

RESEARCH

Open Access



# Integrated analysis of N6-methyladenosine- and 5-methylcytosine-related long non-coding RNAs for predicting prognosis in cervical cancer

Jie Gao<sup>1</sup>, Xiuling Zhang<sup>1</sup>, Anqi Xu<sup>1</sup>, Wei Li<sup>2\*</sup> and HaiYan Gao<sup>1\*</sup>

## Abstract

**Background** N6-methyladenosine (m<sup>6</sup>A) and 5-methylcytosine (m<sup>5</sup>C) play a role in modifying long non-coding RNAs (lncRNAs) implicated in tumorigenesis and progression. This study was performed to evaluate prognostic value of m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs and develop an efficient model for prognosis prediction in cervical cancer (CC).

**Methods** Using gene expression data of TCGA set, we identified m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs. Consensus Clustering Analysis was performed for samples subtyping based on survival-related lncRNAs, followed by analyzing tumor infiltrating immune cells (TIICs). Optimal signature lncRNAs were obtained using lasso Cox regression analysis for constructing a prognostic model and a nomogram to predict prognosis.

**Results** We built a co-expression network of 23 m<sup>6</sup>A-related genes, 15 m<sup>5</sup>C-related genes, and 62 lncRNAs. Based on 9 m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs significantly associated with overall survival (OS) time, two molecular subtypes were obtained, which had significantly different OS time and fractions of TIICs. A prognostic model based on six m<sup>6</sup>A- and m<sup>5</sup>C-related signature lncRNAs was constructed, which could dichotomize patients into two risk subgroups with significantly different OS time. Prognostic power of the model was successfully validated in an independent dataset. We subsequently constructed a nomogram which could accurately predict survival probabilities. Drug sensitivity analysis found preferred chemotherapeutic agents for high and low-risk patients, respectively.

**Conclusion** Our study reveals that m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs are associated with prognosis and immune microenvironment of CC. The m<sup>6</sup>A- and m<sup>5</sup>C-related six-lncRNA signature may be a useful tool for survival stratification in CC and open new avenues for individualized therapies.

**Keywords** N6-methyladenosine, 5-methylcytosine, Long non-coding RNAs, Tumor microenvironment, Risk score, Nomogram

\*Correspondence:

Wei Li  
lw200602722@126.com  
HaiYan Gao  
21448840@qq.com

<sup>1</sup>Department of Laboratory Medicine, The Sixth Affiliated Hospital of Harbin Medical University, 998 Aiyang Street, Songbei District, Harbin 150028, Heilongjiang Province, China

<sup>2</sup>Department of Rheumatism and Immunology, The 2nd Affiliated Hospital of Harbin Medical University, 246 Xuefu Road, Nangang District, Harbin 150001, Heilongjiang Province, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Cervical cancer (CC) is a highly heterogeneous cancer and ranks fourth in incidence and mortality in females worldwide, with cervical squamous cell carcinoma (CESC) as the most common type [1]. CC incidence has experienced an evident increase in recent years because of the availability of effective screening programs [2]. Human papilloma virus (HPV) is a well-documented etiological factor of CC, and thus receiving HPV vaccine remains a primary method for CC prevention [3]. CC patients are commonly treated with combinations of surgical resection, chemotherapy, radiotherapy or other therapies [4]. Nonetheless, the long-term survival and prognosis of advanced CC remains unsatisfactory [5]. Therefore, there is an evident interest in gaining a deeper understanding of molecular mechanisms behind CC carcinogenesis and discovering promising prognostic signatures capable of predicting clinical outcomes of patients reliably and accurately.

Long non-coding RNAs (lncRNAs), a group of non-coding RNAs longer than 200 nucleotides, participate in a variety of biological processes involved in tumorigenesis and progression [6, 7]. A significant body of literature has supported important biological roles of lncRNAs in CC progression, invasion and metastasis, and their potentials as prognostic biomarkers [8, 9]. N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) and 5-methylcytosine (m<sup>5</sup>C) are two important forms of modifications to messenger RNA (mRNAs) and non-coding RNAs (ncRNAs), playing a role in diverse functions, such as RNA splicing and translation. Either m<sup>6</sup>A or m<sup>5</sup>C modification has three classes of components: intracellular methyltransferases (“writers”), demethylases (“erasers”), and signal transducers (“readers”) [10]. m<sup>6</sup>A methyltransferase METTL3 up-regulation has been observed in CC, and may serve as prognostic biomarkers [11]. m<sup>6</sup>A reader YTHDF1 plays an oncogenic role and is related to poor prognosis of CC patients [12]. Moreover, growing evidences are highlighting m<sup>6</sup>A-related genes and m<sup>6</sup>A-related lncRNAs as potential prognostic biomarkers in CC [13, 14]. Substantial evidences have revealed that m<sup>5</sup>C modification is implicated in tumorigenesis, cancer migration, and metastasis of a large number of cancers including CC [15–17]. However, there is a lack of studies on characterization of biological roles of m<sup>5</sup>C modification and m<sup>5</sup>C-related lncRNAs in the pathophysiology of CC.

Biological significance of m<sup>6</sup>A and m<sup>5</sup>C modifications implies the potential of m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs to be used for prognostic purposes in CC. To decipher the regulatory networks of m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs in tumorigenesis and progression of CC and explore their prognostic value, we identified m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs in CESC samples downloaded from The Cancer Genome Atlas (TCGA). Out of them,

we determined survival-related lncRNAs and subsequently, revealed two molecular subtypes. Furthermore, with prognostic lncRNAs we constructed a risk score model and a nomogram to predict prognosis in CESC. In addition, we unraveled associations of m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs with tumor infiltrating immune cells (TIICs) in tumor microenvironment (TME) as well.

## Methods

### Data sources

Gene expression data (log<sub>2</sub> (FPKM+1), Illumina HiSeq 2000 RNA Sequencing platform) of 291 CC samples and corresponding clinicopathological characteristics were downloaded from TCGA database (<https://portal.gdc.cancer.gov/>) and were used as the training set. The validated set of this study used GSE44001 dataset [18] downloaded from gene expression omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>), including gene expression data and survival information of 300 CC samples (Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip).

### Identification of m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs

According to a recently published study [19], 23 m<sup>6</sup>A-related genes and 15 m<sup>5</sup>C-related genes were identified (Supplemental Table 1). Expression levels of genes associated with m<sup>5</sup>C and m<sup>6</sup>A were extracted from the TCGA dataset. Using the `cor` function in R3.6.1 (<http://77.66.12.57/R-help/cor.test.html>), Pearson correlation coefficients (PCC) were calculated between the expression levels of 23 m<sup>6</sup>A-related genes, 15 m<sup>5</sup>C-related genes, and all annotated lncRNAs. The m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs were identified using a threshold of  $|PCC| > 0.3$  and a significance  $p$ -value  $< 0.05$ . Finally, we identified lncRNAs significantly associated with both m<sup>5</sup>C and m<sup>6</sup>A genes and selected those that were significantly correlated with both. A co-expression network between m<sup>5</sup>C and m<sup>6</sup>A genes and their commonly associated lncRNAs was constructed and visualized using Cytoscape 3.6.1 [20] (<https://cytoscape.org/>).

### Consensus clustering analysis

Based on the expression levels of m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs in CESC tumor samples from the TCGA training dataset, univariate Cox regression analysis was performed using the survival package Version 2.41-1 (<http://bioconductor.org/packages/survival/>) in R 3.6.1 to identify lncRNAs significantly associated with survival prognosis, with a significance threshold set at  $P < 0.05$ . Based on expression levels of these survival-related lncRNAs, CC samples in the TCGA set were separated into different subtypes by performing Consensus Clustering Analysis using ConsensusClusterPlus package [21] (version 1.54.0, <http://www.bioconductor.org/packages/>

[release/bioc/html/ConsensusClusterPlus.html](https://www.bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html)) in R language. The parameters were set as  $d$ ,  $\max K=10$ ,  $\text{reps}=50$ ,  $\text{pItem}=0.8$ ,  $\text{pFeature}=1$ ,  $\text{clusterAlg}=\text{"hc"}$ ,  $\text{distance}=\text{"pearson"}$ . Kaplan-Meier (KM) survival curves were plotted for different subtypes using *survival* package.

### Analysis of immune infiltration

**R3.6.1 GSEA** (<http://www.bioconductor.org/packages/release/bioc/html/GSEA.html>) Version 1.36.3, which was based on single sample gene set enrichment analysis (ssGSEA) algorithm was adopted to evaluate the proportion of TIICs in TCGA samples. Subsequently, we performed differential distribution comparisons of TIIC proportions across different subtypes using group t-tests in R3.6.1, with a significance threshold set at a p-value of less than 0.05. Finally, we analyzed the correlation between the expression levels of prognostically significant lncRNAs and the TIIC types that showed significant distribution differences.

### Establishment of m<sup>6</sup>A- or m<sup>5</sup>C-related lncRNA signature for survival prediction

Briefly, out of the survival-related lncRNAs, the optimal prognostic lncRNAs were selected by performing Lasso Cox regression analysis [22] using *lars* package (<https://cran.r-project.org/web/packages/lars/index.html>) in R. The optimal lambda value was determined by performing a 10-fold cross-validation. Based on LASSO coefficients and expression levels of the optimal signature lncRNAs, prognostic model was constructed using the following formula:

$$\text{Risk score (RS)} = \sum \text{Coef}_{\text{lncRNAs}} \times \text{Exp}_{\text{lncRNAs}}$$

Where  $\text{Coef}_{\text{lncRNAs}}$  represents the estimated LASSO coefficient of lncRNAs;  $\text{Exp}_{\text{lncRNAs}}$  represents expression level of lncRNAs.

With median risk score as the cutoff, samples were divided into two risk subgroups in the TCGA set. KM survival curves were plotted for different risk subgroups and compared using log-rank tests. Accuracy, sensitivity and specificity of the prognostic model were assessed using receiver operating characteristic (ROC) curve. Predictive performance of the prognostic score model was validated using GSE44001 dataset.

### Identification of independent prognostic factors and nomogram construction

Clinical factors of patients in the TCGA set were subjected to uni-variable and multi-variable Cox regression analysis to identify independent prognostic factors using *survival* package in R (log-rank p-value < 0.05). Nomogram [23] that could provide a predicted probability was generated using *rms* package (version 5.1-2) in R (<https://cran.r-project.org/web/packages/rms/index.html>).

### Drug sensitivity analysis

Based on gene expression profiles of CC samples in the TCGA set and chemotherapeutic agents downloaded from Genomics of Drug Sensitivity in Cancer (GDSC) [24] database (<https://www.cancerrxgene.org/>), sensitivity of CESC patients in high-risk and low-risk groups to chemotherapeutic drugs was assessed by predicting half-maximal inhibitory concentration (IC<sub>50</sub>) using *pRRophetic* [25] package in R (<http://127.0.0.1:22402/doc/html/Search?objects=1&port=22402>). IC<sub>50</sub> values were compared between different risk groups using Student's t test to assess differential therapeutic effects of chemotherapeutic drugs.

### Function annotation analyses

Differentially expressed genes (DEGs) were screened between different risk groups of the TCGA set using *limma* [26] package in R (<https://bioconductor.org/packages/release/bioc/html/limma.html>) with  $\text{FDR} < 0.05$  and  $|\log_2 \text{FC}| > 0.5$  as the significance cutoff. Gene ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were conducted using The Database for Annotation, Visualization and Integrated Discovery (DAVID) tool (<https://david.ncifcrf.gov/>, version 6.8) with the strict cutoff at  $\text{FDR} < 0.05$ .

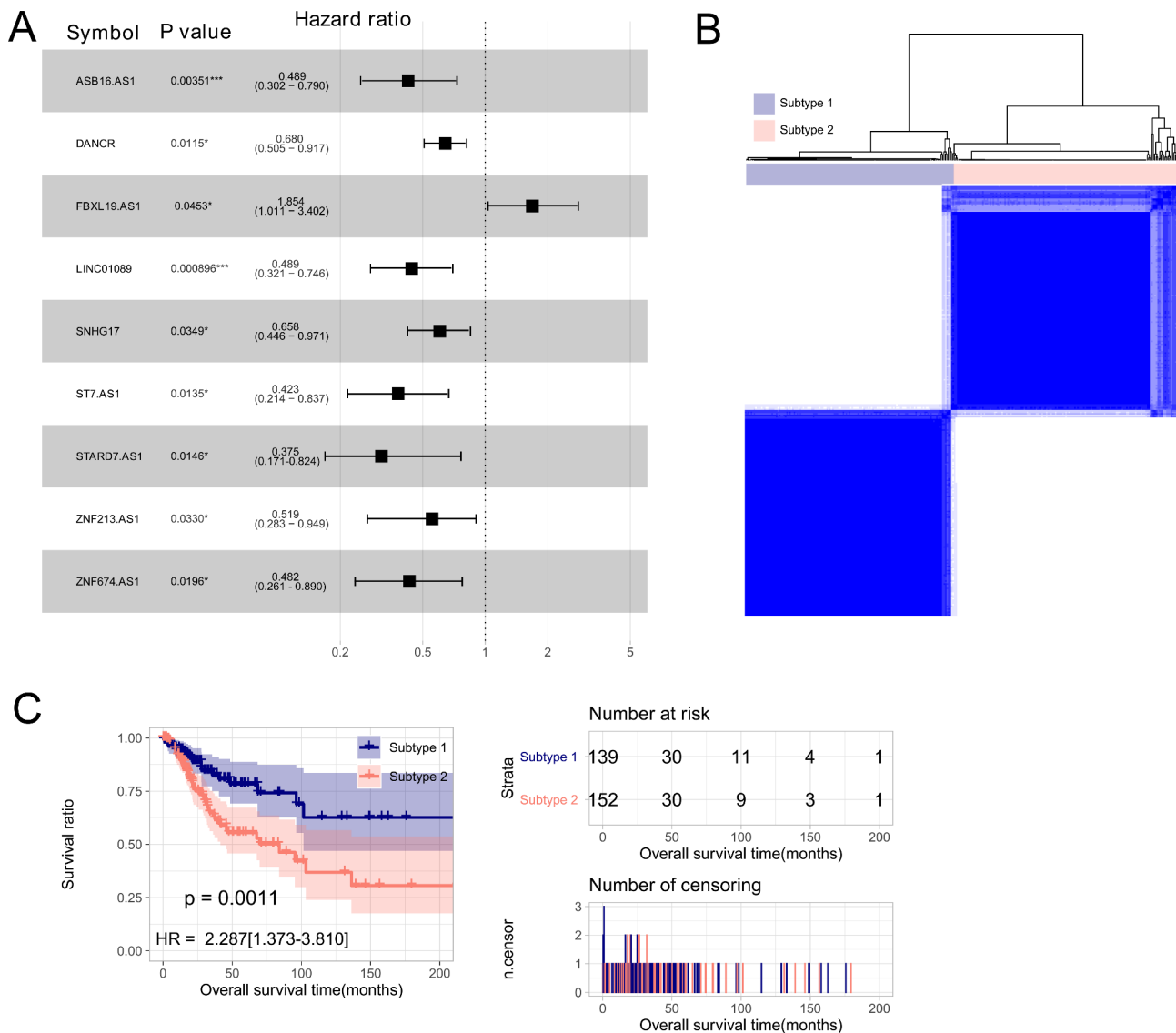
## Results

### A co-expression network was constructed including 23 m<sup>6</sup>A genes, 15 m<sup>5</sup>C genes and 343 lncRNAs

As mentioned in Method section, we recognized 23 m<sup>6</sup>A-related genes and 15 m<sup>5</sup>C-related genes, and extracted 343 common lncRNAs from both the TCGA set and GSE44001 dataset. By performing Pearson correlation analysis, we obtained 202 m<sup>6</sup>A genes-lncRNAs pairs and 264 m<sup>5</sup>C genes-lncRNAs pairs, which shared 62 common lncRNAs. With these m<sup>6</sup>A genes, m<sup>5</sup>C genes and 62 m<sup>6</sup>A and m<sup>5</sup>C-related lncRNAs, a co-expression network was constructed (Figure S1). Out of the 62 m<sup>6</sup>A and m<sup>5</sup>C-related lncRNAs, 9 lncRNAs were significantly related to OS time in uni-variate Cox regression analysis (Fig. 1A).

### Two molecular subtypes were obtained based on nine survival-related lncRNAs

Based on expression levels of the 9 survival-related lncRNAs in the TCGA set, samples were categorized into subtype 1 ( $N=139$ ) and subtype 2 ( $N=152$ ) through performing Consensus Clustering Analysis (Fig. 1B). KM survival curves showed that subtype 1 showed significantly longer OS time than subtype 2 ( $p=0.0011$ ,  $\text{HR}=2.287$  (1.373–3.810), Fig. 1C).



**Fig. 1** Identification of two subtypes with significantly different survival time. **(A)** Forest plot of 9 survival-related lncRNAs in uni-variable Cox regression analysis; **(B)** samples are classified into two subtypes (subtype 1 and 2) by Consensus Clustering Analysis; **(C)** Kaplan- Meier curves of two subtypes

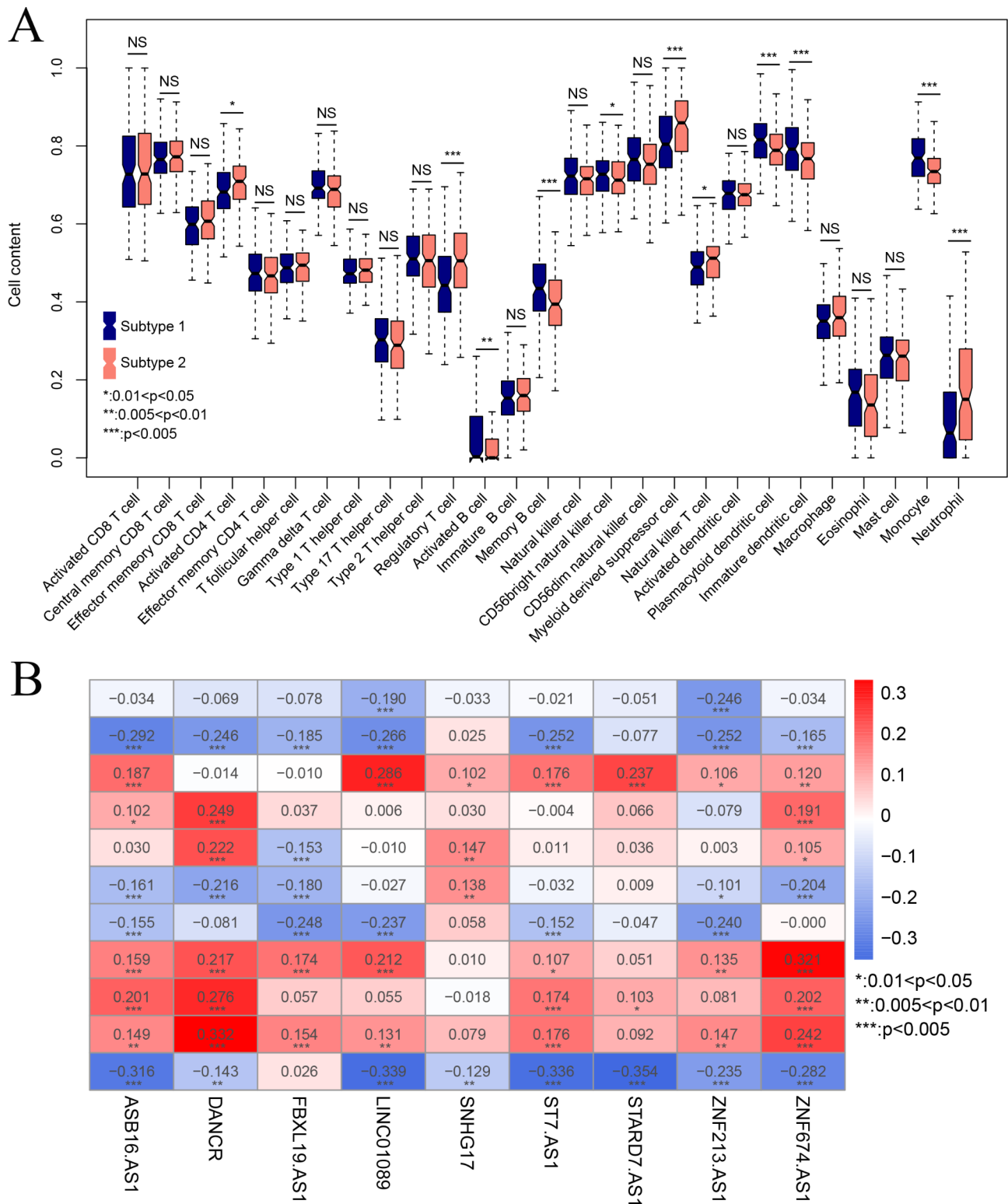
### Characterization of tumor infiltrating immune cells in different subtypes

In order to investigate the characteristics of TME and TIICs of the two molecular subtypes, we analyzed and compared fractions of different TIICs between them. As shown in Fig. 2A, subtype 1 had significantly lower levels of activated CD4+T cells, regulatory T cells, myeloid derived suppressor cells, memory B cells, natural killer T cells and neutrophils, and higher levels of activated B cells, plasmacytoid dendritic cells, CD56+bright natural killer cells, immature dendritic cells, and monocytes ( $p$ -value<0.05). Moreover, PCCs between the 9 survival-related lncRNAs and 11 immune cells were shown in a heatmap (Fig. 2B). Fractions of activated CD4+T cells, and neutrophils were negatively correlated with

expression levels of all nine lncRNAs. On the contrary, fractions of plasmacytoid dendritic cells, and monocytes were positively correlated with expression levels of all nine lncRNAs.

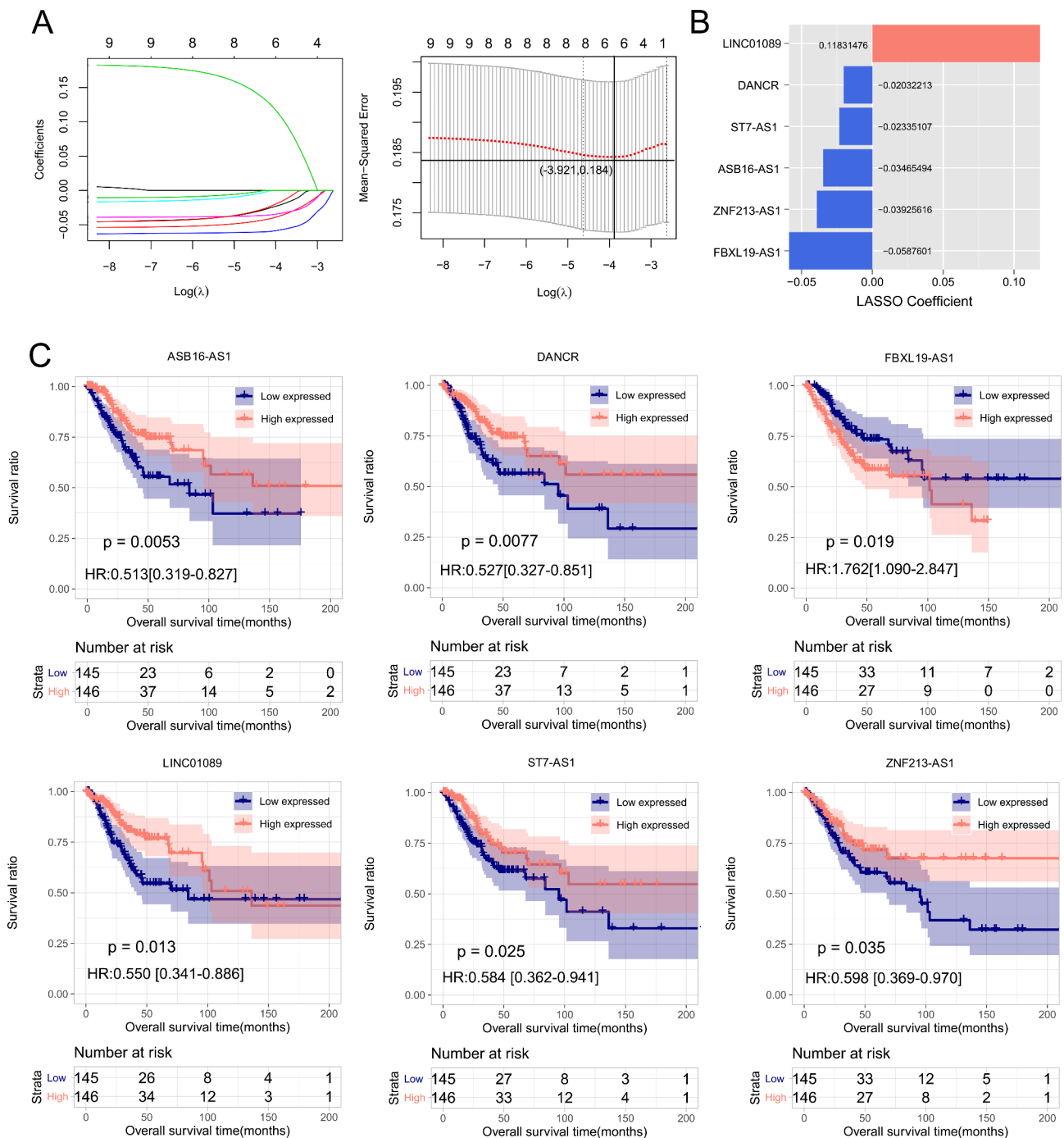
### A prognostic model based on six m<sup>6</sup>A and m<sup>5</sup>C-related lncRNAs was constructed and successfully validated in an independent set

LASSO Cox regression was performed based on the aforementioned nine survival-related lncRNAs. As a result, six optimal lncRNAs were obtained, consisting of ASB16-AS1, DANCR, FBXL19-AS1, LINC01089, ST7-AS1, and ZNF213-AS1 (Fig. 3A-B). Based on LASSO coefficients and expression levels of the six optimal



**Fig. 2** Characterization of tumor infiltrating immune cells in different subtypes. **(A)** Comparative analysis of fractions of tumor infiltrating immune cells between two subtypes; **(B)** Pearson correlation coefficients of 9 survival-related lncRNAs with 11 infiltrating immune cells





**Fig. 3** LASSO Cox regression analysis. **(A)** Performance indicators of LASSO regression analysis **(B)** LASSO coefficients of 6 signature lncRNAs. **(C)** Kaplan-Meier curves for patients stratified by each signature lncRNA. Patients are divided into high-expression group and low-expression groups according to median expression level of a signature lncRNA (ASB16-AS1, DANCR, LINC01089, ST7-AS1, FBXL19-AS1 or ZNF213-AS1)

lncRNAs in the TCGA set, risk score was calculated for each sample using the following formula:

$$RS = (-0.0587601) * Exp_{LINC01089} + (-0.03925616) * Exp_{DANCR} + (-0.02335107) * Exp_{ST7-AS1} + (-0.02032213) * Exp_{ASB16-AS1} + (-0.03465494) * Exp_{ZNF213-AS1} + (0.11831476) * Exp_{FBXL19-AS1}$$

With regard to ASB16-AS1 ( $p=0.0053$ ), DANCR ( $p=0.0077$ ), LINC01089 ( $p=0.013$ ), ST7-AS1 ( $p=0.025$ ), or ZNF213-AS1 ( $p=0.035$ ), high-expressed patients had significantly longer OS time compared to low-expressed patients (Fig. 3C). Conversely, patients with low-expressed FBXL19-AS1 had longer OS time in comparison with patients with high-expressed FBXL19-AS1

( $p=0.019$ , Fig. 3C). In the TCGA set, with median risk score (-0.2913) as cutoff, tumor samples were separated into a high-risk group and a low-risk group (Fig. 4A-B). As depicted in Fig. 4C-D, OS time was significantly longer in low-risk patients compared to high-risk patients ( $p\text{-value}=3.494\text{e-}06$ ,  $\text{HR}=3.310$  (1.937–5.565)), with an AUC of 0.872 (0.841, 0.873). Similarly, the risk score model was applied to GSE44001 dataset. As shown in Fig. 4E-H, patients were dichotomized into high-risk and low-risk patients with significantly different OS time ( $p\text{-value}=0.023$ ,  $\text{HR}=2.145$  (1.094–4.209)) and an AUC of 0.795 (0.873, 0.705).

Sankey diagram was used to analyze the relationships of the above-mentioned two subtypes, two risk subgroups and vital status of CC patients in the TCGA set. The majority of subtype 1 samples were classified into the low-risk group and had better prognosis, while the majority of subtype 2 samples were classified into the high-risk group and had poor prognosis (Fig. 5A). Risk scores were significantly decreased in subtype 1 samples compared to subtype 2 samples (Fig. 5B,  $p\text{-value}<2.22\text{e-}16$ ). These results reveal that the subtyping results based on the 9 survival-related lncRNAs were consistent with the risk stratification results based on the six- $\text{m}^6\text{A}$  and  $\text{m}^5\text{C}$ -related lncRNAs signature.

#### A nomogram was constructed integrating risk score with prognostic factors

Uni-variable and multi-variable Cox regression analysis was carried out for clinical factors of patients in the TCGA set. As shown in Table 1, Pathologic N ( $p\text{-value}=2.19\text{E-}03$ ,  $95\%\text{CI}=1.409\text{--}5.596$ ), Pathologic T ( $p\text{-value}=2.31\text{E-}05$ ,  $95\%\text{CI}=1.371\text{--}2.452$ ), Pathologic Stage ( $p\text{-value}=2.86\text{E-}04$ ,  $95\%\text{CI}=1.195\text{--}1.852$ ), Tobacco smoking history ( $p\text{-value}=6.30\text{E-}03$ ,  $95\%\text{CI}=1.081\text{--}1.904$ ) and RS model ( $p\text{-value}=3.49\text{E-}06$ ,  $95\%\text{CI}=1.937\text{--}5.656$ ) were significantly associated with OS time of CC patients. Furthermore, RS model was an independent prognostic factor ( $p\text{-value}=4.34\text{E-}02$ ,  $95\%\text{CI}=1.027\text{--}5.861$ ).

Combining risk score model with pathologic N, pathologic T, pathologic Stage, and tobacco smoking history, a nomogram was built to predict survival probabilities in CC patients (Fig. 5C). Calibration curves showed that 1-year C-index was 0.771, with a  $p\text{-value}$  of  $1.258\text{e-}02$ ; 3-year C-index=0.743, with a  $p\text{-value}$  of  $2.295\text{e-}05$ ; 5-year C-index=0.772, with a  $p\text{-value}$  of  $6.036\text{e-}09$  (Fig. 5D). Moreover, in ROC curves, 1-year, 3-year and 5-year AUC was 0.856 (0.750, 0.875), 0.875 (0.906, 0.833), 0.869 (0.900, 0.762), separately (Fig. 5E). These results collectively suggest that the nomogram model has high accuracy in predicting 1-, 3-, 5-year survival in CC patients.

#### Differential sensitivity of high-risk and low-risk patients to chemotherapeutic agents

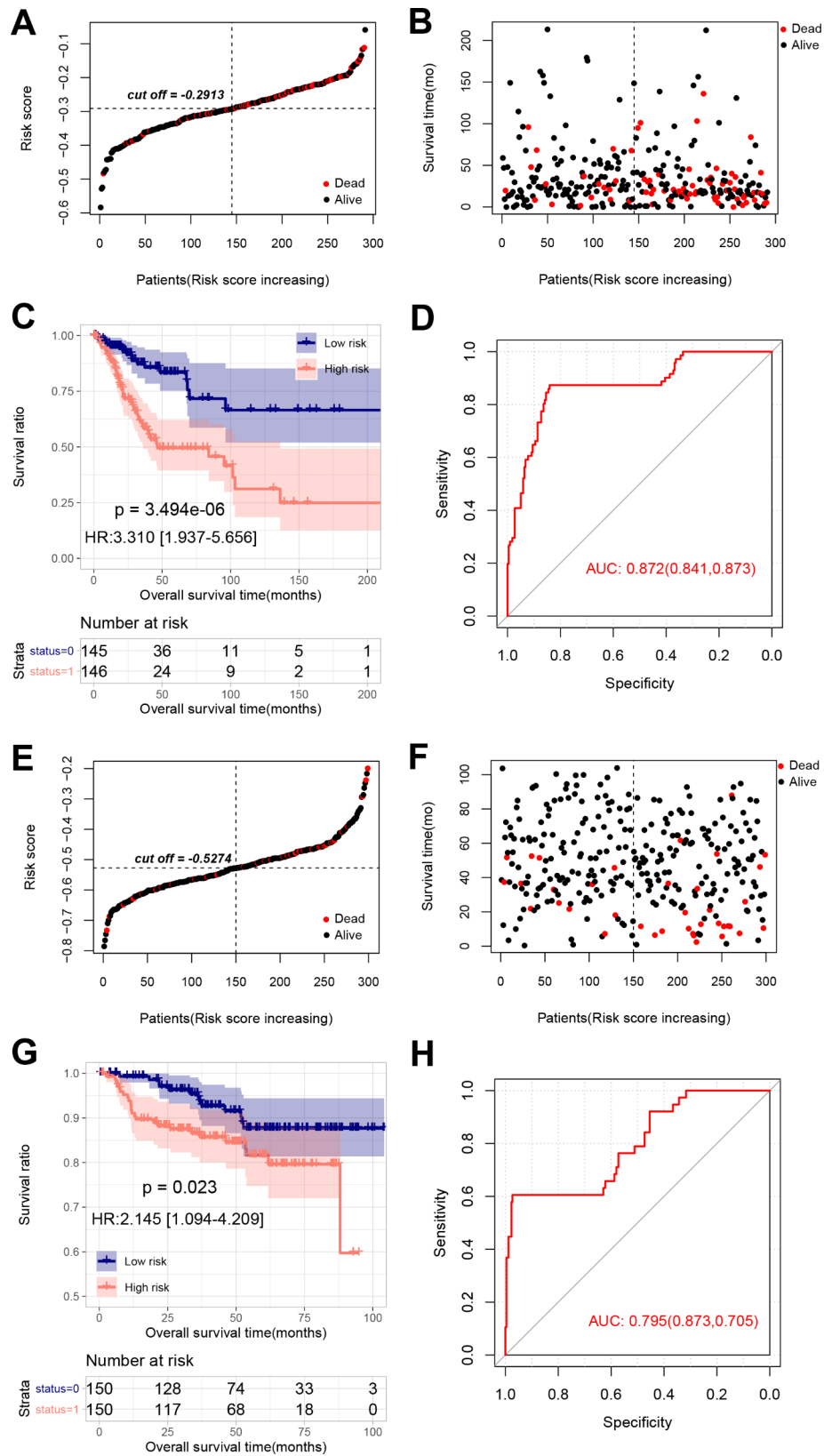
We compared sensitivities of high-risk and low-risk patients to 21 chemotherapeutic agents (Axitinib, AZD6482, AZD7762, AZD8055, Bortezomib, Camptothecin, Cisplatin, Cytarabine, Dasatinib, Erlotinib, Gefitinib, Gemcitabine, GSK269962A, Lapatinib, Nilotinib, Paclitaxel, Rapamycin, Sorafenib, Vinblastine, Vinorelbine, and Vorinostat) in the TCGA set. Low-risk patients were more sensitive to AZD8055 ( $p\text{-value}=0.0012$ ), Rapamycin ( $p\text{-value}=6.2\text{e-}07$ ) and Vorinostat ( $p\text{-value}=0.046$ ), whereas high-risk patients were more sensitive to Dasatinib ( $p\text{-value}=0.00021$ , Fig. 6A).

#### Functional annotation of the DEGs between two risk groups in the TCGA set

A total of 624 DEGs ( $\text{FDR}<0.05$  and  $|\log_2\text{FC}|>0.5$ ) were identified between the two risk subgroups of the TCGA set. These genes were significantly enriched in 25 GO biological processes and 10 KEGG signaling pathways, such as inflammatory response, positive regulation of interleukin-6 production, Calcium signaling pathway, MAPK signaling pathway, and IL-17 signaling pathway. Top 10 BP terms and KEGG signaling pathways are showcased in Fig. 6B.

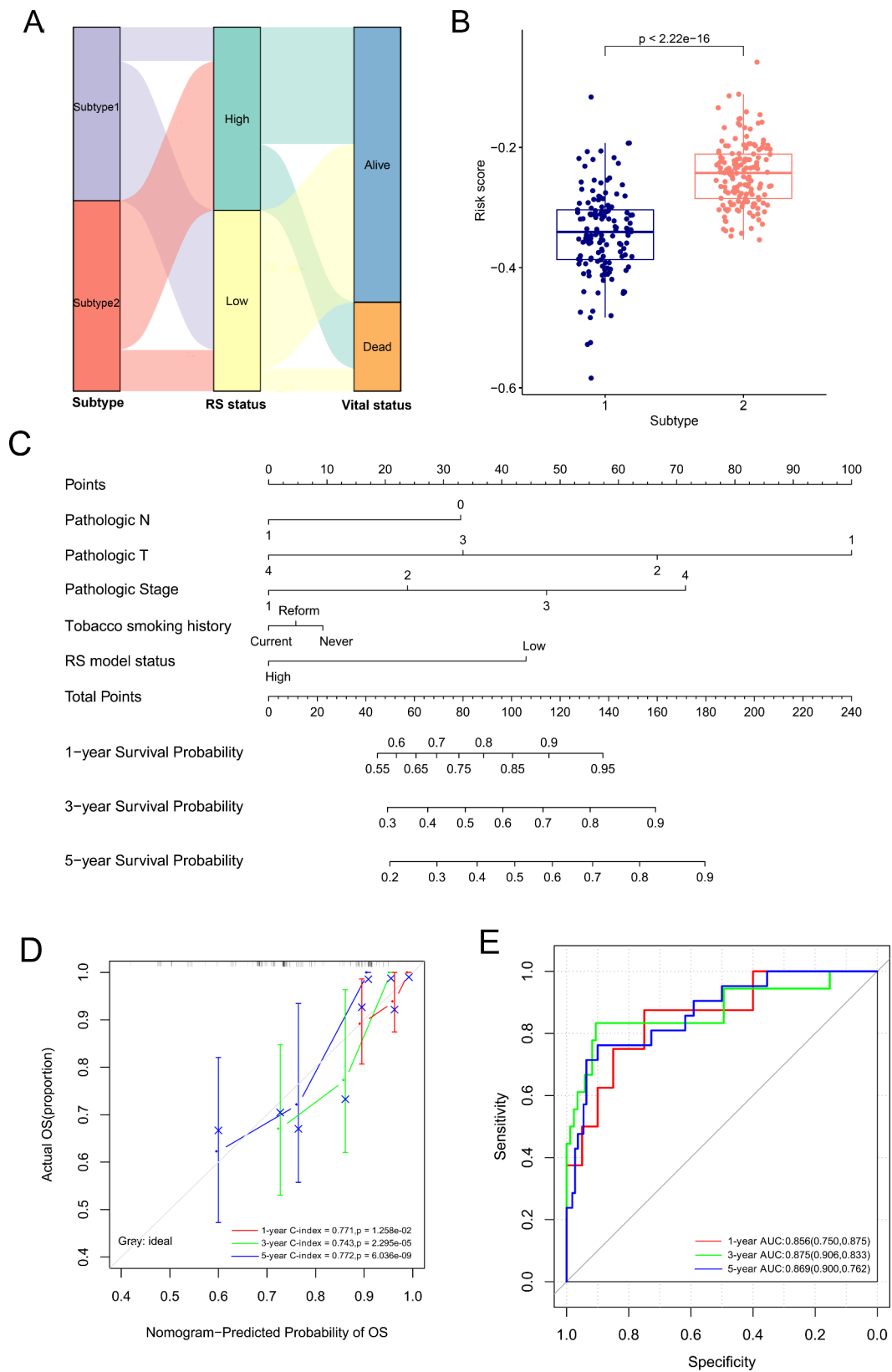
#### Discussion

lncRNAs play critical regulatory roles in progression, metastasis, and prognosis of CC [27]. As two well-studied forms of methylation modifications,  $\text{m}^6\text{A}$  and  $\text{m}^5\text{C}$  modifications that contribute to cancer occurrence and development are gaining increasing attention [28]. Our study underlined the regulatory mechanisms of  $\text{m}^6\text{A}$  and  $\text{m}^5\text{C}$  modifications-related lncRNAs in CC. We identified 62  $\text{m}^6\text{A}$ - and  $\text{m}^5\text{C}$ -related lncRNAs in CC samples, and constructed a co-expression network with these lncRNAs, 23  $\text{m}^6\text{A}$ -related genes and 15  $\text{m}^5\text{C}$ -related genes. Based on nine survival-related lncRNAs we obtained two molecular subtypes with distinctive TME characteristics. Noticeably, we constructed a prognostic model based on six  $\text{m}^6\text{A}$ - and  $\text{m}^5\text{C}$ -related signature lncRNAs, which separated patients into high-risk and low-risk groups with significantly different OS time. Robustness of the prognostic model was successfully validated in an independent cohort. Low-risk patients showed higher sensitivity to AZD8055, Rapamycin and Vorinostat, whereas high-risk patients showed higher sensitivity to Dasatinib. Furthermore, a nomogram based on risk score and prognostic clinical factors was built and showed robust and accuracy performance in predicting prognosis in CC patients. These results suggest that the prognostic model based on six  $\text{m}^6\text{A}$ - and  $\text{m}^5\text{C}$ -related lncRNA is a promising tool for prognosis stratification of CC patients and provide new hints to direct individualized therapies.



**Fig. 4** Performance of the risk model in TCGA set and validation set. Risk score distribution (A), relationship of risk score and survival time (B), KM curves of high-risk and low risk patients (C) and ROC curve (D) of patients in high-risk and low-risk subgroups in the TCGA set. Risk score distribution (E), relationship of risk score and survival time (F), KM curves of high-risk and low risk patients (G) and ROC curve (H) of patients in high-risk and low-risk subgroups in the validation set (GSE44001)





**Fig. 5** Associations of molecular subtypes with risk score and nomogram construction. **(A)** Sankey diagram of molecular subtypes, risk score status and vital status of patients; **(B)** The risk scores of two subtypes were significantly different; **(C)** a nomogram for predicting 1-, 3- and 5-year survival probability in the TCGA set; **(D)** Calibration plots of the nomogram; **(E)** ROC curves for predicting 1-, 3- and 5-year survival probability

**Table 1** Uni-variable and multi-variable Cox regression analysis of clinical factors

Clinical characteristics	Uni-variables cox			Multi-variables cox		
	HR	95%CI	P-value	HR	95%CI	P
Age(years, mean ± SD)	1.017	0.999–1.035	5.68E-02	-	-	-
Pathologic_M (M0/M1)	3.671	1.229–10.96	2.56E-01	-	-	-
Pathologic_N (N0/N1)	2.808	1.409–5.596	2.19E-03	2.222	0.969–5.097	5.95E-02
Pathologic_T (T1/T2/T3/T4)	1.833	1.371–2.452	2.31E-05	1.852	0.890–3.853	9.91E-02
Pathologic_stage (I/II/III/IV)	1.488	1.195–1.852	2.86E-04	0.608	0.305–1.212	1.58E-01
Neoplasm histologic grade (G1/G2/G3)	0.976	0.643–1.535	9.76E-01	-	-	-
Tobacco smoking history (Never/Reform/Current)	1.435	1.081–1.904	6.30E-03	1.172	0.740–1.856	4.98E-01
Prognostic model (High/Low)	3.31	1.937–5.656	3.49E-06	2.454	1.027–5.861	4.34E-02

SD, standard deviation; HR, hazard ratio; CI, confidence interval

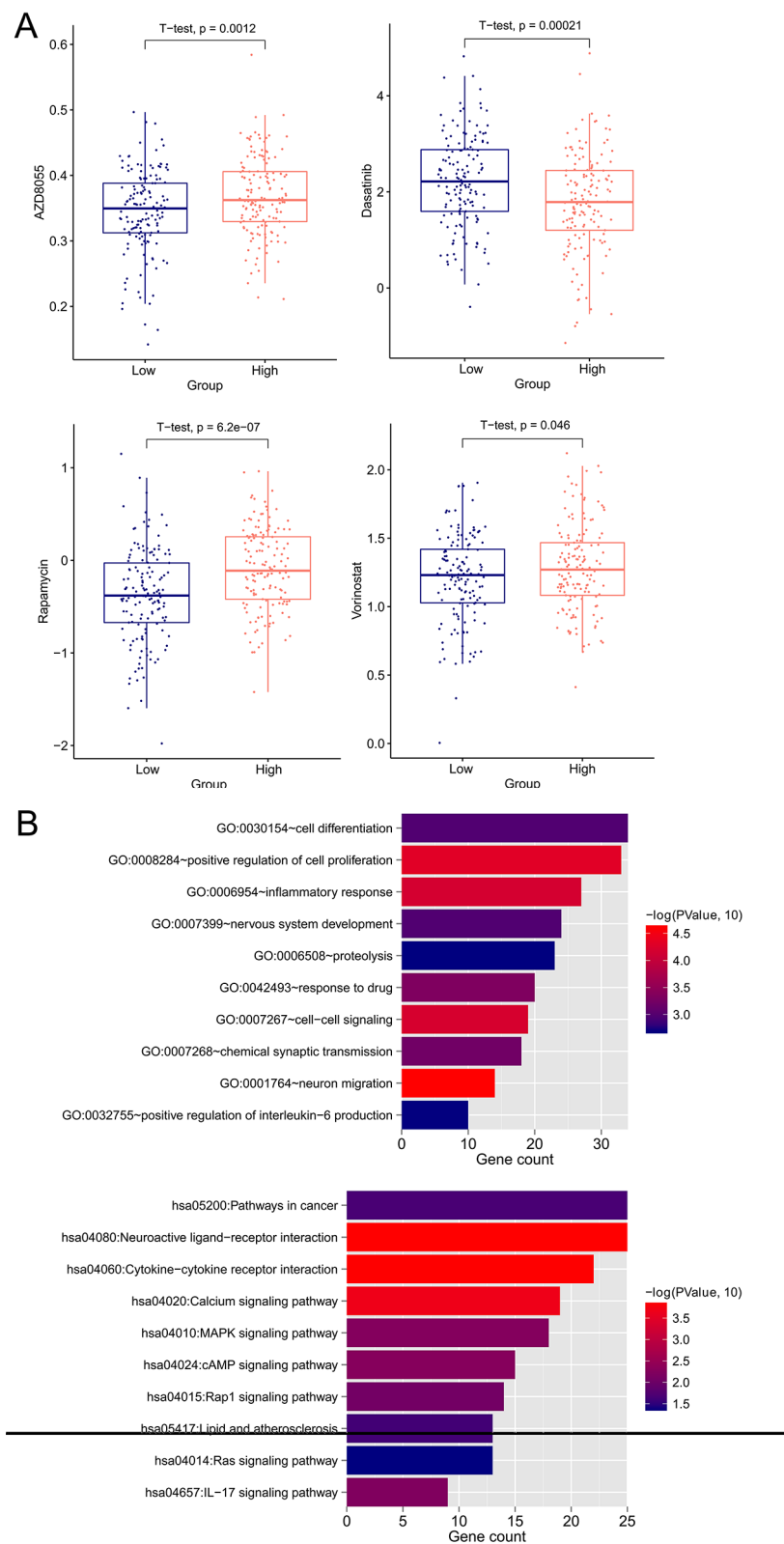
Previous studies suggest that m<sup>6</sup>A RNA methylation regulators play critical roles.

in the malignant progression of CC, and have prognostic implications [29, 30]. Moreover, m<sup>6</sup>A-related lncRNAs may be promising prognostic biomarkers for CC [14]. The m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs signature obtained by our study contained lncRNA ASB16-AS1, DANCR, FBXL19-AS1, LINC01089, ST7-AS1, and ZNF213-AS1. LncRNA ASB16-AS1 is a promising pan-cancer prognostic biomarker, with an association with immune infiltration [31]. There is evidence that lncRNA ASB16-AS1 expression is up-regulated in CC, strengthening cell proliferation, and migration *via* Wnt/ $\beta$ -catenin signal pathway [32]. LncRNA DANCR has prognostic significance in human cancers, and correlates with worse prognosis [33]. Substantial evidences have supported an oncogenic role of lncRNA DANCR in CC, which promotes CC proliferation, metastasis and progression [34, 35]. Emerging studies show that lncRNA FBXL19-AS1 promotes CC proliferation and metastasis [36, 37]. Besides, prognostic potential of lncRNA FBXL19-AS1 has been suggested for CC [38]. LncRNA LINC01089 exerts suppressive effects on the development of CC [39]. LncRNA ST7-AS1 overexpression exhibits positive correlations with shorter OS time, a higher frequency of lymph node metastasis and deeper cervical invasion [40]. Our study consistently found involvement of the 5 m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs in the progression of CC. Elevated expression of lncRNA ZNF213-AS1 plays a role in differentiation and proliferation of acute myeloid leukemia and low-grade gliomas, with an insignificant association with poor prognostic outcome [41]. However, biological roles of lncRNA ZNF213-AS1 in CC remain undefined. Our study indicates that these m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs may serve as prognostic biomarkers for CC.

Our study revealed two molecular subtypes based on express levels of prognostic lncRNAs. The majority of subtype 1 samples had low-risk scores and longer OS time, whereas the majority of subtype 2 samples had high-risk scores and shorter OS time. TME has been recognized as a principal component of CC tumorigenesis

and development, and influences prognosis and treatment of patients [42, 43]. Plasmacytoid dendritic cells play an immunosuppressive role in TME and promote tumor growth [44]. Our study unveiled distinctive TME characteristics of the two subtypes. Subtype 1 had significantly higher levels of activated B cells, CD56+ bright natural killer cells, plasmacytoid dendritic cells, and lowers levels of myeloid derived suppressor cells. It implies that stronger anti-tumor immune function is an important contributor to better prognosis of subtype 1 compared to subtype 2. These findings indicate that m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs could serve as effective prognostic biomarkers and predictors for clinical outcomes and immunotherapeutic responses in CC patients. Additionally, our study found that the low-risk patients were more sensitive to AZD8055, Rapamycin and Voriostat, whereas high-risk patients were more sensitive to Dasatinib. It offers valuable data to facilitate individualized clinical treatments for CC patients. Our study also unraveled that various inflammation-related biological processes and pathways, such as inflammatory response, positive regulation of interleukin-6 production, and IL-17 signaling pathway, as well as Calcium signaling pathway and MAPK signaling pathway might participate in the regulatory mechanisms of m<sup>6</sup>A and m<sup>5</sup>C modifications-related lncRNAs in CC in concordance with previous findings [45–47].

This study has several limitations. First, since the data were analyzed from the TCGA and GEO databases, there is a lack of validation through experiments. Second, the prognostic value of the RiskScore model requires validation with additional external datasets before it can be applied clinically. Finally, the potential molecular mechanisms of these m<sup>6</sup>A/m<sup>5</sup>C-related lncRNAs in cervical cancer remain unclear, and we plan to conduct *in vitro* or *in vivo* experiments in future research to validate our findings.



**Fig. 6** Drug sensitivity and functional enrichment analysis. **(A)** IC50 values of AZD8055, Rapamycin, Vorinostat and Dasatinib between high-risk and low-risk patients in the TCGA set. **(B)** Top 10 GO biological processes and top 10 KEGG signaling pathways

## Conclusion

Taken together, our study comprehensively analyzed interactions of m<sup>6</sup>A and m<sup>5</sup>C modifications with lncRNAs in CC samples, and established a prognostic model based on a signature of six m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs, which could be used to predict outcome of patients. These signature lncRNAs had close associations with TIICs and cancer prognosis, highlighting great promises as prognostic biomarkers and therapeutic targets for CC. Our study expands the knowledge concerning involvement of m<sup>6</sup>A- and m<sup>5</sup>C-related lncRNAs in CC pathogenesis and has clinical implications for risk stratification and personalized therapeutics of patients.

## Abbreviations

(m <sup>6</sup> A)	N <sup>6</sup> methyladenosine
(m <sup>5</sup> C)	5 methylcytosine
(lncRNAs)	long non-coding RNAs
(CC)	cervical cancer
(TIICs)	tumor infiltrating immune cells
(OS)	overall survival
(CESC)	cervical squamous cell carcinoma
(HPV)	Human papilloma virus
(mRNAs)	messenger RNA
(ncRNAs)	non-coding RNAs
(TCGA)	The Cancer Genome Atlas
(TIICs)	tumor infiltrating immune cells
(GEO)	gene expression omnibus
(KM)	Kaplan-Meier
(ROC)	receiver operating characteristic
(GDSC)	Genomics of Drug Sensitivity in Cancer
(DEGs)	Differentially expressed genes
(GO)	Gene ontology
(KEGG)	Kyoto Encyclopedia of Genes and Genomes
(DAVID)	Database for Annotation, Visualization and Integrated Discovery

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s41065-024-00336-w>.

Supplementary Material 1

Supplementary Material 2

## Acknowledgements

None.

## Author contributions

Jie Gao carried out the Conception and design of the research, Xiuling Zhang participated in the acquisition of data. Anqi Xu and Wei Li carried out the Analysis and interpretation of data. HaiYan Gao participated in the design of the study and performed the statistical analysis. Jie Gao drafted the manuscript and revision of manuscript for important intellectual content. All authors read and approved the final manuscript.

## Funding

None.

## Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Declarations

### Ethical approval and informed consent statement

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

Received: 4 July 2024 / Accepted: 5 September 2024

Published online: 16 September 2024

## References

- Sung H, Ferlay J, Siegel RLL, Soerjomataram M, Jemal I, Bray A. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and Mortality Worldwide for 36 cancers in 185 countries. *Cancer J Clin.* 2021;71(3):209–49.
- Pimple SA, Mishra GA. Global strategies for cervical cancer prevention and screening. *Minerva Ginecol.* 2019;71(4):313–20.
- Shami S, Coombs J. Cervical cancer screening guidelines: an update. *JAAPA: Official J Am Acad Physician Assistants.* 2021;34(9):21–4.
- Hill EK. Updates in Cervical Cancer Treatment. *Clin Obstet Gynecol.* 2020;63(1):3–11.
- Ward ZJ, Scott AM, Hricak H, Atun R. Global costs, health benefits, and economic benefits of scaling up treatment and imaging modalities for survival of 11 cancers: a simulation-based analysis. *Lancet Oncol.* 2021;22(3):341–50.
- Lin W, Zhou Q, Wang CQ, Zhu L, Bi C, Zhang S, Wang X, Jin H. lncRNAs regulate metabolism in cancer. *Int J Biol Sci.* 2020;16(7):1194–206.
- Taniue K, Akimitsu N. The functions and unique features of lncRNAs in Cancer Development and Tumorigenesis. *Int J Mol Sci.* 2021;22(2):1–20.
- Yang Q, Al-Hendy A. The Regulatory functions and the mechanisms of Long non-coding RNAs in Cervical Cancer. *Cells* 2022, 11(7).
- He J, Huang B, Zhang K, Liu M, Xu T. Long non-coding RNA in cervical cancer: from biology to therapeutic opportunity. *Biomed Pharmacotherapy = Biomedecine Pharmacotherapie.* 2020;127:1–10.
- Wang T, Kong S, Tao M, Ju S. The potential role of RNA N<sup>6</sup>-methyladenosine in Cancer progression. *Mol Cancer.* 2020;19(1):1–18.
- Wang Q, Guo X, Li L, Gao Z, Su X, Ji M, Liu J. N(6)-methyladenosine METTL3 promotes cervical cancer tumorigenesis and Warburg effect through YTHDF1/HK2 modification. *Cell Death Dis.* 2020;11(10):1–10.
- Wang H, Luo Q, Kang J, Wei Q, Yang Y, Yang D, Liu X, Liu T, Yi P. YTHDF1 aggravates the progression of Cervical Cancer through m(6)A-Mediated Up-Regulation of RANBP2. *Front Oncol.* 2021;11:1–13.
- Condic M, Ralser DJ, Klümper N, Ellinger J, Qureschi M, Egger EK, Kristiansen G, Mustea A, Thiesler T. Comprehensive analysis of N<sup>6</sup>-Methyladenosine (m<sup>6</sup>A) writers, erasers, and readers in Cervical Cancer. *Int J Mol Sci.* 2022;23(13):1–11.
- Zhang H, Kong W, Zhao X, Han C, Liu T, Li J, Song D. N<sup>6</sup>-Methyladenosine-related lncRNAs as potential biomarkers for predicting prognoses and immune responses in patients with cervical cancer. *BMC Genomic data.* 2022;23(1):1–13.
- Nombela P, Miguel-López B, Blanco S. The role of m(6)A, m(5)C and Ψ RNA modifications in cancer: novel therapeutic opportunities. *Mol Cancer.* 2021;20(1):1–30.
- Guo G, Pan K, Fang S, Ye L, Tong X, Wang Z, Xue X, Zhang H. Advances in mRNA 5-methylcytosine modifications: detection, effectors, biological functions, and clinical relevance. *Mol Therapy Nucleic Acids.* 2021;26:575–93.
- Wang L, Zhang J, Su Y, Maimaitiyiming Y, Yang S, Shen Z, Lin S, Shen S, Zhan G, Wang F, et al. Distinct roles of m(5)C RNA methyltransferase NSUN2 in major gynecologic cancers. *Front Oncol.* 2022;12:1–15.
- Zuo Z, Xiong J, Zeng C, Jiang Y, Xiong K, Tao H, Guo Y. Exploration of a robust and prognostic Immune Related Gene signature for cervical squamous cell carcinoma. *Front Mol Biosci.* 2021;8:1–15.
- Song W, Ren J, Xiang R, Yuan W, Fu T. Cross-talk between m(6)A- and m(5)C-Related lncRNAs to Construct a Novel signature and predict the Immune Landscape of Colorectal Cancer patients. *Front Immunol.* 2022;13:1–16.
- Doncheva NT, Morris JHG, Jensen J. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res.* 2019;18(2):623–32.

21. Zhang X, Ren L, Yan X, Shan Y, Liu L, Zhou J, Kuang Q, Li M, Long H, Lai W. Identification of immune-related lncRNAs in periodontitis reveals regulation network of gene-lncRNA-pathway-immunocyte. *Int Immunopharmacol*. 2020;84:1–12.
22. Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. *Biometrical J Biometrische Z*. 2010;52(1):70–84.
23. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol*. 2015;16(4):e173–180.
24. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res*. 2013;41(Database issue):D955–961.
25. Zhu Y, Meng X, Ruan X, Lu X, Yan F, Wang F. Characterization of Neoantigen Load Subgroups in gynecologic and breast cancers. *Front Bioeng Biotechnol*. 2020;8:1–11.
26. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):1–13.
27. Aalijahan H, Ghorbian S. Long non-coding RNAs and cervical cancer. *Exp Mol Pathol*. 2019;106:7–16.
28. Zhu H, Zhu H, Tian M, Wang D, He J, Xu T. DNA methylation and hydroxy-methylation in Cervical Cancer: diagnosis, prognosis and treatment. *Front Genet*. 2020;11:1–12.
29. Pan J, Xu L, Pan H. Development and validation of an m6A RNA methylation Regulator-based signature for prognostic prediction in cervical squamous cell carcinoma. *Front Oncol*. 2020;10:1–8.
30. Ma X, Li Y, Wen J, Zhao Y. m6A RNA methylation regulators contribute to malignant development and have a clinical prognostic effect on cervical cancer. *Am J Translational Res*. 2020;12(12):8137–46.
31. Wu L, Liao W, Wang X, Zhao Y, Pang J, Chen Y, Yang H, He Y. Expression, prognosis value, and immune infiltration of lncRNA ASB16-AS1 identified by pan-cancer analysis. *Bioengineered*. 2021;12(2):10302–18.
32. Liu W, Zhuang R, Feng S, Bai X, Jia Z, Kapora E, Tan W. Long non-coding RNA ASB16-AS1 enhances cell proliferation, migration and invasion via functioning as a ceRNA through miR-1305/Wnt/ $\beta$ -catenin axis in cervical cancer. *Biomed Pharmacotherapy = Biomedecine Pharmacotherapie*. 2020;125:1–10.
33. Liu W, Wang QP, Guo J. Prognostic significance of lncRNA DANCR expression in human cancers: a systematic review and meta-analysis. *Biosci Rep*. 2021;41(8):1–9.
34. Hu C, Han Y, Zhu G, Li G, Wu X. Krüppel-like factor 5-induced overexpression of long non-coding RNA DANCR promotes the progression of cervical cancer via repressing microRNA-145-3p to target ZEB1. *Cell Cycle (Georgetown Tex)*. 2021;20(14):1441–54.
35. Cao L, Jin H, Zheng Y, Mao Y, Fu Z, Li X, Dong L. DANCR-mediated microRNA-665 regulates proliferation and metastasis of cervical cancer through the ERK/SMAD pathway. *Cancer Sci*. 2019;110(3):913–25.
36. Huang X, Shi H, Shi X, Jiang X. LncRNA FBXL19-AS1 promotes proliferation and metastasis of cervical cancer through upregulating COL1A1 as a sponge of miR-193a-5p. *J Biol Res (Thessalonike Greece)*. 2021;28(1):1–14.
37. Wan S, Ni G, Ding J, Huang Y. Long noncoding RNA FBXL19-AS1 expedites cell growth, Migration and Invasion in Cervical Cancer by miR-193a-5p/PIN1 signaling. *Cancer Manage Res*. 2020;12:9741–52.
38. Dai S, Yao D. An immune-associated ten-long noncoding RNA signature for predicting overall survival in cervical cancer. *Translational cancer Res*. 2021;10(12):5295–306.
39. Li S, Han Y, Liang X, Zhao M. LINC01089 inhibits the progression of cervical cancer via inhibiting miR-27a-3p and increasing BTG2. *J Gene Med*. 2021;23(1):1–10.
40. Qi H, Lu L, Wang L. Long noncoding RNA ST7-AS1 Upregulates TRPM7 expression by sponging microRNA-543 to promote Cervical Cancer Progression. *Oncotargets Therapy*. 2020;13:7257–69.
41. Ramilowski JA, Yip CW, Agrawal S, Chang JC, Ciani Y, Kulakovskiy IV, Mendez M, Ooi JLC, Ouyang JF, Parkinson N, et al. Functional annotation of human long noncoding RNAs via molecular phenotyping. *Genome Res*. 2020;30(7):1060–72.
42. Xu F, Shen J, Xu S. Integrated Bioinformatical Analysis identifies GIMAP4 as an Immune-related prognostic Biomarker Associated with Remodeling in Cervical Cancer Tumor Microenvironment. *Front cell Dev Biology*. 2021;9:1–13.
43. Chen Q, Qiu B, Zeng X, Hu L, Huang D, Chen K, Qiu X. Identification of a tumor microenvironment-related gene signature to improve the prediction of cervical cancer prognosis. *Cancer Cell Int*. 2021;21(1):1–16.
44. Zhou B, Lawrence T, Liang Y. The role of Plasmacytoid dendritic cells in cancers. *Front Immunol*. 2021;12:1–10.
45. Liu Y, Li L, Li Y, Zhao X. Research Progress on Tumor-Associated macrophages and inflammation in Cervical Cancer. *Biomed Res Int*. 2020;2020:1–6.
46. Zhou B, Li T, Xie R, Zhou J, Liu J, Luo Y, Zhang X. CircFAT1 facilitates cervical cancer malignant progression by regulating ERK1/2 and p38 MAPK pathway through miR-409-3p/CDK8 axis. *Drug Dev Res*. 2021;82(8):1131–43.
47. Xu J, Pan Q, Ju W. Ras inhibition by zoledronic acid effectively sensitizes cervical cancer to chemotherapy. *Anticancer Drugs*. 2019;30(8):821–7.

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.