Identification of anoikis-related genes signature to predict the prognosis in cervical cancer

Shanping Shi  
Ruijin Hospital, Shanghai Jiao Tong University School of Medicine

Xiaojian Tang  
Ruijin Hospital, Shanghai Jiao Tong University School of Medicine

Jiaqian Huang  
Ruijin Hospital, Shanghai Jiao Tong University School of Medicine

Nan Xiang  
Ruijin Hospital, Shanghai Jiao Tong University School of Medicine

Chen Chen  
Ruijin Hospital, Shanghai Jiao Tong University School of Medicine

Weiwei Feng  
Ruijin Hospital, Shanghai Jiao Tong University School of Medicine

Hua Liu (✉ lh11239@rjh.com.cn)  
Ruijin Hospital, Shanghai Jiao Tong University School of Medicine

Research Article

Keywords: cervical cancer, anoikis-related genes, prognosis, immune cells, drug sensitivity

Posted Date: April 18th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2813411/v1

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

Anoikis is a special programmed cell death mode, and resistance to anoikis is a prerequisite for malignant tumors to acquire invasion and metastasis characteristics. The expression and impact of anoikis-related genes (ARGs) in cervical cancer (CC) are still unknown. The aim of this study is to reveal the prognostic role of ARGs in survival, immune infiltration, and drug sensitivity of CC patients, and to identify potential clinical treatment targets. RNA seq and clinical data of CC patients were downloaded from the TCGA database and GEO database, and gene copy data was downloaded from UCSC. Bioinformatics methods was used to screen differentially expressed ARGs related to prognosis, and conducting data analysis using R software package and Perl software. TISCH database was used to analyze the expression of ARGs in tumor microenvironment (TME) at the single cell level. MMP3 on chromosome 11 is highly expressed in CC tissue and may be a key gene for CC progression. The significant activation of the cycline-cycline receptor interaction, ECM-receptor interaction, JAK-STAT signaling pathway, and focal adhesion pathway may be associated with poor prognosis in CC patients. The decrease in CD8+ T cells and the increase in M0 macrophages may indicate a high-risk prognosis for patients. Bcl-2 inhibitor (ABT-737), axitinib, dihydrorotenone, sorafenib, venetoclax, and nilotinib are optional drugs for early treatment of CC. In the future, ARGs based miRNAs, small molecule drugs/inhibitors, peptide/protein specific therapies, and specific antibodies may be developed for early diagnosis and clinical treatment of CC.

1. Introduction

Most deaths caused by solid tumors are due to tumor recurrence and metastasis rather than the primary lesion [1]. Cervical cancer (CC) is one of the most common malignant tumors in gynecology. According to statistics, over 90% of early CC patients can be cured through surgery or chemotherapy, but CC patients with distant metastasis have poor prognosis at both initial diagnosis and recurrence, with a 4-year overall survival (OS) rate of only 5% [2]. Therefore, identifying effective metastasis related prognostic biomarkers is crucial for early intervention and prognosis prediction of CC.

Anoikis is a special form of programmed cell death induced by adhesion and breakage between cells, extracellular matrix, and other cells. It plays an important role in body development, tissue homeostasis, disease occurrence, and tumor metastasis [3]. The triggering of anoikis mainly occurs through the interaction of two apoptosis pathways, namely the intrinsic pathway (mitochondrial events induced by cellular stress) and the extrinsic pathway (mediated by tumor necrosis factor (TNF) and first cell apoptosis signal (Fas) - ligand) [4, 5]. Apoptosis plays a crucial role in tumor invasion and metastasis. Tumor cells survive by resisting cell apoptosis through autocrine and paracrine mechanisms, and regain the ability to adhere to cells for diffusion, metastasis, and invasion [6]. Previous studies have confirmed that the functional VEGFa/VEGFR2 autocrine and paracrine axes in CC cells can induce tumor cell invasion, metastasis, and resistance to anoikis, making them potential therapeutic targets [7]. Yang Li et al. confirmed that the Hippo-YAP1 axis mediated signaling pathway also plays an important role in the occurrence of anoikis resistance and the promotion of CC progression [8].
In addition, tumor microenvironment (TME) refers to the surrounding microenvironment where tumor cells exist, including surrounding blood vessels, immune cells, fibroblasts, adipocytes, signal molecules and extracellular matrix [9]. Tumor cells in TME can directly invade surrounding tissues or undergo metastasis through blood and lymphatic vessels [10]. Infiltrating cells can induce host immune responses by releasing cytokines or cytokine receptors, thereby regulating the progression of cancer. In addition, tumor infiltrating immune cells are an important part of tumor immune microenvironment (TIME). Depending on the type and stage of the tumor, the invasive immune cells that determine the fate of tumor growth can be tumor promoting cells (such as neutrophils and tumor associated macrophages) and anti-tumor cells (such as cytotoxic CD8+ T cells and natural killer cell). In addition, multidrug resistance (MDR) often leads to poor prognosis [11, 12]. Therefore, a better understanding of TME, immune infiltration status, and drug sensitivity in CC is crucial.

At present, studies have established a model of ARGs to effectively predict the prognosis of patients with colorectal cancer, endometrial cancer, lung adenocarcinoma, etc. [13–15]. This study systematically studied the predictive benefits of ARGs in CC, developed a new prognostic risk model based on 6 prognosis-related ARGs, and further explored the differences in tumor microenvironment according to risk scores. In addition, research on immune infiltration, mutation status, and drug sensitivity may provide a basis for the development of drug therapy targets for CC tumors, which has important value in improving the quality of life of CC patients.

2. Materials And Methods

2.1 Data sources

The gene expression data and corresponding clinical information of cervical cancer samples from were downloaded from The Cancer Genome Atlas (TCGA) database(https://portal.gdc.cancer.gov), including 304 cervical cancer samples and 5 normal samples. The microarray dataset related to cervical cancer (GSE30759) and its corresponding clinical information were downloaded from the in National Center for Biotechnology Information-Gene Expression Omnibus (NCBI-GEO) database (https://www.ncbi.nlm.nih.gov/geo/), including 48 cervical cancer samples and 15 normal samples. 513 ARGs were obtained by downloading and merging "anoikis related genes" as the keyword from Genecard database(https://www.genecards.org/) and Harmonic database(https://maayanlab.cloud/Harmonizome/).

2.2 Identification of the Differential ARGs and Prognosis-related ARGs

Bioinformatics analysis was performed using R software, and the ARG differentially expressed between cervical cancer samples and normal samples in the TCGA database was obtained using the "limma" R package with the filter conditions of logFC 1 and p 0.05. Use the "pheatmap" R package to generate thermal and volcanic maps to visualize differentially expressed ARGs. Combine TCGA and GEO cervical
cancer sample data and perform batch correction to extract the expression matrix of the combined
differential ARG. Based on the combination of clinical information from the "survivor" R package and the
"survival" R package, prognosis related ARGs were screened through Univariate Cox analysis.

2.3 Copy number variation and RNA Editing

The gene copy number data of cervical cancer was download from the UCSC
database(http://xena.ucsc.edu/), by analyzing the mutation rate, genetic locus, and copy number
variation (CNV) of ARGs, then using the "RCirocs" R package to draw a visual loop diagram of copy
number variation. RNA editing data for cervical cancer was downloaded from the SYNAPSE database
(https://www.synapse.org/ ), eliminate data with RNA editing deletion values exceeding 30%, and merge
survival data. Conduct Univariate Cox analysis of RNA editing prognosis using coxPfilter < 0.001 as the
filtering condition.

2.4 Cluster analysis

Conduct sample clustering analysis through the "ConsumusClusterPlus" R package, and determine the
optimal grouping based on the tracking plot and cumulative distribution function (CDF) diagram. Survival
analysis was conducted among different subtypes. In addition, PCA, tSNE, and UMAP methods are used
to determine whether the expression of ARG can distinguish between different types. "Cluster Profiler" R
package was used for gene set enrichment analysis (GSEA) enrichment analysis. The enrichment results
were filtered at p < 0.05 to obtain significant enrichment result pathways, in order to explore the potential
functional pathways between different subtypes.

2.5 Establishment and validation of prognostic ARGs
signature

The patients were randomly divided into a training group (n = 176) and a validation group (n = 176), and
the prognosis related ARG was determined by univariate COX analysis. The "glmnet" R software package
was used to further screen prognostic related genes using the Least absolute shrinkage and selection
operator (LASSO) regression algorithm. 10X cross validation is used to select candidate ARGs and
determine penalty parameters(λ), Corresponds to the minimum value of the partial likelihood deviation. In
addition, the median risk score was used as a threshold to divide all eligible patients into high-risk and
low-risk groups. The Kaplan Meier survival analysis illustrates the difference in survival between high-risk
and low-risk patients. The time dependent receiver operating characteristic (ROC) analysis is used to
evaluate model discrimination performance. Draw differential expression heat maps through the
"pheatmap" package. In addition, we conducted risk scoring on the model to determine the correlation
between risk scores and classification. A nomogram based on identified independent variable factors
was constructed using the "rms" R package and evaluated through ROC and calibration curves.
Subsequently, decision curve analysis (DCA) was used to assess the clinical usefulness of nomogram.

2.6 Immune cell infiltration and drug sensitivity analyses
CIBERSORT and Single Sample GSEA (ssGSEA) scripts are used to calculate the degree of immune infiltration and the proportion of immune cells in different samples for different types of immune cells, and obtain immune cell infiltration results. Spearman rank correlation analysis was used to explore the relationship between risk scores and immune infiltrating cells. Extract drug expression and drug sensitivity from the GDSC2 database, and use the "pRRophic" R package to explore the potential sensitivity of high-risk and low-risk groups to clinical drugs.

2.7 Single cell data analysis

Through the TISCH database (http://tisch.compGenomics.org) selects a GEO dataset (GSE168652) to cluster and annotate cells to determine cell types. By observing the scatter plots of model genes in cells, we can determine the differences of target genes in different cells.

3. Results

3.1 Identification of prognostic related ARGs

After obtaining 513 CC related ARGs, we screened 161 significantly differentially expressed genes, including 76 downregulated genes and 85 upregulated genes (Fig. 1A). The forest map shows 28 ARGs as significant single factor genes (p < 0.01, Fig. 1B), divided into two groups based on HR values: low-risk genes and high-risk genes. Among them, TUBB3, SCRIB, EEF1A1, TLN1 may be high-risk genes associated with poor prognosis. At the same time, the network diagram shows more clearly the risk types of ARGs, and the relationship between them is mainly positive regulation (Fig. 1C).

Due to the frequent loss or increase of chromosomal regions in CC patients, we downloaded copy number variation (CNV) data from the TCGA database to further explore how 28 genes related to anoikis change on the chromosome and the location of each gene on the chromosome (Fig. 1D and 1E). For example, for COL4A2, MMP3, SCRIB, CSK, ABL1, FASN, SLC2A1, ENDOG, LAMC2, RHOB, ROR1, TLN1, PLAU, EDA2R, the frequency of copy number gain was significantly greater than the frequency of copy number loss. On the contrary, the copy number loss frequency of LATS2, CLIC4, ONECUT1, NRAS, ANXA2, TUBB3, EEF1A1, HTRA1, SESN1, LDHA, BCL2, and CDCP1 was greater than the copy number gain frequency. At the same time, high copy number gain frequencies were observed in SLC2A1, ROR1, LAMC2 on chromosome 1, RHOB on chromosome 2, SCRIB on chromosome 8, TLN1, ABL1, ENDOG on chromosome 9, PLAU on chromosome 10, MMP3 on chromosome 11, COL4A2 on chromosome 13, CSK on chromosome 15, FASN on chromosome 17, and EDA2R on chromosome X. High copy number loss frequencies were observed in CLIC4 and NRAS on chromosome 1, CDCP1 on chromosome 3, EEF1A1 and SESN1 on chromosome 6, HTRA1 on chromosome 10, LDHA on chromosome 11, LATS2 on chromosome 13, ANXA2 and ONECUT1 on chromosome 15, TUBB3 on chromosome 16, and BCL2 on chromosome 18. The Manhattan map shows RNA editing with differences of pvalue < 0.001 on chromosomes 1, 9, 10, 13, 19, and X (Fig. 1F).

3.2 Consistent clustering of cervical cancer cell molecular subpopulations using 28 ARGs
In order to better understand the role of ARGs in CC, we conducted consensus clustering on 28 prognostic related genes. As shown in Fig. 2A, when k = 2, the queue is identified as two subtypes of clustering. Overall survival (OS) analysis showed a statistically significant difference in prognosis between the two subtypes (p = 0.017, Fig. 2B). Through PCA analysis, it was found that samples of type A and B could be well distinguished, and the accuracy of this clustering was tested using UMAP and tSNE (Fig. 2C). The expression heatmaps and corresponding clinical pathological features of two subtypes of ARGs showed that the majority of ARGs were highly expressed in type A with poor prognosis, including *MMP3*, *PLAU*, and *LATS2*, which may be unfavorable factors for patient prognosis (Fig. 2D). In addition to exploring the overall distribution of 28 ARGs in the cluster, considering the more significant differences between type A and type B, we applied GSVA software to identify the differential enrichment of KEGG pathways between subtypes (Fig. 2E, Supplement Table 1). The activated ECM-receptor interaction in type A is a key signaling pathway involved in the metastasis and progression of various cancers. Meanwhile, activated focal adhesions play a crucial role in tumor invasiveness and metastasis. In order to investigate the Rank of FoldChange of genes in the entire expression profile, we conducted GSEA enrichment analysis on samples from type A and B and found that the cytokine-receptor interaction pathway, ECM-receptor interaction, JAK-STAT signaling pathway and focal adhesion pathway were active in type A and silent in type B (Fig. 2F).

### Table 1

<table>
<thead>
<tr>
<th>Anoikis-related gene</th>
<th>Coefficient</th>
<th>Hazard ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP3</td>
<td>0.250644648582275</td>
<td>1.178509</td>
</tr>
<tr>
<td>ROR1</td>
<td>0.488002591895893</td>
<td>1.377038</td>
</tr>
<tr>
<td>CSK</td>
<td>-0.722718828762758</td>
<td>0.558379</td>
</tr>
<tr>
<td>ENDOG</td>
<td>1.0275975712907</td>
<td>1.43313</td>
</tr>
<tr>
<td>EDA2R</td>
<td>0.509863187556805</td>
<td>1.536714</td>
</tr>
<tr>
<td>BCL2</td>
<td>-0.450315884136612</td>
<td>0.680152</td>
</tr>
</tbody>
</table>

### 3.3 Gene expression and immune infiltration in subtypes

Merge and intersect the samples with common typing results and gene expression data, and extract genes with differential expression in A and B typing (p < 0.05), ultimately obtaining the expression of 21 significantly different genes in the subgroup. It can be seen that only *IGF1* and *BCL2* are highly expressed in type B, while other ARGs are significantly overexpressed in type A and show significant differences in expression between them (Fig. 3A). The above differentially expressed genes are associated with OS, and therefore may be key molecules affecting the prognosis of CC patients and potential targets for targeted therapy. In addition, there were significant differences in the infiltration level of immune cells. ssGSEA analysis showed a significant difference in the activation ratio of immune cells, including activated
dendritic cells, natural killer T cells, neutrotrophic cells, and type-2 T helper cells, in the expression levels of A and B subtype samples (p < 0.001, Fig. 3B).

3.4 Construction and validation of anoikis-related risk model for cervical cancer

In order to establish characteristics for predicting the survival rate of CC patients, 352 CC patients were randomly divided into a training set and a validation set. In the training set, 6 characteristic ARGs were identified through LASSO regression, cross validation analysis, and multivariate COX regression model, namely MMP3, ROR1, CSK, ENDOG, EDA2R, and BCL2 (Fig. 4A and 4B). At the same time, risk scoring formulas for 6 characteristic ARGs were developed in the training set: risk score = \( \sum_{i=1}^{n} \text{coefficients} \times \text{expression of ARGs (i)} \) (Table 1).

According to the median risk score, patients undergoing training and validation are divided into high-risk and low-risk groups. Under this model, the area under the ROC curve (AUC) of both high and low risk groups is greater than 0.7. The ROC curve constructed by the model exhibits good predictive performance over 1 year, 3 years, and 5 years (Fig. 4C). The Kaplan Meier curve indicates that the high-risk group has a poor prognosis, which was also observed in the validation group (Fig. 4D and 4E). It was observed through the difference in risk scores that there was a significant difference in patient risk scores between types A and B (p < 0.05), with type A having higher risk values and lower survival rates (Fig. 4F and 4G).

3.5 Establishment of a prognostic nomogram for cervical cancer

Considering the impact of clinical factors on cervical cancer models, we combined risk scoring in each sample with clinical information to construct a nomogram. For 50 years old patients in the G3-G4 stage, it is predicted that patients with survival greater than one year will be 0.893%, survival greater than three years will be 0.56%, and survival greater than five years will be 0.445% (Fig. 5A). The calibration curve was used to verify the survival data validation, and it was more accurate to predict the one-year, three-year and five-year survival of patients (Fig. 5B). The cumulative hazard curve showed that the OS risk of CC patients in the column chart gradually increases in a time-dependent manner, and the OS risk of the high-risk group is higher than that of the low-risk group (Fig. 5C). The Decision Curve Analysis (DCA) curve of the TCGA queue indicated that nomogram was a good predictor of patient survival, indicating that the DCA model could guide clinical application (Fig. 5D). All three sets of data indicate that using column charts to predict patient survival is far superior to other clinical traits. In summary, based on risk score column charts can be an effective method for predicting the prognosis of CC patients in clinical practice.

3.6 Risk related gene enrichment analysis and tumor microenvironment
TME plays an important role in tumor development and immunotherapeutic response. Therefore, we further explored the TME status of patients in the high-risk and low-risk groups. Extract samples from the high and low risk groups to obtain the immune cell content of the high and low risk groups. Sort the risk scores from low to high, and display the proportion of different immune cells corresponding to the risk scores (Fig. 6A). As the risk score increased, the expression of M0 macrophages also increased ($r = 0.29$, $p < 0.001$, Fig. 6B). The decrease in CD8+ T cells levels and the increase in M0 macrophage levels may indicate poor prognosis in CC patients ($p < 0.001$, Fig. 6C). The immunocyte related thermogram of CC patients more clearly provides the functional relationship between immune cells in the tumor microenvironment, and better explains the composition of the immune microenvironment. There is significant negative regulation between CD8+ T cells and M0 macrophages (Fig. 6D). The overexpression of ENDOG, MMP3, ROR1, and EDA2 in the high-risk group can be observed through the risk heatmap (Fig. 6E). In addition, these 6 ARGs are closely related to multiple immune cell infiltration (Fig. 6F), and the increase of M0 macrophages and high expression of MMP3 and ROR1 may indicate poor prognosis.

Finally, the scores of different drug sensitivity in cervical cancer were used to predict the potential sensitivity of high-risk and low-risk patients to clinical drugs (Fig. 6G, Supplement Table 2). The low-risk group showed high sensitivity to Bcl-2 inhibitors (ABT-737), axitinib, dihydrorotenone, sorafenib, venetoclax, and nilotinib, It is likely to become a potential therapeutic drug.

3.7 Expression of four ARGs in TME based on a single-cell dataset

We analyzed the distribution of 6 ARGs in different immune cells using the single cell dataset CSEC-GSE168652 from the TISCH database and displayed the distribution and quantity of various cell types (Fig. 7A). MMP3 is highly expressed in fibroblasts and malignant cells, but is less expressed in other cells. BCL2 is mainly highly expressed in CD8Tex, fibroblasts, and smooth muscle cells. CSK is mainly highly expressed in CD8Tex, endothelial cells, mono/micro cells, and smooth muscle cells. EDA2R is almost not expressed in malignant cells and mono/micro cells, but is highly expressed in fibroblasts. ENDOG is highly expressed in malignant cells, CD8Tex, and smooth muscle cells. ROR1 is highly expressed in smooth muscle cells (Fig. 7B). It can be seen from the pie chart that the number and proportion of malignant cells are the highest (Fig. 7C).

4. Discussion

Although great progress has been made in the diagnosis and treatment of CC, the lack of effective prognostic biomarkers makes it difficult to diagnose it in the early stage, and it is expected that the incidence rate will increase in the future. Meanwhile, the prognosis of CC is closely related to the degree of infiltration and metastasis. Therefore, using genes related to metastasis to construct prognostic markers may provide important tools for early intervention. Due to the insufficient number of these biomarkers, there is an urgent need to screen more biomarkers with high predictive ability to be included in the candidate list.
Anoikis is a specific form of cell death that can regulate the biological behavior of various tumors [14]. Anchored growth and epithelial mesenchymal transition (EMT) are two characteristics related to resistance to anoikis, and are also important steps in cancer progression and metastasis. The ability of cancer cells to resist anoikis has now attracted major attention in the scientific community [16]. Disorder of anoikis progression is considered a hallmark of cancer cells, which contributes to tumor invasion, migration, and drug resistance. Previous studies have shown that androgen dependent prostate cancer cells show invasion and metastasis due to the damaged expression of ARGs and their interaction with the TME [17]. Due to the rapid progression of the disease, it is difficult for patients with invasive CC to improve their prognosis in a timely manner through a single targeted pathway or drug therapy. Therefore, it is necessary to screen more transfer related genes as predictive markers. However, there is limited research on the impact of ARGs on the invasiveness, immunotherapy, and drug resistance of CC cells, as well as predicting the prognosis and risk assessment of CC patients.

In this study, we screened differentially expressed ARGs and combined them with clinical survival information to classify high and low risk genes related to prognosis, and observed the frequency and location of risk gene mutations. We found that most significant ARGs are high-risk genes associated with poor prognosis, including \textit{TUBB3, SCRIB, EEF1A1, TLN1}. As a tubulin, \textit{TUBB3} overexpression has been confirmed to be related to the resistance to microtubule targeted anticancer drugs such as taxanes [18]. Fibroblasts with low \textit{SCRIB} expression promote lung cancer cell invasion, which is associated with advanced tumor stage and low survival rate [19]. After binding with tRNA, \textit{EEF1A1} promotes the proliferation and invasion of gastric cancer by inhibiting the downstream molecular pathway of P53, while also inhibiting the apoptosis of gastric cancer cells [20]. These studies indicate that these genes have great significance in combating tumor resistance and invasiveness.

We divided the cervical cancer sample cohort into two subtypes based on the expression of ARGs. We observed that the prognosis of type B was significantly better than that of type A. In order to identify the potential biological pathways of ARGs, through GSEA analysis, we found that the signaling pathways involved in cervical cancer progression in type A were significantly activated, such as ECM-receiver interaction, focal adhesion, JAK-STAT signaling pathway, pathways in cancer, and other pathways. Cytokines play a significant role in immune regulation, which is crucial for human biology and diseases. Based on cytokines, new directions can be opened up as immunotherapeutic drugs [21]. The overview of extracellular matrix receptor interactions is very useful in cancer signaling pathways, and the number of passenger mutations far exceeds that of driver mutations, so they may play an important role in cancer immunotherapy [22]. The most important molecule in focal adhesion is adhesion kinase (FAK), which can phosphorylate p-FAK. Overexpression of FAK is a potential target for anti CC drugs to reduce rapid cell proliferation and invasion by inducing cell apoptosis [22]. Similarly, abnormal JAK-STAT signaling contributes to cancer progression and metastasis. STAT protein plays a crucial role in the development of CC, and inhibition of the JAK-STAT pathway may be crucial for enhancing tumor cell death [23]. In addition, most ARGs are highly expressed in type A, with a significantly higher proportion of natural killer cells, mast cells, neutrophils, and T helper cells in type A compared to type B. These genes, pathways, and immune cells may have important implications for improving the efficacy of targeted therapy.
Afterwards, we further constructed a model containing 6 prognostic related ARGs through LASSO regression analysis and multivariate COX regression analysis, including *MMP3*, *ROR1*, *CSK*, *ENDOG*, *EDA2R*, and *BCL2*. *ENDOG*, *MMP3*, *ROR1*, and *EDA2R* were expressed in the high-risk group. These 6 genes have been proven to be closely related to tumors, and high expression of *ENDOG* is extremely detrimental to the prognosis of tumor patients. Many studies have shown that silencing the expression of *ENDOG* can effectively inhibit cell growth in endometrial cancer, thyroid cancer, and glioblastoma, and *ENDOG* deficiency reduces the proliferation of endometrial tumor cells expressing low PTEN/high p-AKT levels. *ROR1* is expressed on the surface of tumor cells and is a promising immunotherapy target in many epithelial tumors [24]. Previous studies have shown that the silencing of MMP3 inhibits the infiltration and metastasis of cervical cancer cells, and its activation has a negative impact on prognosis [25]. In order to predict the risk groups and clinical characteristics of 1-year, 3-year, and 5-year survival rates, we plotted a column chart to add patient grade, age, and risk score to obtain the total score. Through calibration curves, it can be seen that the column chart has the highest accuracy in predicting the 1-year survival of patients. To evaluate the consistency between the predicted OS of the prognostic model and the actual total survival, a calibration chart was established. The results indicate that nomogram prediction is accurate. Although our riskscore and the nomogram constructed based on it have better predictive performance, considering the heterogeneity between cells, anoikis studies conducted at the single-cell level may more accurately reflect the impact of ARGs on the progression and prognosis of cervical cancer patients. *MMP3*, *ENDOG* and *EDA2R* are mainly expressed in malignant cells.

Immunotherapy is an emerging treatment option for CC patients [26]. Due to insufficient research on immune infiltration and drug sensitivity in CC, patients did not achieve significant results after receiving immunotherapy [27]. Therefore, on the basis of predicting risk genes and evaluating clinical information, we also analyzed immune cell infiltration and drug therapy, discussed the possibility of immunotherapy in CC patients, and explored drug resistance. TMIE has an important impact on tumor metastasis process and targeted treatment effect. We analyzed the proportion of 22 immune cell types in the two subtypes. The significant upregulation of M0 macrophage levels in the high-risk group may indicate poor prognosis in cervical cancer patients. There is a significant negative regulation between CD8+ T cells, follicular helper T cells, and M0 macrophages in immune cell correlation, and the infiltration levels of follicular helper T cells and CD8+ T cells are upregulated in the low-risk group. Macrophages generally exist in the microenvironment of solid tumors, which can help immune escape [28]. According to reports, tumor associated macrophages are one of the most abundant immune cell populations in the stroma of cervical tumors. Their carcinogenic function is mainly attributed to their ability to promote immune evasion and metastasis [29]. *MMP3*, *ROR1*, and risk scoring have a significant positive regulatory relationship with M0 macrophages, which is highly likely to activate M0 macrophage proliferation and increase immune escape.

When investigating the treatment results of risk scores of cervical cancer patients, we found that there was an interaction between drug sensitivity and risk scores: Bcl-2 inhibitors (ABT-737), axitinib, dihydrorotenone, sorafenib, venetoclax and other drugs were more sensitive in high-risk groups. Among them, Bcl-2 inhibitors lack clinical efficacy as a single drug, but the combination of ABT-737 and ErPC3
shows synergistic anti proliferative, anti migration, and apoptotic effects in PC-3 cells [30]. Other studies have shown that axitinib treatment is related to the improvement of survival rate of highly pretreated patients with head and neck cancer. After axitinib, the use of single drug immunocheckpoint inhibitors can observe a significant response rate (remission rate, 45%) [31]. In addition, dihydrorotenone (DHR) induces apoptosis of human plasma cell by inducing endoplasmic reticulum stress, and DHR will destroy the process of tumor cell cycle [32]. Sorafenib is a multi kinase inhibitor, which can promote cell apoptosis, reduce angiogenesis and inhibit tumor cell proliferation [33]. Venetoclax is approved for use in chronic lymphocytic leukemia, and the BCL-3 anti apoptotic protein is the main determinant of Venetoclax resistance [34]. Future research should focus on improving the performance of these drugs or combining them with other therapies for treating CC.

5. Conclusion

In summary, we have identified novel ARGs in cervical cancer that can predict survival outcomes and evaluate the immunotherapy response of cervical cancer patients. This study provides a new perspective on treatment strategies for CC patients. The systematic evaluation of risk scoring can expand our understanding of invasion and help develop more personalized and precise treatment strategies.

Declarations

Author Contributions: Shanping Shi, Hua Liu and Xiaojian Tang initiated the study and designed the experiments. Shanping Shi, Chen Chen and Hua Liu performed data collection and analysis. Xiaojian Tang and Nan Xiang helped with discussion and interpretation of results. Jiaqian Huang and Weiwei Feng wrote the manuscript.

Acknowledgments: Not applicable.

Conflict of Interest: The authors declare no conflict of interest.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

References


9. Arneth B. Tumor Microenvironment. Medicina (Kaunas). 2019; 56:


Figures
Figure 1

Identification of ARGs in CC (A) Heatmap and volcano map for differential analysis between CC tissue and normal tissue. Red indicates an increase, while green/blue indicates a decrease. (B) The forest map shows 28 risk ARGs genes (p<0.05). (C) The prognostic network graph shows the correlation and risk attributes between ARGs. (D) The frequency of copy number variations (CNVs) in ARGs. (E) The
chromosomal location and CNV alteration pattern of ARGs exist. (F) The Manhattan map shows whole genome association (GWAS).

Figure 2

Analysis of CC subtypes related to ARGs. (A) Using consensus clustering to obtain the best consensus matrix with k=2. (B) The difference in OS between two subtype patients (p=0.017). (C) PCA, tSNE, and
UMAP distinguished two subtypes based on the expression of ARGs. (D) Gene expression heatmaps and corresponding clinical pathological characteristics in two subtypes. (E) GSVA analysis enriched differences in the KEGG pathway between subtypes A and B. (F) GSVA analysis showed silenced or activated KEGG pathways in subtypes A and B.

**Figure 3**
Gene expression and immune infiltration patterns in two subtypes. (A) Differences in the expression of ARGs between the two subtypes. (B) The infiltration of immune cells in both subtypes.

Figure 4

Identification of prognostic features related to ARGs. (A) LASSO regression analysis and cross validation identified coefficient profiles of 8 characteristic ARGs (B) and 8 prognostic related genes. (C) The ROC
curves for 1-year, 3-year, and 5-year overall survival, with the abscissa representing the false positive rate (1-specificity) and the ordinate representing the true positive rate (sensitivity). (D) The Kaplan Meier curve shows the different prognoses of the high-risk and low-risk groups during training, validation, and the entire group. (E) Differences in clinical risk scores between the two subtypes. (F) The Sandwich plot shows the correlation between high and low risk and survival status in two subgroups of patients.

**Figure 5**

Nomogram of CC patients. (A) Nomogram based on risk score and clinical pathological features. (B) Calibration chart for column chart validation. (C) The cumulative danger curve represents the probability of survival over time. (D) The nomogram decision curve analysis (DCA) curve of CC patients with 1-year, 3-year, and 5-year OS.
Figure 6

Tumor immune microenvironment. (A) The relative proportion of immune cells in different risk scoring groups. (B) Correlation analysis between risk score and the proportion of M0 macrophages in CC tissue. (C) Differences in immune cell levels between high and low risk groups. (D) The correlation between immune cells. (E) The heat map shows the expression patterns of six ARGs in the high and low risk
groups. (F) The correlation between immune cells and six ARGs. (G) Box plot of drug sensitivity differences between high and low risk groups (p<0.05).

Figure 7

Expression of ARGs in TME related cells. (A) Annotations for all cell types in the GSE168652 dataset. (B) The location and expression of six ARGs in cells. (C) The proportion of each type of cell.

Supplementary Files

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

- SupplementTable1.xlsx
- SupplementTable2.xlsx