The Prognostic Significance of MAFK and its Methylation in Cervical Cancer

Mengjun Zhang
Haerbin medical university cancer hospital

Hao Li
harbin medical university cancer hospital

Yuan Liu
Haerbin Medical University cancer hospital

Siyu Hou
Capital Medical University Affiliated Beijing Shijitan Hospital

Ping Cui
Haerbin medical university cancer hospital

Zhenxing Sun
Haerbin medical university cancer hospital

Yin Yue
Haerbin medical university cancer hospital

Xiuwei Chen (zmj230530@163.com)
Haerbin medical university cancer hospital

Research Article

Keywords: MAFK, biomarker, aberrant methylation, cervical cancer, prognostic evaluation

DOI: https://doi.org/10.21203/rs.3.rs-582069/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

**Background:** The purpose of this study was to determine the value of MAFK as a biomarker of cervical cancer prognosis and to explore its methylation and possible cellular signaling pathways.

**Methods:** We analyzed the cervical cancer data of The Cancer Genome Atlas (TCGA) through bioinformatics, including MAFK expression, methylation, prognosis and genome enrichment analysis.

**Results:** MAFK expression was higher in cervical cancer tissues and was negatively correlated with the methylation levels of five CpG sites. MAFK is an independent prognostic factor of cervical cancer and is involved in the Nod-like receptor signaling pathway. CMap analysis screened four drug candidates for cervical cancer treatment.

**Conclusions:** We confirmed that MAFK is a novel prognostic biomarker for cervical cancer and aberrant methylation may also affect MAFK expression and carcinogenesis. This study provides a new molecular target for the prognostic evaluation and treatment of cervical cancer.

Background

Cervical cancer is the fourth most common cancer after breast, colorectal and lung cancer, with an incidence of 106,000 cervical cancers in China in 2018, accounting for 48,000 deaths. Currently, despite the provision of standardized initial treatment, including surgery, radiotherapy and chemotherapy, there are still multiple limiting factors leading to poor prognosis. The main reason for the above situation is that the incidence of cervical cancer is a very complex process. So far, many scholars agree that persistent Human Papillomavirus (HPV) infection is the main cause of cervical cancer. However, persistent infection with oncogenic HPV is a necessary precursor and driver of cervical carcinogenesis but is not a sufficient pathogenic condition. It is worth noting that in the complex tumorigenesis process, genetics (abnormal gene expression and transcription) and epigenetics (DNA methylation and acetylation) are also important. Therefore, the analysis and exploration of genetic and epigenetic specific biomarkers are urgent tasks for prognosis assessment, recurrence judgment and developing rational treatment strategies for cervical cancer.

The musculoaponeurotic fibrosarcoma family proteins (MAF) are important transcription factors in the nucleus, including small molecular MAF proteins and large molecular MAF proteins, which can regulate the expression of various proteins in cells, cell differentiation and apoptosis. As an important member of the small MAF protein (sMAFs), the oncogene homolog of aponeurotic fibrosarcoma K (MAFK) is a regulator of mammalian gene expression. It does not have any transcription activation domain, so it needs to perform its biological function through homodimerization or heterodimerization with specific bZIP transcription factors. MAFK has been widely reported to be associated with several diseases, incorporates diabetes, neurological disorders, thrombocytopenia and carcinogenesis. In particular, several studies have shown that MAFK is closely related to the occurrence and poor prognosis of pancreatic cancer, breast cancer and osteosarcoma. To sum up, although these studies suggest
that MAFK may promote tumorigenesis, no studies to date have revealed its association with poor prognosis in cervical cancer and the underlying genetic and epigenetic mechanisms.

Recent evidence suggests that epigenetic changes in specific genes may mediate carcinogenesis or predict prognosis(11). DNA methylation is a hot research of epigenetic changes in cervical cancer, and either hypomethylation or hypermethylation is a key event in carcinogenesis. Multiple studies have consistently revealed that hypermethylated CpG islands can repress the transcription of multiple tumor suppressor genes, whereas hypomethylated CpG islands can activate the transcription of oncogenes(12, 13). At present, a series of methylation phenomena in the promoter sequences of genes associated with the occurrence and development of cervical cancer are constantly being discovered. Meanwhile, it is hopeful that, given that DNA methylation is a reversible process, it is a promising therapeutic target.

The emergence and rapid development of a variety of bioinformatics analysis techniques continue to reveal the genetic and epigenetic changes in gene expression. This knowledge is rapidly being transformed into diagnosis and treatment strategies for malignant tumors(14). To gain a comprehensive understanding of the association of MAFK with cervical cancer, various bioinformatics analyses were used to explore its clinical prognosis and molecular mechanisms. Abnormally high expression of MAFK may be closely related to a variety of clinicopathological features, such as clinical stage, tumor tissue grade, TNM stage (tumor invasion, lymph node metastasis and distant metastasis) and poor prognosis. Above may be caused by abnormal methylation in cervical cancer. Therefore, MAFK could be used as a novel oncogene for cervical cancer diagnosis, screening as well as prognostic assessment.

Materials And Methods

1. Data collection

RNA-seq data and Illumina Human Methylation450K array data of cervical cancer patients were obtained from the TCGA database created by the University of North Carolina TCGA genome(https://TCGA-data.nci.nih.gov/tcga/) which included some clinical data such as age, TNM stage, survival time and status in 191 cervical tumors tissues. TCGA database is currently the largest cancer gene information database, including not only comprehensive clinical data but also multi omics data such as gene expression data, miRNA expression data, DNA methylation. The Human Protein Atlas (http://www.proteinatlas.org) was an immunohistochemical database for validation of genes with prognostic value, which explored the differential expression of MAFK between cervical cancer and control groups from the protein level. The detailed clinical information of the corresponding patients is shown in Table S1.

2. Gene set enrichment analysis (GSEA)

GSEA is an ideal bioinformatics analysis tools developed by the research team of MIT and Harvard University's Broad Institute and is used to analyze the cell signaling pathway. The RNA-sequencing data which obtained from the CGGA database were batch corrected and normalized using SVA and LIMMA, and then divided into "H group" (high expression group) or "L group" (low expression group) according to
MAFK expression levels. GSEA (V4.0.3) software were used for enrichment analysis. The number of permutations was set to 1000 times, and the gene set database was set to the Kyoto Encyclopedia of Genes and Genomes (KEGG) cell signaling pathway (P < 0.05 was considered significantly enriched).

3. Co-expression analysis and CMap analysis

The Pearson method is used to perform co-expression analysis on a large number of gene expression data and to construct the correlation between genes and gene functions. According to the correlation coefficient and P value, 10 genes with positive and negative correlation with MAFK were obtained. Connectivity Map (CMap, https://portals.broadinstitute.org/cmap/) is a database constructed by Professor lamb et al for the correlation between small molecule drugs, genes and diseases. It is used to make small molecule drug predictions. Subsequently, the two-dimensional and three-dimensional structure diagrams and chemical formulas of drugs were obtained in the PubChem database.

4. Statistical analysis

R software (version v.3.6.1) was used for statistical data analysis. The gene expression data obtained from TCGA database were divided into high expression group and low expression group according to the expression level of MAFK. Then, Kaplan-Meier method (KM), proportional hazards model (Cox regression) and ROC curve (a time-dependent receiver operating characteristic) were used for survival analysis, univariate and multivariate analysis and the value of MAFK in evaluating the clinical prognosis of cervical cancer. In addition, Wilcox or Kruskal was used to detect the relationship between MAFK expression and clinical features. (P < 0.05 was considered statistically significant.)

Results

1. Abnormally high expression of MAFK in protein level of cervical cancer

Genes are the basic genetic units controlling biological traits, while proteins are the main executors of gene function, and all phenotypes of organisms are primarily manifestations of protein activity. Therefore, we obtained the results of MAFK immunohistochemical staining of cervical cancer tissues from the human protein atlas database that provides information on the tissue and cellular distribution of all 24000 human proteins. It can be seen from Fig. 1 that MAFK showed high expression in cervical cancer tissues compared with normal cervical tissues. In summary, the expression of MAFK in cervical cancer is increased through the protein level.

2. Correlation between the expression level of MAFK and clinical features

191 patients with complete clinical information from the TCGA database were included in this study. The expression of MAFK was positively correlated with clinical stage and histological grade (Figure. 2A, B). In
the methylation data, the methylation level of MAFK was decreased in T stage presented cervical cancer tissues and Tx grade cervical cancer tissues (Figure. 2D).

3. High expression of MAFK is an independent risk factor and has certain evaluation value for cervical cancer prognosis

Firstly, according to the expression level of MAFK in cervical cancer tissues, patients were divided into high expression group and low expression group. Then, Kaplan-Meier survival analysis was performed to explore the potential relationship between MAFK expression and overall survival. The results of survival analysis showed that high expression of MAFK was significantly associated with decreased survival in cervical cancer patients (Figure. 3A, P = 0.009). As shown in Figure. 3A, patients with high MAFK expression had a worse prognosis than those with low MAFK expression. Secondly, ROC curves were drawn to evaluate the prognostic value of MAFK in cervical cancer. ROC analysis showed that the area under the ROC curve (AUC) of three-year and five-year OS were 0.710 and 0.726 (Figure. 3B), indicating that abnormally high expression of MAFK had a certain clinical prognostic value. Thirdly, to further explore the factors associated with poor prognosis in cervical cancer, univariate and multivariate Cox regression analyses were employed. From the univariate results shown in Fig. 2C, we found that high MAFK expression was significantly associated with poor survival (P = 0.009, hazard ratio [HR] = 1.014 (95% CI [1.003–1.024]). In addition, clinical stage (P = 0.004, hazard ratio [HR] = 1.511 (95% CI [1.139–2.003]), pathological N (P < 0.001, hazard ratio [HR] = 1.877 (95% CI [1.315–2.681]), and pathological T (P < 0.001, hazard ratio [HR] = 1.672 (95% CI [1.286–2.172]) were also risk factors for the prognosis of cervical cancer patients. Finally, a multivariate analysis was performed in order to exclude the influence of confounding factors (Fig. 2D). MAFK expression remained independently associated with prognosis (P = 0.004, HR of 1.018 (95% CI [1.006–1.031]). Taken together, we conclude that MAFK is an independent risk factor for poor prognosis in cervical cancer.

4. Determination of MAFK-related cellular signaling pathway by GSEA

To elucidate the molecular mechanism of MAFK involved in the pathological process of cervical cancer, we performed GSEA enrichment analysis. Three cell signaling pathways potentially related to MAFK were predicted. As shown in Fig. 4, MAFK may participate in multiple cancer-related cell signaling pathways, including Nod-Like receptor signaling pathway, Axon guidance signaling pathway and Epithelial cell signaling in Helicobacter pylori infection signaling pathway. These results suggest that MAFK may be involved in these cancer-related cell signaling pathways as an oncogene.

5. Co-expression analysis results and CMap analysis results related to MAFK.

After predicting the molecular mechanisms in which MAFK might be involved, the co-expression correlation between MAFK and other parameters was further explored using Pearson correlation analysis. Co-expression analysis obtained the ten most significant genes positively and negatively correlated with MAFK. As Fig. 5 showed that there were top five genes positively correlated (DNER, FZD4, DUSP4,
GALNT2, CPS1) and top five genes negatively correlated (SPOP, APOBEC3F, KHDRBS1, RGMB, CENPM). Subsequently, the acquired co-expression gene data information was screened through the CMap database for small molecule compounds that might have potential therapeutic effects on cervical cancer. Finally, we searched the PubChem platform for the chemical structure information of drugs, including Azacitidine, Exisulind, Tranylcypromine and Hyanthone (Figure. 6A-D).

6. The expression of MAFK was negatively correlated with the methylation level of MAFK CpG sites

In this study, to explore that specific DNA methylation sites affect the expression of MAFK ultimately leading to poor prognosis in cervical cancer, 191 samples were downloaded from the Cancer Genome Atlas Database (TCGA) and further analyzed. Thirteen MAFK CpG sites in MAFK gene were selected from 191 samples of TCGA cohort, and the methylation level of each site was analyzed. Five of the 13 MAFK CpG sites were selected and analyzed for their correlation with the expression of MAFK (Fig. 7A). The results showed that the methylation levels of cg00939931 (R = −0.24, P = 0.00094), cg03207121 (R = −0.26, P = 0.00024), cg17469039 (R = −0.14, P = 0.046), cg23871276 (R = −0.16, P = 0.027) and cg24647108 (R = −0.19, P = 0.0081) were negatively correlated with the expression of MAFK (Fig. 7B-F).

Discussion

Cervical cancer is the fourth most common female malignant tumor in the world, accounting for 12% of all female cancers and it is worth noting that it is the main cause of female cancer deaths(2). Thus, it is urgent to study the potential molecular mechanism of malignant biological behavior of cervical cancer and find more reliable and promising biomarkers for the diagnosis and prognosis evaluation of cervical cancer.

Based on the status of cervical cancer mentioned above, the abnormally high expression of MAFK in cervical cancer was revealed at the protein level through the HPA database. Subsequently to further explore the effect of MAFK on the prognosis of cervical cancer, a comprehensive bioinformatics analysis was performed in this study. Firstly, through the TCGA database we found that MAFK high expression was significantly associated with clinical stage, histological grade, TNM stage and other related clinical features. Interestingly, several studies are consistent with our findings. For example, Yukari Okita et al confirmed that MAFK, as an oncogene, was closely related to the clinicopathological features of breast cancer and osteosarcoma and could predict poor prognosis(9, 10). Secondly, we found that genes positively correlated with MAFK acted as oncogenes in multiple tumors by co-expression analysis. DNER is an oncogene that leads to the malignant progression of breast cancer by promoting EMT and inhibiting the Wnt/β-Catenin pathway(15). Meanwhile, there are studies confirmed that FZD4 promotes cervical cancer proliferation and invasion in vitro by being regulated by upstream mir-505(16). Therefore, we speculated that MAFK may also play an important role in the malignant progression of cervical cancer as an oncogene. Finally, the survival analysis results indicated that the overexpression of MAFK decreased the overall survival of cervical cancer and had clinical prognostic diagnostic value. In addition, we excluded the influence of random factors by univariate and multivariate analysis, from which we reached
a scientific conclusion that high MAFK expression could serve as an independent risk factor, implying a poor prognosis. Therefore, based on the above studies, it is confirmed that MAFK is a potential oncogene in cervical cancer and can lead to poor prognosis, but its pathological mechanism needs to be deeply explored.

We applied GSEA enrichment method to further understand the pathological mechanism by which MAFK contributes to poor prognosis in cervical cancer. As shown in Fig. 4, MAFK is associated with some signaling pathways involved in cancer initiation and progression. Previous studies have shown that activation of the Nod-like receptor signaling pathway can promote the tumorigenic capacity of tumor cells in breast cancer(17). The role of Axon guidance signaling pathway in pancreatic cancer is indispensable(18). The GSEA enrichment analysis method was conducted by the enrichment profile of differentially expressed genes in some previously validated pathways, which indirectly revealed the mechanism of MAFK in cervical cancer. Many researchers have utilized this approach to identify promising biomarkers for cervical cancer. Similar research ideas and methods have also been used in the studies of Xuan et al(19) and Liu et al(20), which indirectly proves the scientific rationality of our research methods.

Abnormal changes and modifications of epigenetics are as important as genetic abnormalities, and they are also an important part of cancerous and subsequent tumor progression. Therefore, the malignant biological behavior of MAFK gene overexpression and carcinogenesis may also involve epigenetics. Our results showed that the methylation sites were negatively correlated with the expression of MAFK. A great quantity of studies were highly consistent with our study, the occurrence and development of cervical cancer is a multi-factor, multi-stage and multi-gene change process in which aberrant methylation leading to activation of oncogenes is one of its important mechanisms. Taken together, a better understanding of the molecular mechanisms of MAFK will provide new insights into improving the prognosis of cervical cancer, and further exploration of new therapeutic strategies related to MAFK is our ultimate goal.

To achieve the ultimate goal of our cancer therapy and apply our findings to the clinical treatment of cervical cancer, this study explored four small molecule compounds with potential therapeutic effects on cervical cancer by CMap analysis (as shown in Fig. 6). Encouragingly, Azacitidine (AZA) is a DNA methyltransferase inhibitor, and low doses of Azacitidine exert antitumor effects associated with demethylation. This discovery pioneered epigenetic therapy and has been approved by the FDA for the treatment of multiple hematogenous tumors(21). Interestingly, unlike genetic alterations, DNA methylation is reversible, so exploring drugs that alter DNA methylation to treat cervical cancer is a promising new direction. In addition, Exisulind is a selective apoptotic antitumor drug that induces apoptosis in a wide range of precancerous and cancerous tissues. For example, the combination of Exisulind with Docetaxel produced additive or synthetic growth inhibition in the NSCLC cell line(22, 23). It is worth noting that previous studies have evaluated the reliability of the CMap tool for drug prediction(24, 25). In addition, because of the increased cost and time for new drug discovery, modifying existing drug molecules, exploring new indications for old drugs and expanding clinical applications will
accelerate the drug discovery process. Fortunately, our research has contributed greatly to the success of this strategy.

Based on the clinical data from TCGA database, we reveal MAFK as a new potential biomarker for cervical cancer screening in terms of transcriptional histology and DNA methylation, and its abnormally high expression is closely related to malignant progression and poor prognosis of cervical cancer. For the first time, we revealed the potential relationship between MAFK and aberrant methylation, confirming that epigenetic abnormalities have great potential in the diagnosis and treatment of cancer. But our study still had some limitations. There is a lack of M stage data when investigating the correlation between MAFK and TNM stage of cervical cancer, but the lack of some specific clinical data is an inevitable disadvantage of public databases. We have conducted relevant research from multiple omics (genomics and proteomics) and multiple perspectives (genetics and epigenetics) to ensure the comprehensiveness and reliability of our data.

Conclusion

Our study revealed for the first time the relationship between the overexpression of MAFK and the poor prognosis of cervical cancer as well as the possible molecular mechanisms of transcriptional histology and DNA methylation levels. In conclusion, MAFK may become a promising and valuable biomarker for improving the prognosis of cervical cancer and provide new insights into improving the prognosis of cervical cancer while encouraging us to decode the role of the human epigenome in cervical cancer.

Abbreviations

TCGA
The Cancer Genome Atlas; MAF:The musculoaponeurotic fibrosarcoma family proteins; MAFK:homolog of aponeurotic fibrosarcoma K; TNM stage:Tumor invasion, Lymph node metastasis and Distant metastasis; GSEA:Gene set enrichment analysis; KM:Kaplan-Meier method; AZA:Azacitidine;

Declarations

Ethics approval and Consent to participate
Not applicable.

Consent for Publication
Not applicable.

Data Availability Statement
The data can be obtained through the email under reasonable request: zmj230530@163.com.

Conflict of Interest

The author reports no conflicts of interest in this work.

Funding

This work was supported by grants from the National Natural Science Foundation of China (Grant Numbers: 81772274).

Author Contribution section

CXW designed the study. ZMJ and LY reviewed the raw data and confirm the authenticity of all raw data. HSY performed the analysis. LH and CP and SZX collected data. ZMJ drafted the manuscript. CXW and YY revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Thanks for the support of Harbin Medical University Cancer Hospital.

References


Figures
In the HPA database, the immunohistochemical expression of MAFK is different in cervical cancer and normal cervical tissues. It is low staining in normal tissues and strong staining in cervical cancer. (A) Cervix, Uterine normal tissues (ID:1773). (B) Cervical cancer tissues (ID:2480). (C) Cervical cancer tissues (ID:1733). (D) Cervical cancer tissues (ID:3382).
Figure 2

Correlation analysis of different clinical features with MAFK expression and MAFK methylation levels. (Clinical stage, Neoplasm histologic grade, Pathologic T). (A) Clinical stage. (B) Neoplasm histologic grade. (C) MAFK expression by Pathologic T. (D) MAFK methylation level by Pathologic T.
Figure 3

The correlation between MAFK and the poor prognosis of patients with cervical cancer. (A) The Kaplan-Meier survival curve reveals the high expression of MAFK leads to a poor prognosis in cervical cancer, the overall glioma survival curve based on TCGA. (B) ROC curve showed that MAFK had a good prognostic value in cervical cancer. (C-D) Analysis of univariate and multivariate factors affecting the prognosis of patients with cervical cancer. (C) Univariate regression analysis. (D) Multivariate analysis.
Figure 4

GSEA enrichment analysis results of MAFK. (A) Nod-Like receptor signaling pathway. (B) Axon guidance signaling pathway. (C) Epithelial cell signaling in Helicobacter pylori infection signaling pathway.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Correlation coefficient</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPOP</td>
<td>-0.288</td>
<td>3.03E-06</td>
</tr>
<tr>
<td>APOBEC3F</td>
<td>-0.287</td>
<td>3.09E-06</td>
</tr>
<tr>
<td>KHDRBS1</td>
<td>-0.275</td>
<td>8.36E-06</td>
</tr>
<tr>
<td>RGMB</td>
<td>-0.269</td>
<td>1.33E-05</td>
</tr>
<tr>
<td>CENPM</td>
<td>-0.268</td>
<td>1.48E-05</td>
</tr>
<tr>
<td>DNER</td>
<td>0.863</td>
<td>7.17E-77</td>
</tr>
<tr>
<td>FZD4</td>
<td>0.817</td>
<td>2.65E-62</td>
</tr>
<tr>
<td>DUSP4</td>
<td>0.798</td>
<td>1.87E-57</td>
</tr>
<tr>
<td>GALNT2</td>
<td>0.796</td>
<td>4.86E-57</td>
</tr>
<tr>
<td>CPS1</td>
<td>0.769</td>
<td>3.67E-51</td>
</tr>
</tbody>
</table>

Figure 5

Co-expression analysis of MAFK. (A) The ten most significant genes of positive and negative correlating with MAFK. (B) The correlation coefficients and P values of the ten most important positive and negative genes related to MAFK.
Figure 6

Screening of gene therapy drugs for MAFK in the CMap and Pubchem database (Drug name, chemical structure, 2D structure and 3D structure). (A) Azacitidine (B) Exisulind (C) Hyanthone (D) Tranylcypromine.
Figure 7

Correlation between expression of MAFK and methylation levels of MAFK. (A) Thirteen MAFK CpG sites in MAFK gene and the methylation level of each site. (B) cg00939931 (R= -0.24, P = 0.00094). (C) cg03207121 (R= -0.26, P = 0.00024). (C) cg17469039 (R= -0.14, P = 0.046). (E) cg23871276 (R= -0.16, P = 0.027). (F) cg24647108 (R= -0.19, P = 0.0081).
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.docx