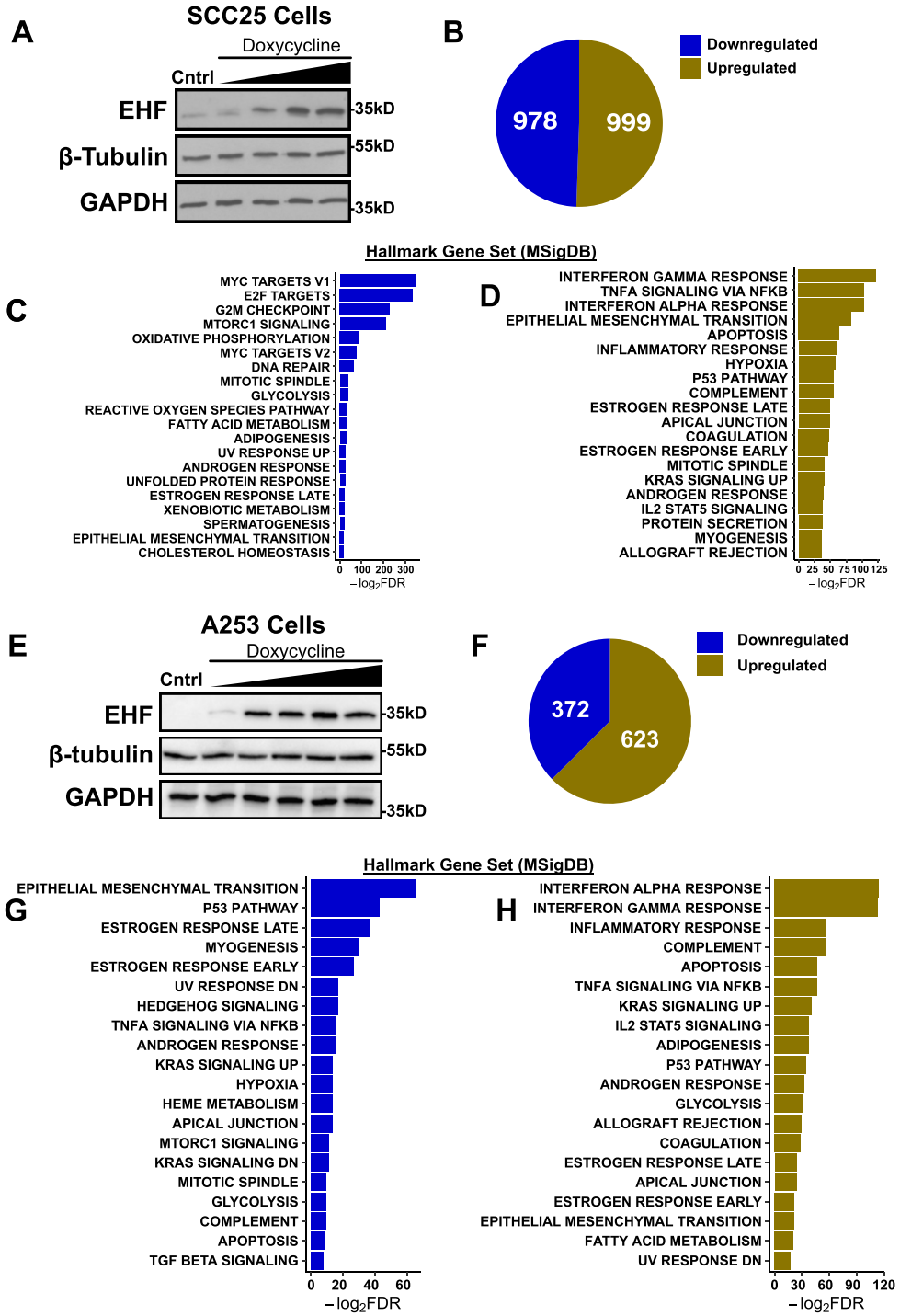


Figure 3. Global Changes in Gene Expression upon EHF Overexpression. (**A**) Western Blot showing dose dependent induction of EHF in SCC25 cells by increasing amounts of of doxycycline treatment. (**B**) Pie chart showing the number of differentially regulated genes upon inducible EHF overexpression in SCC25 cells. (**C**) Gene set enrichment analysis showing processes downregulated by induction of EHF in SCC25 cells. (**D**) Gene set enrichment analysis showing processes upregulated by EHF induction in SCC25 cells. (**E**) Western Blot showing dose dependent induction of EHF in A253 cells by increasing amounts of doxycycline treatment. (**F**) Pie chart showing the number of differentially regulated genes upon inducible EHF overexpression in A253 cells. (**G**) Gene set enrichment analysis showing processes downregulated by induction of EHF in A253 cells. (**H**) Gene set enrichment analysis showing processes upregulated by EHF induction in A253 cells.

EHF overexpression, when compared to control cells that were not treated with doxycycline (Figure 3F & Supplementary Table S1). Gene set enrichment analysis on the downregulated genes showed an enrichment of genes involved in the tumor suppressive p53 signaling pathway, hypoxia, glycolysis and cellular response to ultraviolet radiation (Figure 3G). Conversely, we observed the enrichment of genes involved in inflammatory response and complement systems in the genes upregulated by EHF, similar to what was observed for SCC25 cells (Figure 3H). By integrating results from both cell lines, we identified 351 differentially expressed genes (DEGs) that were common to both datasets (Supplementary Figure S3A). These core EHF-regulated genes largely recapitulated the positive enrichment of an immune response program with the expression of EHF (Supplementary Figure S3B), while processes involved in EMT, mTORC1 signaling, and cellular response to ultraviolet radiation were down regulated (Supplementary Figure S3C). Interestingly, we also noted the positive enrichment of genes involved in drug metabolism, fatty acid oxidation, glycolysis and hypoxia in the upregulated genes (Supplementary Figure S3C).

We reasoned that since overexpression of EHF resulted in the dysregulation of specific pathways, then loss of EHF will likely affect similar pathways and thus validate our overall findings. To test this hypothesis, we used a lentiviral based system to infect CAL-27, an EHF^{high} cell line, with three EHF-targeting shRNAs and a non-targeting shRNA as a control (cntrl). Efficient and specific depletion (KD) of EHF was achieved by two independent shRNAs, sh2 and sh3, as evidenced by western blot analysis which showed robust loss of EHF expression (Figure 4A). Next, we performed RNA-seq experiments to profile the global transcriptomic changes that are unleashed upon the loss of EHF expression. By comparing control cells to EHF-depleted CAL-27 cells, we identified 847 upregulated and 654 downregulated genes that were common between both shRNA 2 (EHF KD1) and shRNA 3 (EHF KD2) (Figure 4B & Supplementary Table S1). Gene set enrichment analysis showed that the loss of EHF led to the downregulation of genes involved in interferon gamma response, p53 signaling, and several metabolic processes (Figure 4C), while the upregulated genes showed enrichment of epithelial to mesenchymal transition (EMT), cell cycle, inflammatory response and cellular response to ultraviolet radiation (Figure 4D). Since the overexpression of EHF decreased clonogenic properties in A253 and SCC25 cells, we anticipated that the loss of EHF would have the opposite effect in CAL-27 cells. Interestingly, CAL-27 cells with attenuated levels of EHF did not show any measurable difference in clonogenicity when compared to control cells (Figure 4E). We suspect that this lack of phenotype in CAL-27 cells might be in part due to the incomplete loss of EHF.

Genome-Wide analysis of EHF binding sites uncovers targets of EHF in HNSCC

To identify the transcriptional targets of EHF, we generated a genome-wide EHF binding profile by ChIP-sequencing using ChIP-grade antibodies directed against EHF in CAL-27 cells which resulted in the identification of 20,273 EHF binding sites (Figure 5A & Supplementary Table S2). *De novo* motif analysis using MEME-ChIP revealed that a significant number (68%) of the binding sites contained the core ETS DNA-Binding motif as the most enriched motif (Figure 5B). In addition to the EHF motif, a number of co-occurring motifs were also identified, these matched with DNA binding elements for AP1, BACH, NFE2L2, and E2F transcription factors amongst others (Supplementary Figure S5). We next performed gene ontology analysis using the Genomic Regions Enrichment of Annotation Tools (GREAT) software (37), which revealed that EHF binds regulatory regions in the neighborhood of genes involved in growth factor signaling, oxidoreductase and dioxygenase activities among other processes (Figure 5C). To examine the distribution pattern of EHF binding sites across the genome, we examined the location of each binding site relative to the transcription start sites (TSS) of genes, and found the enrichment of EHF peaks at regulatory regions distal to the TSS (Figure 5D). Next, we mapped the EHF binding sites to the nearest genes, which resulted in the identification of 11,506 genes that were deemed to be likely direct targets of EHF. Notably, a fraction of the EHF-bound genes (1150 genes) was differentially regulated upon the loss of EHF expression in CAL-27 cells based on RNA-seq results (Figure 5E). Finally, gene set enrichment analysis performed on the top 500 up and downregulated EHF targets in CAL27 cells showed enrichment of EMT, TNF-alpha signaling and glycolysis in the upregulated, and estrogen response, interferon signaling and fatty acid metabolism in the downregulated targets, respectively (Figure 5F). Taken together, these results indicate that EHF is bound to a large number of genomic sites in CAL-27 cells and that it likely mediates transcriptional control of broad range of gene expression programs that are important in HNSCC biology.

A Core EHF-signature reveals key genes involved in metabolic processes as direct EHF targets

We performed a search for a stringent EHF-dependent gene-signature by integrating all the EHF RNA-seq datasets generated from the 3 representative HNSCC cell-lines used in this study (Supplementary Figure S4A). This analysis led to the identification of 132 DEGs common among the three cell lines that were affected by perturbed levels of EHF. After taking into account directionality of gene expression changes, 49 genes survived as members of a core EHF signature. Interestingly, most of the 49 genes were directly regulated by EHF based on ChIP-seq data from the CAL-27 cell line (Supplementary Figure S4B) and predominantly represented key players in metabolic processes such as Reactive Oxygen Species (ROS) scavenging pathways (*SLC3A2, THBS1, SLC9A3R1, AKR1C2, LCN2, ABCG2*) (52–54), glucose metabolism (*SOX2*) (55) and xenobiotic metabolism (*AKR1C1, AKR1C2, SULF2*) (56).

scRNA-seq analysis of primary HNSCC substantiates a role for EHF in regulating cellular redox metabolism *In-vivo*

Our molecular and genomic studies have thus far strongly hinted at a functional role for EHF in broadly regulating metabolic programs in several HNSCC cell lines. To extend

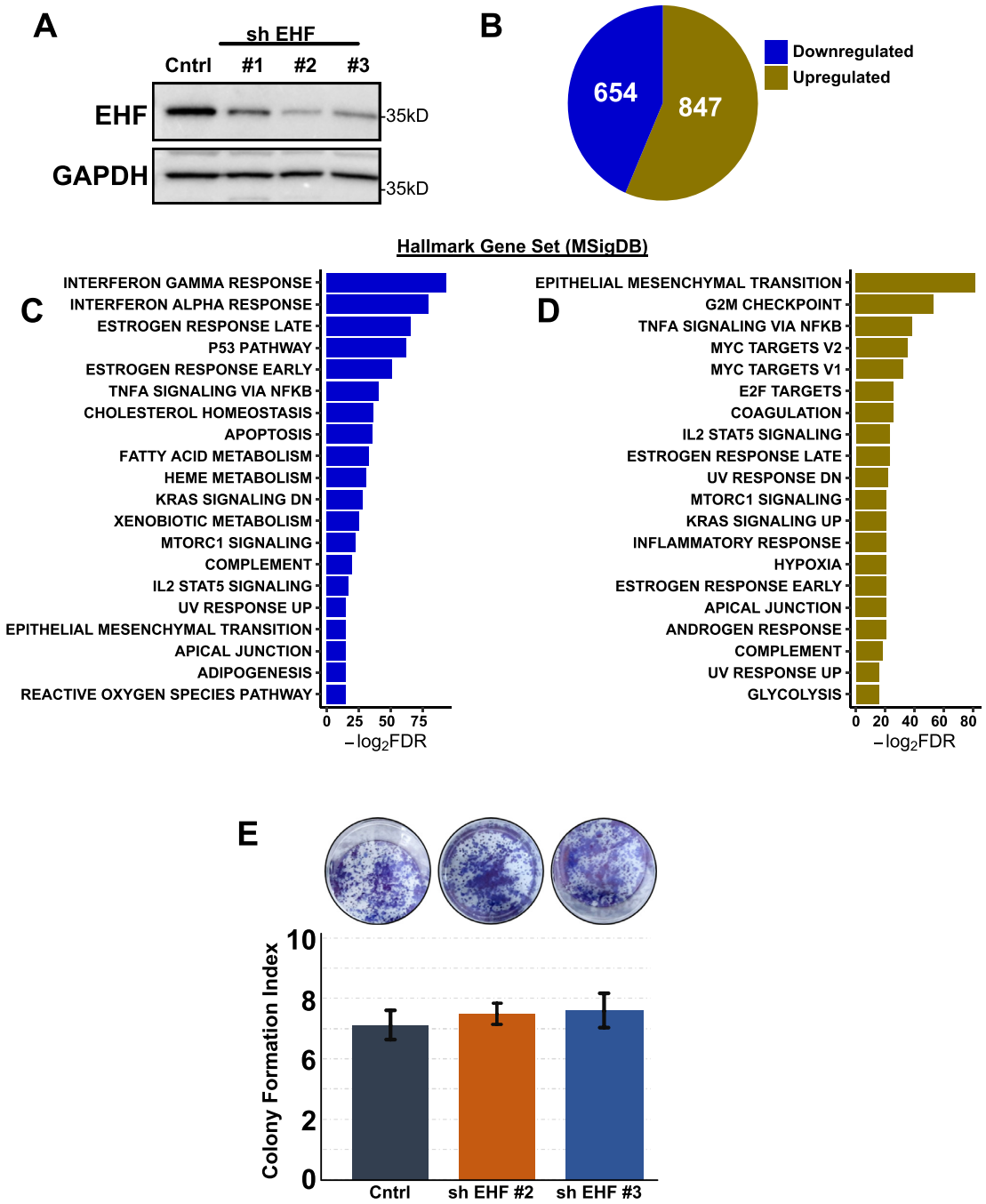


Figure 4. Global Changes in Gene Expression upon Loss of EHF. (**A**) Western blot showing the shRNA mediated downregulation of EHF in CAL-27 cells. (**B**) Pie chart showing the number of dysregulated genes following the loss of EHF. (**C**) Gene set enrichment analysis showing processes downregulated by induction of EHF in CAL-27 cells. (**D**) Gene set enrichment analysis showing processes upregulated by induction of EHF in CAL-27 cells. (**E**) Loss of EHF blocks clonogenicity in CAL-27 cells. (**F**) Loss of EHF blocks spheroid formation in CAL-27 cells

these findings to HNSCC patients, and obtain clinically relevant *in-vivo* data, we next examined the scRNA-seq dataset recently generated from primary HNSCC tissues (21). We reasoned that unlike the bulk RNA-seq data obtained from TCGA in which aggregate expression levels of *EHF* represent the multitude of diverse cell populations, including those residing in the tumor microenvironment, the scRNA-seq dataset could be utilized to specifically parse out the malignant cells of the tumors. We performed Uniform Manifold Approximation and Progression (UMAP) dimension reduction analysis on the malignant cells which reduced 15 independent patient samples to 11 clusters based on similarities in gene expression (Figure 6A). Interestingly, these 11 malignant clusters could be separated into two groups, based on *EHF* expression revealing three (3) EHF^{low} and eight (8) EHF^{high} clusters (Figure 6B), thus confirming the dichotomy we observed in the TCGA dataset. We next carried out differential gene expression analysis between these

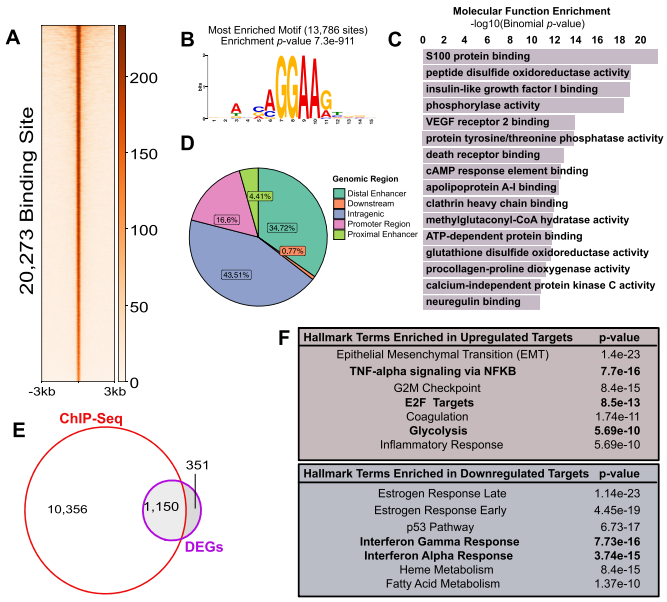


Figure 5. *EHF* ChIP-sequencing Analysis. (**A**) Heatmap depicting global enrichment of ChIP-seq signal centered on *EHF* binding loci in CAL-27 cells. ChIP-seq signal intensity is shown by color shading. (**B**) *de-novo* motif analysis identified core ETS motif as the top, most enriched motif across EHF binding sites. (**C**) Bar plot showing the molecular functions associated with genes that are located near the regulatory regions bound by EHF. (**D**) Pie chart showing the distribution of EHF binding sites across the genome in reference to genes. (**E**) Venn diagram showing the overlap between EHF binding and the DEGs from the loss of EHF in CAL-27 cells. (**F**) MSigDB Hallmark gene set enrichment analysis showing the terms enriched in the top 500 upregulated (top panel) and top 500 downregulated (bottom panel) EHF targets.

two groups and identified genes that were enriched in one cluster compared to the other (Supplementary Table S1). Gene Ontology analysis using several databases revealed the enrichment of energy and drug metabolic processes among the top pathways in the *EHF*^{low} clusters (Figure 6C, E and G). Conversely, the *EHF*^{high} clusters showed an interesting enrichment of activated immune processes (Figure 6D, F and H). These independent analyses reaffirms that the loss of EHF is associated with the enrichment of a specific gene expression program that impinged primarily on tumor metabolic processes.

EHF represses key players in the cellular redox homeostatic program in HNSCC cells

The significant enrichment of genes involved in drug metabolism and NRF2 pathway, which are important processes that regulate redox homeostasis prompted us to probe further into these pathways. First, we focused on Nuclear Factor Erythroid 2 Like 2 (*NFE2L2*) also called *NRF2*, which is an established master regulator of signaling pathways involved in ROS detoxification and cytoprotection, and drives the expression of many genes that are important for these processes (57,58). Given that NRF2 pathway genes were enriched in the *EHF*^{low} cluster (Figures 6G and Figure 7A), we evaluated their expression pattern in our cell-line models. Interestingly, as shown in the heatmaps the NRF2 pathways genes showed an overall trend of downregulation upon *EHF* overexpression in both A253 and SCC25 cells, and increased expression in CAL-27 cells following the loss of EHF (Figure 7B). We followed up on these results by western blot analysis for NRF2 in our inducible and stable overexpression models. Our results showed that the regulation of NRF2 by EHF is somewhat context-dependent, with the short-term induction of EHF resulting in the depletion of NRF2 in both A253 and SCC25 cells (Figure 7C). Stable overexpression of EHF, however resulted in consistent NRF2 repression only in SCC25 cells suggesting that a possible cell-intrinsic mechanism might be at play in A253 cells (Figure 7C). As expected, the overexpression of EHF-Mut did not have any effect on NRF2 protein expression. Consistent with the aforementioned overall repressive effect of EHF on the broad NRF2 regulome, CAL27 cells with diminished EHF levels resulted in the upregulation of NRF2 protein levels (Figure 7D). Finally, to investigate whether EHF directly regulates *NRF2*, we examined our EHF ChIP-seq data and identified binding sites for EHF at the promoter as well as the intragenic regions of the *NRF2* gene suggesting a possible direct mode of regulation (Figure 7E).

The enrichment of drug metabolism in the *EHF*^{low} clusters prompted us to take a closer look at the genes enriched in these processes. We evaluated the expression profile of the drug metabolism genes in our cell line model, and in agreement with the patient data, a good number of these genes were downregulated in both A253 and SCC25 cells following the expression of EHF, but upregulated in CAL-27 cells when EHF is depleted (Figure 8A). We identified three genes, belonging to the aldo-keto reductase (AKR) family showing consistent expression across all three cell lines. We confirmed the expression two of the genes by western blot and showed that EHF blocks the expression of both AKR1C1 and AKR1C2 (Figure 8B). Interestingly, both AKR1C1 and AKR1C2 have been previously reported to mediate cisplatin resistance in head and neck and other cancers (56,59). We also noticed the enrichment of genes involved in glutathione (GSH) metabolism in the EHF^{low} clusters of the single cell data, and proceeded to evaluate the expression of the same genes in our cell line models. Interestingly, many of the genes associated with glutathione metabolism are also part of the drug metabolism pathway, and show an expression pattern consistent with being repressed by EHF in both SCC25 and CAL-27 cells (Figure 8A, C). Notably, we identified that *SLC3A2*, the gene encoding the regulatory subunit of the cystine glutamate antiporter xCT (60,61) is also downregulated by EHF (Supplementary Figure S4). The xCT system is an important transporter for the amino-acid cystine, a rate limiting precursor for glutathione synthesis (62). In addition, glutathione peroxidase 2 (*GPX2*), responsible for recycling reduced glutathione (GSSG) was downregulated in the presence of *EHF* (Figure 8C and Supplementary Figure S4). We verified by western blot that both SLC3A2 and GPX2 protein levels are depleted by EHF, but not EHF (Figure 8D). Conversely, the depletion of EHF in CAL-27 resulted in the upregulation of SLC3A2 but not GPX2, also suggesting a context dependent regulation (Figure 8D). Together, these results suggest that the expression of EHF could potentially deplete cellular glutathione stores. To confirm the dysregulation of glutathione metabolism by EHF, we evaluated

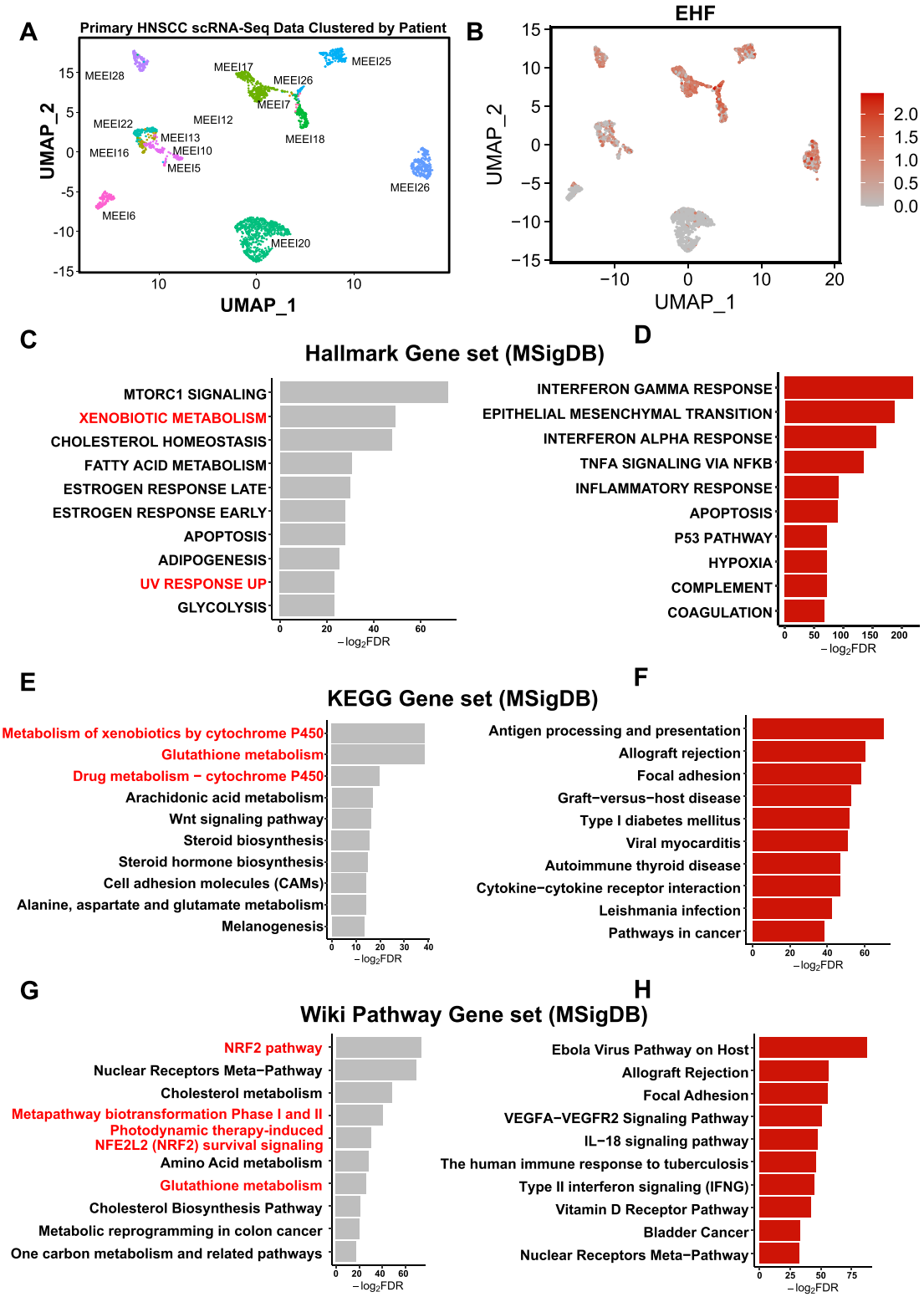


Figure 6. Single-cell RNA-sequencing Dataset Identifies Cluster of Patients with EHF^{high} and EHF^{low} Tumors. (**A**) UMAP clustering analysis reproducing the patient-driven clustering of pre-clinical malignant HNSCC cells from Puram *et al.* 2017. (**B**) Distribution of EHF expression across the malignant HNSCC, clusters patients into EHF^{high} and EHF^{low} subsets. (**C**) and (**D**) Bar plots showing Hallmark processes enriched in EHF^{low} clusters and EHF^{high} clusters, respectively. (**E** and **F**) Bar plots showing KEGG pathway processes enriched in EHF^{low} and EHF^{high} clusters respectively. (**G** and **H**) Bar plot showing wiki pathway processes that are enriched in EHF^{low} and EHF^{high} clusters.

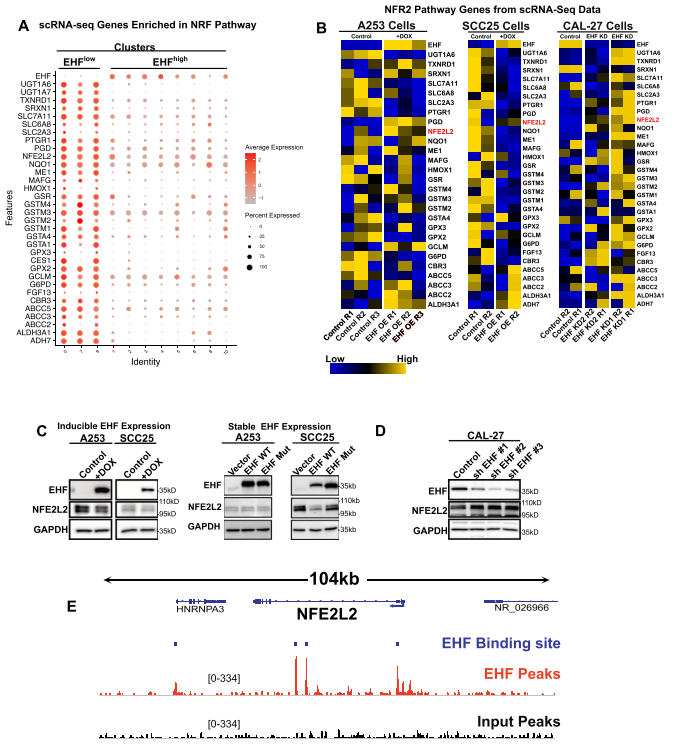


Figure 7. EHF Represses NRF2 and NRF2 Target Genes. (**A**) Dotplot showing the distribution of all the genes from the human HNSCC scRNA-seq data that re enriched in NRF2 pathway. (**B**) Heatmap showing the expression NRF2 pathway genes following the overexpression (A253 and SCC25) or loss (CAL-27) of EHF. Genes missing in any cell line heatmap are not expressed in that particular cell line. (**C**) Western blot showing the expression of NRF2 following inducible and stable EHF expression in both A253 and SCC25 cells. (**D**) Western blot showing the expression of NFR2 following the loss of EHF in CAL-27 cells. (**E**) ChIP-seq data from CAL-27 shows that EHF binds multiple regulatory regions within the *NRF2* gene. Genes missing in any panel are not expressed in that particular cell line.

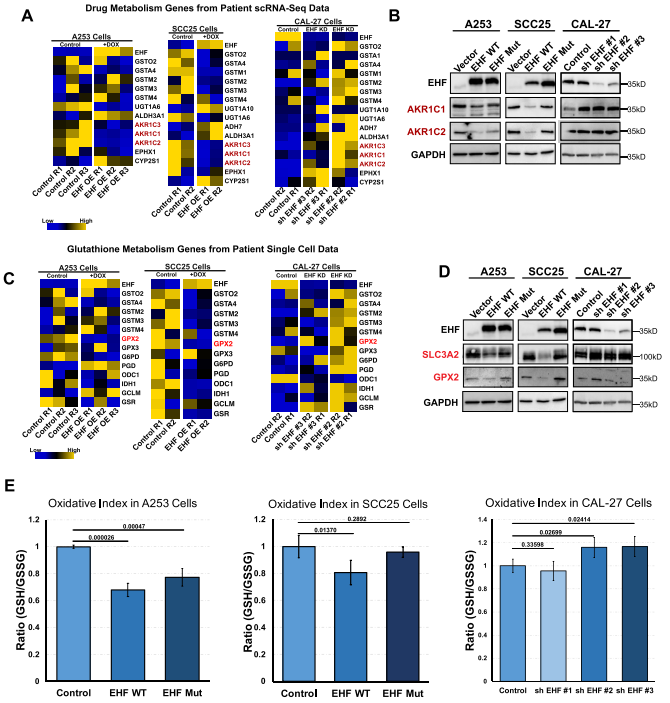


Figure 8. EHF Represses Genes Involved in Drug and Glutathione Metabolism. (**A**) Expression of genes associated with xenobiotic metabolism from Figure 6C, in EHF overexpression (A253 and SCC25) and EHF depleted (CAL-27) cells. Genes missing from any heatmap are not expressed in that particular cell line. (**B**) Western blot validation of the repression of AKR1C1 and AKR1C2 by EHF. (**C**) Expression of genes associated with glutathione metabolism from Figure 6E and G, in EHF overexpressing (A253 and SCC25) and EHF depleted (CAL-27) cells. (**D**) Western blot showing the expression of SLC3A2 and GPX2 following the overexpression and depletion of EHF. (**E**) Quantification of reduced and oxidized glutathione ratios in all three cell lines, showing that EHF expression results in the loss of reduced glutathione (GSH). Experiment was performed in triplicates, and measurements were taken twice. The bar graphs are presented as average values ± standard deviation (error bars).

the ratio of reduced to oxidized glutathione as a surrogate for measuring glutathione metabolism in our cell line models. We observed that in CAL-27 cells, depletion of EHF resulted in increased reduced glutathione levels, while the overexpression of EHF in A253 and SCC25 cells resulted in reduced glutathione levels (Figure 8E), thus confirming that EHF regulates the intracellular levels of glutathione.

SOX2, a direct EHF target shows heterogeneous expression in HNSCC tumor populations that likely represent metabolically different subtypes

Recent studies have identified a role for SOX2 in the maintenance of cellular glutathione stores in SCC (55), potentially via its ability to regulate glucose uptake and metabolism (55,63). Interestingly, gene set enrichment analysis from both patient single-cell data and cell-line data showed the enrichment of both glycolysis and *SOX2* expression in the absence of EHF, suggesting a possible link (Figure 3C, G). We therefore evaluated the expression of *SOX2* in relation to EHF expression in the scRNA-seq datasets. As expected, *SOX2* expression was anti-correlated with *EHF* expression (Figure 9A) while *p63*, another transcription factor that has been shown to regulate glycolysis similar to *SOX2*, (55) was ubiquitously expressed across all clusters. Upon probing our ChIP-seq dataset, we identified an EHF binding site distal to the *SOX2* gene (Figure 9B). We evaluated the expression of SOX2 and EHF by western blot, and found that SOX2 was repressed following the expression of wildtype EHF, but not by EHF-Mut in A253 and SCC25 cell lines, with the effects being more pronounced in the latter as expected (Figure 9C). Similar results were also observed in the doxycycline inducible EHF over-expression system (Figure 9D). In agreement, loss of EHF in CAL-27 cell line, led to a clear upregulation of SOX2, confirming the repressive role of EHF in regulating SOX2 levels (Figure 9D).

EHF and SOX2 exist in distinctly non-overlapping intratumoral compartments within HNSCC tumors that are likely to be metabolically different

To evaluate the protein expression pattern of EHF in human HNSCC patients, we examined EHF expression in a tissue microarray (TMA) (HN438a, US Biomax) consisting of eight (8) normal and forty (40) cancer tumor sections. Our EHF immunostaining analysis revealed weak and dif-

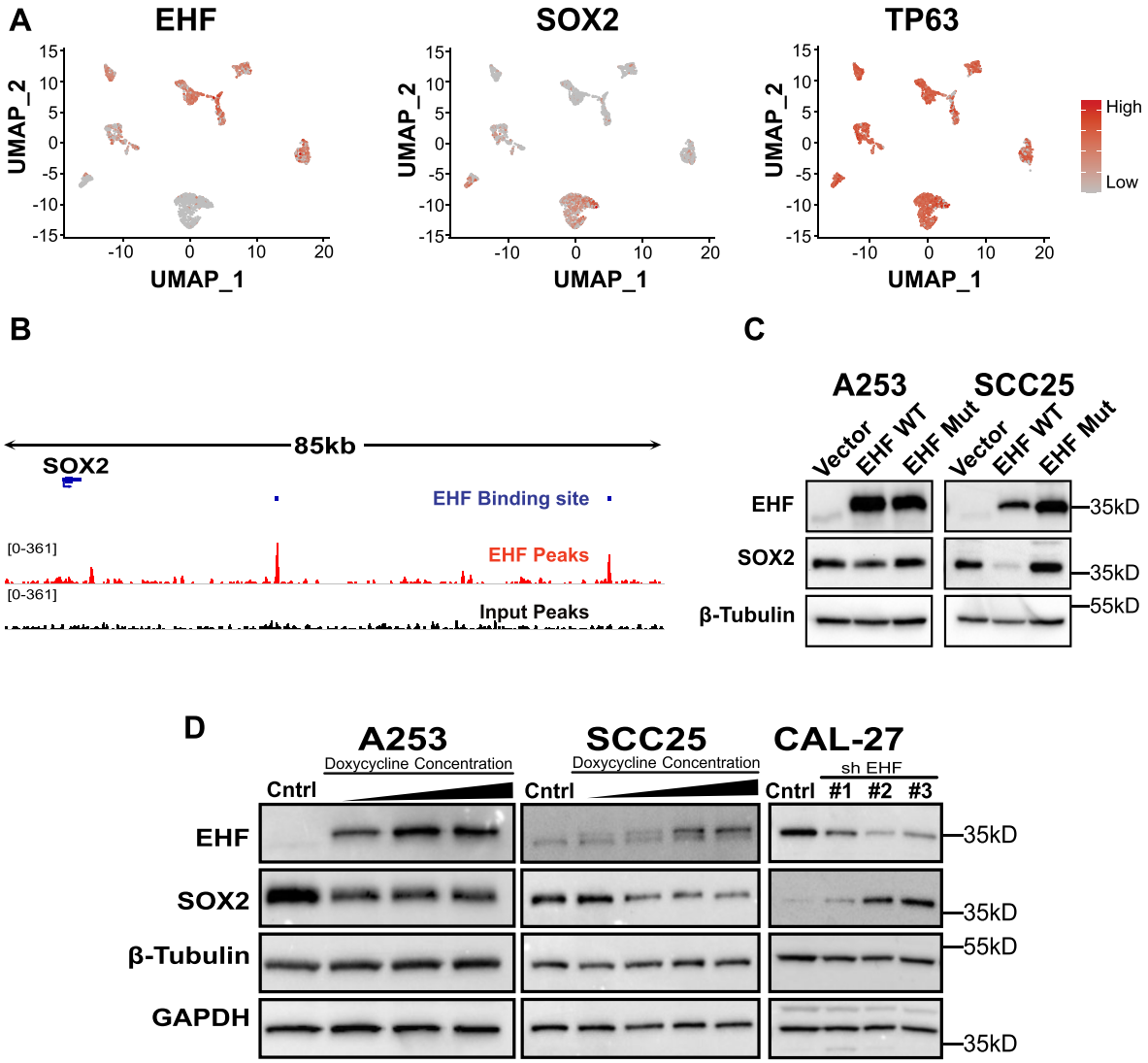


Figure 9. SOX2 Expression is Repressed by EHF. (**A**) Single-cell data showing that the expression of EHF and SOX2 are mutually exclusive, while *TP63* expression is ubiquitous. (**B**) ChIP-sequencing data showing the location of EHF binding sites distal to the *SOX2* gene. (**C**) Western blot showing the repression of SOX2 by stably expressed EHF but not EHF-Mut in A253 and SCC25 cells. (**D**) Western blot showing dose dependent repression of SOX2 by EHF in A253 and SCC25 cells. In contrast, shRNA mediated downregulation of EHF in CAL-27 cells leads to upregulation of SOX2. GAPDH and β-tubulin serve as loading control.

fuse expression of EHF in most of the tumor sections indicating a general loss of EHF expression. We next determined the expression levels SOX2 using an identical tissue microarray section. Interestingly, in tumors where EHF was strongly expressed, we observed a consistent trend of intratumoral regions with EHF^{high} pockets that were bereft of any SOX2 staining (Figure 10A, B).

DISCUSSION

The molecular heterogeneity of HNSCC and the lack of pervasive genomic alterations and mutations that define such tumors pose a significant barrier to developing targeted treatments. Hence, there is a need to better characterize the transcriptional regulatory circuitry that spur the oncogenic and tumor suppressive forces underlying the biology of HNSCC - such efforts might offer therapeutic dividends. Here, we showcase the ETS transcription factor EHF as a novel biomarker that can potentially stratify HNSCC tumors into metabolically active and otherwise. We report for the first time, a comprehensive examination of the transcriptional program driven by EHF in HNSCC and identify a major role for EHF in suppressing metabolic processes in these tumors.

Previous reports have offered conflicting roles for EHF in oral cancers, with one study positing that EHF expression is favorable for the growth and progression of cancer cells (64), while others mark EHF as a tumor suppressor (49,65). By carefully dissecting the TCGA-HNSC dataset, we offer new insights that may be useful in solving this paradox. Our studies reveal that HNSCC tumors can be primarily subdivided into EHF^{high} and EHF^{low} subtypes and that there exists a survival advantage for patients in the TCGA-HNSC database whose tumors expressed high levels of *EHF*. This

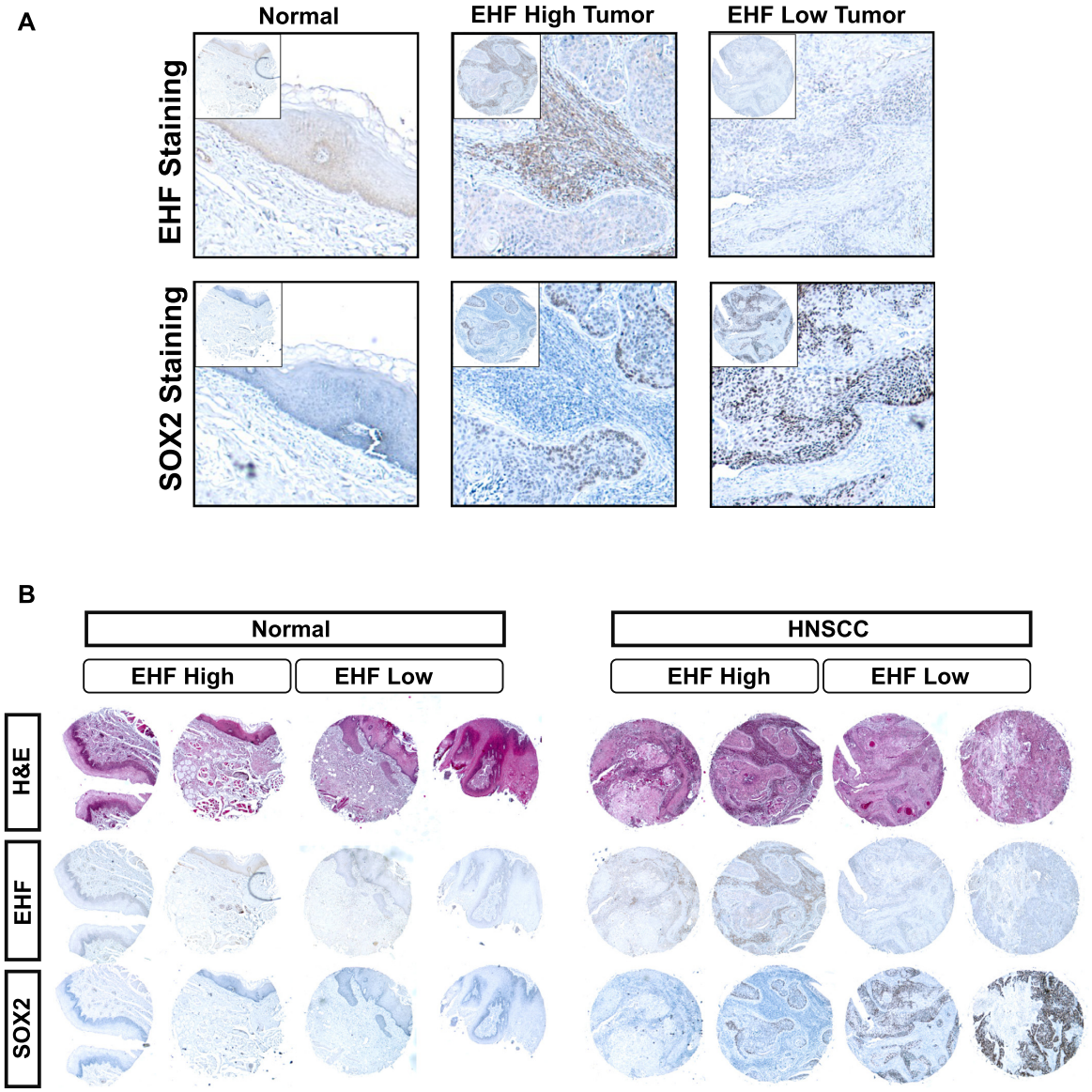


Figure 10. SOX2 and EHF Protein Expression is Variegated and Mutually Exclusive in HNSCC Tumor Samples. (**A**) 20× magnified view of immunostaining of HNSCC TMA, showing the discrete intratumoral areas of expression of EHF and SOX2. Insets whole tumor cores. (**B**) Representative histological sections showing the H&E staining of tumor cores and immunostaining of corresponding cores with EHF and SOX2.

clear dichotomy of *EHF* expression pattern is further recapitulated in a complementary analysis of scRNA-seq based datasets from an independent cohort of pre-clinical HNSCC patients (21), with the caveat that these results are based on a relatively small sample size. The relevance of EHF as a key player in HNSCC is in good agreement with recent findings identifying EHF as a top candidate among the 58 differentially expressed genes between normal and tumor-derived human organoids (66).

To address the underlying molecular mechanisms of EHF function and the likely consequences of *EHF* loss in tumors, we have utilized pre-clinical models of HNSCC and performed comprehensive gain-of-function and loss-of-function studies. Our RNA-seq and ChIP-seq based deciphering of the *EHF* cistrome has provided a rich trove of datasets to broadly examine the biological processes that are governed by this transcription factor. Gene set enrichment analysis, particularly from the scRNA-seq datasets, to a large extent match the data obtained from EHF overexpression and depletion experiments in HNSCC cell lines and have allowed us to hone into conserved and likely important pathways that are driven by EHF. We show that gene expression programs relating to cellular metabolism and metabolic homeostasis are salient features that distinguish the EHF^{high} and EHF^{low} clusters and that many key players within these pathways are direct transcriptional targets of EHF. Our analysis specifically identified the dysregulation of glutathione metabolism, a distinct feature of cancer cells that links cellular survival to metabolism, as being enriched in the EHF^{low} cluster. However, one interesting as-

pect of EHF function that remains unexplored in our current study, is its potential role in influencing the immune microenvironment of HNSCC tumors. Indeed, a persistent theme in our discovery of the EHF-driven pathways also points to immune modulation such as the enrichment of activated immune processes in the *EHF*^{high} clusters of tumors. These observations fit well with published studies showing a role for *EHF* in editing the immune microenvironment and influencing the efficacy of anti-PD1 immunotherapy in pancreatic cancers (51). Given that the loss of EHF in pancreatic cancers leads to a decrease in the number of tumor-infiltrating CD8^{+} T cells, it is tempting to speculate that *EHF*^{low} HNSCC tumors might exhibit similar immune phenotypes and that the altered metabolic state resulting from the loss of *EHF* might facilitate the evasion of immune surveillance.

The wide-ranging expression pattern of EHF across the large cohort of HNSCC tumor samples as well as the pockets of EHF^{high} and EHF^{low} tumor cells observed in the immunostaining of tumor samples speaks to the well-established vagaries of intra-tumor heterogeneity (67–69). We posit that during the evolution and progression of HNSCC, *EHF*^{high} and *EHF*^{low} tumor cells materialize possibly by DNA-methylation driven suppression of regulatory regions that control EHF expression which in turn subverts EHF-gene regulatory networks that affect tumor growth, survival and differentiation among many other oncogenic traits. This notion is in agreement with recent findings showing that ablation of DNA hypermethylation in HNSCC cells blocks the expression of redox metabolic genes and dampens cellular response to oxidative stress (70). Notably, we unearth a strong association between the transcriptional activity of EHF and altered metabolic and ROS pathways. This is exemplified by two key transcription factors, NRF2 and SOX2, which are downstream of EHF regulation and control the redox/metabolic state of the cell and its response to stress. As we show by biochemical, genomic and other assays, EHF acts as broad repressor for many of these genes and pathways and that its loss unleashes the constitutive activation of an NRF2 gene expression program that spans the gamut including coordinated regulation of glutathione metabolism, response to ultraviolet radiation and xenobiotic detoxification. Similarly, the restricted expression pattern of SOX2 to predominantly *EHF*^{low} areas of HNSCC tumors is particularly relevant given the recently discovered role of SOX2 in dictating glucose reliance and metabolic vulnerabilities in squamous cell carcinomas in cooperation with *p63* (55). The variegated expression of SOX2 observed in our study is similar to published results with an independent cohort of HNSCC tissues that demonstrated the heterogeneous expression patterns of SOX2 and EMT-marker vimentin (67). We posit that the occurrence of substantial metabolic heterogeneity across different tumors and even within the same tumor is likely to be quite widespread as exemplified by SOX2 staining of HNSCC tumors in our studies.

Although the tumor suppressive functions of EHF in HNSCC as described here is similar to what has been reported in prostate cancer, where loss of EHF promotes a tumorigenic and stem-like phenotype (71), the role of EHF might be more complex and context-dependent in other cancers. Indeed, EHF has been shown to be overexpressed in primary papillary thyroid cancers and increased expression of EHF via gene amplification has been shown to contribute to the activation of HER family signaling and is associated with poor survival in gastric cancer (72,73). Furthermore, the recent identification of EHF splice variants expressed in some HNSCC cell lines may further complicate its biology (65). We suspect that the broad reach of the role of EHF in cancer may extend to other squamous cancers such as Esophageal adenocarcinoma (EAC), where an interconnected circuitry formed by EHF and three other master TFs; ELF3, KLF5 and GATA6 has been shown to regulate EAC-specific super-enhancers (74,75). Hence, as is often the case in complex gene regulatory processes, EHF is likely to have partners in crime in HNSCC. In this regard, the co-occurrence of AP-1 motifs around the genomic segments that are occupied by EHF in CAL-27 cells is of particular interest in light of recent structural modeling and mutational studies suggesting that EHF and AP-1 (JUN/FOS) use electrostatic repulsion to disfavor simultaneous DNA binding (76). Given the importance of AP-1 factors (77) and EHF (our study) in various facets of HNSCC, a complex interplay between these two key factors might be an interesting future avenue of research. On a similar note, the enrichment of NFE2L2 (NRF2) motifs in EHF occupied genomic regulatory elements also raises an intriguing possibility of molecular feedback and crosstalk between these two transcription factors in modulating the gene regulatory network underlying redox homeostasis. Follow-up studies in such crucial cellular processes will greatly facilitate better molecular stratification of tumors and the development of targeted metabolic inhibitors that can be utilized as therapeutics.

DATA AVAILABILITY

The RNA-Seq and ChIP-seq data generated in this study has been deposited in GEO under accession number GSE168271.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Cancer Online.

ACKNOWLEDGEMENTS

We are especially grateful to Jonathan Bard and Brandon Marzullo for bioinformatics support and the UB Genomics and Bioinformatics Core for the Next Generation sequencing services.

FUNDING

A.O. was supported by the National Institute of Health/National Institute of Dental and Craniofacial Research [DE028480]; A.O. was also partly supported by the State University of New York at Buffalo, School of Dental Medicine, Department of Oral Biology training grant (NIH/NIDCR) [DE023526].

Conflict of interest statement. The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

REFERENCES

1. Siegel,R.L., Miller,K.D. and Jemal,A. (2020) Cancer statistics, 2020. *CA Cancer J. Clin.*, **70**, 7–30.
2. Chow,L.Q.M. (2020) Head and neck cancer. *N. Engl. J. Med.*, **382**, 60–72.
3. Leemans,C.R., Braakhuis,B.J. and Brakenhoff,R.H. (2011) The molecular biology of head and neck cancer. *Nat. Rev. Cancer*, **11**, 9–22.
4. Wuerdemann,N., Wittekindt,C., Sharma,S.J., Prigge,E.S., Reuschenbach,M., Gattenlohner,S., Klussmann,J.P. and Wagner,S. (2017) Risk factors for overall survival outcome in surgically treated human Papillomavirus-negative and positive patients with oropharyngeal cancer. *Oncol. Res. Treat*, **40**, 320–327.
5. Jou,A. and Hess,J. (2017) Epidemiology and Molecular Biology of Head and Neck Cancer. *Oncol. Res. Treat*, **40**, 328–332.
6. Ferris,R.L., Blumenschein,G., Fayette,J., Guigay,J., Colevas,A.D., Licitra,L., Harrington,K., Kasper,S., Vokes,E.E., Even,C *et al.* (2016) Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N. Engl. J. Med.*, **375**, 1856–1867.
7. Ueki,Y., Takahashi,T., Ota,H., Shodo,R., Yamazaki,K. and Horii,A. (2020) Role of programmed death-ligand 1 in predicting the treatment outcome of salvage chemotherapy after nivolumab in recurrent/metastatic head and neck squamous cell carcinoma. *Head Neck*, **42**, 3275–3281.
8. Santuray,R.T., Johnson,D.E. and Grandis,J.R. (2018) New therapies in head and neck cancer. *Trends Cancer*, **4**, 385–396.
9. Fornaro,L., Vasile,E., Aprile,G., Goetze,T.O., Vivaldi,C., Falcone,A. and Al-Batran,S.E. (2018) Locally advanced gastro-oesophageal cancer: recent therapeutic advances and research directions. *Cancer Treat. Rev.*, **69**, 90–100.
10. Huilgol,D., Venkataramani,P., Nandi,S. and Bhattacharjee,S. (2019) Transcription factors that govern development and disease: an achilles heel in cancer. *Genes (Basel)*, **10**, 794–829.
11. Hanahan,D. and Weinberg,R.A. (2000) The hallmarks of cancer. *Cell*, **100**, 57–70.
12. Hanahan,D. and Weinberg,R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646–674.
13. Bradner,J.E., Hnisz,D. and Young,R.A. (2017) Transcriptional addiction in cancer. *Cell*, **168**, 629–643.
14. Seth,A. and Watson,D.K. (2005) ETS transcription factors and their emerging roles in human cancer. *Eur. J. Cancer*, **41**, 2462–2478.
15. Nicholas,T.R., Strittmatter,B.G. and Hollenhorst,P.C. (2019) Oncogenic ETS factors in prostate cancer. *Adv. Exp. Med. Biol.*, **1210**, 409–436.
16. Hsing,M., Wang,Y., Rennie,P.S., Cox,M.E. and Cherkasov,A. (2020) ETS transcription factors as emerging drug targets in cancer. *Med. Res. Rev.*, **40**, 413–430.
17. Luk,I.Y., Reehorst,C.M. and Mariadason,J.M. (2018) ELF3, ELF5, EHF and SPDEF transcription factors in tissue homeostasis and cancer. *Molecules*, **23**, 2191–2216.
18. Gluck,C., Glathar,A., Tsompana,M., Nowak,N., Garrett-Sinha,L.A., Buck,M.J. and Sinha,S. (2019) Molecular dissection of the oncogenic role of ETS1 in the mesenchymal subtypes of head and neck squamous cell carcinoma. *PLoS Genet.*, **15**, e1008250.
19. Yang,Z., Liao,J., Carter-Cooper,B.A., Lapidus,R.G., Cullen,K.J. and Dan,H. (2019) Regulation of cisplatin-resistant head and neck squamous cell carcinoma by the SRC/ETS-1 signaling pathway. *BMC Cancer*, **19**, 485.
20. Yang,H., Schramek,D., Adam,R.C., Keyes,B.E., Wang,P., Zheng,D. and Fuchs,E. (2015) ETS family transcriptional regulators drive chromatin dynamics and malignancy in squamous cell carcinomas. *Elife*, **4**, e10870.
21. Puram,S.V., Tirosh,I., Parikh,A.S., Patel,A.P., Yizhak,K., Gillespie,S., Rodman,C., Luo,C.L., Mroz,E.A., Emerick,K.S. *et al.* (2017) Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell*, **171**, 1611–1624.
22. Wang,Y., Feng,L., Said,M., Balderman,S., Fayazi,Z., Liu,Y., Ghosh,D. and Gulick,A.M. (2005) Analysis of the 2.0 A crystal structure of the protein-DNA complex of the human PDEF Ets domain bound to the prostate specific antigen regulatory site. *Biochemistry*, **44**, 7095–7106.
23. Escamilla-Hernandez,R., Chakrabarti,R., Romano,R.A., Smalley,K., Zhu,Q., Lai,W., Halfon,M.S., Buck,M.J. and Sinha,S. (2010) Genome-wide search identifies Ccnd2 as a direct transcriptional target of Elf5 in mouse mammary gland. *BMC Mol. Biol.*, **11**, 68.
24. Chakrabarti,R., Hwang,J., Andres Blanco,M., Wei,Y., Lukacisin,M., Romano,R.A., Smalley,K., Liu,S., Yang,Q., Ibrahim,T. *et al.* (2012) Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. *Nat. Cell Biol.*, **14**, 1212–1222.
25. Meerbrey,K.L., Hu,G., Kessler,J.D., Roarty,K., Li,M.Z., Fang,J.E., Herschkowitz,J.I., Burrows,A.E., Ciccia,A., Sun,T. *et al.* (2011) The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 3665–3670.
26. Kim,D., Paggi,J.M., Park,C., Bennett,C. and Salzberg,S.L. (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.*, **37**, 907–915.
27. Kim,D., Langmead,B. and Salzberg,S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods*, **12**, 357–360.
28. Pertea,M., Kim,D., Pertea,G.M., Leek,J.T. and Salzberg,S.L. (2016) Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.*, **11**, 1650–1667.
29. Liao,Y., Smyth,G.K. and Shi,W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.
30. Wagner,G.P., Kin,K. and Lynch,V.J. (2012) Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.*, **131**, 281–285.
31. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
32. Rahman,M., Jackson,L.K., Johnson,W.E., Li,D.Y., Bild,A.H. and Piccolo,S.R. (2015) Alternative preprocessing of RNA-sequencing data in the Cancer Genome Atlas leads to improved analysis results. *Bioinformatics*, **31**, 3666–3672.
33. Ogluszka,M., Orzechowska,M., Jedroszka,D., Witas,P. and Bednarek,A.K. (2019) Evaluate Cutpoints: Adaptable continuous data distribution system for determining survival in Kaplan-Meier estimator. *Comput. Methods Programs Biomed.*, **177**, 133–139.
34. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
35. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
36. Liu,T. (2014) Use model-based Analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. *Methods Mol. Biol.*, **1150**, 81–95.
37. McLean,C.Y., Bristor,D., Hiller,M., Clarke,S.L., Schaar,B.T., Lowe,C.B., Wenger,A.M. and Bejerano,G. (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.*, **28**, 495–501.
38. Ramirez,F., Dundar,F., Diehl,S., Gruning,B.A. and Manke,T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.*, **42**, W187–W191.
39. Machanick,P. and Bailey,T.L. (2011) MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics*, **27**, 1696–1697.
40. Kulakovskiy,I.V., Vorontsov,I.E., Yevshin,I.S., Sharipov,R.N., Fedorova,A.D., Rumynskiy,E.I., Medvedeva,Y.A., Magana-Mora,A., Bajic,V.B., Papatsenko,D.A. *et al.* (2018) HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.*, **46**, D252–D259.
41. Gupta,S., Stamatoyannopoulos,J.A., Bailey,T.L. and Noble,W.S. (2007) Quantifying similarity between motifs. *Genome Biol.*, **8**, R24.
42. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
43. Reis,P.P., Waldron,L., Perez-Ordonez,B., Pintilie,M., Galloni,N.N., Xuan,Y., Cervigne,N.K., Warner,G.C., Makitie,A.A., Simpson,C. *et al.* (2011) A gene signature in histologically normal surgical margins is predictive of oral carcinoma recurrence. *BMC Cancer*, **11**, 437.
44. Chen,C., Mendez,E., Houck,J., Fan,W., Lohavanichbutr,P., Doody,D., Yueh,B., Futran,N.D., Upton,M., Farwell,D.G. *et al.* (2008) Gene expression profiling identifies genes predictive of oral

squamous cell carcinoma. *Cancer Epidemiol. Biomarkers Prev.*, **17**, 2152–2162.
45. Stuart,T., Butler,A., Hoffman,P., Hafemeister,C., Papalexi,E., Mauck,W.M. III, Hao,Y., Stoeckius,M., Smibert,P. and Satija,R. (2019) Comprehensive integration of single-cell data. *Cell*, **177**, 1888–1902.
46. Tugores,A., Le,J., Sorokina,I., Snijders,A.J., Duyao,M., Reddy,P.S., Carlee,L., Ronshaugen,M., Mushegian,A., Watanaskul,T. *et al.* (2001) The epithelium-specific ETS protein EHF/ESE-3 is a context-dependent transcriptional repressor downstream of MAPK signaling cascades. *J. Biol. Chem.*, **276**, 20397–20406.
47. Kunderfranco,P., Mello-Grand,M., Cangemi,R., Pellini,S., Mensah,A., Albertini,V., Malek,A., Chiorino,G., Catapano,C.V. and Carbone,G.M. (2010) ETS transcription factors control transcription of EZH2 and epigenetic silencing of the tumor suppressor gene Nkx3.1 in prostate cancer. *PLoS One*, **5**, e10547.
48. Albino,D., Longoni,N., Curti,L., Mello-Grand,M., Pinton,S., Civenni,G., Thalmann,G., D'Ambrosio,G., Sarti,M., Sessa,F. *et al.* (2012) ESE3/EHF controls epithelial cell differentiation and its loss leads to prostate tumors with mesenchymal and stem-like features. *Cancer Res.*, **72**, 2889–2900.
49. Wang,L., Xing,J., Cheng,R., Shao,Y., Li,P., Zhu,S. and Zhang,S. (2015) Abnormal localization and tumor suppressor function of epithelial tissue-specific transcription factor ESE3 in esophageal squamous cell carcinoma. *PLoS One*, **10**, e0126319.
50. Liberzon,A., Birger,C., Thorvaldsdottir,H., Ghandi,M., Mesirov,J.P. and Tamayo,P. (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.*, **1**, 417–425.
51. Liu,J., Jiang,W., Zhao,K., Wang,H., Zhou,T., Bai,W., Wang,X., Zhao,T., Huang,C., Gao,S. *et al.* (2019) Tumoral EHF predicts the efficacy of anti-PD1 therapy in pancreatic ductal adenocarcinoma. *J. Exp. Med.*, **216**, 656–673.
52. Chen,J.K., Zhan,Y.J., Yang,C.S. and Tzeng,S.F. (2011) Oxidative stress-induced attenuation of thrombospondin-1 expression in primary rat astrocytes. *J. Cell. Biochem.*, **112**, 59–70.
53. Makhezer,N., Ben Khemis,M., Liu,D., Khichane,Y., Marzaioli,V., Tlili,A., Mojallali,M., Pintard,C., Letteron,P., Hurtado-Nedelec,M. *et al.* (2019) NOX1-derived ROS drive the expression of Lipocalin-2 in colonic epithelial cells in inflammatory conditions. *Mucosal. Immunol.*, **12**, 117–131.
54. Koppula,P., Zhang,Y., Zhuang,L. and Gan,B. (2018) Amino acid transporter SLC7A11/xCT at the crossroads of regulating redox homeostasis and nutrient dependency of cancer. *Cancer Commun. (Lond)*, **38**, 12.
55. Hsieh,M.H., Choe,J.H., Gadhvi,J., Kim,Y.J., Arguez,M.A., Palmer,M., Gerold,H., Nowak,C., Do,H., Mazambani,S. *et al.* (2019) p63 and SOX2 dictate glucose reliance and metabolic vulnerabilities in squamous cell carcinomas. *Cell Rep.*, **28**, 1860–1878.
56. Chang,W.M., Chang,Y.C., Yang,Y.C., Lin,S.K., Chang,P.M. and Hsiao,M. (2019) AKR1C1 controls cisplatin-resistance in head and neck squamous cell carcinoma through cross-talk with the STAT1/3 signaling pathway. *J. Exp. Clin. Cancer Res.*, **38**, 245.
57. Tonelli,C., Chio,I.I.C. and Tuveson,D.A. (2018) Transcriptional regulation by Nrf2. *Antioxid. Redox. Signal.*, **29**, 1727–1745.
58. He,F., Antonucci,L. and Karin,M. (2020) NRF2 as a regulator of cell metabolism and inflammation in cancer. *Carcinogenesis*, **41**, 405–416.
59. Shirato,A., Kikugawa,T., Miura,N., Tanji,N., Takemori,N., Higashiyama,S. and Yokoyama,M. (2014) Cisplatin resistance by induction of aldo-keto reductase family 1 member C2 in human bladder cancer cells. *Oncol. Lett.*, **7**, 674–678.
60. Shin,C.S., Mishra,P., Watrous,J.D., Carelli,V., D'Aurelio,M., Jain,M. and Chan,D.C. (2017) The glutamate/cystine xCT antiporter antagonizes glutamine metabolism and reduces nutrient flexibility. *Nat. Commun.*, **8**, 15074.
61. Bassi,M.T., Gasol,E., Manzoni,M., Pineda,M., Riboni,M., Martin,R., Zorzano,A., Borsani,G. and Palacin,M. (2001) Identification and characterisation of human xCT that co-expresses, with 4F2 heavy chain, the amino acid transport activity system xc. *Pflugers Arch.*, **442**, 286–296.
62. Lewerenz,J., Hewett,S.J., Huang,Y., Lambros,M., Gout,P.W., Kalivas,P.W., Massie,A., Smolders,I., Methner,A., Pergande,M. *et al.* (2013) The cystine/glutamate antiporter system x(c)(-) in health and disease: from molecular mechanisms to novel therapeutic opportunities. *Antioxid. Redox. Signal.*, **18**, 522–555.
63. Andreucci,E., Pietrobono,S., Peppicelli,S., Ruzzolini,J., Bianchini,F., Biagioni,A., Stecca,B. and Calorini,L. (2018) SOX2 as a novel contributor of oxidative metabolism in melanoma cells. *Cell Commun. Signal*, **16**, 87.
64. Huang,W.C., Jang,T.H., Tung,S.L., Yen,T.C., Chan,S.H. and Wang,L.H. (2019) A novel miR-365-3p/EHF/keratin 16 axis promotes oral squamous cell carcinoma metastasis, cancer stemness and drug resistance via enhancing beta5-integrin/c-met signaling pathway. *J. Exp. Clin. Cancer Res.*, **38**, 89.
65. Sakamoto,K., Endo,K., Sakamoto,K., Kayamori,K., Ehata,S., Ichikawa,J., Ando,T., Nakamura,R., Kimura,Y., Yoshizawa,K. *et al.* (2021) EHF suppresses cancer progression by inhibiting ETS1-mediated ZEB expression. *Oncogenesis*, **10**, 26.
66. Driehuis,E., Kolders,S., Spelier,S., Lohmussaar,K., Willems,S.M., Devriese,L.A., de Bree,R., de Ruiter,E.J., Korving,J., Begthel,H. *et al.* (2019) Oral mucosal organoids as a potential platform for personalized cancer therapy. *Cancer Discov.*, **9**, 852–871.
67. Baumeister,P., Hollmann,A., Kitz,J., Afthonidou,A., Simon,F., Shakhtour,J., Mack,B., Kranz,G., Libl,D., Leu,M. *et al.* (2018) High expression of EpCAM and Sox2 is a positive prognosticator of clinical outcome for head and neck carcinoma. *Sci. Rep.*, **8**, 14582.
68. Mroz,E.A. and Rocco,J.W. (2016) Intra-tumor heterogeneity in head and neck cancer and its clinical implications. *World J. Otorhinolaryngol. Head Neck Surg.*, **2**, 60–67.
69. Gotte,K., Tremmel,S.C., Popp,S., Weber,S., Hormann,K., Bartram,C.R. and Jauch,A. (2005) Intratumoral genomic heterogeneity in advanced head and neck cancer detected by comparative genomic hybridization. *Adv. Otorhinolaryngol.*, **62**, 38–48.
70. Gleneadie,H.J., Baker,A.H., Batis,N., Bryant,J., Jiang,Y., Clokie,S.J.H., Mehanna,H., Garcia,P., Gendoo,D.M.A., Roberts,S. *et al.* (2021) The anti-tumour activity of DNA methylation inhibitor 5-aza-2′-deoxycytidine is enhanced by the common analgesic paracetamol through induction of oxidative stress. *Cancer Lett.*, **501**, 172–186.
71. Albino,D., Civenni,G., Dallavalle,C., Roos,M., Jahns,H., Curti,L., Rossi,S., Pinton,S., D'Ambrosio,G., Sessa,F. *et al.* (2016) Activation of the Lin28/let-7 axis by loss of ESE3/EHF promotes a tumorigenic and stem-like phenotype in prostate cancer. *Cancer Res.*, **76**, 3629–3643.
72. Lv,Y., Sui,F., Ma,J., Ren,X., Yang,Q., Zhang,Y., Guan,H., Shi,B., Hou,P. and Ji,M. (2016) Increased expression of EHF contributes to thyroid tumorigenesis through transcriptionally regulating HER2 and HER3. *Oncotarget*, **7**, 57978–57990.
73. Shi,J., Qu,Y., Li,X., Sui,F., Yao,D., Yang,Q., Shi,B., Ji,M. and Hou,P. (2016) Increased expression of EHF via gene amplification contributes to the activation of HER family signaling and associates with poor survival in gastric cancer. *Cell Death. Dis.*, **7**, e2442.
74. Ma,S., Zhou,B., Yang,Q., Pan,Y., Yang,W., Freedland,S.J., Ding,L.W., Freeman,M.R., Breunig,J.J., Bhowmick,N.A. *et al.* (2021) A transcriptional regulatory loop of master regulator transcription factors, PPARG, and fatty acid synthesis promotes esophageal adenocarcinoma. *Cancer Res.*, **81**, 1216–1229.
75. Chen,L., Huang,M., Plummer,J., Pan,J., Jiang,Y.Y., Yang,Q., Silva,T.C., Gull,N., Chen,S., Ding,L.W. *et al.* (2020) Master transcription factors form interconnected circuitry and orchestrate transcriptional networks in oesophageal adenocarcinoma. *Gut*, **69**, 630–640.
76. Madison,B.J., Clark,K.A., Bhachech,N., Hollenhorst,P.C., Graves,B.J. and Currie,S.L. (2018) Electrostatic repulsion causes anticooperative DNA binding between tumor suppressor ETS transcription factors and JUN-FOS at composite DNA sites. *J. Biol. Chem.*, **293**, 18624–18635.
77. Zhang,M., Hoyle,R.G., Ma,Z., Sun,B., Cai,W., Cai,H., Xie,N., Zhang,Y., Hou,J., Liu,X. *et al.* (2021) FOSL1 promotes metastasis of head and neck squamous cell carcinoma through super-enhancer-driven transcription program. *Mol. Ther.*, **29**, 2583–2600.