Translational Oncology 14 (2021) 100988

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Original Research

# Genetic alterations associated with $^{18}$F-fluorodeoxyglucose positron emission tomography/computed tomography in head and neck squamous cell carcinoma

Sangwon Han[a,1], Jungsu S. Oh[a,1,*], Hyo Sang Lee[b], Jae Seung Kim[a]

[a] Department of Nuclear Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea
[b] Department of Nuclear Medicine, Gangneung Asan Hospital, University of Ulsan College of Medicine, Gangneung, Korea

ARTICLE INFO

*Keywords:*
Squamous cell carcinoma of head and neck
Positron emission tomography computed tomography
Glycolysis
Mutation
Machine learning

ABSTRACT

*Background:* We investigated the relationship between genetic alterations and $^{18}$F-FDG PET/CT findings in head and neck squamous cell carcinoma (HNSC).

*Methods:* Using mRNA-sequences of HNSC samples (480 patients) from the Cancer Genome Atlas (TCGA) portal, gene coexpression networks were constructed via a weighted correlation network analysis (WGCNA) algorithm, and their association with the tumor-to-blood signal ratio on $^{18}$F-FDG PET/CT data (21 patients) was explored. An elastic-net regression model was developed to estimate the PET tumor-to-blood ratio from the gene networks and to derive an FDG signature score ($FDG_{SS}$). The $FDG_{SS}$ was evaluated with regard to clinical variables and general mutational profiles, as well as alterations to oncogenic signaling pathways.

*Findings:* The $FDG_{SS}$ values differed across clinical stages ($p = 0.027$), HPV-status ($p < 0.001$), and molecular subtypes of HNSC ($p < 0.001$). Multivariate Cox regression demonstrated that $FDG_{SS}$ was an independent predictor for overall ($p = 0.019$) and progression-free survival ($p = 0.024$). $FDG_{SS}$ positively correlated with total mutation rate ($p = 0.016$), aneuploidy ($p < 0.001$), and somatic copy number alteration scores ($p < 0.001$). CDKN2A in the cell cycle pathway ($q = 0.014$) and the TP53 gene in the TP53 pathway ($q = 0.005$) showed significant differences between high and low $FDG_{SS}$ patients.

*Conclusion:* $FDG_{SS}$ based on the gene coexpression network was associated with the mutational landscape of HNSC. $^{18}$F-FDG PET/CT is therefore a valuable tool for the in vivo imaging of these cancers, being able to visualize the glucose metabolism of the tumor and allow inferences to be made on the underlying genetic alterations in the tumor.

## Introduction

Genetic alterations are the hallmark of cancer and enable malignant proliferation, angiogenesis, invasion, and resistance to cell death [1]. The differential expression of genes in tumor cells affects their phenotype and ultimately their diverse characteristics such as cancer aggressiveness, treatment response, and patient prognosis [2]. As targeted drugs for specific cancer driver genes or molecular pathways have been investigated and gradually introduced into the clinic, assessments of the genetic alteration status of tumors have become more important for therapeutic management [3]. For example, the addition of epidermal growth factor receptor inhibitors to chemotherapy or radiation therapy regimens has yielded survival benefits in randomized phase III trials [4,5]. Moreover, novel approaches targeting other oncogenic signaling pathways for head and neck squamous cell carcinoma (HNSC) have also been widely investigated [6]

Glycolysis, by which cancer cells convert glucose to lactate in an oxygen-independent manner to generate ATP, instead of performing oxidative phosphorylation, is a hallmark of the metabolic phenotype of cancer. It involves intermediates required for biosynthesizing nucleotides and fatty acids, and reducing agents [7]. A body of evidence indicates that the alteration of multiple signaling pathways resulting from oncogene and tumor suppressor gene mutations can be attributed to this metabolic reprogramming [8,9]. Recent studies reported that this metabolic switch reciprocally facilitates the epigenetic regulation of gene expression which in turn contributes to tumorigenesis [10]. $^{18}$F-fluorodeoxyglucose positron emission tomography/computed tomography ($^{18}$F-FDG PET/CT) enables the non-invasive assessment of increased glucose metabolism in a tumor, and is one of the standard imaging modalities for evaluating HNSC [11], with the semi-quantitative pa-

* Corresponding author.
*E-mail address:* jungsu_oh@amc.seoul.kr (J.S. Oh).
[1] Sangwon Han and Jungsu S. Oh are co-corresponding authors.

https://doi.org/10.1016/j.tranon.2020.100988
Received 10 September 2020; Received in revised form 2 December 2020; Accepted 4 December 2020
1936-5233/© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

rameters of FDG uptake such as maximum standardized uptake value ($SUV_{max}$) known to have significant prognostic value [12]. Accordingly, $^{18}$F-FDG PET/CT parameters are highly likely to be associated with alterations to genes and oncogenic signaling pathways in HNSC lesions. A comprehensive understanding of the link between $^{18}$F-FDG PET/CT and genetic modifications in cancer cells will therefore enhance our understanding of the biology of $^{18}$F-FDG PET/CT and cancer metabolism, knowledge of this relationship is currently lacking. Therefore, in this study, we evaluated such associations using HNSC samples from the Cancer Genome Atlas (TCGA) projects.

## Materials and methods

### *mRNA data from the TCGA-HNSC dataset*

Level three RNA-sequencing data for HNSC patient samples were downloaded from the publicly-available TCGA data portal (https://portal.gdc.cancer.gov/). These data contained 20 531 genes from 520 primary tumors, obtained using an Illumina HiSeq RNASeqV2 (Illumina, San Diego, CA, USA). The raw read counts were normalized and genes with low signal intensity were filtered out with a quantile cut-off of 0.50 using the 'TCGAbiolinks' R package [13]. Clinical data for the patients were obtained via the 'TCGAbiolinks' R package and the cBioPortal website (https://www.cbioportal.org/). Two patients were excluded as they did not have clinical follow-up data available for the determination of overall survival (OS) or progression-free survival (PFS). A further 38 patients with a history of prior treatment or with a previous synchronous malignancy were also excluded. A final cohort of 480 patients with 10 165 mRNA transcripts and available clinical data was analyzed.

### *PET/CT data and analysis*

From the Cancer Imaging Archive (TCIA, https://www.cancerimagingarchive.net/) website, we identified 29 patients in our TCGA-HNSC dataset for whom $^{18}$F-FDG PET/CT data were also available. Of these 29 patients, three with post-operative PET/CT and one for whom the Digital Imaging and Communications in Medicine (DICOM) files were not feasible for semi-quantitative analysis were excluded. An experienced nuclear medicine physician working in a blind manner extracted the maximum standardized uptake value ($SUV_{max}$) of the primary tumor using Metavol software (https://www.metavol.org/). The DICOM files on the TCIA website were acquired using a variety of PET/CT scanners and reconstruction algorithms; therefore, to compensate for multi-scanner issues, cases with a metabolic tumor volume equal to or higher than 4.2 ml according to a threshold of 40% of the $SUV_{max}$ were retrieved, and the maximum tumor-to-blood ratio ($TBR_{max}$) was calculated as the tumor $SUV_{max}$ divided by the SUVmean of the blood pool [14,15]. Finally, the $TBR_{max}$ values on $^{18}$F-FDG PET/CT were obtained for 21 patients.

In these 21 included patients, FDG was injected at a median dose of 514 MBq (range, 307–784 MBq). PET/CT images were acquired 1–2 h after administration of tracer (median, 77 min). The PET images were heterogeneous in terms of scanner modality, acquisition, and reconstruction protocol, as they were mostly acquired as part of routine care (not as part of a controlled clinical trial). Nevertheless, all the PET/CT cameras (Discovery RX [$n = 2$], LS [$n = 4$], ST [$n = 6$], and STE [$n = 7$]; GE Healthcare, Biograph Truepoint [$n = 2$]; Siemens Healthineers) were equipped with ordered subset maximum likelihood (OSEM) iterative reconstruction; although the reconstruction parameters differed across the institutions that supplied the data, most of the iteration, subset, and post-smoothing parameters used were the standard settings for each camera generation (e.g., two iterations, 15 subsets, 6 mm full-width-at-half-maximum [FWHM] Gaussian smoothing). Moreover, most of the scanners were without time-of-flight and point-spread-function modeling, resulting in a voxel size of approximately 4 mm and 7–8 mm FWHM spatial resolution. Therefore, we believe that the spatial resolution and signal-to-noise ratio did not differ substantially according to the different scanners on which the PET/CT data were acquired.

### *Exploration of $TBR_{max}$-associated gene networks by weighted correlation network analysis*

To explore the gene networks that correlated with $TBR_{max}$, weighted gene correlation networks were generated using our gene expression data after log2 transformation with the 'WGCNA' R package [16]. In bioinformatics, weighted correlation network analysis (WGCNA) has been widely applied to find gene modules (clusters of highly correlated genes) and summarize clusters by means of the module eigengene or an intramodular hub gene for relating modules to external traits (https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/). A 'signed network adjacency' matrix was first constructed from the gene expression data from the full study cohort of 480 patients using a soft thresholding power of '6′, which was selected to ensure scale-free topology and provide sufficient node connectivity. The adjacency matrix was then transformed into a 'topological overlap matrix (TOM)' to minimize the effects of noise and spurious associations, and the corresponding dissimilarity was calculated as '1 – TOM'. Hierarchical clustering of the dissimilarity (1 – TOM) was performed using an 'average'-linkage method to produce a clustering tree of genes; the branches of the clustered groups of the genes were highly interconnected. The gene network modules were identified by cutting the branches off the clustering tree using the 'dynamicTreeCut' R package, setting a cut-height value of '0.99′, a deep split of '2′, and a minimum module size of '50′ [15,16].

For the 21 patients with PET/CT data, module eigengenes, defined as the first principal component scores of the gene expression data in a given gene network module, were calculated. A Spearman correlation test was then performed between the $TBR_{max}$ and module eigengenes from each gene network module. Modules with $p$-values $< 0.05$ were defined as $TBR_{max}$-associated gene networks. Hub genes of $TBR_{max}$-associated gene networks were identified according to the Spearman correlation coefficient between the module eigengenes and the expression profiles of individual genes in the respective modules [15]. The functional profiles of $TBR_{max}$-associated gene networks were retrieved by gene ontology (GO) analyses on biological process, pathway, and molecular function using the 'TCGAanalyze_EAcomplete' function in the TCGAbiolinks R package.

### *Elastic-net model for tumor metabolism prediction*

A statistical model using elastic-net regression was built to predict the $TBR_{max}$ values from the transcriptome data for the 21 study patients with available PET/CT data. The training data were log2 transformed mRNA transcripts for the top 10 hub genes from significantly correlated modules, which were preprocessed by subtracting the means and dividing by the standard deviation (centered and scaled). As the number of predictor variables (top 10 hub genes from the $TBR_{max}$-associated gene networks) was relatively large compared with the number of outcome variables ($TBR_{max}$ values for 21 patients), elastic-net regression, a convex combination of ridge and lasso regularized regression, was chosen to deal with overfitting problems. The elastic-net model was fitted using the 'glmnet' and 'caret' R packages to simultaneously tune alpha (mixing percentage between ridge and lasso) and lambda (regularization parameter) values via leave-one-out cross-validation (LOOCV). Root mean square error (RMSE) was used as the loss function for the elastic-net model. After training, the model was applied to the full 480-patient TCGA-HNSC dataset. The output of the model was defined as the FDG signature score ($FDG_{SS}$).

The patients' $FDG_{SS}$ values were compared in terms of human papillomavirus (HPV)-positivity, molecular subtype, clinical stage, and

**Table 1**
Clinical characteristics of all patients ($n = 480$) and patients with PET/CT available ($n = 21$).

| Characteristics | TCGA-HNSC ($n = 480$) | PET/CT ($n = 21$) |
|---|---|---|
| Gender (F:M) | 123:357 (25.6%:74.4%) | 4/17 (19.0%:81.0%) |
| Age at diagnosis (y) | 60.9 ± 11.7 (19–90) | 57.0 ± 14.0 (26–85) |
| Clinical stage | | |
| 1 | 17 (3.6%) | 2 (9.5%) |
| 2 | 88 (18.8%) | 4 (19.0%) |
| 3 | 101 (21.6%) | 3 (14.3%) |
| 4 (a/b/c) | 261 (246/9/6, 55.9%) | 12 (12/0/0, 57.1%) |
| Tumor site | | |
| Hypopharynx | 8 (1.7%) | 0 (0.0%) |
| Larynx | 106 (22.1%) | 5 (23.8%) |
| Oral cavity | 106 (22.1%) | 0 (0.0%) |
| Oropharynx | 50 (10.4%) | 6 (28.6%) |
| Overlapping lesion | 62 (12.9%) | 3 (14.3%) |
| Tongue | 142 (29.7%) | 5 (23.8%) |
| Others* | 6 (1/1/3/1, 1.3%) | 2 (1/1/0/0, 9.5%) |
| Pathologic grade† | | |
| 1 | 60 (12.9%) | 0 (0.0%) |
| 2 | 285 (61.3%) | 13 (61.9%) |
| 3 | 115 (24.7%) | 8 (38.1%) |
| 4 | 5 (1.1%) | 0 (0.0%) |
| Race‡ | | |
| American Indian or Alaska native | 2 (0.4%) | 0 (0.0%) |
| Asian | 11 (2.4%) | 0 (0.0%) |
| Black or African American | 46 (9.8%) | 3 (14.3%) |
| White | 409 (87.4%) | 18 (85.7%) |

Data are expressed as number (proportion) or mean ± standard deviation (range).
*Others include palate (not specified), pharynx (not specified), lips, and mandible.
†,‡Data were available for 465 and 468 patients, respectively.

**Table 2**
Cox proportional hazard model for overall and progression-free survival outcomes in the study patients.

| Outcomes | Variables | Univariate analysis HR (95% CI) | *p* | Multivariate analysis HR (95% CI) | *p* |
|---|---|---|---|---|---|
| Overall survival | Age | 1.02 (1.01–1.04) | <0.001 | 1.02 (1.01–1.04) | <0.001 |
| | Clinical stage | 1.11 (0.94–1.30) | 0.209 | | |
| | Histologic grade | 1.01 (0.82–1.24) | 0.920 | | |
| | $FDG_{SS}$ | 1.14 (1.03–1.25) | 0.009 | 1.12 (1.02–1.23) | 0.019 |
| Progression-free survival | Age | 1.01 (0.99–1.02) | 0.356 | | |
| | Clinical stage | 1.18 (0.99–1.41 | 0.061 | 1.16 (0.97–1.38) | 0.100 |
| | Histologic grade | 1.08 (0.87–1.35) | 0.488 | | |
| | $FDG_{SS}$ | 1.14 (1.03–1.26) | 0.009 | 1.12 (1.02–1.25) | 0.024 |

All variables were regarded as continuous variables.
CI = confidence interval; HR = hazard ratio; NA = not applicable.

pathologic grade using either two-sample *t*-tests or an analysis of variance (ANOVA) test. Univariate and multivariate Cox regression analyses of OS and PFS with the baseline characteristics (age at diagnosis, clinical stage, and histologic grade) were performed. The median $FDG_{SS}$ value was used to stratify the patients into high and low $FDG_{SS}$ groups. Kaplan–Meier curves were generated to compare the OS and PFS between these two groups, and log-rank tests were used to evaluate the statistical significance of any differences.

### *Landscape of the genetic alterations with regard to the $FDG_{SS}$*

The precompiled Mutation Annotation Format (MAF) files containing somatic variant data for the TCGA-HNSC dataset obtained from the National Cancer Institute were downloaded and analyzed using the 'TCGAmutations' and 'maftools' R packages, respectively [17]. Among the 507 patients initially retrieved, we identified 462 patients that matched with our mRNA data. The total number of somatic mutations and non-silent mutations per megabase (Mb) were obtained for each sample. Mutant-allele tumor heterogeneity (MATH) scores, defined as the median absolute deviation of the mutant-allele fraction divided by the median mutant-allele fraction, were derived to infer intra-tumor genetic heterogeneity via the methodology proposed by Mroz et al. [18]. Scores for the somatic copy number alteration (SCNA) burden, including aneuploidy scores (the total sum of chromosome arm-level alteration), the fraction altered (the fraction of bases pairs present in the copy number profiles deviating from the baseline ploidy), and the number of segments (the total number of segments present in the copy number profile of each patient), were also derived using previous reports on TCGA pan-cancer analysis [19–21]. The Pearson correlation test was used to evaluate associations between the $FDG_{SS}$ and mutational scores. The total and non-silent mutation rates and the number of segments were log-transformed because of their right-skewed distributions.

Genetic alterations in the oncogenic signaling pathway were evaluated with regard to the $FDG_{SS}$. Specifically, 10 canonical signaling pathways and their member genes, which are possible key pathways related to cancer drivers or potential therapeutic targets, were assessed via comprehensive analyses of somatic mutations, copy number alterations, epigenetic silencing, and fusions and structural rearrangements of member genes showing frequent genetic alterations [22]. The alteration status was obtained at both the pathway and gene levels: $FDG_{SS}$ values were compared between samples with altered or intact oncogenic signaling pathways using a two-sample *t*-test. The frequency of alterations in the

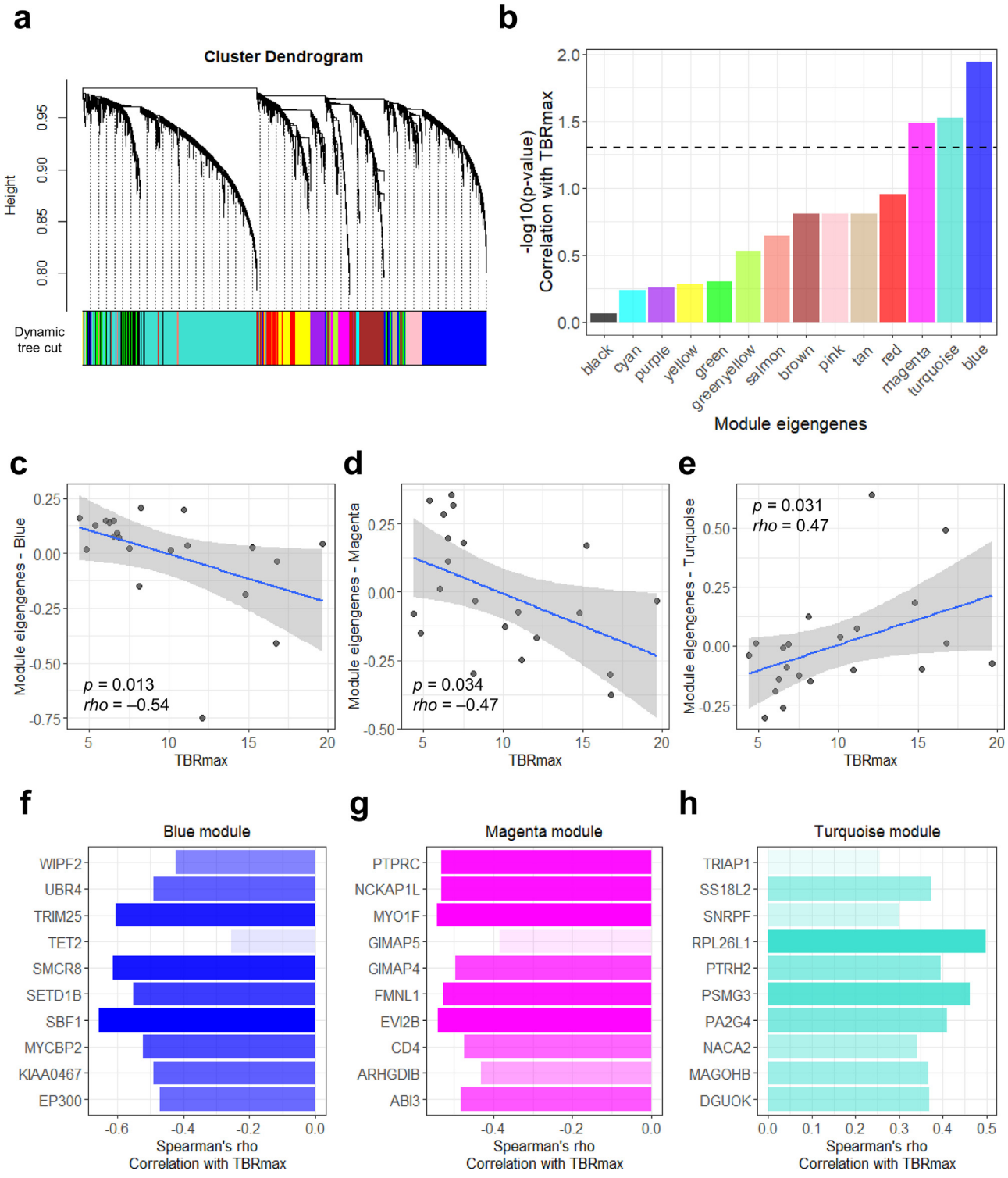


**Fig. 1.** Gene clustering dendrogram generated using WGCNA. The bars represent corresponding coexpression network module colors consisting of highly interconnected genes in the 480 HNSC study samples (a). Bar plot representing the *p*-values from correlation tests between the module eigengenes and $TBR_{max}$. The dotted line denotes a *p*-value of 0.05 (b). Scatter plots showing associations between eigengenes of significantly correlated modules and $TBR_{max}$ (c–e). Bar plots illustrating the correlation coefficients between the log2 transformed expression values of the top 10 hubgenes and $TBR_{max}$ (f–h).

member genes of significant pathways was compared between high and low $FDG_{SS}$ groups using Fisher's exact test. All *p*-values from multiple comparisons were adjusted using false discovery rate (FDR) to compensate for type I errors in the null hypothesis. *P*-values or FDR-adjusted *p*-values (*q*-values) of $< 0.05$ were defined as statistically significant. All statistical analyses were performed using R software (version 3.6.0; R Foundation for Statistical Computing, Vienna, Austria).

## Results

### *$FDG_{SS}$ estimation based on $TBR_{max}$-associated gene networks*

The characteristics of the included patients are presented in Table 1. WGCNA analysis identified 14 coexpression gene network modules (Fig. 1a). Spearman correlation tests between module eigengenes and

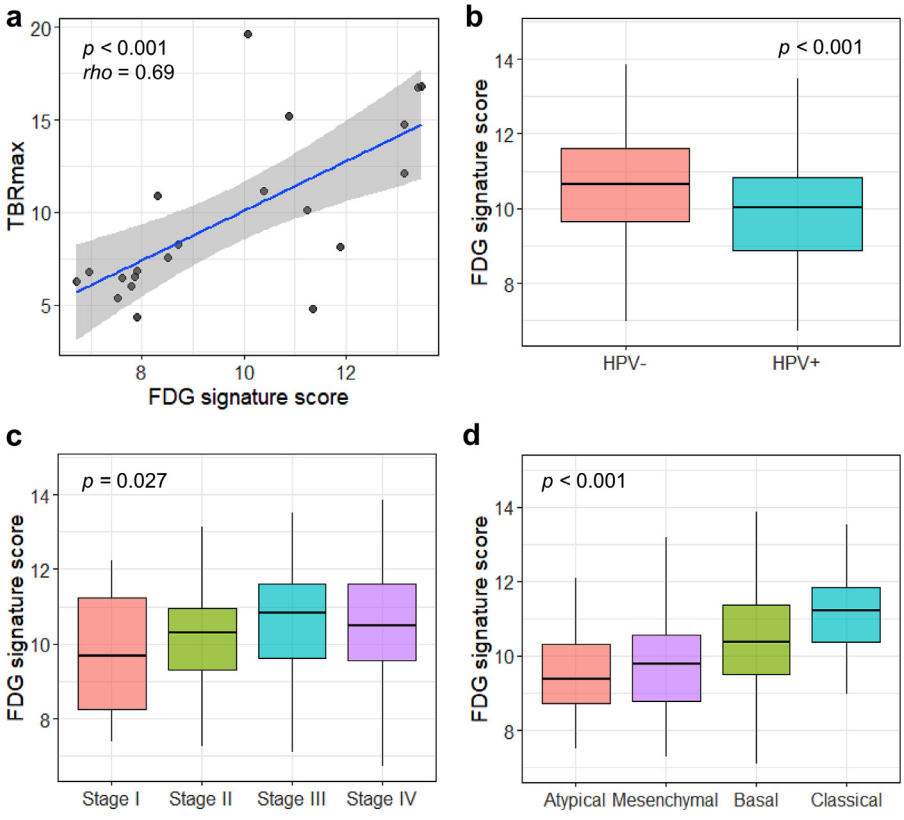


**Fig. 2.** Scatter plots showing correlations between $FDG_{SS}$ and $TBR_{max}$ in the 21 study patients with available $^{18}$F-FDG PET/CT data (a). Boxplots comparing $FDG_{SS}$ with HPV-positivity (b), clinical stage (c), and molecular subtype (d) of the HNSC tumors.

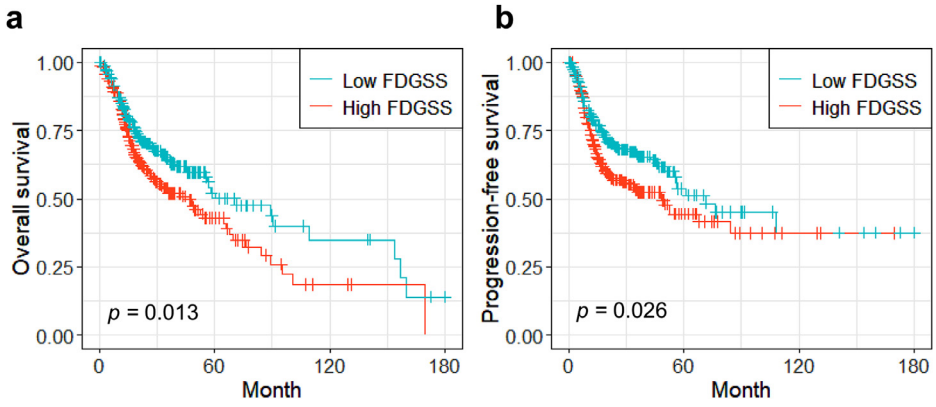


**Fig. 3.** Kaplan–Meier curves comparing overall survival (a) and progression-free survival (b) between patients stratified into high and low groups according to their median $FDG_{SS}$ value.

the $TBR_{max}$ values demonstrated that blue, magenta, and turquoise modules were $TBR_{max}$-associated gene networks (Fig. 1b–e; $p = 0.013$, $rho = -0.54$ for blue; $p = 0.034$, $rho = -0.47$ for magenta; and $p = 0.031$, $rho = 0.47$ for turquoise). The functional profiles of the blue, magenta, and turquoise modules were annotated using GO and included "G-protein coupled receptor pathway", "immune reaction", and "translation" (Supplementary file: Table S1). We selected the top 10 hub genes from each module to build a prediction model for $TBR_{max}$. The correlations between the log2-transformed expression profiles of the hub genes and the $FDG_{SS}$ are shown in Fig. 1f–h.

Elastic-net implementation with LOOCV for $TBR_{max}$ estimation revealed that the lowest RMSE of the predicted $TBR_{max}$ was 3.30 at alpha and lambda values of 0.325 and 1.61, respectively (Supplementary file: Fig. S1). The $FDG_{SS}$, which is predicted via the elastic-net model, showed moderate correlation with $TBR_{max}$ in the 21 patients (Fig. 2a; $p < 0.001$, $rho = 0.69$). The $FDG_{SS}$ was then calculated for all 480 patients from the TCGA-HNSC datasets included in this study. The $FDG_{SS}$ values were significantly higher in the HPV-negative tumors than in the HPV-positive tumors (Fig. 2b; $p < 0.001$). $FDG_{SS}$ also varied significantly across different clinical stages (Fig. 2c; $p = 0.027$) and among the four molecular subtypes of HNSC (Fig. 2d; $p < 0.001$). However, $FDG_{SS}$ did not differ across the pathologic grades of tumor ($p = 0.831$). Cox regression analyses further demonstrated that $FDG_{SS}$ is an independent predictor of OS (Table 2; adjusted HR, 1.12; $p = 0.019$) and PFS (adjusted HR, 1.12; $p = 0.024$). Kaplan–Meier curves indicated that the high-$FDG_{SS}$ group of HNSC cases showed significantly shorter OS and PFS (Fig. 3).

### *Genetic and pathway alteration signature of the $FDG_{SS}$*

A general overview of the somatic mutational profiles of our current study samples with high and low $FDG_{SS}$ is illustrated in Supplementary file: Figure S2. The log-transformed total and non-silent mutation rates showed a positive correlation with $FDG_{SS}$ (Fig. 4a–b; $r = 0.11$, $p = 0.016$ and $r = 0.11$, $p = 0.023$, respectively). The MATH scores were also significantly correlated with $FDG_{SS}$ (Fig. 4c; $r = 0.10$, $p = 0.027$). Regarding the scores for SCNA burden, the $FDG_{SS}$ values were significantly

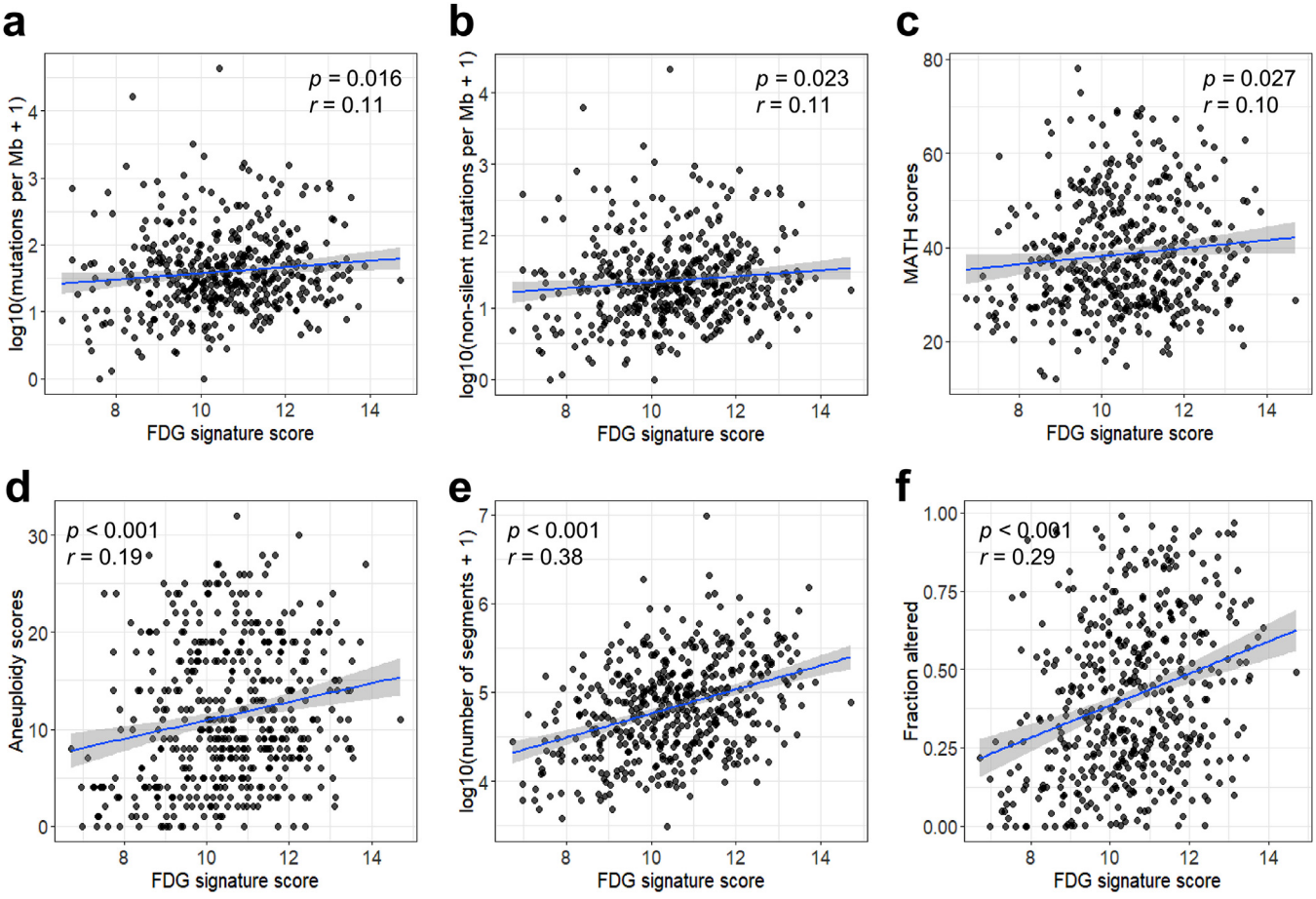


**Fig. 4.** Scatter plots of correlations between $FDG_{SS}$ and the log-transformed total (a) and non-silent mutation rates (b), MATH score (c), aneuploidy score (d), log-transformed number of segments (e), and fraction altered (f).

correlated with the aneuploidy scores (Fig. 4d; $r = 0.19$, $p < 0.001$), log-transformed number of segments (Fig. 4e; $r = 0.38$, $p < 0.001$), and the fraction altered (Fig. 4f; $r = 0.29$, $p < 0.001$).

The $FDG_{SS}$ was next evaluated in relation to alterations in the oncogenic signaling pathways (Fig. 5a). Among the 10 pathways assessed, a significantly higher $FDG_{SS}$ was indicated in samples with altered cell cycle ($q = 0.014$), TP53 ($q = 0.006$), and TGF-beta ($q = 0.024$) pathways than in samples with these pathways intact. The average numbers of alterations in the member genes of each pathway are shown according to high and low $FDG_{SS}$ in Fig. 5b. Specifically, the CDKN2A (cell cycle pathway) and TP53 (TP53 pathway) tumor suppressor genes were differentially altered in patients with a high $FDG_{SS}$ in comparison with cases showing a low $FDG_{SS}$ (Fig. 5c; 73% vs 59%, $q = 0.014$, and 78% vs 64%, $q = 0.005$, respectively). Regarding the TGF-beta pathway, the TGFBR2 and SMAD4 genes showed marginally significant alterations according to the unadjusted $p$-values, but these were not statistically significant following FDR adjustment ($q = 0.296$ for both). In addition, the MGA gene in the MYC pathway showed a significantly higher frequency of alterations in the high-$FDG_{SS}$ group (12% vs 4%, $q = 0.028$).

## Discussion

The major findings of our study are as follows: 1) $TBR_{max}$ is associated with gene networks that function in G-protein coupled receptor signaling, immune responses, and protein translation; 2) the $FDG_{SS}$, which is calculated on the basis of these gene networks, is significantly associated with the molecular subtype and clinical stage of HNSC and patient survival outcomes; 3) the $FDG_{SS}$ is positively correlated with the mutational burden and genetic heterogeneity of HNSC; and 4) specific alterations in the cell cycle, TP53, and TGF-beta pathways are significantly associated with a high $FDG_{SS}$.

Glycolysis is promoted by cell membrane receptor tyrosine kinase and G-protein coupled receptors via several intrinsic pathways, and is upregulated by the increased gene transcription and protein translation of glucose transporters and glycolytic enzymes [23,24]. The immune reactions primarily mediated by T-lymphocytes are closely related to the glucose levels in the tumor microenvironment, as glucose is required for the activation, proliferation, and differentiation of T cells [25]. HPV-negative HNSCs exhibit a higher level of glycolysis and a higher $SUV_{max}$ than the HPV-positive forms [26,27]. A previous study by Ottensmeier et al. indicated that the gene network that negatively correlates with glucose metabolism is highly active in atypical and mesenchymal subtypes of HNSC compared with the basal and classical subtypes [28]. It is also well known that a high $SUV_{max}$ is indicative of a poorer prognosis in HNSC [12]. These earlier findings support our current results that show an association between $FDG_{SS}$ and clinical variables in our HNSC study population.

Our analyses of HNSC cases also found that $FDG_{SS}$ was positively correlated with the genetic instability of tumor cells, as represented by the total and non-silent mutation rates, SCNA scores, and genetic heterogeneity. The total mutational load is known to be closely related to the cancer incidence rate, its response to treatment, and the patient's prognosis [29,30], and also to the metabolic reprograming characterized by an increase in carbohydrate metabolism, which involves comprehensive energy metabolism pathways including glycolysis, pentose phosphate pathways, gluconeogenesis, and glycogen metabolism [31]. The MATH score is a quantitative measure of intra-tumoral genetic het-

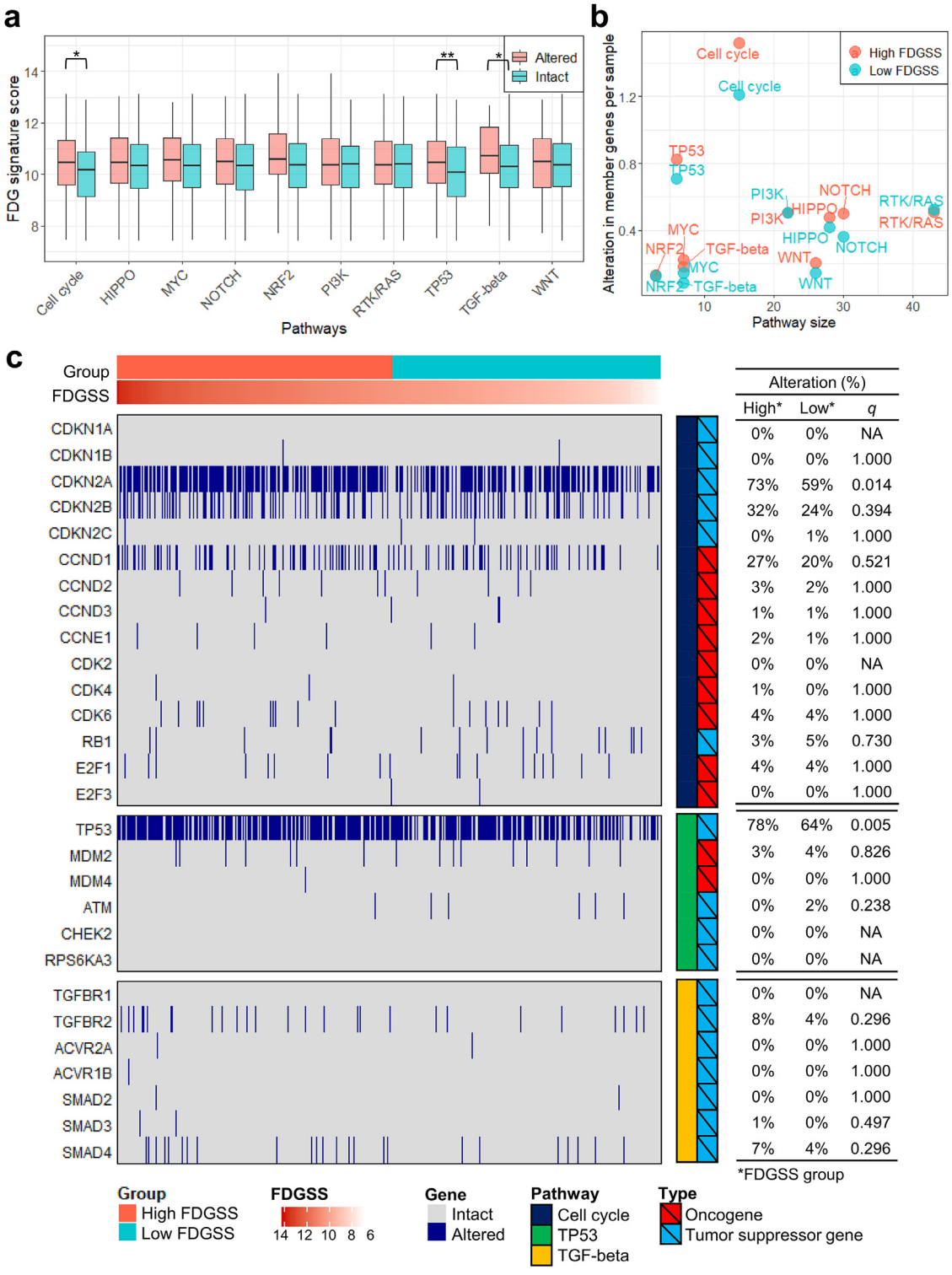


**Fig. 5.** Clustered boxplot showing a comparison of $FDG_{SS}$ with 10 oncogenic signaling pathway alterations; * and ** represent $q$-values of < 0.05 and < 0.01, respectively (a). Clustered scatter plot representing the average number of altered genes of each pathway in high and low $FDG_{SS}$ samples (b). Comprehensive heatmap illustrating member gene alterations in cell cycle, TP53, and TGF-beta pathways, in accordance with the $FDG_{SS}$ group. Each gene is described by its pathway and role in tumor development (right of the heatmap). The frequency of alterations in each gene was compared between the high and low $FDG_{SS}$ groups (c).

erogeneity reflecting the fraction of the whole-exome sequence showing mutant alleles, and is associated with the prognosis of HNSC patients [18] and the risk of metastasis in colon cancer [32]. A previous study of TCGA-HNSC samples reported that MATH scores were strongly associated with metabolic parameters on $^{18}$F-FDG PET/CT, including the metabolic tumor volume and total lesion glycolysis [33]. Aneuploidy and focal copy number alterations are considered to be different classes of SCNA and arise via different underlying mechanisms [19,34]. An abnormal number of chromosomes is associated with a higher pathologic grade and poor clinical outcomes [35]. Chromosomal aberrations cause

a wide range of gene alterations associated with metabolic disruption, a decrease in mitochondrial activity, and an elevated reactive oxygen species level, which in turn contributes to the Warburg effect. Changes in metabolism also increase the error rate in mitosis, causing genetic instability and vice versa [7,36].

The genetic instability that we found to be related to increased FDG uptake may cause dysregulation of canonical oncogenes and signaling pathways, leading to tumorigenesis. We found that a high FDG uptake is related to alterations in the TP53 gene in the TP53 pathway, and to CDKN2A in the cell cycle pathway. TP53 is a key tumor suppressor gene that downregulates glycolytic activity via 1) repression of GLUT1 and GLUT4 expression, 2) upregulation of the TIGAR gene and subsequent lowering of the fructose-2,6-bisphosphate levels in cells, and 3) activation of SCO2 gene expression to promote oxidative phosphorylation [7,8]. The CDKN2A gene encodes the p14ARF protein that is translated via alternative splicing and that stabilizes p53 and inhibits the cell cycle in a TP53-dependent manner [37]. Alterations in TP53 and CDKN2A thus co-occur, as shown in Figure S2. The PI3K/AKT/mTOR pathway is known to play a pivotal role in the Warburg effect via the activation of glucose transporter and hexokinase. However, we found no significant association between $FDG_{SS}$ and PI3K pathway alteration in our current experiments; we speculate that the much lower frequency of alterations in genes of this pathway (including AKT1–3 [0–2%], MTOR [0%], and PTEN [5%]) in our study population may have limited the statistical power of our observations.

This study had several limitations of note. First, our data includes only a small number of patients with available $^{18}$F-FDG PET/CT, and these had heterogeneous tumor origins and HPV subtypes. Our $FDG_{SS}$ calculations may have been subject to confounders or overfitted to the small sample because of the lack of external validation, even though we demonstrated that the clinical and prognostic significance of $FDG_{SS}$ in the whole population was comparable to that of the $TBR_{max}$ reported in previous studies. The TCGA-HNSC dataset is the only publicly-available data source that allows for comprehensive analyses of $^{18}$F-FDG PET/CT, gene expression, and specific gene and pathway alterations in HNSC. We hope our preliminary findings in humans will be further examined in large clinical studies in the near future. Second, the scale of the $FDG_{SS}$ was different from, and smaller than, that of $TBR_{max}$. Hence, our $FDG_{SS}$ data should not be directly interpreted or applied in the same manner as $TBR_{max}$. This limitation may have also decreased the difference between tumors with high and low FDG uptake, and may have therefore caused an underestimation of the statistical significance of our comparison between $FDG_{SS}$ and genetic alterations determined by Pearson's correlations and two-sample *t*-tests. Third, the PET/CT images from the TCIA data portal were obtained using various acquisition protocols, devices, and reconstruction algorithms, all of which can affect $SUV_{max}$. We selected patients with a measurable tumor size to minimize the impact that the resolution differences between the PET scanners would have on $SUV_{max}$, and divided $SUV_{max}$ by the mean blood pool activity to compensate for differences in imaging protocols. However, this will not have completely removed the variability. We speculate that this is one possible explanation for the moderate degree of correlation between $FDG_{SS}$ and $TBR_{max}$.

## Conclusion

Our predicted $TBR_{max}$ values ($FDG_{SS}$), based on the gene coexpression network and elastic-net model, were significantly associated with the clinical characteristics and survival of TCGA-HNSC patients. $FDG_{SS}$ showed positive correlations with total and non-silent mutation rates, genetic heterogeneity, and copy number alterations. A high $FDG_{SS}$ was related to a higher rate of alteration in the cell cycle pathway via the CDKN2A suppressor gene, and in the TP53 oncogenic signaling pathway via the TP53 tumor suppressor gene. The findings from our present comprehensive analysis may support the biologic significance of FDG uptake in HNSC from the perspective of gene and pathway alterations.

## Declaration of Competing Interest

The authors declare no conflicts of interest in relation to this study.

## Funding

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (Ministry of Science and ICT; No. NRF-2020M2D9A1094074).

## Role of the Funding Source

The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author contributions

**Sangwon Han:** conceptualization, data curation, formal analysis, investigation, methodology, software, visualization, and writing the original draft; **Jungsu S. Oh:** conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, and review and editing of the manuscript; **Hyo Sang Lee:** supervision, validation, and reviewing and editing of the manuscript; **Jae Seung Kim:** supervision, validation, and review and editing of the manuscript.

## Acknowledgements

None.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2020.100988.

## References

[1] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (5) (2011) 646–674.
[2] C.H. Chung, J.S. Parker, G. Karaca, J. Wu, W.K. Funkhouser, D. Moore, et al., Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression, Cancer Cell 5 (5) (2004) 489–500.
[3] Sawyers C.J.N. Targeted cancer therapy, *Nature* **432** (7015), 2004, 294–7.
[4] J.A. Bonner, P.M. Harari, J. Giralt, N. Azarnia, D.M. Shin, R.B. Cohen, et al., Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck, The New Eng. J. Med. 354 (6) (2006) 567–578.
[5] J.B. Vermorken, R. Mesia, F. Rivera, E. Remenar, A. Kawecki, S. Rottey, et al., Platinum-based chemotherapy plus cetuximab in head and neck cancer, The New Eng. J. Med. 359 (11) (2008) 1116–1127.
[6] M.J. Echarri, A. Lopez-Martin, R. Hitt, Targeted therapy in locally advanced and recurrent/metastatic head and neck squamous cell carcinoma (LA-R/M HNSCC), Cancers (Basel) 8 (3) (2016).
[7] R.A. Cairns, I.S. Harris, T.W. Mak, Regulation of cancer cell metabolism, Nat. Rev. Cancer 11 (2) (2011) 85–95.
[8] M. Tarrado-Castellarnau, P. de Atauri, M. Cascante, Oncogenic regulation of tumor metabolic reprogramming, Oncotarget 7 (38) (2016) 62726–62753.
[9] G.V. Walker, N. Niikura, W. Yang, E. Rohren, V. Valero, W.A. Woodward, et al., Pretreatment staging positron emission tomography/computed tomography in patients with inflammatory breast cancer influences radiation treatment field designs, Int. J. Radiat. Oncol. Biol. Phys. 83 (5) (2012) 1381–1386.
[10] X. Yu, R. Ma, Y. Wu, Y. Zhai, S. Li, Reciprocal regulation of metabolic reprogramming and epigenetic modifications in cancer, Front. Genet.. 9 (2018) 394.
[11] J. Bussink, C.M. van Herpen, J.H. Kaanders, W.J. Oyen, PET-CT for response assessment and treatment adaptation in head and neck cancer, The Lancet Oncol. 11 (7) (2010) 661–669.
[12] P. Bonomo, A. Merlotti, E. Olmetto, A. Bianchi, I. Desideri, A. Bacigalupo, et al., What is the prognostic impact of FDG PET in locally advanced head and neck squamous cell carcinoma treated with concomitant chemo-radiotherapy? A systematic review and meta-analysis, Eur. J. Nucl. Med. Mol. Imaging 45 (12) (2018) 2122–2138.
[13] A. Colaprico, T.C. Silva, C. Olsen, L. Garofano, C. Cava, D. Garolini, et al., TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data, Nucleic Acids Res. 44 (8) (2016) e71.
[14] M. Soret, S.L. Bacharach, I. Buvat, Partial-volume effect in PET tumor imaging, J. Nucl. Med. 48 (6) (2007) 932–945.
[15] K.J. Na, H. Choi, Tumor metabolic features identified by (18)F-FDG PET correlate with gene networks of immune cell microenvironment in head and neck cancer, J. Nucl. Med. 59 (1) (2018) 31–37.

[16] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinform. 9 (2008) 559.
[17] A. Mayakonda, D.-.C. Lin, Y. Assenov, C. Plass, H.P. Koeffler, Maftools: efficient and comprehensive analysis of somatic variants in cancer, Genome Res. 28 (11) (2018) 1747–1756.
[18] E.A. Mroz, A.D. Tward, R.J. Hammon, Y. Ren, J.W. Rocco, Intra-tumor genetic heterogeneity and mortality in head and neck cancer: analysis of data from the Cancer Genome Atlas, PLoS Med. 12 (2) (2015) e1001786.
[19] A.M. Taylor, J. Shih, G. Ha, G.F. Gao, X. Zhang, A.C. Berger, et al., Genomic and functional approaches to understanding cancer aneuploidy, Cancer Cell 33 (4) (2018) 676–689 .e3.
[20] T.A. Knijnenburg, L. Wang, M.T. Zimmermann, N. Chambwe, G.F. Gao, A.D. Cherniack, et al., Genomic and molecular landscape of DNA damage repair deficiency across the cancer genome atlas, Cell Rep 23 (1) (2018) 239–254 .e6.
[21] V. Thorsson, D.L. Gibbs, S.D. Brown, D. Wolf, D.S. Bortone, T.H. Ou Yang, et al., The Immune Landscape of Cancer, Immunity 48 (4) (2018) 812–830 .e14.
[22] F. Sanchez-Vega, M. Mina, J. Armenia, W.K. Chatila, A. Luna, K.C. La, et al., Oncogenic signaling pathways in the cancer genome atlas, Cell 173 (2) (2018) 321–337 .e10.
[23] P.S. Ward, C.B. Thompson, Signaling in control of cell growth and metabolism, Cold Spring Harb. Perspect. Biol. 4 (7) (2012) a006783.
[24] G. Schneditz, J.E. Elias, E. Pagano, M. Zaeem Cader, S. Saveljeva, K. Long, et al., GPR35 promotes glycolysis, proliferation, and oncogenic signaling by engaging with the sodium potassium pump, Sci Signal. 12 (562) (2019) eaau9048.
[25] C.-.H. Chang, J. Qiu, D. O'Sullivan, D. Buck Michael, T. Noguchi, D. Curtis Jonathan, et al., Metabolic competition in the tumor microenvironment is a driver of cancer progression, Cell 162 (6) (2015) 1229–1241.
[26] A.K. Tahari, K.C. Alluri, H. Quon, W. Koch, R.L. Wahl, R.M. Subramaniam, FDG PET/CT imaging of oropharyngeal squamous cell carcinoma: characteristics of human papillomavirus-positive and -negative tumors, Clin. Nucl. Med. 39 (3) (2014) 225–231.
[27] Y.S. Jung, A.J. Najy, W. Huang, S. Sethi, M. Snyder, W. Sakr, et al., HPV-associated differential regulation of tumor metabolism in oropharyngeal head and neck cancer, Oncotarget 8 (31) (2017) 51530–51541.
[28] C.H. Ottensmeier, K.L. Perry, E.L. Harden, J. Stasakova, V. Jenei, J. Fleming, et al., Upregulated glucose metabolism correlates inversely with CD8+ T-cell infiltration and survival in squamous cell carcinoma, Cancer Res. 76 (14) (2016) 4136–4148.
[29] Spigel D.R., Schrock A.B., Fabrizio D., Frampton G.M., Sun J., He J., et al. Total mutation burden (TMB) in lung cancer (LC) and relationship with response to PD-1/PD-L1 targeted therapies, *J. Clin. Oncol.* 2016;34(15_suppl):9017-9017.
[30] W. Zhang, A. Edwards, E.K. Flemington, K. Zhang, Racial disparities in patient survival and tumor mutation burden, and the association between tumor mutation burden and cancer incidence rate, Sci. Rep. 7 (1) (2017) 13639.
[31] H. Choi, K.J. Na, Pan-cancer analysis of tumor metabolic landscape associated with genomic alterations, Mol. Cancer 17 (1) (2018) 150.
[32] A. Rajput, T. Bocklage, A. Greenbaum, J.H. Lee, S.A. Ness, Mutant-allele tumor heterogeneity scores correlate with risk of metastases in colon cancer, Clin. Colorectal Cancer 16 (3) (2017) e165–ee70.
[33] J. Choi, J.-.A. Gim, C. Oh, S. Ha, H. Lee, H. Choi, et al., Association of metabolic and genetic heterogeneity in head and neck squamous cell carcinoma with prognostic implications: integration of FDG PET and genomic analysis, EJNMMI Res. 9 (1) (2019) 97.
[34] S. Negrini, V.G. Gorgoulis, T.D. Halazonetis, Genomic instability — an evolving hallmark of cancer, Nat. Rev. Molecular Cell Biol. 11 (3) (2010) 220–228.
[35] T. Davoli, H. Uno, E.C. Wooten, S.J. Elledge, Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy, Science 355 (6322) (2017) eaaf8399.
[36] D.L. Newman, S.L. Gregory, Co-operation between aneuploidy and metabolic changes in driving tumorigenesis, Int. J. Mol. Sci., 20 (18) (2019) 4611.
[37] F.J. Stott, S. Bates, M.C. James, B.B. McConnell, M. Starborg, S. Brookes, et al., The alternative product from the human CDKN2A locus, p14(ARF), participates in a regulatory feedback loop with p53 and MDM2, EMBO J. 17 (17) (1998) 5001–5014.