Prognostic lncRNAs, miRNAs, and mRNAs Form a Competing Endogenous RNA Network in CESC

Lang Li

Zhejiang Chinese Medical University

Qiusheng Guo

Zhejiang Chinese Medical University

Gaochen Lan

the First Affiliated Hospital of Zhejiang Chinese Medical University

Fei Liu

Jinhua People's Hospital

Wenwu Wang

the Third Affiliated People's Hospital of Fujian University of Traditional Chinese Medicine

Xianmei Lv (✉ 201911011511059@zcmu.edu.cn)

Quzhou Kecheng Hospital

Research Article

Keywords: mRNA, lncRNA, miRNA, CESC, ceRNA, Prognostic signature model

Posted Date: March 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-352332/v1>

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

[Read Full License](#)

Abstract

Background: Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) tumorigenesis involves a combination of multiple genetic alteration processes. Constructing a survival-associated competing endogenous RNA (ceRNA) network and a multi-mRNA-based prognostic signature model can help us better understand the complexity and genetic characteristics of CESC.

Methods: The RNA-seq data and clinical information of CESC patients were downloaded from The Cancer Genome Atlas (TCGA) database. Differentially expressed mRNAs, lncRNAs and miRNAs were identified by edgeR package. Constructing prognostic model used the differentially expressed RNAs. The Kaplan-Meier method and log-rank test were performed to assess survival rates. The relationships between overall survival (OS) and clinical parameters were evaluated by Cox regression analysis. A survival-associated ceRNA network was constructed by multiMiR package and miRcode database. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis and Gene Ontology (GO) were used to identify the functional role of the ceRNA network in the prognosis of CESC.

Results: Differentially expressed 298 mRNAs, 8 miRNAs, and 29 lncRNAs were significantly associated with the prognosis of CESC. The prognostic signature model based on 4 mRNAs (OPN3, DAAM2, HENMT1, and CAVIN3) was constructed. The prognostic ability was 0.726 for this model. Patients in the high-risk group were significantly associated with worse OS. The KEGG pathways were significantly enriched in the TGF- β and Cell adhesion molecules signaling pathways.

Conclusion: This study identified several potential prognostic biomarkers to construct a multi-mRNA-based prognostic model for CESC.

Background

Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) is one of the most common malignancy worldwide, with more than 570,000 new cases and 274,000 deaths per year [1, 2]. The long-term survival rates of patients with early-stage disease have been greatly improved in recent years. However, 5-year survival rate remains less than 16.8% in recurrent or metastatic patients. At present, the prediction of CESC prognosis mainly depends on the tumour-node-metastasis (TNM) stage. However, the TNM stage is based on anatomical information and does not reflect the biological heterogeneity of CESC. Hence, it is urgent to find novel biomarkers based on transcriptomics that can act as prognostic indicators to guide precise individualized treatment.

Recently, the competing endogenous RNA (ceRNA) hypothesis have provided novel insights into the cancer research field. The ceRNA hypothesis is proposed to be a special pathway to regulate the expression of RNAs [3]. In that hypothesis, miRNA acts as the hub gene that suppresses mRNA translation, while lncRNA competes to bind one or multiple sites of miRNA to suppress the function of miRNA and participates in post-transcription control [4]. Previous studies reported that the ceRNA networks might act as the biomarkers for prognosis in CESC. Song et al [5] constructed a CESC-

associated ceRNA network which consisted of 50 lncRNAs, 81 mRNAs and 18 miRNAs, and found that several RNAs were associated with the prognosis. However, that study did not construct a prognostic model based on the prognostic RNAs for CESC. CESC is a heterogeneous disease with multiple gene alterations and interactions. Hence, it is of great significance to construct a CESC-associated ceRNA network and develop a multi-gene prognosis model based on the ceRNA network.

In this study, we analyzed the relationships between differentially expressed RNAs and overall survival (OS), and constructed a survival-associated ceRNA network and a prognostic signature model based on multi-mRNA for CESC. In addition, we utilized the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis and Gene Ontology (GO) function annotation to further investigate the function of the survival-associated RNAs. This study aimed to provide a novel biomarker to guide personalized medicine and to facilitate understanding of the molecular mechanisms for CESC.

Methods

Data processing

The clinical information and RNA-seq data of CESC patients were downloaded from The Cancer Genome Atlas (<https://portal.gdc.cancer.gov>) (TCGA) database. The inclusion criteria were as follows: (1) Patients with complete clinical information, including T stage, M stage, N stage, pathological stage, survival status, age, and histological type. (2) Patients with complete lncRNA-seq, miRNA-seq, and mRNA-seq data. Finally, 240 CESC samples and 3 adjacent non-tumor samples were examined.

Differentially expressed RNAs analysis

Using the edgeR package of R software to identify differentially expressed mRNAs, lncRNAs, and miRNAs, respectively, between CESC and adjacent non-tumor samples [6, 7]. $P < 0.05$ and the absolute value of fold change $|FC| > 4$ were considered to indicate statistical significance. The expression profiles of lncRNAs, miRNAs, and mRNAs were converted to $\log_2(\text{count} + 1)$ values after normalization by edgeR.

Survival analysis and prognosis-associated signature model construction

Survival R package was used for univariate Cox regression to assess the relationships between differentially expressed RNAs and OS [8]. mRNAs with a $P < 0.001$ in the univariate Cox regression were included into the multivariate Cox regression. $P < 0.05$ was considered to indicate statistical significance.

A mRNA based prognosis signature risk score model was constructed on the basis of a linear combination of the expression level (Exp) multiplied regression coefficient (β) derived from the multivariate cox regression model with the following formula: Risk Score (RS) = $(\text{Exp}_{\text{OPN3}} \times \beta_{\text{OPN3}}) + (\text{Exp}_{\text{DAAM2}} \times \beta_{\text{DAAM2}}) + (\text{Exp}_{\text{HENMT1}} \times \beta_{\text{HENMT1}}) + (\text{Exp}_{\text{CAVIN3}} \times \beta_{\text{CAVIN3}})$.

According to the cut-off of the median RS, CESC patients were divided into low-risk group and high-risk group. The time-dependent receiver operating characteristic (ROC) curve was applied to assess the prognostic accuracy of the model, and the area under curve (AUC) was used to assess the specificity and sensitivity of the model. Using R software (version 4.0.2) and the “survivalROC” package to perform the time-dependent ROC curve analysis.

Analysis of the protein expression levels of the 4 key mRNAs

The Human Protein Atlas (<https://www.proteinatlas.org>) provides tissue and cellular distribution information for approximately 26000 human proteins. The information was obtained via immunoassay techniques (immunohistochemistry, immunofluorescence, and western blotting) to detect proteins expression in 64 cell lines, 48 normal human tissues and 20 tumour tissues [9]. In this study, we compared the expression levels of the 4 key genes (OPN3, DAAM2, HENTM1, and CAVIN3) between normal and CESC tissues using immunohistochemistry images.

Functional annotation and pathway enrichment analysis

Using the ClusterProfiler R package [10] to process KEGG pathway enrichment analysis and GO function annotation. Using the ggplot2 R package [11] to visualize the results of the KEGG and GO analyses. $P < 0.05$ was considered to indicate statistical significance.

ceRNA network construction

We constructed a ceRNA network according to the identified prognostic lncRNAs, miRNAs, and mRNAs. The multiMiR R package [12] and miRcode (<http://www.mircode.org/>) database were applied to predict miRNA-mRNA interactions and lncRNA-miRNA interactions. Finally, Cytoscape (version 3.7.0) software was utilized to construct a ceRNA network based on the lncRNA-miRNA-mRNA axis by combining lncRNA-miRNA interactions with miRNA-target gene interactions. In the ceRNA network, lncRNA and mRNA act as natural miRNA sponges to suppress miRNA function by binding one or multiple sites of miRNA.

Statistical analysis

Using the t-test to access the relationships between different groups and the gene expression profiles. The Kaplan-Meier method and log-rank test were performed to assess survival rates. Univariate Cox regression and multivariate Cox regression were used to analyze the relationships between the Clinicopathological parameters and OS. R software (version 4.0.2) was used to plot the figures and statistical analyses. $P < 0.05$ was considered to indicate statistical significance.

Results

Differentially expressed RNAs analysis

19545 mRNAs, 2713 miRNAs, and 13977 lncRNAs were extracted from transcriptome data. The whole expression landscape of mRNAs, miRNAs, and lncRNAs between 3 adjacent normal samples and 240 tumor samples are presented in Fig. 1a, b, c, respectively. In total, we identified 1398 downregulated and 1495 upregulated mRNAs (Fig. 1d), 5 downregulated and 15 upregulated miRNAs (Fig. 1e), 493 downregulated and 658 upregulated lncRNAs (Fig. 1f), and an additional file show this in more detail. [see additional file 1]

Survival-associated RNAs

The relationships between the differentially expressed lncRNAs, miRNAs, and mRNAs and OS were evaluated in 240 CESC patients. The univariate Cox regression revealed that 298 mRNAs, 8 miRNAs, and 129 lncRNAs were significantly associated with OS, respectively. The top 15 mRNAs, miRNAs, and lncRNAs ranked by p-value are presented in Fig. 2a, b, c, respectively, and an additional file show this in more detail. [see additional file 2]

Functional annotation and pathway enrichment analysis

GO functional annotation and KEGG pathway enrichment were utilized to investigate the biological functions of survival-associated RNAs. GO functional annotation included biological process (BP), cellular component (CC), and molecular function (MF). The top 10 GO terms and KEGG pathways are listed in Fig. 3. GO BP analysis showed that the targeted RNAs were significantly enriched in the epithelial cell proliferation and endothelium development (Fig. 3a). For MF analysis, the targeted RNAs were significantly enriched in the BMP binding receptor and fattening acid synthase activity (Fig. 3b). CC analysis revealed enrichment of ruffle and catenin complex (Fig. 3c). KEGG Pathway enrichment analysis revealed that the targeted genes were significantly enriched in the TGF- β and Cell adhesion molecules signaling pathways (Fig. 3d). An additional file show this in more detail. [see additional file 3]

Survival-associated ceRNA network

A comprehensive survival-associated lncRNA-miRNA-mRNA ceRNA network was constructed by combining the lncRNA-miRNA interactions with the miRNA-mRNA interactions. The ceRNA network includes 24 lncRNAs (COL18A1-AS1, SNHG14, EMX2OS, LINC00479, LINC00299, WDFY3-AS2, WT1-AS, MEG3, EPB41L4A-AS1, CYP1B1-AS1, CLRN1-AS1, GK-AS1, LINC00466, FAM66C, LINC00319, SACS-AS1, DLEU2, JAZF1-AS1, PVT1, DIO3OS, FRMD6-AS2, MIAT, LINC00484, and CRNDE), 6 miRNAs (miR-4664, miR-143, miR200a, miR-210, miR-126, and miR-200b), and 34 mRNAs (ZBTB47, FASN, OPN3, LIPG, IL1A, MPDZ, FOXP3, BMP6, GRB7, PHLDA2, MAP3K20, CIDEA, MMP1, SAMD5, ELOVL3, HENMT1, PTPN22, CAVIN3, DAAM2, HOXA1, KRT222, CHAF1B, PPP1R14A, SQLE, CHAF1A, RMI2, MAP7, POU4F1, MCAM, DLL4, ASF1B, PCDH10, PCDH18, and FBN1) (Fig. 4).

Predictive model for overall survival

To construct a mRNA based prognosis signature model for CESC, 8 mRNAs of $P < 0.0001$ in the univariate Cox regression were included into the multivariate Cox regression analysis. Finally, 4 mRNAs (OPN3, DAAM2, HENMT1, and CAVIN3) were identified. The information of those mRNAs are presented in Table 1. In addition, we evaluated the transcript and protein levels of above 4 genes between adjacent normal and CESC tissues using t-test and immunohistochemistry (IHC). The transcript levels of OPN3 ($P = 0.013$) and HENMT1 ($P = 0.0095$) were significantly upregulated, while the transcript levels of DAAM2 ($P < 0.0001$) and CAVIN3 ($P < 0.0001$) were significantly downregulated in CESC tissues compared with adjacent normal tissues (Fig. 5). On the other hand, the Human Protein Atlas database was used to evaluate the protein expression levels of above 4 genes in adjacent normal and CESC tissues by IHC. Based on the immunohistochemical staining images, the protein expression levels of OPN3 and HENMT1 were higher, while the protein expression level of CAVIN3 was lower in CESC samples than those in normal samples, which was consistent with the transcriptomics results (Fig. 6).

Table 1
The information of 4 mRNAs significantly associated with OS in CESC

Gene	Ensemble ID	Location	β	HR	P
OPN3	ENSG00000054277	Chr1(241,593,124-	0.2625	1.3002	0.0006
DAAM2	ENSG00000146122	241,640,369)	0.2724	1.3131	0.014
HENM1	ENSG00000162639	Chr6 (397,923 ,66-	-0.0794	0.9236	0.0046
CAVIN3	ENSG00000170955	399,048,70)	0.0202	1.0204	0.0122
		Chr1(108,648,295-			
		108661474)			
		Chr11(631,894,6-			
		632,050,1)			
β , Regression coefficient; HR, Hazard ratio.					

A mRNA based prognosis signature RS model was constructed based on above 4 mRNAs. The RS was calculated with the following formula: $RS = (Exp_{OPN3} \times \beta_{OPN3}) + (Exp_{DAAM2} \times \beta_{DAAM2}) + (Exp_{HENMT1} \times \beta_{HENMT1}) + (Exp_{CAVIN3} \times \beta_{CAVIN3})$. The “Exp” value represents the expression level and the “ β ” value represents regression coefficient derived from the multivariate cox regression model.

The RS of each patient were calculated according to the above formula. Then, CESC patients were divided into low-risk group and high-risk group according to the cut-off of the median RS. Using t-test to compare the expression levels of the 4 mRNAs between low-risk group and high-risk group. The expression levels of OPN3 ($P < 0.0001$), DAAM2 ($P < 0.0001$), and CAVIN3 ($P < 0.0001$) were higher, while the expression levels of HENMT1 ($P < 0.0001$) was lower in the high-risk group than those in the low-risk group (Fig. 7). Figure 8 shows the performance of the mRNA-based model. Figure 8a shows the rank of

patients according to the RS. The scatter plot presents the OS of CESC patients decrease along with the increase of RS (Fig. 8b). The heat map shows the expression level of HENMT1 decrease along with the increase of RS, while the expression levels of OPN3, DAAM2 and CAVIN3 increase (Fig. 8c).

The Kaplan-Meier method and log-rank test were performed to identify the relationship between different RS groups and OS, and the ROC curve was used to exam the sensitivity and specificity of the model. CESC patients in the high-risk group were significantly associated with shorter OS ($P < 0.0001$) (Fig. 9a). The AUC of the RS revealed that the model showed prognostic assessment ability was 0.726 (Fig. 9b).

Several clinical parameters were found to have some prognosis value with univariate analysis, such as M stage ($P = 0.02$), N stage ($P = 0.01$), T stage ($P = 0.001$), pathological stage ($P = 0.001$) and the signature RS ($P = 1.5e-6$); however, only the signature RS remained statistically significant with the confirmation by multivariate analysis (HR = 6.35, $P = 0.01$, Table 2).

Table 2
Univariate and multivariate Cox regression analysis of characteristics and Signature RS in CESC

Clinical Parameter	N	Univariate Analysis	Multivariate Analysis
		HR (95% CI) P	HR (95% CI) P
Age	192	1.79(0.84–3.8) 0.13	--
≤ 65	48	3.55(1.2-10.55) 0.02	0 (0-Inf) 1
> 65	107	2.6(1.32–5.11) 0.01	1.51(0.56–4.06) 0.42
M stage	10	3.58(1.89–6.77) 0.001	3.58(1.89–6.77) 0.16
M0	125	5.77(3.01–11.05) 0.001	1.22(0.17–8.67) 0.84
M1	53	1.13(0.55–2.31) 0.74	--
N stage	200	0.22(0.12–0.41) 1.5e-6	6.35(1.66–24.3) 0.01
N0	26		
N1+	125		
T stage	53		
T1-T2	47		
T3-T4	193		
Pathological Stage	120		
I-II	120		
III-IV			
Histological Types			
Adenosqua-mous			
Squamous			
Signature RS			
High			
Low			
N, Number of patients; HR, Hazard ratio; CI, Confidence interval.			

Discussion

In this study, distinct mRNAs, lncRNAs, and miRNAs were identified to further understand the molecular events related to CESC prognosis. In addition, a prognostic ceRNA network provides new insights for the prognosis of CESC.

CESE tumorigenesis involves a combination of multiple genetic alteration processes rather than simply due to a single gene alteration. However, almost all studies focused on a single or a single cluster 'driver' gene of CESC so far. Daniel et al [13] reported that Keratin-17 is a prognostic biomarker of CESC. Li et al [14] reported that FAM83A is a potential biomarker regulated by miR-20, which promotes the development of CESC through the PI3K/AKT/mTOR signaling pathway. At present, no single pivotal driver gene or gene cluster was reported to be superior to evaluate the prognosis of CESC. Moreover, TNM stage, as a major prognostic indicator, is based on anatomical information and does not reflect the biological heterogeneity of CESC. Hence, it is of great significance to identify the effective biomarkers and construct a multi-mRNA-based model based on survival-associated biomarkers to predict the prognosis of CESE.

Other studies constructed a ceRNA network of CESC, but there are some limitations. Song et al [5] constructed a CESC-associated ceRNA network which consisted of 50 lncRNAs, 81 mRNAs and 18 miRNAs, and found that several RNAs were associated with the prognosis. However, that study did not construct a prognostic model based on the prognostic RNAs for CESC. Another study conducted by Chen et al [15] also constructed a CESC-associated ceRNA network, however, the study only assessed the relationship between OS and a single gene. Although some studies have constructed a CESC-associated ceRNA and simultaneously developed a prognostic model, they did not use ROC curve to evaluate the predictive ability of the model. Prognostic models based on multiple mRNAs could provide more accurate predictions than single gene. Our study constructed a ceRNA network and identified several novel potential prognostic RNAs for CESC. Moreover, a prognostic model was constructed based on the prognostic mRNAs including OPN3, DAAM2, HENMT1, and CAVIN3. The evaluation ability was 0.726 of this model with ROC.

Previous studies developed several prognostic models based on multiple genes for CESC. Meng et al [16] developed a CESC prognosis model based on DSG2, ITM2A, CENPM, RIBC2, and MEIS2. Liu et al [17] constructed a multi-mRNA prognosis model which were comprised of ITGA5, HHEX, and S1PR4. Similarly, our study also constructed a CESC-associated prognosis model based on OPN3, DAAM2, HENMT1, and CAVIN3, and the evaluation ability was 0.726 of this model with ROC. Moreover, in our study, the Cox proportional hazard model was used, and we corrected for the following confounding factors that can greatly affect prognosis: T stage, N stage, M stage, and pathological stage. The RS signature of this model was significantly associated with the clinical outcome of CESC patients. These results indicate that the model we constructed is an independent prognostic factor.

The KEGG pathways analysis in the ceRNA network indicated that the targeted RNAs are significantly enriched in the TGF-beta signaling pathway and Cell adhesion molecules pathway. TGF-beta is a critical cancer-associated signaling pathway, which is involved in proliferation, apoptosis, differentiation, migration, and epithelia-mesenchymal transition (EMT) of cancer [18, 19]. TGF-beta signaling pathway plays vital role in CESC as well. Deng et al [20] reported that CD36 and TGF- β interact to promote the EMT of CESC. Yang et al [21] found that downregulation of SEMA4C could inhibit EMT, invasion, and metastasis of CESC via inhibiting TGF-beta. Cell adhesion molecules pathway plays a critical role in the development of CESC. Carvalho et al [22] reported that L1 cell adhesion molecule expression is

associated with a poor prognosis. The biological process analysis revealed that the main disturbed biological process by those survival-associated RNAs in the ceRNA network is the epithelial cell proliferation. The cytokinetic homeostasis is controlled by the balance of cell proliferation and apoptosis. Previous studies have reported that excessive cell proliferative activity is a precancerous lesion in some types of tumors. T Obara indicated that the epithelial cell proliferation activity is significantly stepwise increased from normal gallbladder mucosa to cancer [23]. However, different pathways are usually perturbed by different molecules and thus need to be further investigated in the laboratory.

Our study still has some limitations. Firstly, although the effects of OPN3, DAAM2, HENTM1, and CAVIN3 in our prognosis model on tumorigenesis have been reported in other types of cancer [24–27], the exact effects on CESC have yet to be fully elucidated and need to be verified by experiments. Secondly, the research data came from a single online database, and another independent cohort is needed to verify above results in the future. Thirdly, the information of cervical cancer patients from the TCGA should be assessed with another experimental method.

Conclusion

In conclusion, this study constructed a ceRNA network that provides novel insights for conducting the study of gene complexity in CESC. This study may provide novel biomarkers and assist in the design of molecular targeted therapy for CESC.

Abbreviations

CESC: Cervical squamous cell carcinoma and endocervical adenocarcinoma; TCGA: The Cancer Genome Atlas; ceRNA: competitive endogenous RNA; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology; ROC: Receiver operating characteristic; AUC: Area under curve; |FC|: Absolute value of fold change; RNA-seq: RNA sequence; LncRNA: Long non-coding RNA; miRNA: MicroRNA; mRNA: Message RNA; EMT: Epithelia-mesenchymal transition; OS: Overall survival

Declarations

Ethics approval consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets used during the current study are available from TCGA (<https://portal.gdc.cancer.gov>) and The Human Protein Atlas (<https://www.proteinatlas.org>). Analyzed data is available within

supplementary information or from the authors upon reasonable request.

Competing interests

No conflict of interest exists in the manuscript.

Funding

This study was sponsored by "the Postgraduate Scientific Research Fund" of Zhejiang Chinese Medical University.

Authors' contributions

LL performed the data analysis. LL, GCL, and QSG prepared the manuscript. GCL, WWW, and FL revised the manuscript. GCL acquired the funding. XML designed the study.

Acknowledgements

We thank those researchers who shared the data and R packages for this study.

References

1. Paavonen J: Human papillomavirus infection and the development of cervical cancer and related genital neoplasias. *Int J Infect Dis* 2007, 11 Suppl 2:S3-9.
2. 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 Cancer J Clin* 2018, 68(6):394–424.
3. Thomson DW, Dinger ME: Endogenous microRNA sponges: evidence and controversy. *Nat Rev Genet* 2016, 17(5):272–283.
4. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP: A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 2011, 146(3):353–358.
5. Song J, Ye A, Jiang E, Yin X, Chen Z, Bai G, Zhou Y, Liu J: Reconstruction and analysis of the aberrant lncRNA-miRNA-mRNA network based on competitive endogenous RNA in CESC. *J Cell Biochem* 2018, 119(8):6665–6673.
6. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, 26(1):139–140.
7. McCarthy DJ, Chen Y, Smyth GK: Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 2012, 40(10):4288–4297.
8. Terry M. Therneau PMG: Modeling Survival Data: Extending the Cox Model. Springer, New York; 2000.

9. 9. Wu ZH, Cai F, Zhong Y: Comprehensive Analysis of the Expression and Prognosis for GBPs in Head and neck squamous cell carcinoma. *Sci Rep* 2020, 10(1):6085.
10. 10. Yu G, Wang LG, Han Y, He QY: clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 2012, 16(5):284–287.
11. 11. Wickham H: ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York; 2016.
12. 12. Ru Y, Kechris KJ, Tabakoff B, Hoffman P, Radcliffe RA, Bowler R, Mahaffey S, Rossi S, Calin GA, Bemis L *et al*: The multiMiR R package and database: integration of microRNA-target interactions along with their disease and drug associations. *Nucleic Acids Res* 2014, 42(17):e133.
13. 13. Mockler D, Escobar-Hoyos LF, Akalin A, Romeiser J, Shroyer AL, Shroyer KR: Keratin 17 Is a Prognostic Biomarker in Endocervical Glandular Neoplasia. *Am J Clin Pathol* 2017, 148(3):264–273.
14. 14. Rong L, Li H, Li Z, Ouyang J, Ma Y, Song F, Chen Y: FAM83A as a Potential Biological Marker Is Regulated by miR-206 to Promote Cervical Cancer Progression Through PI3K/AKT/mTOR Pathway. *Front Med (Lausanne)* 2020, 7:608441.
15. 15. Chen W, Chen X, Wang Y, Liu T, Liang Y, Xiao Y, Chen L: Construction and Analysis of lncRNA-Mediated ceRNA Network in Cervical Squamous Cell Carcinoma by Weighted Gene Co-Expression Network Analysis. *Med Sci Monit* 2019, 25:2609–2622.
16. 16. Meng H, Liu J, Qiu J, Nie S, Jiang Y, Wan Y, Cheng W: Identification of Key Genes in Association with Progression and Prognosis in Cervical Squamous Cell Carcinoma. *DNA Cell Biol* 2020, 39(5):848–863.
17. 17. Liu J, Nie S, Li S, Meng H, Sun R, Yang J, Cheng W: Methylation-driven genes and their prognostic value in cervical squamous cell carcinoma. *Ann Transl Med* 2020, 8(14):868.
18. 18. Syed V: TGF- β Signaling in Cancer. *J Cell Biochem* 2016, 117(6):1279–1287.
19. 19. Xie F, Ling L, van Dam H, Zhou F, Zhang L: TGF- β signaling in cancer metastasis. *Acta Biochim Biophys Sin (Shanghai)* 2018, 50(1):121–132.
20. 20. Deng M, Cai X, Long L, Xie L, Ma H, Zhou Y, Liu S, Zeng C: CD36 promotes the epithelial-mesenchymal transition and metastasis in cervical cancer by interacting with TGF- β . *J Transl Med* 2019, 17(1):352.
21. 21. Yang L, Yu Y, Xiong Z, Chen H, Tan B, Hu H: Downregulation of SEMA4C Inhibit Epithelial-Mesenchymal Transition (EMT) and the Invasion and Metastasis of Cervical Cancer Cells via Inhibiting Transforming Growth Factor-beta 1 (TGF- β 1)-Induced Hela cells p38 Mitogen-Activated Protein Kinase (MAPK) Activation. *Med Sci Monit* 2020, 26:e918123.
22. 22. Mancusi de Carvalho JP, Salim RC, Carvalho FM, Nogueira Dias Genta ML, Baracat EC, Carvalho JP: L1 cell adhesion molecule (L1CAM) in stage IB cervical cancer: distinct expression in squamous cell carcinomas and adenocarcinomas. *J Clin Pathol* 2020, 73(11):748–753.
23. 23. Obara T, Tanno S, Fujii T, Izawa T, Mizukami Y, Yanagawa N, Ura H, Kohgo Y: Epithelial cell proliferation and gene mutation in the mucosa of gallbladder with pancreaticobiliary malunion and cancer. *J Hepatobiliary Pancreat Surg* 1999, 6(3):229–236.

24. Xu C, Wang R, Yang Y, Xu T, Li Y, Xu J, Jiang Z: Expression of OPN3 in lung adenocarcinoma promotes epithelial-mesenchymal transition and tumor metastasis. *Thorac Cancer* 2020, 11(2):286–294.
25. Zhu W, Krishna S, Garcia C, Lin CJ, Mitchell BD, Scott KL, Mohila CA, Creighton CJ, Yoo SH, Lee HK *et al*: Daam2 driven degradation of VHL promotes gliomagenesis. *Elife* 2017, 6.
26. Sun G, Ni K: The Role of Cavin3 in the Progression of Lung Cancer and Its Mechanism. *Biomed Res Int* 2020, 2020:6364801.
27. Hoff AM, Alagaratnam S, Zhao S, Bruun J, Andrews PW, Lothe RA, Skotheim RI: Identification of Novel Fusion Genes in Testicular Germ Cell Tumors. *Cancer Res* 2016, 76(1):108–116.

Figures

Figure 1

The differentially expressed RNAs between CESC and adjacent normal tissues. Heat map of the whole expression landscape of (a) mRNAs, (b) miRNAs, and (c) lncRNAs. Volcano plot of the differentially expressed (d) mRNAs, (e) miRNAs and (f) lncRNAs. Red represents high expression and green indicates low expression. Black shows no significant difference in expression.

Figure 2

Forest plot of the hazard ratios of the top 15 survival-associated (a) mRNAs, (b) miRNAs, and (c) lncRNAs. A hazard ratio > 1 represents the high-risk RNAs and a hazard ratio < 1 represents the protective RNAs.

Figure 3

The enrichment analysis of the survival-associated RNAs. (a) Bubble plot of enriched GO BP terms. (b) Bubble plot of enriched GO MF terms. (c) Bubble plot of enriched GO CC terms. (d) Bubble plot of enriched KEGG pathways

Figure 4

The prognostic ceRNA network in CESC. The connection indicates the interactions among lncRNA, miRNA, or mRNA.

Figure 5

Comparison of the transcript levels of the 4 key mRNAs between CESC and normal tissues. The transcript levels of (a) OPN3 and (c) HENMT1 were significantly upregulated, while the transcript levels of (b)

DAAM2 and (d) CAVIN3 were significantly downregulated in CESC tissues compared with normal tissues.
 * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

Figure 6

Representative immunohistochemistry images of the 4 key mRNAs in CESC and adjacent normal tissues. The protein expression levels of OPN3 and HENMT1 were higher, while the protein expression level of CAVIN3 was lower in CESC samples than those in normal samples (Human Protein Atlas).

Figure 7

Comparison of the transcript levels of the 4 key mRNAs between low-risk group and high-risk group. The expression levels of (a) OPN3, (b) DAAM2, and (d) CAVIN3 were higher, while the expression level of (c) HENMT1 was lower in the high-risk group than those in the low-risk group. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

Figure 8

Performance of the prognostic model in classifying patients into low-risk and high-risk groups. (a) The risk score distribution in the four-mRNA signature model. (b) The risk score distribution of survival status in the four-mRNA signature model. (c) Heat map of the 4 key mRNAs expression in the low-risk and high-risk groups. The color from blue to red shows a trend from low expression to high expression.

Figure 9

Kaplan-Meier and ROC curves for the four-mRNA signature in CESC. (a) Kaplan-Meier survival curves showing OS outcomes according to relative high-risk and low-risk patients. (b) Time-dependent ROC curves analysis for survival prediction by the four-mRNA signature.

Supplementary Files

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

- [Additionnalfile1.xlsx](#)
- [Additionalfile2.xlsx](#)

- [AdditionalFile3.xlsx](#)