Significance of *TP53* Mutational Status-Associated Signature in the Progression and Prognosis of Endometrial Carcinoma

Wancheng Zhao  
Shengjing Hospital of China Medical University

Fangfang Bi  
Shengjing Hospital of China Medical University

Xue Pan  (✉ panx@sj-hospital.org)  
Shengjing Hospital of China Medical University

Research

**Keywords:** endometrial carcinoma, TP53 mutation, signature, prognosis, TMB

**Posted Date:** September 20th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-820719/v1

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

Background: TP53 mutations are associated with poor outcome for patients with endometrial carcinoma (EC). However, to date, there have been no studies focused on the construction of TP53 mutational status-associated signature in EC. In this study, we aim to conduct a TP53 mutation associated prognostic genes signature for EC.

Methods: Hence, we explored the mutational landscape of TP53 in patients with EC based on the simple nucleotide variation data downloaded from The Cancer Genome Atlas (TCGA) database. Differential expression analysis and least absolute shrinkage and selection operator (LASSO)–Cox analysis was used to establish TP53 mutation associated prognostic genes signature. The overall survival rate between the high-risk and low-risk groups was compared by Kaplan-Meier (K-M) method.

Results: We found that the TP53 mutation was associated with poor outcome, older age, lower BMI, and higher grade and stage of EC in patients. A TP53 mutational status-associated signature was established based on transcriptome profiling data. Moreover, the patients in TCGA database were categorized into high- and low-risk groups. Kaplan–Meier (K-M) analysis indicated that the patients in the high-risk group have poor survival outcome. Furthermore, receiver operating characteristic (ROC) curves confirmed the robust prognostic prediction efficiency of the TP53 mutational status-associated signature. Finally, the prognostic ability was successfully verified in the other two datasets from cBioPortal database as well as in 60 clinical specimens.

Conclusion: In summary, our research constructed a powerful TP53 mutational status-associated signature that could be a potential novel prognostic biomarker and therapeutic target for EC.

1. Introduction

Endometrial carcinoma (EC), which originates from the endometrial epithelium, is the second most common malignancy of the female reproductive system [1]. Although surgery, chemotherapy, and immunotherapy for EC have led to some improvement in the clinical outcome [2], patient mortality rate is still high [3]. Therefore, it is of practical clinical significance to further explore the pathogenesis of EC at the molecular level, and to evaluate and predict the survival rate of patients by studying the prognostic signature of EC [4, 5].

Tumor protein P53 (TP53), located on the short arm of chromosome 17 (17p13.1) [6], has a broad spectrum of mutations in human cancers, including allelic loss, deletion, insertion, and point mutations [7]. Chromosomal deletion of TP53 gene is associated with the occurrence, chemotherapy resistance, and poor prognosis of many tumors [8]. Notably, about 70–80% of mutations in the TP53 gene are missense mutations caused by the substitution of a single nucleotide, which consequently, changes the corresponding amino acid residues. This change, especially the change of arginine residues, can significantly affect TP53 gene activity [9]. Moreover, the TP53 protein is inactivated in more than half of tumors [10]. A mutant TP53 protein not only loses its tumor suppressor function, but may also acquire a
functional expression similar to an oncogene, promoting the occurrence and development of cancer [11]. Therefore, TP53 could potentially be a novel biomarker of tumor prognosis and an effective therapeutic target.

Previous studies have confirmed the prognostic value of TP53 mutations in EC [12–14]. However, so far, no studies have focused on the construction of a TP53 mutational status-associated signature in EC. Hence, this study aimed to construct a TP53 mutational status-associated signature based on The Cancer Genome Atlas (TCGA) database.

2. Results

2.1 TP53 mutational status in EC

Based on the TP53 mutation data from the TCGA dataset, we found that the mutation frequency of the TP53 was 37%. As expected, K-M analysis confirmed that the patients with TP53 mutation had poor survival outcome (Fig. 1A, P < 0.001). Moreover, Figs. 1B–1E revealed that the TP53 mutation was related to poor outcome, older age, lower BMI, and higher grade and stage of EC in patients (P < 0.05).

2.2 Construction of the TP53 mutational status associated signature

A total of 1058 DEGs were identified based on p < 0.05 and log2|FC| > 1 screening standard (Supplementary table 2; Figs. 2A,B). Univariate Cox regression preliminarily screened 50 DEGs associated with the prognosis of patients with EC (Supplementary table 3, P < 0.001), then Lasso regression analysis identified 17 key DEGs (Figs. 2C, D). Of these, nine were further screened by forward stepwise method, then a TP53 mutational status-associated signature was constructed. Moreover, multivariate Cox regression analysis was used to calculate the risk score of each EC patient, with the following formula:

Risk score = 0.0012 x exp ERBB2 + 0.1121 x exp GLOD5 -0.0555 x exp KCNK6 + 0.0024 x exp MAL + 0.0029 x exp MUCL1 + 0.2903 x exp OR2W3 + 0.0471 x exp RBP2 + 0.7545 x exp STAC + 0.3902 x exp ZNF829 (Table 1). The difference analysis between TP53 mutation and TP53 wildtype group revealed that ERBB2, GLOD5, KCNK6, MAL, OR2W3, STAC and ZNF829 were expressed at higher levels in the TP53 mutation group, while MUCL1 and RBP2 were expressed at higher levels in the TP53 wildtype group (Figs. 3A–3I).
Table 1
Regression coefficients of the nine TP53 mutational status-associated prognostic genes.

<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>ERBB2</td>
<td>0.001239</td>
<td>1.001239</td>
<td>1.000481</td>
<td>1.001998</td>
<td>0.00135</td>
</tr>
<tr>
<td>GLOD5</td>
<td>0.112093</td>
<td>1.118617</td>
<td>1.058151</td>
<td>1.182539</td>
<td>7.70E-05</td>
</tr>
<tr>
<td>KCNK6</td>
<td>-0.05554</td>
<td>0.945978</td>
<td>0.899089</td>
<td>0.995311</td>
<td>0.032261</td>
</tr>
<tr>
<td>MAL</td>
<td>0.002416</td>
<td>1.002419</td>
<td>1.001031</td>
<td>1.003809</td>
<td>0.000632</td>
</tr>
<tr>
<td>MUCL1</td>
<td>0.002946</td>
<td>1.002951</td>
<td>1.001758</td>
<td>1.004145</td>
<td>1.22E-06</td>
</tr>
<tr>
<td>OR2W3</td>
<td>0.290267</td>
<td>1.336784</td>
<td>1.132684</td>
<td>1.577661</td>
<td>0.000595</td>
</tr>
<tr>
<td>RBP2</td>
<td>0.047118</td>
<td>1.048245</td>
<td>0.998892</td>
<td>1.100037</td>
<td>0.055503</td>
</tr>
<tr>
<td>STAC</td>
<td>0.754457</td>
<td>2.126455</td>
<td>1.56286</td>
<td>2.893294</td>
<td>1.57E-06</td>
</tr>
<tr>
<td>ZNF829</td>
<td>0.390186</td>
<td>1.477256</td>
<td>1.112411</td>
<td>1.961761</td>
<td>0.007017</td>
</tr>
</tbody>
</table>

2.3 Evaluation of the TP53 mutational status-associated signature

We then evaluated and validated the prognostic efficacy of the TP53 mutational status-associated signature in both training and validation datasets. The risk scores and survival status of each patient with EC have been determined (Figs. 4A, D, G, J). K-M analysis showed that patients in the low-risk group have longer survival time than in the high-risk group (Figs. 4B, E, H, K, P < 0.001). Moreover, ROC analysis showed that the overall survival rates of EC patients at 1-, 3-, and 5-year in the TCGA dataset were 0.775, 0.762, and 0.738, respectively (Fig. 4C). In the ucec_tcga_pan_can_atlas_2018 dataset, the overall survival rates at 1-, 3-, and 5-year were 0.764, 0.831, and 0.858, respectively (Fig. 4F), whereas in the ucec_tcga_pub dataset, the overall survival rates at 1-, 3-, and 5-year were 0.886, 0.878, and 0.890, respectively (Fig. 4I). Figure 4L revealed that the 1-, 3-, and 5-year overall survival rates (AUC) in clinical specimens were 0.925, 0.851 and 0.826, respectively.

2.4 Independent prognostic value of the TP53 mutational status-associated signature

With overall survival as the dependent variable, the risk score calculated by the TP53 mutational status-associated signature, age, BMI, pathological stage, and grade in TCGA dataset were used as covariates for univariate and multivariate Cox regression analyses. TP53 mutational status-associated signature has significant prognostic value, indicating that it could be used as an independent prognostic factor for EC patients (Fig. 5A, 5B, P < 0.001). We then investigated the relationship between the risk score and the
clinopathological characteristics in TCGA dataset, and found that patients with older age and higher EC grade and stage were more distributed in the high-risk group (Fig. 5C–F, P < 0.001).

2.5 Establishment of the nomogram model based on the TP53-Associated Signature

We successfully constructed a nomogram based on the expression levels of the above nine DEGs. After the clinicians input the expression values of nine genes for a specific EC patient into the nomogram, the corresponding score values in the score scale were obtained, and the resulting score values were added into the total score scale. Finally, a vertical line was drawn on the survival scale to estimate the survival rates at 1-, 3-, and 5-year (Fig. 6A). Calibration curves showed that the predicted survival rates of patients with EC were in good agreement with the actual survival rates at 1-, 3-, and 5-year (Fig. 6B–6D). Moreover, DCA results showed that the nomogram had high net income (Fig. 6E).

2.6 Mutational landscape associated with the TP53 mutational status-associated signature

TMB is defined as the number of tumor-specific mutations per million coding region bases [15]. Figures 7A–7C revealed that the patients with TP53 wild-type and those in the low-risk group had higher TMB values. Moreover, the Sankey diagram showed the relationship between risk score, TP53 mutational status, TMB, and survival status (Fig. 7D). Finally, we investigated the mutational landscape associated with the TP53 mutational status-associated prognostic signature and found that PTEN had higher mutation frequency in the high-risk group, while TP53, PPP2R1A, PIK3CA and MUC16 had low mutation frequencies (Fig. 7E).

3. Discussion

EC is the most common type of cancer in the female reproductive system [16]. In recent years, the understanding of EC has deepened, and some achievements have been made in the treatment and prognostic assessment of EC. However, there has still not been a breakthrough in treatment strategies, and individualized treatment of EC still faces great challenges. Previous studies have reported that the TP53 mutation is associated with poor outcome of patients with EC, which was confirmed in our research [17, 18]. However, to date, there are still no relevant studies on the development of a TP53 mutational status-associated signature. In our study, a TP53 mutational status-associated signature with powerful predictive potential in TCGA dataset was constructed, and verified its potential using two datasets from the cBioPortal database, as well as in 60 clinical specimens, indicating that this could be a novel prognostic biomarker and therapeutic target for EC.

This TP53 mutational status-associated signature was constructed using LASSO–Cox analyses of identified key DEGs, which included ERBB2, GLOD5, KCNK6, MAL, MUCL1, OR2W3, RBP2, STAC, and
To explore how these genes are involved in the development of EC, we reviewed the previous studies.

Erb-B2 receptor tyrosine kinase 2 (ERBB2), also known as HER2, is a member of the ERBB family [19]. ERBB2, as a proto-oncogene, has been confirmed to be upregulated in EC tissues and is related to poor prognosis [20]. Several targeted therapies for ERBB2, such as trastuzumab, pertuzumab, and lapatinib, have been used in the clinical setting [21]. Potassium channel subfamily K member 6 (KCNK6) is the background potassium channel belonging to the potassium channel family of double pore domain. KCNK6 is upregulated in thyroid carcinoma and breast cancer and is related to the proliferation, invasion and migration of breast tumor cells [22, 23]. Myelin and lymphocyte protein (MAL) encodes T lymphocyte maturation-related proteins and plays a role in T cell differentiation. Down-regulated MAL, as a tumor suppressor gene, was associated with a variety of human epithelial malignancies [24]. A study revealed that the MAL can be used for the early diagnosis of EC [25]. Mucin-like 1 (MUC1), also known as SBEM, is a breast-specific gene that is associated with the occurrence, progression, prognosis, and chemotherapy response of breast cancer [26]. OR2W3, which belongs to the ORS gene family, has been revealed to be related to the progression of breast cancer [27]. Retinoblastoma-binding protein 2 (RBP2) belongs to the JARID protein family, and is responsible for histone demethylase (HDM) activity. As a chromatin-modifying enzyme, it has been shown to be involved in the development and progression of a variety of cancers [28]. Src homology three (SH3) and cysteine rich domain (STAC) encodes a cysteine-rich protein containing the SH3 domain, which is mainly expressed in neurons and may be involved in neuron-specific signal transduction [29]. So far, no relevant studies have been found on the GLD5 and ZNF829 genes. Although most of these genes have not been previously reported in EC, they have been found to play an important role in the development of other tumors.

To evaluate and validate the prognostic value of the TP53 mutational status-associated signature in both the training and validation datasets, as well as in clinical specimens, an ROC curve at 1, 3 and 5 years was plotted. We found that the mean of AUC value was more than 0.80, indicating that the TP53 mutational status-associated signature have a powerful prognostic ability.

4. Conclusion

In summary, we conducted and validated a TP53 mutational status-associated signature with robust predictive potential. To our knowledge, this is the first study to do so. The TP53 mutational status associated signature could potentially be used as a novel prognostic biomarker and therapeutic target for EC.

5. Materials And Methods

5.1 Data acquisition
A dataset with 529 patients with EC including simple nucleotide variation, transcriptome profiling datasets and clinical data was acquired from TCGA, and was utilized as the training dataset (https://portal.gdc.cancer.gov/). The ucec_tcga_pan_can_atlas_2018 and ucec_tcga_pub datasets, which contained information of 527 and 331 patients with EC, respectively, along with their transcriptome profiling datasets and corresponding clinical information, were downloaded from cBioPortal database (http://www.cbioportal.org/study/summary?id=ucec_tcga) and used as validation datasets. The deadline for the dataset was June 2020. Inclusion criteria for sample screening included: (1) Primary endometrial cancer confirmed by pathology without any preoperative radiotherapy or chemotherapy; and (2) Prognostic information was complete without deletion.

### 5.2 Specimen collection

A total of 60 patients with EC that were admitted to the Obstetrics and Gynecology department of the Shengjing Hospital of China Medical University from January 2016 to December 2016 were selected as the research objects. The ages of patients ranged from 26 to 76 years, and the average age was 56.22 ± 10.51 years. In terms of FIGO stage, 26 cases were at stage I, 12 were at stage II, 14 cases were at stage III, and 8 cases were at stage IV. In terms of pathological grades, there were 21, 14, and 25 cases at G1, G2, and G3, respectively.

### 5.3 Identification of differentially expressed genes

The “edgeR” package in R was used to screen differentially expressed genes (DEGs) between patients with $TP53$ mutation or not. Inclusion criteria was log(|FC|) > 1 and p < 0.05. The “ggplot2” package in R was used to draw the volcano map, and the “ComplexHeatmap” package was used to draw the heat map in order to show the differential expression in patients with $TP53$ mutation or not.

### 5.4 Construction and validation of the $TP53$ mutational status-associated signature

The “Survival” package in R was performed to obtain the DEGs associated with prognostic value according to univariate Cox regression analysis. DEGs with significant prognostic value (P < 0.001) were screened to establish the $TP53$ mutational status associated signature using LASSO–multivariate Cox analysis. The risk score for each patient was calculated using the following formula: risk score = Σ (regression coefficient × gene expression of DEGs). The median value of risk score was used to classify the patients into high- and low-risk groups, and the K-M and Log-rank method was used to compare the overall survival outcome between the two groups. A receiver operating characteristic (ROC) curve was plotted evaluate the prognostic ability of the $TP53$ mutational status-associated signature at different time endpoints using the "Survival" and "timeROC " in R software. In addition, to evaluate the predictive performance of the $TP53$ mutational status-associated signature, we verified our analysis results using the ucec_tcga_pan_can_atlas_2018 and ucec_tcga_pub datasets from the cBioPortal database, as well as in 60 clinical specimens.

### 5.5 RT-qPCR analysis
Trizol was used to extract the Total RNA of EC samples (Takara, Japan), then complimentary DNA (cDNA) was synthesized using the PrimeScript RT kit (Takara, Japan), following the manufacturer’s instructions. The cDNA was amplified using SYBR Premix ExTaq kit (Takara, Japan), and mRNA hydrolevel was detected using an ABI Prism 7000 fluorescence quantitative PCR assay (Applied Biosystems, Waltham, MA, USA). GAPDH was used as the internal reference, and mRNA expression in the TP53 mutational status-associated signature was calculated by $2^{-\Delta\Delta CT}$ method. The sequences of primers used for RT-qPCR was displayed in Supplementary table 1. Then, we established a TP53 mutational status-associated signature as per the method used for the training dataset based on the expression level of mRNAs. The risk scores of 60 clinical specimens were obtained, and specimens were classified into high- and low-risk groups.

5.6 Establishment of the nomogram model based on the TP53-associated signature

To maximize clinical decision making, the "rms" package in the R software was used to conduct a nomogram based on the expression level of genes in the TP53 mutational status-associated signature. Calibration and Decision Curve Analysis (DCA) curves were constructed to evaluate the correction degree and the efficacy of the histogram for prognosis prediction in different sample sets, respectively.

5.7 Statistical analysis

Univariate and multivariate Cox regression analysis were used to evaluate the independence of TP53 mutational status-associated signature for EC. The total number of mutations in the DNA of cancer cells (TMB) of the patients with EC in TCGA database was calculated by Perl software. Wilcoxon rank-sum test was used for comparative analysis between two groups. All statistical analyses were performed using the R language (version 3.6.2). Bilateral test $P < 0.05$ was considered statistically significant.

Abbreviations


Declarations

Ethics approval and consent to participate

This study was approved by the ethics committee of the ShengJing Hospital of China Medical University, oral 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**

Xue Pan, Fangfang Bi, Wancheng Zhao conceived and designed the study. Xue Pan, Fangfang Bi, Wancheng Zhao developed the methodology. Xue Pan, Fangfang Bi, Wancheng Zhao analyzed and interpreted the data. Xue Pan, Fangfang Bi, Wancheng Zhao wrote, reviewed, and/or revised the manuscript.

**Acknowledgments**

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

**References**


21. Minami Y, Shimamura T, Shah K, LaFramboise T, Glatt KA, Liniker E, Borgman CL, Haringsma HJ, Feng W, Weir BA *et al.* The major lung cancer-derived mutants of ERBB2 are oncogenic and are


**Figures**
Figure 1

TP53 mutational status in EC patients from the TCGA dataset. (A) Kaplan-Meier (K-M) analysis confirmed that the patients with TP53 mutation have poor survival outcome. (B–E) TP53 mutation was related to poor outcome, older age, lower BMI, and higher levels of grade and stage of patients with EC.
Figure 2

Construction of the TP53 mutational status-associated signature. (A) Heat map showing the top 40 DEGs between patients with TP53 mutation or not. (B) Volcano plot showing the DEGs between patients with TP53 mutation or not. (C) Regression coefficient of 17 DEGs based on the Lasso model. (D) Variables were screened by 10-fold cross-validation method in the Lasso model. When the number of DEGs variables was 17, the \( \lambda \) value corresponding to the minimum partial likelihood deviation was obtained.
Figure 3

(A–I) The relative expression level of nine prognostic genes (ERBB2, GLOD5, KCNK6, MAL, MUCL1, OR2W3, RBP2, STAC, and ZNF829) between patients with TP53 mutation or not.
Figure 4

Evaluation of the TP53 mutational status-associated signature in training and validation datasets, as well as in clinical specimens. (A–C) Risk score, survival status, K–M curve and ROC curve in TCGA, (D–F) ucec_tcga_pan_can_atlas_2018, (G–I) ucec_tcga_pub dataset, (J–L) clinical specimens.
Figure 5

Independent prognostic value of the TP53 mutational status-associated signature. (A) Univariate Cox regression analyses. (B) Multivariate Cox regression analyses. (C–E) Patients with older age, and higher EC grade and stage were more distributed in the high-risk group.
Figure 6

Construction of a nomogram model. (A) A nomogram model predicting survival in EC patients at 1-, 3-, and 5-year. (B–D) The calibration curve at 1, 3, 5 years. (E) Decision curve analysis (DCA) curve evaluated the clinical benefit of the nomogram model.
Figure 7
Mutational landscape associated with the TP53 mutational status-associated signature. (A) Differences in TMB between the high- and low-risk groups. (B) Differences in TMB between EC patients with TP53 mutation or not in the high-risk group. (C) Differences in TMB between EC patients with TP53 mutation or not in the low-risk group. (D) Sankey diagram showed the relationship between risk score, TP53 mutational status, TMB and survival status. (E) Mutational landscape associated with the TP53-associated signature.

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.
• Supplementarytable1.docx
• Supplementarytable2.docx
• Supplementarytable3.docx