Original Article

Low expression of developing brain homeobox 2 (Dbx2) may serve as a biomarker to predict poor prognosis in endometrial cancer

Xinlu Zhang1*, Yaping Wang1*, Shujun Zhao1,2, Qiaohong Qin1,2, Min Zhang1, Yi Jiang1, Hai Zhu1, Hongyu Li1,2

1Gynecologic Oncology, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou 450000, Henan, P. R. China; 2Zhengzhou Key Laboratory of Gynecological Oncology, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou 450000, Henan, P. R. China. *Equal contributors.

Received December 3, 2020; Accepted February 2, 2021; Epub May 15, 2021; Published May 30, 2021

Abstract: Objective: For investigating Dbx2’s expression in endometrial cancer (EC) and its effect on prognosis of patients with EC. Methods: A comparison was performed in the Cancer Genome Atlas (TCGA) database in terms of the expression profiling of EC and the survival data. To obtain differential expression genes (DEGs), Volcano plot and Venn analysis were adopted. DEGs function was performed by carrying out the GO annotation analysis (GO) and gene set enrichment analysis (GSEA). In clinical EC samples, PCR was applied to the verification of Dbx2’s expression. Results: Dbx2 was a downregulated expression in tumor tissues. Dbx2 can have a poor prognosis role in EC by regulating the apoptotic signaling pathway and the immune pathway. Lower expression of Dbx2 was related to lymph node metastasis and FIGO stage. Conclusion: Dbx2 is downregulated in endometrial cancer, which serves as a biomarker to predict poor prognosis.

Keywords: Endometrial cancer, Dbx2, prognosis, biomarker, bioinformatics analysis

Introduction

As a prevailing malignancy, endometrial cancer (EC) causes the fourth highest cancer-related mortality among females worldwide [1]. About 95% of the early-stage EC patients survive for more than five years. However, the prognosis of advanced, recurrent, or metastatic EC patients remains unsatisfactory [2-4]. Traditional surgery, radiotherapy, and chemotherapy have made some progress in the treatment of endometrial cancer, but there is still some work to be done. The development of new therapeutic targets and effective prognostic biomarkers has become necessary to identify the prognosis of EC and to offer individualized options of treatment.

With the development of molecular biology, genetics research brings new hope for cancer treatment. In previous studies, some promising prognostic biomarkers have been developed for the treatment of endometrial cancer. One study demonstrated that the enhanced expression of KLKS-8 in EC can be partly explained by the poorer prognosis of EC patients. It identified KLKS-8 as an independent factor for the prognosis of EC patients in terms of the overall survival (OS) [5]. Four miRNAs (miR-4758, miR-876, miR-142, and miR-190b), as independent prognostic factors, were used to classify EC patients into two groups, low and high-risk, to provide survival risk prediction for EC patients [6]. In endometrioid endometrial carcinoma, HMGA1 overexpression may be a prognostic biomarker, which opened a new door for cancer treatment and diagnosis [7]. However, there is still no research where mRNA is directly screened according to prognosis. Therefore, we attempted to classify the patients by prognosis and to screen the mRNA profiles of the groups, respectively.

Materials and methods

Microarray data

The TCGA database was used to collect microarray data. A total of 66 samples of RNA-se-
Dbx2 may serve as a biomarker in EC

sequencing (RNA-seq) data were involved in this study, of which 42 tumor samples were collected from EC patients and 24 normal tissues were obtained from healthy donors.

**Identification of DEGs**

The collected microarray data were split into 2 groups corresponding to two different categories, respectively. First, it was based on patient prognosis. Then, it was based on the expressions of genes in both normal endometrium mucosa tissues and tumor tissues. Under the support of Venn diagrams as a web tool, the visualization of the differentiated gene expressions in each tissue was realized. In doing so, the specific genes having correlation with the prognosis of EC patients were identified.

**GO annotation analysis**

The Gene Ontology (GO) project was employed to carry out functional analysis of the differentially expressed genes in our study [8].

**PPI network construction**

Through the Search Tool for the Retrieval of Interacting Genes (STRING, https://string-db.org) online database, the establishment of a PPI network was completed for the assessment of proteins' interaction [9].

**Gene set enrichment analysis (GSEA)**

On the GSEA software, GSEA was conducted to identify the pathways in two different Dbx2 mRNA level groups. The genome extracted from the Molecular Signatures Database (MSigDB) was used in our study. The data obtained from TCGA was applied for GSEA analysis and the pathways with an FDR <0.05 were treated as significant.

**Clinical specimens**

The patients at the Third Affiliated Hospital of Zhengzhou University (Zhengzhou, China) between June 2013 and August 2015 were targeted to obtain normal and tumor endometrial tissues, each with the number of 40. All endometrial cancer cases we used in our study were all type I endometrial cancer. The pathological types of endometrial cancer cases we used were all endometrial adenocarcinoma. The tumor specimens as well as adjacent noncancerous tissues were obtained. We immediately froze the specimens in liquid nitrogen and kept them at -80°C after resection. The approval for the whole research was obtained from the Institutional Ethics Committee of the Third Affiliated Hospital of Zhengzhou University (Ethics approval number: (2019) medical review NO. 87). The clinical characteristics and parameters of patients are indicated in Table 1. We collected informed consent from all subjects with follow-up information available.

**RNA extraction and quantitative real-time polymerase chain reaction (RT-qPCR)**

Trizol solution (Invitrogen, Waltham, MA, USA) was applied to the extraction of total RNA. PrimeScript RT reagent kit (Takara Bio, Otsu,
Dbx2 may serve as a biomarker in EC

Shiga, Japan) was used for the conversion of RNA into cDNA. We then performed RT-qPCR analysis with specific primers and SYBR Green qPCR Master Mix (Takara, Japan) in line with the manufacturer’s instructions through the ABI 7500 system. The qPCR amplification con-

Figure 1. Expression differences of genes associated with prognosis in endometrial cancer. A. Patients were divided into 2 groups according to the expression of RNA-seq data from TCGA database. B. The volcano plots of RNA-seq data from the TCGA database and the x-axis represents the log2 transformed of fold change ratios; the y-axis is the log10 transformed adjusted p-value. The red colored dots represented the DEGs based on fold change >2. Herein, the volcano plot indicated the different genes when comparing the long OS patient group with the short OS patient group. C. The plot of 78 up-regulated and 102 down-regulated genes based on volcano analysis. D. Hierarchical clustering analysis of the candidate genes associated with the survival data obtained from TCGA database. E. Hierarchical clustering analysis of the RNA-seq data of different genes in 42 CA samples and 24 normal endometrial mucosa samples. F. With a threshold of P<0.05, false discovery rate (FDR) <0.05 and fold change >2, DEGs were selected by volcano plot when comparing 42 CA samples with 24 normal endometrial mucosa samples from TCGA database. G. The plot showed 263 up-regulated and 56 down-regulated genes based on the above-mentioned volcano analysis. H. Venn diagram representing the distribution of DEGs in different groups. 4 DEGs were expressed in both CA and short survival patients.
Dbx2 may serve as a biomarker in EC

**Table 2. Expression of 4 genes associated with a poor endometrial cancer prognosis in TCGA**

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene ID</th>
<th>Description</th>
<th>Style</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNNI3</td>
<td>817943</td>
<td>troponin I type 3</td>
<td>down</td>
</tr>
<tr>
<td>TNNI2</td>
<td>779979</td>
<td>troponin I type 2</td>
<td>up</td>
</tr>
<tr>
<td>Dbx2</td>
<td>772847</td>
<td>developing brain homeobox 2</td>
<td>down</td>
</tr>
<tr>
<td>CYP2A6</td>
<td>803764</td>
<td>cytochrome P450, family 2, subfamily A, polypeptide 6</td>
<td>down</td>
</tr>
</tbody>
</table>

**Figure 2.** A lower expression of Dbx2 was observed in TCGA database to predict poor prognosis. A-D. mRNA expression of DEGs in cancer vs control from patients of TCGA database (**, P<0.01; ***, P<0.001; ****, P<0.0001). E-H. Kaplan-Meier curve of DEGs was provided by patients from TCGA data.

After the first procedure of 5 min at 94°C, we conducted amplification for 30-second denaturation of 30 cycles at 94°C, followed by 30-second annealing at 58 °C and 30-second elongation at 72°C. After the final cycle, we performed terminal elongation (5 min at 72°C) and stored the samples at 4°C. The adopted primers included 5'-GGAGGCCAAAGG-GTCATCATCTC-3' sense primer and 5'-GAGGGGCCCATCCACAG-TCTTC T-3' antisense primer for GAPDH, 5'-ACTCTAATTCC-AAAGCTCAGAGG-3' sense primer and 5'-GGCAAGTTCGGTTCGTTT-3' antisense primer for Dbx2. GAPDH was taken as a way to impose internal control. The relative expression of the target gene and the GAPDH were calculated by comparing Ct (2^ΔΔCt) values [10].

**Statistical analysis**

All bioinformatics analyses were carried out on R software (Version 3.4; R Foundation for Statistical Computing, Vienna, Austria). Prism 7 (GraphPad Software Inc., La Jolla, USA) and SPSS 19.0 (SPSS Inc., Chicago, IL, USA) were used for data analysis. The use of χ^2 test aimed at the comparison of clinical-pathologic factors. Continuous variables were compared by one-way variance (ANOVA) analysis followed by Tukey’s post hoc test or unpaired students’ test. The log-rank test and Kaplan-Meier curves were used to assist with survival analysis. How Dbx2 expression was associated with clinical parameters was determined by the construction of univariate and multivariate logistic regression models. The probability values of P<0.05 were identified as statistical significance.
Dbx2 may serve as a biomarker in EC

Figure 3. Gene set enrichment analysis (GSEA) of Dbx2. A. GO analysis showed Dbx2 was closely related to the numerous signaling pathway in EC. B-F. Multi-GSEA for Dbx2-associated signaling pathways.
Dbx2 may serve as a biomarker in EC

Results

Gene expression analysis

The TCGA data portal (https://cancergenome.nih.gov/) was used to collect EC samples, involving 24 normal endometrium tissues and 42 tumor samples. The raw data and clinical information of these EC patients were downloaded to conduct later analyses.

Identification of specific genes signature

For the identification of the underlying genes bearing correlation with the poor survival rates of EC patients, the patients from TCGA database were split into 2 groups based on OS, including short (<1000 days) OS and long (>1000 days) OS (Figure 1A). Through Volcano plot and hierarchical cluster, we identified a total of 180 DEGs (fold change >2), among which 102 were down-regulated and 78 were up-regulated (Figure 1B-D). Subsequently, compared to normal tissues, the up-regulation of 263 genes and down-regulation of 56 genes were found in tumor tissues (Figure 1E-G) (P<0.05, FDR <0.05, FC >2, respectively). The Venn diagram showed 4 crossed DEGs from downregulated genes of the above two screening methods (Figure 1H). The expression of 4 DEGs was analyzed through TCGA databases, which were shown in Table 2.

Low expression and poor prognosis of DEGs verified in database

To help screen the prognostic significance of the identified DEGs, we verified the DEGs’ expression in the database. A total of 4 genes (Dbx2, CYP2A6, TNNI2, and TNNI3) were confirmed as candidates of DEGs by TCGA database (Figure 1H). Three DEGs (Dbx2, TNNI3, and CYP2A6) were downregulated in the tumor tissues of EC patients in the database (Figure 2B-D). There was no diagnostic significance observed in the expression of CYP2A6 among EC patients according to the Kaplan-Meier curves (Figure 2G; P=0.2567). Therefore, 2 DEGs (Dbx2 and TNNI3) were confirmed as closely associated with the prognosis of EC patients. As shown in the Kaplan-Meier survival analysis, the lower-level expression of Dbx2 displayed a close correlation with the poor survival rates of EC patients (Figure 2H). Additionally, lower expression of Dbx2 was observed in tumor tissues (Figure 2D). In this case, Dbx2 attracted more attention from us and prompted the further analysis using Dbx2. Therefore, Dbx2 was identified as a candidate gene for the poor prognosis of EC patients.

Significant GOs and pathways

As suggested GO analyze, it is speculated that Dbx2 was closely related to the numerous signaling pathway, including the regulation of apoptotic signaling pathway, KAPPAB transcription factor activity, the negative regulation of innate immune response, the negative regulation of stem cell proliferation, type I interferon production, and interferon GAMA mediated signaling pathway (Figure 3).

A PPI network of the genes

The PPI network consists of 11 nodes capable of interaction and 23 edges. It was verified that IPX6, LHX1, ISL1, OLIG2, ISL2, and IRX3 were closely correlated with each other and they were confirmed as hub proteins (Figure 4).

Validation of clinical patients

The present research revealed a significant reduction of Dbx2 in EC tumor tissues (Figure 5A). The correlation between the expression of Dbx2 and clinic-pathological variables was conducted. The data suggested the patients in advanced FIGO stage and low grade have a significantly lower expression of Dbx2 (Figure 5B, 5C). The association between a poorer prognosis and the suppressed expression of Dbx2 in tumor tissues was identified (P=0.017) (Figure 5D), as was a close correlation of FIGO stage...
Dbx2 may serve as a biomarker in EC

Figure 5. Dbx2 is a predicator of poor prognosis. A. The mRNA expression level of Dbx2 in different groups were exhibited. B. The mRNA expression levels of Dbx2 were reflected in the correlation with FIGO stage. C. The mRNA levels of Dbx2 were compared in respect of histological differentiation. D. Impact of Dbx2 expression on the overall survival rates in clinical patients (n=40) (**, P<0.01; ***, P<0.001; ****, P<0.0001).

(P=0.001) and lymph node metastasis (P=0.013) with the reduced expression of Dbx2 (Table 3). We speculated Dbx2 as a prognostic gene having a correlation with the progression of tumor in EC patients.

The logistic regression analysis was conducted to establish whether Dbx2 was useful in predicting the prognosis of EC or not. The association of EC patients’ survival rates with advanced FIGO stage (OR=3.256; P=0.011), lymph node metastasis (OR=4.892; P=0.021), and abnormal CA125 level (OR=9.939; P=0.024) was revealed by univariate analysis. The advanced FIGO stage (OR=4.84; P=0.020) and lymph node metastasis (OR=6.90; P=0.010) were found from multivariate analyses as independent predictors of prognosis (Table 4). Overall, the above findings have demonstrated the crucial role of Dbx2 expression in the prediction of prognosis for EC patients.

Discussion

High-throughput technologies were increasingly applied. Many novel biomarkers and therapeutic targets have been explored by analyzing the transcriptomes, using a highly significant tool to assist with cancer studies [11]. Patients were classified before receiving individualized treatment according to the expression levels of specific genes, making diagnosis and treatment of cancer more efficient [12, 13]. However, effective tumor markers and therapeutic targets of endometrial cancer remain unclear due to the limitations of research. Our research will contribute to predicting the prognosis of EC patients and offering individualized options for treatment [14].

In our study, the novel genes correlated with poor prognosis in EC were determined through the microarray analysis of mRNA profiling data obtained from the TCGA database. First, the gene expressions between normal and tumor tissues and a comparison between short OS groups and long OS in the TCGA database were systematically analyzed. The data indicated that three DEGs (Dbx2, TNNI3, and CYP2A6) of EC patients were down-regulated significantly. Kaplan-Meier curves presented that CYP2A6 has no prognosis value in EC. In this case, Dbx2 attracted more attention and prompted us to perform further analysis using it. Subsequently, in clinical EC samples, Dbx2 expression was verified by RT-qPCR, and our results were consistent with the above. GO and GSEA were adopted, and the correlations between candidate targets were studied to evaluate the role of Dbx2 in promoting disease progression, especially in the tumor microenvironment. As revealed by GO analysis, Dbx2 is involved in numerous signaling pathways. We speculated the close correlation of Dbx2 with poor prognosis in EC patients through the participation in the apoptotic signaling pathway and the immune pathway. The PPI network of the genes exhibited that Dbx2 as a center target was closely correlated with IPX6, LHX1, ISL1, OLIG2, ISL2, and IRX3. It can play a vital role in multiple gene regulatory pathways, which needs to be further verified.

Dbx (developing brain homeobox gene) was first reported in 1992 in a 12.5-day embryonic mouse telencephalon cDNA library [15, 16].
The Dbx2 gene encodes various homeodomain-containing proteins that contain 251 amino acids. This evolutionarily conserved protein largely promotes the differentiation of the central nervous system in vertebrates [17]. To date, two Dbx genes (Dbx1 and Dbx2) discovered in zebrafish are found to play a crucial role in neural patterning and development [18]. The published studies on Dbx2 functions focused on neural differentiation and patterning, mainly including neuronal modulation, signal transduction, driving brain evolution, and spinal cord regeneration [19-23]. The Dbx family’s essential role in different tumor tissues has been revealed in recent years. In hepatocellular carcinoma (HCC) tissues, the up-regulation of Dbx2 activated the sonic hedgehog (Shh) pathway to promote migration, invasion, and proliferation of HCC cells [24]. Breezy et al. observed an overrepresentation of homeobox family genes (DBX1, NKX2-6, and SIX6) in the whole genome DNA methylation signature of HER2-positive breast cancer [25]. The methylation level of DBX1 in hepatocellular carcinoma was significantly less compared to healthy controls, as verified by Zhang et al. [26]. It still remains undetermined whether different types of cancer, tissue samples, or testing methods lead to different levels of gene regulation. The consistent results of our clinical data with the TCGA database support low Dbx2 expression as a novel candidate biomarker gene of poor prognosis in endometrial carcinoma. However, the regulatory mechanism requires further validation to completely reveal the Dbx2-targeted therapy in EC.

TNNI2 and TNNI3 were responsible for distal arthrogryposis (DA), and muscle type-specific and developmental regulations in both the heart ventricle and the atrium [27, 28]. Showing a close correlation with peritoneal metastasis, high TNNI2 expression was considered an independent risk marker for peritoneal recurrence after curative gastrectomy [29]. Using high-throughput RNA sequencing, SYT8/TNNI2 was detected as a novel fusion transcript in bladder cancer [30]. Therefore, we hypothesized the essential role of TNNI3 and TNNI2 in cancer, which should become the future hotspots of tumor research.
Dbx2 may serve as a biomarker in EC

Table 4. Logistic regression model analysis of patients with endometrial cancer

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Univariate</th>
<th>Multivariate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Age (&lt;57 vs ≥57)</td>
<td>1.048</td>
<td>0.443-2.479</td>
</tr>
<tr>
<td>Tumor size (&lt;30 mm vs ≥30 mm)</td>
<td>1.935</td>
<td>0.631-5.934</td>
</tr>
<tr>
<td>Lymph node metastasis (no vs yes)</td>
<td>3.256</td>
<td>1.460-10.125</td>
</tr>
<tr>
<td>FIGO stage (I/II vs III/IV)</td>
<td>4.892</td>
<td>1.138-11.348</td>
</tr>
<tr>
<td>Histological differentiation (poor-medium vs well)</td>
<td>1.531</td>
<td>0.435-6.236</td>
</tr>
<tr>
<td>Myometrial invasion (surface vs deep)</td>
<td>4.118</td>
<td>1.081-21.156</td>
</tr>
<tr>
<td>Menopause (no vs yes)</td>
<td>1.179</td>
<td>0.486-2.856</td>
</tr>
<tr>
<td>CEA (&lt;5 vs ≥5)</td>
<td>0.785</td>
<td>0.331-1.860</td>
</tr>
<tr>
<td>CA125 (&lt;70 vs ≥70)</td>
<td>9.939</td>
<td>1.156-13.538</td>
</tr>
<tr>
<td>CA199 (&lt;35 vs ≥35)</td>
<td>0.659</td>
<td>0.224-1.940</td>
</tr>
<tr>
<td>CA724 (&lt;6.9 vs ≥6.9)</td>
<td>1.031</td>
<td>0.363-2.930</td>
</tr>
</tbody>
</table>

Note: *: P<0.05. CEA, carcinoembryonic antigen; CA, carbohydrate antigen; OR, odds ratio; 95% CI, 95% confidence interval.

Genetic polymorphisms in the CYP2A6 gene have been associated with variations in enzyme activity. The higher the CYP2A6 activity related to smoke, the more extensive it was exposed to higher levels of carcinogens, which raised the risk level of lung cancer [31]. CYP2A6 polymorphisms are correlated with the efficacy of S-1 in the adjuvant setting for gastric cancer [32]. However, CYP2A6 has no prognosis value in EC, as revealed by the Kaplan-Meier curves in our data. Different cancer types and the enzymes involved in the gene may not be excluded. LHX1 is a significant transcription factor that combines with NaB to cause cell death in colorectal cancer [33]. ISL2 modulates angiogenesis and promotes cell proliferation and malignant transformation in oligodendroglioma [34]. OLIG2 negative staining differentiated the types of soft tissue tumors, further supporting that these tumors have a different stem cell of origin [35]. These genes will be validated in our future studies.

There are certain limitations that our study is subject to and they are worthy of further discussion. First, the present research failed to verify Dbx2 expression endometrial cancer from the protein level. Second, this study involved no data obtained from in vivo animal models to validate the existing results. Third, this study failed to explore the upstream regulators of Dbx2 and the regulatory roles played by Dbx2 in the downstream of some signaling pathway in EC. There is no doubt that further investigation and experimental validation are needed to find the role played by Dbx2 in EC.

In conclusion, the down-regulation of Dbx2 in EC and its close correlation with poor prognosis were found. Dbx2 has a poor prognostic role in EC by regulating apoptotic signaling pathway and the immune pathway. The lower Dbx2 expression was found to have a close correlation with FIGO stage and lymph node metastasis. Therefore, Dbx2 is a therapeutic target and prognostic bio-marker for EC.

Acknowledgements

We extend our gratitude to all people making contributions to this study. This paper was supported by the Medical Science and Technology Offensive Joint Development Project of Henan Province, China (No. 2018020220).

Disclosure of conflict of interest

None.

Address correspondence to: Hongyu Li, Gynecologic Oncology, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou 450000, Henan, P. R. China. Tel: +86-371-13939077898; Fax: +86-371-13939077898; E-mail: lihongyu99@qq.com

References

Dbx2 may serve as a biomarker in EC


Dbx2 may serve as a biomarker in EC


