An 8-gene signature predicts the prognosis of cervical cancer following radiotherapy

FEI XIE¹, DAN DONG², NA DU³, LIANG GUO⁴, WEIHUA NI¹, HONGYAN YUAN¹, NANNAN ZHANG¹, JIANG JIE¹, GUOMU LIU⁵ and GUIXIANG TAI¹

¹Department of Immunology, College of Basic Medical Science, Jilin University; Departments of ²Obstetrics and Gynecology, ³Infections, ⁴Pathology and ⁵Nephrology, The First Hospital of Jilin University, Changchun, Jilin 130021, P.R. China

Received July 11, 2018; Accepted May 10, 2019
DOI: 10.3892/mmr.2019.10535

Abstract. Gene expression and DNA methylation levels affect the outcomes of patients with cancer. The present study aimed to establish a multigene risk model for predicting the outcomes of patients with cervical cancer (Cerc) treated with or without radiotherapy. RNA sequencing training data with matched DNA methylation profiles were downloaded from The Cancer Genome Atlas database. Patients were divided into radiotherapy and non-radiotherapy groups according to the treatment strategy. Differently expressed and methylated genes between the two groups were identified, and 8 prognostic genes were identified using Cox regression analysis. The optimized risk model based on the 8-gene signature was defined using the Cox's proportional hazards model. Kaplan-Meier survival analysis indicated that patients with higher risk scores exhibited poorer survival compared with patients with lower risk scores (log-rank test, P=3.22x10⁻⁷). Validation using the GEO44001 gene set demonstrated that patients in the high-risk group exhibited a shorter survival time compared with the low-risk group (log-rank test, P=3.01x10⁻³). The area under the receiver operating characteristic curve values for the training and validation sets were 0.951 and 0.929, respectively. Cox regression analyses indicated that recurrence and risk status were risk factors for poor outcomes in patients with Cerc treated with or without radiotherapy. The present study defined that the 8-gene signature was an independent risk factor for the prognosis of patients with Cerc. The 8-gene prognostic model had predictive power for Cerc prognosis.

Introduction
Cervical cancer (Cerc) is a leading cause of cancer-associated mortality in women worldwide (1,2). Surgery in combination with chemotheraphy and radiotherapy is the most common strategy for Cerc treatment. Radiotherapy significantly improves Cerc patient prognosis (3). However, the overall survival of patients with Cerc diagnosed at advanced stages remains poor, with the 5-year survival rate ≤50%, despite advanced surgical protocols and diagnostic methods (4).

The general prognosis criteria of the International Federation of Gynecology and Obstetrics (FIGO) classification do not include all prognostic factors, including histologic subtypes and lymph node metastasis, which are effective for the prediction of Cerc prognosis (5). Molecular markers and clinical parameters are crucial for the prediction of clinical outcomes and deciding treatment strategies (2,5). In addition, the identification of biomarkers associated with radiotherapy response is of great importance for understanding the molecular mechanisms of Cerc and developing novel strategies.

Radiotherapy significantly benefits patients with Cerc (3). The methylation status in the promoters of a number of genes is associated with patient outcomes after radiotherapy (6-8). For example, Dunn et al (6) demonstrated that the O⁶-methylguanine-DNA-methyltransferase (MGMT) promoter methylation level was positively associated with the progression-free survival and overall survival of patients with glioblastomas treated with temozolomide and radiotherapy (6). Huang et al (7) indicated that the combined Ras association domain family member (RASSF) 1A/RASSF2A methylation level was negatively correlated with the disease-free survival (DFS) of radiotherapy-treated squamous cell carcinoma. Widschwendter et al (9) revealed that the methylated myoblast determination protein 1 (MYOD1) in Cerc was associated with poor DFS (9).

An increasing number of studies have indicated the prognostic power of gene signatures for disease prognosis, metastasis and recurrence. Okayama et al (10) identified a 4 gene signature with prediction power for stage I lung cancer prognosis (10); Cheng et al (11) described an 8-gene classifier
with predictive power for locoregional recurrence of breast cancer in patients post-mastectomy (11). Therefore, the predictive power of multigene signatures for disease development may be of great clinical interest. A 12-gene classifier has been used for the clinical diagnosis of low and high metastasis of in uveal melanoma (12,13). In addition, the DNA methylation level is a significant factor in disease development (6-8). However, to the best of our knowledge, there have been few studies investigating gene methylation signatures for prognosis in radiotherapy-treated patients with CerC.

The present study was designed to explore a novel risk model for predicting outcome of patients with CerC by analyzing RNA sequencing (RNA-seq) data in combination with matched DNA methylation profiles from The Cancer Genome Atlas (TCGA) database. A multigene risk model that predicted the outcomes of patients with CerC treated with or without radiotherapy was identified.

Materials and methods

TCGA and Gene Expression Omnibus (GEO) dataset. Training data were downloaded from TCGA database (https://gdc-portal.nci.nih.gov/) in June 2018. A total of 307 mRNA-seq profiles (Illumina Hiseq2000) and 312 DNA methylation profiles (Illumina Infinium Human Methylation 450 BeadChip) were downloaded. Paired mRNA-seq and methylation data were included in the present study. Clinical features including age, pathologic stage and grade, and survival rate of patients with CerC were extracted and used for subsequent analysis.

Validation dataset GSE44001 (GPL14951 Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip) (14) was downloaded from the National Center of Biotechnology Information GEO database (http://www.ncbi.nlm.nih.gov/geo/). The GSE44001 dataset consists of 300 patients with primary early CerC (FIGO stage I-II). Prognostic data were available for the training and validation sets. The study design presented in Fig. 1.

Identification of differentially expressed and methylated genes. Samples from the TCGA training set were assigned into two groups according to radiotherapy treatment (with and without radiotherapy). Differentially expressed genes (DEGs) and differentially methylated genes (DMGs) between the two groups were identified using Linear Models for Microarray Data (Limma) package (version 3.43.7) in R (https://bioconductor.org/packages/release/bioc/html/limma.html). A false discovery rate (FDR) <0.05 and |log2 fold change (Fc)| >0.263 (>1.2 Fc) were set as the cutoffs. DEGs with differential methylation were selected.

Correlation analysis between gene expression and methylation level. Pearson's correlation between gene expression and the methylation level of DEGs was calculated using CorrTest function (https://stat.ethz.ch/R-manual/R-devel/library/stats/html/corrtest.html) in R. DEGs with significantly correlated expression and methylation level (P<0.05) were included as candidate genes for subsequent analysis.

Selection of prognostic DEGs. Univariate Cox regression analysis in survival package of R (version 2.41.3; https://CRAN.R-project.org/package=survival) (15) was used to screen the DEGs and DMGs associated with the prognosis of patients with CerC. P<0.05, determined by a Kaplan-Meier log-rank test, was defined as the significant cutoff value.

Definition and validation of prognostic risk model. Cox's proportional hazards (Cox-PH) model based on the L1-regularized least absolute shrinkage and selection operator regression algorithm in the penalized package (version 0.9.50) was used for optimizing a prediction model with a linear gene signature (16). The optimized parameter 'lamba' was obtained by 1,000 rounds of cross-validated likelihood (cvl) circular calculation. The risk score of each sample was defined as the linear combination of prognostic gene expression level and Cox-PH regression coefficient: Risk score = ∑ coefgene x Expression (Methylation) gene. Patients were assigned into high-risk and low-risk groups according to the median value of risk score. The overall survival difference between the two groups was evaluated using Kaplan-Meier and log-rank methods in survival package of R (version 2.41.3; https://CRAN.R-project.org/package=survival). The GSE44001 dataset was used to validate the performance and predictive power of the prognostic risk model. The area under the time-independent receiver operating characteristic curve (AUC) was used for evaluation (2).

Selection and stratification analyses of potential clinical prognostic factors. The independent prognostic risk factors among clinical variables in TCGA patients were selected using univariate and multivariate Cox regression analysis in survival package of R (https://CRAN.R-project.org/package=survival). P<0.05, determined by a Kaplan-Meier log-rank test, was set as the significant cutoff value. Stratification analysis was performed for patients with and without radiotherapy, with a significant threshold of P<0.05, as determined by a log-rank test.

Bioinformatic analysis of prognostic DEGs. Patients within the training set were assigned into high-risk (samples with higher risk scores than the median) and low-risk (samples with lower risk scores than the median) groups according to the computed risk scores. DEGs between the two groups (FDR <0.05 and llogFCI >0.263) were identified using Limma package in R. Hierarchical clustering analysis of DEGs was analyzed using the Pheatmap package (version 1.0.8; https://CRAN.R-project.org/package=pheatmap) and Cox's proportional hazards (Cox-PH) model based on the L1-regularized least absolute shrinkage and selection operator regression algorithm in the penalized package. The GSEA (19,20) was performed to identify the Kyoto Encyclopedia of Genes and Genomes (KEGG) (21) pathways significantly (P<0.05) associated with DEGs between the two groups.

Statistical analysis. Continuous clinical variables, including age and overall survival, are presented as the mean ± standard deviation (SD), and differences between groups were analyzed using Student's t-test. Differences in categorical variables, including mortality and pathological characteristics, between two groups were analyzed using Fisher's exact test. Univariate Cox regression analysis was employed for the identification of independent prognostic genes, and a two-step Cox regression analysis was used to identify independent prognostic factors.
among clinical variables. In the stratified analysis, prognostic
differences between the high-risk and low-risk patients stratifi-
cation analysis were analyzed using Kaplan-Meier survival
analysis. All analyses were performed in R (version 3.4.1;
https://www.r-project.org/), and P<0.05 was considered to
indicate a statistically significantly difference.

Results

Baseline characteristics of patients with CerC. A total of
238 patients with CerC with paired mRNA-seq and DNA
methylation profiles from TCGA were used in the present study.
Table I describes the baseline characteristics of the included
238 patients. A total of 64 and 174 patients were assigned
into non-radiotherapy and radiotherapy groups, respectively.
Significant differences in age (P=3.12x10⁻²), pathologic n
stage (P=2.15x10⁻³), pathologic T stage (P=1.90x10⁻⁵), patho-
logic stage (P=9.84x10⁻⁵), new tumor incidence (recurrence;
P=2.20x10⁻¹⁶) and therapy strategy (P=3.16x10⁻¹⁰) were
observed between patients with and without radiotherapy.
There was no difference in overall survival and survival rate
between the two groups (Table I).

Identification of DEGs and DMGs in patients with CerC.
There were 1,488 DEGs and 2,888 DMGs identified between
the two groups (Fig. 2); the majority of the DEGs (62.28%,
1,016/1,488) were upregulated and the majority of the DMGs
(58.14%, 1,679/2,888) were hypomethylated by radiotherapy,
compared with the non-radiotherapy group (Fig. 2A and B).
There were 203 overlapping genes, with an overall negative
correlation between average expression and methylation
levels (Cor=-0.4045; P=2.16x10⁻⁹; Fig. 2C). Pearson's corre-
lation analysis identified 107 genes (including 83 up- and
24 downregulated genes, Table SI) with negatively correlated
expression and methylation levels.

Identification of prognostic genes. Using univariate Cox
regression analysis in the survival package of R, a total of 25
prognostic DEGs and 21 prognostic DMGs were identified
from the aforementioned 107 genes (Table SII), including
8 overlapped genes, which were identified to be candidate
genes associated with the prognosis of patients with CerC
(Fig. 3A). The optimal 8-gene matrix was obtained using
Cox-PH model (max ‘lambda’=5.8598, max cvl=-367.5751;
Fig. 3B). The Cox-PH regression coefficients are indicated

Figure 1. Flow diagram of the study design. KM, Kaplan-Meier; Lasso, least absolute shrinkage and selection operator. PCC, Pearson correlation coefficient
analysis.
XIE et al: AN 8-GENE RISK MODEL FOR CERVICAL CANCER

Table I. Baseline characteristics of The Cancer Genome Atlas patients with cervical cancer treated with or without radiotherapy.

<table>
<thead>
<tr>
<th>Clinical characteristics</th>
<th>Without radiotherapy (N=64)</th>
<th>With radiotherapy (N=174)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years, mean ± SD</td>
<td>45.09±11.49</td>
<td>49.01±14.19</td>
<td>3.12x10⁻²a</td>
</tr>
<tr>
<td>Pathologic M (M0/M1/NA)</td>
<td>29/134</td>
<td>57/9/108</td>
<td>1.65x10⁻¹b</td>
</tr>
<tr>
<td>Pathologic N (N0/N1/NA)</td>
<td>45/7/12</td>
<td>58/35/81</td>
<td>2.15x10⁻³b</td>
</tr>
<tr>
<td>Pathologic T (T1/T2/T3/T4/NA)</td>
<td>47/9/0/4/4</td>
<td>62/46/17/5/44</td>
<td>1.90x10⁻⁵b</td>
</tr>
<tr>
<td>Pathologic stage (I/II/III/IV/NA)</td>
<td>49/8/2/4/1</td>
<td>78/47/29/16/4</td>
<td>9.84x10⁻⁵b</td>
</tr>
<tr>
<td>Pathologic grade (1/2/3/4/NA)</td>
<td>6/27/26/0/5</td>
<td>10/79/66/11/18</td>
<td>6.92x10⁻¹b</td>
</tr>
<tr>
<td>Smoking (reformed/current/never/NA)</td>
<td>7/12/38/7</td>
<td>33/43/86/12</td>
<td>2.06x10⁻²b</td>
</tr>
<tr>
<td>New tumor (yes/no/-)</td>
<td>51/12/1</td>
<td>31/142/1</td>
<td>2.20x10⁻⁶b</td>
</tr>
<tr>
<td>Targeted molecular therapy (yes/no/NA)</td>
<td>5/25/34</td>
<td>127/37/10</td>
<td>3.16x10⁻¹⁰b</td>
</tr>
<tr>
<td>Death (dead/alive)</td>
<td>15/49</td>
<td>51/123</td>
<td>4.17x10⁻¹b</td>
</tr>
<tr>
<td>Overall survival months, mean ± SD</td>
<td>35.11±43.03</td>
<td>38.75±38.92</td>
<td>5.54x10⁻¹a</td>
</tr>
</tbody>
</table>

*Student's t test. *Fisher's exact test. NA, not available; SD, standard deviation.

The mRNA expression value of each gene was associated with the poor survival of patients with hypomethylation of SYT13, demonstrated that the expression and methylation levels of these 8 genes were potential independent risk factors for prognosis in patients with cerv. These data demonstrated that the 8-gene signature had performance and predictive power for outcomes of patients with CerC.

Prognostic value of clinical variables. A two-step Cox regression analyses (univariate and multivariate) were used to define the potential prognostic values of clinical variables, including age, pathologic stage and grade, smoking, radiotherapy, recurrence and risk status, in patients from the TCGA data. Table III demonstrates that 3 independent risk factors, including pathologic stage (hazard ratio (HR)=2.386; 95% confidence interval (CI), 1.097-5.192; P=0.0284), new tumor (recurrence; HR=7.333; 95% CI, 1.833-12.235; P=3.21x10⁻⁹) and risk status (HR=1.359; 95% CI, 1.702-8.905; P=1.28x10⁻⁸) were of prognostic value for the outcomes of patients with CerC. Kaplan-Meier survival analysis determined the prognostic potential of pathologic stage and tumor recurrence. As presented in Fig. 6, there was a significantly shorter overall survival time in patients with advanced (III-IV) pathological stages (P=3.21x10⁻³, log-rank test; Fig. 6A) and recurrence (P=9.22x10⁻¹⁵, log-rank test; Fig. 6B), compared with patients in early stage disease without recurrence.

Establishment and evaluation of the risk model. The mRNA prognostic model based on the combination of Cox-PH regression coefficients and gene expression levels was established as: Risk score=(-0.6241 x Exp_{CCDC136} + (-1.2037) x Exp_{ABCG2} + (-1.0168) x Exp_{CYP26A1} + (-1.1094) x Exp_{CCDC136} + (0.6923) x Exp_{ABC2} + (0.6709) x Exp_{TFMD233} + (0.7099) x Exp_{CCDC136} + (1.5623) x Exp_{TMEM233}). According to the median risk score, 238 patients with CerC in the training set were assigned into high-risk and low-risk groups. Kaplan-Meier survival analysis indicated that patients with low-risk scores exhibited longer overall survival compared with patients with high-risk scores (P=3.22x10⁻⁷, log-rank test; Fig. 5A). The AUC was 0.929 (Fig. 5D). These results demonstrated that the 8-gene signature had performance and predictive power for outcomes of patients with CerC.

Stratification analysis for risk factors associated with radiotherapy. To additionally confirm the risk factors associated with radiotherapy, stratified analysis for patients with radiotherapy and without radiotherapy was performed. A two-step Cox regression analyses indicated that pathologic N stage (HR=4.247; 95% CI, 1.3651-6.216; P=1.25x10⁻²), pathologic stage (HR=2.275; 95% CI, 1.052-3.868; P=4.53x10⁻²), recurrence (HR=3.841; 95% CI, 1.332-10.22; P=2.27x10⁻³) and risk status (HR=5.110; 95% CI, 1.578-6.547; P=6.51x10⁻²) were risk factors for radiotherapy-treated patients, whereas recurrence (HR=4.665; 95% CI, 2.367-9.463; P=1.58x10⁻³) and risk status (HR=7.546; 95% CI, 1.177-8.364; P=3.30x10⁻²) were risk factors for patients treated without radiotherapy (Table IV). These results demonstrated that recurrence and 8-gene signature risk status were independent risk factors for predicting the prognosis of patients with CerC.
Identification of DEGs and KEGG pathways associated with risk status of patients with CerC. To define the gene profiles between patients with high and low risk status, 490 DEGs were identified (Table SIII) in the high-risk group, compared with the low-risk group (Fig. 7A), including 313 upregulated DEGs (63.88%, including CXCL5, SYT13, FOXC2, ITGB3 and TMEM233) and 177 downregulated DEGs (36.18%, including CYP26A1 and TNNI3) in the high-risk group.

Figure 2. DEGs and DMGs between patients with and without radiotherapy. (A) Volcano plot of the DEGs (left) and the kernel density curve plot (right). (B) Volcano plot of the DMGs (left) and the kernel density curve plot (right). Green lines indicate the thresholds of FDR<0.05 (horizontal) and |log₂FC| >0.263 (vertical), respectively. (C) Identification of DEGs and DMGs levels between the two groups (right), and the Pearson's correlation analysis for average expression and methylation levels of the 203 genes. DEGs, differentially expressed genes; DMGs, differentially methylated genes; FDR, false discovery rate; FC, fold change.
Fig. 7B demonstrates the markedly altered expression profiles of these DEGs in patients with low and high risk scores. GSEA KEGG pathway analysis indicated that these genes (including CYP26A1 and CXCL5) were associated with pathways including 'ECM Receptor Interaction', 'Retinol Metabolism', 'Focal Adhesion', 'Hedgehog Signaling Pathway', 'NOD-like Receptor Signaling Pathway' and 'Chemokine Signaling Pathway' (Table V).

**Discussion**

Identification of molecular biomarkers associated with radiotherapy may aid in devising strategies for improving radiotherapy response (22). In the present study, a large-scale analysis of RNA-seq from TCGA Cervical Cancer samples, in combination with matched DNA methylation profiles, was performed, and an 8-gene risk model was identified (CCDC136, ABCG2, CYP26A1, TNNI3, CXCL5, SYT13, FOXC2, and TMEM233) to predict the risk status of patients with Cervical Cancer. Among these 8 genes, 4 hypermethylated genes (CCDC136, ABCG2, CYP26A1 and TNNI3) were positively associated with the overall survival of patients with Cervical Cancer, and 4 hypomethylated genes (SYT13, FOXC2, CXCL5 and TMEM233) were negatively associated with the overall survival of patients with Cervical Cancer.

The 4 hypermethylated genes (CCDC136, ABCG2, CYP26A1 and TNNI3) had previously been identified to be dysregulated in various human cancer tissues (23-26). Among these, TNNI3 is an angiogenesis inhibitor responsible for the inhibition of endothelial cell tube formation (27,28). Kern et al (28) suggested that metastasis was decreased in a mouse model of pancreatic cancer in response to troponin I treatment, compared with control mice. Downregulated troponin I inhibits cancer cell proliferation, as it is required for tumor growth (29). CCDC136, also known as nasopharyngeal carcinoma-associated gene 6, is located at chromosome 7q31-32. It is commonly deleted in a number of types of malignant human cancer, and has been recognized to function as a putative tumor suppressor in gastric and nasopharyngeal carcinoma (23,30).
Wei et al. (30) suggested that CCDC136 negatively regulated the Wnt/β-catenin signaling pathway in zebrafish embryos. Wnt/β-catenin signaling is oncogenic and confers cancer cell proliferation, drug resistance and metastasis in various types of human cancer, including ovarian cancer and cervix (31-34) (Fig. 8). In the present study, it was identified that CCDC136 and TNNI3 were upregulated in the radiotherapy group, compared with the non-radiotherapy group. This may be associated with the decreased angiogenesis and downregulated Wnt/β-catenin signaling in patients treated with radiotherapy, which in turn is associated with the lower recurrence and improved prognosis observed in the radiotherapy group.

ABCG2 encodes a multidrug transporter protein, breast cancer resistance protein (BCRP), which contributes to drug resistance in cancer cell lines and tumors (35,36). It has been reported that ABCG2 is downstream of Wnt/β-catenin signaling and is responsible for chemoresistance (37). Downregulated BCRP/ABCG2 is common in tumor tissues, including cervix, which may function in tumorigenesis by promoting the accumulation of genotoxins and nitric oxide (24,38,39). In addition, ABCG2 promoter methylation has been described in multiple...
Table III. Cox regression analyses for the prognostic value of clinical variables.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P-value</td>
</tr>
<tr>
<td>Age (≤45/&gt;45)</td>
<td>1.014 (0.996‑1.032)</td>
<td>0.138</td>
</tr>
<tr>
<td>Pathologic M (M0/M1)</td>
<td>4.101 (1.356‑12.41)</td>
<td>0.0677</td>
</tr>
<tr>
<td>Pathologic N (N0/N1)</td>
<td>2.923 (1.440‑5.932)</td>
<td>1.86x10^3</td>
</tr>
<tr>
<td>Pathologic T (T1/T2/T3/T4)</td>
<td>1.986 (1.473‑2.678)</td>
<td>2.69x10^6</td>
</tr>
<tr>
<td>Pathologic stage (I/II/III/IV)</td>
<td>1.594 (1.277‑1.989)</td>
<td>2.20x10^5</td>
</tr>
<tr>
<td>Pathologic grade (1/2/3/4)</td>
<td>0.943 (0.612‑1.452)</td>
<td>7.89x10^1</td>
</tr>
<tr>
<td>Smoking (reformed/current/never)</td>
<td>0.984 (0.719‑1.346)</td>
<td>9.18x10^1</td>
</tr>
<tr>
<td>New tumor (yes/no)</td>
<td>5.637 (3.446‑9.22)</td>
<td>9.22x10^15</td>
</tr>
<tr>
<td>Targeted molecular therapy (yes/no)</td>
<td>0.953 (0.547‑1.659)</td>
<td>8.64x10^1</td>
</tr>
<tr>
<td>Risk status (high/low)</td>
<td>3.736 (2.177‑6.411)</td>
<td>3.22x10^7</td>
</tr>
</tbody>
</table>

HR, hazard ratio; CI, confidence interval.

Figure 5. Kaplan-Meier survival analysis and ROC curves for the 8-gene signature in patients with cervical cancer. (A) Kaplan-Meier survival analysis and (B) ROC curve analysis for patients in TCGA training set (n=238). (C) Kaplan-Meier survival analysis and (D) ROC curve analysis for patients in TCGA GSE44001 validation set (n=300) based on the 8-gene signature risk model. ROC, receiver operating characteristic; TCGA, The Cancer Genome Atlas; AUC, area under the curve.
myeloma tissues (40). The demethylation of ABCG2 increases its expression and enhances multidrug resistance in cancer cells (40,41). CYP26A1 is an oncogenic protein in breast cancer, cervical squamous neoplasia, ovarian cancer, and head and neck cancer (25,26,42). CYP26A1 is a metabolizing enzyme for retinoic acids (RAs) (43). RAs induce the differentiation of various types of stem cells (43), and the RA receptor γ (RARγ) is associated with the Akt/NF-κB and Wnt/β-catenin signaling pathways in tumorigenesis (44). Yasuhara et al (45) suggested that RARγ enhances and inhibits Wnt/β-catenin signaling in RA-free and RA-treated conditions, respectively. Demethylation and hypermethylation of CYP26A1 had been
Table V. Gene Set Enrichment Analysis of the KEGG pathways associated with differentially expressed genes between patients with cervical cancer with high- and low-risk scores.

<table>
<thead>
<tr>
<th>KEGG term</th>
<th>ES</th>
<th>NES</th>
<th>NOM P-value</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECM receptor interaction</td>
<td>0.7695</td>
<td>1.2517</td>
<td>1.88x10^-2</td>
<td>LAMA1, COL11A1, ITGB3, IBSP</td>
</tr>
<tr>
<td>Retinol Metabolism</td>
<td>-0.8045</td>
<td>-1.2455</td>
<td>2.05x10^-2</td>
<td>ADH7, CYP26A1, CYP26C1, UGT2A1</td>
</tr>
<tr>
<td>Focal adhesion</td>
<td>0.6963</td>
<td>1.2640</td>
<td>2.14x10^-2</td>
<td>LAMA1, COL11A1, ITGB3</td>
</tr>
<tr>
<td>Hedgehog signaling pathway</td>
<td>-0.7371</td>
<td>-1.2183</td>
<td>2.34x10^-2</td>
<td>WNT3A, BMP7</td>
</tr>
<tr>
<td>NOD-like receptor signaling pathway</td>
<td>0.7413</td>
<td>1.1071</td>
<td>3.98x10^-2</td>
<td>CXCL2, IL6, IL1B</td>
</tr>
<tr>
<td>Chemokine signaling pathway</td>
<td>0.5071</td>
<td>0.9989</td>
<td>4.72x10^-2</td>
<td>CXCL2, CXCL6, ADCY1, CXCL3, CXCL5</td>
</tr>
</tbody>
</table>

KEGG, Kyoto Encyclopedia of Genes and Genomes; ECM, extracellular matrix; ES, enrichment score; NES, normalized enrichment score; NOM, nominal.

Figure 7. DEGs between patients with high- and low-risk scores. (A) Volcano plot differentially expressed genes (Fdr <0.05 and |log2 FC| >0.263). (B) The heat map of the DEGs in The Cancer Genome Atlas patients. DEGs, differentially expressed genes; FDR, false discovery rate; FC, fold change.

Figure 8. Graphical presentation of the radiotherapy-relevant 8-gene signature in cervical cancer. Green and red points indicate downregulated and upregulated genes in the radiotherapy group, respectively. SYT13, synaptotagmin XIII gene; CCDC136/ NAG6, coiled-coil domain containing 136 gene; TNNI3, cardiac troponin I gene; CY26A1, cytochrome P450 26A1 gene; RA, retinoic acid; RARγ, RA receptor γ; ABCG2/BCRP, ATP binding cassette subfamily G member 2 gene; CXCL5, epithelial neutrophil-activating peptide-78 gene; FOXC2, Forkhead 1 gene.
demonstrated in the CYP26A1-positive T47D cell line, which exhibits low rates of metastasis, and the CYP26A1-negative T47D cell line, which exhibits high rates of metastasis, respectively (46). It has been suggested that increased methylation levels in the CYP26A1 promoter is associated with poor survival in patients with prostate cancer (47). In the present study, the expression levels of ABCG2 and CYP26A1 were downregulated and upregulated, respectively, in patients in the radiotherapy group compared with the non-radiotherapy group. These two genes were identified to be positively associated with the prognosis of patients with CerC, and their hypermethylation was correlated with poor survival. In addition, it was also observed that CYP26A1 was associated with the ‘Retinol Metabolism’ GSEA KEGG pathway, which was associated with RA metabolism in cancer cells (48). These results suggested the complex roles of these genes in response to radiotherapy, and their potential prognostic value.

Among the 4 hypomethylated genes (SYT13, FOXC2, CXCL5 and TMEM233), SYT13, FOXC2 and CXCL5 have been demonstrated to be associated with tumorigenesis. FOXC2 is a downstream target of the Akt/NF-κB signaling pathway and is critical for tumor metastasis (49). The inhibition of FOXC2 results in the suppression of tumor metastasis and chemoresistance in lung cancer cells, nasopharyngeal carcinomas and CerC cells (49-51). Synaptotagmins are a family of Ca2+ sensors that function in promoting membrane fusion (52,53). Overexpression of synaptotagmin has previously been described in human cancer (54-57). Kanda et al (58) demonstrated that SYT13 was upregulated in gastric cancer and was associated with metastatic status. CXCL5 is a CXC-type chemokine, and is involved in angiogenesis and associated with poor prognosis in cancer patients (59-62). In addition, CXCL5 expression activated the Akt/NF-κB and Wnt/β-catenin signaling pathways (63-65). FOXC2 and CXCL5 were upregulated in all patients treated with radiotherapy compared with patients without radiotherapy, and their expression was associated with poor survival in patients with CerC. These demonstrated the potential prognostic value of SYT13, FOXC2 and CXCL5 for predicting patients with high risk status or poor outcomes.

In conclusion, a significant difference in survival was observed between the patients with CerC with high- and low-risk scores according to the 8-gene signature. The AUC and survival analysis in the training and validation set revealed the performance and predictive power of the 8-gene signature risk model for predicting survival of patients with CerC. Cox regression analysis indicated that the 8-gene signature was an independent risk factor for the prognosis of patients with CerC. Validation with more and larger clinical cohorts may additionally verify the potential prognostic value of the 8-gene signature in patients with CerC.

Acknowledgements

Not applicable.

Funding

The present study was supported by grants from the Double Tenth Engineering of Major Research Project of Jilin Provincial Science and Technology Department (grant no. 20140201012YY) and the Major Development Programs for New Drugs of the Chinese Academy of Sciences during the 12th Five-Year Plan Period (grant no. 2011ZX09102-001-36).

Availability of data and materials

The results published here are in part based upon data generated by the TCGA Research Network (https://www.cancer.gov/tcga). The Gene Expression Omnibus dataset GSE44001 is available at http://www.ncbi.nlm.nih.gov/geo/. All data generated during this study are included in this published article.

Authors' contributions

FX, DD and GT were responsible for the conception and design of the research. ND, LG, WN, HY, NZ, JJ and GL acquired and analyzed the data. FX and DD drafted the manuscript. LG, WN, HY, NZ, JJ and GL revised important intellectual content. All authors agreed with the final revision.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Cancer Genome Atlas Research Network; Albert Einstein College of Medicine; Analytical Biological Services; Barretos Cancer Hospital; Baylor College of Medicine; Beckman Research Institute of City of Hope; Buck Institute for Research on Aging; Canada’s Michael Smith Genome Sciences Centre; Harvard Medical School; Helen F, et al.: Integrated genomic and molecular characterization of cervical cancer. Nature 543: 378-384, 2017.


47. Liu LY: Association of tissue promoter methylation levels of APc, RASSF1A, CYP26A1 and TBX15 with prostate cancer progression (unpublished PhD thesis). University of Toronto (Canada), 2012.


52. Liu LY: Association of tissue promoter methylation levels of APc, RASSF1A, CYP26A1 and TBX15 with prostate cancer progression (unpublished PhD thesis). University of Toronto (Canada), 2012.


This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.