Oral Oncology 89 (2019) 14–22

Contents lists available at ScienceDirect

Oral Oncology

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

Obesity and genes related to lipid metabolism predict poor survival in oral squamous cell carcinoma

Qinchao Hu[a,b,1], Jianmin Peng[a,b,1], Xijuan Chen[a,b], Huan Li[c,e], Ming Song[d,e], Bin Cheng[a,b,*], Tong Wu[a,b,*]

[a] Department of Oral Medicine, Hospital of Stomatology, Sun Yat-sen University, Guangzhou, China
[b] Guangdong Provincial Key Laboratory of Stomatology, Guanghua School of Stomatology, Sun Yat-sen University, Guangzhou, China
[c] Department of Intensive Care Unit, Sun Yat-sen University Cancer Center, Guangzhou, China
[d] Department of Head and Neck Surgery, Sun Yat-sen University Cancer Center, Guangzhou, China
[e] State Key Laboratory of Oncology in South China, Collaborative Innovation Center of Cancer Medicine, Guangzhou, China

ARTICLE INFO

Keywords:
Obesity
Lipid metabolism
Oral squamous cell carcinoma
Prognosis

ABSTRACT

Objectives: Obesity is an important risk factor for several malignancies, but its effect on oral squamous cell carcinoma (OSCC) prognosis is controversial. We aimed to disclose the association between obesity and the OSCC outcome, and explore the potential of some lipid metabolism-related genes as biomarkers for prognostic prediction.

Materials and methods: A total of 576 patients diagnosed as T1/2N0M0 OSCC without prediagnosis weight loss was included in this retrospective study. These patients were grouped according to body mass index (BMI). The univariate and multivariate analysis were used to compare the progression-free survival (PFS) and disease specific survival (DSS) between groups. Propensity score matching (PSM) was adopted to minimize confounders. Data from Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) were employed to analyze the potential of some lipid metabolism-related genes for OSCC prognosis prediction.

Results: The PFS ($P = 0.023$) and DSS ($P = 0.047$) were poorer in obese patients than in normal weight ones. Obesity was an independent risk factor for PFS (Hazard Ratio = 2.016, 95% Confidence Interval 1.101–3.693, $P = 0.023$) and DSS (Hazard Ratio = 2.022, 95% Confidence Interval 1.040–3.932, $P = 0.038$). Furthermore, the PSM matched cohort analysis revealed that obesity was associated with poor prognosis of OSCC patients. Finally, 72 dysregulated lipid metabolism-related genes were identified in OSCC, and a combining signature of *TGFB1*, *SPP1*, and *SERPINE1* was defined as a biomarker for prognostic prediction.

Conclusions: Obesity is an independent risk factor for T1/2N0M0 OSCC, and a combining signature of *TGFB1*, *SPP1*, and *SERPINE1* may be applied to predict prognosis of OSCC patients.

Introduction

Obesity has become a worldwide problem threatening human health. Epidemiological evidence demonstrates that obesity is an important risk factor contributing to the carcinogenesis of esophagus, gallbladder, pancreas, colorectum, kidney, female breast (postmenopausal), endometrium and ovary [1,2]. However, the connection between obesity and oral squamous cell carcinoma (OSCC) is controversial. Some studies indicated that obesity was an independent risk factor for OSCC patients [3,4], while others argued that obesity or overweight predicted superior survival [5–9]. The unexpected protective role of obesity in cancer survival is called “obesity paradox”. This phenomenon could be attributed to the confounding, collider stratification bias, improper use of body mass index (BMI) and so on [10]. To avoid these interferential factors, only early-stage OSCC patients without prediagnosis weight loss were included in this study. Patients

Abbreviations: OSCC, oral squamous cell carcinoma; PFS, progression-free survival; DSS, disease specific survival; BMI, body mass index; PSM, propensity score matching; GEO, gene expression omnibus; TCGA, The Cancer Genome Atlas; HR, hazard ratio; CI, confidence interval; HNSCC, head and neck squamous cell carcinoma; RMA, robust multi-array average; DEGs, differentially expressed genes; GO, Gene Ontology; BP, biological processes; PPI, protein-protein interaction; ROC, receiver operating characteristic; EMT, epithelial-mesenchymal transition

* Corresponding authors at: Department of Oral Medicine, Hospital of Stomatology, Sun Yat-sen University, 56 Lingyuan Road West, Guangzhou 510060, China.
E-mail addresses: chengbin@mail.sysu.edu.cn (B. Cheng), wutong23@mail.sysu.edu.cn (T. Wu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.oraloncology.2018.12.006
Received 12 July 2018; Received in revised form 7 December 2018; Accepted 10 December 2018
Available online 14 December 2018
1368-8375/

were stratified by a BMI criterion which was reported to be suitable for Chinese population [11]. Furthermore, the propensity score matching (PSM) method was employed to control confounders and reduce selection bias to make the results more convincing.

The mechanism of obesity on OSCC prognosis is largely unknown. Obesity is a metabolic dysfunction disease, characterized by adipose tissue accumulation, lipids elevation, and insulin resistance [12]. In terms of the connection between obesity and cancer, studies mainly focused on lipid metabolism disorder and inflammatory responses under obese condition. In microenvironment, obesity-related genes and lipid signaling were dysregulated, which exacerbated the lipid metabolism disturbance and finally facilitated tumor proliferation and progression in obese patients [13–15]. Therefore, disorder of lipid metabolism and adipokine levels were suggested to be important pathogenetic factors linking obesity and cancer prognosis. To explore the expression profiles of lipid metabolism-related genes in OSCC, the public datasets from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) database were reanalyzed. The predictive values of lipid metabolism-related signatures in OSCC were also investigated.

Materials and methods

Patients and study design

After being approved by the Ethics Committee of Sun Yat-sen University Cancer Center, a total of 3166 patients with pathological confirmation of OSCC between 1997 and 2015 were collected. Prior informed consents were obtained from patients. The inclusion criteria of patients were as follow: (1) at pathological stage of T1/2N0M0, (2) received curative-intent surgical resection after diagnosis, (3) with information of height and weight at the time of diagnosis for BMI calculating. The patients were excluded on the basis of the following criteria: (1) a history of treatment for OSCC in other institutions, (2) a history of chemotherapy for other malignancy or radiotherapy for head and neck cancer, (3) prediagnosis weight loss, (4) follow-up time of less than 6 months. Finally, 576 patients were enrolled in this study.

Clinical data collection and follow-up

Clinicopathologic data including age, sex, height and weight (measured at the time of diagnosis by clinical staffs), alcohol and tobacco abuse history, diabetes mellitus status, tumor site, histologic grade, and tumor stage were extracted from the electronic medical record and summarized in Table 1. TNM stages were classified according to the 7th edition of the AJCC cancer staging manual [16]. BMI was calculated as weight in kilograms divided by height in meters squared (kg/m^2) and categorized using a criterion as low body weight ($< 18.5\,kg/m^2$), normal weight ($18.5–23.9\,kg/m^2$), overweight ($24.0–27.9\,kg/m^2$), obesity ($\geq 28\,kg/m^2$), which was reported to be optimal for Chinese [11]. The endpoints of this study were progression-free survival (PFS) and disease specific survival (DSS). PFS was defined as the time from surgery to first recurrence or metastasis of primary cancer and DSS was defined as the time from surgery to death due to OSCC.

Datasets

Two repositories (the GEO and TCGA databases) were used to identify the lipid metabolism-related genes significantly dysregulated in OSCC. Raw datasets GSE30784, GSE42743 and normalized dataset GSE31056 were download from NCBI GEO databases. The expression profiles of GSE30784 and GSE42743 were based on GPL570 (Affymetrix Human Genome U133 Plus 2.0 Array), and GSE31056 was based on GPL10526 (Affymetrix GeneChip Human Genome HG-U133 Plus 2 Array). The mRNA sequencing data and clinical information were downloaded from TCGA data portal. A total of 520 head and neck squamous cell carcinoma (HNSCC) cases were obtained with both sequencing data and corresponding clinical information. Of these, 331 OSCC patients including 331 tumor samples and 31 matched normal samples were selected for subsequent analysis. In addition, GSE41613 from GEO was employed as a validation cohort to confirm our findings.

Data processing and differential expression analyses

Raw data was pretreated by robust multi-array average (RMA) approach for normalization and background correction. Gene annotation was performed using affy R package. To analyze the differentially expressed genes (DEGs) in GSE30784 (167 tumors *vs.* 45 normal samples), GSE42743 (74 tumors *vs.* 29 normal samples), GSE31056 (23 tumors *vs.* 24 normal samples), and TCGA (331 tumors *vs.* 31 normal samples), limma package in R was performed. The significant differentially expressed mRNAs were considered as $\log_2|FC| > 1.0$ and $P < 0.05$. The common DEGs among GSE30784, GSE42743, GSE31056, and TCGA data sets were identified by intersect function in R.

Enrichment analysis and PPI network of lipid metabolism-related genes

Function enrichment analysis of common DEGs were performed on the Gene Ontology (GO; http://www.geneontology.org). A threshold of $P < 0.05$ was set to indicate significant function categories. For analysis in functional level, the DEGs were enriched in lipid metabolism-related biological processes (BP). Then, the genes enriched in these processes were extracted. Heatmap was constructed to show their expression profiles in GSE30784, GSE42743, GSE31056, and TCGA datasets. To analyze the protein-protein interaction (PPI), lipid metabolism-related DEGs were upload to STRING online database (http://string-db.org) for the retrieval of interacting genes, and the networks were visualized by Cytoscape software (http://www.cytoscape.org/). The top ten highly connected genes according to the degree were considered to be hub genes and used for overall survival prediction analysis.

Statistical analysis

To examine the statistical differences across different BMI groups, the clinical characteristics were compared using Kruskal-Wallis test, χ^2 test or Fisher's exact test, where appropriate. Survival curves were plotted by the Kaplan–Meier method and compared by the log-rank test. A Cox proportional hazard model was performed to evaluate the association between each covariate and survival in univariate and multivariate analysis. Hazard ratios (HRs) were reported with 95% confidence intervals (CIs) and *P* values. To balance the baselines and minimize the selection bias of retrospective analysis, propensity score matching analysis was performed. Logistic regression model was conducted to develop a propensity score for each patient based on the covariates including age, gender, smoking history, drinking history, diabetes mellitus status, tumor site, histologic grade, and T classification. A 1-to-1 matching (without replacement) by propensity score was performed with the match tolerance 0.05. Survival analyses for the matched groups were performed as described above.

TCGA was used as training cohorts to study the association between survival and hub gene expression. Receiver operating characteristic (ROC) curves were used to determine the optimal cutoff, with the highest score of both sensitivity and specificity, to divide patients into high expression or low expression groups. The log-rank test was then performed. Furthermore, statistically significant prognostic genes were used as a whole to construct a model to predict prognosis. A risk score was calculated by multivariate Cox regression analysis based on the gene expression. Patients were stratified into high risk group and low risk group based on the ROC curve method. Survival analysis was performed by Kaplan–Meier method and log-rank test. To validate the

Table 1
Baseline characteristics of enrolled patients.

Characteristic	All (N = 576)	Low body weight (N = 53)	Normal weight (N = 336)	Overweight (N = 154)	Obese (N = 33)	*P* value
Age, y						0.798
Median (range)	54 (20–89)	54 (20–81)	54.5 (21–89)	55.5 (23–84)	50 (29–87)	
Sex						0.513
Male	335 (58.2%)	26 (49.1%)	198 (58.9%)	90 (58.4%)	21 (63.6%)	
Female	241 (41.8%)	27 (50.9%)	138 (41.1%)	64 (41.6%)	12 (36.4%)	
Smoking						0.741
No	358 (62.2%)	33 (62.3%)	205 (61.0%)	101 (65.6%)	19 (57.6%)	
Yes	218 (37.8%)	20 (37.7%)	131 (39.0%)	53 (34.4%)	14 (42.4%)	
Drinking						0.069
No	486 (84.4%)	41 (77.4%)	280 (83.3%)	139 (90.3%)	26 (78.8%)	
Yes	90 (15.6%)	12 (22.6%)	56 (16.7%)	15 (9.7%)	7 (21.2%)	
Diabetes mellitus						0.345
No	549 (95.3%)	53 (100%)	319 (94.9%)	145 (94.2%)	32 (97.0%)	
Yes	27 (4.7%)	0 (0%)	17 (5.1%)	9 (5.8%)	1 (3.0%)	
Tumor site						0.004
Tongue	454 (78.8%)	37 (69.8%)	272 (81.0%)	116 (75.3%)	29 (87.9%)	
Mouth floor	29 (5.0%)	9 (17.0%)	14 (4.2%)	5 (3.2%)	1 (3.0%)	
Bucca cavioris	25 (4.3%)	3 (5.7%)	13 (3.9%)	9 (5.8%)	0 (0%)	
Gingiva	2 (4.2%)	1 (1.9%)	18 (5.4%)	4 (2.6%)	1 (3.0%)	
Retromolar	3 (0.5%)	0 (0%)	0 (0%)	2 (1.3%)	1 (3.0%)	
Palate	22 (3.8%)	3 (5.7%)	7 (2.1%)	11 (7.1%)	1 (3.0%)	
Lip	19 (3.3%)	0 (0%)	12 (3.6%)	7 (4.5%)	0 (0%)	
Histologic grade						0.316
Well	397 (68.9%)	31 (58.5%)	227 (67.6%)	115 (74.7%)	24 (72.7%)	
Moderate	153 (26.6%)	20 (37.7%)	92 (27.4%)	34 (22.1%)	7 (21.2%)	
Poor	26 (4.5%)	2 (3.8%)	17 (5.1%)	5 (3.2%)	2 (6.1%)	
T classification						0.096
T1	285 (49.5%)	27 (50.9%)	172 (51.2%)	65 (42.2%)	21 (63.6%)	
T2	291 (50.5%)	26 (49.1%)	164 (48.8%)	89 (57.8%)	12 (36.4%)	

prognostic signatures, GSE41613 was download and analyzed as mentioned above. Statistical analysis was performed by SPSS 23.0 software (IBM, Armonk, NY) and two-tailed value of $P < 0.05$ was considered as statistically significant.

Results

Patients and matching

Among the 3166 OSCC patients, 576 of them were included in this study. The baseline characteristics of patients are listed in Table 1. The median age at OSCC diagnosis was 54 years, and males accounted for 58.2%. Only minority of patients were current or former smokers (n = 218, 37.8%), drinkers (n = 90, 15.6%), or diabetics (n = 27, 4.7%). Most primary tumors were located at tongue (n = 454, 78.8%), with well differentiation (n = 397, 68.9%). The proportions of T1 and T2 classification were 49.5% and 50.5%, respectively. Patients were divided into 4 groups, low body weight (n = 53), normal weight (n = 336), overweight (n = 154) and obesity (n = 33), according to BMI at the time of diagnosis. In the obese patients, 63.6% of them were at T1 stage, while the proportions in low body weight, normal weight, and overweight patients were 50.9%, 51.2%, and 42.2%, respectively. There were no significant differences in the distribution of age, gender, smoking history, drinking history, diabetes mellitus status and histologic grade among the four groups. The distribution of tumor site location was significantly different among the four groups, as shown in Table 1.

Furthermore, to balance the baseline of patients and reduce selection bias, PSM method was performed. Finally, 33 normal weight patients were matched with 33 obese patients for further analysis. As shown in Table 2, the two matched cohorts had comparable characteristics.

Table 2
Baseline characteristics of the propensity score-matched cohort.

Characteristic	Normal Weight (N = 33)	Obese (N = 33)	*P* value
Age, y			
Median (range)	55 (30–79)	50 (29–87)	0.529
Sex			0.138
Male	15 (45.5%)	21 (63.6%)	
Female	18 (54.5%)	12 (36.4%)	
Smoking			0.614
No	21 (63.6%)	19 (57.6%)	
Yes	12 (36.4%)	14 (42.4%)	
Drinking			0.757
No	27 (81.8%)	26 (78.8%)	
Yes	6 (18.2%)	7 (21.2%)	
Diabetes mellitus			1.000
No	33 (100.0%)	32 (97.0%)	
Yes	0 (0.0%)	1 (3.0%)	
Tumor site			0.705
Tongue	31 (93.9%)	29 (87.9%)	
Mouth floor	0 (0.0%)	1 (3.0%)	
Gingiva	2 (6.1%)	1 (3.0%)	
Retromolar	0 (0.0%)	1 (3.0%)	
Palate	0 (0.0%)	1 (3.0%)	
Histologic grade			1.000
Well	25 (75.8%)	24 (72.7%)	
Moderate	7 (21.2%)	7 (21.2%)	
Poor	1 (3.0%)	2 (6.1%)	
T classification			0.428
T1	24 (72.7%)	21 (63.6%)	
T2	9 (27.3%)	12 (36.4%)	

Poorer prognosis in obese patients

The median follow-up time was 64 months (range, 1–283 months) in the 576 enrolled patients. The 1-, 5-, and 10-year PFS rates were 78.2%, 70.3%, and 42.3% for obese patients and 86.4%, 78.5%, and 73.3% for normal weight patients, respectively. The PFS was significantly poorer in obesity group compared with the normal weight group ($P = 0.023$, Fig. 1A). However, the overweight ($P = 0.551$) and low body weight patients ($P = 0.380$) had similar PFS compared with normal weight patients. In addition, obese patients had poorer DSS than normal weight patients (5-year DSS rates, 60.3% *vs.* 80.7%; $P = 0.047$, Fig. 1B), while

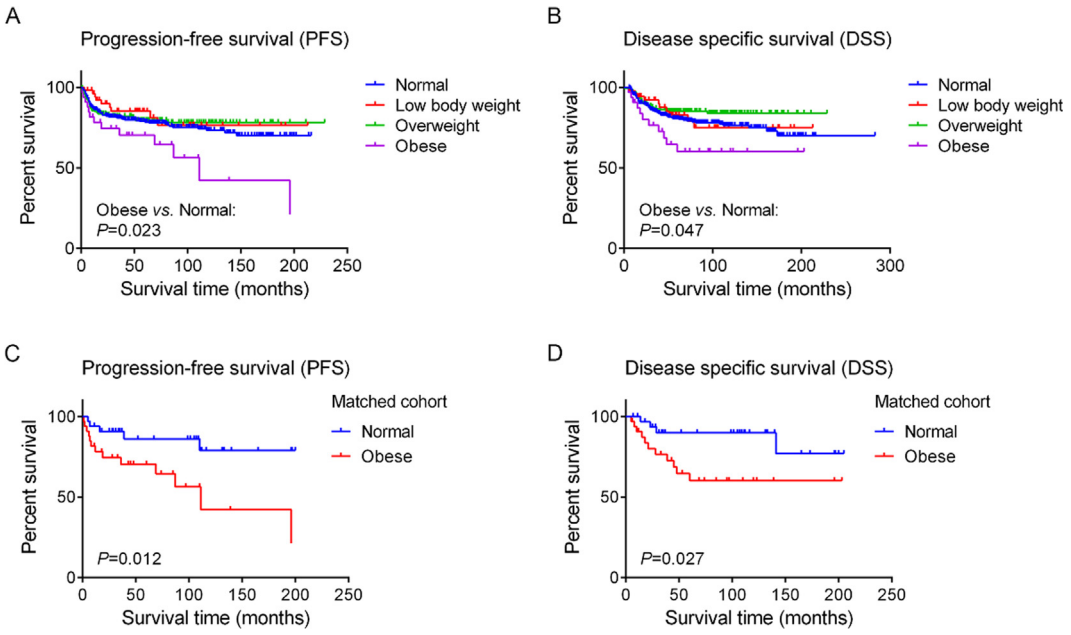


Fig. 1. Kaplan–Meier survival curves for PFS and DSS in the whole cohort (A, B) and PSM matched cohort (C, D). Survival curves are compared by the log-rank test, and $P < 0.05$ is considered as statistically significant. PFS, progression-free survival; DSS, disease specific survival; PSM, propensity score matching.

Table 3
Univariate and multivariate analyses of risk factors for progression-free survival (PFS) and disease specific survival (DSS).

Characteristic	PFS			DSS		
	Univariate Analysis (*P* value)	Multivariate analysis		Univariate Analysis (*P* value)	Multivariate analysis	
		HR (95% CI)	*P* value		HR (95% CI)	*P* value
BMI						
Normal	Ref.	Ref.		Ref.	Ref.	
Low body weight	0.380	0.707 (0.348–1.439)	0.339	0.776	0.798 (0.399–1.596)	0.523
Overweight	0.551	0.878 (0.573–1.347)	0.552	0.116	0.625 (0.381–1.027)	0.064
Obese	0.023	2.016 (1.101–3.693)	0.023	0.047	2.022 (1.040–3.932)	0.038
Age						
< 50y	Ref.	Ref.		Ref.	Ref.	
≥50y	0.505	0.821 (0.563–1.197)	0.305	0.014	1.455 (0.928–2.280)	0.102
Sex						
Male	Ref.	Ref.		Ref.	Ref.	
Female	0.939	1.022 (0.640–1.631)	0.927	0.037	0.852 (0.501–1.450)	0.556
Smoking						
No	Ref.	Ref.		Ref.	Ref.	
Yes	0.781	1.192 (0.714–1.991)	0.501	0.036	1.224 (0.710–2.110)	0.468
Drinking						
No	Ref.	Ref.		Ref.	Ref.	
Yes	0.796	0.843 (0.467–1.523)	0.572	0.135	1.037 (0.582–1.848)	0.901
Diabetes mellitus						
No	Ref.	Ref.		Ref.	Ref.	
Yes	0.009	2.458 (1.293–4.671)	0.006	0.019	2.037 (0.998–4.158)	0.051
Tumor site						
Tongue	Ref.	Ref.		Ref.	Ref.	
Mouth floor	0.627	0.752 (0.290–1.949)	0.557	0.278	1.006 (0.434–2.334)	0.988
Bucca cavioris	0.945	1.113 (0.443–2.793)	0.820	0.127	1.580 (0.712–3.506)	0.261
Gingiva	0.245	0.593 (0.183–1.923)	0.384	0.312	0.343 (0.082–1.434)	0.143
Retromolar	0.770	0.685 (0.086–5.490)	0.722	0.470	1.133 (0.136–9.458)	0.908
Palate	0.551	1.282 (0.548–2.996)	0.566	0.017	1.664 (0.803–3.446)	0.171
Lip	0.070	0.206 (0.028–1.492)	0.118	0.386	0.453 (0.110–1.864)	0.273
Histologic grade						
Well	Ref.	Ref.		Ref.	Ref.	
Moderate	0.019	1.645 (1.112–2.435)	0.013	0.246	1.155 (0.727–1.835)	0.541
Poor	0.317	1.441 (0.649–3.199)	0.370	< 0.001	2.288 (1.289–4.804)	0.007
T classification						
T1	Ref.	Ref.		Ref.	Ref.	
T2	0.975	1.107 (0.773–1.587)	0.579	0.005	1.708 (1.131–2.578)	0.011

Abbreviations: HR, hazard ratio; CI, confidence interval; Ref, reference.

Table 4
Multivariate analysis of progression-free survival (PFS) and disease specific survival (DSS) in the propensity score-matched cohort.

Characteristic	PFS			DSS		
	HR	95% CI	*P* value	HR	95% CI	*P* value
BMI						
Normal	Ref.			Ref.		
Obese	4.612	1.493–14.247	0.008	3.848	1.091–13.576	0.036
Age						
< 50y	Ref.			Ref.		
≥50y	0.923	0.303–2.812	0.888	1.120	0.289–4.347	0.869
Sex						
Male	Ref.			Ref.		
Female	0.814	0.167–3.963	0.799	0.536	0.074–3.880	0.537
Smoking						
No	Ref.			Ref.		
Yes	1.317	0.251–6.917	0.745	1.142	0.152–8.555	0.897
Drinking						
No	Ref.			Ref.		
Yes	1.978	0.446–8.779	0.370	2.611	0.454–15.026	0.282
Diabetes mellitus						
No	Ref.			Ref.		
Yes	–	–	0.995	–	–	0.998
Tumor site						
Tongue	Ref.			Ref.		
Mouth floor	–	–	0.990	1.356	0.126–14.545	0.801
Gingiva	3.414	0.344–33.888	0.294	8.382	0.680–103.336	0.097
Retromolar	–	–	0.988	–	–	0.981
Palate	–	–	0.990	2.318	0.153–35.072	0.544
Histologic grade						
Well	Ref.			Ref.		
Moderate	0.538	0.112–2.538	0.439	0.743	0.150–3.685	0.717
Poor	0.896	0.095–8.482	0.924	2.091	0.206–21.192	0.533
T classification						
T1	Ref.			Ref.		
T2	1.475	0.521–4.173	0.464	1.249	0.362–4.311	0.725

Abbreviations: HR, hazard ratio; CI, confidence interval; Ref, reference.

the differences were not statistically significant when we focused on the overweight and low body weight patients ($P = 0.116$ and 0.776, respectively, compared with normal weight patients). To identify whether obesity was an independent risk factor after adjusted by other variables, multivariate analysis was adopted. Results confirmed that obese patients took a 2.016-fold higher risk to progress (HR = 2.016, 95% CI 1.101–3.693, $P = 0.023$) and 2.022-fold higher risk to die (HR = 2.022, 95% CI 1.040–3.932, $P = 0.038$, Table 3), when compared with normal weight patients.

In the propensity score-matched cohort, similar results were found. The obese patients had poorer PFS (5-year PFS rates, 70.3% *vs.* 86.2%; $P = 0.012$, Fig. 1C) and DSS (5-year DSS rates, 60.3% *vs.* 89.8%; $P = 0.027$, Fig. 1D) than the matched normal weight patients. Multivariate analysis further revealed that obesity was an independent risk factor for PFS (HR = 4.612, 95% CI 1.493–14.247, $P = 0.008$) and DSS (HR = 3.848, 95% CI 1.091–13.576, $P = 0.036$, Table 4).

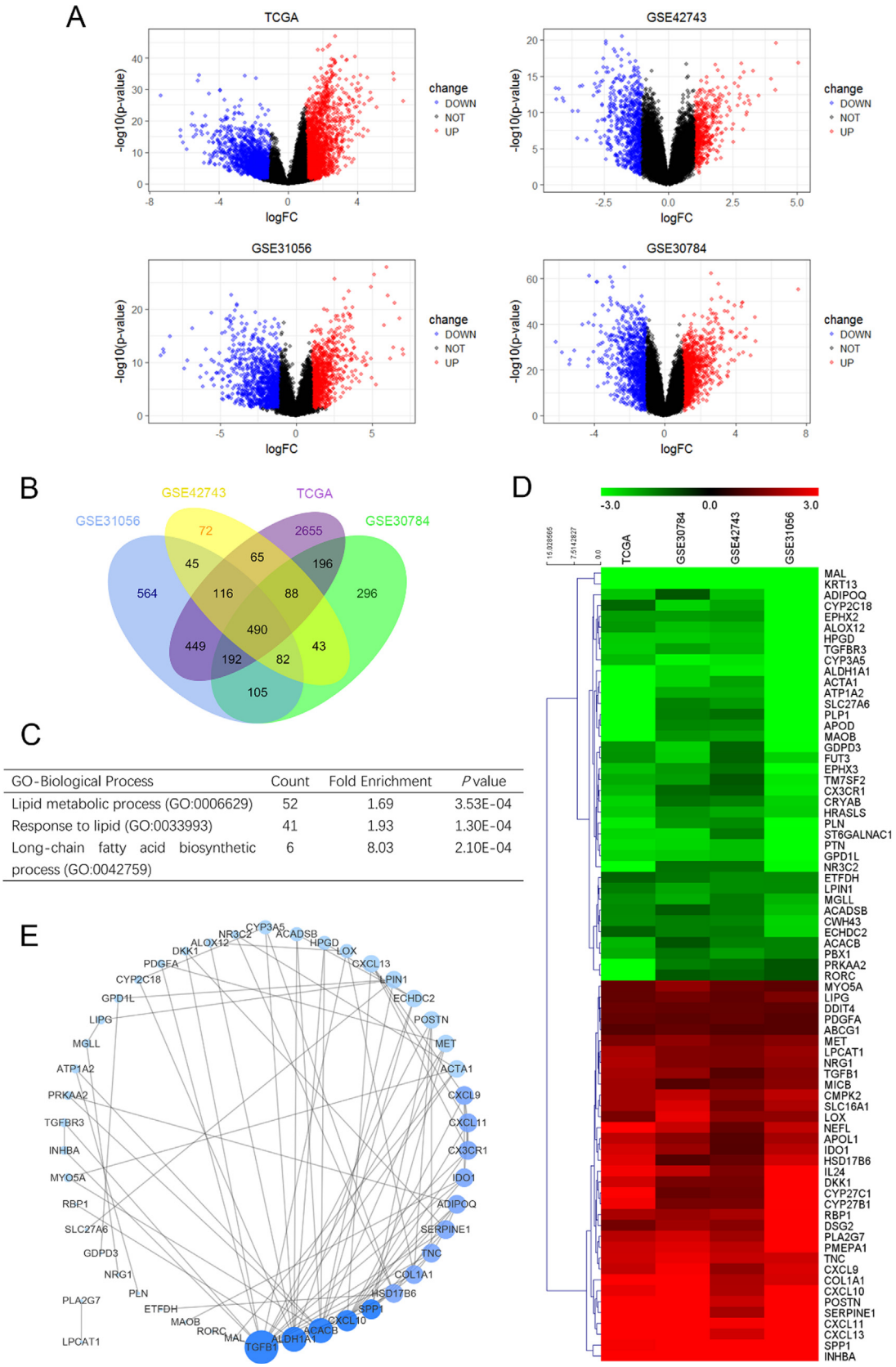
Lipid metabolism-related genes screening

This retrospective study revealed that obesity was a significant factor for poorer prognosis of OSCC, thus we suspected that the lipid metabolism-related genes might be abnormally expressed in tumors and associated with the survival of OSCC patients. Therefore, four datasets including GSE30784, GSE42743, GSE31056, and TCGA were used to find out the dysregulated genes in OSCC. The volcano plots in Fig. 2A displayed the DEGs between tumor and normal tissues in four datasets. In total, 1492 genes from GSE30784, 1001 genes from GSE42743, 2043 genes from GSE31056 and 4251 genes from TCGA intersected, and 490 common DEGs were identified (Fig. 2B). The common DEGs were performed for further functional analysis based on Gene Ontology. As indicated in Fig. 2C, DEGs were enriched in the "biological processes (BP)" categories including lipid metabolic process (GO:0006629), response to lipid (GO:0033993) and long-chain fatty acid biosynthetic process (GO:0042759), which were related to obesity or lipid metabolism. Altogether, there were 72 lipid metabolism-related genes enriched in these categories, and a heatmap was constructed to visualize their expression profiles (Fig. 2D). A total of 38 genes downregulated and 35 genes upregulated significantly in four independent cohorts concurrently. The PPI network of the lipid metabolism-related genes was shown in Fig. 2E, suggesting the top ten highly connected genes were *TGFB1, ALDH1A1, ACACB, CXCL10, SPP1, HSD17B6, COL1A1, TNC, SERPINE1,* and *ADIPOQ*.

Identification and validation of lipid metabolism-related signatures for survival prediction

Of the ten highly connected genes, ROC curve method was employed to set the cutoff point for each gene to divide 331 OSCC patients from TCGA into two groups: high expression group and low expression group, followed by Kaplan-Meier survival analysis. This procedure identified 3 prognostic genes including *TGFB1, SPP1,* and *SERPINE1* ($P < 0.05$, Fig. 3A). According to the survival plots, high expression level of *TGFB1, SPP1* and *SERPINE1*, were considered as poorer prognostic biomarkers compared with the low level ones. Subsequently, a risk score was calculated for each patient, and ROC method was employed to determine the optimal cutoff (Fig. 3B). Using the selected threshold, TCGA patients were classified into low-risk (n = 112) and high-risk (n = 219) groups. Log-rank test indicated that OSCC patients in the high-risk group had significantly shorter overall survival than those in the low-risk group ($P < 0.001$, Fig. 3C).

To confirm our findings, GSE41613 containing 97 OSCC patients was used as the validation dataset. Similarly, the results suggested that patients with high expression level of *TGFB1, SPP1,* and *SERPINE1* had poorer survival than those with low level ($P < 0.05$, Fig. 3D). The



GO-Biological Process	Count	Fold Enrichment	*P* value
Lipid metabolic process (GO:0006629)	52	1.69	3.53E-04
Response to lipid (GO:0033993)	41	1.93	1.30E-04
Long-chain fatty acid biosynthetic process (GO:0042759)	6	8.03	2.10E-04

Fig. 2. Screening of the lipid metabolism-related genes dysregulated in OSCC. (A) Volcano plot of genes that are differentially expressed between tumor and normal tissues in TCGA, GSE42743, GSE31056, and GSE30784. Blue and red color represent significant downregulated and upregulated genes, respectively. (B) Venn diagram of overlapping DEGs among four datasets. (C) Gene ontology analysis of the 490 common DEGs, indicating significant enriched lipid metabolism-related terms in biological processes. (D) Lipid metabolism-related genes were extracted from the indicated categories to construct a heatmap of expression profiles, suggesting 38 genes downregulated and 35 genes upregulated in four independent datasets concurrently. (E) PPI network of lipid metabolism-related genes. OSCC, oral squamous cell carcinoma; DEGs, differentially expressed genes; PPI, protein-protein interaction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

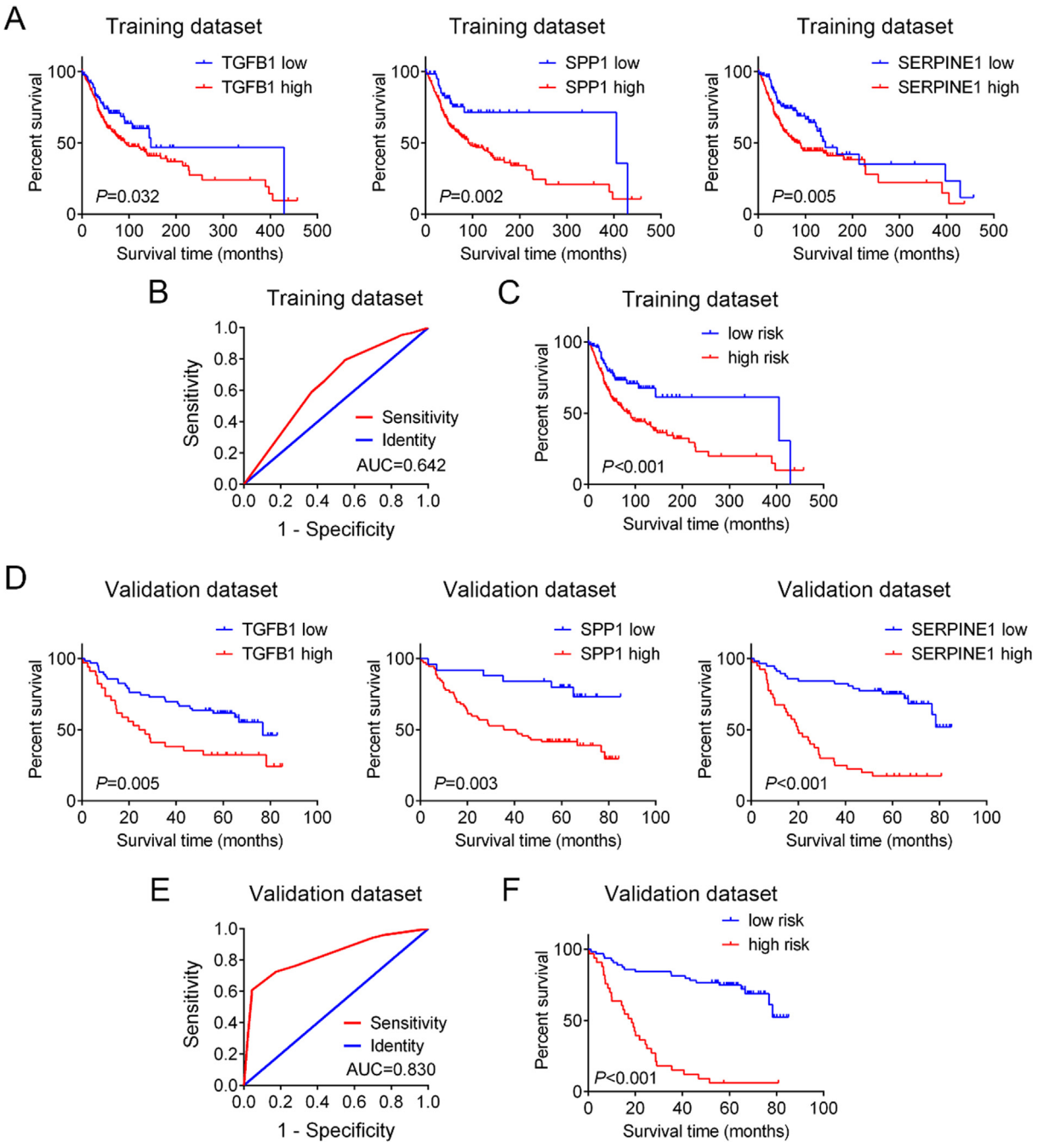


Fig. 3. Identification and validation of lipid metabolism-related signatures for survival prediction. (A) High expression of *TGFB1, SPP1,* and *SERPINE1* predict poorer overall survival in the training dataset (TCGA). (B) Patients in the TCGA cohort are divided into high risk and low risk groups according to the ROC curve. (C) Survival analysis shows that the high risk group exhibits poorer overall survival than the low risk group. (D, E, F) The survival prediction of these three genes is confirmed in the validation dataset (GSE41613). Survival curves are compared by the log-rank test, and $P < 0.05$ is considered as statistically significant. TCGA, The Cancer Genome Atlas; ROC, receiver operating characteristic.

optimal cutoff of risk score was determined by ROC method (Fig. 3E). The patients were divided and subjected to Kaplan-Meier survival analysis. In consistence with the results in training dataset, survival of the high risk group was significantly worse than that of the low risk group ($P < 0.001$, Fig. 3F).

Discussion

Recently, divergent results arose regarding the issue whether obesity is a protective or dangerous factor for OSCC outcome [3–9]. To clarify this discrepancy, we conducted this retrospective study and found that obese OSCC patients exhibited poorer PFS and DSS when compared with normal weight patients, which was further confirmed in the propensity score-matched cohort. Moreover, several lipid metabolism-related genes dysregulated in OSCC, including *TGFB1, SPP1,* and *SERPINE1*, were identified for prognostic prediction.

The plausible association between obesity and improved survival, termed “obesity paradox”, could be interpreted by several aspects. The first one is reverse causation [10]. This phenomenon appears because weight loss is a consequence rather than a cause of cancer. For OSCC, weight loss usually occurs before clinical recognition of cancer especially in late-stage patients [17]. The obese or overweight patients who

suffer from weight loss may have normal BMIs and be allocated to the reference group (normal weight), thus attenuate obesity-related death. To avoid this deviation, only patients in early-stage without pre-diagnosis weight loss were included in our study. And the results suggested that obese patients had poorer PFS and DSS. Improper use of BMI is the second cause of diverse results [18]. The classic BMI criterion is derived from western populations, which is not suitable for Asians who tend to have lower BMIs [19]. Thus, we adopted a modified BMI criterion recommended for Chinese population to make the grouping more accurate in this study [11]. Finally, collider stratification bias and detection bias also contribute to the obesity paradox [10,18]. It suggested that confounders, including smoking, sex, age, race, and tumor stage, had profound influences on the connection between obesity and cancer, which might lead to reversed conclusions. For example, obesity was a protective factor for OSCC in smokers, while this effect vanished in non-smokers [5,20,21]. Considering these confounders, the PSM method was further employed in our research to balance the baselines of patients and reduce selection bias. As a result, obesity was an independent risk factor for OSCC, which was comparable to our previous results. To sum up, after well designed and methodologically control of confounding, our study strongly supported that obesity negatively affected the prognosis of OSCC.

The mechanisms underlying the association of obesity and OSCC outcome are still largely unknown. In obesity, signaling molecules and lipid metabolites secreted by adipocytes and adipose tissue are now recognized as crucial factors for cancer progression [15,22]. Our retrospective study suggested that poorer outcome of OSCC was associated with obesity, implicated the involvement of lipid metabolism dysfunction and adipokine dysregulation in OSCC. Therefore, we explored the DEGs related to lipid metabolism in OSCC using GEO and TCGA datasets. Among them, we found that several genes of adipokines and cytokines were differentially expressed, such as *SPP1*, *SERPINE1*, and *TGFB1*, which participated in lipid metabolism and carcinogenesis as well. In the obese state, cancer cells stimulated peritumoral adipocytes to produce adipokine SPP1, which in turn facilitated tumor proliferation and invasiveness [23]. TGFB1 was engaged in tumor fibrosis that mediated by adipose stromal cells (ASCs), playing an important role in extracellular matrix remodeling and epithelial-mesenchymal transition (EMT) process [24,25]. Adipokine SERPINE1 was also involved in EMT and metastasis in several cancers [26–28]. These evidence suggested that disorder of lipid metabolism in OSCC might increase the tumor metastatic ability, which was consistent with previous studies [29,30]. In addition, genes associated with fatty acid β-oxidation, *ACACB* and *ADIPOQ*, was also dysregulated in OSCC. It implies that tumor may increase lipid utilization to provide energy for its rapid proliferation and progression.

To date, key biomarkers of the obese phenotype to predict clinical prognosis or treatment response are insufficient [25]. Therefore, we tested the prognostic prediction value of the dysregulated hub genes and put forward a three-genes-based signature (*TGFB1*, *SPP1*, and *SERPINE1*) as a predictive biomarker. This signature exhibited good predictive power not only in the training cohort but also in the validation cohort. These results would provide new insights to understand the pathogenesis linking lipid metabolism and OSCC survival, as well as to uncover potentially novel treatment targets.

There are several limitations in this study. Although multivariate analysis and PSM method are efficient for minimizing the impact of recorded confounders, the unrecorded confounders in early-stage OSCC that associated with survival are not addressed, which is determined by the retrospective design of this study. Therefore, prospectively studies are needed for further confirmation. In addition, as these data come from the Chinese population, whether the finding is generalizable to other races, such as Caucasians, needs further investigation. We would encourage expanding the study cohort to multicenter and multiracial to reveal the impact of obesity on OSCC thoroughly and generally.

Collectively, our retrospective study confirmed that obesity was associated with poorer PFS and DSS in early-stage OSCC patients, even after using PSM method to reduce selection bias and minimizing study confounders. Moreover, a combining signature of three lipid metabolism-related genes (*TGFB1, SPP1,* and *SERPINE1*) exhibited potential for clinical utility in OSCC prognosis prediction. These data may support further studies on obesity and cancer, including mechanism analysis, clinical surveillance and potential treatment targets.

Role of the funding source

This work was supported by the National Natural Science Foundation of China (No. 81630025, 81600878), the Natural Science Foundation of Guangdong Province (No. 2016A030310217, 2017A030311033), the Medical Scientific Research Foundation of Guangdong Province (No. A2016096).

Conflict of interest statement

None declared.

References

[1] Bhaskaran K, Douglas I, Forbes H, Dos-Santos-Silva I, Leon DA. Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. Lancet 2014;384:755–65.
[2] Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K. Body fatness and cancer-viewpoint of the IARC working group. N Engl J Med 2016;375:794–8.
[3] Iyengar NM, Kochhar A, Morris PG, Morris LG, Zhou XK, Ghossein RA, et al. Impact of obesity on the survival of patients with early-stage squamous cell carcinoma of the oral tongue. Cancer 2014;120:983–91.
[4] Iyengar NM, Ghossein RA, Morris LG, Zhou XK, Kochhar A, Morris PG, et al. White adipose tissue inflammation and cancer-specific survival in patients with squamous cell carcinoma of the oral tongue. Cancer 2016;122:3794–802.
[5] Karnell LH, Sperry SM, Anderson CM, Pagedar NA. Influence of body composition on survival in patients with head and neck cancer. Head Neck 2016;38(Suppl 1):E261–7.
[6] Albergotti WG, Davis KS, Abberbock S, Bauman JE, Ohr J, Clump DA, et al. Association of pretreatment body mass index and survival in human papillomavirus positive oropharyngeal squamous cell carcinoma. Oral Oncol 2016;60:55–60.
[7] Zou L, Liu T, Yang A. Metabolic syndrome is associated with better prognosis in patients with tongue squamous cell carcinoma. Chin J Cancer 2015;34:184–8.
[8] Arthur AE, Peterson KE, Rozek LS, Taylor JM, Light E, Chepeha DB, et al. Pretreatment dietary patterns, weight status, and head and neck squamous cell carcinoma prognosis. Am J Clin Nutr 2013;97:360–8.
[9] Park SM, Lim MK, Shin SA, Yun YH. Impact of prediagnosis smoking, alcohol, obesity, and insulin resistance on survival in male cancer patients: National Health Insurance Corporation Study. J Clin Oncol 2006;24:5017–24.
[10] Park Y, Peterson LL, Colditz GA. The plausibility of obesity paradox in cancer-point. Cancer Res 2018;78:1898–903.
[11] Zhou BF. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults–study on optimal cut-off points of body mass index and waist circumference in Chinese adults. Biomed Environ Sci 2002;15:83–96.
[12] Moore LL, Chadid S, Singer MR, Kreger BE, Denis GV. Metabolic health reduces risk of obesity-related cancer in framingham study adults. CancerEpidemiol Biomarkers Prev 2014;23:2057–65.
[13] Al-Jawadi A, Moussa H, Ramalingam L, Dharmawardhane S, Gollahon L, Gunaratne P, et al. Protective properties of n-3 fatty acids and implications in obesity-associated breast cancer. J Nutr Biochem 2018;53:1–8.
[14] Balaban S, Shearer RF, Lee LS, van Geldermalsen M, Schreuder M, Shtein HC, et al. Adipocyte lipolysis links obesity to breast cancer growth: adipocyte-derived fatty acids drive breast cancer cell proliferation and migration. CancerMetab 2017;5:1.
[15] Nieman KM, Romero IL, Van Houten B, Lengyel E. Adipose tissue and adipocytes support tumorigenesis and metastasis. Biochim Biophys Acta 2013;1831:1533–41.
[16] Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 2010;17:1471–4.
[17] Datema FR, Ferrier MB, Baatenburg DJR. Impact of severe malnutrition on short-term mortality and overall survival in head and neck cancer. Oral Oncol 2011;47:910–4.
[18] Lennon H, Sperrin M, Badrick E, Renehan AG. The obesity paradox in cancer: a review. Curr Oncol Rep 2016;18:56.
[19] Zheng W, McLerran DF, Rolland B, Zhang X, Inoue M, Matsuo K, et al. Association between body-mass index and risk of death in more than 1 million asians. N Engl J Med 2011;364:719–29.
[20] Nieto A, Sanchez MJ, Martinez C, Castellsague X, Quintana MJ, Bosch X, et al. Lifetime body mass index and risk of oral cavity and oropharyngeal cancer by smoking and drinking habits. Br J Cancer 2003;89:1667–71.

[21] Franceschi S, Dal Maso L, Levi F, Conti E, Talamini R, La Vecchia C. Leanness as early marker of cancer of the oral cavity and pharynx. Ann Oncol 2001;12:331–6.
[22] Park J, Morley TS, Kim M, Clegg DJ, Scherer PE. Obesity and cancer-mechanisms underlying tumour progression and recurrence. Nat Rev Endocrinol 2014;10:455–65.
[23] Ribeiro RJ, Monteiro CP, Cunha VF, Azevedo AS, Oliveira MJ, Monteiro R, et al. Tumor cell-educated periprostatic adipose tissue acquires an aggressive cancer-promoting secretory profile. CellPhysiol Biochem 2012;29:233–40.
[24] Park J, Euhus DM, Scherer PE. Paracrine and endocrine effects of adipose tissue on cancer development and progression. Endocr Rev 2011;32:550–70.
[25] Donohoe CL, Lysaght J, O'Sullivan J, Reynolds JV. Emerging concepts linking obesity with the hallmarks of cancer. Trends Endocrinol Metab 2017;28:46–62.
[26] Zhang W, Xu J, Fang H, Tang L, Chen W, Sun Q, et al. Endothelial cells promote triple-negative breast cancer cell metastasis via PAI-1 and CCL5 signaling. FASEB J 2018;32:276–88.
[27] Lin X, Lin BW, Chen XL, Zhang BL, Xiao XJ, Shi JS, et al. PAI-1/PIAS3/Stat3/miR-34a forms a positive feedback loop to promote EMT-mediated metastasis through Stat3 signaling in Non-small cell lung cancer. Biochem Biophys Res Commun 2017;493:1464–70.
[28] Strong AL, Ohlstein JF, Biagas BA, Rhodes LV, Pei DT, Tucker HA, et al. Leptin produced by obese adipose stromal/stem cells enhances proliferation and metastasis of estrogen receptor positive breast cancers. Breast Cancer Res 2015;17:112.
[29] Pascual G, Avgustinova A, Mejetta S, Martín M, Castellanos A, Attolini CS, et al. Targeting metastasis-initiating cells through the fatty acid receptor CD36. Nature 2017;541:41–5.
[30] Lai Y, Liu H, Chiang W, Chen T, Chu LJ, Yu J, et al. MiR-31-5p-ACOX1 axis enhances tumorigenic fitness in oral squamous cell carcinoma via the promigratory prostaglandin E2. Theranostics 2018;8:486–504.