The Proteomic Landscape of Growth Factor Signaling Networks Associated with FAT1 Mutations in Head and Neck Cancers

Zhengjia Chen, Emory University
Chao Zhang, Emory University
Jianhong Chen, Roswell Park Comprehensive Cancer Center, Buffalo
Dongsheng Wang, Emory University
Jieqi Tu, University of Illinois at Chicago
Carter Van Waes, National Institute on Deafness and Other Communication Disorders, NIH, Bethesda
Nabil Saba, Emory University
Zhao G Chen, Emory University
Zhong Chen, National Institute on Deafness and Other Communication Disorders, NIH, Bethesda

Journal Title: CANCER RESEARCH
Volume: Volume 81, Number 17
Publisher: AMER ASSOC CANCER RESEARCH | 2021-09-01, Pages 4402-4416
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1158/0008-5472.CAN-20-3659
Permanent URL: https://pid.emory.edu/ark:/25593/vx86c

Final published version: http://dx.doi.org/10.1158/0008-5472.CAN-20-3659
Accessed August 10, 2024 4:17 PM EDT
The Proteomic Landscape of Growth Factor Signaling Networks Associated with FAT1 Mutations in Head and Neck Cancers

Zhengjia Chen1,2, Chao Zhang3, Jianhong Chen4, Dongsheng Wang5, Jieqi Tu1, Carter Van Waes6, Nabil F Saba5, Zhuo G. Chen5,*, Zhong Chen6,*

1. Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL 60612
2. Biostatistics Shared Resource Core, University of Illinois Cancer Institute, Chicago, IL 60612
3. Biostatistics and Bioinformatics Shared Resource of Winship Cancer Institute, Emory University, Atlanta, GA 30322
4. Department of Cancer Prevention and Control, Roswell Park Comprehensive Cancer Center, Buffalo, NY 14263
5. Department of Hematology and Medical Oncology, Emory University School of Medicine, Atlanta, GA 30322
6. Head and Neck Surgery Branch, National Institute on Deafness and Other Communication Disorders, NIH, Bethesda, MD 20892

Abstract

FAT1 is frequently mutated in head and neck squamous cell carcinoma (HNSCC), but the biological and clinical effects of FAT1 mutations in HNSCC remain to be fully elucidated. We investigated the landscape of altered protein and gene expression associated with FAT1 mutations and clinical outcomes of HNSCC patients. FAT1 mutation was stratified with clinical information from The Cancer Genome Atlas HNSCC databases with more than 200 proteins or phosphorylated sites. FAT1 mutation was significantly more prevalent among HPV(−), female, and older patients and was enriched in oral, larynx, and hypopharynx primary tumors. FAT1 mutation was also significantly associated with lower FAT1 gene expression and increased protein expression of HER3_pY1289, IRS1, and CAVEOLIN1. From an independent International Cancer Genome

*Corresponding Authors: Zhong Chen, Head and Neck Surgery Branch, National Institute on Deafness and Other Communication Disorders, NIH, Bethesda, MD 20892, USA, Phone: 301-435-2073; Fax: 301-594-4643; chenz@nidcd.nih.gov; Zhuo Georgia Chen, Department of Hematology and Medical Oncology, Emory University School of Medicine, Atlanta, GA. Phone: 404-778-3977; Fax: 404-778-5520; gzchen@emory.edu.

Author Contributions

Conception and design: Zhong Chen, Zhuo G. Chen, Zhengjia Chen, Nabil F Saba
Provision of study material: Zhong Chen, Zhengjia Chen, Chao Zhang, Jianhong Chen, Zhuo G Chen
Collection and assembly of data: Chao Zhang, Jianhong Chen, Zhengjia Chen, Dongsheng Wang, Jieqi Tu
Data analysis and interpretation: Chao Zhang, Jianhong Chen, Zhengjia Chen, Zhong Chen, Zhuo G Chen, Nabil F Saba, Dongsheng Wang, Jieqi Tu
Manuscript writing: Zhong Chen, Zhuo G Chen, Zhengjia Chen, Jianhong Chen, Nabil F Saba, Carter Van Waes
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors
Disclosure of Potential Conflict of Interest
All authors declare having no conflict of interest.
Consortium dataset, FAT1 mutation in oral cancer co-occurred with top mutated genes TP53 and CASP8. Poorer overall survival or progression-free survival was observed in patients with FAT1 mutation or altered HER3_pY1289, IRS1, or CAVEOLIN. Pathway analysis revealed dominant ERBB/neuregulin pathways mediated by FAT1 mutations in HNSCC, and protein signature panels uncovered the heterogeneity of patient subgroups. Decreased pEGFR, pHER2, and pERK and upregulated pHER3 and HER3 proteins were observed in two FAT1 knockout HNSCC cell lines, supporting that FAT1 alterations lead to altered EGFR/ERBB signaling. In squamous cancers of the lung and cervix, a strong association of FAT1 and EGFR gene expression was identified. Collectively, these results suggest that alteration of FAT1 appears to involve mostly HPV(−) HNSCC and may contribute to resistance to EGFR-targeted therapy.

Keywords
FAT1 mutation; Reverse Phase Protein Array (RPPA); The Cancer Genomics Atlas (TCGA); clinical outcomes; HPV status; head and neck cancer

Introduction

Head and neck squamous cell carcinoma (HNSCC) is a devastating and often fatal disease, which significantly disturbs patients’ vital upper aero-digestive function and is the sixth most common cancer worldwide (1). Despite advances in multimodality treatments, including recent advances in immune-therapy, the overall 5-year survival rate for patients with curable HNSCC is only 61% (http://seer.cancer.gov/). Tumor human papillomavirus (HPV) status remains the strongest predictor of survival, with HPV-unrelated disease carrying a 5-year overall survival (OS) rate of less than 50% (2). Significant gaps in understanding the biology of disease continue to be the main reason behind the paucity of effective therapeutic interventions.

The challenges in effectively treating HNSCC are attributed to its extreme heterogeneity as far as anatomic locations and genetic aberrations (3–6). These genetic alterations, especially gene mutations, accumulate during the growth of HNSCC, complicating our understanding of disease biology and limiting the effectiveness of targeted approaches and ultimately the prognosis of patients. The Cancer Genome Atlas (TCGA) has provided the most comprehensive characterization to date of the genomic and proteomic landscape (5,6). In TCGA dataset, the mutation rate of FAT1 is ~23% in HNSCC, ranking FAT1 as the second most mutated gene after TP53 and supporting its important role in HNSCC biology (5). In HPV-negative (HPV−) HNSCC, the FAT1 mutation rate is as high as 28% with many truncations (7), suggesting that wild-type FAT1 serves as a tumor suppressor gene in this disease (4). The FAT1 mutation rate in HNSCC is the highest among major solid tumors, making its investigation of primary interest in this disease (www.cbioportal.org). The effect of FAT1 mutation on HNSCC malignant phenotypes has not been extensively investigated, and little is known about its clinical implications, highlighting the significance of our findings.

FAT1 is a cadherin-like protein family member, and a large type 1 transmembrane protein of 4588 amino acid residues. It has 34 cadherin repeats, a laminin G domain, and five
epidermal growth factor (EGF)-like repeats in the extracellular region, followed by a transmembrane region and a C-terminal cytoplasmic domain containing a PDZ-binding motif (8). Early studies identified FAT1 as an ortholog of the Drosophila fat gene family, and loss of fat leads to cell cycle dysregulation and hyperproliferation in larval imaginal discs (9). Recently, FAT1 mutation was identified in human cancers and indicated to contribute to Wnt (Wingless and Int-1) activation, consistent with observations that FAT1 serves as a tumor suppressor in human cancer cells (10). In addition, FAT1 mutant was found to inactivate the Hippo regulatory complex, leading to activation of YAP1 (Yes-associate protein 1) signaling in HNSCC (7,11). However, comprehensive understanding of the mechanisms underlying FAT1 mutation and its associated growth factor-mediated activation of signaling pathways in cancer progression, particularly in HNSCC, remain limited.

We herein report our comprehensive proteomic analysis identifying altered protein expression and activation status associated with FAT1 mutation in HNSCC using the Reverse Phase Protein Arrays (RPPA) dataset from TCGA. This platform represents a powerful approach with globally functional proteomics, including more than 200 validated antibodies detecting proteins or phosphorylation sites indicating activation status (Cancer Proteome Atlas, TCPA, http://tcpaportal.org). The selection of this panel of antibodies was based on antigens representing molecules with critical functions in cancer biology, such as growth factors, signaling molecules, transcription factors, cell cycle, and apoptosis modulators. Using TCGA genomic, expression, and RPPA datasets, we compared differentially expressed proteins and phosphorylated sites in HNSCC tissues associated with FAT1 mutation status using systematic biostatistics and bioinformatics analyses. We have identified a panel of differentially expressed proteins, stratified by FAT1 mutation status, involved in the activation of growth factors and signaling pathways with potential effect on cancer cell proliferation, cell cycle regulation, apoptosis, and immunomodulation. Furthermore, FAT1 mutation and associated protein expression were associated with a statistically significant difference in patients’ OS and progression free survival (PFS), indicating a potential prognostic value for these biomarkers. Our novel findings from proteomic analysis associated with FAT1 mutation therefore provide important information as to potential biomarkers that are linked to FAT1 and that deserve future validation in HNSCC.

Materials and Methods
TCGA genomic and proteomic databases and patient information
A flow chart of data extraction and analyses is presented in Supplemental Figure S1A–C. TCGA specimens were obtained with written informed consent under an Institutional Review Board (IRB) approved protocol in accordance with recognized ethical guidelines (6,12,13). FAT1 mutation, gene expression, protein expression, and associated clinical features of TCGA HNSCC cohort were downloaded from the Broad Institute Firebrowse website (http://firebrowse.org/). These datasets include level3 RNASeq data (presented as log2 transformed RNA-Seq by Expectation Maximization [RSEM]), normalized protein levels of Reverse Phase Protein Array (RPPA)(12), and curated MAF files. Patient information is summarized in Supplemental Table S1. In brief, these databases contain 489
cases of HNSCC, with 348 wild-type and 141 cases showing \textit{FAT1} alterations, including mutations and other genetic variations (28.8%). There are 329 patient samples with RPPA data, including 96 with altered \textit{FAT1} (29.2%) and 233 with wild-type \textit{FAT1} status.

**Survival analysis of patients with \textit{FAT1} mutation and associated protein expression**

We examined the correlation of protein expression with patients’ OS and PFS. OS was calculated as the time from date of diagnosis to death or last patient contact. PFS was calculated as the time from date of diagnosis to date of disease progression, death, or last contact, whichever was earliest. Biomarkers were further dichotomized by the optimal cutoff point (high [$\geq$ cutoff point] versus low [$<$ cutoff point]), which corresponded to the most significant relationship with OS or DFS based on log-rank statistic (14; [http://www2sascom/proceedings/sugi28/261-28pdf](http://www2sascom/proceedings/sugi28/261-28pdf)).

**Univariate and multivariate survival analysis**

In the univariate survival analysis, Kaplan-Meier survival estimates were calculated for each group of patients dichotomized by each biomarker’s optimal cutoff point and compared between the two groups with the log-rank test among all patients, respectively (15). For the most significant biomarkers (such as HER3, CYCLINE, IRS, RET-pY905, cMYC, P16INK4A, etc.), we further investigated whether \textit{FAT1} can modify their effects on survival by subgroup analyses.

**Ingenuity Pathway Analysis (IPA)**

IPA is an all-in-one, annotated, web-based software application that enables analysis, integration, and understanding of data from gene expression and proteomics ([analysis.ingenuity.com](http://analysis.ingenuity.com)). IPA allows searches for targeted information on genes, proteins, chemicals, and drugs, which could build interactive models of experimental systems. The software is backed by the Ingenuity Knowledge Base of highly structured, detail-rich biological and chemical findings. Detailed analysis is described in Supplemental Materials and Methods.

**Heatmap and hierarchical clustering analysis**

Unsupervised clustering analysis was performed using 20 signature protein molecules after extraction and normalization from the HPV(−) patient group, by comparing RPPA protein levels between patient groups with \textit{FAT1} wild-type versus mutant (p < 0.1). The 329 total patient samples with RPPA data were subjected to hierarchical clustering by Manhattan distance in the column, Euclidean distance in row, and Ward.D linkage method was performed using the Pheatmap package (version 1.0.12) of R software. Based on cluster formation in both row and column directions, samples with associated signature protein molecules were pooled and assigned subcluster numbers. For each subcluster, protein scores were calculated using overall samples’ protein levels measured by RPPA, compared using one-way ANOVA and reported as p-values. Box and whisker plots of median distributions of protein levels for each subcluster were made using GraphPad Prism (Version 6.1). PFS in various patient subgroups were tested for differences using Kaplan-Meier analysis. Survival curves for each patient subgroup were created using GraphPad Prism (Version 6.1).
Results

Correlation of FAT1 mutations with clinical characteristics of HNSCC patients

The majority of genetic alterations of the FAT1 gene in cancers are mutations, while only a low percentage of copy number variations (CNV), predominantly deep deletions, has been observed (Supplemental Figure S2). HNSCC (including both HPV status) has the highest FAT1 mutation rate compared with other major cancer types, revealed by the PanTCGA project. We analyzed FAT1 mutation status in HNSCC using TCGA genomic and proteomic datasets and investigated associations with patients’ clinical characteristics (Table 1). TCGA datasets include 489 total cases with complete genomic information, 329 of which have protein data from RPPA. FAT1 mutation status is defined by TCGA, including point mutation, deletion, indel insertion, translocation, and fusion. FAT1 mutation rates and clinical characteristics in the genomic and proteomic datasets showed similar frequencies and distributions (Table 1, Supplemental Table S1). We observed that the FAT1 mutation rate is significantly higher in females than in males in both genomic (38.76% vs. 25.28%, p=0.004) and proteomic (39.36% vs. 25.11%, p=0.01) datasets. Patients with FAT1 mutations tend to be older with a mean age of 65 years compared with 59 years among patients with FAT1 wild-type (p<0.001 in both datasets). Interestingly, among patients with FAT1 mutations, case numbers were statistically higher in the oral cavity (OC) and larynx and hypopharynx (LH) than in oropharynx (OP), and the FAT1 mutation rate was higher in HPV(−) than HPV(+) patients in both datasets. There was no statistically significant association of FAT1 mutation with race, alcohol usage, smoking status, disease stage, or lymph node status (Table 1).

FAT1 mutation status associated with gene expression under different clinical conditions

FAT1 gene mutation status (left panels) and mRNA expression (right panels) were compared after stratification by HPV status (Figure 1A), primary tumor location (Figure 1B), disease stage (Figure 1C), and lymph node status (Figure 1D). More patients with FAT1 mutations and higher FAT1 gene expression levels were observed among HPV(−) than HPV(+) patients. FAT1 gene expression levels were significantly lower in patients with FAT1 mutations than wild-type for both HPV statuses (Figure 1A). Comparing tumors from different primary locations, the percentage of patients with FAT1 alterations was higher among patients with OC and LH (mainly HPV-negative) than in OP (mainly HPV-positive) tumors. FAT1 gene expression levels were significantly lower in OC and LH tumors with mutation than wild-type, but were similar between FAT1 wild-type and mutant OP tumors. Wild-type FAT1 gene expression in OP tumors was significantly lower than in OC and LH (Figure 1B). The percentages of patients with FAT1 alterations were similar across different disease stages (including both HPV status), but there were higher case numbers of stage
IV tumors. For each disease stage, \textit{FAT1} gene expression was significantly lower among patients with mutant \textit{FAT1} than wild-type. However, no statistical significance was observed when comparing among different stages within both \textit{FAT1} genetic status (Figure 1C). The percentages of patients with \textit{FAT1} mutation were similar in tumors with different lymph nodes statuses (including both HPV status). \textit{FAT1} gene expression was lower in patients with mutant status, except for N3 stage, which could be due to the lower number of cases. There was no difference in \textit{FAT1} gene expression among patients with different lymph node stages within each \textit{FAT1} genetic status (Figure 1D).

\textbf{\textit{FAT1} mutation status in oral cancer from the ICGC dataset and association with other top mutated genes in HNSCC from TCGA dataset}

We analyzed an independent dataset of oral cancer from the ICGC project, which mainly focused on genetic mutations in 178 donors from India (https://dcc.icgc.org, Supplemental Figure S3A) (16, 17). In this dataset, the frequency of \textit{FAT1} mutation was further enriched compared to TCGA at 37%, ranking #3 among these patients. Interestingly, the top two mutated genes were \textit{TP53} and \textit{CASP8}, and the \textit{CASP8} mutation rate was unusually high at 42.1% (Supplemental Figure S3B). In addition, we studied the association of \textit{FAT1} genetic alterations with other top mutated genes in HNSCC from TCGA dataset (Supplemental Figure S3C–D). While \textit{CASP8} mutation was only seen in 13% of cases, this alteration was highly enriched in the group with \textit{FAT1} genetic alterations. \textit{FAT1} genetic alterations co-occurred with \textit{CASP8} or \textit{TP53}, and this association was statistically significant (Supplemental Figure S3D–F).

\textbf{Association of \textit{FAT1} mutation with expression and activation of growth factor receptors and signaling molecules}

To understand if \textit{FAT1} mutation is associated with the expression of other proteins, we compared two patient groups with different \textit{FAT1} mutation status for their protein expression assayed by RPPA. We analyzed a total of 329 patient samples, including 287 HPV(−) and 35 HPV(+) tumors, which have complete clinical information (Table 1, Supplemental Table S1). Supplemental Table S2.1 presents protein analysis data for total patients with both HPV status, including 233 patients with wild-type \textit{FAT1} and 96 patients with mutated \textit{FAT1}, using Student’s T-test under a relaxed significance level of 0.1 as a screening threshold. In Table 2, we present the protein names used in RPPA with official gene names and gene symbols, chromosome locations, and a brief description of their functions and potential interacting molecules. We ranked these proteins and corresponding gene expression from TCGA RNAseq data according to statistically significant differences between different \textit{FAT1} mutation status. We also included a few molecules with near statistically significant p values ≥0.05 (Table 2, Supplemental Table S2 and S3).

We observed that the molecules most significantly associated with \textit{FAT1} mutation are related to growth factor receptors and their downstream signaling modulators under the RPPA protein selection criteria. In patient samples with \textit{FAT1} mutation, we observed increased \textit{HER3}_pY1289 (p=0.004), insulin receptor substrate 1 (IRS1, p=0.017), and cellular membrane protein CAVEOLIN 1 (p=0.017). Vascular endothelial growth factor receptor 2 exhibited a p value barely over the cutoff (VEGFR2, p=0.054). PD-L1 was
included due to its established clinical importance as a therapeutic target and CMYC given its significant genetic amplification (5). Among these proteins, HER3_pY1289 and IRS1 were significantly or near significantly upregulated in HPV(−) HNSCC (p=0.011 and 0.06, respectively), but not in HPV(+) HNSCC (p>0.1). Upregulation of HER3_pY1289 and VEGFR2 was only observed at the protein level, while upregulation of IRS1 and CAVEOLIN 1 was also observed at the mRNA level. We identified 12 proteins that were significantly increased in tumors with wild-type FAT1 status (Table 2). The immune checkpoint molecule PDCD1 (PD1) was increased in total patient samples with wild-type FAT1 (p=0.054), and in HPV(−) cases (p=0.051).

Since RPPA data did not include the protein expression of immune signatures for CD3+, CD4+, CD8+ T cells, and Treg cells, we analyzed their immune gene signatures and related inflammatory cellular and cytokine markers using TCGA RNAseq data. We did not find any significant difference in most immune gene expression signatures associated with FAT1 mutation in either total patients’ samples or HPV(−) samples (Supplemental Table S4.1 and S4.2). However, we observed decreased gene expression of immune cytokine IL-2 and growth factors FGF1 and FGF2, and increased NRG1 (neuregulin, protein ligand for ERBB2/3) and TGFβ1 in total patient samples with FAT1 mutation (Supplemental Table S4.1 and Supplemental Figure S4). In HPV(−) patient samples with FAT1 mutation, NRG1 exhibited increased expression with statistical significance, while the CD274 (PDL1) expression was marginally increased (p value=0.06), consistent with its protein expression (Supplemental Table S4.2).

**FAT1 mutation and related proteins contribute to HNSCC patient survival**

To further investigate the clinical significance of FAT1 mutation and associated proteins, we analyzed OS and PFS of total and HPV(−) patients using both univariate and multivariable survival analyses, presented as median survival months. From the HNSCC TCGA dataset, we observed a trend towards decreased PFS of patients with FAT1 mutation alone, which approached statistical significance (Supplemental Figure S5, p=0.06). To explore the covariates that could influence the effect of mutant FAT1 status on survival, we performed survival analysis, presented as Kaplan-Meier plots, after stratification by protein expression and FAT1 mutation status. This revealed that mutant FAT1 together with HER3_pY1289, IRS1, and CYCLINE2 were significantly associated with worse OS (Figure 2A). Patients with lower expression of HER3_pY1289 exhibited better OS, while FAT1 mutation status alone showed a lesser impact. In patients with wild-type FAT1, IRS1 expression strongly affected OS. Patients with lower IRS1 expression exhibited the best survival, while patients with mutant FAT1 status had intermediate OS. In contrast, patients with mutant FAT1 status and lower expression of CYCLINE2 exhibited better OS than patients with lower CYCLINE2 expression and wild-type FAT1 status (Figure 2A). In HPV(−) patients, higher IRS and lower RET_pY905 were dominantly and significantly associated with worse OS (Figure 2B). Higher P16INK4A and lower CMYC were significantly associated with better PFS in patients with wild-type FAT1 (Figure 2C). Pairwise comparison of the survivals for Figure 2 are presented in Supplemental Table S5.1. Altered expression of several molecules did not significantly affect survival under the analysis with stratified FAT1, but exhibited statistical significance in a pairwise analysis in each group (Supplemental Figure S6A–D, Cancer Res. Author manuscript; available in PMC 2022 March 01.
Supplemental Table S5.2). Univariate and multivariate statistical analyses of OS using COX model are presented in Supplemental Table S6, PFS in Supplemental Table S7 (18).

We then tested these molecules showing significant association with FAT1 mutation using univariate analysis of Kaplan-Meier plots, but without stratification by FAT1 mutation status. We observed that higher expression levels of HER3_pY1289, CYCLINE2, and IRS1 were significantly associated with worse OS among all patients (Supplemental Figure S7A). HPV(−) patients exhibited worse OS with lower expression of RET_pY905 and mTOR_pS2448, and higher IRS (Supplemental Figure S7B). Among the total patients, lower expression of P16INK4A and higher expression of CAVEOLIN1, IRS1, and CMYC were significantly associated with worse PFS (Supplemental Figure S7C). Among HPV(−) patients, only lower expression of mTOR_pS2448 was significantly associated with worse PFS (Supplemental Figure S7D).

**Potential signaling pathways and networks mediated by FAT1 alterations**

To understand potential signaling pathways and networks mediated by FAT1 alterations in HNSCC, we performed IPA using 31 molecules identified and combined from total and HPV(−) patient samples. As expected, the top pathways identified by IPA included Molecular Mechanisms of Cancer, ERBB Signaling, Neuregulin Signaling, and EGF Signaling, ranked by high statistical significance and positive z-scores (Supplemental Figure S8A). The molecules involved in these pathways are presented in Supplemental Table S8. Upstream analysis predicted TP53 and MYC as essential transcriptional regulators, and ERBB2 as the potentially critical receptor that regulates downstream pathways (Figure 3A, Supplemental Figure S8A). Supporting these predictions, we observed the co-occurrence of TP53 and FAT1 genetic alterations in the HNSCC TCGA dataset (Supplemental Figure S3F). Altered CMYC protein expression and ERBB2 phosphorylation were associated with FAT1 mutation (Table 2, Supplemental Table S2), of which increased CMYC protein expression also contributed to worse patient survival (Supplemental Figure S7C, Figure 2C). Growth factors and cytokines, such as TGFB1, IL-6, and FGF2, were identified as potential upstream regulators (Figure 3B). These predictions are supported by RNA-seq data in TCGA analysis (Supplemental Figure S4). We analyzed the downstream cellular functional categories regulated by the proteins identified, and observed a decrease in apoptosis activity and increase in cell cycle progression and proliferation activities (Figure 3C). Furthermore, we uncovered potential therapeutic drugs which could target these molecules, including sirolimus/rapamycin (mTOR inhibitor) (19), curcumin (anti-inflammatory dietary supplement) (20), and tanespimycin (17-AAG, HSP90 inhibitor) (21) (Supplemental Figure S8B, S9).

**Validation of FAT1 affected signaling pathways**

To validate if there is a functional relationship between FAT1 inactivation and observed major signaling pathways, we knocked out the FAT1 gene in two HNSCC cell lines, MDA686TU and UMSCC-1P, using CRISPR-Cas9 technology, and established both parental (WT) and FAT1-knockout (KO) HNSCC cell line pairs (Figure 3D). Western blot analysis illustrated decreased pEGFR, pHER2 and pERK proteins, indicating inactivation of the EGFR signaling axis. pHER3 and HER3 were consistently elevated but to a lower degree
after FAT1 protein KO. These experimental results support the causal relationship of FAT1 mutation and altered signaling pathways identified in this study.

As several upstream signaling cytokines and growth factors were predicted by IPA (Figure 3B) and confirmed by gene expression analysis (Supplemental Table S4, Figure S4), we further analyzed the protein expression of IL-6 and TGFβ in supernatants from FAT1 WT and KO MDA686TU and UMSCC-1P cells by ELISA (Supplemental Figure S10). We confirmed elevations of IL-6 and TGFβ proteins in FAT1 KO HNSCC cell lines, inferred above by IPA and gene expression data.

**Tumor heterogeneity classified by FAT1 mutation and associated proteins**

To understand the tumor heterogeneity associated with FAT1 mutation and related protein expression or phosphorylation, we performed hierarchical cluster analysis using a panel of 20 differentially expressed proteins listed in Table S2.2. We observed two subgroups of HNSCC patients within each group of patients with FAT1 mutation (Figure 4A) or wild-type status (Figure 4B). The two observed patient subgroups were represented by hierarchical clusters horizontally on the top of the heatmaps, without identified significant association with HPV status or primary tumor locations. Two major protein expression patterns were revealed by the hierarchical links on the left side within each heatmap. In patients with FAT1 mutation (A), the protein group presented at the top of the cluster included all phosphorylated proteins, such as TUBERING_pT1462, GSK_pS9, AKT_pS473, SHC_pY317, EGFR_pY1068, HER2_pY1248, RET_PY905, CJUN_pS73, while the protein cluster at the bottom included most total proteins with only three phosphorylated proteins, HER3_pY1289, SRC_pY527, and MTOR_pS2448. Similar protein clusters were observed in patients with FAT1 wild-type (B), with the only exception of two phosphorylated proteins, MTOR_pS2448 and SRC_pY527. Next, we divided the heatmaps into four quadrants (labeled I-IV at the tops and bottoms), and calculated the total expression scores of each protein clustering group (Figure 4C, D). The expression scores were statistically different among the four clusters from both heatmaps (p<0.001). When we compared the scores between protein clusters, all pairwise comparisons exhibited statistical significance, with one exception of cluster III vs. IV from the FAT1 mutant group (Figure 4C). Furthermore, we did not observe an overall benefit in PFS when comparing patient group I vs. II separated by the protein clusters (Figure 4E, F). However, when particular PFS time points were analyzed, we observed a statistically significant improvement in survival for patient subgroup I regardless of FAT1 status, that subgroup I patients overexpressed EGFR_pY1068, HER2_pY1248 and other phosphorylated proteins. In patients with FAT1 mutant status, the only survival benefit was observed at day 1200 (3.3 years, Figure 4E), while for patients with FAT1 wild-type status, the survival benefit was observed relatively later at days 2500 (6.8 years), 3000 (8.2 years), and 4000 (11 years) (Figure 4F, Supplemental Table S9).

**FAT1 genetic and expression alterations in other squamous cancer types**

Our previous publication revealed that squamous cell cancers (SCC) share many common genetic alterations (6). We explored the broader relevance of FAT1 alterations in HNSCC and other squamous cancer types, such as lung and cervical SCCs, and identified a strong
association between \textit{FAT1} and \textit{EGFR} gene expression (Figure 5A–C). This association ranked \#1 in lung SCC and \#4 in cervical cancer but \#24 in HNSCC. The association between gene expression of \textit{FAT1} and these genes in lung and cervical SCC is consistent with the data in HNSCC, where \textit{FAT1} gene mutations and deletions are associated with decreased \textit{FAT1} gene expression (Figure 5D–F). In addition, we observed that \textit{FAT1} genetic alterations in HNSCC and lung SCC are associated with a high mutation count, while in lung SCC, \textit{FAT1} genetic alterations are also associated with aneuploidy score (Figure 5G–I). Furthermore, we displayed \textit{FAT1} and \textit{EGFR} gene expression by individual HNSCC, lung SCC, and cervical cancer patient samples by Oncoprint (Supplemental Figure S11A–C) and observed that \textit{EGFR} gene expression was also associated with copy number alterations (Supplemental Figure S11D–F).

\section*{Discussion}

In this study, we characterized the global proteomic landscape associated with \textit{FAT1} mutation in more than 300 HNSCC tissue samples from TCGA with available RPPA data (12). The \textit{FAT1} mutation rate in HNSCC is 23–28\% (5,7), and is the highest among other major cancer types including lung and cervical SCC (11). The frequency of \textit{FAT1} mutations is further enriched for oral SCC in the independent ICGC data set. Together, these observations support a critical role and unique contribution of \textit{FAT1} to oncogenesis of a subset of head and neck cancers. Following \textit{TP53}, \textit{FAT1} is the second most commonly mutated gene in HNSCC, especially HPV(−) tumors, underlining its pathological significance (5). \textit{TP53} and \textit{FAT1} mutations co-occur in HNSCC. Given the challenges in targeting p53 in HNSCC, pathway(s) modulated by mutant \textit{FAT1} could have impact on disease outcome and serve as additional targets for prevention or therapy for a significant proportion of patients with \textit{TP53} alterations (22). The unique alteration pattern of \textit{FAT1} implicates its functional importance in SCC, but with less understood mechanisms than the well-studied \textit{TP53} and \textit{CDKN2A} (\textit{p16}) (6). Furthermore, \textit{FAT1} mutations include a high percentage of truncation and nonsense mutation types, which could eliminate its expression and abolish the entire function of the \textit{FAT1} molecule. Accordingly, tumors with \textit{FAT1} mutation exhibited significantly lower \textit{FAT1} gene expression, which supports its functional role as a tumor suppressor.

Our most significant and novel finding is the identification of a broader proteomic landscape of predominant onco genes and tumor suppressors and potentially druggable targets associated with mutant or wild-type \textit{FAT1}. We observed increased expression of a network of proteins that includes cell surface receptors, such as HER3\_pY1289, VEGFR2, and PDL1, plus IGFR signaling mediator IRS1, and cell cycle modulator CMYC, in more than 90 HNSCC patient samples with \textit{FAT1} mutation (Supplemental Figures S12, S13). Although the highly mutated \textit{FAT1} molecule itself has not been identified as a therapeutic target in HNSCC, due to its nature as a tumor suppressor, these surface receptors and signaling molecules associated with \textit{FAT1} mutation are well-investigated therapeutic targets with available targeted agents and possible links to immune-related pathways. Specifically, our data showed that in both total and HPV(−) patients, HER3\_pY1289 was upregulated in \textit{FAT1} mutated HNSCC (Table 2 and Supplemental Table S2). HER3 is one of the EGFR family members and HER3\_pY1289 is the activated form that transduces signaling after...
partnering with other EGFR family members (23, 24). We found significant correlation between FAT1 and EGFR expression in lung, cervical and head and neck SCC, among which this correlation was ranked #1 in lung SCC. We also observed that FAT1 knockout could reduce EGFR expression and signaling in HNSCC cell lines, indicating the potential impact of FAT1 in EGFR signaling. In addition, IRS1 (25), a key regulator of IGF-1R, is also upregulated in FAT1 mutated HNSCC. Cross-talk between HER3 and IGF-R1(26) signals could synergistically activate ERK/MAPK, PI3K/AKT, and RAS/RAF pathways and promote cell proliferation/survival, protein synthesis, and cell cycle through CMYC (27) (Supplemental Figures S12, S13). Among other associated proteins, CAVEOLIN (28,29) can alter cell shapes, and together with VEGFR2 (30), enhance cell angiogenesis, stemness, migration, and invasion potential (Supplemental Figure S13). These findings are consistent with a recent publication suggesting that FAT1 contributes to epithelial-to-mesenchymal transition (EMT), tumor cell stemness, and metastasis in both skin and lung cancer models (31). Furthermore, expression of a different checkpoint protein, PD1, was higher in the tumor microenvironment with wild-type FAT1, which could induce immunotolerance and suppress anti-tumor immunity (32).

IPA analysis revealed dominant upstream regulators, including TP53, CMYC, and ERBB2, as potentially activated transcription factors and cell surface molecules, plus upregulated growth factors and proinflammatory cytokines, such as TGFβ1 (33), IL-6 (34,35), and FGF2 (36). These predictions were confirmed by gene expression data from TCGA RNA-seq dataset (Supplemental Figure S4 and Supplemental Table S4), and/or the protein data tested by ELISA in HNSCC FAT1 knockout cell lines (Supplemental Figure S10). Increased TGF1β expression in the FAT1 altered group was predicted by IPA, and confirmed by gene expression analysis from TCGA datasets and protein expression in cell lines by ELISA. Although increased IL6 was not observed in the current gene expression analysis, moderate upregulation of IL-6 was observed in the supernatant of FAT1 knockout cells by ELISA. Though we did not observe FAT1 alteration affecting immune CD3+, CD4+, CD8+ T cells, and Treg cells, we observed altered growth factors and pro-inflammatory cytokines, which could modulate the immune response in the tumor microenvironment. These predicted signaling pathways and regulatory networks and in vitro experimental results warrant further investigation regarding their mechanistic association with FAT1 mutation and clinical implication.

HNSCC is remarkably heterogeneous, partially due to the complex anatomical structures and different primary tumor locations. Two significant etiologies of tobacco usage and HPV infection represent two major types of diseases, which activate distinct biological pathways in carcinogenesis. We found that the FAT1 mutation rate is significantly higher in HPV(−) than HPV(+) patients in this cohort, consistent with a previous publication (7). In that study, the FAT1 mutation rate was reported as up to 28% in HPV(−) versus 2.8% in HPV(+) patients, suggesting that FAT1 mutation may play more critical roles in HPV(−) than HPV(+) HNSCC. It is not surprising that the FAT1 mutation rate is higher in OC and LH than in OP, since most OP tumors are HPV(+). Significantly, patients with FAT1 mutation were older and female, which could be compounded by HPV(−) status. Also, we identified a group of proteins differentially expressed in FAT1 mutated HPV(+) cases, which were distinct from those proteins identified in HPV(−) cases (Supplemental Table S2.3).
However, we cannot conclude the possible role of *FAT1* mutation in HPV(+) HNSCC due to the small number of cases, and this should be validated in a larger cohort.

The heterogeneity of HNSCC is also influenced by the large variety of genomic, expression, and proteomic alterations that drive carcinogenesis. This study observed two major subgroups of samples exhibiting different protein signatures within each group classified by different *FAT1* mutation status. Hierarchical clustering analysis of protein expression revealed two distinct clusters of protein molecules that were mutually exclusive in HNSCC subgroups with either *FAT1* mutant or wild-type status (Figure 4). We observed mutually exclusive protein expression and activation, such as phosphorylation of EGFR/HER2 in one subgroup of HNSCC, versus activated HER3 and IGFR signaling IRS protein in the different patient subgroups regardless of *FAT1* mutation status. The underlying mechanisms and clinical impacts of these protein clusters shown by hierarchical heatmaps remain unclear. These observations are consistent with our and other previous studies showing that HER3 activation is one of the primary mechanisms that bypasses EGFR blockade (23, 37). Literature suggests that HER3 might interact with IGF-1R to activate common downstream signaling (38), contributing to cetuximab resistance and impacting clinical outcomes (39).

One of the clinically significant observations from this study comes from the survival analysis of the protein biomarkers and their relationship with the significant impact of *FAT1* mutations through statistical analysis (Figure 2 and Supplemental Figure S6). *FAT1* mutation status exhibited a stronger impact on PFS than OS. Our observations of the impact of mutant *FAT1* on PFS are supported by a previous study on HNSCC patients in Taiwan, which reported a 29% mutation rate of *FAT1* and significant correlations with lymph node status and worse disease-free survival (DFS), but not OS (40). Although we did not find that *FAT1* mutations alone were associated with either PFS or OS in univariate analysis of this cohort, statistical significance could be diminished by the long survival time of follow-up for >200 months (>16 years) in this cohort, where other causes affect survival. There was a survival benefit in PFS among all patients with *FAT1* wild-type status which approached significance (p=0.06; Supplemental Figure S2B), which could reach statistical significance if shorter-term survival of 5 years was analyzed. It is worth noting that the PFS analysis at earlier time points revealed statistical significance, and that patients with activated HER3_pY1289 and other protein overexpression exhibited worse PFS. Furthermore, an earlier effect on PFS was seen among patients with mutated *FAT1* status than those with *FAT1* wild-type (Figure 4E, F). These observations of survival benefits may help classify HNSCC patients with differential treatment plans and prognostic outcomes, which warrant future clinical investigations. Better stratification of HNSCC patients based on these alterations could be instrumental in addressing the therapeutic dilemma that has surrounded patients with HPV-unrelated disease, who exhibit TP53 alteration and resistance to EGFR antagonists. Exploring immune-related correlates to these pathways may unlock critical information instrumental in the design of future combinatorial approaches in patients with HPV-unrelated disease.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.
Acknowledgments

This study was supported by grants from a pilot Winship Invests Grant from Emory University Winship Cancer Institute to GZC and NFS, and NIDCD intramural projects Z01-DC-00016, 73, 74 to ZC. The research reported in this publication was partially supported by the University of Illinois Cancer Center Biostatistics Shared Resource Core (BSRCC). We thank Li Jia, PhD, Bioinformatician from National Institutes of Health Library for her analysis of proteomic datasets of head and neck cancers. We thank Mr. Jianwei Chen who participated in the study of TCGA datasets of head and neck cancer and analyzing the association of the FAT1 mutation with pathogenesis of head and neck cancer. We also thank Drs. Anthea Hammond, Gregory B. Lesinski, Shi-Yong Sun (Emory University) for their suggestions and critical reading of the article. The content is solely the authors’ responsibility and does not necessarily represent the official views of the National Institutes of Health.

References

Significance:

Integrative bioinformatics and statistical analyses reveal a panel of genes and proteins associated with FAT1 mutation in HNSCC, providing important insights into prospective clinical investigations with targeted therapies.
Figure 1. Comparison of clinical characteristics and prognostic factors associated with FAT1 mutation status through genomic data analysis in HNSCC patients

Clinical characteristics were extracted for 489 HNSCC patients from TCGA datasets of genomic information associated with FAT1 mutation status. The categorical covariates on the left panels are presented as the case number and the percentage of total patients (%).

(A) Comparison of FAT1 mutation frequency (%) in HNSCC patients differing in HPV status. p-value < 0.001 by Fisher’s Exact test. (B) Primary tumor location. The frequency of FAT1 mutation is statistically significantly lower in OP compared with other locations. p-value < 0.001 by Chi-Square test. (C) Tumor stage. (D) Lymph node status. The right
panels present *FAT1* RNA gene expression (RSEM) by RNAseq through TCGA datasets. (A) Statistical significance is indicated by * when comparing *FAT1* gene expression between patients with *FAT1* wild-type versus mutant status, # comparing between HPV(−) patients versus HPV(+) patients. (B) * Comparing between the groups with *FAT1* wild-type versus mutant status; # comparing between OP versus OC, + comparing between OP versus LH. (C) * Comparing between *FAT1* wild-type versus mutant status among different stages, # comparing Stage III versus Stage I; + comparing Stage III versus Stage II. p-value <0.05, Student’s T–test.
Figure 2. Kaplan-Meier survival analysis of significant OS and PFS of HNSCC patients after stratification by overall FAT1 mutation status and protein expression. Kaplan-Meier OS analysis was performed for total HNSCC patients (n=329, A) and HPV(−) patients (n=287, B), and PFS was analyzed for total patients (n=309, C) based on the data availability. The survival curves are presented for those that reached the overall statistical significance after stratification by FAT1 mutation status and protein expression. Median survival months are presented. Logrank test was performed for statistical analysis, and p values are presented at the upper left corners. The pairwise comparison of any two patient
groups was highlighted with *, based on the statistical p-value <0.05. “Mutation” in the legend indicates FAT1 mutation.
Figure 3. Ingenuity Pathway Analysis (IPA) of proteins associated with \textit{FAT1} mutation status. Molecules with significant difference between patients with FAT1 wild-type and mutant status were extracted from total and HPV(−) patients, and the fold changes in protein values were calculated based on \textit{FAT1} wild-type versus mutant status. (A) Predicted upstream receptor and transcription regulators. (B) Predicted upstream regulators of cytokines and growth factors. (C) Predicted cellular function. All those categories are highly statistically significant, p<0.001 (Fisher's Exact test provided by IPA). The protein names were converted to gene names for IPA analysis. Legends from IPA are presented at the right.
upper corner. Pink/Red indicates increased expression in FAT1 wild-type patients, and Green indicates increased expression in patients with mutant FAT1. Additional legends are presented in Supplemental Figure S8. (D) FAT1 in HNSCC cell lines MDA686TU and UMSCC-1P was knocked out using CRISPR-Cas9 technique. To validate the effect of FAT1 knockout on the EGFR family and related signaling pathways, western blot analysis was performed using whole cell lysates from each of the FAT1 wild-type (WT) and knockout (KO) cell line pairs. Relative densitometrical value (KO to WT ratio) is indicated under each protein band, which is representative of 3 separate experiments.
Figure 4. Hierarchical clustering analysis of RPPA protein expression to classify patient subgroups and tumor heterogeneity signatures associated with \textit{FAT1} mutation. Unsupervised clustering analysis was performed using 20 signature proteins after extraction and normalization. The hierarchical clustering heatmap revealed four molecular subtypes among the total patients with \textit{FAT1} mutant (96 patients, panel A), or wild-type (233 patients, panel B), respectively. The intensity of protein expression is indicated on the heatmaps as red for relatively high protein level, and blue for relatively low protein levels compared with tumor means. Clinical characterization of HNSCC samples is indicated at the top of the heatmap, such as HPV status (N, HPV negative, red; P, HPV positive, blue); and tumor
primary site (TSS, tissue source sites, OC, oral cavity; LX larynx; OP, oropharynx). The four protein signatures are labeled as I (blue), II (green), III (purple), and IV (red). Overall scores were calculated based upon protein levels and compared among the four subclusters within each of the two heatmaps (C for FAT1 mutant, and D for wild-type, respectively). Boxplot denoting the four subclusters of proteins is shown as I (blue), II (green), III (purple), and IV (red), as presented on the heatmaps. Medians are represented by the thick black lines in the middle of the boxplots; bars represent the 25th and 75th percentile values. *p values presented on the upper left in each panel were calculated by one-way ANOVA test for the four subcluster scores. To compare the score values of each two clusters, student T-test was performed: * indicates p-value < 0.05 of all other clusters compared with cluster I, # indicates comparison with cluster II, and + indicates comparison with cluster III.

HNSCC patients were separated into subgroups I and II, corresponding to the I/III and II/IV subclusters as presented in the heatmaps (A, B). Patient subgroup I and II are presented in red and blue color, respectively for PFS curves in (E) for patients with FAT1 mutation, and (F) for patients with FAT1 wild-type. PFS was compared using the Kaplan-Meier test. * indicates statistical significance (p=0.006) of PFS of patients with FAT1 mutation at 1200 days (E), and # indicates p<0.001 for patients with FAT1 wild-type status at 2500, 3000, and 4000 days.
Figure 5. FAT1 gene expression is associated with CNV and EGFR gene expression in HNSCC, lung SCC and cervical SCC.

FAT1 and EGFR gene expression are highly associated with HNSCC (A), lung SCC (B), and cervical SCC (C), with rank order analyzed from TCGA datasets. The associations of FAT1 gene expression with copy number variation (CNV) in HNSCC (D), lung SCC (E), and cervical SCC (F) are shown with statistical significance calculated by Spearman coefficient test with p values presented. FAT1 genetic alterations are associated with mutation count in HNSCC (G) and lung SCC (H), and associated with the aneuploidy score in lung SCC (I). P-values were calculated by Chi-Square/Kruskal-Willis test.
## Table 1.
Comparison of clinical characteristics and prognostic factors associated with FAT1 mutation status through genomic or proteomic data analysis in HNSCC patients

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Patients with Genomic Data (N=489)</th>
<th>Patients with RPPA Data (N=329)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wild type N=348</td>
<td>Mutation N=141</td>
<td>Wild type N=233</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>79 (61.24)</td>
<td>50 (38.76)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>269 (74.72)</td>
<td>91 (25.28)</td>
</tr>
<tr>
<td>Age</td>
<td>Mean</td>
<td>59</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>59</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>19</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>87</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Std Dev</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Race</td>
<td>White</td>
<td>298 (71.46)</td>
<td>119 (28.54)</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>8 (80.00)</td>
<td>2 (20.00)</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>32 (69.57)</td>
<td>14 (30.43)</td>
</tr>
<tr>
<td>Alcohol</td>
<td>No</td>
<td>106 (68.39)</td>
<td>49 (31.61)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>234 (72.45)</td>
<td>89 (27.55)</td>
</tr>
<tr>
<td>Smoker</td>
<td>Former</td>
<td>134 (66.67)</td>
<td>67 (33.33)</td>
</tr>
<tr>
<td></td>
<td>Current</td>
<td>123 (74.55)</td>
<td>42 (25.45)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>84 (74.34)</td>
<td>29 (25.66)</td>
</tr>
<tr>
<td>Stage</td>
<td>I</td>
<td>19 (76.00)</td>
<td>6 (24.00)</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>47 (68.12)</td>
<td>22 (31.88)</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>54 (71.05)</td>
<td>22 (28.95)</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>179 (70.47)</td>
<td>75 (29.53)</td>
</tr>
<tr>
<td>Subsite</td>
<td>LH</td>
<td>79 (66.39)</td>
<td>40 (33.61)</td>
</tr>
<tr>
<td></td>
<td>OC</td>
<td>200 (67.8)</td>
<td>95 (32.2)</td>
</tr>
<tr>
<td></td>
<td>OP</td>
<td>69 (92.00)</td>
<td>6 (8.00)</td>
</tr>
<tr>
<td>Lymph Nodes</td>
<td>N0</td>
<td>105 (66.46)</td>
<td>53 (33.54)</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td>109 (73.15)</td>
<td>40 (26.85)</td>
</tr>
<tr>
<td></td>
<td>N2</td>
<td>46 (71.88)</td>
<td>18 (28.13)</td>
</tr>
<tr>
<td></td>
<td>N3</td>
<td>10 (83.33)</td>
<td>2 (16.67)</td>
</tr>
<tr>
<td>HPV</td>
<td>HPV−</td>
<td>280 (67.31)</td>
<td>136 (32.69)</td>
</tr>
<tr>
<td></td>
<td>HPV+</td>
<td>68 (93.15)</td>
<td>5 (6.85)</td>
</tr>
</tbody>
</table>

1. Clinical characteristics were extracted from TCGA HNSCC datasets with genomic and proteomic (RPPA) information to calculate the prognostic factors associated with FAT1 mutation status using univariate analysis.
2. 489 HNSCC patients with genomic data were extracted from TCGA datasets. FAT1 mutation status included all genetic alterations of FAT1 molecules, including point mutation, deletion, indel insertion, translocation, and fusion.
3. 329 HNSCC patients with proteomic data were extracted from TCPA datasets (The Cancer Proteome Atlas).

*Cancer Res.* Author manuscript; available in PMC 2022 March 01.
4. The non-parametric p-value was calculated by Chi-Square test for most of the parameters, except for race, HPV status, and lymph node information, which were calculated by the Fisher’s exact test due to the small sample sizes in certain categories. Two sample t-test was calculated for patient age variation. Std Dev: standard deviation. The statistical significance with p value < 0.05 is highlighted in bold.

5. Categorical covariates are presented as the case number and the percentage of total patients (%), except for patient age information.

6. Abbreviation of HNSCC tumor primary anatomic location: LH, larynx and hypopharynx; OC, oral cavity; OP, oropharynx.

7. Lymph node category: N0: no regional lymph node metastasis; N1: metastasis in 1-3 axillary lymph nodes; N2: 4-9 lymph nodes; N3: 10 or more lymph nodes.

8. HPV status was extracted from PanSCC TCGA project, and supported by the HPV virus sequencing data.
Table 2.
Univariate association of FAT1 mutation with protein and gene expression in HNSCC different in HPV status

<table>
<thead>
<tr>
<th>RPPA Protein Name/Gene Name/Symbol</th>
<th>Chromosome Location</th>
<th>Function</th>
<th>Patient Subgroups P Value</th>
<th>mRNA Expression Total/HPV(−)/(+)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Increased protein expression in FAT1 mutant tumors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HER3_pY1289 ERBB2 RECEPTOR</td>
<td>12q13.2</td>
<td>Phosphorylated protein of HER3, a member of EGFR tyrosine kinase family (EGFR, ERBB2, NRG1, AREG)</td>
<td>0.004 0.011 &gt;0.10</td>
<td>No/No/NA 5</td>
</tr>
<tr>
<td>IRS1 INSULIN RECEPTOR SUBSTRATE 1</td>
<td>2q36.3</td>
<td>A substrate of insulin receptor tyrosine kinase, participant in insulin signaling (PI3K/AKT, MAPK)</td>
<td>0.017 0.060 &gt;0.10</td>
<td>0.0007/No/NA</td>
</tr>
<tr>
<td>CAVEOLIN1 CAVEOLIN 1 CAV1</td>
<td>7q31.2</td>
<td>Integral membrane protein, the main component of the flask-like invaginations of the plasma membrane (G proteins, ITGα9, FYN, GRB2, SHC, AXIN1, RAS, MAPK)</td>
<td>0.017 &gt;0.10 0.003</td>
<td>0.004/NA/0.017</td>
</tr>
<tr>
<td>VEGFR2 KINASE INSERT DOMAIN RECEPTOR; KDR</td>
<td>4q12</td>
<td>Binds VEGF to stimulate endothelial cell mitogenesis and migration, enhance permeability, and promote angiogenesis and vasculogenesis (VEGFA, VEGFC)</td>
<td>0.054 &gt;0.10 0.083</td>
<td>No/NA/No</td>
</tr>
<tr>
<td>PDL1 PROGRAMMED CELL DEATH 1 LIGAND 1 CD274</td>
<td>9p21.4</td>
<td>A type 1 transmembrane protein that expressed by myeloid and cancer cells. The binding of PD-L1 to the checkpoint molecule PD-1 transmits an inhibitory signal and reduces the proliferation of antigen-specific T-cells, while simultaneously reducing apoptosis in Treg cells (PD1)</td>
<td>&gt;0.10 0.063 0.053</td>
<td>NA/No/0.002</td>
</tr>
<tr>
<td>CMYC MYC PROTOONCOGENE MYC</td>
<td>8q24.21</td>
<td>An oncogene and transcription factor which is often constitutively expressed and leads to the increased expression of many genes involved in cell cycle and proliferation (MAX, E2Fs, p21, BCL2)</td>
<td>0.072 &gt;0.10 &gt;0.10</td>
<td>No/NA/NA</td>
</tr>
<tr>
<td><strong>Increased protein expression in wild type FAT1 tumors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASNS ASPARAGINE SYNTHETASE ASNS</td>
<td>7q21.3</td>
<td>An enzyme catalyzes and forms asparagine, involved in cell cycle, nutrient response, and drug resistance</td>
<td>0.001 0.005 0.096</td>
<td>0.03/0.02/No</td>
</tr>
<tr>
<td>SRC_pY527 V-SRC AVIAN SARCOMA VIRAL ONCOGENE SRC</td>
<td>20q11.23</td>
<td>The Src family of protein tyrosine kinase, phosphorylation reduces its PTK activity (ERK1/2, YAP/YES, NRG1, Erb4, KRAS, RAC)</td>
<td>0.006 0.001 &gt;0.10</td>
<td>No/NA/NA</td>
</tr>
<tr>
<td>P16INK4A CYCLIN-DEPENDENT KINASE INHIBITOR 2A CDKN2A</td>
<td>9p21.3</td>
<td>Tumor suppressor modulates two cell cycle regulator p53 and RB1 pathways</td>
<td>0.007 0.085 &gt;0.10</td>
<td>0.007/No/NA</td>
</tr>
<tr>
<td>RET_pY905 RARRANGED DURING TRANSFECTION PROTOONCOGENE RET</td>
<td>10q11.21</td>
<td>A receptor tyrosine kinases, cell-surface molecule transduces signals for cell growth and differentiation. (p38 MAPK, BCL2, BCLXL)</td>
<td>0.019 0.011 &gt;0.10</td>
<td>No/NA/NA</td>
</tr>
<tr>
<td>SCD1 STEAROYL-CoA DESATURASE SCD</td>
<td>10q24.31</td>
<td>An iron-containing enzyme, catalyzes synthesis of unsaturated fatty acids (PPARG, PI3K, MAPK)</td>
<td>0.027 0.033 &gt;0.10</td>
<td>0.019/0.037/NA</td>
</tr>
<tr>
<td>RPPA Protein Name/Gene Name/Symbol</td>
<td>Chromosome Location</td>
<td>Function</td>
<td>Patient Subgroups P Value</td>
<td>mRNA Expression Total/HPV(−)/(+)</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>MTOR_pS2448</td>
<td>1p36.22</td>
<td>A highly conserved protein kinase, involved in a key Regulator TORC1 for cell growth and proliferation and mRNA translation, and TORC2 for actin cytoskeletal rearrangement, cell survival, and cell cycle progression (PI3K, PTEN, YAP, etc)</td>
<td>0.042 0.010 &gt;0.10</td>
<td>0.097/No/NA</td>
</tr>
<tr>
<td>CYCLINE2</td>
<td>8q22.1</td>
<td>Plays a role in the G1/S portion of the cell cycle and accelerates G1 and regulates cell cycle; Aberrant expression can lead to cancer. (CDK2, p27, p21)</td>
<td>0.045 &gt;0.10 0.054</td>
<td>0.0001/NA/No</td>
</tr>
<tr>
<td>SHC_pY317</td>
<td>1q21.3</td>
<td>Encodes a signaling and transforming protein containing Src homology 2 and 3 (SH2 and SH3) domains, and activates receptor tyrosine kinase to the RAS pathway. Overexpression is linked to cancer mitogenesis, carcinogenesis, and metastasis. (EGFR, RAS, ERK, ROS)</td>
<td>&gt;0.10 0.012 0.015</td>
<td>NA/No/0.068</td>
</tr>
<tr>
<td>HER2_pY1248</td>
<td>17q12</td>
<td>Phosphorylated protein of the human EGFR family, which amplification or over-expression plays an important role in the development and progression of breast and other cancers. An important biomarker and target of cancer therapy (EGFR, HER3, HER4, MAPK, PI3K, PKC, STAT)</td>
<td>&gt;0.10 0.014 0.001</td>
<td>NA/No/No</td>
</tr>
<tr>
<td>CJUN_pS73</td>
<td>1p32.1</td>
<td>A subunit forms the AP-1 early response transcription factor, regulated by growth factors, pro-inflammatory cytokines, oxidative, and cellular stress, UV irradiation, ERK and JNK pathways (Fos, FosL, JunB, JunD)</td>
<td>&gt;0.10 0.035 0.009</td>
<td>NA/No/No</td>
</tr>
<tr>
<td>PAI1</td>
<td>7q22.1</td>
<td>PAI1 is a serine protease inhibitor that functions as the principal inhibitor of tissue plasminogen activator and urokinase, the activators of plasminogen and hence fibrinolysis. Promote cell dissemination and induce tumor vascularization and metastasis. (tPA, uPA)</td>
<td>&gt;0.10 0.038 &gt;0.10</td>
<td>NA/0.004/NA</td>
</tr>
<tr>
<td>PDCD1</td>
<td>2q37.3</td>
<td>A cell surface checkpoint molecule of the B7/CD28 superfamily, and acts as an inhibitory molecule on T cells after interacting with its ligands PDL1 and PDL2</td>
<td>0.058 0.051 &gt;0.10</td>
<td>No/No/NA</td>
</tr>
</tbody>
</table>

1. HNSCC samples were extracted from TCGA RPPA datasets. The total samples with FAT1 WT/MT are 233/96; in HPV(−) samples: 197/90; in HPV(+) samples: 31/4.

2. Molecule names in bold letters show the protein names used in RPPA.

3. Upper case text shows the full name presented OMIM (Online Mendelian Inheritance in Man). The underlined letter indicates the gene symbol.

4. Annotation of the chromosome location and function are from information provided by OMIM. The molecules in () are those molecules related to the RPPA proteins listed, through signaling pathways, protein interaction, gene regulation, or downstream molecules based on OMIM and publications.

5. "No" indicated P Value >0.10, and “NA” indicates mRNA expression was not analyzed based on no significant expression of protein levels.