Original Article

Analysis of methylation-driven genes for predicting the prognosis of patients with oral squamous cell carcinoma

Jun Chen[1#], Zejun Dong[2,3#], Biaodong Li[1], Zhiliang Nie[1], Jiaxuan Qiu[1]

[1]Department of Oral and Maxillofacial Surgery, The First Affiliated Hospital, Jiangxi Medical College, Nanchang University, Nanchang, China; [2]Innovation Center for Diagnostics and Treatment of Thalassemia, Nanfang Hospital, Southern Medical University, Guangzhou, China; [3]Department of Medical Genetics, School of Basic Medical Sciences, Southern Medical University, Guangzhou, China

Contributions: (I) Conception and design: J Qiu; (II) Administrative support: J Chen, Z Dong; (III) Provision of study materials or patients: J Chen, Z Dong; (IV) Collection and assembly of data: B Li, Z Nie; (V) Data analysis and interpretation: J Chen, Z Dong; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]The authors contributed equally to this work.

Correspondence to: Jiaxuan Qiu, MD, PhD. Department of Oral and Maxillofacial Surgery, The First Affiliated Hospital, Jiangxi Medical College, Nanchang University, No. 17, Yongwaizheng Street, Donghu District, Nanchang 330006, China. Email: qiujiaxuan@163.com.

Background: Oral squamous cell carcinoma (OSCC) is a highly aggressive malignancy that is characterized by early distant metastasis and poor prognosis. DNA methylation plays an important role in the etiology and pathogenesis of OSCC. This study aimed to identify methylation-driven genes through bioinformatics analysis as potential biomarkers for early diagnosis and prognostic assessment of OSCC.

Methods: Methylation data, RNA sequencing (RNA-seq) data and clinical prognosis information of OSCC patients were retrieved from The Cancer Genome Atlas (TCGA) database. The R packages MethylMix were employed to analyze the correlation between methylation status and corresponding gene expression in tumor and normal tissues to obtain methylation-driven genes. Univariate Cox regression analysis was developed to further screen methylation-driven genes associated with the prognosis of OSCC patients. Subsequently, multivariate Cox regression analysis was utilized to construct a linear prognostic risk prediction model. Furthermore, a combined survival analysis integrating methylation and gene expression was performed to investigate the prognostic value.

Results: A total of 374 differentially expressed methylation-driven genes were identified. Seven methylation-driven genes (*BST2*, *KRT15*, *ZNF134*, *NT5E*, *GSTA7P*, *NAPRT*, and *GOLPH3L*) were found to be significantly associated with patient prognosis. Additionally, four methylation-driven genes (*BST2*, *KRT15*, *ZNF134* and *NAPRT*) were used to construct a linear prognostic risk prediction model for OSCC patients. Furthermore, a combined Kaplan-Meier survival analysis revealed that three methylation-driven genes (*ZKSCAN7*, *MFF*, *ZNF134*) alone can be used as independent prognostic markers or drug targets.

Conclusions: Our findings facilitate a better understanding of molecular mechanisms of OSCC and provide potential biomarkers of early diagnosis, precision treatment and prognosis evaluation.

Keywords: Oral squamous cell carcinoma (OSCC); The Cancer Genome Atlas (TCGA); methylation-driven genes; biomarkers; overall survival rate

Submitted Dec 14, 2023. Accepted for publication Apr 28, 2024. Published online Jun 20, 2024.
doi: 10.21037/tcr-23-2303
View this article at: https://dx.doi.org/10.21037/tcr-23-2303

Introduction

Head and neck cancer (HNC) ranks as the seventh most prevalent malignant tumor globally (1). The majority of HNC cases are head and neck squamous cell carcinomas (HNSCC) (2). Among HNSCC, oral squamous cell carcinoma (OSCC) represents the most common histological subtype, with an annual incidence exceeding 300,000 cases (3). OSCC is characterized by its high malignancy, aggressiveness, and propensity for distant metastasis, leading to a poor prognosis (4). Early diagnosis of OSCC is challenging due to its insidious onset, resulting in a significant proportion of patients being diagnosed at advanced stages (5). Despite advancements in diagnostic and treatment approaches over the past few decades, the five-year overall survival rate remains below 50 percent (6). Timely diagnosis of OSCC enables surgical resection and improves patient survival rates (7). Additionally, the identification of specific prognostic markers and the establishment of a reliable prognostic assessment system are crucial for accurate prediction of patient life expectancy (8).

Epigenetics, a subdiscipline of genetics, investigates heritable changes in gene expression without alterations in the nucleotide sequence (9). It encompasses various mechanisms such as DNA methylation, histone modification, and regulation by non-coding RNAs (10). DNA methylation, one of the extensively studied epigenetic modifications, has been implicated in the occurrence and progression of numerous malignant tumors (11-13). Consequently, an increasing number of researchers are focusing on DNA methylation studies, aiming to discover specific methylation markers that can enhance early diagnosis rates and predict tumor prognosis (14-17).

The Cancer Genome Atlas (TCGA) database, renowned as the largest cancer genome repository globally, not only has genetic variant and epigenetic data, but also offers comprehensive and standardized clinical information of patients (18). MethylMix, an R-based algorithm, integrates DNA methylation data from both normal and tumor samples and correlates them with gene expression data to identify potential methylation-driven genes (19). In this study, we extracted DNA methylation and RNA sequencing (RNA-Seq) data of OSCC from the TCGA database and employed the MethylMix algorithm to determine specific methylation-driven genes associated with OSCC. Subsequently, we identified seven methylation-driven genes significantly associated with patient prognosis using univariate Cox regression analysis. Furthermore, four methylation-driven genes were selected via multivariate Cox regression analysis to construct a linear prognostic risk assessment model. Lastly, we conducted a combined survival analysis of DNA methylation and gene expression to shed light on the exploration of novel methylation-driven genes as diagnostic markers and prognostic indicators for OSCC. We present this article in accordance with the TRIPOD reporting checklist (available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2303/rc).

Highlight box

Key findings

- A linear prognostic risk assessment model based on methylated-driven genes effectively predicts the prognosis of oral squamous cell carcinoma (OSCC) patients.

What is known and what is new?

- DNA methylation plays an important role in the process of tumorigenesis and development, but few studies have explored the impact of methylation-driven genes on the prognosis of OSCC patients.
- Based on OSCC methylation-driven genes, an effective linear prognostic risk assessment model for OSCC patients was established.

What is the implication, and what should change now?

- A linear prognostic risk assessment model based on methylation-driven genes can effectively predict the prognosis of OSCC patients. In addition, specific methylation-driven genes can facilitate early diagnosis and targeted therapy for OSCC patients.
- Given the manuscript's findings on OSCC-related methylation-driven genes, early diagnostic reagents and targeted drugs for methylation-driven genes should be investigated.

Methods

Data collection

The data for this study were all obtained from the TCGA HNSCC datasets (https://portal.gdc.cancer.gov/). Among them, the DNA methylation data in the TCGA database were generated using the Illumina Infinium Human Methylation450k platform comprising 390 samples (34 normal samples and 356 cancer samples). The transcriptome profiling data were retrieved from 32 normal samples and 341 tumor samples from the IlluminaHiSeq RNASeq platform. Additionally, clinical data from 307 OSCC patients with complete information were downloaded. The study was conducted in accordance with

the Declaration of Helsinki (as revised in 2013).

Identification of methylation-driven genes

TCGA transcriptome data were normalized and analyzed using the limma R package in R software to identify differentially expressed genes (DEGs). The R package MethylMix was used to integrate and analyze methylation data and gene expression data based on the Wilcoxon rank sum test to identify methylation-driven genes. Three types of datasets were analyzed in MethylMix as input: DNA methylation data from the normal group, DNA methylation data from the tumor group, and messenger RNA (mRNA) expression data from the tumor group. Subsequently, DNA methylation data and mRNA expression data were integrated using R for further analysis. In this study, 341 OSCC samples with matched DNA methylation and mRNA expression data as well as 34 non-OSCC samples with DNA methylation data were integrated for correlation analysis. The analysis was conducted using the Wilcoxon rank sum test, with significance thresholds set at $\log_2$fold change (FC) >0, P<0.05, and Cor <–3. Next, we calculated by the correlation between DNA methylation and mRNA expression. Methylation-driven genes were identified using the β-mixed model. Clustering of the methylation-driven genes was performed using R software.

Functional enrichment and pathway analysis of methylation-driven genes

Functional-enrichment analysis of methylation-driven genes including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed using R. A statistical significance threshold at P<0.05 was applied for both GO and KEGG pathway enrichment analyses. The significant results were visualized using the GOCircle as well as GOChord plotting of GOplot R package.

Construction of prognostic risk assessment model in OSCC

DNA methylation and mRNA expression data and the corresponding clinical data of OSCC were merged using a Perl script. Univariate Cox regression analysis was performed to further screen methylation-driven genes correlated with the prognosis of OSCC patients. Subsequently, multivariate Cox regression analysis was employed to construct the prognostic risk assessment model. OSCC patients were divided into high- and low-risk cohorts based on the median risk score. Kaplan-Meier survival curves for the two groups were generated using the survival R packages to estimate the predictive ability of the risk model.

Independent prognostic analysis

A total of 307 OSCC patients with complete clinical information were retrieved from the TCGA database. The patients' survival status was recorded as 0 for survival and 1 for death. Clinical grades were categorized as follows: G1 as 1, G2 as 2, and G3 as 3. Similarly, the tumor, node, metastasis (TNM) clinical stages, a system frequently utilized in medical oncology for cancer staging based on tumor extent (T), regional lymph node involvement (N), and distant metastasis presence (M), were assigned numerical values: stage I as 1, stage II as 2, stage III as 3, and stage IV as 4. Perl scripts were utilized to integrate the clinical information and the corresponding risk values. Subsequently, the integrated data underwent univariate and multivariate Cox independent prognostic analyses in R. Statistical significance was determined at a significance level of P<0.05. Receiver operating characteristic (ROC) curves were plotted for each factor using R.

Combined survival analysis of DNA methylation and gene expression

To further investigate the effect of the methylation status and gene expression level of the identified methylation-driven key genes on OSCC patient prognosis, a combined survival analysis of DNA methylation and gene expression was performed using the survival R package. Statistical significance was determined using the log-rank test, with a significance level of P<0.01.

Statistical analysis

All statistical analyses were performed in R software (3.5.0). Methylation data and transcriptome data were analyzed using the MethylMix R package to identify methylation-driven genes based on the Wilcoxon rank sum test, P<0.05 was considered statistically significant. Statistical significance in combined survival analysis of DNA methylation and gene expression was determined using the log-rank test, and P<0.01 was considered statistically significant.

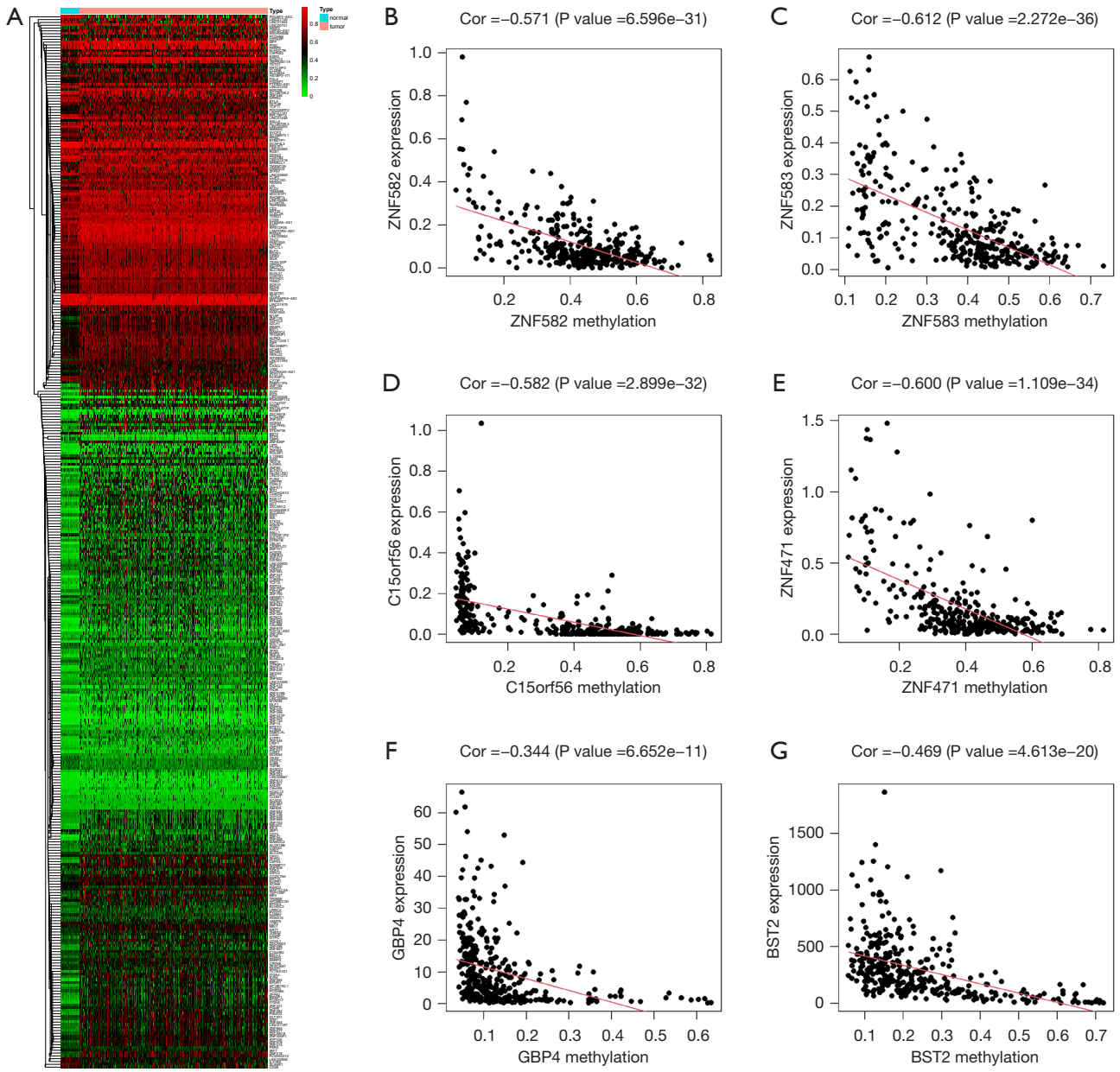

Figure 1 Heat map of methylation-driven genes and correlation between DNA methylation and mRNA expression. (A) The hierarchical clustering heat map of methylation-driven genes in OSCC. The color ranges from green to red in order to indicate low to high level of gene methylation. (B-G) The correlation between the degree of methylation of methylation-driven genes and their gene expression level in OSCC. Cor, correlation; mRNA, messenger RNA; OSCC, oral squamous cell carcinoma.

Results

Identification of methylation-driven genes in OSCC

A total of 374 methylation-driven genes were identified using MethylMix R package as shown in https://cdn.amegroups.cn/static/public/tcr-23-2303-1.xlsx (logFC >0, P<0.05, Cor <−0.3). To visualize the differential expression of methylation-driven genes in each sample, a heat map was generated using R software (*Figure 1A*). Six representative methylation-driven genes (*ZNF582*, *ZNF583*, *C15orf56*, *ZNF471*, *GBP4*, and *BST2*) (*Figure 1B-1G*) demonstrated a significant negative correlation between the degree of methylation and gene

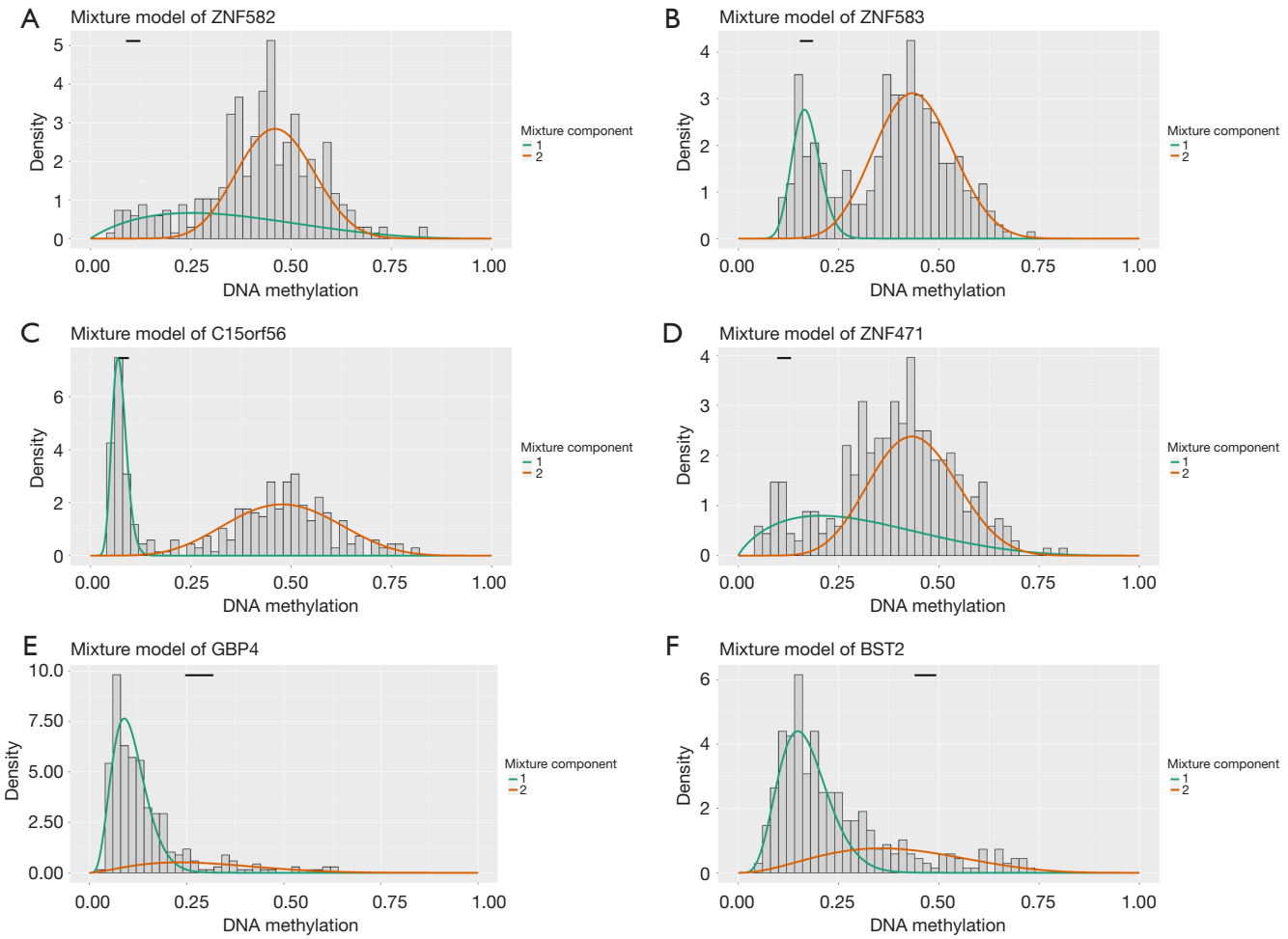

Figure 2 Methylation degree distribution of methylation-driven genes in OSCC samples and normal samples. The histogram in the figure illustrates the methylation distribution of methylation-driven genes in tumor samples. The green curve and red curve represent the distribution of methylation levels of methylation-driven genes in cancer samples, and the black horizontal line indicates the distribution of methylation-driven genes in normal samples (A-F). OSCC, oral squamous cell carcinoma.

expression level (P<0.05). The distribution of methylation degrees is shown in *Figure 2*. *ZNF582* (*Figure 2A*), *ZNF583* (*Figure 2B*), *C15orf56* (*Figure 2C*), and *ZNF471* (*Figure 2D*) were hypermethylated in OSCC samples and hypomethylated in normal samples. Conversely, *GBP4* (*Figure 2E*) and *BST2* (*Figure 2F*) exhibited hypomethylation in OSCC patients and hypermethylation in normal patients.

Functional enrichment and pathway analysis of methylation-driven genes in OSCC

To further explore the molecular functions and mechanisms of methylation-driven genes in OSCC, we conducted functional enrichment and pathway analysis using R (https://cdn.amegroups.cn/static/public/tcr-23-2303-2.xlsx, https://cdn.amegroups.cn/static/public/tcr-23-2303-3.xlsx). The results revealed the enrichment of methylation-driven genes in various pathways, biological processes (BPs) and cellular components (CCs). Functional analysis revealed that eight GO terms (cell-cell adhesion via plasma-membrane adhesion molecules, sinoatrial node development, positive chemotaxis, positive regulation of interferon-gamma production, defense response to virus, defense response to symbiont, positive regulation of B cell proliferation, positive regulation of neuroblast proliferation) displayed significant differences (P<0.05) (*Figure 3A*). The top 40 methylation-driven genes associated with the most significant GO terms were shown in the GOChord plot (*Figure 3B*). Additionally,

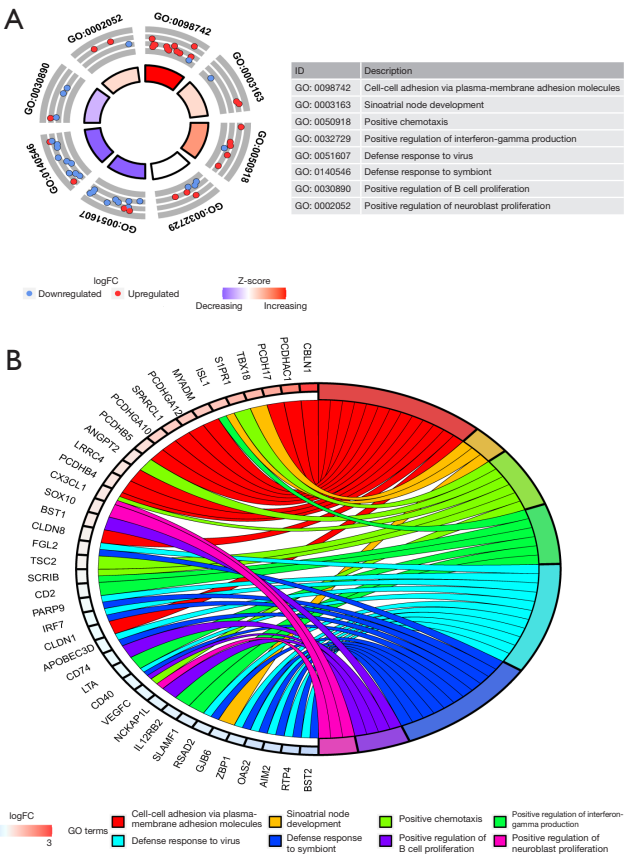

Figure 3 Functional enrichment analysis of methylation-driven genes in OSCC. (A) The outer circle represents the relative expression levels (logFC) of methylation-driven genes between cancerous and normal samples in each GO term. Red dots represent genes that are hypermethylated in cancerous tissues and blue dots represent genes that are hypomethylated in cancerous tissues. The inner circle illustrates the overall trend of up- and down-regulation of methylation-driven genes in each GO term, with blue to red indicating the down- and up-regulation trends in that sequence. (B) The circle shows the methylation-driven genes enriched in the top 8 GO terms. GO, Gene Ontology; FC, fold change; OSCC, oral squamous cell carcinoma.

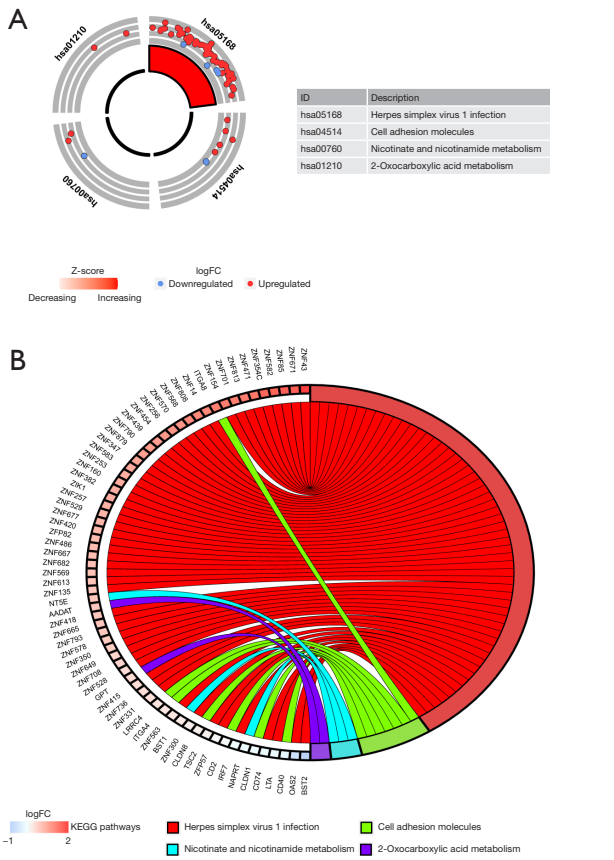

Figure 4 Pathway analysis of methylation-driven genes in OSCC. (A) The outer circle represents the relative expression levels (logFC) of methylation-driven genes between cancerous and normal samples in each KEGG term. Red dots represent genes that are hypermethylated in cancerous tissues and blue dots represent genes that are hypomethylated in cancerous tissues. The inner circle illustrates the overall trend of up- and down-regulation of methylation-driven genes in each KEGG term, with blue to red indicating the down- and up-regulation trends in that sequence. (B) The circle shows the methylation-driven genes enriched in the statistically significant different KEGG terms. FC, fold change; KEGG, Kyoto Encyclopedia of Genes and Genomes; OSCC, oral squamous cell carcinoma.

the KEGG analysis revealed four statistically significant pathways (*Figure 4A*), with "hsa05168 Herpes simplex virus 1 infection" being the most significantly enriched pathway. Genes associated with this enriched pathway are displayed in the KEGGChord plot (*Figure 4B*).

Construction of prognostic risk assessment model in OSCC

Seven methylation-driven genes significantly associated with patient prognosis were identified using univariate Cox regression analysis (*Table 1*). These seven methylation-driven genes were then subjected to multivariate Cox regression analysis to produce a set of four genes used to construct a linear prognostic risk assessment model (*Table 2*). To better understand the differences in methylation levels of these four methylation-driven genes in the high-risk and low-risk groups, a heat map of these four methylation-driven genes was drawn using R (*Figure 5A*). The prognostic

Table 1 Univariate Cox regression analysis

Gene	HR	HR.95L	HR.95H	P value
BST2	0.255504	0.068855	0.948112	0.04
KRT15	0.073547	0.008566	0.631473	0.02
ZNF134	0.056086	0.008011	0.39268	0.004
NT5E	4.123583	1.107377	15.35515	0.03
GSTA7P	0.427264	0.187175	0.975316	0.04
NAPRT	5.812477	1.050916	32.14805	0.04
GOLPH3L	0.342956	0.128327	0.916559	0.03

HR, hazard ratio; HR.95L, lower bound of hazard ratio at 95% confidence level; HR.95H, upper bound of hazard ratio at 95% confidence level.

Table 2 Multivariate Cox regression analysis

Gene	Coef	HR	HR.95L	HR.95H	P value
BST2	–1.30561	0.271008	0.072773	1.009235	0.05
KRT15	–3.19275	0.041059	0.004929	0.342016	0.003
ZNF134	–2.98134	0.050725	0.006984	0.368397	0.003
NAPRT	2.156957	8.644795	1.498416	49.87433	0.02

Coef, coefficient; HR, hazard ratio; HR.95L, lower bound of hazard ratio at 95% confidence level; HR.95H, upper bound of hazard ratio at 95% confidence level.

risk index = 2.156957 × expression value of *NAPRT* + (–1.30561) × expression value of *BST2* + (–3.19275) × expression value of *KRT15* + (–2.98134) × expression value of *ZNF134*. Patients were divided into high-risk (n=147 samples) and low-risk (n=148 samples) groups according to the median prognostic risk index (*Figure 5B*). The results showed that significantly more patients died in the high-risk group than in the low-risk group (*Figure 5C*). Kaplan-Meier survival curve analysis of patients in the high-risk and low-risk cohorts revealed that the overall survival rate of patients in the high-risk group was lower than that of patients in the low-risk group, and the difference was statistically significant (P<0.05) (*Figure 6A*). The time-dependent ROC curve was plotted to assess the predictive efficacy of this prognostic risk assessment model, and the results showed that the area under the curve (AUC) of the 5-year overall survival rate of patients predicted by this prediction risk model was 0.703 (*Figure 6B*).

Independent prognostic risk assessment model analysis

The results of univariate and multivariate Cox regression analysis revealed that age, TNM clinical stage, and prognostic risk value were significant independent prognostic factors for clinical prognosis (P<0.001) (*Figure 7A*,*7B*). The hazard ratios for age, TNM clinical stage, and risk score were 1.035 [95% confidence interval (CI): 1.017–1.054], 1.617 (95% CI: 1.294–2.019), and 2.057 (95% CI: 1.549–2.730), respectively. The area under the ROC curve (AUC) for predicting 5-year survival using age, TNM stage, and risk values were 0.597, 0.654, and 0.703, respectively (*Figure 7C*).

Combined survival analysis of DNA methylation and gene expression

The combined Kaplan-Meier survival curve analysis of methylation and gene expression revealed a significant association between three methylation-driven genes and the overall survival rate of OSCC patients (P<0.01) (*Figure 8*). Patients with hypermethylation and low expression of *ZKSCAN7* demonstrated a lower overall survival rate compared to those with hypomethylation and high expression (*Figure 8A*). Conversely, patients with hypermethylation and low expression of *ZNF134* and *MFF*

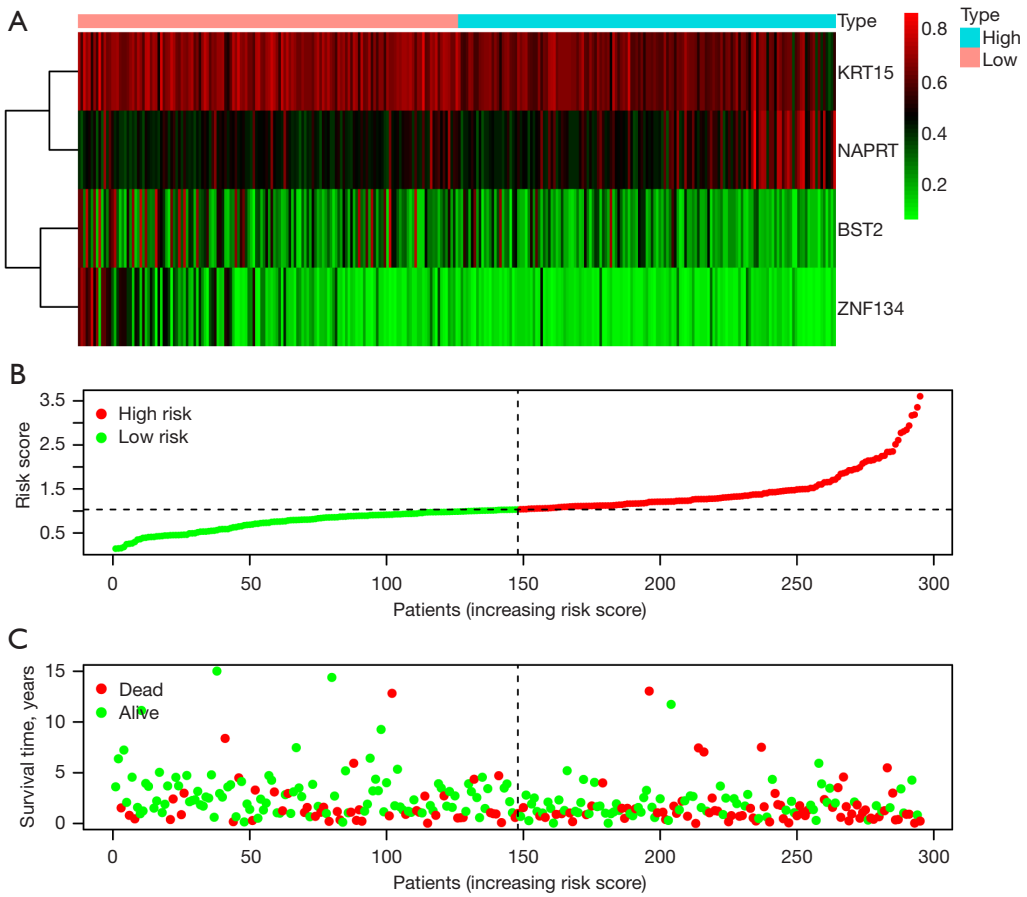

Figure 5 Heat map of these four methylation-driven genes (*NAPRT*, *BST2*, *KRT15* and *ZNF134*) and the prognostic evaluation of a linear prognostic risk assessment models constructed by these four methylation-driven genes. (A) The hierarchical clustering heat map of the four methylation-driven genes (*NAPRT*, *BST2*, *KRT15* and *ZNF134*) in OSCC. The color ranges from green to red in order to indicate low to high level of gene methylation. (B) Patient risk score distribution. The green dot represents low risk score, the red dot represents high risk score. (C) Relationship between patient survival status and risk score. The green dot represents survival, the red dot represents death. OSCC, oral squamous cell carcinoma.

exhibited a higher overall survival rate compared to those with hypomethylation and high expression (*Figure 8B*,*8C*).

Discussion

OSCC is characterized by a high degree of malignancy, invasiveness and susceptibility to distant metastases, resulting in a poor prognosis (4). Moreover, the insidious onset of OSCC makes early diagnosis difficult (5). Therefore, early diagnosis and effective prognostic evaluation of OSCC are crucial.

In the present study, we obtained 374 methylation-driven genes and established a prognostic risk assessment model that could effectively predict patients' prognostic survival by comprehensively analyzing OSCC methylation data and transcriptome data in the TCGA database through bioinformatics technology. Furthermore, three methylation-driven genes, *ZKSCAN7*, *MFF* and *ZNF134*, were found to be closely associated with patient survival by combined survival analysis of DNA methylation and gene expression. These findings hold potential implications for the early diagnosis, targeted therapy, and prognostic evaluation of OSCC patients.

DNA methylation plays a significant role in tumorigenesis and development, with many tumors exhibiting tumor suppressing genes (TSGs) promoter hypermethylation or oncogenes promoter hypomethylation (12). Therefore, more and more scholars are devoted to the research related

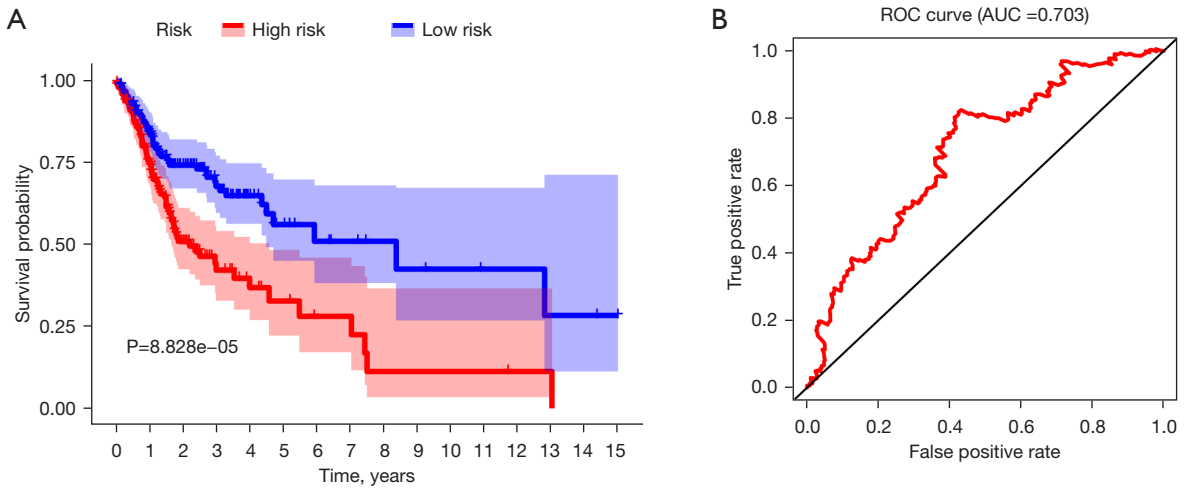

Figure 6 Kaplan-Meier survival curve and ROC curve for the linear prognostic risk assessment model constructed by four methylation-driven genes (*NAPRT*, *BST2*, *KRT15* and *ZNF134*). (A) Kaplan-Meier curve analysis of prognostic risk assessment model for overall survival of OSCC patients. (B) Time-dependent ROC curve of risk-prognosis models to predict 5-year survival of patients. ROC, receiver operating characteristic; AUC, area under the curve; OSCC, oral squamous cell carcinoma.

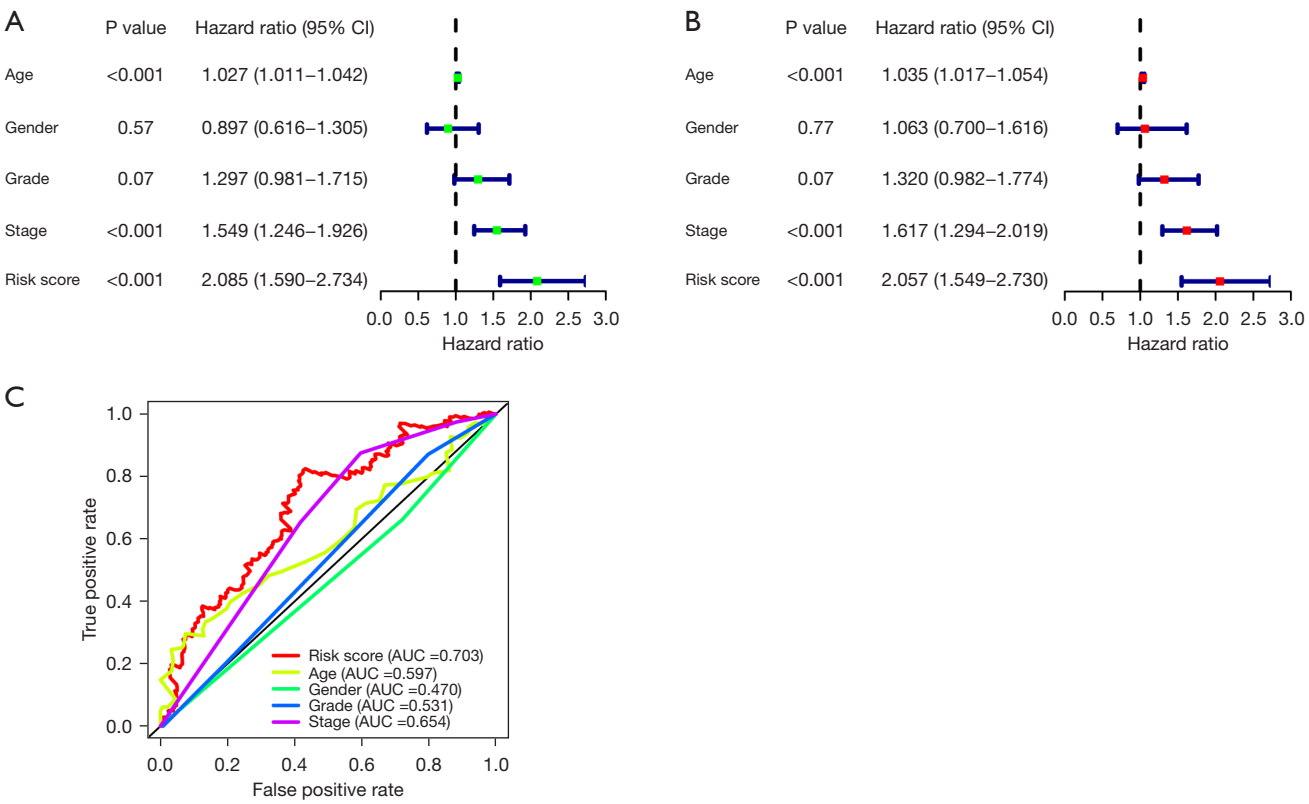

Figure 7 Independent prognostic analysis and multi-index ROC curves. (A) Forest plot of univariate Cox regression analysis. (B) Forest plot of multivariate Cox regression analysis. (C) Time-dependent ROC curve of multi-index (risk score, age, gender, clinical TNM stage and clinical grade) to predict 5-year survival of patients. CI, confidence interval; AUC, area under the curve; ROC, receiver operating characteristic; TNM, tumor, node, metastasis.

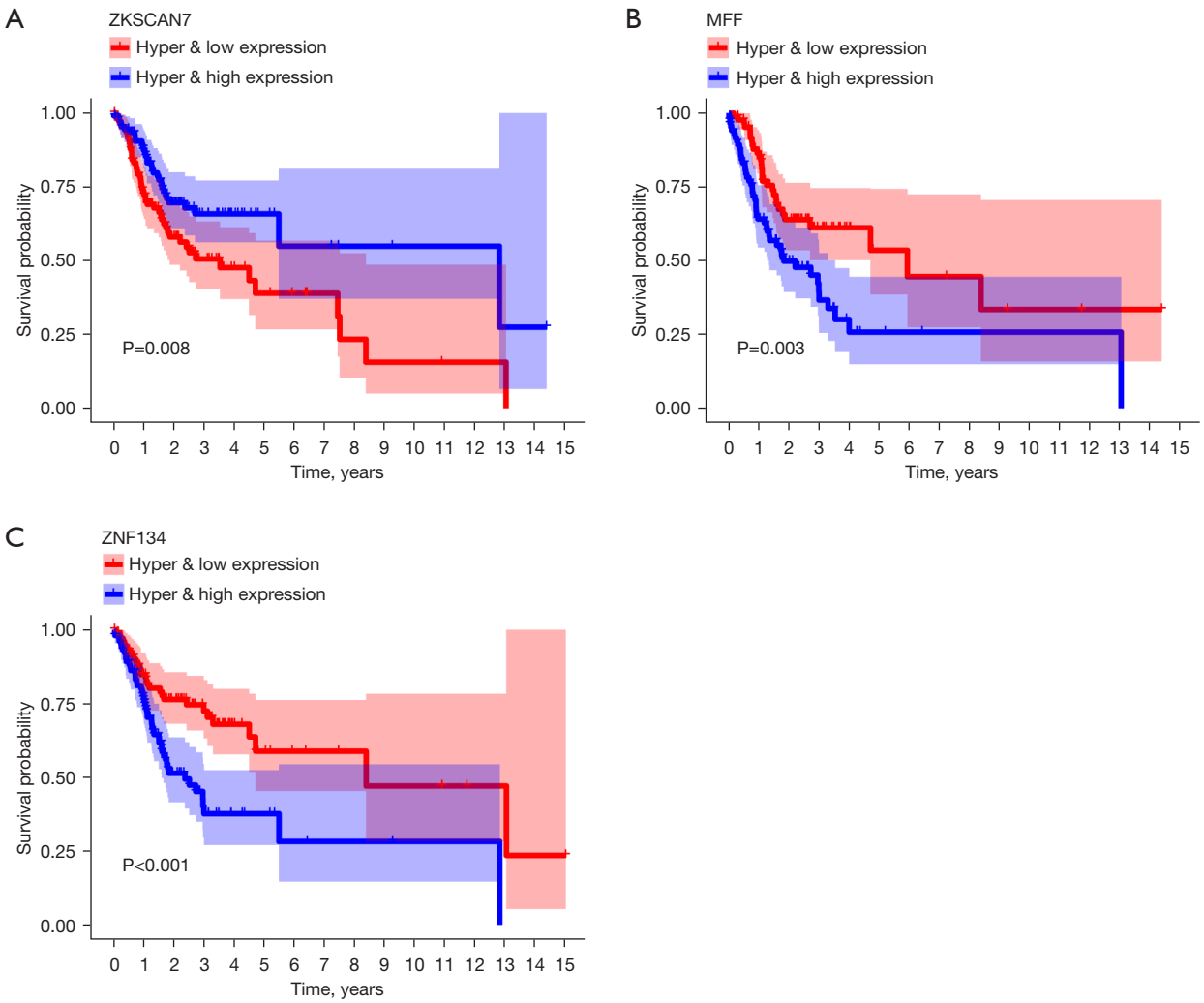

Figure 8 Kaplan-Meier survival analysis curves for combined methylation and gene expression. (A-C) The combination survival analysis of three genes (*ZKSCAN7*, *MFF*, *ZNF134*).

to DNA methylation and tumor. Meanwhile, with the emergence of public databases and the advancement of bioinformatics technology, the research on methylation-driven genes and tumors has also become a hot topic among scholars. For example, Li *et al.* identified through bioinformatics analysis that four methylation-driven long non-coding RNAs (lncRNAs) and eight methylation-driven mRNAs have potential as biomarkers for predicting the prognosis of lung adenocarcinoma (20). In addition, Lu *et al.* identified that three methylation-driven genes, *ABCD1*, *CCDC8*, and *FBXO17*, were significantly associated with the prognosis of esophageal squamous cell carcinoma based on the study of the TCGA database (21). However, no similar reports have been seen in OSCC. Moreover, in this study, the combined survival analysis of DNA methylation and gene expression of methylation-driven genes was also done, and three genes, *ZKSCAN7*, *MFF*, and *ZNF134*, were found to be closely associated with the prognostic survival of OSCC patients. These findings identified potential markers for early diagnosis, prognostic evaluation and targeted therapy in OSCC patients.

GO and KEGG pathway enrichment analyses of methylation-driven genes showed that cell-cell adhesion through plasma-membrane adhesion molecules were most significantly enriched in GO analysis, and interestingly, in KEGG pathway enrichment analysis, cell adhesion molecules also ranked as the second most significant term after herpes simplex virus 1 infection. The above results

suggest that differentially methylation-driven genes play a crucial role in regulating cell-cell adhesion in OSCC. Cell adhesion is a fundamental process in multicellular organisms and is essential for the development, maintenance and function of tissues (22). It has been reported that abnormal cell adhesion is closely associated with tumorigenesis and progression (23-25). Normal cells are suppressed by adhesion in a particular environment, i.e., they maintain an appropriate spacing between the same cells that are in the cellular matrix (26,27). In tumors, cells may lose this inhibition, leading to cell aggregation and abnormal proliferation (28,29). Tumor cells typically exhibit enhanced migratory and invasive capabilities, which are associated with their altered ability to adhere to surrounding cells and stroma (30). Loss of adhesion to normal tissues may make it easier for tumor cells to invade surrounding tissues and form metastases in other parts of the body (31). Therefore, our finding suggests that the occurrence and development of OSCC may be closely related to the process of cell adhesion and provides a specific direction for us to further investigate the mechanism of OSCC occurrence and development.

The combined survival analysis of DNA methylation and mRNA suggested that the overall survival of OSCC patients in the ZKSCAN7 hypermethylated and low expression group was lower than that in the ZKSCAN7 hypomethylated and high expression group. This suggests that ZKSCAN7 may be a TSG in OSCC. ZKSCAN7 encodes a protein belonging to the zinc finger family of proteins, which contains KRAB and SCAN (SRE-ZBP, CTfin51, AW-1, and number 18 complementary DNA) domains. The KRAB domain is commonly associated with transcriptional repression of genes, while the SCAN domain is associated with nucleic acid binding (32,33). No article has been published on *ZKSCAN*7 gene. This suggests that ZKSCAN7 may be a novel TSG in OSCC, which provides us with new ideas for future research. However, combined survival analysis of DNA methylation and mRNA showed that the overall survival of OSCC patients in the hypermethylated and low expression group of the *MFF* and *ZNF134* genes was significantly higher than that of their hypomethylated and high expression group.

MFF is a gene associated with the process of mitochondrial fission, which encodes a protein involved in the regulation of mitochondrial fission (34). Mitochondria play a key role in regulating cellular energy metabolism, cell death and oxidative stress (35). Abnormal mitochondrial morphology and function may lead to abnormal cell growth and metabolism, which may be associated with tumorigenesis and progression (36). Mitochondrial dynamics play a crucial role in tumorigenicity and malignancy of various types of cancers by promoting the tumor-initiating potential of cancer cells (37). A number of studies have reported anticancer effects through direct inter- or indirect modulation of MFF-induced mitochondrial fission process in tumors. Tang *et al.* reported in their study that overexpression of MFF was significantly upregulated in liver cancer-initiating cells (LCICs), promoting mitochondrial fission and enhancing stemness and tumor initiation in non-LCICs. Meanwhile, TBX19/PRMT1 complex-mediated up-regulation of MFF promotes mitochondrial fission and tumor-initiating ability in liver cancer cells, thus identifying PRMT1 as a viable therapeutic target for liver cancer (38). Seo *et al.* reported in their study that mitochondrial fission modulation was performed using a peptidomimetic agent that disrupts the MFF-VDAC1 (voltage-dependent anion channel-1) complex and showed anticancer activity in a variety of tumor models (39). In this study, we found that the overall survival of OSCC patients in the hypomethylated and highly expressed MFF group was significantly higher than that in the hypermethylated and low-expression group, suggesting that MFF may be a key gene in the development of OSCC, and may serve as a new therapeutic target and prognostic assessment marker for OSCC.

A previous study has reported that ZNF134 is hypermethylated and low-expressed in TET2-mutated diffuse large B-cell lymphoma (DLBCL) (40). Pan *et al.* also reported that ZNF134 is hypermethylated and low-expressed in HNSCC, and that the overall survival rate of patients with hypermethylated ZNF134 was higher than that of the hypomethylated group (41). Interestingly, our findings are consistent with those reported by Pan *et al.* In this paper, OSCC patients with ZNF134 hypermethylation and low expression had a better prognosis. Due to the hereditary and reversible character of epigenetics, the three genes *ZKSCAN*7, *MFF* and *ZNF134* could be potential therapeutic targets for OSCC.

In conclusion, in our study, a total of 374 methylated-driven genes were identified in association with OSCC. Additionally, a prognostic risk assessment model was constructed for OSCC and three genes (*ZKSCAN*7, *MFF* and *ZNF134*) significantly associated with patient prognosis were discovered. Nevertheless, it is important to acknowledge the limitations of this study. The findings were solely derived from bioinformatics analysis, without validation of the prognostic risk assessment model in a clinical setting. Additionally, three methylation-driven

genes associated with OSCC prognosis were based on bioinformatics analysis without experimental validation. In the follow-up study, we will try to verify the clinical significance of this prognostic risk assessment model by clinical samples. Meanwhile, the detailed biological functions and molecular mechanisms of the three genes (*ZKSCAN7*, *MFF* and *ZNF134*) in OSCC will be explored by *in vitro* and *in vivo* experiments.

Conclusions

Our findings established a crucial bioinformatics foundation and pertinent theoretical framework to facilitate subsequent early diagnosis, targeted therapy, and prognostic assessment in OSCC patients.

Acknowledgments

Funding: This work was supported by the grants from National Natural Science Foundation of China (No. 82260194 to J.C., B.L., Z.N. and J.Q.) and the Central Government Guides the Local Science and Technology Development Fund (No. 20221ZDG020068 to J.C., B.L., Z.N. and J.Q.).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2303/rc

Peer Review File: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2303/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2303/coif). J.C., B.L., Z.N. and J.Q. report that this work is supported by the grants from National Natural Science Foundation of China (No. 82260194) and the Central Government Guides the Local Science and Technology Development Fund (No. 20221ZDG020068). The other author has no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Mody MD, Rocco JW, Yom SS, et al. Head and neck cancer. Lancet 2021;398:2289-99.
2. Bhat GR, Hyole RG, Li J. Head and neck cancer: Current challenges and future perspectives. Adv Cancer Res 2021;152:67-102.
3. Tan Y, Wang Z, Xu M, et al. Oral squamous cell carcinomas: state of the field and emerging directions. Int J Oral Sci 2023;15:44.
4. Kumari P, Debta P, Dixit A. Oral Potentially Malignant Disorders: Etiology, Pathogenesis, and Transformation Into Oral Cancer. Front Pharmacol 2022;13:825266.
5. Tilakaratne WM, Ekanayaka RP, Warnakulasuriya S. Oral submucous fibrosis: a historical perspective and a review on etiology and pathogenesis. Oral Surg Oral Med Oral Pathol Oral Radiol 2016;122:178-91.
6. Radaic A, Kamarajan P, Cho A, et al. Biological biomarkers of oral cancer. Periodontol 2000 2023. [Epub ahead of print]. doi: 10.1111/prd.12542.
7. Madhura MG, Rao RS, Patil S, et al. Advanced diagnostic aids for oral cancer. Dis Mon 2020;66:101034.
8. Gyanchandani A, Shukla S, Vagha S, et al. Diagnostic Utility of Cytokeratin 17 Expression in Oral Squamous Cell Carcinoma: A Review. Cureus 2022;14:e27041.
9. Li Y. Modern epigenetics methods in biological research. Methods 2021;187:104-13.
10. Gluckman PD. Epigenetics and metabolism in 2011: Epigenetics, the life-course and metabolic disease. Nat Rev Endocrinol 2011;8:74-6.
11. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. Lancet 2018;392:777-86.
12. Nishiyama A, Nakanishi M. Navigating the DNA methylation landscape of cancer. Trends Genet 2021;37:1012-27.

13. Papanicolau-Sengos A, Aldape K. DNA Methylation Profiling: An Emerging Paradigm for Cancer Diagnosis. Annu Rev Pathol 2022;17:295-321.
14. Müller D, Győrffy B. DNA methylation-based diagnostic, prognostic, and predictive biomarkers in colorectal cancer. Biochim Biophys Acta Rev Cancer 2022;1877:188722.
15. Luo H, Wei W, Ye Z, et al. Liquid Biopsy of Methylation Biomarkers in Cell-Free DNA. Trends Mol Med 2021;27:482-500.
16. Jung G, Hernández-Illán E, Moreira L, et al. Epigenetics of colorectal cancer: biomarker and therapeutic potential. Nat Rev Gastroenterol Hepatol 2020;17:111-30.
17. Terp SK, Stoico MP, Dybkær K, et al. Early diagnosis of ovarian cancer based on methylation profiles in peripheral blood cell-free DNA: a systematic review. Clin Epigenetics 2023;15:24.
18. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015;19:A68-77.
19. Gevaert O. MethylMix: an R package for identifying DNA methylation-driven genes. Bioinformatics 2015;31:1839-41.
20. Li R, Yang YE, Yin YH, et al. Methylation and transcriptome analysis reveal lung adenocarcinoma-specific diagnostic biomarkers. J Transl Med 2019;17:324.
21. Lu T, Chen D, Wang Y, et al. Identification of DNA methylation-driven genes in esophageal squamous cell carcinoma: a study based on The Cancer Genome Atlas. Cancer Cell Int 2019;19:52.
22. Fields RD, Itoh K. Neural cell adhesion molecules in activity-dependent development and synaptic plasticity. Trends Neurosci 1996;19:473-80.
23. Ang HL, Mohan CD, Shanmugam MK, et al. Mechanism of epithelial-mesenchymal transition in cancer and its regulation by natural compounds. Med Res Rev 2023;43:1141-200.
24. Janiszewska M, Primi MC, Izard T. Cell adhesion in cancer: Beyond the migration of single cells. J Biol Chem 2020;295:2495-505.
25. Cavallaro U, Christofori G. Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. Nat Rev Cancer 2004;4:118-32.
26. Satyanarayanajois SD. Cell adhesion molecules: structure, function, drug design, and biomaterials. Curr Pharm Des 2008;14:2126-7.
27. Laffón A, González-Amaro R. Cell adhesion molecules: an overview. Br J Rheumatol 1995;34:1101-2.
28. Makgoba MW, Bernard A, Sanders ME. Cell adhesion/signalling: biology and clinical applications. Eur J Clin Invest 1992;22:443-53.
29. Trzpis M, McLaughlin PM, de Leij LM, et al. Epithelial cell adhesion molecule: more than a carcinoma marker and adhesion molecule. Am J Pathol 2007;171:386-95.
30. Kessenbrock K, Plaks V, Werb Z. Matrix metalloproteinases: regulators of the tumor microenvironment. Cell 2010;141:52-67.
31. Harjunpää H, Llort Asens M, Guenther C, et al. Cell Adhesion Molecules and Their Roles and Regulation in the Immune and Tumor Microenvironment. Front Immunol 2019;10:1078.
32. Collins T, Stone JR, Williams AJ. All in the family: the BTB/POZ, KRAB, and SCAN domains. Mol Cell Biol 2001;21:3609-15.
33. Ecco G, Imbeault M, Trono D. KRAB zinc finger proteins. Development 2017;144:2719-29.
34. Christen M, Gutierrez-Quintana R, Vandenberghe H, et al. Mitochondrial fission factor (MFF) frameshift variant in Bullmastiffs with mitochondrial fission encephalopathy. Anim Genet 2022;53:814-20.
35. Murphy E, Ardehali H, Balaban RS, et al. Mitochondrial Function, Biology, and Role in Disease: A Scientific Statement From the American Heart Association. Circ Res 2016;118:1960-91.
36. Zong WX, Rabinowitz JD, White E. Mitochondria and Cancer. Mol Cell 2016;61:667-76.
37. Tan YQ, Zhang X, Zhang S, et al. Mitochondria: The metabolic switch of cellular oncogenic transformation. Biochim Biophys Acta Rev Cancer 2021;1876:188534.
38. Tang M, Yang M, Wu G, et al. Epigenetic Induction of Mitochondrial Fission Is Required for Maintenance of Liver Cancer-Initiating Cells. Cancer Res 2021;81:3835-48.
39. Seo JH, Chae YC, Kossenkov AV, et al. MFF Regulation of Mitochondrial Cell Death Is a Therapeutic Target in Cancer. Cancer Res 2019;79:6215-26.
40. Liu P, Jiang W, Zhao J, et al. Integrated analysis of genome wide gene expression and DNA methylation microarray of diffuse large B cell lymphoma with TET mutations. Mol Med Rep 2017;16:3777-82.
41. Pan Y, Song Y, Cheng L, et al. Analysis of methylation-driven genes for predicting the prognosis of patients with head and neck squamous cell carcinoma. J Cell Biochem 2019;120:19482-95.

Cite this article as: Chen J, Dong Z, Li B, Nie Z, Qiu J. Analysis of methylation-driven genes for predicting the prognosis of patients with oral squamous cell carcinoma. Transl Cancer Res 2024;13(6):2892-2904. doi: 10.21037/tcr-23-2303

