ADH7, miR-3065 and LINC01133 are associated with cervical cancer progression in different age groups

SHENGDI DING[1], XIAOHONG HUANG[1], JUNMEI ZHU[1], BING XU[1],
LIMIN XU[2], DONGHUA GU[3] and WENYUAN ZHANG[1]

[1]Department of Obstetrics and Gynecology; [2]Central Laboratory; [3]Department of Pathology, Huzhou Central Hospital,
Huzhou, Zhejiang 313000, P.R. China

Abstract. The aim of the present study was to identify potential therapeutic targets that serve crucial roles in the progression of cervical cancer. Clinical data, RNA sequencing (RNAseq)-counts and micro (mi)RNA data regarding cervical squamous cell carcinoma were retrieved from The Cancer Genome Atlas, and analyses were performed using the University of California Santa Cruz database. RNAseq and miRNA data were stratified into 3 groups (according to the patients' age), and genes were re-annotated and preprocessed prior to Mfuzz time clustering analysis. Subsequently, enrichment analyses were performed in order to identify differentially expressed mRNAs (DEmRNAs) and a protein-protein interaction analysis network was constructed. miRNA-gene, miRNA-lncRNA, and long non-coding (lnc)RNA-mRNA pairs were collected and the lncRNA-miRNA-mRNA competing endogenous (ce)RNA network was established. Further enrichment analyses were performed in order to identify crucial mRNAs in the ceRNA network. Finally, survival and drug association analyses were implemented. A total of 269 DEmRNAs [including alcohol dehydrogenase 7 (*ADH7*), vestigial-like family member 3 (*VGLL3*) and cytochrome P450, family 26, subfamily B, polypeptide 1 (*CYP26B1*)], 274 DElncRNAs (including LINC01133) and 16 DEmiRNAs (including miR-3065 and miR-330) were identified. There were 102 lncRNAs, 15 miRNAs, 15 mRNAs and 522 interaction pairs in the ceRNA network. In particular, *ADH7* was regulated by miR-3065, and miR-3065 interacted with LINC01133 in the ceRNA network. Furthermore, *ADH7* and *CYP26B1* were enriched in the retinoic acid metabolic process and the retinol metabolism pathway. *ADH7* and *VGLL3* were significantly associated with the cervical cancer survival rate. *ADH7*, *VGLL3*, *CYP26B1*, miR-3065, miR-330, miR-499a and LINC01133 play pivotal roles in the progression of cervical cancer in different age groups.

Introduction

Cervical cancer is the fourth most prevalent cancer in women worldwide, behind colon, breast and lung cancer, and represents a leading cause of tumor-associated mortality amongst women in developing countries (1). Notably, it is estimated that ~570,000 cervical cancer cases occurred worldwide in 2018 (2). Human papillomavirus (HPV) infection is a persistent infection that is associated with the tumorigenesis of cervical cancer; in particular, HPV type 16 (HPV16), which is present in >50% of patients (3). Previous studies have revealed that the distribution and prevalence of HPV genotypes differs depending on age cohort (4-6). Furthermore, HPV is a sexually transmitted infection and 75-80% of sexually active individuals are likely to be infected in their lifetime (7). Hence, the prevalence of both HPV and cervical cancer among women may be associated with age (8). For example, females aged 25-44 years have a demonstrated ability to clear or prevent HPV infection and, consequently, the infection rate of HPV in these women is lower than other age groups (9). However, over the past 10 years the prevalence of HPV has increased among gynecological outpatients aged between 20 and 24 years (8). This may be due to the long-term use of contraceptives, early coitus and an overall increase in stress levels (10,11). Therefore, research on differences in cervical cancer between various age groups may provide clinical guidance to help address the increasing prevalence of cervical cancer in young women.

There are currently several treatments for patients with cervical cancer, including radiotherapy, chemotherapy, interventional therapy and radical hysterectomy, and the overall survival rate of patients with cervical cancer has improved

Correspondence to: Dr Wenyuan Zhang, Department of Obstetrics and Gynecology, Huzhou Central Hospital, 198 Hongqi Road, Huzhou, Zhejiang 313000, P.R. China
E-mail: zhangwenyuan1976@163.com

Abbreviations: DEmRNAs, differentially expressed mRNAs; ADH7, alcohol dehydrogenase 7; VGLL3, vestigial-like family member 3; CYP26B1, cytochrome P450, family 26, subfamily B, polypeptide 1; HPV, human papillomavirus; miRNAs, microRNAs; lncRNAs, long noncoding RNAs; ceRNA, competing endogenous RNA; TCGA, The Cancer Genome Atlas; KEGG, Kyoto Encyclopedia of Genes and Genomes; HCC, hepatocellular carcinoma

Key words: cervical cancer, retinoic acid, ceRNA network, survival analysis

to a certain extent (12,13). However, lymphatic metastasis, postoperative recurrence and other toxic side effects, including secondary infection, radiation damage and malnutrition all have an adverse effect on therapeutic outcomes (14). Therefore, it is essential to investigate novel and effective prognostic indicators and therapeutic targets to improve cervical cancer therapy (15).

MicroRNAs (miRNAs) are a group of endogenous short noncoding RNAs consisting of ~22 nucleotides that cause the degradation of mRNA by targeting the 3' untranslated region (16,17). At present, several miRNAs have been demonstrated to influence the progression of cervical cancer. For instance, miR-152 expression is upregulated in the peripheral blood of patients with cervical intraepithelial neoplasia, serving as an early diagnostic and prognostic biomarker (18). Furthermore, miR-4429 sensitizes cervical cancer cells to irradiation via the targeting of RAD51 recombinase (19). Additionally, increasing evidence has demonstrated that the long non-coding (lnc)RNAs serve crucial roles in the migration, differentiation, invasion and proliferation of cells (20,21). Notably, several lncRNAs that regulate various cellular biological processes and influence the oncogenesis of cervical cancer have been identified (22). For example, the lncRNA CRNDE mediates the tumorigenesis of cervical cancer via inhibition of p53 upregulated modulator of apoptosis expression (PUMA) (23).

Recently, it has been demonstrated that certain lncRNAs and mRNA can co-operate with miRNAs, downregulating the expression of miRNAs by forming a competing endogenous (ce)RNA network (24). Hence, the present study hypothesized that lncRNAs may regulate mRNA expression via co-operation with miRNAs in cervical cancer. In the present study, the clinical data, RNA sequencing (RNA-seq)-counts and miRNA data in different age groups were integrated to further investigate the molecular mechanisms underlying cervical cancer development, resulting in the identification of several potential therapeutic targets of cervical cancer.

Materials and methods

Data source and preprocessing. The Cancer Genome Atlas (TCGA; https://cancergenome.nih.gov) is the largest database of cancer gene information, containing data regarding 29 common cancer types, 39 cancer subtypes and >14,500 cancer samples. The data describe various characteristics, including miRNA, RNA-seq, single nucleotide polymorphisms, copy number variations and methylation patterns, as well as detailed clinical information. The University of California Santa Cruz Genome Browser (UCSC; http://xena.ucsc.edu) (25) serves as a tool to view the recently assembled human genome and performs standardized processing of data from TCGA, the International Cancer Genome Consortium (ICGC) and Therapeutically Applicable Research to Generate Effective Treatments (TARGET) databases.

In the present study, clinical (n=317), RNAseq-counts (n=309) and microRNA data (n=312) were downloaded from TCGA-cervical squamous cell carcinoma (CESC) dataset (26). Subsequently, the data were pre-processed via matching, selecting and deleting data using the UCSC Xena platform. The RNAseq data were then divided into 3 groups according to the patients' age: 1-40 years (84 samples),

41-60 years (157 samples) and >60 years (65 samples). Similarly, the miRNA data were assigned into 3 groups according to the patients' age (as aforementioned), and the number of samples in each group was consistent with the RNAseq data.

TCGA genes re-annotation and data preprocessing. The HUGO Gene Nomenclature Committee (HGNC; http://www.genenames.org) is responsible for providing unique, standard and widespread symbols for human ncRNA genes, protein-encoding genes and other genes. For each gene in the human genome, there is a unique ID and symbol listed in the HGNC database. Gene information of all symbols in the human genome were retrieved from the HGNC database for further data preprocessing. mRNAs and lncRNAs were distinguished from RNA-seq data for subsequent analysis.

The ENSG (European Neuroblastoma Study Group) number of genes in the RNAseq expression profiles from HGNC corresponded to the gene symbol (27,28). Genes that did not match were deleted. The clinical data were used for the subsequent survival analysis if they matched the RNA-seq expression profile. Analogously, the miRNAs were used for subsequent analysis if the samples were matched with RNA-seq expression profile samples and clinical samples.

RNA-seq Mfuzz time clustering analysis. mRNA, lncRNA and miRNA expression matrices were obtained from TCGA database, and the mean of the expression matrix was calculated according to each age group. mRNA data was eliminated if the standard deviation (SD) was <0.3; otherwise, lncRNA and miRNA were eliminated if their SD was <0.1. The Mfuzz package in R software (version 3.2.3) (https://www.r-project.org) was applied to perform soft clustering analysis of noise resistance to the degeneration of DEGs (29). In this process, the Fuzzy C-Means clustering method was adopted to conduct clustering analysis based on the variation of time and the change in expression levels. Finally, multiple clustering results were obtained and those with consistent expression changes were placed in the same group. DEmRNAs, DElncRNAs, and DEmiRNAs were identified in the group according to the difference in expression value (P<0.05 was used as the cut-off value).

Enrichment analyses. The clusterProfiler (version 3.2.11; http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html) package of R software was used to conduct functional Gene Ontology (http://geneontology.org) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway (30) functional enrichment analyses for the DEmRNAs. P<0.05 was considered to indicate a statistically significant difference.

Protein-protein interaction (PPI) analysis. The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING; version 10.0; http://www.string-db.org/) database was used to search and predict a PPI network. The database can be used for >2,000 species, containing 9.6 million different proteins and 13.8 million types of protein interactions. Notably, the PPI network included direct physical interactions as well as indirect functional correlations (31). According to the RNA-seq Mfuzz time clustering analysis, mRNAs that were significantly

Table I. Mfuzz analysis result for cervical cancer.

Data type	Cluster(s), n	Group	Cluster numbers	Gene number(s)
DEmRNA	14	Upregulated	1	30
		Downregulated	6	239
DElncRNA	11	Upregulated	2	131
		Downregulated	2	143
DEmiRNA	16	Upregulated	4	8
		Downregulated	5	8

DE, differentially expressed; lncRNA, long non-coding RNA; miRNA, microRNA.

associated with the age of patients with cervical cancer were selected and used to analyze the interaction between the mRNAs identified in the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (http://www.string-db. org). mRNA data with a standard deviation value <0.3 was eliminated. P<0.05 was considered to indicate a statistically significant difference.

Co-expression analysis of DEmRNAs and DElncRNAs. mRNA and lncRNA expression matrices were obtained via RNA-seq Mfuzz time clustering analysis and co-expression was calculated using the Pearson correlation coefficient method (6). Using DEmRNAs and DElncRNAs data, the co-expression association was calculated and the mRNA-lncRNA co-expressed pairs were obtained. P<0.05 and an absolute correlation coefficient value |r| >0.4 were considered to indicate a significant correlation. An r-value >0 was considered to indicate of a positive correlation, whereas r <0 was considered as a representative of a negative correlation. The correlation strength was highest when the |r|-value was closest to 1.

Prediction of lncRNA and mRNA targeted by miRNA. miRWalk (32,33), RNA22 (34), miRanda (http://miranda.org. uk), Targetscan (http://www.targetscan.org/vert_72), miRDB (http://www.mirdb.org) and miRMap tools (http://mirmap. ezlab.org) were applied to predict mRNAs targeting miRNA. The miRWalk, miRanda, Targetscan and RNAhybrid tools were also used for predicting lncRNAs targeting miRNA. The target mRNAs and lncRNAs were arbitrarily selected if they were predicted by >6 and >2 retrieval tools among the aforementioned tools, respectively. The parameter settings in the miRWalk tool were selected as miRNA and official symbol. Species was set as human and the database retrieval association was set as 'OR' across all databases.

ceRNA network construction and mRNA enrichment analysis. The miRNA-mRNA, miRNA-lncRNA and co-expressed lncRNA-mRNA pairs were identified as aforementioned. Subsequently, lncRNAs and target genes that were associated with the same miRNA were selected, and the lncRNA-miRNA-mRNA ceRNA network was constructed using Cytoscape software (version 3.4.0) (35,36). Thereafter, functional analyses for crucial mRNAs located in the ceRNA network were conducted as aforementioned. P<0.05 was considered to indicate a statistically significant difference.

Survival analysis. The optimal cut-off of key genes was obtained according to the expression value, survival time and survival state based on Survminer package (version 0.4.2) (https://www. rdocumentation.org/packages/survminer/versions/0.4.2) of R software. Genes were considered to be upregulated if the expression value was greater than the optimal cut-off value; conversely, genes lower than the cut-off value were considered to be downregulated. Subsequently, the survival analysis and survival test were performed utilizing survival in the R package (version 2.42-6). The significance cut-off was set as P<0.05. mRNAs with a significant correlation between survival analysis and prognosis were selected and the significant results of survival analysis were plotted on the Kaplan-Meier curve using SPSS software (version 16.0).

Drug association analysis. The prediction of targeted drugs was performed using the Drug-Gene Interaction (DGIdb) online database (http://www.dgidb.org/search_interactions) based on the pre-identified mRNAs located in the ceRNA network and using the following parameter settings, FDA approved and antineoplastic. The interaction network of drug-target genes was constructed using Cytoscape software.

Statistical analysis. Statistical analyses were performed using SPSS software (version 16.0). The Pearson correlation coefficient method was used to analyze co-expression between DEmRNAs and DElncRNAs. Data are presented as the mean ± standard deviation. The independent-samples student t-test was used to compare data between two groups and the χ^2 test was used to compare the count data groups. P<0.05 was considered to indicate a statistically significant difference.

Results

Mfuzz time clustering analysis of cervical cancer progression in different age groups. The design ideas and analysis process are detailed in Fig. S1. The Mfuzz analysis results for cervical cancer are presented in Table I. The results show that a total of 14, 11 and 16 clusters were identified for DEmRNA, DElncRNA, and DEmiRNA data, respectively. There were 269 DEmRNAs; 30 upregulated [including alcohol dehydrogenase 7 (*ADH7*)] and 239 downregulated [including vestigial-like family member 3 (*VGLL3*) and cytochrome P450, family 26, subfamily B, polypeptide 1 (*CYP26B1*)]. A total of 274 DElncRNAs were identified; 131 upregulated

Table II. Hub nodes in the protein-protein interaction network.

| Gene symbol	Degree	Betweenness	Closeness
CSF2	14.0	4759.71880	0.052211303
KLK3	11.0	4843.13800	0.051845074
WNT3A	11.0	4234.45750	0.051328503
HIST1H2BD	10.0	187.71416	0.050535080
HIST1H2BJ	10.0	187.71416	0.050535080
HIST1H2BO	10.0	187.71416	0.050535080
HIST1H2BM	10.0	187.71416	0.050535080
FOXJ1	9.0	3495.36670	0.049985297
CDKN2A	9.0	4913.64550	0.052372150
HOXB1	9.0	1055.40840	0.050073640
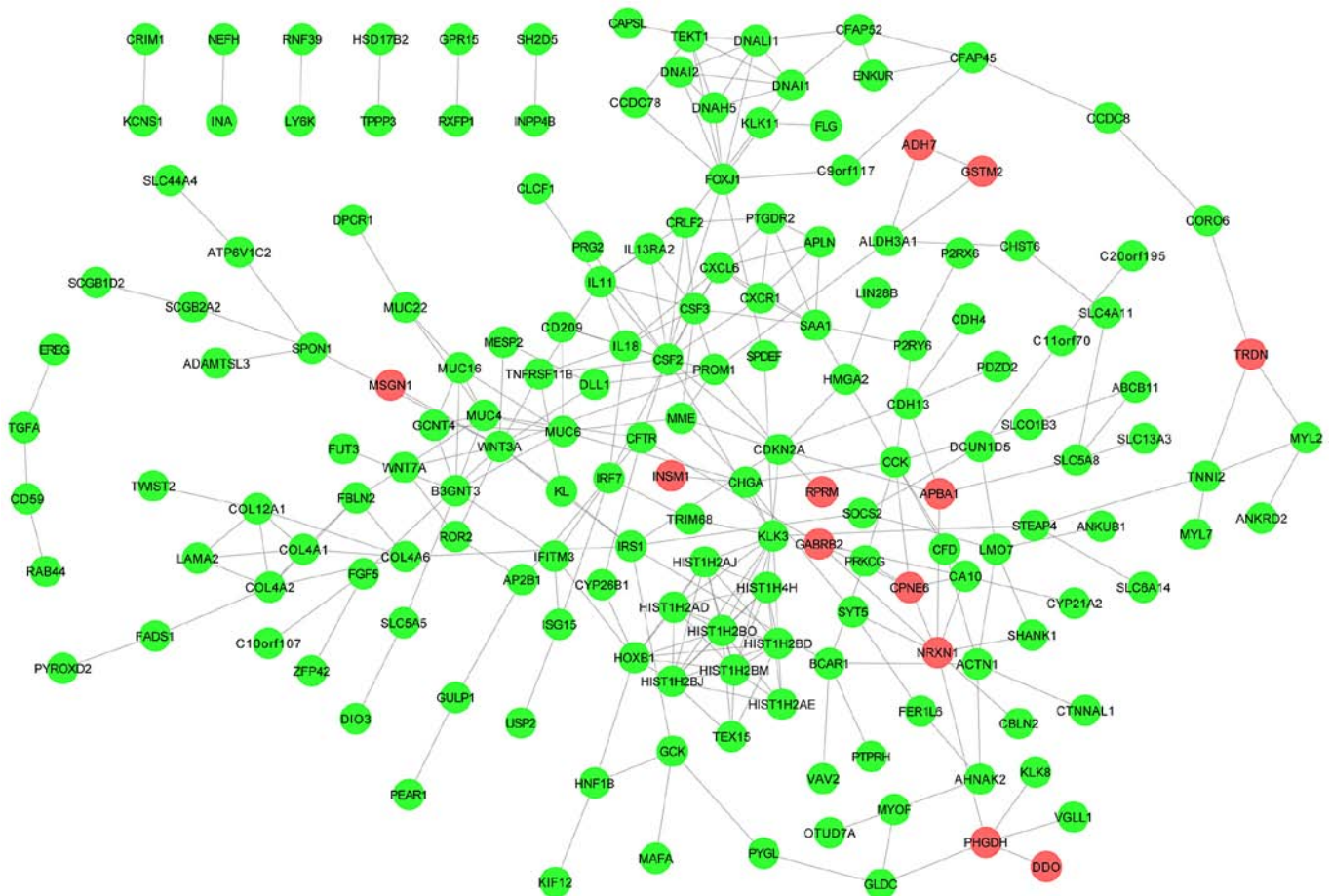
Figure 1. PPI network for DEmRNAs. A PPI network was constructed for the pre-identified DEmRNAs using the Search Tool for the Retrieval of Interacting Genes/Proteins database. Red and green circles represent the upregulated and downregulated mRNA, respectively. There are 292 interaction pairs and 173 nodes in the PPI network. DEmRNAs, differentially expressed mRNAs; PPI, protein-protein interaction.

(e.g. LINC01133) and 143 downregulated, and 16 DEmiRNAs were discovered; 8 upregulated (including miR-499a) and 8 downregulated (including miR-3065 and -330) (Table SI). In particular, *ADH7* expression was upregulated in patients >60 years, while *VGLL3* and *CYP26B1* expression were downregulated in the older age groups (Table SII).

PPI network with age of DEmRNAs in cervical cancer. A total of 292 interaction pairs and 173 nodes identified in the PPI network (Fig. 1). The hub nodes are presented in Table II.

Enrichment analyses with age change of DEmRNAs in cervical cancer. A total of 57 functional terms were obtained
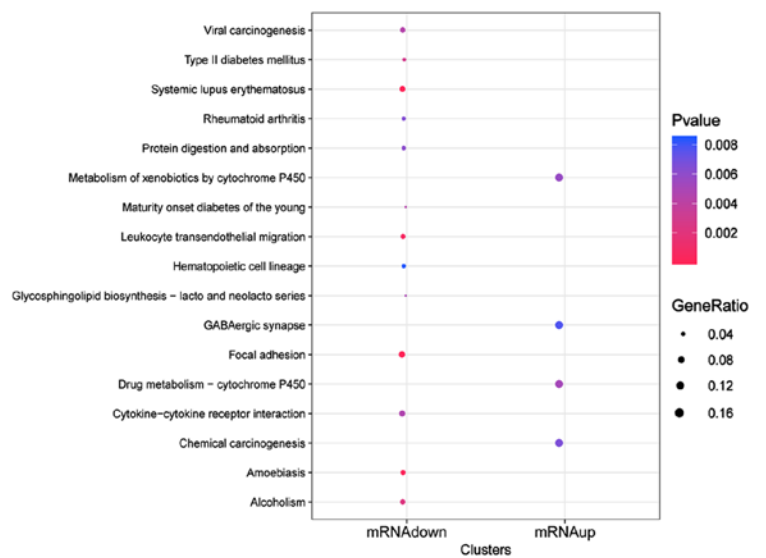
Figure 2. R package clusterProfiler was applied to perform functional GO and KEGG pathway enrichment analyses for the DEmRNAs. The horizontal axis refers to downregulated or upregulated mRNAs and the vertical axis represents the detailed terms of (A) GO-BP function or (B) KEGG pathway. The size of the dot represents the percentage of the total number of enriched mRNAs. The color gradient correlates to significance, whereby red indicates a higher significance. A total of 57 functional terms are obtained for DEmRNAs (panel A). DEmRNAs are enriched in 17 KEGG pathways (panel B). GO, Gene Ontology; BP, biological processes; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEmRNAs, differentially expressed mRNAs.

for the DEmRNAs (Fig. 2A). For upregulated DEmRNAs, the top 10 enriched functions are listed in Table III, such as 'cofactor binding', 'modified amino acid binding' (eg., *ADH7*, etc.) and 'oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as an acceptor'. Similarly, the top 10 enriched functions for downregulated DEmRNAs included 'extracellular matrix structural constituent', 'receptor ligand activity' and 'growth factor receptor binding'. DEmRNAs were enriched in 17 KEGG pathways (Fig. 2B). Upregulated DEmRNAs were enriched in 4 KEGG pathways; 'GABAergic synapse', 'drug metabolism-cytochrome P450', 'metabolism of xenobiotics by cytochrome P450' and 'chemical carcinogenesis'

Table III. GO enrichment analysis for differentially expressed mRNAs.

A, Upregulated

Term	Description	P-value	Key genes
GO:0048037	Cofactor binding	4.77×10^{-03}	CYP4Z1, DDO
GO:0072341	Modified amino acid binding	5.07×10^{-03}	CPNE6, GSTM2
GO:0016616	Oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	1.28×10^{-02}	ADH7, PHGDH
GO:0140161	Monocarboxylate:sodium symporter activity	1.42×10^{-02}	SLC6A12
GO:0033130	Acetylcholine receptor binding	1.56×10^{-02}	NRXN1
GO:0043295	Glutathione binding	1.56×10^{-02}	GSTM2
GO:0016614	Oxidoreductase activity, acting on CH-OH group of donors	1.63×10^{-02}	ADH7, PHGDH
GO:1900750	Oligopeptide binding	1.70×10^{-02}	GSTM2
GO:0015355	Secondary active monocarboxylate transmembrane transporter activity	1.84×10^{-02}	SLC6A12
GO:0019841	Retinol binding	2.12×10^{-02}	ADH7

B, Downregulated

Term	Description	P-value	Key genes
GO:0005201	Extracellular matrix structural constituent	2.56×10^{-06}	MUC4, FBLN2
GO:0048018	Receptor ligand activity	2.83×10^{-05}	IL18, SAA1
GO:0070851	Growth factor receptor binding	1.39×10^{-04}	CSF2, FGF5
GO:0045504	Dynein heavy chain binding	4.77×10^{-04}	DNAI2, DNAI1
GO:0005125	Cytokine activity	9.35×10^{-04}	IL18, WNT7A
GO:0008509	Anion transmembrane transporter activity	1.86×10^{-03}	ABCB11, AQP6
GO:0015081	Sodium ion transmembrane transporter activity	2.44×10^{-03}	ABCB11, NALCN
GO:0045503	Dynein light chain binding	2.44×10^{-03}	DNAI2, DNAH5
GO:0008083	Growth factor activity	2.61×10^{-03}	CLCF1, CSF2
GO:0022804	Active transmembrane transporter activity	3.41×10^{-03}	ATP2C2, SLC13A3

(Table IV). Downregulated DEGs were enriched in 13 KEGG pathways, (Table IV).

ceRNA network construction. In total, 632 mRNA-lncRNA interactions were identified according to the aforementioned method, including 169 mRNAs and 106 lncRNAs (Table SIII). A total of 1,841 pairs of miRNA-mRNA interactions (consisting of 1,518 mRNAs and 13 miRNAs) were also identified, followed by a total of 15 overlapped mRNAs and 7 corresponding miRNAs derived from the miRNA-mRNA and mRNA-lncRNA interactions. In addition, 90,026 pairs of miRNA-lncRNA interactions were extracted, including 15,069 lncRNAs and 16 miRNAs. Similarly, 102 overlapped lncRNAs and 15 corresponding miRNAs were extracted from those miRNA-lncRNA and lncRNA-mRNA interactions. Finally, the lncRNA-miRNA-mRNA ceRNA network was established utilizing those 102 lncRNAs, 15 miRNAs, 15 mRNAs and 522 interaction pairs (Fig. 3).

Enrichment analysis with age change of crucial mRNAs in cervical cancer. Enrichment analysis for crucial mRNA located in the ceRNA network was implemented; the results identified 20 functional terms (Fig. 4A) and 5 KEGG pathways (Fig. 4B). Here, the top 10 enriched functions of DEmRNAs are listed in Table VA, such as the 'retinoic acid metabolic process' (e.g., *ADH7*, *CYP26B1*, etc.), 'regulation of stem cell population maintenance' and 'endodermal cell differentiation'. In addition, the top 10 enriched KEGG pathways for DEmRNAs are listed in Table VB, including 'retinol metabolism', 'glycosaminoglycan degradation' and maturity onset diabetes of the young.

Crucial mRNA survival analysis and drug association analysis. In total, the survival data of 265 samples were obtained and the results are presented in Tables VI and SIV. In particular, the crucial mRNAs *ADH7* (Fig. 5A) and *VELL3* (Fig. 5B) were significantly associated with cervical cancer survival rate (P<0.05). The prediction of targeted drugs was performed for *ADH7* and *VGLL3*. While no associated drugs were revealed to be associated with *VGLL3*, two drugs (CHEMBL1161866 and ethanol) were associated with *ADH7*. This allows for development of intervention targets for *ADH7*.

Table IV. KEGG pathway enrichment analysis for differentially expressed mRNAs.

A, Upregulated

Term	Description	P-value	Key genes
hsa00982	Drug metabolism-cytochrome P450	5.15×10^{-03}	ADH7, GSTM2
hsa00980	Metabolism of xenobiotics by cytochrome P450	5.72×10^{-03}	ADH7, GSTM2
hsa05204	Chemical carcinogenesis	6.63×10^{-03}	ADH7, GSTM2
hsa04727	GABAergic synapse	7.77×10^{-03}	GABRB2, SLC6A12

B, Downregulated

Term	Description	P-value	Key genes
hsa04510	Focal adhesion	4.94×10^{-05}	VAV2, ACTN1
hsa05322	Systemic lupus erythematosus	5.25×10^{-05}	HIST1H2AE, HIST1H2BO
hsa05146	Amoebiasis	2.19×10^{-04}	CSF2, PRKCG
hsa04670	Leukocyte transendothelial migration	5.99×10^{-04}	VAV2, ACTN1
hsa05034	Alcoholism	2.27×10^{-03}	HIST1H2AE, HIST1H2BO
hsa04930	Type II diabetes mellitus	2.87×10^{-03}	MAFA, IRS1
hsa04950	Maturity onset diabetes of the young	4.44×10^{-03}	MAFA, GCK
hsa05203	Viral carcinogenesis	4.47×10^{-03}	CDKN2A, IRF7
hsa04060	Cytokine-cytokine receptor interaction	4.68×10^{-03}	IL18, CLCF1
hsa00601	Glycosphingolipid biosynthesis-lacto and neolacto series	4.95×10^{-03}	FUT3, B3GNT3
hsa04974	Protein digestion and absorption	6.13×10^{-03}	MME, COL4A6
hsa05323	Rheumatoid arthritis	6.42×10^{-03}	IL18, CSF2
hsa04640	Hematopoietic cell lineage	8.37×10^{-03}	CSF2, CSF3

GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Discussion

Cervical cancer represents a leading cause of tumor-associated mortality among women, seriously threatening the health of females (2). It has been revealed that the prevalence of HPV and cervical cancer among women may be associated with their age (8). Thus, it is essential to investigate the molecular mechanism underlying cervical cancer in different ages. In total, 14 DEmiRs, 11 DElncRNAs and 16 DEmiRNAs were identified based on the different age groups in the present study. Additionally, 102 lncRNAs, 15 miRNAs and 15 mRNAs were confirmed in the ceRNA network. In particular, *ADH7* was regulated by miR-3065, and miR-3065 was associated with LINC01133 in the ceRNA network. Furthermore, DEmRNAs located in the ceRNA network were involved in 5 KEGG pathways, such as 'retinol metabolism' (*ADH7* and *CYP26B1*). Analogously, DEmRNAs were enriched in several functions, such as the 'terpenoid metabolic process', 'stem cell population maintenance' and 'retinoic acid metabolic process' (*ADH7* and *CYP26B1*). Furthermore, *ADH7* and *VGLL3* were significantly associated with cervical cancer survival rate.

It has been revealed that retinoids have the ability to regulate cellular differentiation and proliferation, as well as epithelial-to-mesenchymal transition via specific nuclear retinoic-acid receptors (37,38). Retinoids are considered as inhibitors of cervical cancer progression due to their interaction with the HPV proteins E6 and E7 (38). Furthermore, Gariglio *et al* (39) demonstrated that retinoid deficiency is associated with cervical squamous metaplasia. In a recent study, lecithin retinol acyltransferase and retinol dehydrogenase 12, which are associated with retinol metabolism, have been identified to suppress the apoptosis, migration and invasiveness of cervical cancer cells (40). These data from the present study indicated that genes correlated with retinol metabolism may influence cervical cancer progression. Whereas *CYP26B1* was downregulated in cervical cancer samples, *ADH7* was upregulated. Notably, in the current study, both *CYP26B1* and *ADH7* were revealed to influence the retinol metabolism pathway and the retinoic acid metabolic process. *ADH7* is as a member of the alcohol dehydrogenase family and, thus, may be involved in the synthesis of retinoic acid, which is a hormone crucial for cellular differentiation (41). It has been reported that the upregulation of *CYP26B1* may result in the progression of cervical malignancies (42). Following survival analysis, *ADH7* was revealed to be significantly associated with the cervical cancer survival rate. Therefore, it was speculated that *ADH7* and *CYP26B1* participate in cervical cancer progression via the regulation of the retinol metabolism pathway and the retinoic
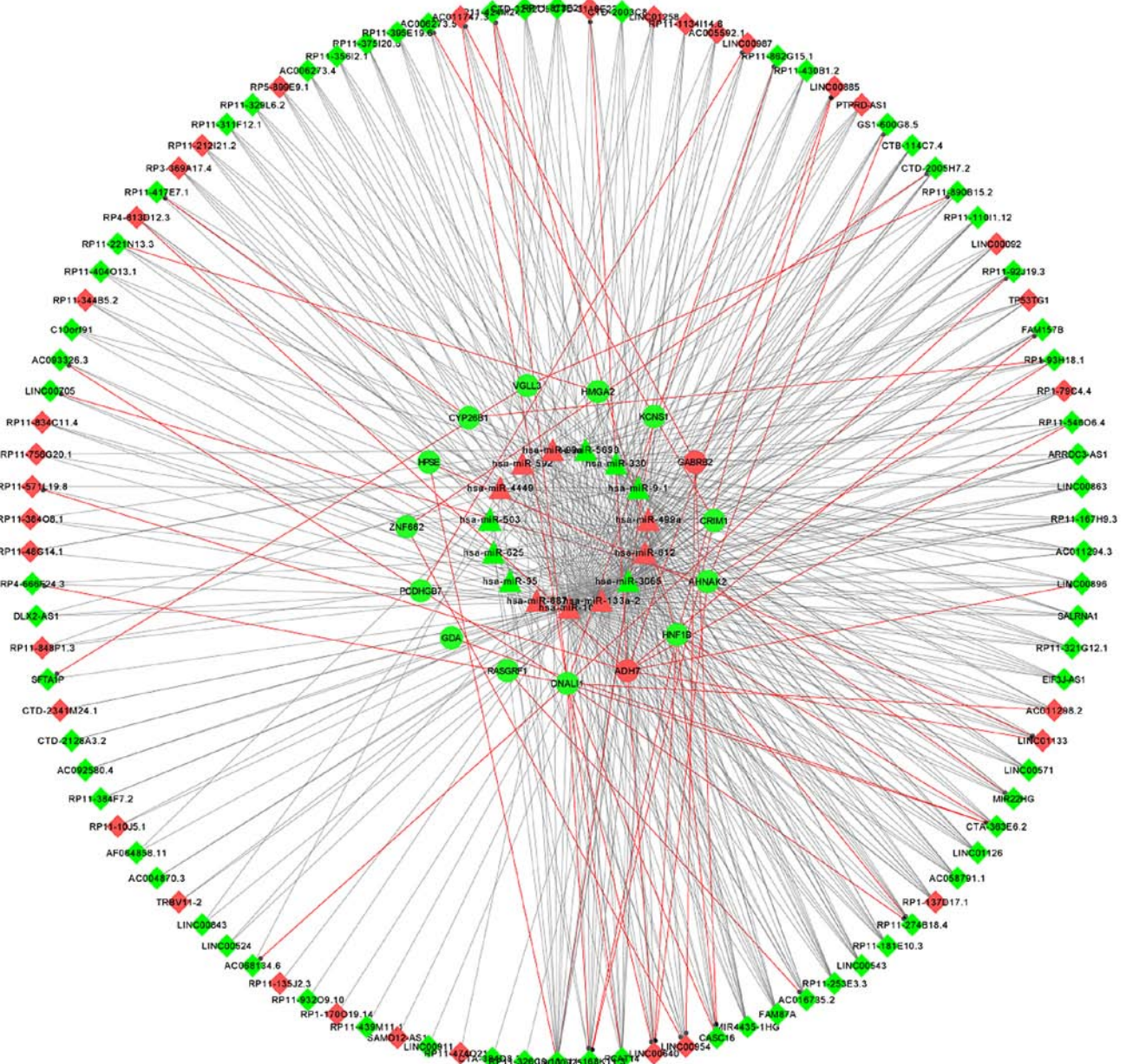
Figure 3. ceRNA regulatory network comprising miRNAs, mRNAs and lncRNAs. miRNA-gene, miRNA-lncRNA and co-expressed lncRNA-mRNA pairs were identified and the lncRNA-miRNA-mRNA ceRNA network was constructed. Thereafter, functional analyses for crucial mRNAs located in the ceRNA network was performed. Red and green colors represent upregulated and downregulated genes, respectively. Triangles, circles and diamonds represent miRNAs, mRNAs and lncRNAs, respectively. The red line and the grey line refer to the pairs of lncRNA-mRNA interactions and miRNA-mRNA interactions, respectively. ceRNA, competing endogenous RNA; miRNA, microRNA; lncRNA, long non-coding RNA.

acid metabolic process. In particular, *ADH7* expression levels gradually increased with an increase in patient age, while *CYP26B1* expression gradually declined with a decrease in patient age, indicating that *ADH7* and *CYP26B1* are correlated with the differential progression of cervical cancer between different age groups.

It has been revealed that miR-499a influences the tumorigenesis of several different types of cancer, including breast cancer (43) and hepatocellular carcinoma (44). In addition, miR-499a contributes to the migration and proliferation of hepatitis B virus-associated hepatocellular carcinoma (HCC) cells (45). Recently, miR-499a-5p has been observed to promote migration, proliferation and epithelial-to-mesenchymal transition in metastatic lung cancer cell lines, via

modulation of the mTOR pathway (46); however, the functions of miR-499 in cervical cancer development have not yet been fully elucidated. In the present study, miR-299a expression was upregulated in cervical cancer samples and was revealed to downregulate *CYP26B1* expression. Thus, miR-499a may be associated with the tumorigenesis of cervical cancer via regulation of *CYP26B1*.

A previous study has suggested that downregulation of lncRNA WT1 inhibited cervical carcinoma cell migration and invasiveness via the regulation of miR-330-5p, indicating that miR-330-5p may influence cervical cancer progression (47). Although, to the best of our knowledge, there is currently no evidence that vestigial-like family member 3 (*VGLL3*) is associated with cervical cancer progression, the present study
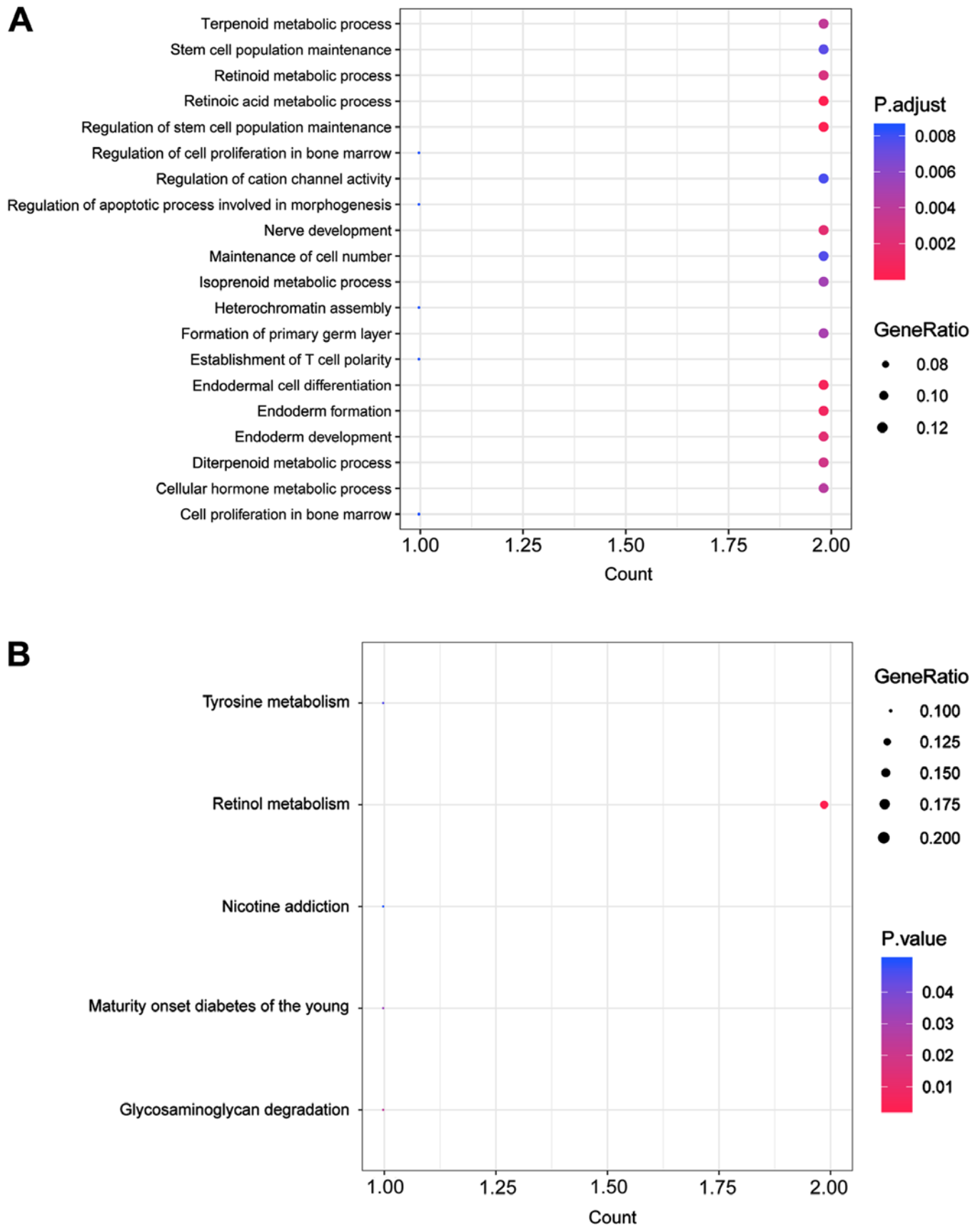
Figure 4. Enrichment analysis for crucial mRNA located in the ceRNA network was implemented. The R package clusterProfiler was applied to perform functional GO and KEGG pathway enrichment analyses for the DEmRNAs. The horizontal axis refers to the count of mRNAs, and the vertical axis represents the detailed terms of (A) GO-BP function or (B) KEGG pathway. The size of the dot represents the percentage of the total number of enriched mRNAs. The color gradient correlates to significance, whereby red indicates a higher significance. ceRNA, competing endogenous RNA; GO, Gene Ontology; BP, biological processes; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEmRNAs, differentially expressed mRNAs.

Table V. Enrichment analysis for crucial differentially expressed mRNAs.

A, GO analysis

Term	Description	P-value	Count	Key genes
GO:0042573	Retinoic acid metabolic process	1.84×10^{-04}	2	ADH7, CYP26B1
GO:2000036	Regulation of stem cell population maintenance	2.3×10^{-04}	2	HNF1B, HMGA2
GO:0035987	Endodermal cell differentiation	6.83×10^{-04}	2	HNF1B, HMGA2
GO:0001706	Endoderm formation	9.75×10^{-04}	2	HNF1B, HMGA2
GO:0021675	Nerve development	1.90×10^{-03}	2	AHNAK2, GABRB2
GO:0007492	Endoderm development	2.00×10^{-03}	2	HNF1B, HMGA2
GO:0001523	Retinoid metabolic process	2.58×10^{-03}	2	ADH7, CYP26B1
GO:0016101	Diterpenoid metabolic process	3.00×10^{-03}	2	ADH7, CYP26B1
GO:0006721	Terpenoid metabolic process	3.76×10^{-03}	2	ADH7, CYP26B1
GO:0034754	Cellular hormone metabolic process	4.18×10^{-03}	2	ADH7, CYP26B1

B, KEGG pathway analysis

Term	Description	P-value	Count	Key genes
hsa00830	Retinol metabolism	3.08×10^{-03}	2	ADH7, CYP26B1
hsa00531	Glycosaminoglycan degradation	2.39×10^{-02}	1	HSPE
hsa04950	Maturity onset diabetes of the young	3.26×10^{-02}	1	HNF1B
hsa00350	Tyrosine metabolism	4.49×10^{-02}	1	ADH7
hsa05033	Nicotine addiction	4.98×10^{-02}	1	GABRB2
hsa00071	Fatty acid degradation	5.46×10^{-02}	1	ADH7
hsa00010	Glycolysis/Gluconeogenesis	8.33×10^{-02}	1	ADH7
hsa00982	Drug metabolism-cytochrome P450	8.33×10^{-02}	1	ADH7
hsa00982	Drug metabolism-cytochrome P450	8.80×10^{-02}	1	ADH7

Table VI. Survival analysis for differentially expressed mRNAs.

Genes	P-value	High median	Low median
ADH7	2.44×10^{-02a}	136.2000000	N/A
AHNAK2	5.20×10^{-01}	101.5333333	N/A
CRIM1	5.15×10^{-01}	N/A	101.53333330
CYP26B1	7.59×10^{-01}	103.2333333	101.53333330
DNALI1	3.05×10^{-01}	136.2000000	101.53333330
GABRB2	8.53×10^{-01}	103.2333333	136.20000000
GDA	6.01×10^{-01}	103.2333333	N/A
HMGA2	8.94×10^{-01}	136.2000000	101.53333330
HNF1B	6.79×10^{-01}	103.2333333	136.20000000
HPSE	5.82×10^{-02}	N/A	103.23333330
KCNS1	2.00×10^{-01}	N/A	96.26666667
PCDHGB7	5.13×10^{-01}	136.2000000	101.53333330
RASGRF1	6.53×10^{-01}	NA	96.26666667
VGLL3	2.87×10^{-02a}	84.0000000	N/A
ZNF662	7.24×10^{-01}	103.2333333	136.2

N/A, not applicable. The data was missing within the database or incomputable.

indicated that *VGLL3* is regulated by miR-330 and was significantly correlated with cervical cancer survival rate. Therefore, it is speculated that *VGLL3* may influence cervical cancer carcinogenesis via the regulation of miR-330 expression.
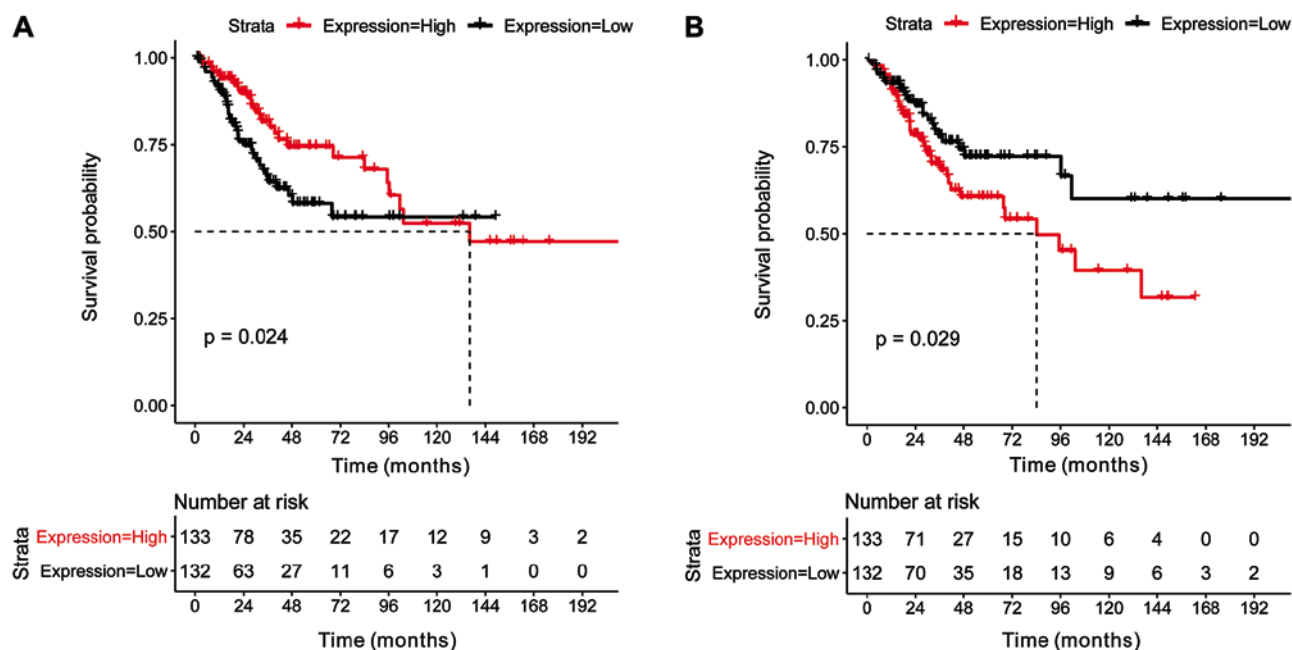
Figure 5. Survival curves of (A) ADH7 and (B) VGLL3 clinical data from The Cancer Genome Atlas database were used to perform the survival analysis. mRNAs with a significant association between survival analysis and prognosis were selected and the significant results of survival analysis were plotted on the survival curve. The horizontal axis represents survival time and the vertical axis represents the survival rate. Red refers to the high-expression samples, while black refers to the low-expression samples. ADH7, alcohol dehydrogenase 7; VGLL3, vestigial-like family member 3.

Notably, *VGLL3* expression gradually declined with decreasing patient age, suggesting that *VGLL3* may be associated with the progression of cervical cancer in different age groups.

Numerous studies have indicated that the lncRNA LINC01133 serves as a regulator of various cancer types; for example, LINC01133 was revealed to suppress the invasiveness, migration and proliferation of nasopharyngeal carcinoma cells, via interaction with Y box binding protein 1 (48). Furthermore, Zheng *et al* (49) demonstrated that LINC01133 contributes to the progression of HCC via activation of the PI3K/AKT pathway. It has also been identified that LINC01133 was upregulated in lung squamous cell cancer cells and was associated with poorer overall survival outcomes (50). Finally, LINC01133 was determined to be associated with the survival time of patients with cervical squamous cell carcinoma (26). In the present study, LINC01133 was upregulated in cervical cancer samples and was demonstrated to interact with miR-3065-5p. Currently, the role of miR-3065 in cervical cancer progression is yet to be elucidated. However, a previous study indicated that miR-3065-5p expression is associated with the tumorigenesis of clear cell renal cell carcinomas (51). Notably, in the present study miR-3065 expression was downregulated in cervical cancer samples and was revealed to regulate *ADH7* expression. Hence, LINC01133 may serve as a ceRNA, via interaction with miR-3065, to upregulate *ADH7* expression, resulting in the progression of cervical cancer.

Overall, *ADH7*, *VGLL3*, *CYP26B1*, miR-3065, miR-330, miR-499a and LINC01133 may influence the progression of cervical cancer in different age groups. These genes may therefore represent novel targets for the diagnosis or treatment of patients with cervical cancer. There limitations of the present study include insufficient sample size, lack of clinical practice and lack of research on the mechanism of intermolecular interactions. For future studies, the predicted results should be validated using experimental data, and the molecular mechanism of cervical cancer carcinogenesis requires further clarification.

Acknowledgements

Funding

Availability of data and materials

The datasets used and/or analyzed during the present study are available from the corresponding author upon reasonable request.

Authors' contributions

SD, XH, JZ, BX, LX, DG and WZ contributed to the design of the present study. SD, XH and WZ drafted the initial manuscript. JZ, BX, LX and DG interpreted the data. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Cotto KC, Wagner AH, Feng YY, Kiwala S, Coffman AC, Spies G, Wollam A, Spies NC, Griffith OL and Griffith M: DGIdb 3.0: A redesign and expansion of the drug-gene interaction database. Nucleic Acids Res 46: D1068-D1073, 2018.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68: 394-424, 2018.
3. de Lourdes Mora-García M, López-Cisneros S, Gutiérrez-Serrano V, García-Rocha R, Weiss-Steider B, Hernández-Montes J, Sánchez-Peña HI, Ávila-Ibarra LR, Don-López CA, Muñóz-Godínez R, *et al*: HPV-16 infection is associated with a high content of CD39 and CD73 ectonucleotidases in cervical samples from patients with CIN-1. Mediators Inflamm 7: 4651627, 2019.
4. Kjær SK, Munk C, Junge J and Iftner T: Carcinogenic HPV prevalence and age-specific type distribution in 40,382 women with normal cervical cytology, ASCUS/LSIL, HSIL, or cervical cancer: What is the potential for prevention? Cancer Causes Control 25: 179-189, 2014.
5. Bruni L, Diaz M, Castellsagué X, Ferrer E, Bosch FX and de Sanjosé S: Cervical human papillomavirus prevalence in 5 continents: Meta-analysis of 1 million women with normal cytological findings. J Infect Dis 202: 1789-1799, 2010.
6. Wheeler CM, Hunt WC, Cuzick J, Langsfeld E, Pearse A, Montoya GD, Robertson M, Shearman CA and Castle PE; New Mexico HPV Pap Registry Steering Committee: A population-based study of human papillomavirus genotype prevalence in the United States: Baseline measures prior to mass human papillomavirus vaccination. Int J Cancer 132: 198-207, 2013.
7. Zhao P, Liu S, Zhong Z, Hou J, Lin L, Weng R, Su L, Lei N, Hou T and Yang H: Prevalence and genotype distribution of human papillomavirus infection among women in northeastern Guangdong province of China. BMC Infect Dis 18: 204, 2018.
8. Luo G, Sun X, Li M, Liu T, Hu G, He Y, Mao L, Yan L, Xie L, Zou H and Luo X: Cervical human papillomavirus among women in Guangdong, China 2008-2017: Implication for screening and vaccination. J Med Virol 91: 1856-1865, 2019.
9. Wu C, Zhu X, Kang Y, Cao Y, Lu P, Zhou W, Zhou H, Zhang Y and Song Y: Epidemiology of humanpapilloma virus infection among women in Fujian, China. BMC Public Health 18: 95, 2017.
10. Zhao FH, Tiggelaar SM, Hu SY, Xu LN, Hong Y, Niyazi M, Gao XH, Ju LR, Zhang LQ, Feng XX, *et al*: A multi-center survey of age of sexual debut and sexual behavior in Chinese women: Suggestions for optimal age of human papillomavirus vaccination in China. Cancer Epidemiol 36: 384-390, 2012.
11. Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, de Sanjose S, Naucler P, Lloveras B, Kjaer S, *et al*: Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: Joint European cohort study. BMJ 337: 969-972, 2008.
12. Biewenga P, van der Velden J, Mol BW, Stalpers LJ, Schilthuis MS, van der Steeg JW, Burger MP and Buist MR: Prognostic model for survival in patients with early stage cervical cancer. Cancer 117: 768-776, 2011.
13. Kumar L, Pramanik R, Kumar S, Bhatla N and Malik S: Neoadjuvant chemotherapy in gynaecological cancers - Implications for staging. Best Pract Res Clin Obstet Gynaecol 29: 790-801, 2015.
14. Li N, Guo X, Liu L, Wang L and Cheng R: Molecular mechanism of miR-204 regulates proliferation, apoptosis and autophagy of cervical cancer cells by targeting ATF2. Artif Cells Nanomed Biotechnol 47: 2529-2535, 2019.
15. Zhao YB, Wang JH, Chen XX, Wu YZ and Wu Q: Values of three different preoperative regimens in comprehensive treatment for young patients with stage Ib2 cervical cancer. Asian Pac J Cancer Prev 13: 1487-1489, 2012.
16. Beresneva EV, Rykov SV, Hodyrev DS, Pronina IV, Ermilova VD, Kazubskaia TP, Braga EA and Loginov VI: Methylation profile of group of miRNA genes in clear cell renal cell carcinoma; involvement in cancer progression. Genetika 49: 366-375 (In Russian).
17. White NM, Khella HW, Grigull J, Adzovic S, Youssef YM, Honey RJ, Stewart R, Pace KT, Bjarnason GA, Jewett MA, *et al*: MiRNA profiling in metastatic renal cell carcinoma reveals a tumour-suppressor effect for miR-215. Br J Cancer 105: 1741-1749, 2011.
18. Yang D and Zhang Q: MiR-152 may function as an early diagnostic and prognostic biomarker in patients with cervical intraepithelial neoplasia and patients with cervical cancer. Oncol Lett 17: 5693-5698, 2019.
19. Sun H, Fan G, Deng C and Wu L: MiR-4429 sensitized cervical cancer cells to irradiation by targeting RAD51. J Cell Physiol 23: 185-193, 2020.
20. Mercer TR, Dinger ME and Mattick JS: Long non-coding RNAs: Insights into functions. Nat Rev Genet 10: 155-159, 2009.
21. Sun L, Luo H, Liao Q, Bu D, Zhao G, Liu C, Liu Y and Zhao Y: Systematic study of human long intergenic non-coding RNAs and their impact on cancer. Sci China Life Sci 56: 324-334, 2013.
22. Peng L, Yuan X, Jiang B, Tang Z and Li GC: LncRNAs: Key players and novel insights into cervical cancer. Tumour Biol 37: 2779-2788, 2016.
23. Zhang JJ and Fan LP: Long non-coding RNA CRNDE enhances cervical cancer progression by suppressing PUMA expression. Biomed Pharmacother 117: 108726, 2019.
24. Yang XZ, Cheng TT, He QJ, Lei ZY, Chi J, Tang Z, Liao QX, Zhang H, Zeng LS and Cui SZ: LINC01133 as ceRNA inhibits gastric cancer progression by sponging miR-106a-3p to regulate APC expression and the Wnt/β-catenin pathway. Mol Cancer 17: 126, 2018.
25. Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, Roskin KM, Schwartz M, Sugnet CW, Thomas DJ, *et al*: The UCSC genome browser database. Nucleic Acids Res 31: 51-54, 2003.
26. Mao X, Qin X, Li L, Zhou J, Zhou M, Li X, Xu Y, Yuan L, Liu QN and Xing H: A 15-long non-coding RNA signature to improve prognosis prediction of cervical squamous cell carcinoma. Gynecol Oncol 149: 181-187, 2018.
27. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Girón CG, *et al*: Ensembl 2018. Nucleic Acids Res 46: 754-761, 2018.
28. Hubbard T, Barker D, Birney E, Cameron G, Chen Y, Clark L, Cox T, Cuff J, Curwen V, Down T, *et al*: The ensembl genome database project. Nucleic Acids Res 30: 38-41, 2002.
29. Kumar L and E Futschik M: Mfuzz: A software package for soft clustering of microarray data. Bioinformation 2: 5-7, 2007.
30. Kanehisa M and Goto S: KEGG: Kyoto encyclopaedia of genes and genomes. Nucleic Acids Res 8: 27-30, 2000.
31. Bolstad BM, Irizarry RA, Astrand M and Speed TP: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 22: 185-193, 2003.
32. Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, Winsor GL, Hancock RE, Brinkman FS and Lynn DJ: InnateDB: Systems biology of innate immunity and beyond-recent updates and continuing curation. Nucleic Acids Res 41: D1228-D1233, 2013.
33. Dweep H and Gretz N: MiRWalk2.0: A comprehensive atlas of microRNA-target interactions. Nat Methods 12: 697, 2015.
34. Megiorni F, Cialfi S, Dominici C, Quattrucci S and Pizzuti A: Synergistic post-transcriptional regulation of the cystic fibrosis transmembrane conductance regulator (CFTR) by miR-101 and miR-494 specific binding. PLoS One 6: e26601, 2011.
35. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK and Kjems J: Natural RNA circles function as efficient microRNA sponges. Nature 495: 384-388, 2013.
36. Paraskevopoulou MD and Hatzigeorgiou AG: Analyzing MiRNA-LncRNA interactions. Methods Mol Biol 1402: 271-286, 2016.
37. Sporn MB and Roberts AB: Role of retinoids in differentiation and carcinogenesis. Cancer Res 43: 3034-3040, 1983.
38. Abu J, Batuwangala M, Herbert K and Symonds P: Retinoic acid and retinoid receptors: Potential chemopreventive and therapeutic role in cervical cancer. Lancet Oncol 6: 712-720, 2005.
39. Gariglio P, Gutiérrez J, Cortés E and Vázquez J: The role of retinoid deficiency and estrogens as cofactors in cervical cancer. Arch Med Res 40: 449-465, 2009.
40. Liu P, Chen N and Yang X: The effect of retinol metabolism related genes LRAT/RDH12 in the progress of squamous cervical cancer cell line. Prog Obstet Gynecol 2018.
41. Jelski W, Chrostek L, Zalewski B and Szmitkowski M: Alcohol dehydrogenase (ADH) isoenzymes and aldehyde dehydrogenase (ALDH) activity in the sera of patients with gastric cancer. Dig Dis Sci 53: 2101-2105, 2008.

42. Osanai M and Lee GH: Increased expression of the retinoic acid-metabolizing enzyme CYP26A1 during the progression of cervical squamous neoplasia and head and neck cancer. BMC Res Notes 7: 1-9, 2014.
43. Kabirizadeh S, Azadeh M, Mirhosseini M, Ghaedi K and Tanha HM: The SNP rs3746444 within mir-499a is associated with breast cancer risk in Iranian population. J Cell Immunother 2: 95-97, 2016.
44. Wang XH, Wang FR, Tang YF, Zou HZ and Zhao YQ: Association of miR-149C>T and miR-499A>G polymorphisms with the risk of hepatocellular carcinoma in the Chinese population. Genet Mol Res 13: 5048-5054, 2014.
45. Xiang Z, Wang S and Xiang Y: Upregulated MicroRNA499a by hepatitis B virus induced hepatocellular carcinogenesis via targeting MAPK6. PLoS One 9: e111410, 2014.
46. He S, Li Z, Yu Y, Zeng Q, Cheng Y, Ji W, Xia W and Lu S: Exosomal miR-499a-5p promotes cell proliferation, migration and EMT via mTOR signaling pathway in lung adenocarcinoma. Exp Cell Res 379: 203-213, 2019.
47. Cui L, Nai M, Zhang K, Li L and Li R: lncRNA WT1-AS inhibits the aggressiveness of cervical cancer cell via regulating p53 expression via sponging miR-330-5p. Cancer Manag Res 11: 651-667, 2019.
48. Zhang W, Du M, Wang T, Chen W, Wu J, Li Q, Tian X, Qian L, Wang Y, Peng F, et al: Long non-coding RNA LINC01133 mediates nasopharyngeal carcinoma tumorigenesis by binding to YBX1. Am J Cancer Res 9: 779-790, 2019.
49. Zheng YF, Zhang XY and Bu YZ: LINC01133 aggravates the progression of hepatocellular carcinoma by activating the PI3K/AKT pathway. J Cell Biochem 120: 4172-4179, 2019.
50. Zhang J, Zhu N and Chen X: A novel long noncoding RNA LINC01133 is upregulated in lung squamous cell cancer and predicts survival. Tumor Biol 36: 7465-7471, 2015.
51. Müller S and Nowak K: Exploring the miRNA-mRNA regulatory network in clear cell renal cell carcinomas by next-generation sequencing expression profiles. Biomed Res Int 2014: 948408-948408, 2014.