

# Development and Validation of an Immune-related Prognostic Signature in Cervical Cancer

**Rongjia Su**

Shanghai First People's Hospital: Shanghai Jiaotong University First People's Hospital

**Chengwen Jin**

Shandong University School of Medicine: Shandong University Cheeloo College of Medicine

**Hualei Bu**

Shandong University Qilu Hospital

**Xiaoyun Wang**

Shanghai First People's Hospital: Shanghai Jiaotong University First People's Hospital

**Menghua Kuang**

Shanghai First People's Hospital: Shanghai Jiaotong University First People's Hospital

**Jiangdong Xiang**

Shanghai First People's Hospital: Shanghai Jiaotong University First People's Hospital

**chengjuan jin** (✉ [201815053@mial.sdu.edu.cn](mailto:201815053@mial.sdu.edu.cn))

Shanghai First People's Hospital: Shanghai Jiaotong University First People's Hospital

<https://orcid.org/0000-0001-6559-811X>

---

## Primary research

**Keywords:** Cervical cancer, Weighted gene co-expression network analysis (WGCNA), Least absolute shrinkage and selection operator (LASSO), Prognostic signature, Gene set enrichment analysis (GSEA)

**Posted Date:** January 29th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-153941/v1>

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

[Read Full License](#)

---

# Abstract

**Background:** Cervical cancer is the fourth most frequently gynecological malignancy across the world. Immunotherapies have proved to improve prognosis of cervical cancer. However, few studies on immune-related prognostic signature had been reported in cervical cancer.

**Methods:** Raw data and clinical information of cervical cancer samples were download from TCGA and UCSC Xena website. Immunophenoscore of immune infiltration cells in cervical cancer samples was calculated through ssGSEA method using GSVA package. WGCNA, Cox regression analysis, LASSO analysis and GSEA analysis were performed to classify cervical cancer prognosis and explore the biological signaling pathway.

**Results:** There were 8 immune infiltration cells associated with prognosis of cervical cancer. Through WGCNA, 153 genes from 402 immune-related genes were significantly correlated with prognosis of cervical cancer. A 15-gene signature demonstrated powerful predictive ability in prognosis of cervical cancer. GSEA analysis showed multiple signaling pathways containing PD-L1 expression and PD-1 checkpoint pathway differences between high risk and low risk groups.

**Conclusions:** The 15-gene signature was associated with multiple immune cells and immune infiltration in tumor microenvironment. Furthermore, the 15-gene signature is an effective potential prognostic classifier in the immunotherapies and surveillance of cervical cancer.

## Introduction

According to estimates from GLOBOCAN 2018, cervical cancer was the fourth most common cancer among women, with approximately 570,000 new cases and 311,000 deaths[1]. Cure rate for earliest stages are more than 90%, while locally advanced lesions treated with multimodality therapy can only achieve 65% cure rate in stage II lesions and 55% in stage III lesions[2]. Therapies to improve survivorship are in desperate need for cervical cancer especially locally advanced disease.

With rapid development of precision medicines, novel therapeutic strategies, especially immunotherapies, have been proposed to significantly improve clinical outcomes of cervical cancer[3, 4]. Harnessing an antitumor immune response has been a fundamental strategy in cancer immunotherapy. A paradigm shift has appeared in cancer immunotherapy: from traditional immune enhancement with low objective responses and frequent adverse events to more effective and less toxic reactions immune normalization[5, 6].

Cancer immunotherapies, such as administration of the cytokine IL-2, adoptive cell transfer (ACT) and the checkpoint modulators CTLA-4 and PD-1 have proved effective in clinical practice[6]. Blockade of the checkpoint modulators CTLA-4 and PD-1 starts the field of immune normalization in immunotherapy. Upregulated PD-1 in tumor microenvironment inhibits an effector T cell antitumor immune response, and therapies blocking this pathway have proven effective against multiple tumor types[5]. Anti-PD therapy

perform antitumor immunity mainly through the following three principles: 1) targeting a tumor-induced immune escape mechanism; 2) selectively modulating immunity in the tumor microenvironment; 3) resetting immunity in the tumor microenvironment[5]. It has been proved that patients with high mutation burden and burden of potential neoepitopes benefit more from immunological checkpoint blockade[7–11]. Immunotherapy plays a dispensable role in management of cervical cancer. In KEYNOTE-158, pembrolizumab has been approved by the US Food and Drug Administration for use in advanced cervical cancer with progressive disease either during or after chemotherapy. The objective response rate was 26.3% with a disease control rate of 68.4% for immunotherapy in cervical cancer[4]. Only small proportion patients benefited from immunotherapy, this proportion will hopefully increase with better patient selection and combination therapy. So it is urgent to find potential biomarkers for prediction of response to checkpoint immunotherapy and the rationale for the use of checkpoint immunotherapy.

In the present study, we qualify immune cells infiltration in cervical cancer and analyze the correlation between immune cells and cancer prognosis. Hub genes regulating prognosis through immune infiltration in cervical cancer were identified by weighted gene co-expression network analysis (WGCNA) and least absolute shrinkage and selection operator (LASSO). It was suggested by infiltrated immune cells and pathway enrichment analysis that our immune-related signature was closely related to tumor prognosis and could predict response of immunotherapy. Finally, a robust immune-related prognostic signature based on transcriptomics in cervical cancer was constructed and validated.

## **Materials And Methods**

### **Data Source and Processing**

The gene expression profiles and clinical information of cervical cancer were downloaded from the Cancer Genome Atlas (TCGA) Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>). Patients with pathologically confirmed cervical cancer and complete information about transcriptomics overall survivals (OS) was included in this study. Finally, a total of 304 primary cervical cancer samples, 2 metastatic cervical cancer samples and 3 normal cervix samples from TCGA was analyzed in our study. Details on clinical information of the included samples were summarized in Table 1. A workflow of this study was indicated in Fig. 1.

Table 1  
Clinical characteristics of 309 samples from TCGA

	<b>Alive (N = 235)</b>	<b>Dead (N = 74)</b>	<b>Overall (N = 309)</b>
<b>Age (years)</b>			
Mean (SD)	47.4 (13.2)	50.8 (15.4)	48.2 (13.8)
Median [Min, Max]	46.0 [20.0, 88.0]	48.0 [21.0, 79.0]	46.0 [20.0, 88.0]
<b>Pathology</b>			
Adenocarcinoma	24 (10.2%)	5 (6.8%)	29 (9.4%)
Adenosquamous carcinoma	4 (1.7%)	1 (1.4%)	5 (1.6%)
Basaloid squamous cell carcinoma	1 (0.4%)	0 (0%)	1 (0.3%)
Endometrioid adenocarcinoma	3 (1.3%)	0 (0%)	3 (1.0%)
Mucinous adenocarcinoma	12 (5.1%)	5 (6.8%)	17 (5.5%)
Papillary squamous cell carcinoma	1 (0.4%)	0 (0%)	1 (0.3%)
Squamous cell carcinoma	190 (80.9%)	63 (85.1%)	253 (81.9%)
<b>Stage</b>			
I	125 (53.2%)	38 (51.4%)	163 (52.8%)
II	60 (25.5%)	11 (14.9%)	71 (23.0%)
III	35 (14.9%)	11 (14.9%)	46 (14.9%)
IV	8 (3.4%)	14 (18.9%)	22 (7.1%)
Missing	7 (3.0%)	0 (0%)	7 (2.3%)
<b>Sample</b>			
Metastatic	1 (0.4%)	1 (1.4%)	2 (0.6%)
Primary Tumor	232 (98.7%)	72 (97.3%)	304 (98.4%)
Solid Tissue Normal	2 (0.9%)	1 (1.4%)	3 (1.0%)

## Infiltration of immune cells

Tumor-infiltrating immune cells can be quantified from RNA sequencing data of human tumors using bioinformatics approaches. Single-sample gene set enrichment analysis (ssGSEA) calculated and qualified the infiltration of immune cells through RNA-seq data. The infiltration of 28 immune cells was obtained. Immunophenoscore (normalized enrichment score, NES) of each TCGA cervical cancer sample

were calculated through ssGSEA. The ssGSEA ranks the genes by their absolute expression in a sample and computes Enrichment score by integrating the differences between the empirical cumulative distribution functions of the gene ranks[12].

## **Survival analysis**

Univariate cox regression analysis was performed to identify the association between cervical cancer survival and immunophenoscore of infiltrated immune cells. Forest plot was drawn to demonstrate the influence of infiltrated immune cells on survival. The best separation statistic was performed by use of “survminer” package which divides gene expression into high groups and low groups based on best separation. Next, Kaplan-Meier curve was made to further analyze relationship between survival and infiltration immune cells.

## **Construction of weighted gene co-expression network analysis (WGCNA)**

Weighted correlation network analysis (WGCNA) is an algorithm for finding genetic interactions in a weighted manner. It is used to build a gene co-expression networks to mine network modules closely associated with clinical traits through systematic biological method[13]. In this study, immunophenoscore of the infiltrated immune cells was regarded as target clinical traits. Genes expressing NA were removed. The top 25% genes with most median absolute deviation (MAD) used as a robust measure of variability were selected for WGCNA analysis.

Using the WGCNA function adjacency, an adjacency matrix is constructed by computing the Pearson correlation between all pairs of genes in the selected sample. Genes were divided into different gene modules based on the dissimilarity measure. A hierarchical clustering tree was constructed with different branches of the tree representing different gene modules.

The WGCNA network was built. There were 8 immune cells associated with cervical cancer survival. The magenta module was the module that most related to infiltration of immune cells. A total of 402 genes contained in magenta module were analyzed for survival using the standardized expression data FPKM. Finally, we found that 153 genes were associated with survival ( $p < 0.05$ ) in magenta module.

## **Screening hub genes by LASSO and survival analysis**

The 153 genes with the highest correlation associated with survival was picked out from WGCNA analysis. Then hub genes were further screened from the 153 genes by the use of LASSO. Survival analysis of hub genes was performed using “survival” and “survminer” packages to verify whether differently expressed genes affected tumor prognosis. Data of cervical cancer from UCSC Xena was downloaded and applied to compare the expression of hub genes in normal cervix and cervical cancer.

## **Construction of prognostic scoring model**

The TCGA cervical cancer patients were randomly divided into two groups by stage: the training group 70% and the test group 30% by the use of caret R package. Finally, a total of 199 patients were used as training group and 74 patients were regarded as test group. The selected key genes for support vector machine analysis were used to fit a LASSO Cox-Proportional Hazards (Cox-PH) model for selecting an optimal panel of predictive genes with penalized package of R[14]. Optimal lambda value was computed through a 10 cross-validations. Next, Cox-PH coefficients and infiltration of immune cells levels of these selected genes were used to calculate prognostic score as follows:

$$\text{Risk Score} = \sum (\text{coef}_{\text{gene}} * \text{immunophenoscore}_{\text{gene}}),$$

where  $\text{Coef}_{\text{gene}}$  and  $\text{immunophenoscore}_{\text{gene}}$  suggest Cox-PH coefficient and immunophenoscore level of a gene, respectively.

## Validation of prognostic scoring model

We calculated risk score of every patient in the training and test group by the model. We separated the training group and test group into a high-risk group and a low-risk group with the median risk score as cutoff respectively. Kaplan-Meier curve was applied to obtain and compare the overall survival time of two risk groups. ROC (operating characteristic curve) was performed to evaluate the predictive accuracy of the model.

## Comparison of 8 immune cell subtypes between high risk and low risk groups

To explore the differences of immune cell subtypes between high risk and low risk groups, the 8 immune cell subtypes associated with overall survival in cervical cancer were assessed in test and train cohorts. Mann-Whitney U test was used to compare differences in immune cell subtypes in the high risk and low risk groups.

## Gene set enrichment analysis (GSEA)

All 304 primary cervical cancer patients were evaluated by prognostic scoring model and then divided into high risk group and low risk group based on the standard cutoff. The global gene expression was analyzed and displayed with volcanic maps using the limma package in R. GSEA was conducted respectively to search "All gene sets" enriched in the samples with high risk group and low risk group. The differentially expressed genes were enriched in PD-L1 pathway functional sets.

## Results

### Quantify immune cell infiltration and survival analysis

We used ssGSEA to quantify mRNA data for immune cell infiltration. Finally, 28 infiltrating immune cells were included. Immunophenoscores of 304 primary cervical cancer samples, 2 metastatic cervical cancer samples and 3 normal cervix samples for 28 immune cells were calculated and demonstrated in Fig. 2A.

Univariate Cox regression analysis was performed to identify association between immune cells and overall survival of cervical cancer, the results are shown in Fig. 2B. It indicated that 8 immune cells were correlated with OS significantly. Moreover, these 8 immune cells (activated B cell, activated CD8 T cell, eosinophil, monocyte, activated CD4 T cell, effector memory CD8 T cell, immature B cell, plasmacytoid dendritic cell) were all protective factors for OS. Kaplan-Meier curve validated that high expression of these 8 immune cells were related to longer overall survival time suggested worse prognosis (Fig. 3).

## **The weighted gene co-expression network analysis (WGCNA) construction and key module identification**

The top 25% of the gene expression of variance was screened by the quartile of gene expression level, and 609 genes were screened out to construct co-expressed gene networks and the sample dendrogram and trait heatmap are constructed (Fig. 4A). In this study, the power of  $\beta = 8$  was selected to ensure a scale-free network (Fig. 4B, C) (scale free  $R^2 = 0.9$ , slope = -1.65).

After merging the modules with the high similarity of feature genes in the gene cluster dendrogram through a cut-line (0.25) (Fig. 5A), 9 modules were identified by the dynamic tree cut method. The clustering dendrograms of genes were shown in Fig. 5B. A heat map illustrating the correlation between immunophenoscores of infiltrated immune cells and key genes in the module was created (Fig. 5C).

It was obviously that magenta module was the most correlated with immune infiltration in the heat map. So magenta module was selected as the clinical significant module for further analysis. We found that the correlation coefficient between magenta module and 11 immune infiltration cells was  $\geq 0.7$  which suggested strong correlation. This indicates that the genes in this module are most relevant to tumor overall survival. All 402 genes in magenta module were analyzed for survival of cervical cancer. Through univariate cox regression survival analysis, 153 genes found to be significantly correlated with prognosis of cervical cancer were selected for further analysis.

## **Screening the key genes and survival analysis**

To develop a prognostic scoring model based on WGCNA, the 153 key genes were used to fit the LASSO Cox-PH model. Using parameter lambda (0.038) obtained upon performing 1000 cross-validations, a combination of 15 genes was obtained. These 15 genes were as follows: LAG3, CD74, CCL22, CH25H, OLR1, MIAT, BATF, IKZF3, TRARG, ACSL6, C11orf21, GTSF1, APOL1, CD1C, LINC00158 (Table 2, Fig. 6). Risk score of each sample in the training set was calculated. Then the samples in the training cohort were divided into high- and low-risk groups according to median risk score. As demonstrated in Fig. 6, the high-risk group patients in the training cohort has shorter overall survival time than the low-risk group ( $p < 0.001$ , HR: 5.1, 95%CI: 3.4–7.5), with an AUC of 0.803 in 1 year, 0.809 in 3 year, 0.800 in 5 year. Next, the 15- gene signature was then verified by the validation cohort from TCGA. Consistently, the high-risk group patients in the validation cohort had worse prognosis than the low-risk group ( $p < 0.001$ , HR: 1.3, 95%CI: 1.1–1.4) and the AUC was 0.845 in 1 year, 0.703 in 3 year and 0.761 in 5 year (Fig. 7). All these suggested

that the 15- gene signature associated with immune infiltration is able to predict prognosis in cervical cancer patients.

Table 2  
Fifteen genes from LASSO analysis

Gene	HR	95% CI	p.value	Lasso Coefficient
TRARG1	660	(0.00027-1.6e + 09)	0.39	4.7797999
ACSL6	3.4	(1.4–8.6)	0.008	0.653170468
C11orf21	1.3	(1.1–1.5)	0.00022	0.128185283
GTSF1	1.1	(1.0-1.2)	0.002	0.081672124
CD74	1.0	(1.0–1.0)	0.03	-8.18E-05
APOL1	1.0	(1.0–1.0)	0.016	0.001359139
OLR1	1.0	(1.0–1.0)	0.02	0.008057682
MIAT	1.0	(1.0–1.0)	5.00E-04	0.01207222
BATF	0.96	(0.93-1.0)	0.025	-0.006690799
LAG3	0.93	(0.86–0.99)	0.029	-0.012618
CCL22	0.90	(0.79-1.0)	0.073	-0.003652351
CH25H	0.88	(0.77-1.0)	0.053	-0.017915424
CD1C	0.78	(0.63–0.97)	0.027	-0.100839181
IKZF3	0.76	(0.63–0.93)	0.0061	-0.086985406
LINC00158	0.00062	(6.7e-07-0.58)	0.034	-0.708766884

## Immune cell subtypes between high risk and low risk groups

The expression levels of the 15 genes in test and validation cohorts are shown in Fig. 8A. Different immune scores had differential overall survival in cervical cancer samples. In both test cohorts and train cohorts, these 8 immune cell subtypes (activated B cell, activated CD8 T cell, eosinophil, monocyte, activated CD4 T cell, effector memory CD8 T cell, immature B cell, plasmacytoid dendritic cell) expressed differentially in high risk and low risk groups. (Fig. 8B-C).

## GSEA analysis

All 304 primary cervical cancer samples were divided into high-risk group and low-risk group based on the 15-gene signature. To identify potential function of the 15 key genes, GSEA was conducted respectively to

search “All gene sets” enriched in all 304 samples. GSEA showed 33 significant KEGG pathway associated with risk score, including PD-L1 expression and PD-1 checkpoint pathway in cancer (Fig. 9, Table 3). These results validated that the key genes in clinical significant module were mainly involved in the regulation of immune system. From GSEA, the 15-gene signature obviously participated in regulation of PD-1/PD-L1 pathway in cervical cancer.

Table 3  
Significant KEGG pathways from GSEA analysis

<b>ID</b>	<b>Description</b>	<b>P.value</b>	<b>p.adjust</b>	<b>q.values</b>
hsa04060	Cytokine-cytokine receptor interaction	3.29E-05	0.000668	0.000554
hsa04062	Chemokine signaling pathway	3.40E-05	0.000668	0.000554
hsa05152	Tuberculosis	3.42E-05	0.000668	0.000554
hsa04514	Cell adhesion molecules	3.45E-05	0.000668	0.000554
hsa05322	Systemic lupus erythematosus	3.49E-05	0.000668	0.000554
hsa04659	Th17 cell differentiation	3.52E-05	0.000668	0.000554
hsa04061	Viral protein interaction with cytokine and cytokine receptor	3.53E-05	0.000668	0.000554
hsa04660	T cell receptor signaling pathway	3.53E-05	0.000668	0.000554
hsa04640	Hematopoietic cell lineage	3.53E-05	0.000668	0.000554
hsa04658	Th1 and Th2 cell differentiation	3.54E-05	0.000668	0.000554
hsa04662	B cell receptor signaling pathway	3.56E-05	0.000668	0.000554
hsa05416	Viral myocarditis	3.63E-05	0.000668	0.000554
hsa05320	Autoimmune thyroid disease	3.64E-05	0.000668	0.000554
hsa04672	Intestinal immune network for IgA production	3.64E-05	0.000668	0.000554
hsa05340	Primary immunodeficiency	3.67E-05	0.000668	0.000554
hsa05330	Allograft rejection	3.67E-05	0.000668	0.000554
hsa05310	Asthma	3.70E-05	0.000668	0.000554
hsa04630	JAK-STAT signaling pathway	6.88E-05	0.001174	0.000974
hsa05145	Toxoplasmosis	0.000526	0.008506	0.007058
hsa04612	Antigen processing and presentation	0.000647	0.009924	0.008234
hsa05140	Leishmaniasis	0.000717	0.010486	0.008701
hsa05321	Inflammatory bowel disease	0.001157	0.01606	0.013326
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	0.001203	0.01606	0.013326
hsa05169	Epstein-Barr virus infection	0.001524	0.01949	0.016172
hsa05170	Human immunodeficiency virus 1 infection	0.001867	0.022921	0.019019
hsa04950	Maturity onset diabetes of the young	0.002213	0.026135	0.021685

ID	Description	P.value	p.adjust	q.values
hsa03320	PPAR signaling pathway	0.002867	0.032598	0.027048
hsa04390	Hippo signaling pathway	0.003763	0.040506	0.03361
hsa03010	Ribosome	0.003912	0.040506	0.03361
hsa04080	Neuroactive ligand-receptor interaction	0.004057	0.040506	0.03361
hsa04020	Calcium signaling pathway	0.00416	0.040506	0.03361
hsa05166	Human T-cell leukemia virus 1 infection	0.004222	0.040506	0.03361
hsa04940	Type I diabetes mellitus	0.00475	0.044185	0.036663

## Discussion

Cervical cancer is the fourth most frequently diagnosed cancer and the fourth leading cause of cancer death in women across the world[15]. Cervical cancer is a major challenge for global health[1]. The incidence and mortality vary widely with geographic location. Approximately 90% of cervical cancers occur in developing countries that lack organized screening and HPV vaccination[4]. In lower human development index settings, cervical cancer ranks second in incidence and mortality[15]. There are more than half a million new cervical cancer cases diagnosed and over 300 000 deaths worldwide each year[4]. Cervical cancer represents a major public health problem even in developed countries such as United States. There were 13800 new cases and 4290 deaths estimated in 2020 in United States[16]. Moreover, cervical cancer was the second leading cause of death in women aged 20–39 years in United States[16]. Increasing incidence of cervical adenocarcinoma and attenuation of earlier declines for cervical squamous cell carcinoma emphasize the importance of the need for improved therapeutic options to reduce the burden of cervical cancer of the U.S[17].

Cervical cancer in early stage remains a curable disease that could be treated by surgical resection and concurrent chemoradiation[2]. Women with bulky or locally advanced cervical cancer (IIB-IVA) when lesions are not amenable to high cure rates with surgery or radiation (RT) account for the majority of cervical cancer death[2]. The most important risk factors affecting prognosis of cervical cancer are stage, status of the lymph nodes, tumor volume, depth of tumor invasion into the cervical stroma, and LVSI[18]. The 5-year overall survival decreased significantly with rising stages, with FIGO stage IA, IB, IIB and IIIB were 90%, 83%, 70%, 42% respectively[19]. A substantial percentage of advanced cervical cancer patients will undergo recurrence and poor prognosis[2, 20]. Recurrence rate of cervical cancer fluctuates from 10–74%[21]. Tumor stage was one of the most pivotal factors related to recurrence. The recurrence rate by stage was as follows: 10% for stage IB, 17% for stage IIA, 23% for stage IIB, 42% for stage III and 74% for stage IVA[22]. Prognosis of metastatic and recurrent cervical cancer was extremely poor with a median survival time of 12 months[23]. Improved therapeutic options in management of cervical cancer, especially locally advanced cervical cancer is in urgent need.

High-risk subtypes of the human papilloma virus (HPV) are the cause of cervical cancer[24]. Viral oncoproteins E6 and E7 leads to dysregulation of p53 and HIF-1 alpha, thus affecting cell cycle proteins and VEGF expression[24–27]. HPV is highly immunogenic and elicits immune responses in humans, thus immune might play important roles in carcinogenesis of cervical cancer. Incidence of cervical cancer was substantial declined in countries with high HPV vaccine coverage[28–30]. Immunotherapy fights against tumor cells through activating endogenous immune response which seems to switch on the new frontier of the anticancer treatment[31]. Immunotherapy has different approaches, such as active immunotherapy (vaccine), passive immunotherapy (adoptive cellular transfer, antibodies and cytokines) and immunomodulation (cyclooxygenase 2 inhibitor)[31]. Discovery of immune checkpoint such as CTLA-4 and PD-1 plays an indispensable role in the development of cancer immunotherapy. It was surprising that immune checkpoint inhibitors anti-CTLA-4 and anti-PD-1 displayed enormous success in solid tumors[32]. Similar to other solid tumors, novel immunotherapeutic approaches, such as immune checkpoint inhibitors have shown promising results in cervical cancer[4]. Based on high response observed in KEYNOTE-158, the US Food and Drug Administration approved pembrolizumab inhibiting PD-1 for use in advanced cervical cancer with progressive disease either during or after chemotherapy[33].

Immunotherapy has shown encouraging results in cervical cancer treatment. Thus, implementing immunotherapeutic approaches earlier in advanced cervical cancer would seem to be most appropriate. However, the objective response rate with anti-PD-1/PD-L1 monotherapy is hovering at 20%. Moreover, immune-related toxicities and severe adverse effects can occur during PD-1/PD-L1 blockade therapy[34]. PD-1/PD-L1 blockade therapy does not demonstrate efficacy in almost 80% cervical cancer patients, suggesting the potential mechanism of PD-1/PD-L1 in immunotherapy remain to be further clarified. So new immune checkpoint inhibitors or comprehensive understanding of specific mechanism underlying PD-1/PD-L1 regulation in carcinogenesis is in urgent need.

Immune infiltration of cervical cancer determines the immune activation of tumor microenvironment and is related with clinical outcome of patients. In this study, immunophenoscore of 28 infiltrated immune cell in TCGA cervical cancer was calculated. Univariate Cox regression analysis showed that activated B cell, activated CD8 T cell, eosinophil, monocyte, activated CD4 T cell, effector memory CD8 T cell, immature B cell, plasmacytoid dendritic cell were strongly associated with overall survival of cervical cancer. WGCNA revealed that magenta module was the most relevant module to immune infiltration. In total, we identified 10 kinds of infiltrated immune cells with strong correlation with magenta module. These were as follows: activated B cell (R = 0.88), activated CD8 T cell (R = 0.82), activated dendritic cell (R = 0.79), central memory CD4 T cell (R = 0.78), effector memory CD8 T cell (R = 0.85), immature B cell (R = 0.93), macrophage (R = 0.72), MDSC (R = 0.85), regulatory T cell (R = 0.71), T follicular helper cell (R = 0.72), type 1 T helper cell (R = 0.83).

By the use of WGCNA, we found magenta module showed the highest correlation with the immune infiltration of cervical cancer. The magenta module contained 609 immune-related genes. There were 153 genes picked up from the magenta module for further analysis after taking intersection of prognosis and immune infiltration. Through LASSO analysis, 15 immune-related genes (LAG3, CD74, CCL22, CH25H,

OLR1, MIAT, BATF, IKZF3, TRARG, ACSL6, C11orf21, GTSF1, APOL1, CD1C, LINC00158) were included in prognostic classifier. Except CTLA-4 and PD-1/PD-L1 targeted cancer immunotherapy, LAG3 (lymphocyte activation gene-3, CD223) is the third inhibitory receptor to be targeted in the clinic[35]. CD74 (invariant chain) plays a dispensable role in process of immune systems that it participates in antigen presentation, B-cell differentiation and inflammatory signaling. CD74 has the potential to be a therapeutic target in cancer and autoimmune disease[36]. The chemokine CCL22 promoted regulatory T cell communication with dendritic cells to control immunity and was associated with poor prognosis[37, 38]. CH25H produces 25-hydroxycholesterol, which inhibited tumor-derived extracellular vesicles uptake and correlated with prognosis in melanoma patients[39]. OLR1 (oxidized LDL receptor 1) was a possible link between obesity, dyslipidemia and cancer. OLR1 played carcinogenic role by activating NF- $\kappa$ B pathway to promote proliferation, migration and inhibit apoptosis and de novo lipogenesis[40]. MIAT lncRNA was overexpressed in a number of malignancies and caused poor prognosis[41–43]. BATF was an important transcription factor regulating differentiation of early effector CD8 + T cells[44] and was a prognostic indicator for patients with colon cancer[45]. IKZF3 promoted growth of multiple tumors to cause poor prognosis[46].

The degree of immune infiltration significantly affected cervical cancer survival. In this study, the expression of 15 genes affected immune infiltration scores. Patients with higher immune scores had better prognosis. This study implied that prognosis value of risk score was associated with immune system. GSEA analysis showed that the 15-gene prognostic signature was obviously associated with PD-L1 expression and PD-1 checkpoint pathway in cancer. The prognostic signature could provide basis for potential immunotherapy in the future. However, the study has several limitations. First, no vitro or vivo molecular experiment was performed to verify our analysis, Second, our study was a retrospective study. So prospective study is in need to validate the findings of our study.

## Conclusions

In conclusion, we successfully constructed a 15-gene prognostic signature with powerful predictive function. Differences in the overall survival of high and low risk groups are implicated in immune infiltration, tumor microenvironment, PD-L1 expression and PD-1 checkpoint pathway. These findings revealed the underlying mechanism of immunotherapy and provided basis for cervical cancer pathogenesis and clinical treatment.

## Declarations

### Acknowledgements

We appreciate researchers who have developed and maintained public databases such as TCGA, UCSC which will accelerate the understanding and treatment of human cancer.

### Authors' contributions

S-RJ is responsible for experimental design. J-CW and B-HL are responsible for instrument operation. W-XY and K-MH are responsible for data analysis. X-JD and J-CJ are for providing overall ideas. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the National Natural Science Foundation of China (82002736).

## Availability of data and materials

The raw data supporting the conclusions of this article sourced the Cancer Genome Atlas (TCGA) Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>).

## Consent for publication

Not applicable.

## Competing interest

The authors declare that they have no competing interests.

## Declarations

Not applicable.

## References

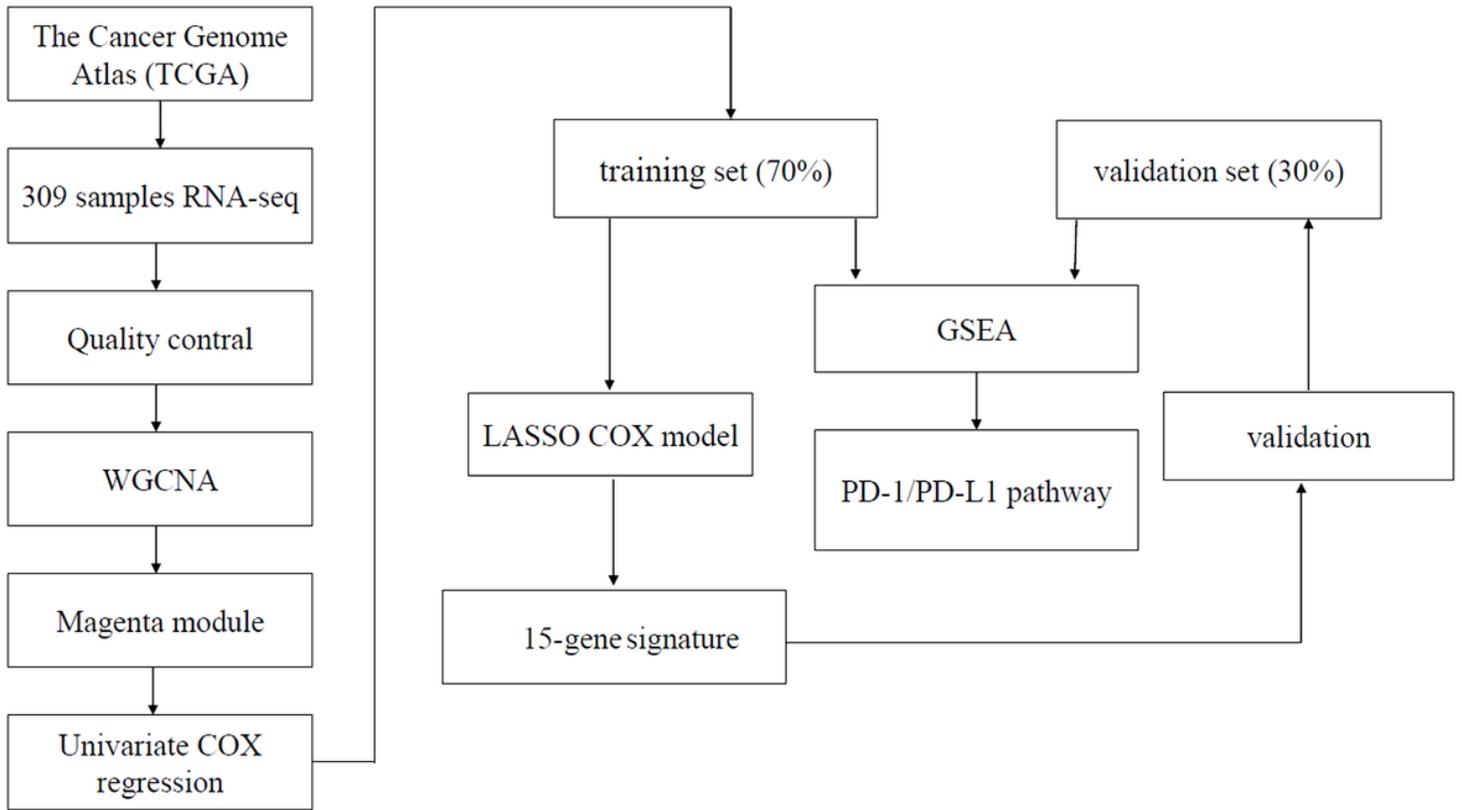
1. Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, Bray F: **Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis.** *The Lancet Global health* 2020, **8**(2):e191-e203.
2. Monk BJ, Tewari KS, Koh WJ: **Multimodality therapy for locally advanced cervical carcinoma: state of the art and future directions.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2007, **25**(20):2952-2965.
3. Chung HC, Ros W, Delord JP, Perets R, Italiano A, Shapira-Frommer R, Manzuk L, Piha-Paul SA, Xu L, Zeigenfuss S *et al*: **Efficacy and Safety of Pembrolizumab in Previously Treated Advanced Cervical Cancer: Results From the Phase II KEYNOTE-158 Study.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2019, **37**(17):1470-1478.
4. Cohen PA, Jhingran A, Oaknin A, Denny L: **Cervical cancer.** *Lancet (London, England)* 2019, **393**(10167):169-182.
5. Sanmamed MF, Chen L: **A Paradigm Shift in Cancer Immunotherapy: From Enhancement to Normalization.** *Cell* 2018, **175**(2):313-326.
6. Tran E, Robbins PF, Rosenberg SA: **'Final common pathway' of human cancer immunotherapy: targeting random somatic mutations.** *Nature immunology* 2017, **18**(3):255-262.

7. Asaoka Y, Ijichi H, Koike K: **PD-1 Blockade in Tumors with Mismatch-Repair Deficiency.** *The New England journal of medicine* 2015, **373**(20):1979.
8. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, Berent-Maoz B, Pang J, Chmielowski B, Cherry G *et al.*: **Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma.** *Cell* 2016, **165**(1):35-44.
9. Powles T, Eder JP, Fine GD, Braiteh FS, Loriot Y, Cruz C, Bellmunt J, Burris HA, Petrylak DP, Teng SL *et al.*: **MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer.** *Nature* 2014, **515**(7528):558-562.
10. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS *et al.*: **Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer.** *Science (New York, NY)* 2015, **348**(6230):124-128.
11. van Rooij N, van Buuren MM, Philips D, Velds A, Toebes M, Heemskerk B, van Dijk LJ, Behjati S, Hilkmann H, El Atmioui D *et al.*: **Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2013, **31**(32):e439-442.
12. Finotello F, Trajanoski Z: **Quantifying tumor-infiltrating immune cells from transcriptomics data.** *Cancer immunology, immunotherapy : CII* 2018, **67**(7):1031-1040.
13. Langfelder P, Horvath S: **WGCNA: an R package for weighted correlation network analysis.** *BMC bioinformatics* 2008, **9**:559.
14. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F: **Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.** *International journal of cancer* 2015, **136**(5):E359-386.
15. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.** *CA: a cancer journal for clinicians* 2018, **68**(6):394-424.
16. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2020.** *CA: a cancer journal for clinicians* 2020, **70**(1):7-30.
17. Islami F, Fedewa SA, Jemal A: **Trends in cervical cancer incidence rates by age, race/ethnicity, histological subtype, and stage at diagnosis in the United States.** *Preventive medicine* 2019, **123**:316-323.
18. Peters WA, 3rd, Liu PY, Barrett RJ, 2nd, Stock RJ, Monk BJ, Berek JS, Souhami L, Grigsby P, Gordon W, Jr., Alberts DS: **Concurrent chemotherapy and pelvic radiation therapy compared with pelvic radiation therapy alone as adjuvant therapy after radical surgery in high-risk early-stage cancer of the cervix.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2000, **18**(8):1606-1613.
19. Sturdza A, Pötter R, Fokdal LU, Haie-Meder C, Tan LT, Mazon R, Petric P, Šegedin B, Jurgenliemk-Schulz IM, Nomden C *et al.*: **Image guided brachytherapy in locally advanced cervical cancer: Improved pelvic control and survival in RetroEMBRACE, a multicenter cohort study.** *Radiotherapy and*

- oncology : journal of the European Society for Therapeutic Radiology and Oncology* 2016, **120**(3):428-433.
20. Lontos M, Kyriazoglou A, Dimitriadis I, Dimopoulos MA, Bamias A: **Systemic therapy in cervical cancer: 30 years in review.** *Critical reviews in oncology/hematology* 2019, **137**:9-17.
  21. Peiretti M, Zapardiel I, Zanagnolo V, Landoni F, Morrow CP, Maggioni A: **Management of recurrent cervical cancer: a review of the literature.** *Surgical oncology* 2012, **21**(2):e59-66.
  22. Perez CA, Grigsby PW, Nene SM, Camel HM, Galakatos A, Kao MS, Lockett MA: **Effect of tumor size on the prognosis of carcinoma of the uterine cervix treated with irradiation alone.** *Cancer* 1992, **69**(11):2796-2806.
  23. Elit LM, Hirte H: **Management of advanced or recurrent cervical cancer: chemotherapy and beyond.** *Expert review of anticancer therapy* 2014, **14**(3):319-332.
  24. Wardak S: **Human Papillomavirus (HPV) and cervical cancer.** *Medycyna doświadczalna i mikrobiologia* 2016, **68**(1):73-84.
  25. Burd EM: **Human papillomavirus and cervical cancer.** *Clinical microbiology reviews* 2003, **16**(1):1-17.
  26. Monk BJ, Willmott LJ, Sumner DA: **Anti-angiogenesis agents in metastatic or recurrent cervical cancer.** *Gynecologic oncology* 2010, **116**(2):181-186.
  27. Nakamura M, Bodily JM, Beglin M, Kyo S, Inoue M, Laimins LA: **Hypoxia-specific stabilization of HIF-1alpha by human papillomaviruses.** *Virology* 2009, **387**(2):442-448.
  28. Brotherton JM, Fridman M, May CL, Chappell G, Saville AM, Gertig DM: **Early effect of the HPV vaccination programme on cervical abnormalities in Victoria, Australia: an ecological study.** *Lancet (London, England)* 2011, **377**(9783):2085-2092.
  29. Drolet M, Bénard É, Pérez N, Brisson M: **Population-level impact and herd effects following the introduction of human papillomavirus vaccination programmes: updated systematic review and meta-analysis.** *Lancet (London, England)* 2019, **394**(10197):497-509.
  30. Gargano JW, Zhou F, Stokley S, Markowitz LE: **Human papillomavirus vaccination in commercially-insured vaccine-eligible males and females, United States, 2007-2014.** *Vaccine* 2018, **36**(23):3381-3386.
  31. Ventriglia J, Paciolla I, Pisano C, Cecere SC, Di Napoli M, Tambaro R, Califano D, Losito S, Scognamiglio G, Setola SV *et al*: **Immunotherapy in ovarian, endometrial and cervical cancer: State of the art and future perspectives.** *Cancer treatment reviews* 2017, **59**:109-116.
  32. Balar AV, Weber JS: **PD-1 and PD-L1 antibodies in cancer: current status and future directions.** *Cancer immunology, immunotherapy : CII* 2017, **66**(5):551-564.
  33. Marabelle A, Fakih M, Lopez J, Shah M, Shapira-Frommer R, Nakagawa K, Chung HC, Kindler HL, Lopez-Martin JA, Miller WH, Jr. *et al*: **Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study.** *The Lancet Oncology* 2020, **21**(10):1353-1365.

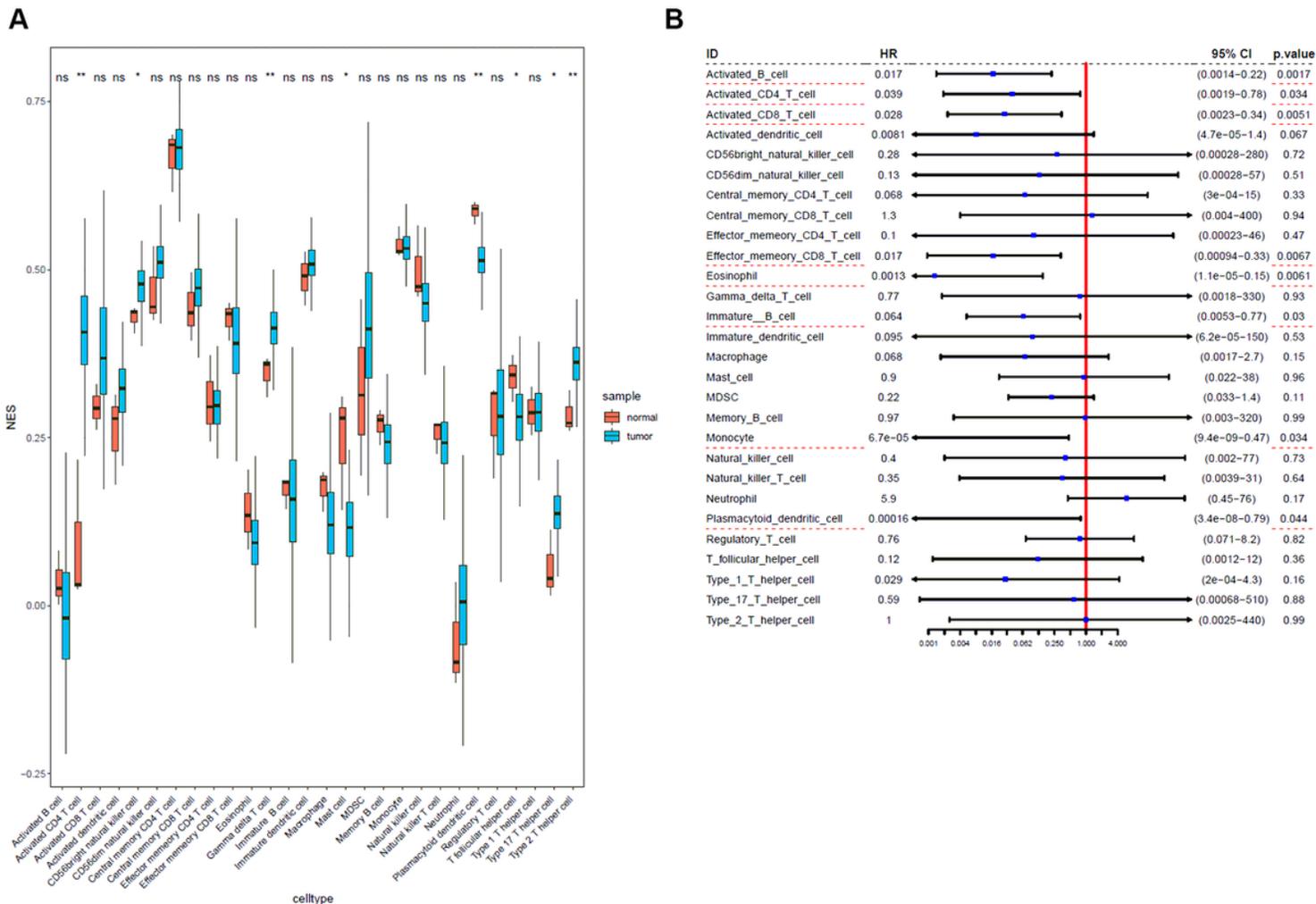
34. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, Powderly JD, Carvajal RD, Sosman JA, Atkins MB *et al*: **Safety, activity, and immune correlates of anti-PD-1 antibody in cancer.** *The New England journal of medicine* 2012, **366**(26):2443-2454.
35. Andrews LP, Marciscano AE, Drake CG, Vignali DA: **LAG3 (CD223) as a cancer immunotherapy target.** *Immunological reviews* 2017, **276**(1):80-96.
36. Borghese F, Clanchy FI: **CD74: an emerging opportunity as a therapeutic target in cancer and autoimmune disease.** *Expert opinion on therapeutic targets* 2011, **15**(3):237-251.
37. Rapp M, Wintergerst MWM, Kunz WG, Vetter VK, Knott MML, Lisowski D, Haubner S, Moder S, Thaler R, Eiber S *et al*: **CCL22 controls immunity by promoting regulatory T cell communication with dendritic cells in lymph nodes.** *The Journal of experimental medicine* 2019, **216**(5):1170-1181.
38. Wang Q, Schmoeckel E, Kost BP, Kuhn C, Vattai A, Vilsmaier T, Mahner S, Mayr D, Jeschke U, Heidegger HH: **Higher CCL22+ Cell Infiltration is Associated with Poor Prognosis in Cervical Cancer Patients.** *Cancers* 2019, **11**(12).
39. Ortiz A, Gui J, Zahedi F, Yu P, Cho C, Bhattacharya S, Carbone CJ, Yu Q, Katlinski KV, Katlinskaya YV *et al*: **An Interferon-Driven Oxysterol-Based Defense against Tumor-Derived Extracellular Vesicles.** *Cancer cell* 2019, **35**(1):33-45.e36.
40. Khaidakov M, Mitra S, Kang BY, Wang X, Kadlubar S, Novelli G, Raj V, Winters M, Carter WC, Mehta JL: **Oxidized LDL receptor 1 (OLR1) as a possible link between obesity, dyslipidemia and cancer.** *PLoS one* 2011, **6**(5):e20277.
41. Huang X, Gao Y, Qin J, Lu S: **lncRNA MIAT promotes proliferation and invasion of HCC cells via sponging miR-214.** *American journal of physiology Gastrointestinal and liver physiology* 2018, **314**(5):G559-g565.
42. Sha M, Lin M, Wang J, Ye J, Xu J, Xu N, Huang J: **Long non-coding RNA MIAT promotes gastric cancer growth and metastasis through regulation of miR-141/DDX5 pathway.** *Journal of experimental & clinical cancer research : CR* 2018, **37**(1):58.
43. Liu W, Wang Z, Wang C, Ai Z: **Long non-coding RNA MIAT promotes papillary thyroid cancer progression through upregulating LASP1.** *Cancer cell international* 2019, **19**:194.
44. Kurachi M, Barnitz RA, Yosef N, Odorizzi PM, Dilorio MA, Lemieux ME, Yates K, Godec J, Klatt MG, Regev A *et al*: **The transcription factor BATF operates as an essential differentiation checkpoint in early effector CD8+ T cells.** *Nature immunology* 2014, **15**(4):373-383.
45. Dai L, Cui X, Zhang X, Cheng L, Liu Y, Yang Y, Fan P, Wang Q, Lin Y, Zhang J *et al*: **SARI inhibits angiogenesis and tumour growth of human colon cancer through directly targeting ceruloplasmin.** *Nature communications* 2016, **7**:11996.
46. Liu XP, He L, Zhang QP, Zeng XT, Liu SQ: **Baicalein Inhibits Proliferation of Myeloma U266 Cells by Downregulating IKZF1 and IKZF3.** *Medical science monitor : international medical journal of experimental and clinical research* 2018, **24**:2809-2817.

## Figures



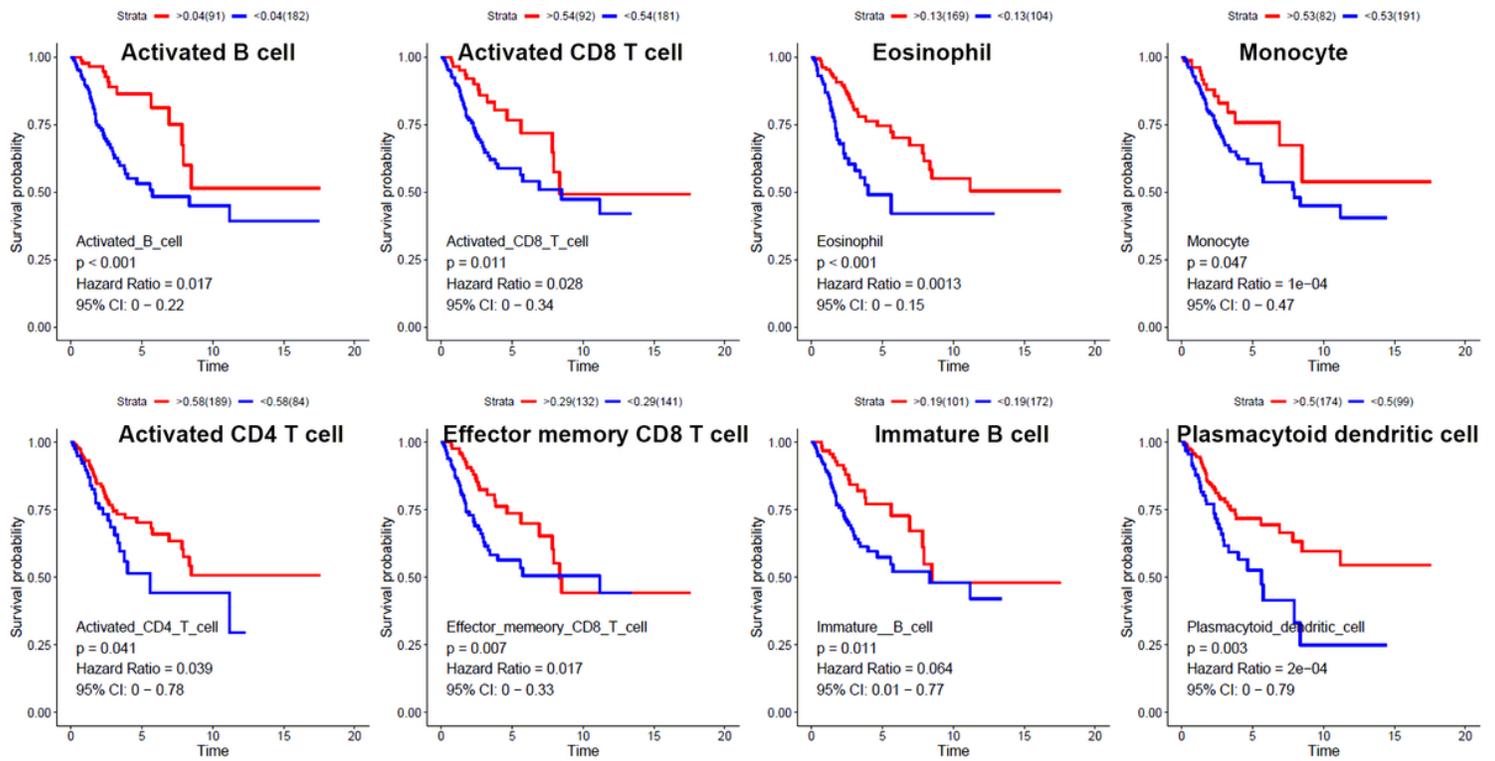
**Figure 1**

Flowchart of the 15-gene signature prognostic model construction process.



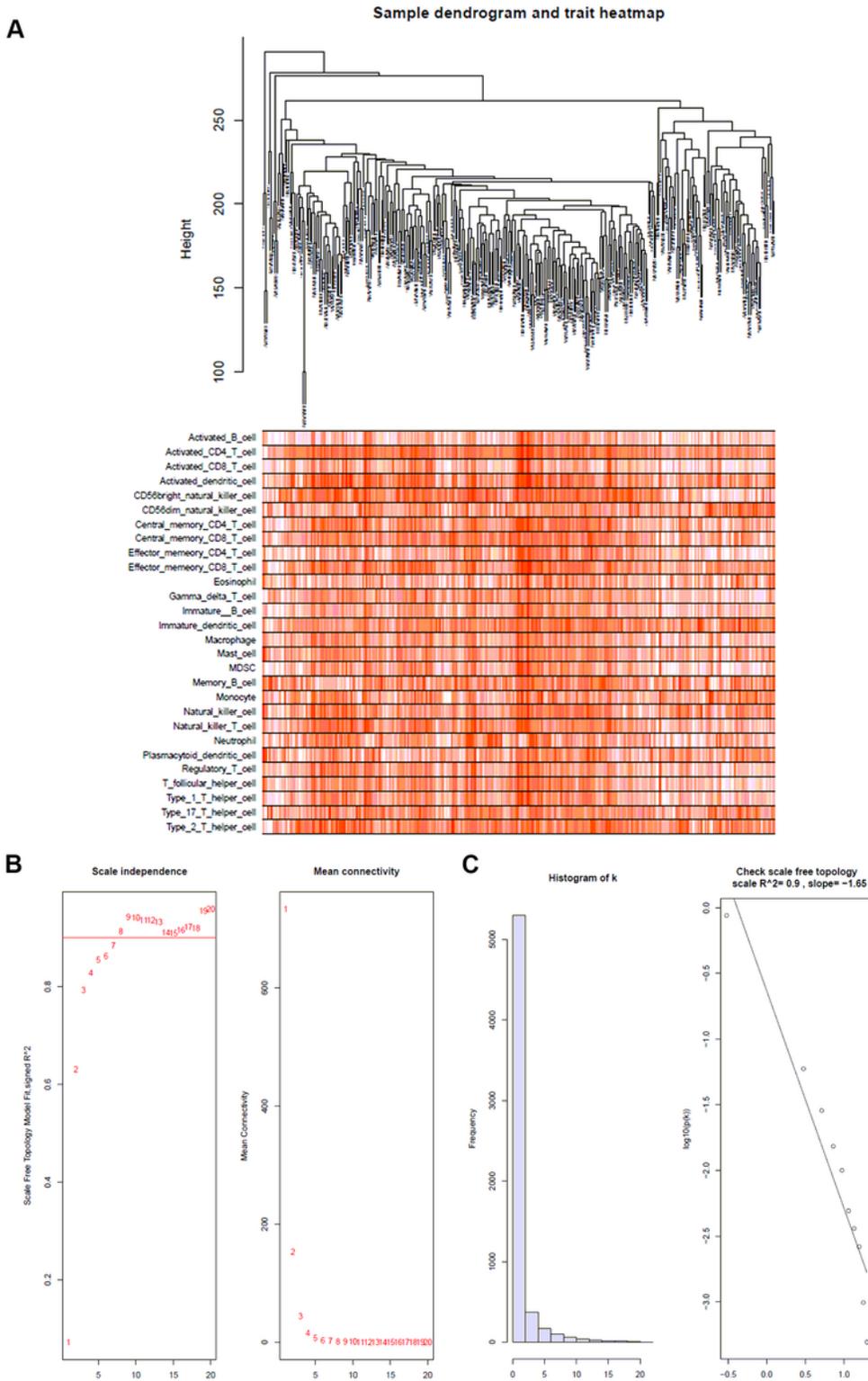
**Figure 2**

Immunophenoscores and hazard ratios of 28 immune cells in cervical cancer. (A) Boxplot of immunophenoscores from 28 immune cells TCGA cervical cancer and normal samples. (B) Forest plot of hazard ratios in each infiltrated immune cell for OS. Infiltrated immune cell subtypes significantly associated with overall survival were underlined with a red dashed line.



**Figure 3**

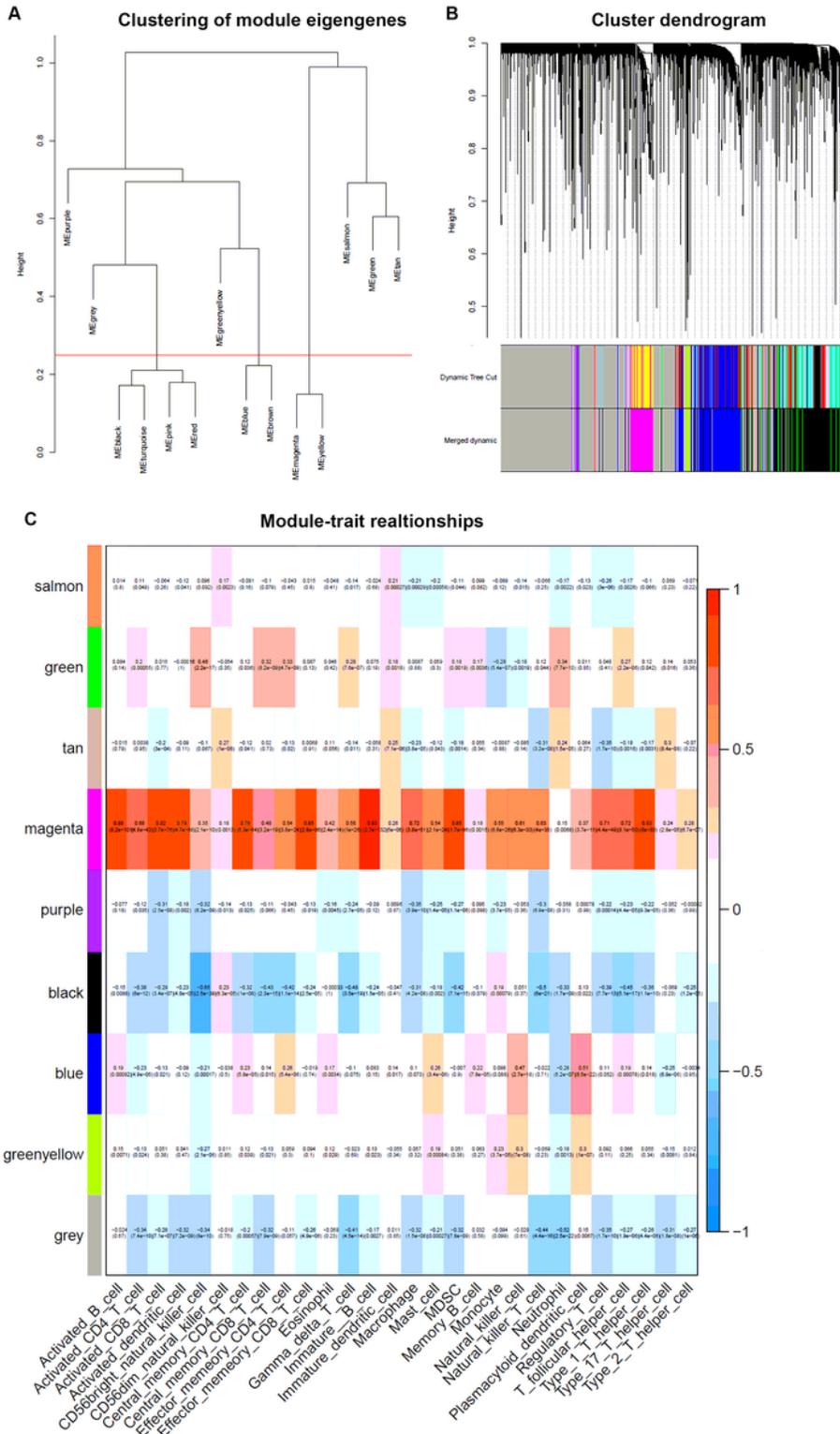
Kaplan-Meier curves for overall survival of 8 infiltrated immune cells. Activated B cell, activated CD8 T cell, eosinophil, monocyte, activated CD4 T cell, effector memory CD8 T cell, immature B cell and plasmacytoid dendritic cell were associated with overall survival of cervical cancer.



**Figure 4**

Sample tree and trait heatmap of 28 immune cells. (A) Sample dendrogram and trait heatmap of 28 immune cells. The clinical trait information is immunophenoscores of 28 infiltrated immune cells in each sample. (B) Analysis of the scale-free fit index for various soft-thresholding powers ( $\beta$ ) and the mean connectivity for soft threshold powers. The soft-thresholding power in the WGCNA was determined based on a scale-free  $R^2$  ( $R^2 = 0.90$ ). The left panel presents the relationship between the soft-threshold and

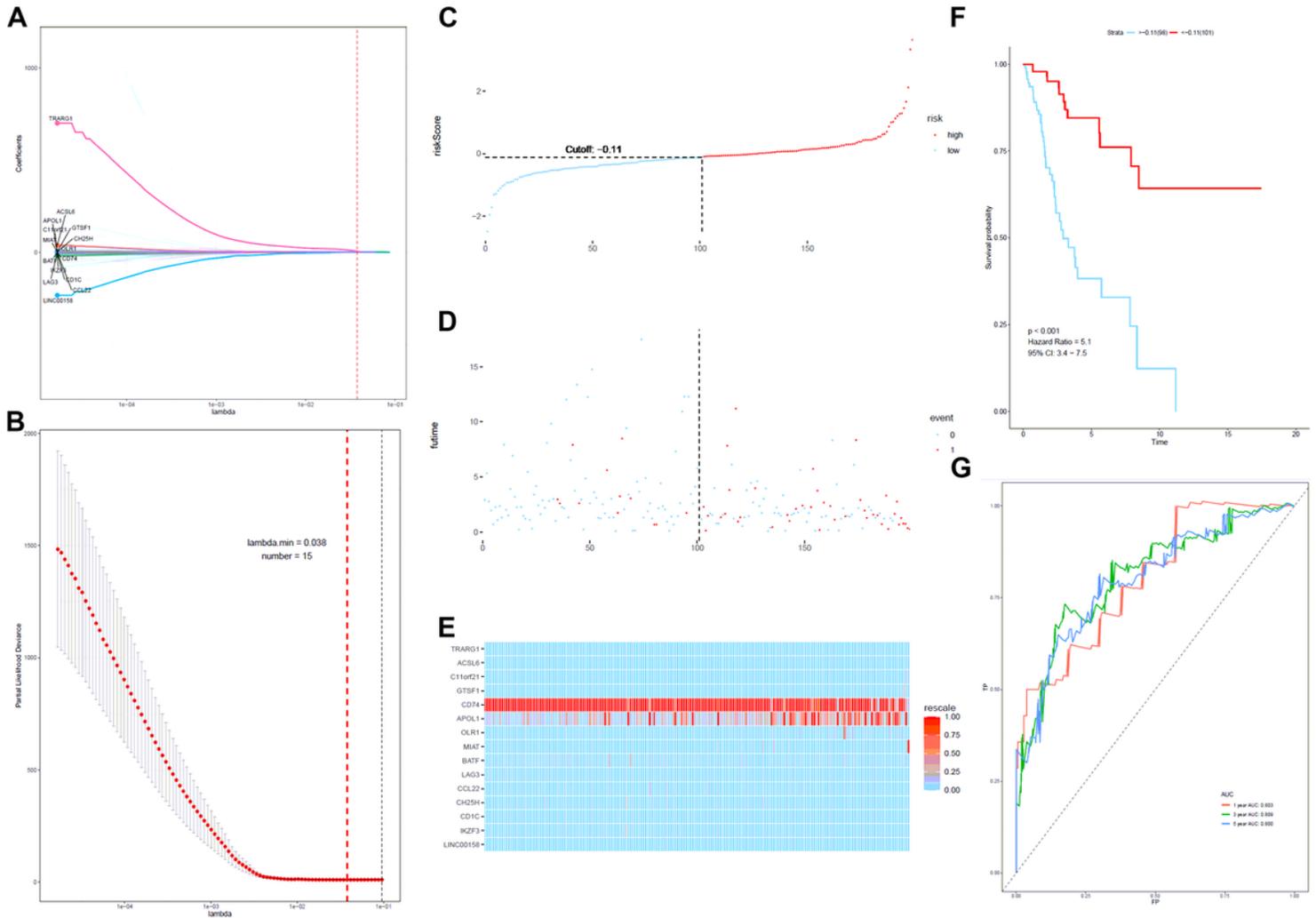
scale-free R2. The right panel presents the relationship between the soft-threshold and mean connectivity. (C) Histogram of connectivity distribution when  $\beta = 8$  and checking the scale free topology when  $\beta = 8$ .



**Figure 5**

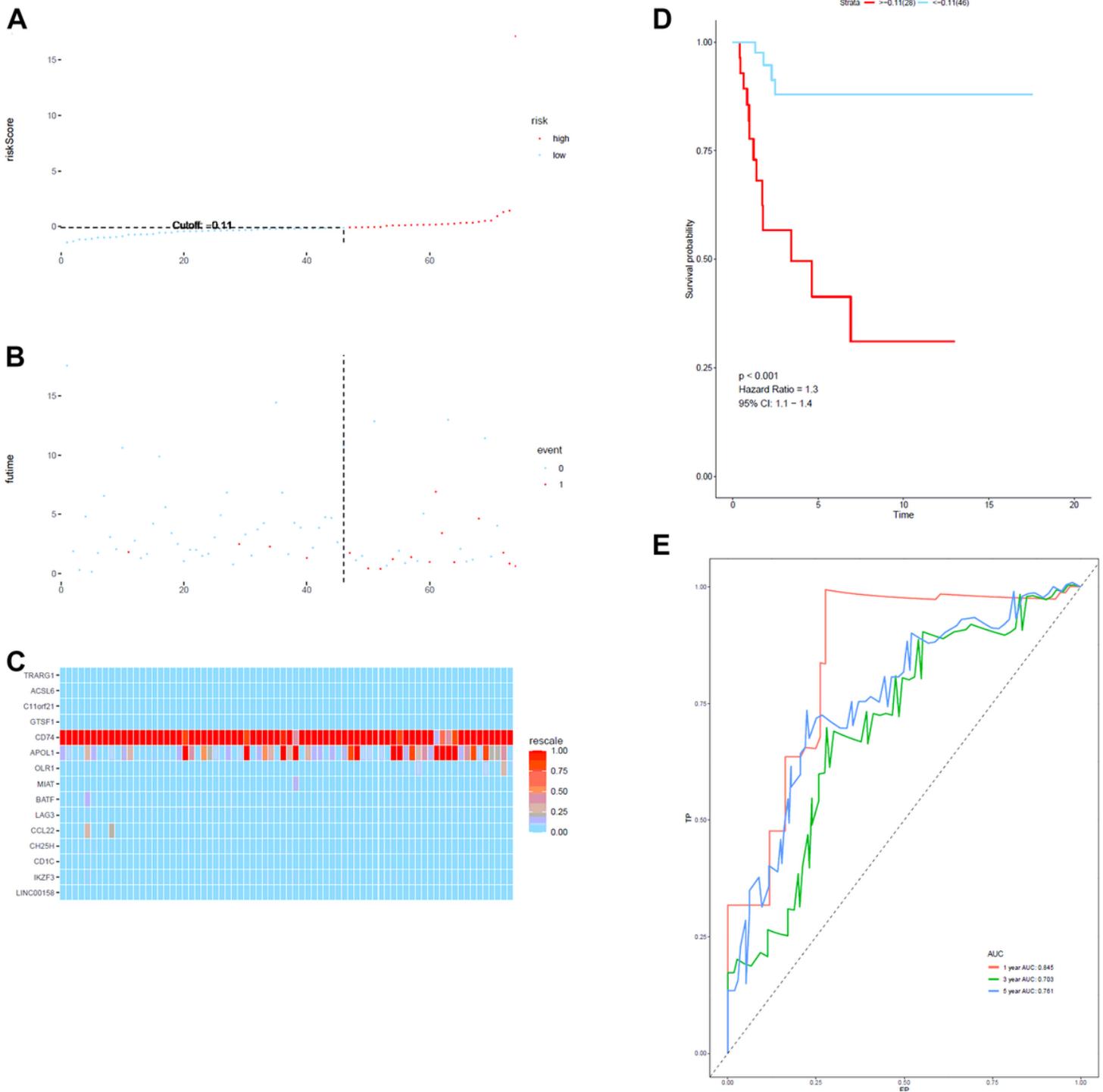
Weighted gene co-expression network in cervical cancer. (A) Clustering of module eigengenes. A cut-line  $\approx 0.25$  was selected for module dendrogram and merged some modules according to dissimilarity of estimated module eigengenes. (B) A dendrogram of the differentially expressed genes clustered based on

different metrics. Each branch in the figure represents one gene, and every color below represents one co-expression module. (C) A heatmap showing the correlation between the gene module and clinical traits. The correlation coefficient in each cell represented the correlation between gene module and the clinical traits. The magenta module showed the highest correlated with the infiltrated immune cells.



**Figure 6**

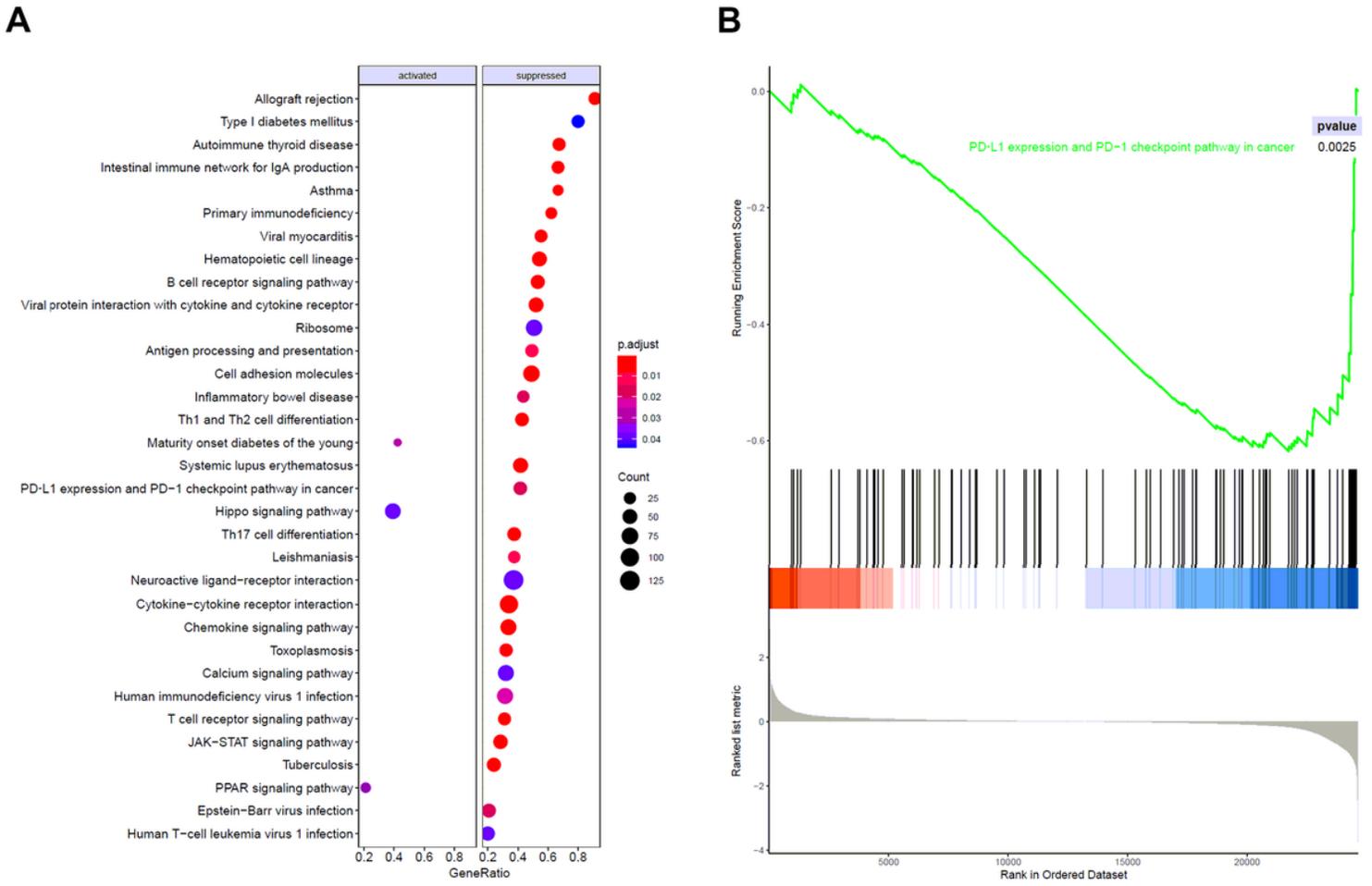
Construction of the prognostic classifier. (A-B) Determination of the number of factors by the LASSO analysis. (C) The distribution of risk score in test cohort. (D) The survival duration and status of patients in test cohort. (E) A heatmap of immune related genes in the classifier in test cohort. (F) Kaplan-Meier curve for cervical cancer patients in test cohort. (G) ROC curve for cervical cancer in test cohort.



**Figure 7**

The distribution of time-dependent ROC curves and Kaplan-Meier survival based on the integrated classifier in the train cohort. (A) The distribution of risk score in train cohort. (B) The survival duration and status of patients in train cohort. (C) A heatmap of immune related genes in the classifier in train cohort. (D) Kaplan-Meier curve for cervical cancer patients in train cohort. (E) ROC curve for cervical cancer in train cohort.





**Figure 9**

GSEA analysis and PD-L1 pathway. (A) GSEA analysis demonstrated 33 KEGG pathways associated with risk score in cervical cancer. (B) PD-L1 expression and PD-1 checkpoint pathway in cancer was one of the 33 pathways that significantly related with risk score.