Cytl1 is a Key Target for Chemoradiotherapy Resistance in Locally Advanced Cervical Cancer: Based on Bioinformatics Analysis

Dong meilian
Zhengzhou University First Affiliated Hospital
Fengsen Liu
Zhengzhou University First Affiliated Hospital
Xiaobin Gu
Zhengzhou University First Affiliated Hospital
Yin Mi
Zhengzhou University First Affiliated Hospital
Yonggang Shi (fccshiyg@zzu.edu.cn)
Zhengzhou University First Affiliated Hospital

Research

Keywords: Uterine cervical neoplasms, risk score, Prognosis, cytl1

Posted Date: October 18th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-960348/v1

Cytl1 is a key target for chemoradiotherapy resistance in locally advanced cervical cancer: based on bioinformatics analysis

Meilian Dong[#1], Fengsen Liu[#2], Xiaobin Gu[1], Yin Mi[1], Yonggang Shi[1]*

[1]Department of Radiation Oncology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China

[2]Biotherapy Center, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China

[#]Contributed equally

*Corresponding author:

Yonggang Shi, Department of Radiation Oncology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China

Tel: 86-0371-67966841

E-mail: fccshiyg@zzu.edu.cn

Abstract

Background: Locally advanced cervical cancer (LACC), FIGO stage equal or more than IB1 is treated with chemotherapy and external beam radiotherapy followed by brachytherapy. However, treatment failure and distant metastasis still remain a major problem. Bioinformatics analysis was used to comprehensively evaluate the intrinsic resistance to treatment in LACC. We selected cytokine-like protein 1 (Cytl1), a novel cytokine, for validation by immunochemistry assays (IHC).

Materials and methods: The Gene Expression Omnibus (GEO) and the cancer genome atlas (TCGA) were analyzed to identify significant prognostic differentially expressed genes (DEGs). The CIBERSORT and ESTIMATE algorithms was performed to validate the genetic risk score model. A total of 113 tissues were evaluated by IHC. Kaplan-Meier curves and the log-rank test were applied for survival analysis.

Results: One thousand and twenty-one genes, consisting of 663 upregulated and 358 downregulated genes, were screened out in the GSE56363 cohort. KEGG and GO pathway enrichment methods was used to analyze the abundant DEGs to find the common characteristics. A genetic risk score model which formed by fifteen genes was established from GSE56363 and validated in the TCGA discovery set. We performed the CIBERSORT and ESTIMATE algorithms to find the difference of immune related signature between the high- and low-risk group to quantify the prognostic value of the fifteen-gene risk model. Among the 22 immune cell types, CD8 T cells, resting CD4 T cells, resting mast cells, activated mast cells, monocytes, neutrophil granulocyte were significantly different between high- and low-risk group. Immune-checkpoints (LAG3 and PDCD1) were highly expressed, and associated with poor prognosis in LACC. IHC demonstrated that cytl1

overexpression was positively associated with poor overall survival (P = 0.01). Cytl1 expression was closely correlated with $CD8^+$T cell low-expression in LACC (r = 0.889, P = 0.066).

Conclusion: This study identified a novel fifteen-gene signature biomarker for predicting the treatment effect of LACC. Moreover, the current results demonstrated that cytl1was overexpressed and associated with poor prognosis in LACC. Cytl1 may be a potential prognostic marker and novel therapeutic target for LACC.

Keywords: Uterine cervical neoplasms, risk score, Prognosis, cytl1

Introduction

Cervical cancer is the fourth leading cause of cancer-related death in women worldwide[1, 2]. According to GLOBOCAN 2018 cancer statistics, there were estimated 569,847 new cases and 311,365 deaths from cervical cancer worldwide in 2018[2]. For early-stage (IA to IIA1) cervical cancer, radical surgery or radiotherapy is recommended treatment method[3]. In contrast to patients with early-stage cervical cancer who have access to two choices, concurrent chemoradiotherapy (CCRT) is efficient for improving prognosis in patients who diagnosed at advanced stage[4]. However, the treatment failure still remains a major problem in local advanced cervical cancer (LACC) due to intrinsic resistance. Therefore, there is an urgent need to identify new prognostic factors and models that could distinguish patients which were unfavorable prognoses LACC. The diagnosis of baseline in individual patients who with unfavorable prognoses could improve the cancer treatment by the avoidance of inefficient therapy and optimize the therapy schedule.

High-throughput platforms for analyzing genetic changes during tumorigenesis, such as microarrays, gained more and more attentions in medical oncology research[5, 6]. The secondary analysis of these high-throughput data could comprehensively identify hundreds of differentially expressed genes (DEGs) involved in the development and progression of various cancers[7-9]. Therefore, this study uses bioinformatics analysis[10] to study the mechanism of LACC during CCRT. Recently, there were few studies focus on the mechanism of radiotherapy resistance in LACC during CCRT. In previous studies, one study did not distinguish the pathological type[11], or others focused on hypoxia[12, 13]. Therefore, in this study, only cervical squamous cell carcinoma was included for research,

and all differentially expressed genes (DEGs) were analyzed. A risk predictive model was constructed and independently validated in a testing cohort. And the predictive power of the genetic risk score model was validated from multiple aspects. Interestingly, in the present study, we firstly identified one DEG, namely CYTL1 in cervical cancer. Through retrospective analysis and IHC assay, we found that CYTL1 was a prognostic factor of LACC and correlated with immune cell infiltration during CCRT. Our results indicated that the expression-based gene signature for predicting survival in LACC patients, which is of great importance for making clinical decisions with regard to the optimal treatment regimen. Maybe, CYTL1 is a potential therapeutic target in cervical cancer.

Materials and methods

Patients and Data sources

Keywords “uterine cervical neoplasm”, “gene expression” and “chemoradiotherapy” were used for searching. One mRNA expression dataset, with the accession number of GSE56363[14]. satisfying our criteria, was downloaded from the GEO database, including 21 patients, who was defined as complete response (CR) or non-complete response (NCR) (partial response and stable disease) by clinically evaluated the tumor response at 6 months after the end of the CRT treatment, and used for discovery cohort. Gene expression data and related clinical information of LACC patients in the TCGA project were obtained from the CBioPortal (http://www.cbioportal. org), which was used for validation cohort in order to validate the prognostic potential of the genetic risk score.

Paraffin specimens of cervical tissue were obtained from biopsy or surgically removed

cervical tissues of inpatients in the first affiliated hospital of Zhengzhou university. A total of 139 paraffin specimens of cervical tissues (normal cervical tissue of 26 patients and local advanced cervical squamous carcinoma of 113 patients) of hospitalized patients from 2013 to 2015 were selected to detect the expression level and prognostic value of cytl1. Pathological sections were diagnosed by pathologists and composed by the following three groups: cervical cancer patients with 6-month complete response (CR group, n=81), non-complete response (NCR group, n=32), normal cervical tissue (control group, n=26). The study was approved by the Ethics Committee of the first affiliated hospital of Zhengzhou university.

Identification of differential expression genes

To identify DEGs (differential expression genes) that were differentially expressed between CR and NCR, we used R (with limma version 3.34.8 package, Auckland, New Zealand) to apply significance analysis of microarray in the GEO dataset GSE56363.The cut-off criteria were an adjusted a logFC value of ≥ 2and *p*-value of < 0.05 for DEGs in NCR samples compared with CR tissues. We also developed a heat map of DEG expression using HemI1.0.1 software.

The mRNA-Seq data from TCGA was produced using the Illumina HiSeq 2000 platform and processed by the RNAseqV2 pipeline, which used MapSplice for alignment and RSEM for quantification. Information about LACC stage of the dataset was assessed according to the TNM stages specified by the International Federation of Gynecology and Obstetrics (FIGO) Surgical Staging of Cancer of the Cervix Uteri (2009).

GO and KEGG pathway enrichment analysis of DEGs

Bioinformatics resource websites can systematically synthesize biological function annotation information for large-scale genes or proteins and find the most significantly enriched biological annotation. In this study, after the screened differential genes are uploaded to DAVID online software, the corresponding species are determined, and the functional annotation in DAVID software is selected to obtain the biological process, molecular function, cell composition and biological signal pathway analysis (KEGG) in which the differential genes participate respectively. $P<0.05$ was considered the level of statistical significance.

Establishment and verification the Logistics prediction model

Multivariate Cox survival analysis was used to create the prognostic index model for NCR patients of LACC. We create the Cox regression model, and calculated the risk score for each patient based on the individual gene expression levels of the screened genes.

Survival risk value was the sum of the product of every gene expression level and its corresponding regression coefficients, and then regarding the median as zero boundary points, samples were divided into high- and low-risk group. At last, we performed to evaluate the model by leave-one-out cross-validation (LOOCV). Time-dependent Area under Curve (AUC) regularly was used for the prediction model of some survival data or the time factor needs to be added. The predictive performance of the 15-gene signature was further validated in the testing data set and entire data set. In addition, according to Cox multivariate hazard analyses, predictive power of the genetic risk score model as same as the other clinical information in predicting the prognosis of patients with LACC.

Consensus Clustering for Tumor-Infiltrating Immune Cells and immune checkpoints

We quantified the infiltration levels for distinct immune cells and differential expression of immune checkpoints in LACC by using the two computational algorithms[15], CIBERSORT and ESTIMATE and employing the LM22 signature. We evaluated the immune and stromal score for each sample in high- and low-risk group by ESTIMATE.

Immunohistochemistry (IHC)

We performed the immunohistochemistry on FFPE 4-μm thick paraffin specimens of cervical tissue sections, using a standard protocol. To facilitate sectioning, we placed the tissue wax chips on a low temperature table to cool for about 30 minutes, and bake chips in a 600°C thermostat for 30 minutes. The chips were soaked in xylene for 10 minutes to dewax, twice, then immerse in 100%, 95%, and 70% gradient alcohol for 3 minutes, and washed with PBS for 3×5 minutes. We processed the slices using the boiling process for 5 minutes, cooled naturally, and rinsed with PBS 3 times, 5 minutes each time. After cleaning, 3% H_2O_2 blocking solution dropwise was added and incubated for 20 minutes. Then we washed the chips with PBS 3 times, 5 minutes each time. 50μL of the primary antibody (1:100) rabbit anti-human monoclonal antibody dropwise was added to the slices, and incubated for 12h in a 4°C humidifier in the refrigerator. 50μL of HRP-labeled secondary antibody was dropped on the slices, placed in a humidifier, and incubated at 37°C for 30 minutes. The newly prepared DAB color-developing liquid is dropped on the slices, and the time and degree of color-developing are observed and controlled under a microscope. Two pathologists who did not know the clinical research data read the pictures under the

microscope to avoid interference of the immunohistochemistry results. The tissue sections were observed sequentially under microscope, and 3 fields of view were randomly selected for each section and photographed (×200 times, ×400 times). The average value was used as the staining result. The criterion was a combination of staining intensity (Intensity, I) and percentage of positive cells (Percentage, P). The calculation results were described as follow: CYTL1, CD4, CD8 negative (-), CYTL1, CD4, CD8 weakly positive (+), CYTL1, CD4, CD8 moderately positive (++), CYTL1, CD4, CD8 strongly positive (+++).

Statistical analysis

In this study, time event was defined as CR or NCR by clinically evaluated the tumor response at 6 months after the end of the CRT treatment. Kaplan-Meier is mostly used to analyze the relationship between gene expression and time event, the statistical significance was determined with the log-rank test based on the χ^2 distribution, and the survival differences were compared using the log-rank test. Differences were considered statistically significant for $P<0.05$. Survival analysis and the survival curve were performed with R. Univariate and multivariate analyses were performed with the Cox proportional hazards regression model. Student's *t*-test was performed to analyze the differences between groups.

Results

Identification of differential expression genes

The data identified based on the GEO dataset needs to be pre-processed before the difference analysis. There are special cases such as missing values or genes corresponding

to multiple probes. After background correction, standardization and summarization, it can be seen that the processed chip data value distribution effect is better. Further analysis of the independent cohort, 1021 DEGs (663 upregulated genes and 358 downregulated genes) were demonstrated from GSE56363 dataset (Figure 1).

KEGG and GO pathway enrichment analysis for the DEGs

KEGG and GO enrichment methods was used to analyze the abundant DEGs to find the common characteristics. The pathway of these DEGs enriched mainly in focal adhesion and multiple tumors, as shown in Figure 2A. The biological process of gene ontology enriched mainly in cellular component disassembly, extracellular structure organization, organelle fission, and extracellular matrix organization (Figure 2B). The cellular component of gene ontology enriched mainly in adherens junction, extracellular matrix, collagen−containing extracellular matrix, and focal adhesion (Figure 2C). The molecular function of gene ontology also contained extracellular matrix structural constituent (Figure 2D). These results maybe alert us that these DEGs mainly correlated with EMT (endothelial-mesenchymal transition) pathways in cancer.

Construction and Validation of the DEG-based prognostic signature

Firstly, univariate Cox regression analysis of the DEGs was performed individually to find which genes were strongly associated with survival. Thirty-eight DEGs correlated with prognosis, as shown in Table 1. Furthermore, we performed multivariate Cox regression analysis based on the results of univariate analysis. As shown in Table 2, fifteen genes were identified: Fascin-1 (FSCN1), NDC1, TNFRSF18, SH3BP5, GPC1, MT3, FBLN5,

IQSEC2, CYTL1, NDN, CKB, HIC1, KDELR2, RALGDS, SLCO3A1. Then, a risk model was established based on the Cox regression coefficient and the risk score of every patient. The formula was as follows: prognostic index (PI)= 0.006063 × relative expression of FSCN1 + 0.104381 × relative expression of NDC1 + 0.019972 × relative expression of TNFRSF18 + 0.162985 × relative expression of SH3BP5 + 0.009462 × relative expression of GPC1 + 0.038881 × relative expression of MT3 + 0.055523 × relative expression of FBLN5 + 0.136231 × relative expression of IQSEC2 + 0.087325 × relative expression of CYTL1 + 0.083397 × relative expression of NDN + (-0.00657) × relative expression of CKB + (-0.32014) × relative expression of HIC1 + 0.006355 × relative expression of KDELR2+(-0.03246) × relative expression of RALGDS + (-0.05878) × relative expression of SLCO3A1. To verify the risk prediction model, another microarray dataset from the TCGA dataset was acquired, which included detailed gene expression information and survival time of patients with LACC. And, a total of 304 patients were obtained. We randomly divided these patients into training and testing cohort in a ratio of 1:1. Next, we used Kaplan–Meier survival analysis to evaluate the prognostic impact of model on both high- and low-risk group in the test group. We found that the association between risk score and survival time of patients was statistically significant, with the low-risk patients (n=76) showing a substantial advantage in OS time compared with the high-risk patients (n=76) (Figure 3A, D). A cluster analysis of each patient genes according to the risk scores was performed, to explore the true contributions of the fifteen genes. The visual diagram and the heat map were shown in Fig. 3B and 3C. Then, we tested this model in the validation cohort (n=152) and total cohort (n=304). The risk scores of each patient were calculated, and the high-risk scores predicted by our model were significantly worse than those of low-

risk scores (Figure 4A, D, E). The visual diagram and the heat map of validation cohort were shown in Figure 4B and 4C. In addition, according to Cox multivariate hazard analyses, predictive power of the genetic risk score model as same as the other clinical information in predicting the prognosis of patients with LACC, as shown in Figure 5. Compared with other clinicopathological factors, the model is an independent prognostic factor, as shown in Figure 5. The time-dependent AUC curve indicates that the model has a good predictive performance, as shown in Figure 6.

Variance of immune related signature validate the fifteen-gene risk prediction model

The tumor microenvironment (TME) contains immune cells and stromal cellular elements which contain tumor-infiltrating immune cells. The crucial role in therapeutic response of the tumor-infiltrating immune cells has been revealed in many researches. So, we performed the CIBERSORT and ESTIMATE[16] algorithms to find the difference of immune related signature between the high- and low-risk group in the total set to quantify the prognostic value of the fifteen-gene risk model. A heatmap of the relative proportion of the 22 immune cells in each LACC sample were assessed and presented in Figure 7A. Among the 22 immune cell types, CD8 T cells, resting CD4 T cells, resting mast cells, activated mast cells, monocytes, neutrophil granulocyte were significantly different between high- and low-risk group, as shown in figure 7B. Resting CD4 T cells, monocytes, mast cells activated, and neutrophils were positively correlated with the fifteen-gene risk prediction model. CD8 T cells and mast cells resting were negatively correlated with fifteen-gene risk prediction model (Figure 7B).

Immunotherapy with immune-checkpoint inhibitors is a breakthrough in recent years in a

variety of tumor types, including cervical cancer[17, 18]. Overexpression of the immune checkpoints, like anti-PD-1 and anti-PD-L1, has demonstrated a crucial mechanism of immune-surveillance escape. Immune checkpoint inhibitors have clear efficacy in many cancers[19]. For further verification of the prognostic value of the fifteen-gene risk model, we run a univariate Cox regression to analysis the difference expression of immune-checkpoint in high- and low-risk group. According to the results, CD274, TIGIT, LAG3, PDCD1, CTLA4, HAVCR2, PDCD1LG2 were highly expressed. Furthermore, the high expression of LAG3 and PDCD1 was associated with poor prognosis in cervical cancer patients, as shown in Figure 8A and 8B.

Expression Level of CYTL1

IHC showed that positive rate of CYTL1 is 100% (strongly positive, 18.8%; moderately positive, 50%; weakly positive, 32.2%) in the NCR group. Relatively, the expression situation is that 23.5% of moderately positive, 48.1% of weakly positive, and 28.4% of negative in the CR group. The positive rate of CYTL1 in NCR group is significantly higher than CR group (P = 0.01; Figures9A、B、C). Kaplan-Meier survival analysis showed that the prognosis of patients with high CYTL1 expression was poorer than low CYTL1 expression (P = 0.023, Figures 9E).

The correlation between CYTL1 and immune or inflammatory cell infiltration

We used TIMER database to analysis the Spearman's correlations between CYTL1 expression level and tumor-infiltrating cells. The positive correlation between cells included activated CD4+ T cell (Act_CD4; P =0.046), CD8+ T cell (P =1.672e-04),

regulatory cell (Treg; P =2.823e-05), B cell (P =0.009), activated dendritic cells (Act_DC; P =0.035), Resting mast cell (P =0.006), Activated natural killer cells (P =0.035), as shown in Figure 10A-E. We verify the correlation between CYTL1 and CD4+ T cell, CD8+ T cell by IHC. The infiltration of CD8+ T cells is higher in the CR group than NCR group. The correlation between CYTL1 and the infiltration of CD8+ T cells was significant (P = 0.066, Rs = 0.889, Figure 10J、L). The correlation between CYTL1 and the infiltration of CD4+ T cells was not significant (Figure 10I、K).

Discussion

Cervical cancer is a common cancer type in females worldwide, with 85% of cases occurring in developing countries[20]. Moreover, cervical cancer is the fourth leading cause of cancer death in women[2]. Therefore, the research on cervical cancer has important social value and economic benefits. As we all know that, the treatment is based on chemoradiotherapy treatment for LACC. Unfortunately, about half of the patients with LACC will recurrent or metastasize after complete therapy of first 2 years. Overcoming resistance of chemoradiotherapy is still a challenge. If we can distinguish patients with baseline resistance and unfavorable prognoses, inefficient therapy will be deprived, or other potentially active treatments will be regarded. Maybe, the therapeutic effect of cancer will be improved finally. However, how to select patients individually remains controversial in clinical practice. Therefore, it is critical to investigate predictive molecules or related model for extreme resistance of chemoradiotherapy in patients with LACC. Recently, systematically combination of bioinformatics algorithm has been used to find out prediction methods of driver genes and events which responsible for rapid development or

treatment resistance of different cancer types. The findings highlight the ability of computational approaches in these effects[21]. In cervical cancer, relatively study which employed published bioinformatics algorithms was rarely seen, especially in LACC. Herein, we employed bioinformatics to determine risk model to forecast chemoradiotherapy resistance in patients with LACC.

In the present study, 1021 differentially expressed genes, consisting of 663 upregulated and 358 downregulated genes were identified based on microarray analysis. Subsequently, through a comprehensive study of data mining, bioinformatics and subsistence analysis, we identified 15 DEGs that are associated with survival of patients with LACC, namely FSCN1，NDC1, TNFRSF18, SH3BP5, GPC1, MT3, FBLN5, IQSEC2, CYTL1, NDN, CKB, HIC1, KDELR2, RALGDS, and SLCO3A1 in the training set. Functional analysis demonstrated that these DEGs were mainly involved in the focal adhesion, leukocyte transendothelial migration, cytokine-cytokine receptor interaction, pathways in cancer, HTLV-1 infection, chemokine signaling pathway, glutathione metabolism, sphingolipid signaling pathway, platelet activation, proteoglycans in cancer, rap1 signaling pathway. Focal adhesion and cytokine-cytokine receptor interaction were function in invasion and chemoradio-resistance of tumor cells[22]. Fariborz Soroush, et al. indicate that neutrophil-endothelial interaction of murine cells was not a good predictor of their interactions in human cells[23]. Hans-Åke Fabricius[24], et al. reviewed the role of platelet cell for the direct and indirect platelet-tumor cell contact during metastasis formation in human tumors. It is conceivable that the overexpression or downregulation of these DEGs in those pathways can affect the course of treatment-resistant. According to the Cox coefficients derived from K-M curves, we developed a 15-gene based risk score from GSE56363. We also

independently validated the risk model in TCGA. To our knowledge, this is the first study using the bioinformatic approach to construct a prognostic model about CCRT in LACC.

Previous studies using whole exome sequencing analysis have revealed that several recurrent genomic alterations involved in the tumorigenesis of cervical cancer[25]. A recent study identified SHKBP1, ERBB3, CASP8, HLA-A and TGFBR2 as novel significantly mutated genes in cervical cancer[26]. Furthermore, amplification in immune targets including CD274 and PDCD1LG2 were found, which was related to the response to lapatinib[27]. Growing evidence have shown that abnormally expressed genes in cervical cancer, were strongly associated with development of this disease and could be regarded as the potential prognostic factors, such as MAP3K3, FOXO1, RHOB, DIRAS1, RERG, RAP2C, and MEF2C[28]. In the present study, the DEGs are extensively involved in the tumorigenesis. For example, FBLN5 is significantly down-regulated in ovarian carcinoma and acts as a tumor suppressor by inhibiting the migration and invasion of ovarian cancer cells[29]. Hara's study revealed that the overexpression of GPC1 was associated with chemoresistance to cis-Diammineplatinum (II) dichloride (CDDP) and poor prognosis in esophageal squamous cell carcinoma[30]. GPC1 is a promising theragnostic platform for the effective treatment of cervical cancer[31]. Furthermore, the overexpression of SH3BP5 was correlated with advanced tumor stage in diffuse large B-cell lymphoma[32]. Wang's work suggests that RILP suppresses the invasion of breast cancer cells by interacting with RalGDS[33]. Emerging evidences indicate that Fascin-1 (FSCN1) may possess causal roles in the development of several types of cancers and serves as a novel biomarker of aggressiveness in certain carcinomas, for example, triple-negative breast cancer (TNBC), esophageal cancer, small cell lung cancer, hepatocellular carcinoma, and cervical cancer[27,

[34-37]. The DEGs found in this study were in line with previous evidence in other types of cancer.

Meanwhile, the prognostic capability of this signature was explored. The differences were significant between high- and low-risk patients in survival curves of the training, testing, and total sets. Furthermore, the multivariate Cox regression analysis indicated that the 15-gene-based prediction model was a predictive factor for LACC, compared to clinicopathological factors, which included age, stage, grade, N, T. Time-dependent AUC (Area Under roc Curve) indicated that the 15-gene-based prediction model was a firm risk tool.

Many studies explored the safety and efficacy of immunotherapy in metastatic and recurrent cervical cancer. Recently, clinicians paid close attention to immunotherapy in concurrent with standard CCRT or as an adjuvant therapy in LACC gradually. However, the study about optimal timing for immune therapy in relation to CCRT was scarcity. The 15-gene-based prediction model which we made was focused on the tumor-infiltrating immune cells and checkpoint between the high- and low-risk group in LACC. The resting mast cells and CD8 T cells were low in the 15-gene-based prediction model high-risk group and associated with poorer OS. Resting CD4 T cells, monocytes, mast cells activated, and neutrophils were high in the 15-gene-based prediction model high-risk group and associated with poorer OS. The result consistented with other study that CD8+ T-cell infiltration may be a prognostic biomarker for patients with cervical cancer receiving radiotherapy[38] or immune checkpoint inhibitors[16, 38]. Lirong Yang, et al. reported that mast cells were significantly related to OS of patients with cervical cancer[39]. In our study, the high expression of LAG3 and PDCD1 was associated with poor prognosis in cervical

cancer patients. Rui Li[40] et al. conducted a prospective clinical trial to assess the tumor immune microenvironment during CCRT of patients with LACC, and they suggested that immune checkpoint inhibitors were administered before CCRT maybe more effective. These results validated the prognostic value of our 15-gene-based prediction model partially from other side.

Interestingly, in the present study, we firstly identified one DEG, namely CYTL1 in cervical cancer. CYTL1 is a small widely expressed secreted protein lacking significant primary sequence homology to any other known protein[26]. Previous evidence showed that CYTL1 had cytokine-like properties with some overlap and some distinction from the chemokines[26]. The DNA methylation pattern of the CYTL1 promoter region was significantly different between early and advanced stages of SCC[41]. It had been found that abnormal DNA methylation of specific genes could cause various cancers and treatment resistance. In addition, by reviewing the literature, we found that CYTL1 induces the EMT in human coronary artery endothelial cells via activation of the TGF-β-SMAD signaling pathway[42]. It provides a direction for us to study CYTL1-mediated CRT resistance of cervical cancer. So, CYTL1 was selected for further analysis. We tested the expression level of CYTL1 and the association with the progress of the LACC by IHC. High CYTL1 expression indicated poor prognosis in LACC during CCRT. As mentioned before, we wanted to explore about optimal timing for immune therapy in relation to CCRT. Therefore, we detected the correlation between CYTL1 and T cell (CD4+ and CD8+) in the TIMER database. And the study verified the result by IHC. In our study, we demonstrated that CYTL1 was involved in the development of cervical cancer and had prognostic value. Maybe, CYTL1 is a potential therapeutic target in cervical cancer. The detailed role of

CYTL1 needs further investigations.

In conclusion, we identified 15 DEGs which are associated with survival outcomes of cervical cancer patients. The bioinformatics and subsistence analysis of these DEGs provided further insight into the role of DEGs in the development of cervical cancer. In addition, the 15-gene-based model was formulated, which should be helpful for cervical cancer personalized treatment and prognostication. By reviewing the literature, we chose CYTL1 for further analysis. Through retrospective analysis and IHC assay, we found that CYTL1 was a prognostic factor of LACC and correlated with immune cell infiltration during CCRT. Maybe, CYTL1 is a potential therapeutic target in cervical cancer. Actually, current research is far from achieving the ultimate goal of discovering a complete prediction model or factor. The ongoing trials and more prospective clinical trial will hopefully clarify the challenges and efficacy of the prediction factors in clinical practice.

Auther contributions DM and LF take responsibility for the whole work and plan the whole study. DM and GX contribute to draft the manuscript together. LF and YM prepare the tables and the figures. SY contribute to review the related literature.

Finding This work was supported by the Joint Co-construction Project in Medical Science of Henan province (No. LHGJ20200359).

Declaration

All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under

review at any other publication.

Acknowledgements

The authors thank Yongliang Yuan for his support in the graphic realization of the figures.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Not applicable.

References

[1] Kimondo F C, Kajoka H D, Mwantake M R, et al. Knowledge, attitude, and practice of cervical cancer screening among women living with HIV in the Kilimanjaro region, northern Tanzania[J]. Cancer Rep (Hoboken), 2021:e1374.

[2] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin,

2018,68(6):394-424.

[3] Cibula D, Potter R, Planchamp F, et al. The European Society of Gynaecological Oncology/European Society for Radiotherapy and Oncology/European Society of Pathology guidelines for the management of patients with cervical cancer[J]. Radiother Oncol, 2018,127(3):404-416.

[4] Yoo W, Kim S, Huh W K, et al. Recent trends in racial and regional disparities in cervical cancer incidence and mortality in United States[J]. PLoS One, 2017,12(2):e172548.

[5] Zhou L, Du Y, Kong L, et al. Identification of molecular target genes and key pathways in hepatocellular carcinoma by bioinformatics analysis[J]. Onco Targets Ther, 2018,11:1861-1869.

[6] Liu F, Li H, Chang H, et al. Identification of hepatocellular carcinoma-associated hub genes and pathways by integrated microarray analysis[J]. Tumori, 2015,101(2):206-214.

[7] Deng Y, Yuan W, Ren E, et al. A four-methylated LncRNA signature predicts survival of osteosarcoma patients based on machine learning[J]. Genomics, 2021,113(1 Pt 2):785-794.

[8] Liao X, Zhu G, Huang R, et al. Identification of potential prognostic microRNA biomarkers for predicting survival in patients with hepatocellular carcinoma[J]. Cancer Manag Res, 2018,10:787-803.

[9] Zhao K, Li Z, Tian H. Twenty-gene-based prognostic model predicts lung adenocarcinoma survival[J]. Onco Targets Ther, 2018,11:3415-3424.

[10] Lee E, Moon A. Identification of Biomarkers for Breast Cancer Using Databases[J]. J Cancer Prev, 2016,21(4):235-242.

[11] Halle C, Andersen E, Lando M, et al. Hypoxia-induced gene expression in chemoradioresistant cervical cancer revealed by dynamic contrast-enhanced MRI[J]. Cancer Res, 2012,72(20):5285-5295.

[12] Fjeldbo C S, Julin C H, Lando M, et al. Integrative Analysis of DCE-MRI and Gene Expression Profiles in Construction of a Gene Classifier for Assessment of Hypoxia-Related Risk of Chemoradiotherapy Failure in Cervical Cancer[J]. Clin Cancer Res, 2016,22(16):4067-4076.

[13] Jonsson M, Fjeldbo C S, Holm R, et al. Mitochondrial Function of CKS2 Oncoprotein Links Oxidative Phosphorylation with Cell Division in Chemoradioresistant Cervical Cancer[J].

Neoplasia, 2019,21(4):353-362.

[14] Balacescu O, Balacescu L, Tudoran O, et al. Gene expression profiling reveals activation of the FA/BRCA pathway in advanced squamous cervical cancer with intrinsic resistance and therapy failure[J]. BMC Cancer, 2014,14:246.

[15] Zhou R, Zhang J, Zeng D, et al. Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I-III colon cancer[J]. Cancer Immunol Immunother, 2019,68(3):433-442.

[16] Yang S, Wu Y, Deng Y, et al. Identification of a prognostic immune signature for cervical cancer to predict survival and response to immune checkpoint inhibitors[J]. Oncoimmunology, 2019,8(12):e1659094.

[17] Borcoman E, Le Tourneau C. Pembrolizumab in cervical cancer: latest evidence and clinical usefulness[J]. Ther Adv Med Oncol, 2017,9(6):431-439.

[18] Wang J, Li Z, Gao A, et al. The prognostic landscape of tumor-infiltrating immune cells in cervical cancer[J]. Biomed Pharmacother, 2019,120:109444.

[19] Ventriglia J, Paciolla I, Pisano C, et al. Immunotherapy in ovarian, endometrial and cervical cancer: State of the art and future perspectives[J]. Cancer Treat Rev, 2017,59:109-116.

[20] Jemal A, Bray F, Center M M, et al. Global cancer statistics[J]. CA Cancer J Clin, 2011,61(2):69-90.

[21] Wang T, Ruan S, Zhao X, et al. OncoVar: an integrated database and analysis platform for oncogenic driver variants in cancers[J]. Nucleic Acids Res, 2021,49(D1):D1289-D1301.

[22] Gaianigo N, Melisi D, Carbone C. EMT and Treatment Resistance in Pancreatic Cancer[J]. Cancers (Basel), 2017,9(9).

[23] Soroush F, Tang Y, Mustafa O, et al. Neutrophil-endothelial interactions of murine cells is not a good predictor of their interactions in human cells[J]. FASEB J, 2020,34(2):2691-2702.

[24] Fabricius H A, Starzonek S, Lange T. The Role of Platelet Cell Surface P-Selectin for the Direct Platelet-Tumor Cell Contact During Metastasis Formation in Human Tumors[J]. Front Oncol, 2021,11:642761.

[25] Ojesina A I, Lichtenstein L, Freeman S S, et al. Landscape of genomic alterations in cervical carcinomas[J]. Nature, 2014,506(7488):371-375.

[26] Tomczak A, Singh K, Gittis A G, et al. Biochemical and biophysical characterization of cytokine-like protein 1 (CYTL1)[J]. Cytokine, 2017,96:238-246.

[27] Wang C Q, Li Y, Huang B F, et al. EGFR conjunct FSCN1 as a Novel Therapeutic Strategy in Triple-Negative Breast Cancer[J]. Sci Rep, 2017,7(1):15654.

[28] Gao C, Zhou C, Zhuang J, et al. MicroRNA expression in cervical cancer: Novel diagnostic and prognostic biomarkers[J]. J Cell Biochem, 2018,119(8):7080-7090.

[29] Heo J H, Song J Y, Jeong J Y, et al. Fibulin-5 is a tumour suppressor inhibiting cell migration and invasion in ovarian cancer[J]. J Clin Pathol, 2016,69(2):109-116.

[30] Hara H, Takahashi T, Serada S, et al. Overexpression of glypican-1 implicates poor prognosis and their chemoresistance in oesophageal squamous cell carcinoma[J]. Br J Cancer, 2016,115(1):66-75.

[31] Matsuzaki S, Serada S, Hiramatsu K, et al. Anti-glypican-1 antibody-drug conjugate exhibits potent preclinical antitumor activity against glypican-1 positive uterine cervical cancer[J]. Int J Cancer, 2018,142(5):1056-1066.

[32] Kobayashi K, Yamaguchi M, Miyazaki K, et al. Expressions of SH3BP5, LMO3, and SNAP25 in diffuse large B-cell lymphoma cells and their association with clinical features[J]. Cancer Med, 2016,5(8):1802-1809.

[33] Wang Z, Zhou Y, Hu X, et al. RILP suppresses invasion of breast cancer cells by modulating the activity of RalA through interaction with RalGDS[J]. Cell Death Dis, 2015,6:e1923.

[34] Akanuma N, Hoshino I, Akutsu Y, et al. MicroRNA-133a regulates the mRNAs of two invadopodia-related proteins, FSCN1 and MMP14, in esophageal cancer[J]. Br J Cancer, 2014,110(1):189-198.

[35] Koay M H, Crook M, Stewart C J. Fascin expression in cervical normal squamous epithelium, cervical intraepithelial neoplasia, and superficially invasive (stage IA1) squamous carcinoma of the cervix[J]. Pathology, 2014,46(5):433-438.

[36] Liu Y, Hong W, Zhou C, et al. miR-539 inhibits FSCN1 expression and suppresses hepatocellular carcinoma migration and invasion[J]. Oncol Rep, 2017,37(5):2593-2602.

[37] Yang X, Lei P, Huang Y, et al. MicroRNA-133b inhibits the migration and invasion of non small

cell lung cancer cells via targeting FSCN1[J]. Oncol Lett, 2016,12(5):3619-3625.

[38] Mori Y, Sato H, Kumazawa T, et al. Analysis of radiotherapy-induced alteration of CD8(+) T cells and PD-L1 expression in patients with uterine cervical squamous cell carcinoma[J]. Oncol Lett, 2021,21(6):446.

[39] Yang L, Yang Y, Meng M, et al. Identification of prognosis-related genes in the cervical cancer immune microenvironment[J]. Gene, 2021,766:145119.

[40] Li R, Liu Y, Yin R, et al. The Dynamic Alternation of Local and Systemic Tumor Immune Microenvironment During Concurrent Chemoradiotherapy of Cervical Cancer: A Prospective Clinical Trial[J]. Int J Radiat Oncol Biol Phys, 2021.

[41] Kwon Y J, Lee S J, Koh J S, et al. Genome-wide analysis of DNA methylation and the gene expression change in lung cancer[J]. J Thorac Oncol, 2012,7(1):20-33.

[42] Zhu S, Kuek V, Bennett S, et al. Protein Cytl1: its role in chondrogenesis, cartilage homeostasis, and disease[J]. Cell Mol Life Sci, 2019,76(18):3515-3523.

Figure legents

Figure 1: 1021 DEGs (663 upregulated genes and 358 downregulated genes) were demonstrated from GSE56363 dataset.

Figure 2: KEGG and GO pathway enrichment analysis for the DEGs.

The pathway of these DEGs enriched mainly in focal adhesion and multiple tumors, as shown in Figure 2A. The biological process of gene ontology enriched mainly in cellular component disassembly, extracellular structure organization, organelle fission, and extracellular matrix organization (Figure 2B). The cellular component of gene ontology enriched mainly in adherens junction, extracellular matrix, collagen−containing extracellular matrix, and focal adhesion (Figure 2C). The molecular function of gene

ontology also contained extracellular matrix structural constituent (Figure 2D). These results maybe alert us that these DEGs mainly correlated with EMT (endothelial-mesenchymal transition) pathways in cancer.

Figure 3: Construction and Validation of the DEG-based prognostic signature.
We found that the association between risk score and survival time of patients was statistically significant, with the low-risk patients (n=76) showing a substantial advantage in OS time compared with the high-risk patients (n=76) (Figure 3A, D). A cluster analysis of each patient genes according to the risk scores was performed, to explore the true contributions of the fifteen genes. The visual diagram and the heat map were shown in Fig. 3B and 3C.

Figure 4: This model was tested in the validation cohort and total cohort. The risk scores of each patient were calculated, and the high-risk scores predicted by our model were significantly worse than those of low-risk scores (Figure 4A, D, E). The visual diagram and the heat map of validation cohort were shown in Figure 4B and 4C.

Figure 5: Compared with other clinicopathological factors, the model is an independent prognostic factor.

Figure 6: The time-dependent AUC curve indicates that the model has a good predictive performance.

Figure 7: A heatmap of the relative proportion of the 22 immune cells in each LACC sample were assessed and presented in Figure 7A. Among the 22 immune cell types, CD8 T cells, resting CD4 T cells, resting mast cells, activated mast cells, monocytes, neutrophil granulocyte were significantly different between high- and low-risk group, as shown in figure 7B. Resting CD4 T cells, monocytes, mast cells activated, and neutrophils were positively correlated with the fifteen-gene risk prediction model. CD8 T cells and mast cells resting were negatively correlated with fifteen-gene risk prediction model (Figure 7B).

Figure 8: A univariate Cox regression was run to analysis the difference expression of immune-checkpoint in high- and low-risk group. According to the results, CD274, TIGIT, LAG3, PDCD1, CTLA4, HAVCR2, PDCD1LG2 were highly expressed. Furthermore, the high expression of LAG3 and PDCD1 was associated with poor prognosis in cervical cancer patients, as shown in Figure 8A and 8B.

Figures 9: Expression Level of CYTL1 by IHC.

The positive rate of CYTL1 in NCR group is significantly higher than CR group (P = 0.01; Figures9A、B、C). Kaplan-Meier survival analysis showed that the prognosis of patients with high CYTL1 expression was poorer than low CYTL1 expression (P = 0.023, Figures 9E).

Figure 10: The correlation between CYTL1 and immune or inflammatory cell infiltration. The positive correlation between cells included activated CD4+ T cell (Act_CD4; P

=0.046), CD8+ T cell (P =1.672e-04), regulatory cell (Treg; P =2.823e-05), B cell (P =0.009), activated dendritic cells (Act_DC; P =0.035), Resting mast cell (P =0.006), Activated natural killer cells (P =0.035), as shown in Figure 10A-E. We verify the correlation between CYTL1 and CD4+ T cell, CD8+ T cell by IHC. The infiltration of CD8+ T cells is higher in the CR group than NCR group. The correlation between CYTL1 and the infiltration of CD8+ T cells was significant (P = 0.066, Rs = 0.889, Figure 10J、L). The correlation between CYTL1 and the infiltration of CD4+ T cells was not significant (Figure 10I、K).

Figures

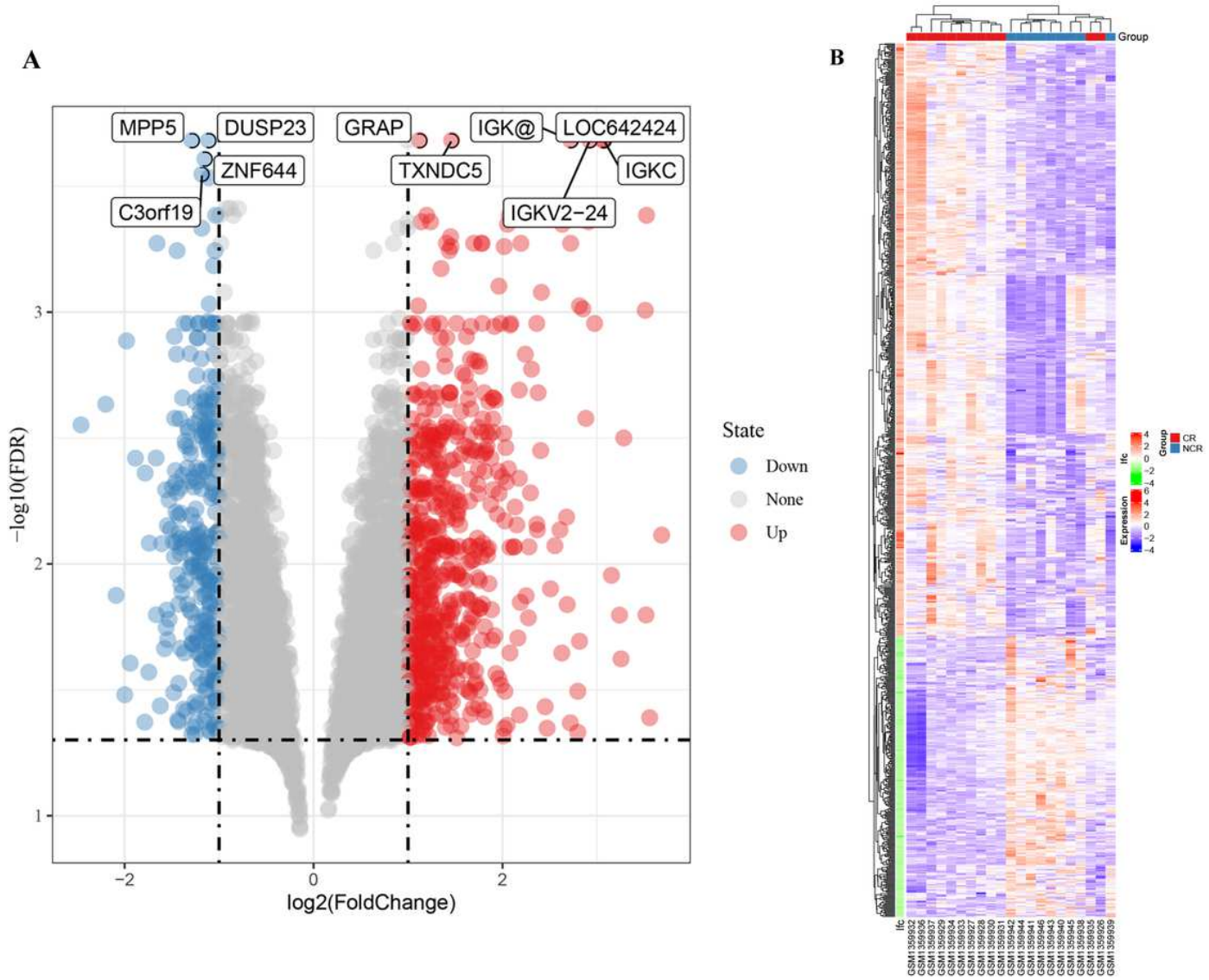


Figure 1

1021 DEGs (663 upregulated genes and 358 downregulated genes) were demonstrated from GSE56363 dataset.

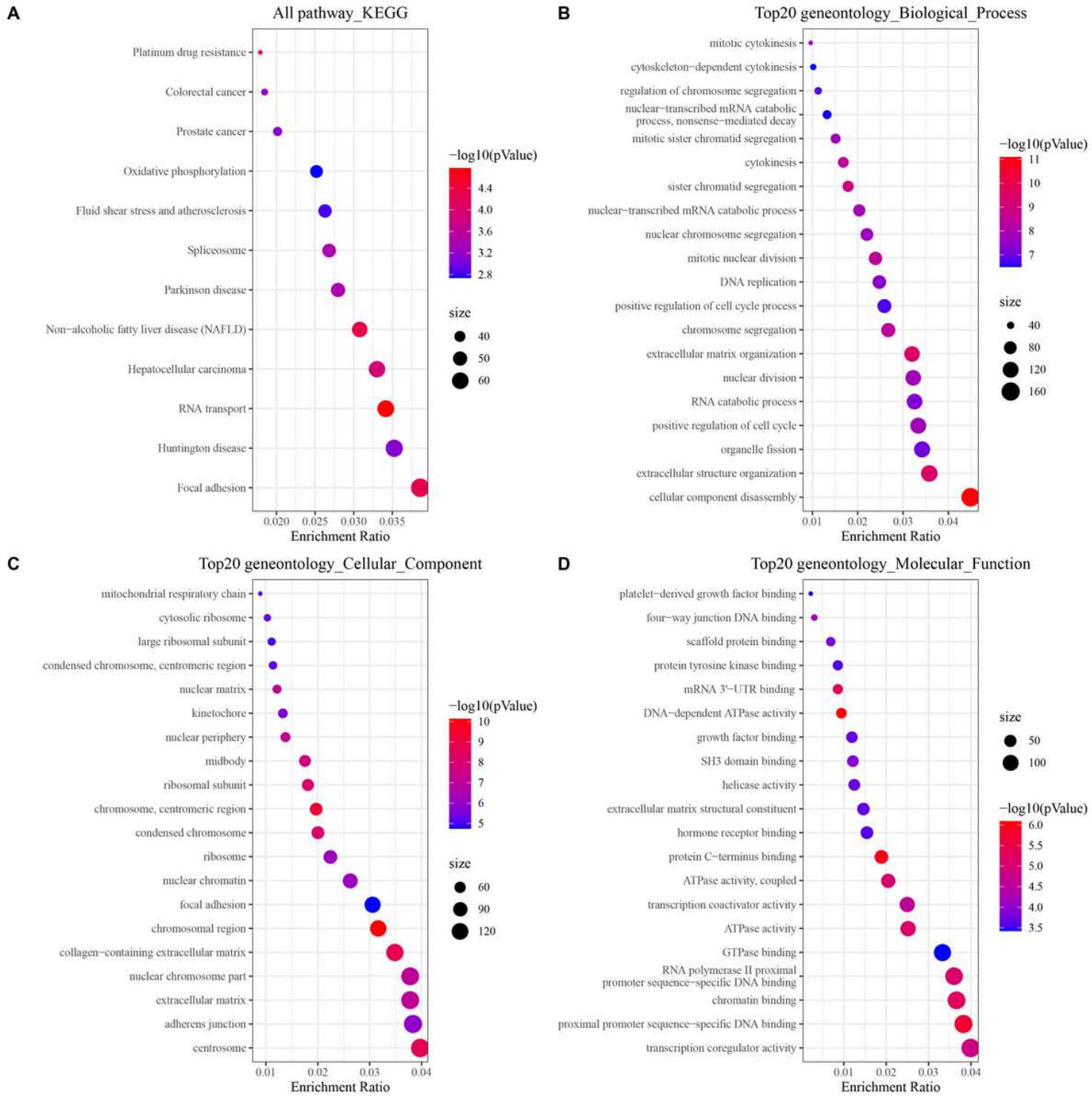


Figure 2

KEGG and GO pathway enrichment analysis for the DEGs. The pathway of these DEGs enriched mainly in focal adhesion and multiple tumors, as shown in Figure 2A. The biological process of gene ontology enriched mainly in cellular component disassembly, extracellular structure organization, organelle fission, and extracellular matrix organization (Figure 2B). The cellular component of gene ontology enriched mainly in adherens junction, extracellular matrix, collagen−containing extracellular matrix, and focal adhesion (Figure 2C). The molecular function of gene ontology also contained extracellular matrix

structural constituent (Figure 2D). These results maybe alert us that these DEGs mainly correlated with EMT (endothelial-mesenchymal transition) pathways in cancer.

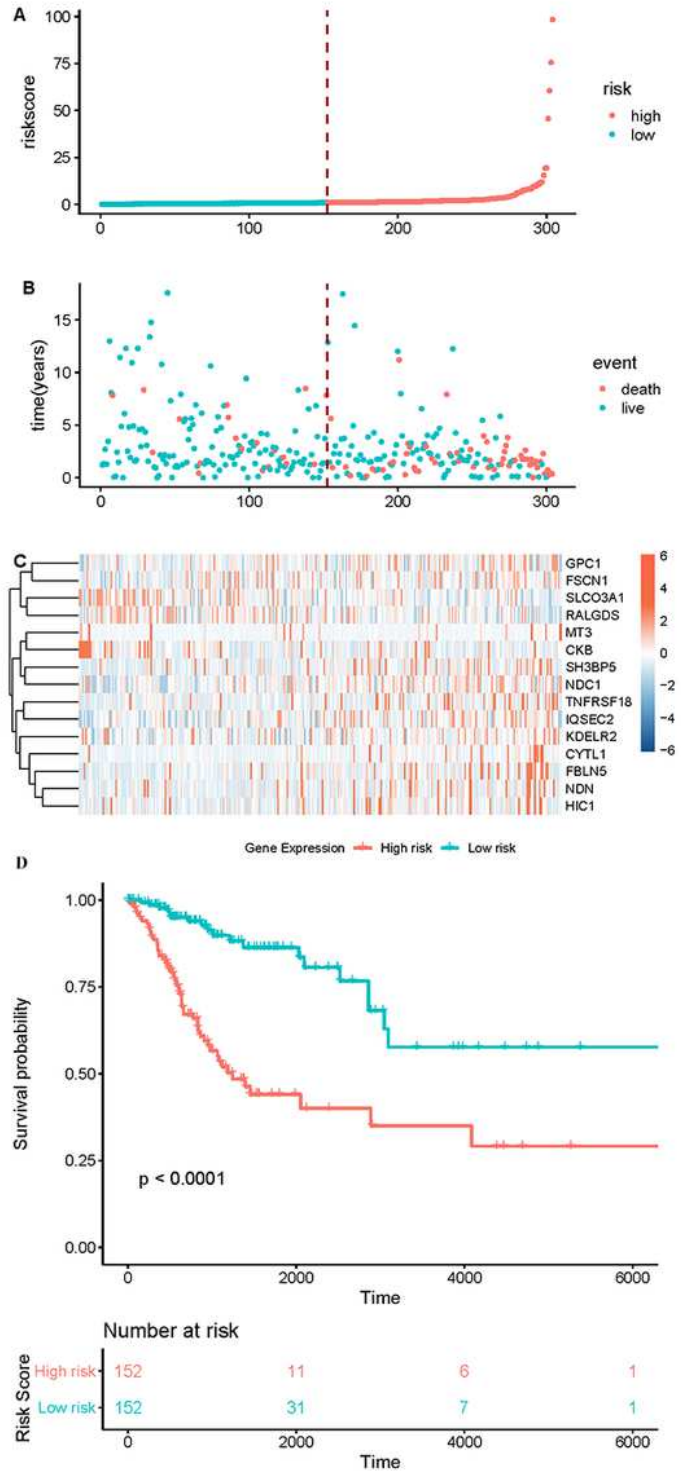


Figure 3

Construction and Validation of the DEG-based prognostic signature. We found that the association between risk score and survival time of patients was statistically significant, with the low-risk patients (n=76) showing a substantial advantage in OS time compared with the high-risk patients (n=76) (Figure

3A, D). A cluster analysis of each patient genes according to the risk scores was performed, to explore the true contributions of the fifteen genes. The visual diagram and the heat map were shown in Fig. 3B and 3C.

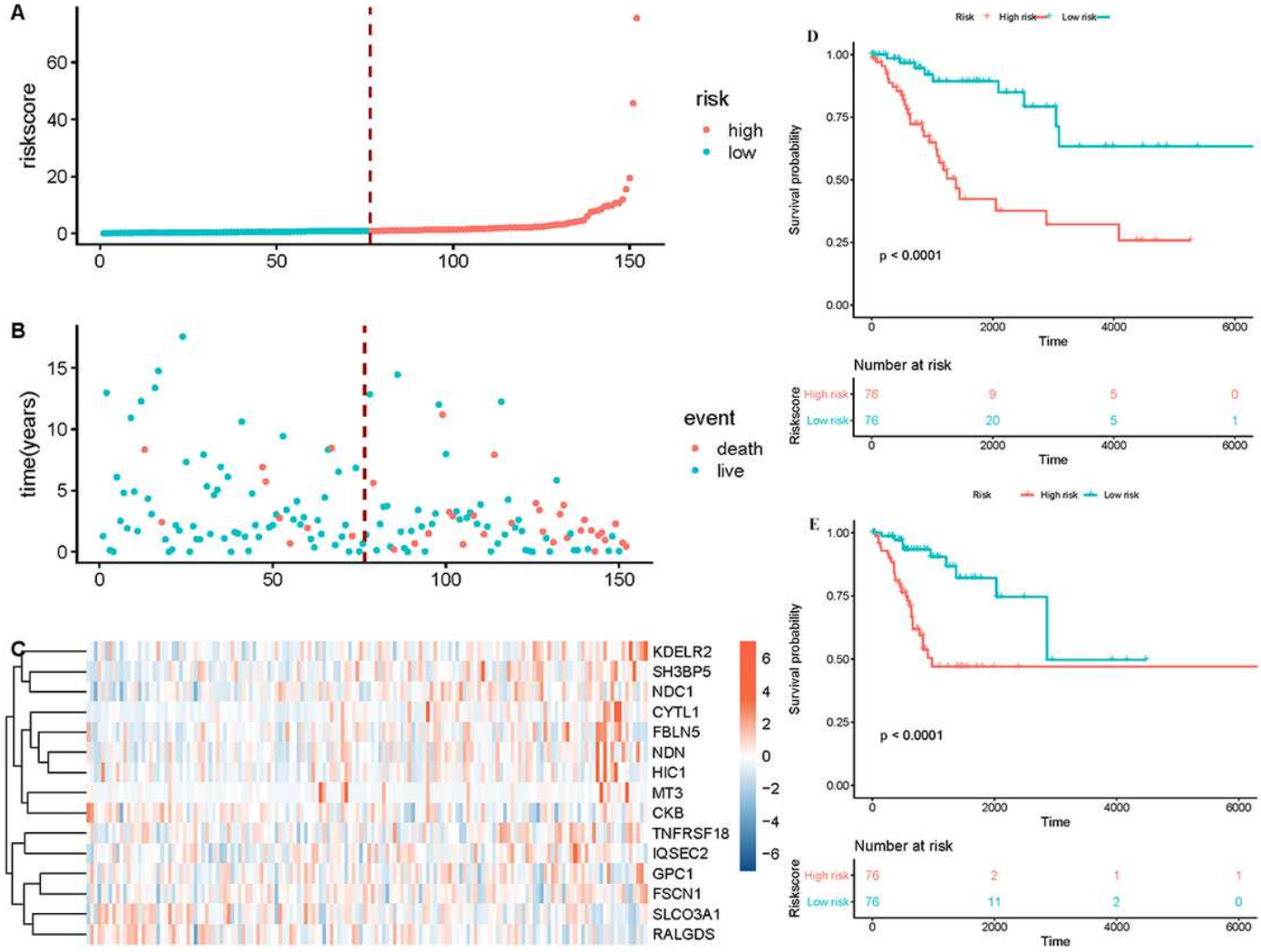


Figure 4

This model was tested in the validation cohort and total cohort. The risk scores of each patient were calculated, and the high-risk scores predicted by our model were significantly worse than those of low-risk scores (Figure 4A, D, E). The visual diagram and the heat map of validation cohort were shown in Figure 4B and 4C.

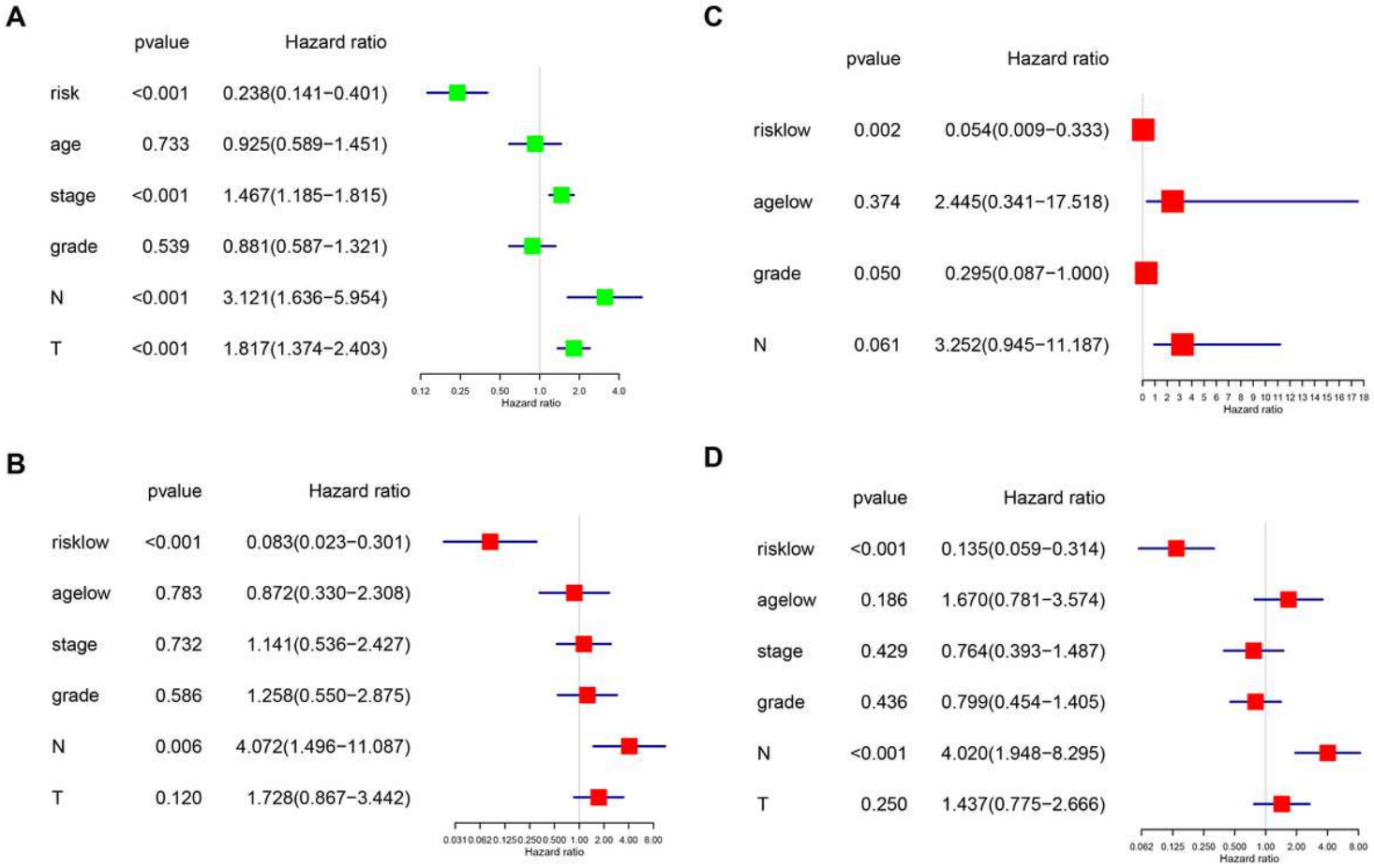


Figure 5

Compared with other clinicopathological factors, the model is an independent prognostic factor.

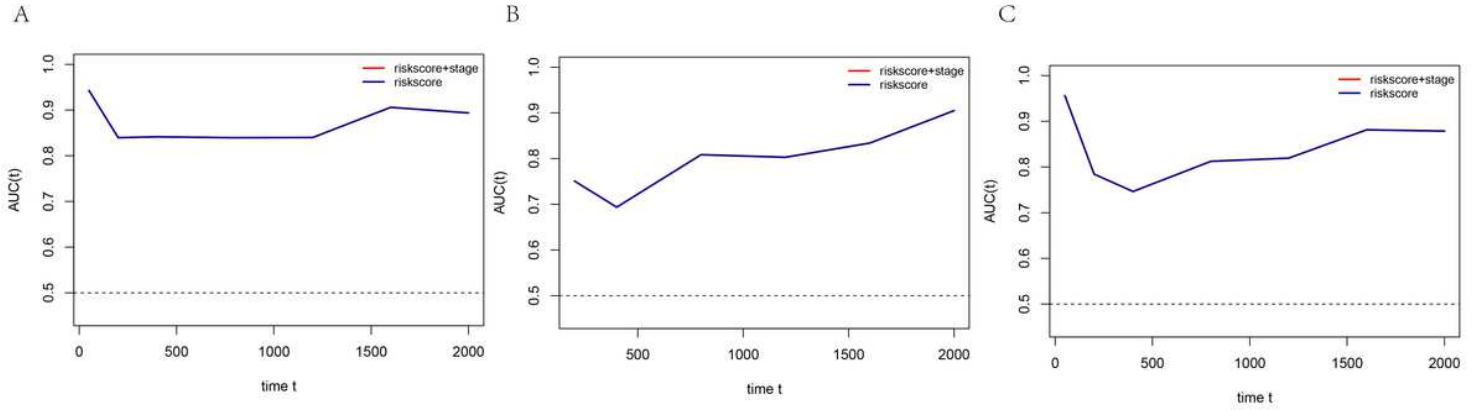


Figure 6

The time-dependent AUC curve indicates that the model has a good predictive performance.

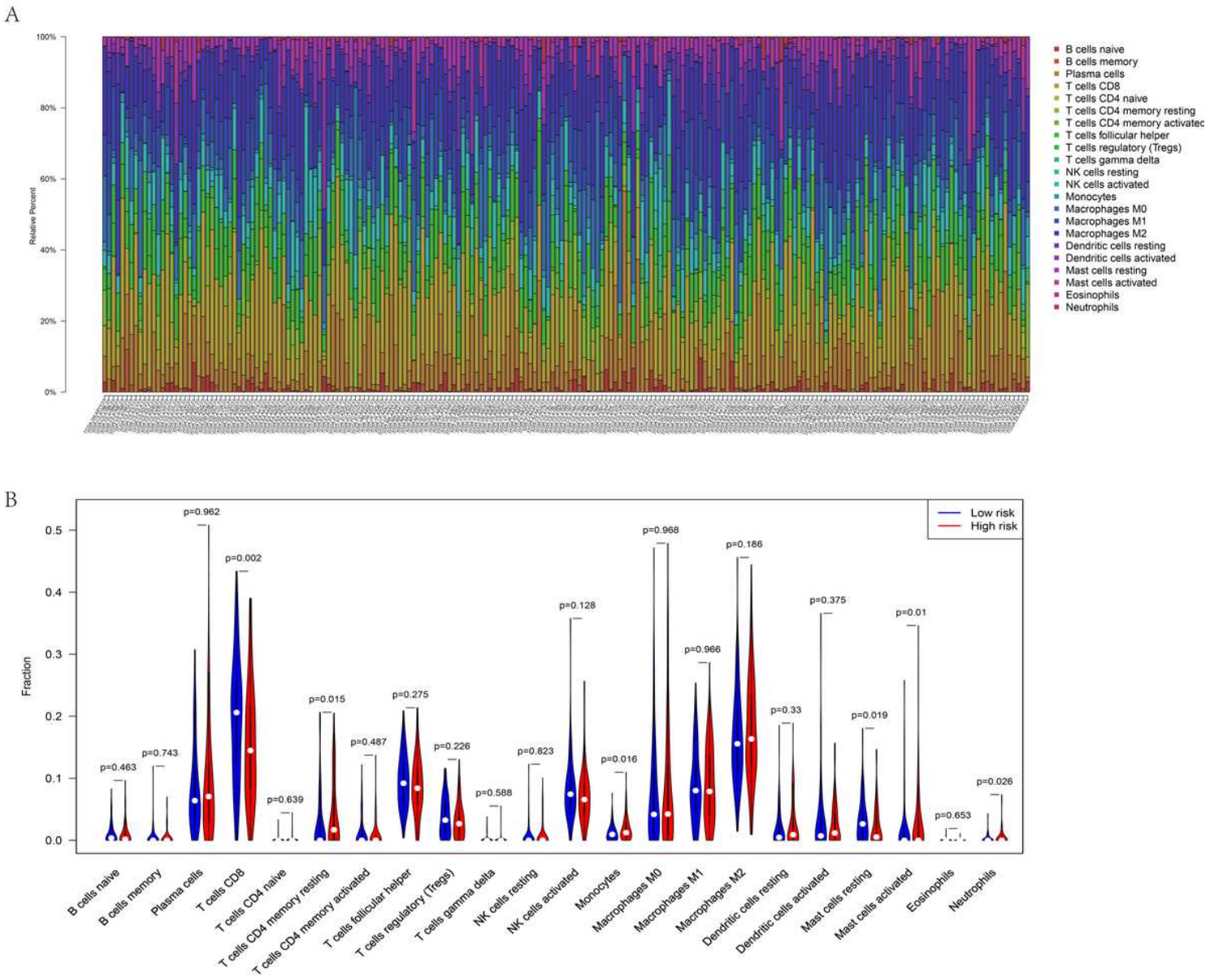


Figure 7

A heatmap of the relative proportion of the 22 immune cells in each LACC sample were assessed and presented in Figure 7A. Among the 22 immune cell types, CD8 T cells, resting CD4 T cells, resting mast cells, activated mast cells, monocytes, neutrophil granulocyte were significantly different between high- and low-risk group, as shown in figure 7B. Resting CD4 T cells, monocytes, mast cells activated, and neutrophils were positively correlated with the fifteen-gene risk prediction model. CD8 T cells and mast cells resting were negatively correlated with fifteen-gene risk prediction model (Figure 7B).

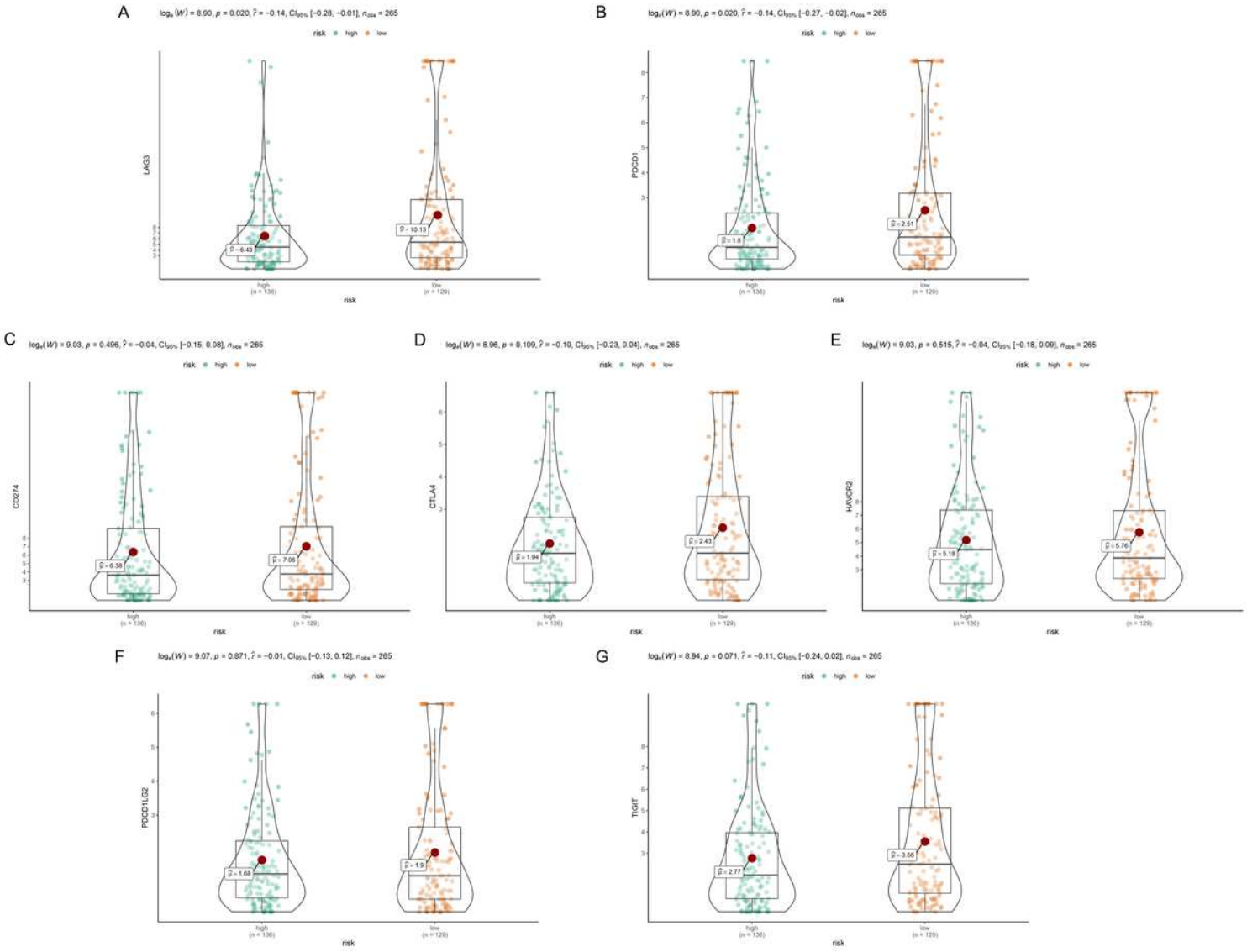
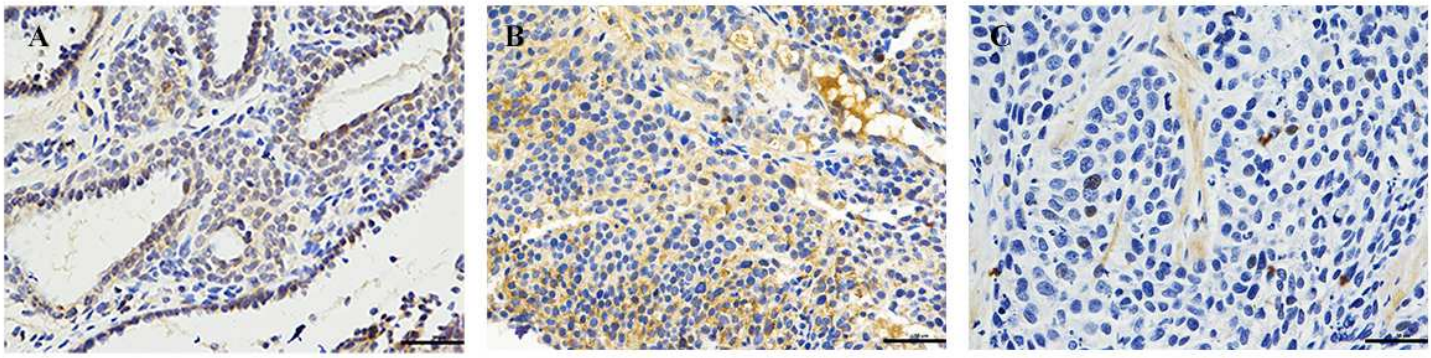


Figure 8

A univariate Cox regression was run to analysis the difference expression of immune-checkpoint in high- and low-risk group. According to the results, CD274, TIGIT, LAG3, PDCD1, CTLA4, HAVCR2, PDCD1LG2 were highly expressed. Furthermore, the high expression of LAG3 and PDCD1 was associated with poor prognosis in cervical cancer patients, as shown in Figure 8A and 8B.



D

Characteristics	n (%)	Survive rate of 5 years	Tumor-related death
FIGO 2009 stage			
IB2–II	80	55	25
III-IVA	33	24	9
Bulky			
No	34	23	11
Yes	79	56	23
Lymph nodes			
(+)	55	35	20
(-)	58	44	14

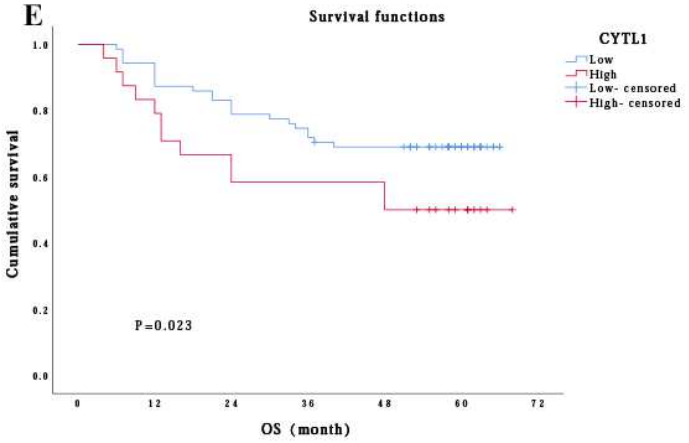
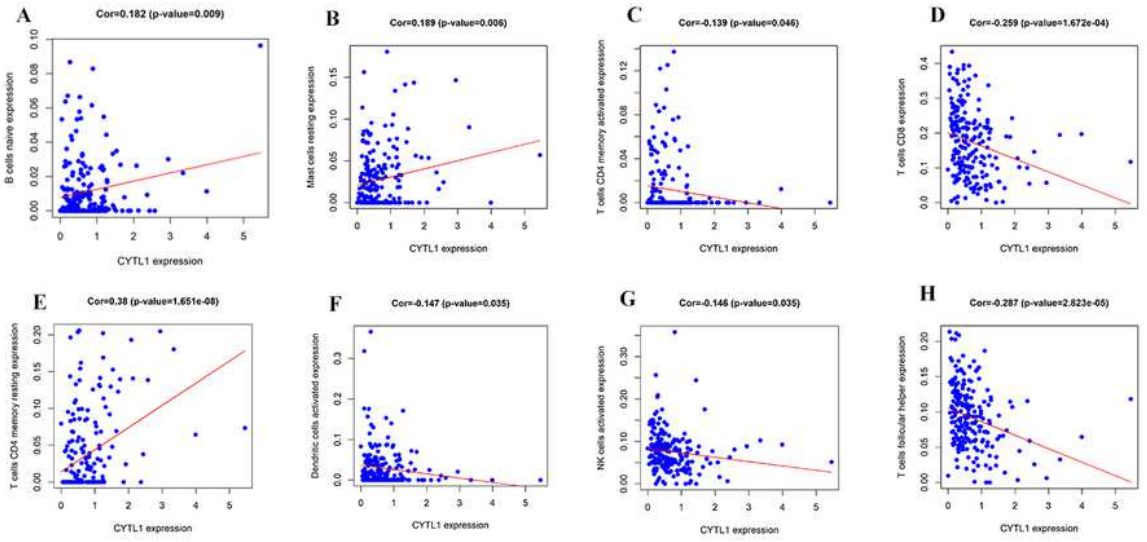


Figure 9

Expression Level of CYTL1 by IHC. The positive rate of CYTL1 in NCR group is significantly higher than CR group (P = 0.01; Figures9A□B□C). Kaplan-Meier survival analysis showed that the prognosis of patients with high CYTL1 expression was poorer than low CYTL1 expression (P = 0.023, Figures 9E).



I

Group	N	Expression of CD4				Rs	P
		(-)	low+	1+	2+		
CR	77	10	6	54	7	0.751	0.286
N Control	36	6	9	21	0		

J

Group	N	Expression of CD8				Rs	P
		(-)	low+	1+	2+		
CR	77	7	9	42	19	0.889	0.066
N Control	36	6	12	18	0		

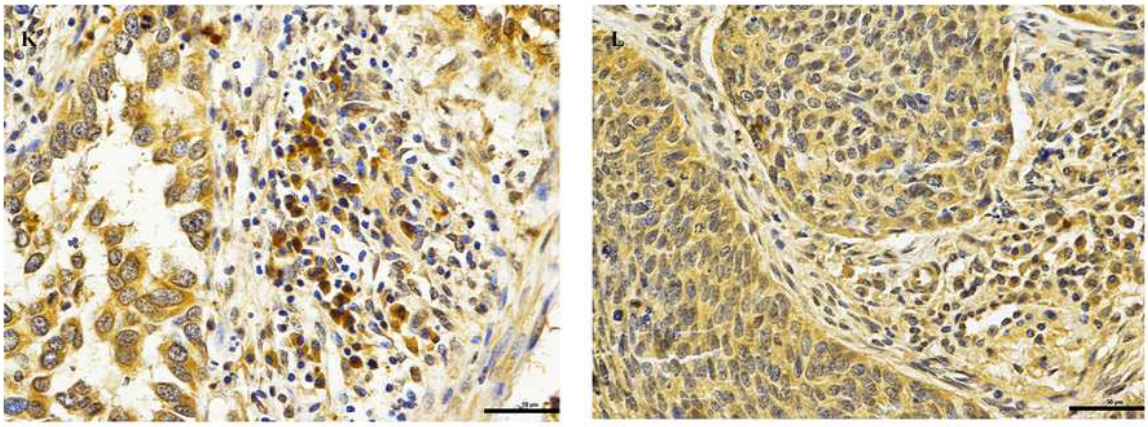


Figure 10

The correlation between CYTL1 and immune or inflammatory cell infiltration. The positive correlation between cells included activated CD4+ T cell (Act_CD4; P =0.046), CD8+ T cell (P =1.672e-04), regulatory cell (Treg; P =2.823e-05), B cell (P =0.009), activated dendritic cells (Act_DC; P =0.035), Resting mast cell (P =0.006), Activated natural killer cells (P =0.035), as shown in Figure 10A-E. We verify the correlation between CYTL1 and CD4+ T cell, CD8+ T cell by IHC. The infiltration of CD8+ T cells is higher in the CR

group than NCR group. The correlation between CYTL1 and the infiltration of CD8+ T cells was significant (P = 0.066, Rs = 0.889, Figure 10J□L). The correlation between CYTL1 and the infiltration of CD4+ T cells was not significant (Figure 10I□K).