Significance of PTEN Mutational Status Associated Gene Signature in the Progression and Prognosis of Endometrial Carcinoma

Qing Hu  
Shengjing Hospital of China Medical University

Jun Wang  
Shengjing Hospital of China Medical University

Lina Ge  
Shengjing Hospital of China Medical University

Ying Wu (✉️ wuy@sj-hospital.org)  
Shengjing Hospital of China Medical University  https://orcid.org/0000-0003-3801-4638

Primary research

**Keywords:** PTEN mutation, signature, endometrial carcinoma, prognosis, TMB

**Posted Date:** August 3rd, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-724760/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.  
[Read Full License](https://creativecommons.org/licenses/by/4.0/)
Abstract

Background

PTEN mutation had been reported to be involved in the development and prognosis of endometrial carcinoma (EC). However, a prognostic genes signature associated with PTEN mutational status has not been developed. In this study, we aim to conduct a PTEN mutation associated prognostic genes signature for EC.

Methods

We obtained the simple nucleotide variation and transcriptome profiling data from The Cancer Genome Atlas database as training data. Lasso Cox regression algorithm was used to establish PTEN mutation associated prognostic genes signature. The overall survival rate of the high-risk and low-risk groups was determined by Kaplan-Meier (K-M) method. The accuracy of risk score prediction was tested by ROC curve.

Results

K-M curves revealed that the EC patients with PTEN mutation have favorable survival outcome. Differential expression analysis between the EC patients with PTEN mutation and PTEN wild identified 224 differential expression genes (DEGs). Eighty-four DEGs with prognostic value was fitted into least absolute shrinkage and selection operator (LASSO)–Cox analysis and a seven PTEN mutation associated prognostic genes signature with robust prognostic ability was constructed, which was successfully validated in the other two datasets from cBioPortal database as well as 60 clinical specimens. Furthermore, the PTEN mutation associated prognostic genes signature had been proved to be an independent prognostic predictor for EC. Remarkably, the EC patients in high-risk group were characterized with higher stages and grades as well as lower tumor mutation burden of EC, with poor survival outcome. Collectively, the PTEN mutation associated prognostic genes signature was a favorable prognostic biomarker for EC.

Conclusion

In summary, we conducted and validated a prognostic predictor for EC associated with PTEN mutational status, which may be used as favorable prognostic biomarkers and therapeutic targets for EC.

Background

Endometrial carcinoma (EC) is a common malignancy of the female reproductive system, the incidence of which is increasing [1]. EC is a heterogeneous tumor, and the prognosis of patients is closely related to the tumor grade and stage. Early, accurate, and effective diagnosis is helpful to improve the prognosis of EC patients [2]. Surgery and postoperative radiotherapy are routine methods for treating this condition,
but there is still a lack of effective treatment for recurrent or progressive EC [3]. Therefore, there is an urgent need to further explore biological markers for the prognostic prediction of EC.

PTEN, a recognized tumor suppressor gene, is one of the most common mutated genes in human tumors, and can be detected in a variety of tumor tissues, including EC [4]. Kong et al.[5] discovered that the mutation rate of PTEN in EC was the highest compared with other tumors, and 37 %~61% of EC patients had PTEN gene mutation. A study [6] identified that PTEN deficient endometrial epithelial cells are more likely to convert to complex atypical hyperplasia in response to estrogen stimulation and thus develop into EC, so PTEN deficiency is generally considered to be an early event in EC development. Another study [7] has also confirmed that PTEN function loss is an early cancerous event, showing a higher frequency of PTEN mutations in precancerous or stage I tumors than in advanced or even metastatic EC, and that PTEN mutations are associated with better outcomes. Thus, PTEN mutations in endometrial hyperplasia may serve as an early warning of increased cancer risk [8].

In view of the important role of PTEN mutation in the progression and prognosis of EC. We fully revealed the mutational landscape of PTEN in EC and conducted a PTEN mutational status associated prognostic genes signature to predict the prognosis of EC based on The Cancer Genome Atlas (TCGA) database. External validation in the other two datasets from cBioPortal database as well as 60 clinical specimens was performed to prove the prognostic ability. Furthermore, we revealed the relationship between the PTEN mutational signature and stage, grade as well as tumor mutation burden (TMB) of EC. The PTEN mutation associated prognostic genes signature can be used as a favorable prognostic biomarker for EC.

Materials And Methods

Data acquisition

The simple nucleotide variation, transcriptome profiling datasets and corresponding clinical information of 529 EC patients downloaded from TCGA database (https://portal.gdc.cancer.gov/) was considered as training dataset. The transcriptome profiling datasets and corresponding clinical information of validation datasets ucec_tcga_pan_can_atlas_2018 and ucec_tcga_pub included 527 and 331 EC patients, respectively, were obtained from cBioPortal database (http://www.cbioportal.org/study/summary?id=ucec_tcga). The clinical information included age, BMI, stage, grade, overall survival time, and survival status.

Specimen collection

A total of 60 patients with primary endometrial cancer who were admitted to the Department of Gynecology and Obstetrics of Shengjing Hospital from January 2016 to January 2017 were selected. Patients in the primary endometrial cancer group were aged from 35 to 60 years, with an average age of 54.5 years. Patients with FIGO stage I were 30 cases, stage II were 15 cases, stage III~IV were 15 cases. Patients with G1 were 15 cases, G2 were 17 cases, G3 were 28 cases. All patients underwent staging surgery for endometrial cancer, removing pelvic lymph nodes and abdominal para-aortic lymph nodes.
The patient was confirmed as endometrial cancer by pathologists and had not received chemotherapy or radiotherapy before surgery. This study was approved by the ethics committee of the ShengJing Hospital of China Medical University, and informed consent was obtained from all patients.

Identification of Differentially Expressed Genes

The “limma” package was used to indentify the differentially expressed genes (DEGs) between the EC patients with PTEN mutation and PTEN wild. The screening criteria of DEGs was log\(|\text{FC}|>2\) and \(p<0.05\) and the results were visualized Heatmaps and volcanic maps.

Construction and Validation of a PTEN Mutational Status Associated Prognostic Signature

The “survival” package was applied to screen the DEGs with prognostic value by univariate Cox regression analysis based on the screening criteria \(p<0.05\). Next, the key DEGs with prognostic value were further selected by Least absolute shrinkage and selection operator (LASSO) regression and stepwise regression analyses. Multivariate Cox regression was used to calculate the regression coefficients of the key DEGs with prognostic value and conduct a PTEN mutational status associated prognostic signature. The risk score for each EC patient was calculated using the following formula: risk score = exp1 x \(\beta_1\) + exp2 x \(\beta_2\) + ... + expn x \(\beta_n\) (expn represents the expression value of each key DEGs with prognostic value, and \(\beta_n\) represents the regression coefficient) [9]. The EC patients were classified into the high- and low-risk groups based on the median risk score. Kaplan–Meier (K-M) method and log-rank test was used to evaluate the survival between high- and low-risk groups. Time-dependent and -independent receiver operating characteristic (ROC) curves was used to evaluate the prognostic ability of the PTEN mutational status associated prognostic signature. Furthermore, we validated the PTEN mutational status associated prognostic signature in ucec_tcga_pan_can_atlas_2018 and ucec_tcga_pub respectively. Finally, univariate and multivariate Cox regression analyses to determine whether PTEN associated signature has independent prognostic value in both testing and validation datasets.

Real-time qPCR

Real-time qPCR was used to detect the relative expression levels of PTEN associated genes in the 60 EC tissues. Total RNA of EC samples was extracted by Trizol Reagent (Invitrogen, USA). Retro-transcription cDNA was manipulated, and RT-qPCR was performed using SYBR Premix Exaq (Takara, Japan). GAPDH was selected as an internal reference to detect the relative expression level of PTEN associated genes in EC tissues based on the 2-\(\Delta\Delta\text{Ct}\) method. The sequences of primers used for RT-qPCR are presented in Supplementary Table 1. Next, we conducted a PTEN associated signature based on the relative expression level of PTEN associated genes to verify our bioinformatics analysis results. K-M curve was used to evaluate the survival between high- and low-risk groups. ROC curves were used to evaluate the prognostic ability of the PTEN associated signature.

Construction of a nomogram model based on the PTEN- Associated Signature
The "rms" and "survival" package was used to construct a nomogram model based on the expression level of the PTEN-associated signature. Consistency between actual and predicted survival rates was assessed by calibration curves. Decision Curve Analysis (DCA) was used to evaluate the accuracy of the prognostic prediction model [10, 11].

Statistical Analysis

The “maftools” package was used to reveal the PTEN mutational status in EC. Perl was used calculate the TMB of the patients with EC from TCGA database. Wilcoxon rank sum test was used for comparative analysis between two groups. DEGs related to overall survival (OS) were screened out by univariate Cox regression analysis. Lasso Cox regression algorithm was used to establish the risk prognosis model. The overall survival rate of the high-risk and low-risk groups was determined by K-M method. The accuracy of risk score prediction was tested by ROC curve. Univariate and multivariate Cox regression analyses were used to assess whether PTEN associated signature has independent prognostic value. The above statistics were analyzed by R (version 3.6.3) software. The test level P < 0.05 was considered statistically significant.

<table>
<thead>
<tr>
<th>id</th>
<th>coef</th>
<th>HR</th>
<th>HR.95L</th>
<th>HR.95H</th>
<th>pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDPD2</td>
<td>0.202281</td>
<td>1.224192</td>
<td>1.0066</td>
<td>1.48882</td>
<td>0.042781</td>
</tr>
<tr>
<td>GRB7</td>
<td>0.002907</td>
<td>1.002911</td>
<td>1.000964</td>
<td>1.004862</td>
<td>0.003377</td>
</tr>
<tr>
<td>KCNK9</td>
<td>0.287543</td>
<td>1.333148</td>
<td>1.113335</td>
<td>1.596361</td>
<td>0.001761</td>
</tr>
<tr>
<td>MUC3A</td>
<td>0.124442</td>
<td>1.132517</td>
<td>1.051765</td>
<td>1.219468</td>
<td>0.000976</td>
</tr>
<tr>
<td>MYT1</td>
<td>0.449528</td>
<td>1.567573</td>
<td>1.155313</td>
<td>2.126944</td>
<td>0.003887</td>
</tr>
<tr>
<td>RPS6KA6</td>
<td>0.265031</td>
<td>1.303471</td>
<td>0.941477</td>
<td>1.804652</td>
<td>0.110342</td>
</tr>
<tr>
<td>TSPYL5</td>
<td>0.035858</td>
<td>1.036509</td>
<td>1.013534</td>
<td>1.060005</td>
<td>0.001716</td>
</tr>
</tbody>
</table>

Table 1 Regression coefficients of the seven PTEN mutational status associated prognostic genes.

Results

PTEN mutational status in Endometrial Carcinoma

The mutational landscape of EC in TCGA database was visualized by Horizontal histogram using the “maftools” package and PTEN have the higher mutation frequency (64%; Figure 1). Kaplan-Meier curves revealed that the EC patients with PTEN mutation have longer survival time than the patients with PTEN wild (p<0.001; Figure 1B). In addition, percentchart showed that the PTEN mutation occur more frequently in the younger, lower stage and grade patients (p<0.001; Figure 1C-E).
Identification of Differentially Expressed Genes and Construction of the PTEN Mutational Status Associated Prognostic Signature

Considering the power prognostic value of PTEN mutational status, we aimed to conduct a PTEN mutational status associated prognostic signature to predict the prognosis of EC. Firstly, the “limma” package identified 224 DEGs (37 upregulated genes and 187 downregulated genes) between the EC patients with PTEN mutation or not based on the screening criteria of log\(|FC|>2\) and \(p<0.05\) (Figure 2A, B; Supplementary Table 2). Eighty-four DEGs with prognostic value were selected by univariate Cox regression analysis based on the screening criteria \(p<0.05\) (Supplementary Table 3). LASSO–Cox analysis was performed based on the 84 DEGs and a PTEN mutation associated prognostic signature with seven genes was constructed (Figure 2C, D). Riskscore of each EC patient was calculated using the following formula: risk score = 0.2023 x expGDPD2 + 0.0029 x expGRB7 + 0.2875 x expKCNK9 + 0.1244 x expMUC3A + 0.4495 x expMYT1 + 0.2650 x expRPS6KA6 + 0.0359 x expTSPYL5 (Table 1). The difference analysis showed that all of the seven genes were higher expressed in PTEN wild group than in PTEN mut group (Figure 2E-2K).

Evaluation and Validation of the PTEN Mutational Status Associated Prognostic Signature

We then evaluated and validated the prognostic ability of the PTEN mutational status associated prognostic signature in both training and validation datasets. Firstly, the risk score, survival status of the EC patients was shown in Figure 3A, D, G, J. Then, the EC patients were classified into the high- and low-risk groups based on the median risk score. K-M curve indicated that the patients in high-risk group have poor survival outcome (Figure 3B, E, H, K). The area under the curve (AUC) value at 1, 3, 5 years of the ROC curve in TCGA dataset was 0.706, 0.694 and 0.662, respectively (Figure 3C). The AUC value at 1, 3, 5 years of the ROC curve in ucec_tcga_pan_can_atlas_2018 dataset was 0.753, 0.804 and 0.853, respectively (Figure 3F). The AUC value at 1, 3, 5 years of the ROC curve in ucec_tcga_pub dataset was 0.888, 0.862 and 0.859, respectively (Figure 3I). The AUC value at 1, 3, 5 years of the ROC curve in our clinical specimens was 0.910, 0.806 and 0.782, respectively (Figure 3L). All of the above results highlighted the robust predictive potential of the PTEN mutational status associated prognostic signature.

Independent Prognostic Value of the PTEN Mutational Status Associated Prognostic Signature and Its Relationship with Clinicopathological Characteristics

To determine whether PTEN associated signature has independent prognostic value in TCGA dataset, univariate and multivariate Cox regression analyses was performed. Both univariate and multivariate Cox regression analyses indicated that stage, grade and risk score were related to the prognosis of EC patients (Figure 4A, B). Furthermore, in TCGA dataset, percentchart showed that the high-risk was associated with the younger, higher stage and grade of EC patients \((p<0.05; \text{Figure 4C-4E})\).

Construction of a nomogram model based on the PTEN-Associated Signature
The "rms" and "survival" package was used to construct a nomogram model based on the expression levels of the seven PTEN-associated signature to predict the survival rates of EC patients at 1, 3, and 5 years (Figure 5A). The calibration curves at 1, 3, and 5 years revealed high consistency between the actual and predicted survival rates, suggesting the powerful predictive performance of the nomogram model (Figure 5B–5D). DCA curve indicated that the prognostic ability of the model was accuracy (Figure 5E).

Mutational Landscape Associated with the PTEN Mutational Status Associated Prognostic Signature

Tumor mutation burden (TMB) refers to the total number of replacement and insertion/deletion mutations in each base group in the coding region of the evaluated gene exon in the genome of a tumor cell [12]. Figure 6A-6C revealed that the patients in high-risk group and those with PTEN mutation have higher TMB value. Sankey diagram showed the relationship between risk score, PTEN mutational status, TMB and survival status (Figure 6D). Finally, we investigate the mutational landscape associated with the PTEN mutational status associated prognostic signature and found that PTEN higher mutation frequency in high-risk group (Figure 6E).

Discussion

PTEN mutation is the most frequent type of mutation in EC [13]. Tumor suppressor PTEN has been shown to interact with cell adhesion complexes and stabilize cell junctions, thereby reducing invasion and metastasis of a range of cancer cells including EC [14-16]. Previous studies have been keen to investigate the influence of PTEN mutation on the progression and prognosis of endometrial cancer, but few focused on the development of PTEN mutational status associated prognostic signature [17-19]. In the present study, simple nucleotide variation, transcriptome profiling datasets were downloaded and analyzed. We found that PTEN have the higher mutation frequency in EC patients and PTEN mutation was associated with younger, lower stage and grade as well as favorable survival outcome of EC patients. We then conducted and verified a PTEN mutational status associated prognostic signature, which was associated with the malignant progression and prognosis of EC patients. We believe that this PTEN mutation-associated signature is a novel way to predict prognosis and evaluate efficacy of EC, and may become a new target for treatment of EC patients in the future.

In this study, a PTEN mutation-associated signature including seven genes GDPD2, GRB7, KCNK9, MUC3A, MYT1, RPS6KA6 and TSPYL5 selected by LASSO–Cox analysis were conducted. We reviewed previous studies and found that RPS6KA6 and TSPYL5 had been reported to be associated with the occurrence and development of EC. Ribosomal S6 kinase 4 (RSK4) as a tumor suppressor gene also known as RPS6KA6 has been shown to be significantly down-regulated in multiple malignancies including breast, colon, kidney, ovarian, and acute myeloid leukemia [20-27]. It can inhibit tumor cell proliferation, invasion, epithelial mesenchymal transformation [28, 29]. A study [30] reported that hypermethylation of RSK4 in EC results in a lower expression level of RSK4 in EC than in normal endometrial tissue. Lower RSK4 methylation was associated with higher EC grade. Testis-specific protein,
Y-encoded-like 5 (TSPYL5) is a member of the TSPYL family. According to existing studies, TSPYL5 expression is deleted or down-regulated in many tumors [31]. As a new tumor suppressor gene, TSPYL5 is closely related to the malignant progression and prognosis of tumors [32-34]. It had been reported that TSPYL5 was associated with the tumor differentiation, cell cycle and survival of EC [35, 36]. Although the role of GRB7, KCNK9, MUC3A, and MYT1 in EC has not been reported, their role in other tumors has been investigated. Growth factor receptor bound protein 7 (GRB7) is an important bridging protein that is involved in physiological and pathological processes such as embryonic development, angiogenesis, metabolic regulation and tumorigenesis by binding to tyrosine kinase receptors (RTKs). GRB7 has been reported to be involved in cell proliferation, migration, invasion, prognosis, and tumor-associated angiogenesis of a variety of tumors [37-39]. TASK-3 (KCNK9), as a member of the K2P potassium channel family, had been reported to be overexpressed in a variety of tumor tissues, such as breast cancer, gastric adenocarcinoma, ovarian cancer and lung adenocarcinoma and closely related to the progression of tumor [40-43]. MUC3A is a mucin cluster located on chromosome 7q22. As a tumor suppressor gene, MUC3A has been found to be low expressed in a variety of tumors, and is involved in malignant progression and prognosis of tumors [44, 45]. Myelin Transcription Factor 1 (MyT1) is mainly expressed in developing central nervous system cells, which mediate the proliferation and differentiation of oligodendrocytes and the formation of Myelin sheath in nerve cells [46]. Recent studies have shown that MYT1 is involved in the malignant progression of gastric cancer, liver cancer and glioblastoma [47-49]. Glycerophosphodiester phosphodiesterase 2 (GDPD2) is mainly involved in lipid metabolism and its effect on endometrial carcinoma and other tumors has not been reported [50].

To further evaluated the prognostic ability of the PTEN associated signature in both training and validation datasets as well as clinical specimens, ROC curve was plotted and AUC was calculated. We found that the mean of AUC more than 0.78, suggesting that the prognostic ability of PTEN associated signature was power. Moreover, univariate and multivariate Cox regression analyses revealed that the PTEN associated signature was an independent prognostic predictor for EC. Therefore, we believe that the PTEN associated signature has the potential to be a promising clinical prognostic tool of EC.

Conclusions

In summary, we conducted and validated a prognostic predictor for EC associated with PTEN mutational status. The PTEN mutation associated prognostic genes signature may be used as favorable prognostic biomarkers and therapeutic targets for EC.

Declarations

Ethics approval and consent to participate

This study was approved by the ethics committee of the ShengJing Hospital of China Medical University, and informed consent was obtained from all patients.
Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable.

Authors’ contributions

Ying Wu, Jun Wang, Lina Ge, and Qing Hu conceived and designed the study. Ying Wu, Jun Wang, Lina Ge, and Qing Hu developed the methodology. Ying Wu, Jun Wang, Lina Ge, and Qing Hu analyzed and interpreted the data. Ying Wu and Qing Hu wrote, reviewed, and/or revised the manuscript.

Acknowledgments

We thank the authors who provided TCGA and cBioPortal public datasets.

References


6. Malentacchi F, Turrini I, Sorbi F, Proietto E, Castiglione F, Fambrini M, Petraglia F, Pillozzi S, Noci I: Pilot investigation of the mutation profile of PIK3CA/PTEN genes (PI3K pathway) in grade 3 endometrial...


12. Bi F, Chen Y, Yang Q: **Significance of tumor mutation burden combined with immune infiltrates in the progression and prognosis of ovarian cancer.** *Cancer Cell Int* 2020, **20**:373.


**Figures**
Figure 1

PTEN mutational status of endometrial carcinoma in The Cancer Genome Atlas (TCGA) dataset. (A) Mutational landscape of endometrial carcinoma (EC) in TCGA dataset. (B) Kaplan-Meier curves revealed that the EC patients with PTEN mutation have favorable survival outcome. (C-E) Percent chart showing the PTEN mutation occur more frequently in the younger, lower stage grade EC patients.
Figure 2

Identification of differentially expressed genes (DEGs) and construction of the PTEN mutational status associated prognostic signature. (A) Heat map showing the top 40 DEGs between the EC patients with PTEN mutation or not. (B) Volcano plot showing the DEGs between the EC patients with PTEN mutation or not. (C) LASSO coefficient profiles of the 84 DEGs with prognostic value. (D) Nine PTEN mutation prognostic genes obtained from LASSO regression based on the ten-fold cross-validation and the
minimum criterion. (E-K) The relative expression level of seven prognostic genes (GDPD2, GRB7, KCNK9, MUC3A, MYT1, RPS6KA6 and TSPYL5) between the EC patients with PTEN mutation or not.

Figure 3

Evaluation and validation of the PTEN mutational status associated prognostic signature in training and validation dataset. Risk score, survival status, K-M curve and ROC curve in TCGA (A-C), ucec_tcga_pan_can_atlas_2018 (D-F) and ucec_tcga_pub dataset (G-I). clinical specimens (J-L).
Figure 4

Independent Prognostic Value of the PTEN Mutational Status Associated Prognostic Signature and Its Relationship with Clinicopathological Characteristics. (A) Univariate Cox regression analyses. (B) Multivariate Cox regression analyses. (C-E) Percent chart showing the high-risk was associated with the younger, higher stage and grade of EC patients.
Figure 5

Construction of a nomogram model. (A) A nomogram for predicting the 1-, 3-, and 5-year overall survival rates of EC patients. (B-D) The calibration curve at 1, 3, 5 years. (E) DCA curve evaluated the accuracy of the nomogram model.
Figure 6

Mutational landscape associated with the PTEN mutational status associated prognostic signature. (A) Differences in tumor mutation burden (TMB) between high- and low-risk group. (B) Differences in TMB between PTEN mutation or not of the patients in high-risk group. (C) Differences in TMB between PTEN mutation or not of the patients in low-risk group. (D) Sankey diagram showed the relationship between risk score, PTEN mutational status, TMB and survival status. (E) Mutational landscape associated with PTEN associated signature.

Supplementary Files

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

- IRB12.pdf
- Supplementarytable1.docx
• Supplementarytable2.docx
• Supplementarytable3.docx