RESEARCH   **Open Access**

Bioinformatics analysis of differentially expressed genes and pathways in the development of cervical cancer

Check for updates

Baojie Wu[*] and Shuyi Xi

Abstract

Background: This study aimed to explore and identify key genes and signaling pathways that contribute to the progression of cervical cancer to improve prognosis.

Methods: Three gene expression profiles (GSE63514, GSE64217 and GSE138080) were screened and downloaded from the Gene Expression Omnibus database (GEO). Differentially expressed genes (DEGs) were screened using the GEO2R and Venn diagram tools. Then, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed. Gene set enrichment analysis (GSEA) was performed to analyze the three gene expression profiles. Moreover, a protein–protein interaction (PPI) network of the DEGs was constructed, and functional enrichment analysis was performed. On this basis, hub genes from critical PPI subnetworks were explored with Cytoscape software. The expression of these genes in tumors was verified, and survival analysis of potential prognostic genes from critical subnetworks was conducted. Functional annotation, multiple gene comparison and dimensionality reduction in candidate genes indicated the clinical significance of potential targets.

Results: A total of 476 DEGs were screened: 253 upregulated genes and 223 downregulated genes. DEGs were enriched in 22 biological processes, 16 cellular components and 9 molecular functions in precancerous lesions and cervical cancer. DEGs were mainly enriched in 10 KEGG pathways. Through intersection analysis and data mining, 3 key KEGG pathways and related core genes were revealed by GSEA. Moreover, a PPI network of 476 DEGs was constructed, hub genes from 12 critical subnetworks were explored, and a total of 14 potential molecular targets were obtained.

Conclusions: These findings promote the understanding of the molecular mechanism of and clinically related molecular targets for cervical cancer.

Keywords: Bioinformatics analysis, Cervical intraepithelial neoplasia, Cervical cancer, Differentially expressed genes, Functional enrichments

Background

Human papillomavirus (HPV) infection is a primary cause of cervical cancer and led to 311,365 deaths in 2018 [1]. Cervical intraepithelial neoplasia (CIN) is a potentially premalignant transformation of squamous cells of the cervix. From normal to CIN (N-CIN) and ultimately to cancer (CIN-CC), cervical cancer is a continuous and evolving process [2, 3]. In addition to HPV vaccination, early diagnosis and treatment can reduce the mortality rate of cervical cancer. In recent years, the prognosis of cervical cancer has been a concern. The identification of biomarkers correlated with the diagnosis and prognosis of cervical cancer is of great importance. With the

* Correspondence: w3522021987@163.com; zrwbj@walvax.com
Shanghai Zerun Biotechnology Co., Ltd., Pilot Department, Building 9, 1690 Zhangheng Road Pudong, Shanghai 201203, China

development of bioinformatics, more studies have focused on the signaling and metabolic pathways of cervical cancer, data mining and validation of related biomolecular targets. The aim of this study was to explore and identify the key genes and signaling pathways contributing to the progression of cervical cancer to improve prognosis. An integrated bioinformatics analysis was performed to select differentially expressed genes (DEGs) and hub genes and to investigate their protein–protein interaction (PPI) networks, related prognostic signatures, functional annotations and potential prognostic value. This study may offer better insight into potential molecular mechanisms to explore preventive and therapeutic strategies.

Methods

Data processing

CIN is an important process in the development of cervical cancer. The gene expression profiles related to CIN progression were retrieved and downloaded from the Gene Expression Omnibus (GEO) database of the National Center for Biotechnology Information (NCBI). The retrieval formula was as follows: ("cervical intraepithelial neoplasia" [MeSH Terms] OR cervical intraepithelial neoplasia [All Fields]) AND "*Homo sapiens*" [porgn] AND ("Expression profiling by array" [Filter] AND ("0001/01/01" [PDAT]: "2020/11/27" [PDAT])). Three expression profile microarray datasets (GSE63514, GSE64217 and GSE138080) were selected and downloaded from the GEO database for analysis.

GSE63514 [4] is an expression profile based on the GPL570 platform (Affymetrix Human Genome U133 Plus 2.0 Array) and contains samples of normal cervical epithelium, CIN and cervical squamous epithelial cancer. GSE64217 is an expression profile based on the GPL10558 platform (Illumina HumanHT-12 V4.0 expression beadchip), and GSE64217 was provided by the Indian Institute of Technology Kharagpur, School of Medical Science and Technology, Multimodal Imaging and Computing for Theranostics. It contains samples of normal cervical mucosa, CIN and cervical squamous cell carcinoma (CESC). GSE138080 [5] is an expression profile based on the GPL4133 platform (Agilent-014850 Whole Human Genome Microarray 4x44K G4112F) and contains samples of normal cervical squamous epithelium, CIN and CESC. Samples of the three microarray expression profiling datasets were classified and analyzed according to the progression of cervical cancer. The workflow of this study is indicated in Fig. 1.

Analysis of microarray datasets

According to the progression of cervical cancer, in the subsequent analysis, samples of each dataset were divided into three groups: N-CIN, CIN-CC, and N-CC. The GEO2R tool was used to analyze the three

Fig. 1 Flowchart of the integrated analysis
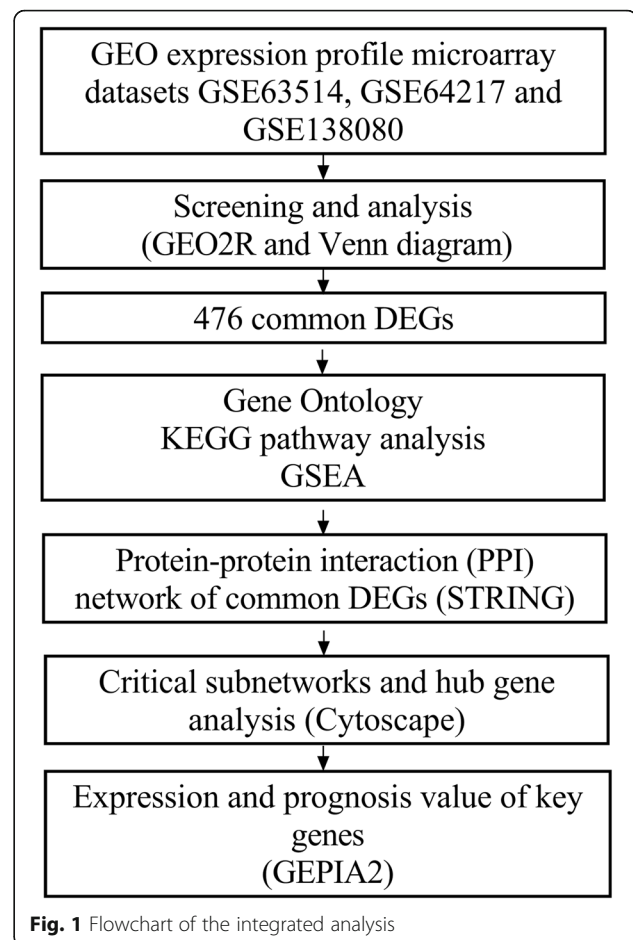
expression datasets [6, 7]. Normalization and log2 conversion were carried out for each dataset to filter out the DEGs of the three datasets, and the DEGs are displayed as volcano plots. The filtering conditions were as follows: |log2-fold change| ≥ 1 and adjusted *P*-value (adj. *P*) < 0.05. Then, the Venn diagram tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) was used to compare and analyze the results of the intersection analysis. Based on the intersection of datasets, the final DEGs were obtained. In this study, we selected DEGs according to the intersection of at least two expression profile datasets to avoid the disadvantages of a single dataset and then integrated the results for further biological function analysis.

Functions and pathways of DEGs

DAVID Bioinformatics Resources 6.8 was utilized to distinguish and enrich the biological attributes, such as biological processes, cellular components, molecular functions and pathways, of important DEGs [8] (https://david.ncifcrf.gov/). Moreover, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway [9] and Gene Ontology (GO) enrichment

analyses were used to identify the significant pathways. $P < 0.05$ was set as the cutoff criterion for significant enrichment.

Enrichment analysis of key pathways and core genes

Gene set enrichment analysis (GSEA) was used to determine the key pathways and core genes during the development of cervical cancer [10, 11]. Enrichment analyses were conducted to determine whether a series of a priori-defined biological processes were enriched. The enriched pathways were arranged in the order of their normalized enrichment scores, and those with $P < 0.01$ were chosen for further analysis. The results of the GSEA of different expression profile datasets were intersected to obtain the common significant KEGG pathways, and core gene sets were analyzed.

Construction of the PPI network

The STRING database is an online search tool used to analyze known proteins and predict PPI networks, including direct and indirect interactions between proteins and their functional correlations [12] (https://string-db.org/). Molecular interactions and PPI networks can promote the exploration of molecular targets, signaling and metabolic pathways, and network functions involved in the progression of cervical lesions. Therefore, the STRING database was used to construct the PPI network.

Critical subnetworks and hub genes

Hub genes play an important role in biological processes. Based on the PPI network, hub genes were screened according to network topology. Cytoscape software (version 3.8.2, cytoHubba and MCODE plug-ins) was used to discover the key targets or subnetworks of complex networks [13–15]. The critical subnetworks and hub genes during the development of cervical cancer were analyzed.

Expression and prognostic value of hub genes

As cervical cancer is a complex disease, its etiopathogenesis involves compound gene expression and multiple interactions. GEPIA2 (http://gepia2.cancer-pku.cn/) was used to analyze the expression of multiple hub genes and the prognostic value of the selected hub genes in CESC [16]. GEPIA2 can be used to perform principal component analysis (PCA) of genes and presents results as 2D plots. To evaluate the potential clinical value of hub genes more comprehensively, PCA dimensionality reduction was performed. Multiple gene comparison in different cancer species also provided a reference for the prognostic evaluation of different genes in cervical lesions. Moreover, the functional enrichment of prognostic genes was demonstrated.
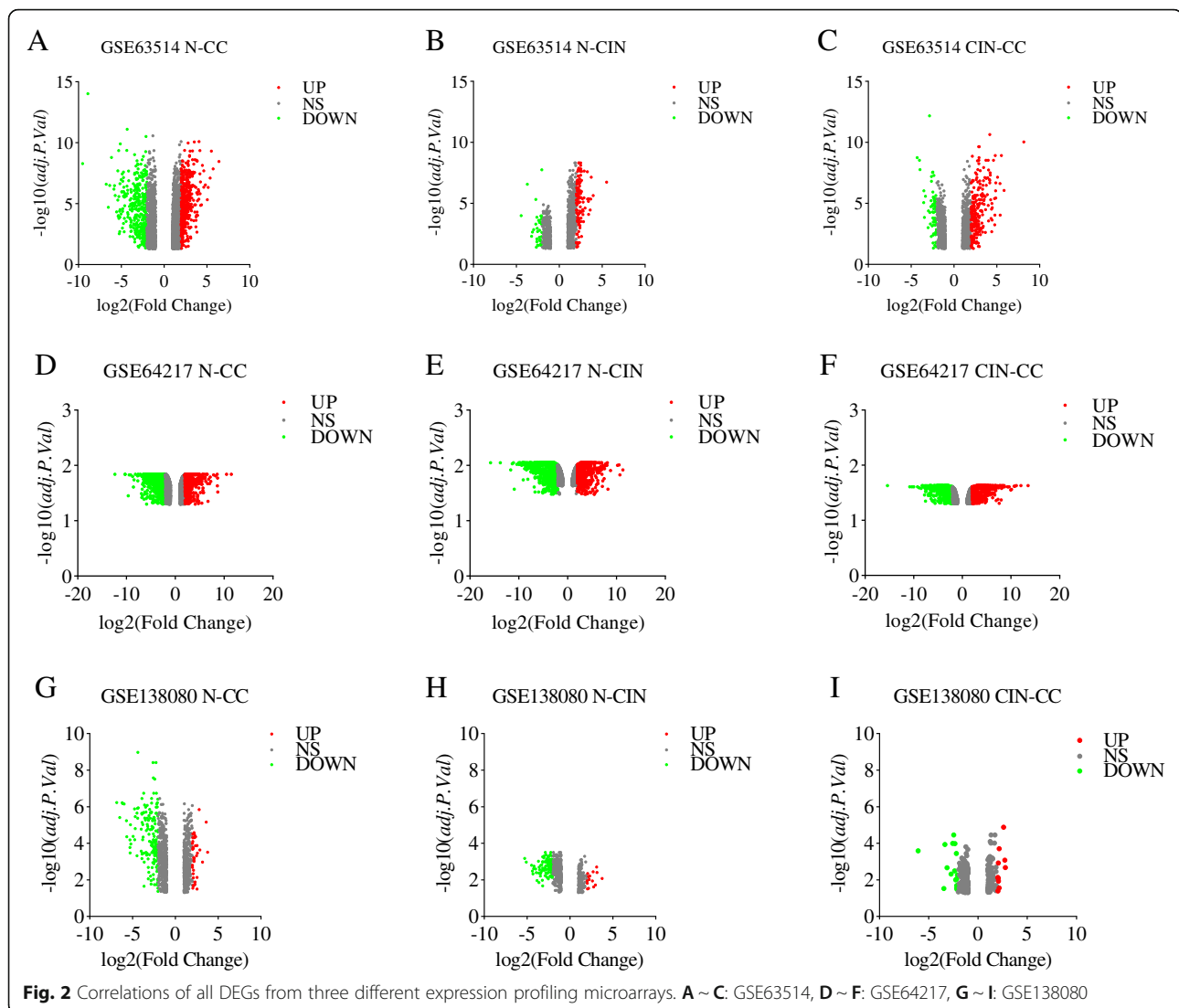
Results

Identification of DEGs

Volcano plots (Fig. 2) showed the correlation of all DEGs from the three expression profiling microarrays. After comprehensive analysis and screening with the Venn diagram tool (Fig. 3), 63 upregulated genes were identified in the N-CIN group, and *CDKN2A* [17, 18] was shared among the three datasets. Among the 56 downregulated genes, *EMP1, CRISP2, ALOX12, DMKN, ZBED2, PPP1R3C, CDA* and *CRCT1* were shared among the three datasets. There were 83 upregulated genes and 7 downregulated genes in the CIN-CC group. Furthermore, 155 upregulated genes in the N-CC group, including *TCAM1P, MCM2, HS6ST2, AIM2, CDKN2A, RFC4, PLOD2, APOC1* and *CENPF* [19], were shared among the three datasets. Among the 208 downregulated genes, 51 were shared among the three datasets: *TMPRSS11B, BBOX1, ZSCAN18, ENDOU, KLK8, ANKRD35, A2ML1, CRYAB, MAL, VSIG10L, ECM1, SPINK5, TM7SF2, SPIN K7, PRSS27, RBP7, PRSS3, ACPP, HPGD, CWH43, RHCG, SCEL, TP53I3, SPRR2C, CRABP2, HCG22, DMKN, PRSS2, CLIC3, SPNS2, LCE3D, FUT3, RDH12, CRNN, CEACAM7, LYNX1, MYZAP, KRTDAP, NDRG4, SLC5A1, GPX3, PPP1R3C, SLURP1, SLC24A3, THSD4, PSCA, CDA, FAM3D, CFD, HOPX,* and *CRCT1*. Finally, after deletion of duplicate genes, a total of 476 DEGs were screened (|log2-fold change| ≥ 1 and adj. $P < 0.05$): 253 genes with upregulated expression and 223 genes with downregulated expression.

GO and KEGG pathway enrichment analyses

A total of 476 DEGs were uploaded to DAVID for GO/KEGG analyses. The terms of each GO category are provided in Additional file 1: Table S1, Table S2 and Table S3. Most DEGs were enriched in the biological processes keratinization, epidermis development, DNA replication, mitotic division, cell cycle, proteolysis, regulation of cell proliferation, cell cycle, and related activity of enzymes; the cellular components extracellular space, cornified envelope, extracellular exosome, and extracellular region; and the molecular functions serine-type endopeptidase activity, serine-type peptidase activity, structural molecule activity, cysteine-type and endopeptidase inhibitor activity. The results of KEGG pathway enrichment are shown in Table 1.

GSEA enrichment of expression datasets

Expression datasets (GSE63514, GSE64217 and GSE138080) were subjected to GSEA (version 4.1.0). Then, key pathways and core related genes were obtained. The results showed 5 common KEGG pathways in the N-CIN group ($p < 0.01$), namely, DNA mismatch repair (*PCNA* [20], *EXO1* [21], *POLD1, MSH6,* and *LIG1*), the cell cycle (*MCM3, MCM5* [22], *CDC6,*

Fig. 2 Correlations of all DEGs from three different expression profiling microarrays. **A ~ C**: GSE63514, **D ~ F**: GSE64217, **G ~ I**: GSE138080
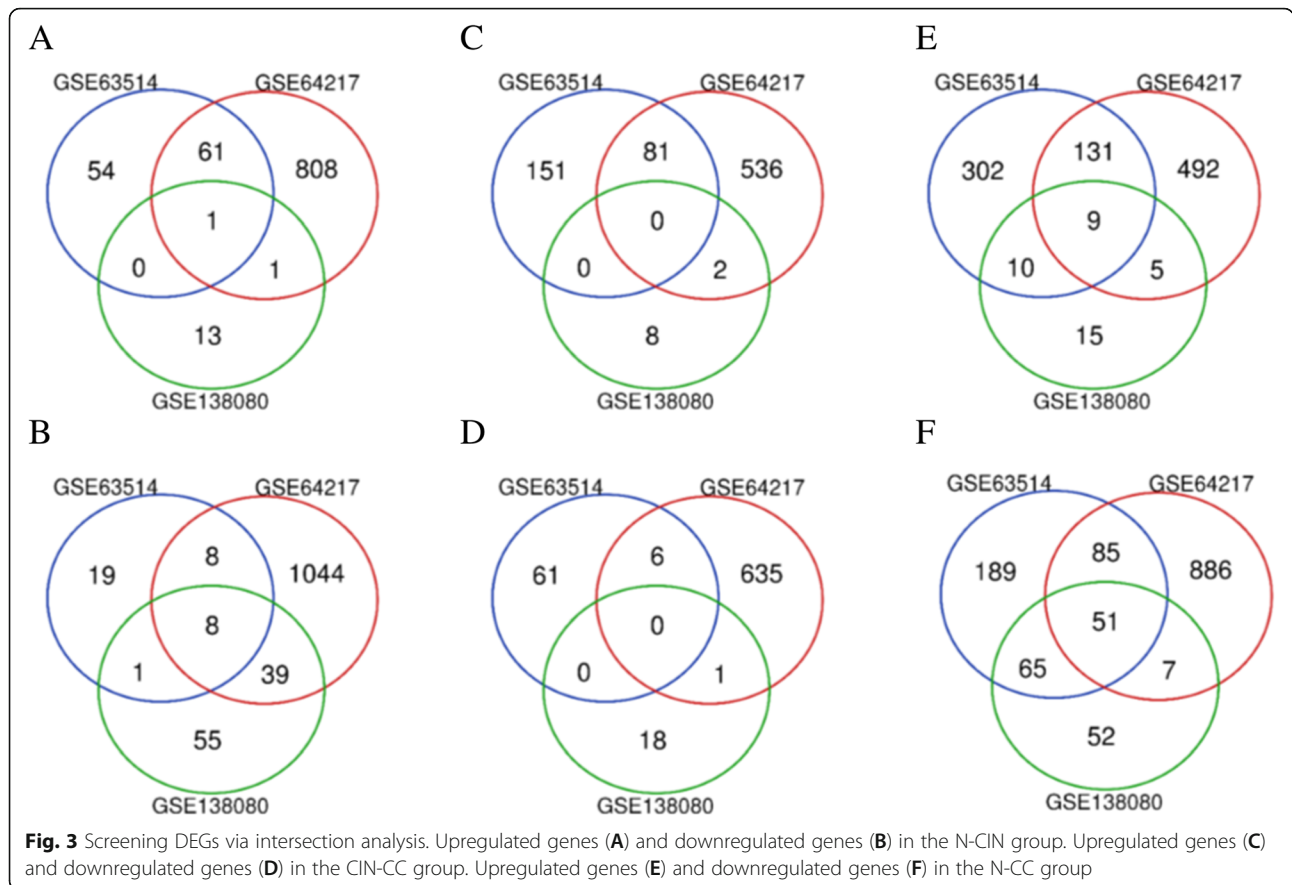
MCM6 [23], *CHEK2*, *PKMYT1*, *CDC7*, *RBL1*, *WEE1*, *CCNA2* [24], and *TTK*), DNA replication (*PCNA*, *POLD1*, *DNA2*, *MCM3*, *MCM5*, *PRIM2*, *POLE2*, *MCM6*, *LIG1*, *RNAEH2A*, and *PRIM1*), cysteine and methionine metabolism (*LDHC*), and nucleoside exception repair (*PCNA*, *POLD1*, *POLE2*, and *LIG1*). In addition, there were 6 common KEGG pathways in the CIN-CC group ($p < 0.01$), namely, the adipocytokine signaling pathway, small cell lung cancer (*ITGA6* [25], *PIAS3*, and *LAMC2* [26]), pathways in cancer, the Toll-like receptor signaling pathway, graft versus host disease, and the TGF-beta signaling pathway [27]. There were 9 common KEGG pathways in the N-CC group ($p < 0.01$), namely, the cell cycle (*PCNA*, *CDC25B*, *MCM3*, *MCM5*, *CDC6* [28], *GSK3B*, *MCM6*, *CHEK2*, *PKMYT1*, *CDC20* [29], *PTTG1*, *SMAD3*, *CCNB1* [30], *RBL1*, *CDC7* [23], *WEE1*, *CDK2* [31], *CCNA2*, and *TTK*), nucleotide excision repair, the Toll-like receptor signaling pathway, prion diseases,

spliceosome, DNA replication, proteasome, colorectal cancer, and pancreatic cancer. Moreover, GSEA showed that DNA mismatch repair (N-CIN), small cell lung cancer (CIN-CC), and the cell cycle (N-CC) were the most significantly enriched pathways ($P < 0.01$, FDR < 0.05), and snapshots of the enrichment analysis are shown in Figs. 4, 5 and 6.

PPI network construction and hub genes of cervical lesions

A total of 476 DEGs were uploaded to the STRING database to construct the PPI network (minimum required interaction score: highest confidence 0.900, k-means clustering: number of clusters 3). The PPI network is shown in Fig. 7. There were 415 nodes and 931 edges in the network (PPI enrichment $P < 1.0E-16$). The functional enrichment analysis in the PPI network included 387 GO terms, 5 KEGG pathways, 64 Reactome

Fig. 3 Screening DEGs via intersection analysis. Upregulated genes (**A**) and downregulated genes (**B**) in the N-CIN group. Upregulated genes (**C**) and downregulated genes (**D**) in the CIN-CC group. Upregulated genes (**E**) and downregulated genes (**F**) in the N-CC group

pathways, and 111 protein domains. Significantly enriched functions of the network are shown in Fig. 8. According to these results, most of the proteins were distributed among the following aspects: polymorphism, glycoprotein, signal, disease, disabled bond, and secreted. These genes express proteins and then interact functionally in the PPI network, revealing their role in the progression of cervical cancer.

On the basis of the PPI network, the critical subnetworks were extracted with Cytoscape software. The

Table 1 KEGG pathway enrichment analysis of DEGs (DAVID)

Term	Count	*P*-value
hsa04110: Cell cycle	14	7.05E-07
hsa03030: DNA replication	7	5.88E-05
hsa05323: Rheumatoid arthritis	8	0.001555896
hsa04115: p53 signaling pathway	7	0.001840686
hsa05219: Bladder cancer	5	0.007981073
hsa04610: Complement and coagulation cascades	6	0.01080063
hsa05146: Amoebiasis	7	0.017024795
hsa04060: Cytokine-cytokine receptor interaction	11	0.019861501
hsa04114: Oocyte meiosis	7	0.020898334
hsa00590: Arachidonic acid metabolism	5	0.030571729

critical subnetworks and hub genes of the cervical lesions are shown in Fig. 9. The results showed that the hub genes constituted the key network of cervical carcinogenesis, and the genes were scored with the cyto-Hubba plug-in. Among the results of different algorithms, the common hub genes with high scoring were as follows: *NUSAP1, TOP2A, KIF2C, NDC80, ASPM, KIF20A, CDK1, KIF11, BIRC5, MCM2,* and *CHEK1*. Then, the MCODE plug-in was also used to analyze the PPI network, and 12 regions (subnetworks) closely related to the PPI network were found and separated. These regions might represent the molecular complex, and the results from the MCODE analysis included these high scoring hub genes and supported the results of the regional analysis (Table 2). These genes mainly affected cell division, the cell cycle, keratinocyte differentiation [32], DNA replication, mismatch repair, nucleoside exception repair, cytokine–cytokine receptor interaction, the chemokine signaling pathway and arachidonic acid metabolism during the progression of cervical cancer.

Expression and prognostic value of hub genes

GEPIA2 analysis showed that these hub genes were highly expressed in CESC tissues but weakly expressed in normal tissues (Fig. 10). Furthermore, to obtain more
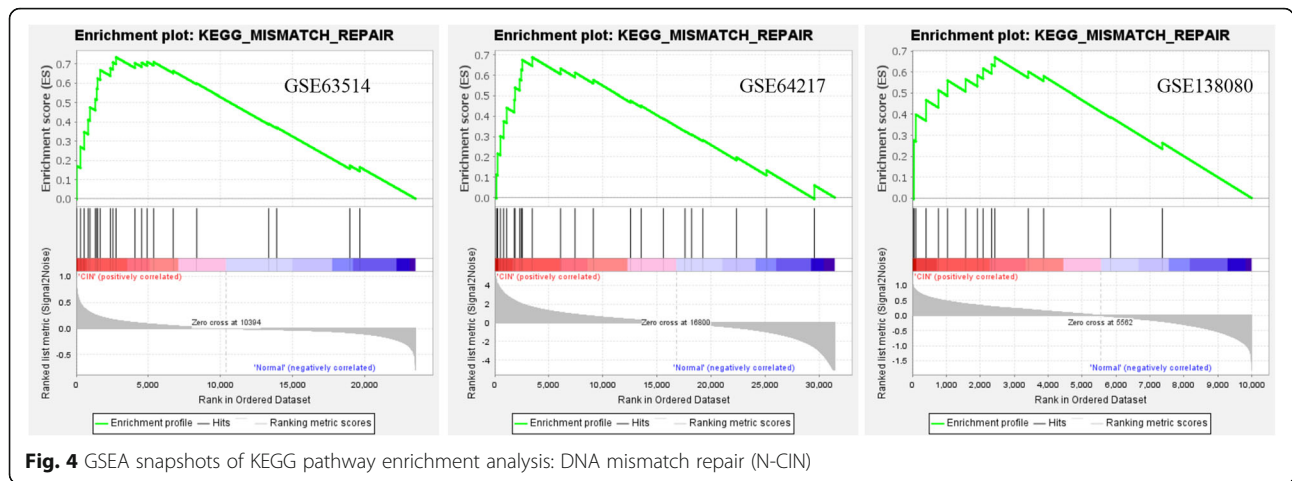
Fig. 4 GSEA snapshots of KEGG pathway enrichment analysis: DNA mismatch repair (N-CIN)

comprehensive prognostic information, hub genes from 12 subnetworks (Table 2) were subjected to survival analysis. Overall survival analysis by GEPIA2 indicated that the high expression of *MCM2, TOP2A, BLM, RMI2* [33], *EXO1, RFC4, PSCA, KNTC1, CDC45* and *GINS2* [34] in cervical lesions was correlated with an improved prognosis (Fig. 11). However, the high expression of *CXCL8* [35], *TNFAIP6, CXCL5* and *CDA* in cervical cancer tissues was associated with a poor patient prognosis (Fig. 12).

The tissue-specific expression of the hub genes in different cancer types is shown in Fig. 13 as an interactive heat map. A heat map was used to analyze the expression of the target genes in different tumor samples. Compared with other genes, *MCM2, TOP2A, CDC45, KNTC1, RFC4* and *RMI2* were highly expressed in CESC tumor tissues and might be better indicators of prognosis. PPI modeling (STRING) and enrichment analysis were performed on 14 target genes (Fig. 14). Gene functions according to the summaries of the Gene database (NCBI) are shown in Table 3. These findings might provide target genes for the prognosis and treatment of cervical cancer.

Discussion

Cervical cancer is one of the most common cancers among women worldwide. After persistent infection with high-risk HPV, the progression of inflammation to CIN to cancer often takes a long time. CIN is an important precancerous lesion of cervical cancer. At present, the prevention of cervical cancer depends mainly on HPV vaccines and HPV screening. Therefore, we need to not only prevent and monitor the development from normal cervical tissue or cells to CIN but also treat and block the development from CIN to cancer. Based on the analysis of multiple datasets, this study deepened our understanding of the molecular mechanism of cervical carcinogenesis and identified key prognostic genes. Potential biomarkers and target genes can be used to diagnose progressive disease before it leads to cancer.

In the present study, three expression profile datasets (GSE63514, GSE63217 and GSE138080) were downloaded from the GEO database. According to the development of cervical cancer, the samples were divided into three groups: N-CIN, CIN-CC and N-CC. Intersection analysis of these groups made the screening
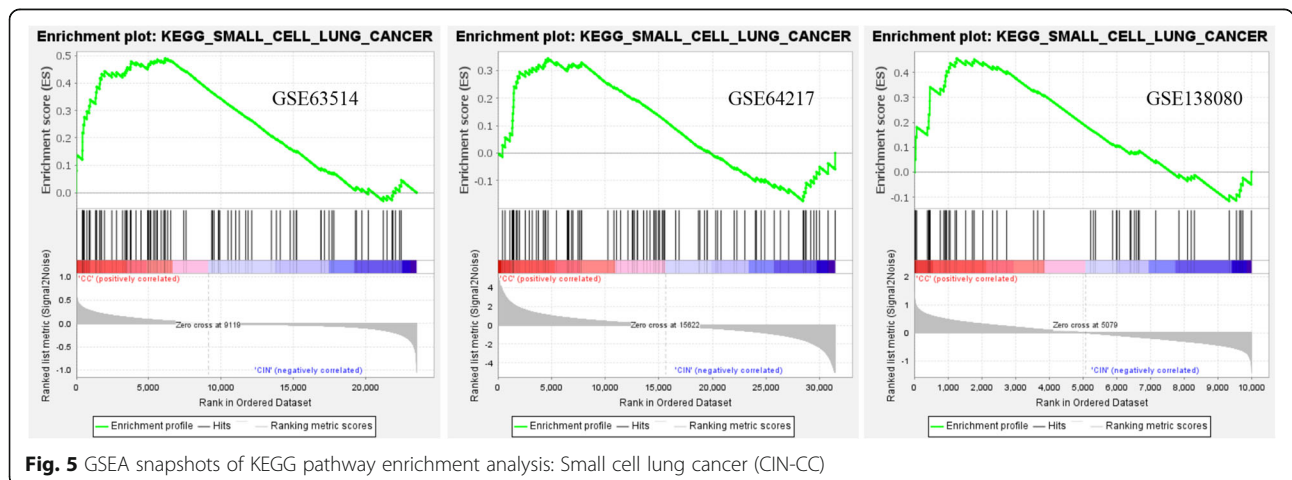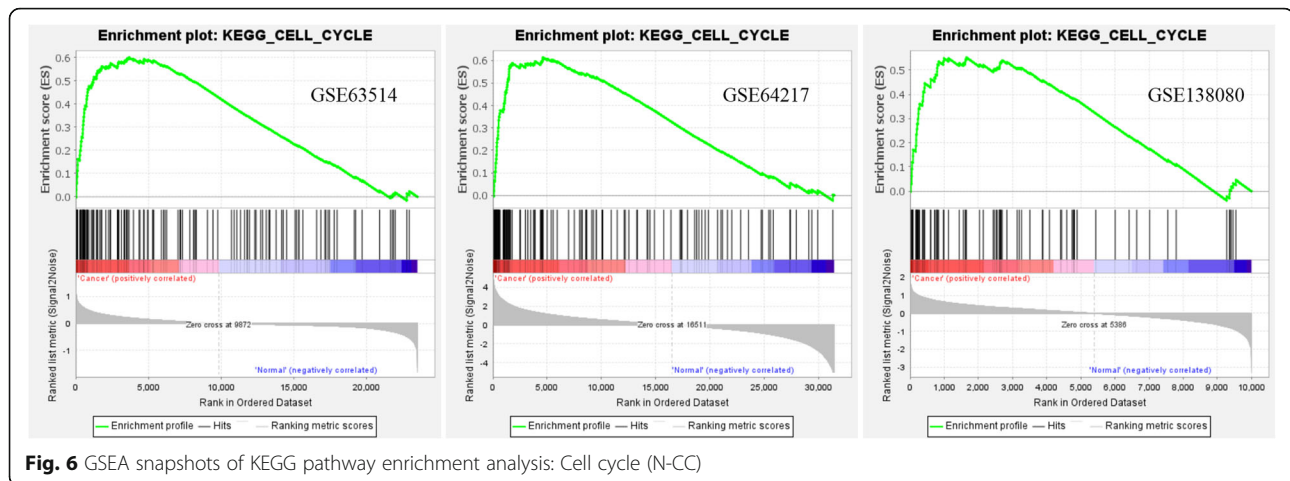

Fig. 5 GSEA snapshots of KEGG pathway enrichment analysis: Small cell lung cancer (CIN-CC)

Fig. 6 GSEA snapshots of KEGG pathway enrichment analysis: Cell cycle (N-CC)

and validation of DEGs more reliable. A total of 476 DEGs were screened: 253 upregulated genes and 223 downregulated genes. These genes also regulate a number of biological pathways in the progression of cervical lesions, such as the cell cycle, DNA replication, rheumatoid arthritis, the p53 signaling pathway, bladder cancer, complement and coagulation cascades, amoebiasis, cytokine-cytokine receptor interaction, oocyte meiosis, and arachidonic acid metabolism. These pathways need to be further studied.

Furthermore, GSEA suggested that the DNA mismatch repair pathway plays an important role in the N-CIN process. Most importantly, the small cell lung cancer pathway might play an important role in the CIN-CC process. The cell cycle pathway might play an important role in the N-CC process. These results provide a better understanding of the molecular pathways associated with the development of the disease from normal cervical epithelium to CIN to cervical cancer. In the N-CIN stage, the enriched molecular pathways were DNA replication, nucleoside exception repair, DNA mismatch repair and specific amino acid metabolism (cysteine and methionine). With the development of cervical epithelial neoplasia, the enriched molecular pathways in the CIN-CC stage were small cell lung cancer, the adipocytokine signaling pathway, pathways in cancer, the Toll-like receptor signaling pathway, graft versus host disease and the TGF-beta signaling pathway, which play important roles in the occurrence and development of cancer. In the whole process of cervical lesions, we may need to pay close attention to pathways such as the cell cycle, DNA replication, DNA repair and pathways in cancer. Our results indicate the importance of these pathways in the occurrence and development of cervical cancer. We need to pay attention to the roles of these pathways in cervical carcinogenesis and further study their interactions.

Notably, the PPI network related to cervical lesions was composed of functional proteins that interacted with each other to participate in biological signal transmission, gene expression regulation, energy and material metabolism, cell cycle regulation and cell division. Eleven genes were identified as hub genes from 12 critical subnetworks of cervical cancer. Critical subnetworks might strongly contribute to the occurrence and development of cervical cancer and have high diagnostic value. Further analysis showed that *NUSAP1* [36], *TOP2A, KIF2C* [37], *NDC80* [23], *ASPM* [21, 38], *KIF20A* [39], *CDK1* [19, 38], *KIF11* [40, 41], *BIRC5* [42], *MCM2* and *CHEK1* [40, 41, 43] were high scoring hub genes and showed significantly upregulated expression in cervical cancer tissues compared with normal tissues ($P < 0.01$). By analyzing the prognostic values of the hub genes from the 12 subnetworks, a total of 14 potential molecular targets (*MCM2* [23, 35, 44–46], *TOP2A* [19, 35, 40], *BLM, RMI2, EXO1, RFC4, PSCA, KNTC1, CDC45, GINS2, CXCL8, TNFAIP6, CXCL5,* and *CDA*) were obtained and annotated.

These 23 key regulatory genes were enriched in different pathways, such as the cell cycle and mismatch repair, and play important roles in the occurrence and development of diseases. Our DAVID analysis showed that *CDK1, CHEK1, MCM2* and *CDC45* were involved in the cell cycle pathway. We found that the cell cycle pathway was very important to the progression of cervical cancer, which is worthy of further research. Mine KL et al. [47] proved that the cell cycle may be the main driving factor of cervical cancer. Van Dam et al. [48] showed that deregulation of the cell cycle is a major component of cervical cancer biology. Y Luo et al. [49] showed that *CDK1* might play an important role in regulating the genetic network related to the occurrence, development and metastasis of cervical cancer. Relevant studies also showed that the progression of cervical cancer can be affected by regulating the cell cycle, which highlights the biological significance of the cell cycle in cervical cancer [50–53].

Fig. 7 PPI network of the DEGs
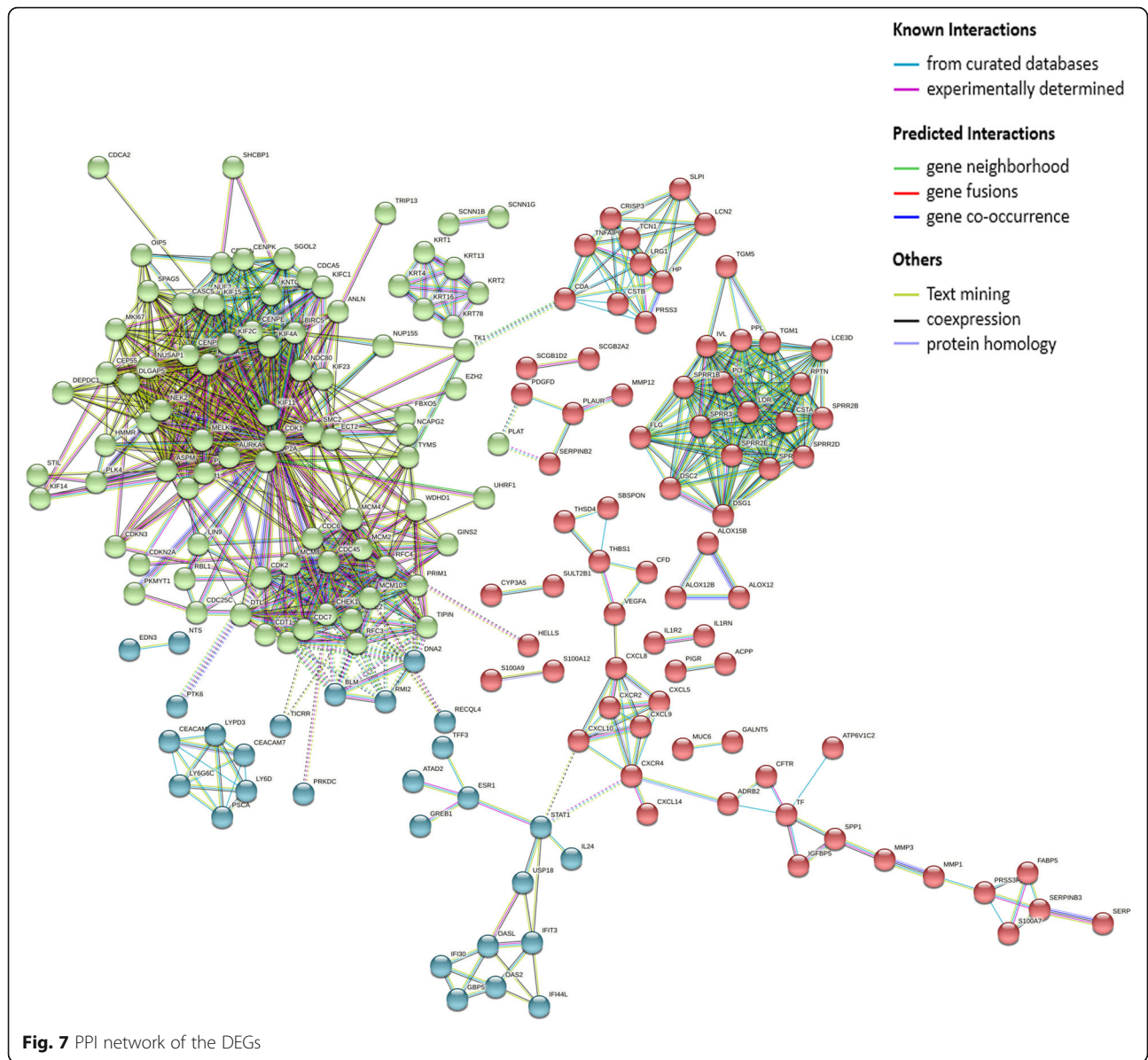
In addition, Liang Zhao et al. [43] revealed that the imbalance of CHEK1 and CDKN2A further promoted the proliferation of cancer cells by affecting the response of cell cycle checkpoints to DNA damage. Therefore, further research on the cell cycle and its related genes is of great significance. Moreover, we found that CDK1 and CHEK1 were involved in the p53 signaling pathway. The p53 signaling pathway is involved in the proliferation and apoptosis of cervical cancer cells and has high prognostic value.

The DNA replication pathway was also very important in the progression of cervical cancer in this study. Mitali Das et al. [54] showed that MCM2–7 was significantly enriched in DNA replication, and high MCM2–7 expression promoted the malignant proliferation of cervical cancer cells. MCM2 was studied in a variety of human malignant tumors and was related to the histopathological grade of many [55–58]. Zheng J et al. [59] showed the diagnostic value of MCM2 immunocytochemical staining in cervical lesions and its relationship with HPV infection. The expression of RFC4 may also be associated with tumor progression and poor patient survival and is considered to be one of the main driving factors of the cervical cancer cell cycle network [47]. Dan Liu et al. [60] showed that RFC4 was related to DNA replication and cell proliferation in cervical cancer. The abnormal expression of RFC4 might be related to the progression of cervical cancer [61]. Similarly, our results indicated that MCM2 and RFC4 play an

Fig. 8 Functional enrichment in the PPI network of DEGs
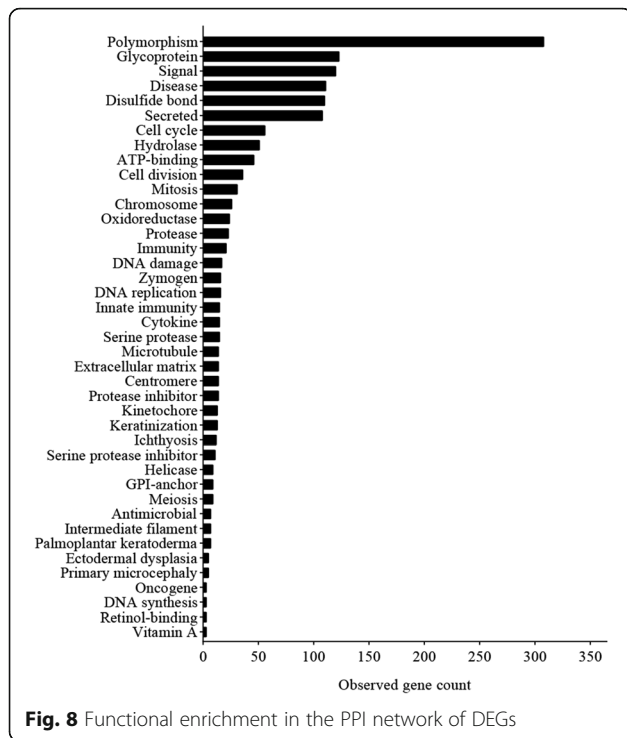
important role in the DNA replication pathway. In particular, *CXCL8* and *CXCL5* were important regulatory genes in our study. They are involved in rheumatoid arthritis and the cytokine–cytokine receptor interaction pathway in cervical carcinogenesis. CXCL8 was also associated with amoebiasis and the bladder cancer pathway. Ruiling Yan et al. [62] studied the clinical and prognostic value of *CXCL8* in patients with cervical cancer. The abnormal expression of *CXCL5* contributes to the tumorigenicity of cervical cancer [63]. Our results also showed that *RFC4* and *EXO1* are involved in DNA mismatch repair and are very important to cervical lesions.

However, due to the complexity of cervical cancer, its molecular mechanism needs to be further studied, and key regulatory genes and pathways will be continuously understood. We can track the influence of some key regulatory genes in cervical cancer through related research. For example, Beiwei Yu et al. [64] showed that *TOP2A* and *CENPF* are synergic master regulators that are activated in cancer. Jinhui Liu et al. [33] showed that *RMI2* was a novel key gene in CESC. Huan Chen et al. [65] showed that the *KNTC1* gene may be related to the pathophysiology of cervical cancer and may be one of the markers
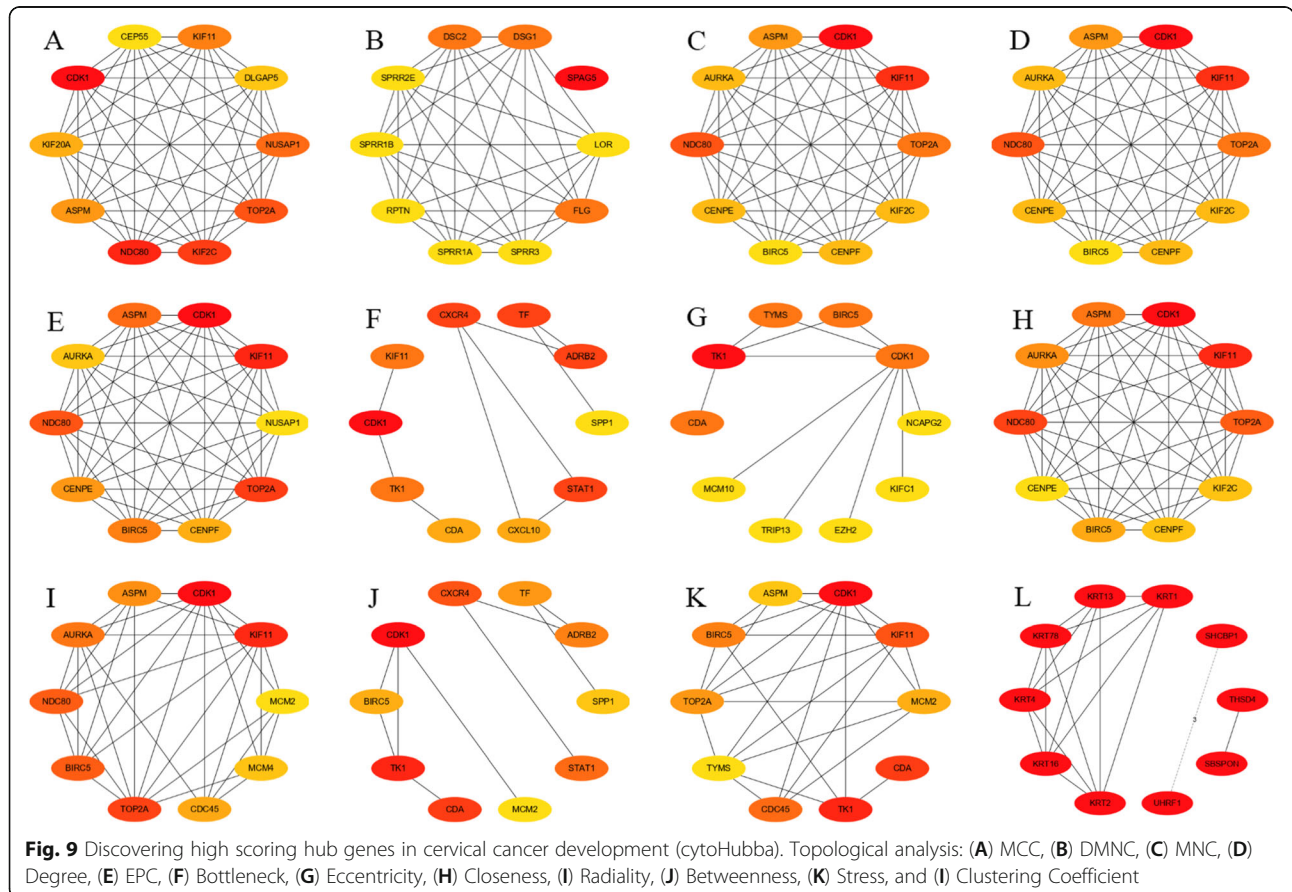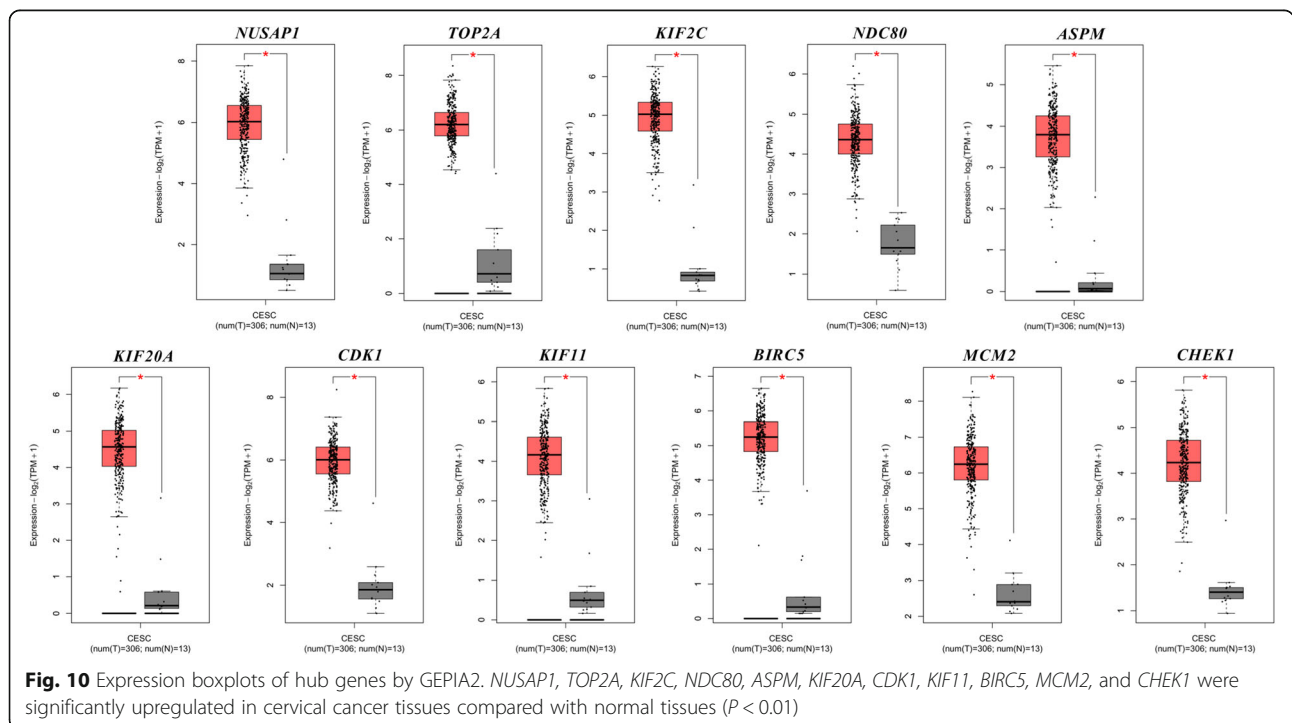

Fig. 9 Discovering high scoring hub genes in cervical cancer development (cytoHubba). Topological analysis: (**A**) MCC, (**B**) DMNC, (**C**) MNC, (**D**) Degree, (**E**) EPC, (**F**) Bottleneck, (**G**) Eccentricity, (**H**) Closeness, (**I**) Radiality, (**J**) Betweenness, (**K**) Stress, and (**l**) Clustering Coefficient

Table 2 Discovering regions closely related to the PPI network (MCODE)

Cluster	Score (Density*#Nodes)	Nodes	Edges	Genes
1	21.478	24	247	TOP2A, KIF15, PRC1, SPAG5, KIF11, BIRC5, KIF23, MELK, DLGAP5, KIF2C, NUF2, MKI67, CENPE, NDC80, NEK2, AURKA, ASPM, CENPF, KIF20A, CDK1, TTK, CEP55, NUSAP1, KIF4A
2	16.625	17	133	RPTN, SPRR1B, SPRR2E, SPRR2D, LCE3D, DSC2, SPRR2B, DSG1, PI3, TGM1, CSTA, IVL, LOR, FLG, SPRR3, SPRR1A, PPL
3	9.111	10	41	CHEK1, MCM4, CDC45, MCM10, PRIM1, CDC7, CDT1, CDK2, POLE2, MCM8
4	8	8	28	CRISP3, PRSS3, HP, LRG1, CSTB, TCN1, TNFAIP6, CDA
5	6	6	15	SGOL2, CENPN, CASC5, CENPK, CDCA5, KNTC1
6	6	6	15	KRT78, KRT2, KRT13, KRT1, KRT16, KRT4
7	6	6	15	CXCR2, CXCR4, CXCL9, CXCL5, CXCL8, CXCL10
8	6	6	15	LY6G6C, CEACAM7, PSCA, CEACAM5, LYPD3, LY6D
9	4	4	6	RFC4, EXO1, TIPIN, RFC3
10	4	4	6	PRSS3P2, SERPINB3, FABP5, S100A7
11	3.429	8	12	DNA2, WDHD1, GINS2, DTL, RMI2, CDC6, BLM, MCM2
12	3	3	3	ALOX12B, ALOX12, ALOX15B

for the early diagnosis of cervical precancerous lesions. The abnormal expression of GINS2 can inhibit cell proliferation and tumorigenicity, as well as cell migration and invasion [66]. KIF20A expression is related to the overall survival rate of patients with early CESC and its progression [67]. At present, research on these key regulatory genes and related molecular pathways is lacking. The interactions of these genes and pathways in the progression of

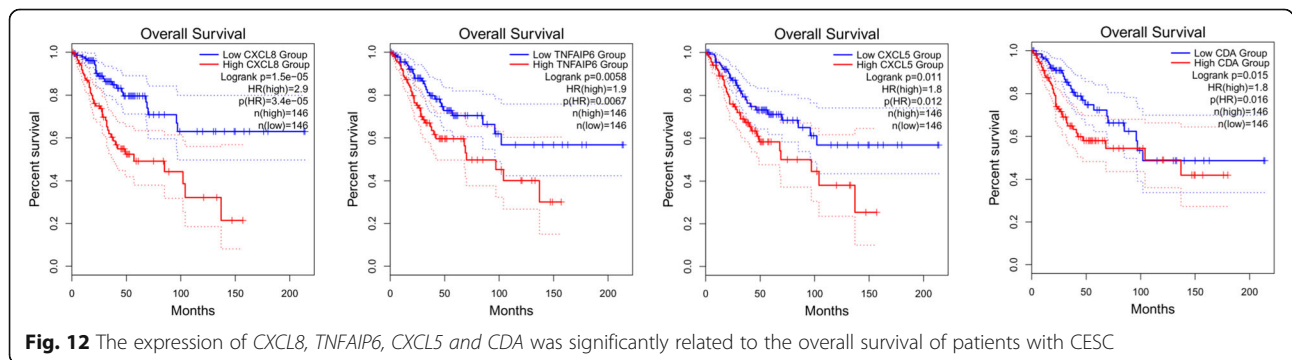cervical cancer still need attention, research and verification.

Gene bioinformatics can provide a possible molecular targeting mechanism for the prevention and treatment of cervical diseases. Functional studies of candidate genes from public databases may lead to a better understanding of the development of cervical cancer. We divided the expression datasets into three groups according to the different stages of cervical

Fig. 10 Expression boxplots of hub genes by GEPIA2. *NUSAP1, TOP2A, KIF2C, NDC80, ASPM, KIF20A, CDK1, KIF11, BIRC5, MCM2,* and *CHEK1* were significantly upregulated in cervical cancer tissues compared with normal tissues (P < 0.01)

Fig. 11 The expression of *MCM2, TOP2A, BLM, RMI2, EXO1, RFC4, PSCA, KNTC1, CDC45* and *GINS2* was significantly related to the overall survival of patients with CESC

lesions to comprehensively investigate the progression of cervical cancer. Intersection analysis of multiple expression profiles avoids the limitation of a single dataset and has repeatability and reliability. However, the limitation of this study was that the datasets were obtained from three different chip platforms, and as there were differences in data quality, the results were easily affected, resulting in bias. Our results may provide effective targets for the treatment of cervical cancer, but the effect on prognosis requires follow-up data, and further studies are needed to explore these key genes and important pathways.
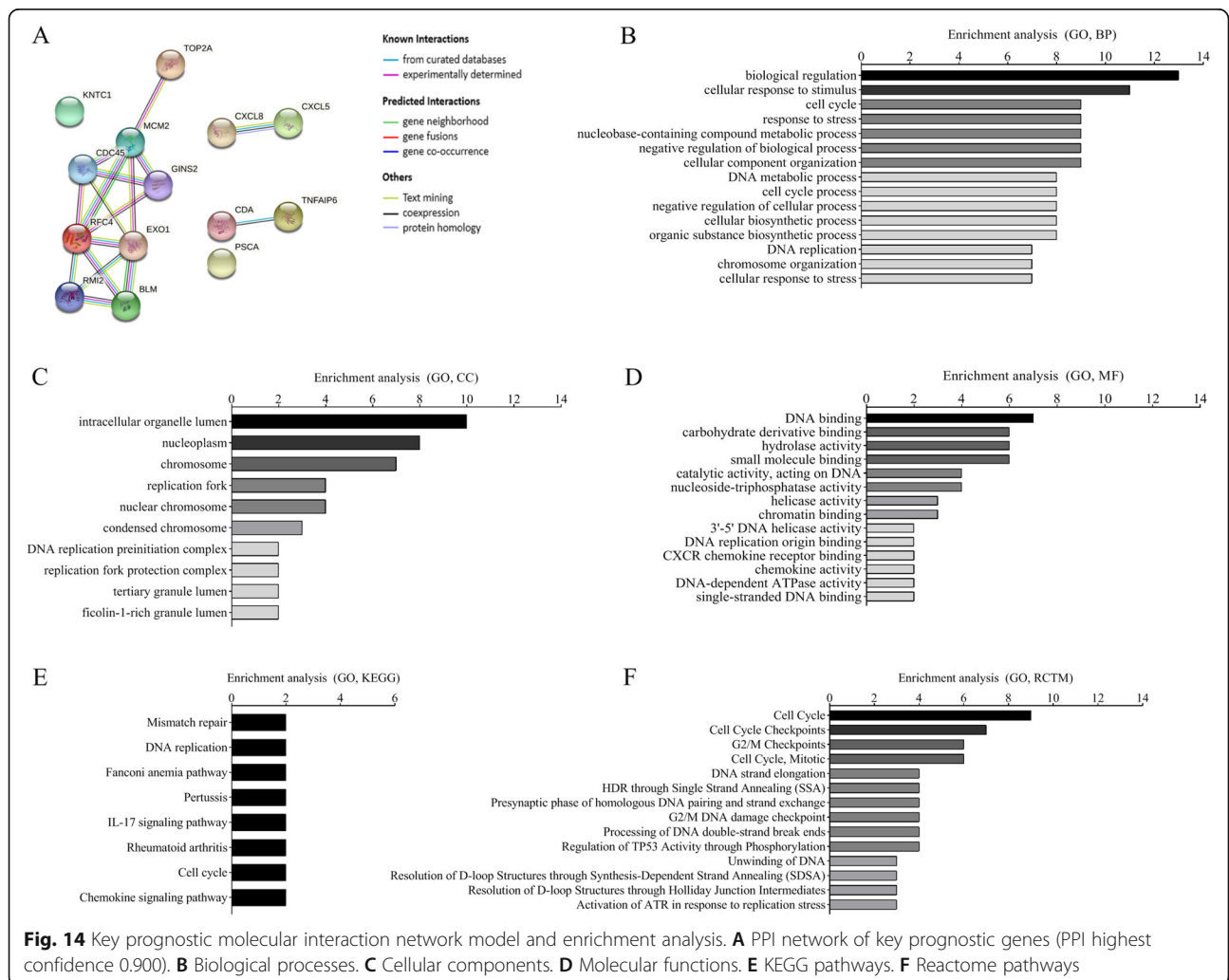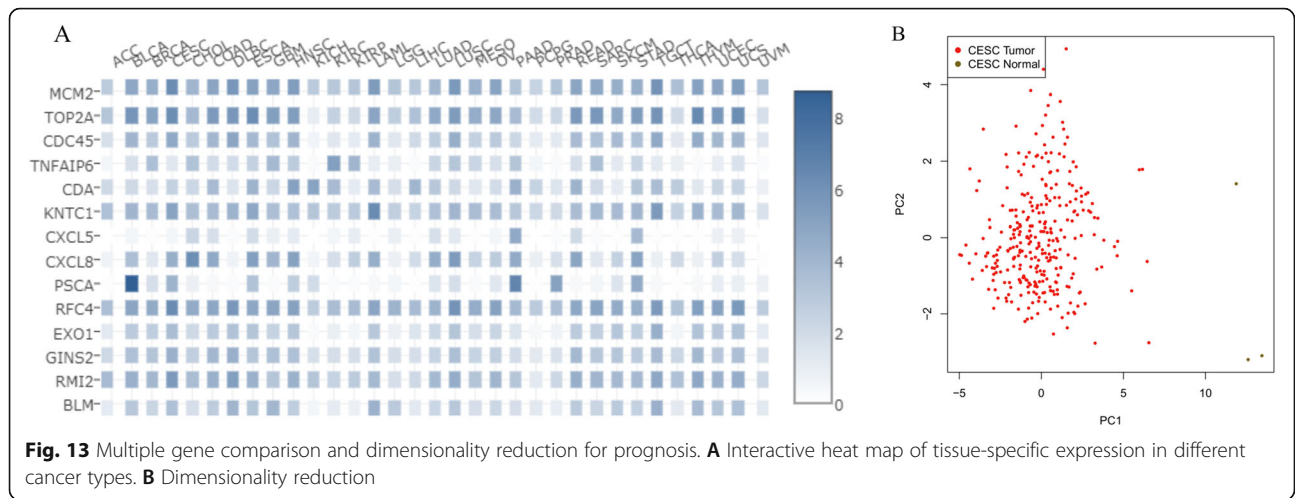

Fig. 12 The expression of *CXCL8, TNFAIP6, CXCL5 and CDA* was significantly related to the overall survival of patients with CESC

Fig. 13 Multiple gene comparison and dimensionality reduction for prognosis. **A** Interactive heat map of tissue-specific expression in different cancer types. **B** Dimensionality reduction

Fig. 14 Key prognostic molecular interaction network model and enrichment analysis. **A** PPI network of key prognostic genes (PPI highest confidence 0.900). **B** Biological processes. **C** Cellular components. **D** Molecular functions. **E** KEGG pathways. **F** Reactome pathways

Table 3 Functional analysis of possible prognostic genes for the progression of cervical cancer

Gene	Function
MCM2	Minichromosome maintenance complex component 2. MCM2 is involved in the initiation of eukaryotic genome replication. It may be involved in the formation of replication forks and in the recruitment of other DNA replication-related proteins, and it regulates the helicase activity of the complex.
TOP2A	DNA topoisomerase II a. These findings might provide target genes for the prognosis and treatment of cervical carcinoma. TOP2A controls and alters the topologic state of DNA during transcription. This nuclear enzyme is involved in processes such as chromosome condensation, chromatid separation, and the relief of torsional stress that occurs during DNA transcription and replication.
BLM	BLM RecQ-like helicase. This Bloom-associated helicase unwinds a variety of DNA substrates including Holliday junctions and is involved in several pathways contributing to the maintenance of genome stability.
RMI2	RecQ-mediated genome instability 2. RMI2 plays a role in homologous recombination-dependent DNA repair and is essential for genome stability.
EXO1	Exonuclease 1. EXO1, with 5′ to 3′ exonuclease activity as well as RNase H activity, interacts with Msh2, which is involved in mismatch repair and recombination.
RFC4	Replication factor C subunit 4. The elongation of primed DNA templates by DNA polymerase delta and DNA polymerase epsilon requires the accessory proteins proliferating cell nuclear antigen (PCNA) and replication factor C (RFC).
PSCA	Prostate stem cell antigen. PSCA is highly expressed in the prostate and also expressed in the bladder, placenta, colon, kidney, and stomach. This gene is upregulated in a large proportion of prostate cancers and is also detected in cancers of the bladder and pancreas.
KNTC1	Kinetochore-associated 1. KNTC1 ensures proper chromosome segregation during cell division.
CDC45	Cell division cycle 45. CDC45, an essential protein required for the initiation of DNA replication, is important for the early steps of DNA replication in eukaryotes.
GINS2	GINS complex subunit 2. The GINS complex is essential for the initiation of DNA replication.
CXCL8	C-X-C motif chemokine ligand 8. CXC is a major mediator of the inflammatory response. This protein is also secreted by tumor cells and promotes tumor migration, invasion, angiogenesis and metastasis. This chemokine is also a potent angiogenic factor.
TNFA IP6	TNF alpha-induced protein 6. TNFAIP6 is a member of the hyaluronan-binding protein family. Its hyaluronan-binding domain is involved in extracellular matrix stability and cell migration. This protein is important in the protease network associated with inflammation.
CXCL5	C-X-C motif chemokine ligand 5. Chemokines, which recruit and activate leukocytes, promote angiogenesis and remodel connective tissues. It plays a role in cancer cell proliferation, migration, and invasion.
CDA	Cytidine deaminase. CDA is one of several deaminases responsible for maintaining the cellular pyrimidine pool.

Conclusions

In conclusion, a comprehensive bioinformatics analysis of DEGs and pathways involved in the occurrence and development of cervical lesions was performed, and we explored and obtained key regulatory genes and pathways contributing to the progression of cervical cancer to improve prognosis. Moreover, these results may promote the understanding of molecular mechanisms and clinically related molecular targets for prognosis in cervical cancer and provide new insight into the occurrence and development of cervical cancer.

Abbreviations
GEO: Gene Expression Omnibus; DEGs: Differentially expressed genes; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; GSEA: Gene set enrichment analysis; PPI: Protein–protein interaction; HPV: Human papillomavirus; CIN: Cervical intraepithelial neoplasia; N-CIN: From normal to cervical intraepithelial neoplasia; CIN-CC: From cervical intraepithelial neoplasia to cervical cancer; NCBI: National Center for Biotechnology Information; N-CC: From normal to cervical cancer; CESC: Cervical squamous cell carcinoma or endocervical adenocarcinoma; PCA: Principal component analysis; PCNA: Proliferating cell nuclear antigen; RFC: Replication factor C

Supplementary Information
The online version contains supplementary material available at https://doi. org/10.1186/s12885-021-08412-4.

Additional file 1: GO enrichment analysis of DEGs in Biological Processes, Cellular Components and Molecular Functions (DAVID). **Table S1.** GO enrichment analysis of DEGs in Biological Processes (DAVI D). **Table S2.** GO enrichment analysis of DEGs in Cellular Components (DAVID). **Table S3.** GO enrichment analysis of DEGs in Molecular Functions (DAVID).

Authors' information
Corresponding author and first author: Baojie Wu (June 1987-), male, senior engineer of the Pilot Department, Master's degree in biochemistry and molecular biology, mainly engaged in optimization of the purification process, amplification of the purification process, and bioinformatics data mining. Email: w3522021987@163.com, zrwbj@walvax.com. Second author: Shuyi Xi (February 1978-), male, department head, mainly engaged in

management of the Pilot Department. Email: xishuyi@walvax.com. The authors are committed to the study of HPV vaccines.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The author declares that he/she has no competing interests.

References
1. 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. https://doi.org/10.3322/caac.21492.
2. Wang R, Pan W, Jin L, Huang W, Li Y, Wu D, et al. Human papillomavirus vaccine against cervical cancer: opportunity and challenge. Cancer Lett. 2020;471:88–102. https://doi.org/10.1016/j.canlet.2019.11.039.
3. Pinto AP, Crum CP. Natural history of cervical neoplasia: defining progression and its consequence. Clin Obstet Gynecol. 2000;43(2):352–62. https://doi.org/10.1097/00003081-200006000-00015.
4. den Boon JA, Pyeon D, Wang SS, Horswill M, Schiffman M, Sherman M, et al. Molecular transitions from papillomavirus infection to cervical precancer and cancer: role of stromal estrogen receptor signaling. Proc Natl Acad Sci U S A. 2015;112(25):E3255–64. https://doi.org/10.1073/pnas.1509322112.
5. Babion I, Miok V, Jaspers A, Huseinovic A, Steenbergen RDM, van Wieringen WN, et al. Identification of deregulated pathways, key regulators, and novel miRNA-mRNA interactions in HPV-mediated transformation. Cancers (Basel). 2020;12(3):700.
6. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res. 2013;41(Database issue):D991–5. https://doi.org/10.1093/nar/gks1193.
7. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002; 30(1):207–10. https://doi.org/10.1093/nar/30.1.207.
8. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. 2007;35(Web Server issue):W169–75.
9. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30. https://doi.org/10.1093/nar/28.1.27.
10. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50. https://doi.org/10.1073/pnas.0506580102.
11. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet. 2003;34(3):267–73. https://doi.org/10.1038/ng1180.
12. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRI NG v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607–d613. https://doi.org/10.1093/nar/gky1131.
13. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504. https://doi.org/10.1101/gr.1239303.
14. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003; 4(1):2. https://doi.org/10.1186/1471-2105-4-2.
15. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014;8 Suppl 4(Suppl 4):S11.
16. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45(W1):W98–w102. https://doi.org/10.1093/nar/gkx247.
17. Dasgupta S, Ewing-Graham PC, Swagemakers SMA, van der Spek PJ, van Doorn HC, Noordhoek Hegt V, et al. Precursor lesions of vulvar squamous cell carcinoma - histology and biomarkers: a systematic review. Crit Rev Oncol Hematol. 2020;147:102866. https://doi.org/10.1016/j.critrevonc.2020.102866.
18. Leal SM Jr, Gulley ML. Current and emerging molecular tests for human papillomavirus-related neoplasia in the genomic era. J Mol Diagn. 2017; 19(3):366–77. https://doi.org/10.1016/j.jmoldx.2017.01.006.
19. Zhao Q, Li H, Zhu L, Hu S, Xi X, Liu Y, et al. Bioinformatics analysis shows that TOP2A functions as a key candidate gene in the progression of cervical cancer. Biomed Rep. 2020;13(4):21. https://doi.org/10.3892/br.2020.1328.
20. Mathevet P, Frappart L, Hittelman W. Cervix dysplasias: study of Rb and p53 gene expression and correlation with mitotic activity. Gynecol Obstet Fertil. 2000;28(1):44–50.
21. Wen X, Liu S, Cui M. Effect of BRCA1 on the concurrent Chemoradiotherapy resistance of cervical squamous cell carcinoma based on transcriptome sequencing analysis. Biomed Res Int. 2020;2020:3598417.
22. Zhou YH, Fan WF, Deng J, Xi HL. Establishment and analysis of the prediction model for cervical squamous cell carcinoma. Eur Rev Med Pharmacol Sci. 2017;21(22):5042–8. https://doi.org/10.26355/eurrev_201711_13816.
23. Suman S, Mishra A. An interaction network driven approach for identifying biomarkers for progressing cervical intraepithelial neoplasia. Sci Rep. 2018; 8(1):12927. https://doi.org/10.1038/s41598-018-31187-x.
24. Guo W, Yu H, Zhang L, Chen X, Liu Y, Wang Y, et al. Effect of hyperoside on cervical cancer cells and transcriptome analysis of differentially expressed genes. Cancer Cell Int. 2019;19(1):235. https://doi.org/10.1186/s12935-019-0953-4.
25. Manavi M, Hudelist G, Fink-Retter A, Gschwandtler-Kaulich D, Pischinger K, Czerwenka K. Gene profiling in pap-cell smears of high-risk human papillomavirus-positive squamous cervical carcinoma. Gynecol Oncol. 2007; 105(2):418–26. https://doi.org/10.1016/j.ygyno.2006.12.032.
26. Yang C, Xu X, Jin H. Identification of potential miRNAs and candidate genes of cervical intraepithelial neoplasia by bioinformatic analysis. Eur J Gynaecol Oncol. 2016;37(4):469–73.
27. Iancu IV, Botezatu A, Goia-Ruşanu CD, Stănescu A, Huică I, Nistor E, et al. TGF-beta signalling pathway factors in HPV-induced cervical lesions. Roum Arch Microbiol Immunol. 2010;69(3):113–8.
28. Zong S, Liu X, Zhou N, Yue Y. E2F7, EREG, miR-451a and miR-106b-5p are associated with the cervical cancer development. Arch Gynecol Obstet. 2019;299(4):1089–98. https://doi.org/10.1007/s00404-018-5007-y.
29. Rajkumar T, Sabitha K, Vijayalakshmi N, Shirley S, Bose MV, Gopal G, et al. Identification and validation of genes involved in cervical tumourigenesis. BMC Cancer. 2011;11(1):80. https://doi.org/10.1186/1471-2407-11-80.
30. Li S, Liu N, Piao J, Meng F, Li Y. CCNB1 expedites the progression of cervical squamous cell carcinoma via the regulation by FOXM1. Onco Targets Ther. 2020;13:12383–95. https://doi.org/10.2147/OTT.S279951.
31. Wu K, Yi Y, Liu F, Wu W, Chen Y, Zhang W. Identification of key pathways and genes in the progression of cervical cancer using bioinformatics analysis. Oncol Lett. 2018;16(1):1003–9. https://doi.org/10.3892/ol.2018.8768.
32. Klymenko T, Gu Q, Herbert I, Stevenson A, Iliev V, Watkins G, et al. RNA-Seq Analysis of differentiated keratinocytes reveals a massive response to late events during human papillomavirus 16 infection, including loss of epithelial barrier function. J Virol. 2017;91(24):e01001–17.
33. Liu J, Nie S, Gao M, Jiang Y, Wan Y, Ma X, et al. Identification of EPHX2 and RMI2 as two novel key genes in cervical squamous cell carcinoma by an integrated bioinformatic analysis. J Cell Physiol. 2019;234(11):21260–73. https://doi.org/10.1002/jcp.28731.

34. Qiu HZ, Huang J, Xiang CC, Li R, Zuo ED, Zhang Y, et al. Screening and discovery of new potential biomarkers and small molecule drugs for cervical Cancer: a bioinformatics analysis. Technol Cancer Res Treat. 2020;19: 1533033820980112.

35. Yang HJ, Xue JM, Li J, Wan LH, Zhu YX. Identification of key genes and pathways of diagnosis and prognosis in cervical cancer by bioinformatics analysis. Mol Genet Genomic Med. 2020;8(6):e1200. https://doi.org/10.1002/mgg3.1200.

36. Xie Q, Ou-Yang W, Zhang M, Wang H, Yue Q. Decreased expression of NUSAP1 predicts poor overall survival in cervical Cancer. J Cancer. 2020; 11(10):2852–63. https://doi.org/10.7150/jca.34640.

37. van Dam PA, Rolfo C, Ruiz R, Pauwels P, Van Berckelaer C, Trinh XB, et al. Potential new biomarkers for squamous carcinoma of the uterine cervix. ESMO Open. 2018;3(4):e000352. https://doi.org/10.1136/esmoopen-2018-000352.

38. Xue JM, Liu Y, Wan LH, Zhu YX. Comprehensive analysis of differential gene expression to identify common gene signatures in multiple cancers. Med Sci Monit. 2020;26:e919953.

39. Wu DM, Shi J, Liu T, Deng SH, Han R, Xu Y. Integrated analysis reveals down-regulation of SPARCL1 is correlated with cervical cancer development and progression. Cancer Biomark. 2018;21(2):355–65. https://doi.org/10.3233/CBM-170501.

40. Yuan Y, Shi X, Li B, Peng M, Zhu T, Lv G, et al. Integrated analysis of key microRNAs /TFs /mRNAs/ in HPV-positive cervical cancer based on microRNA sequencing and bioinformatics analysis. Pathol Res Pract. 2020; 216(6):152952. https://doi.org/10.1016/j.prp.2020.152952.

41. Yi Y, Liu Y, Wu W, Wu K, Zhang W. Reconstruction and analysis of circRNA-miRNA-mRNA network in the pathology of cervical cancer. Oncol Rep. 2019; 41(4):2209–25. https://doi.org/10.3892/or.2019.7028.

42. Bourmenskaya O, Shubina E, Trofimov D, Rebrikov D, Sabdulaeva E, Nepsha O, et al. Host gene expression profiling of cervical smear is eligible for cancer risk evaluation. J Clin Pathol. 2013;66(4):282–5. https://doi.org/10.1136/jclinpath-2012-201313.

43. Zhao L, Zhang Z, Lou H, Liang J, Yan X, Li W, et al. Exploration of the molecular mechanisms of cervical cancer based on mRNA expression profiles and predicted microRNA interactions. Oncol Lett. 2018;15(6):8965–72. https://doi.org/10.3892/ol.2018.8494.

44. Papasavvas E, Kossenkov AV, Azzoni L, Zetola NM, Mackiewicz A, Ross BN, et al. Gene expression profiling informs HPV cervical histopathology but not recurrence/relapse after LEEP in ART-suppressed HIV+HPV+ women. Carcinogenesis. 2019;40(2):225–33. https://doi.org/10.1093/carcin/bgy149.

45. Kaur G, Balasubramaniam SD, Lee YJ, Balakrishnan V, Oon CE. Minichromosome maintenance complex (MCM) genes profiling and MCM2 protein expression in cervical Cancer development. Asian Pac J Cancer Prev. 2019;20(10):3043–9. https://doi.org/10.31557/APJCP.2019.20.10.3043.

46. Xu Z, Zhou Y, Shi F, Cao Y, Dinh TLA, Wan J, et al. Investigation of differentially-expressed microRNAs and genes in cervical cancer using an integrated bioinformatics analysis. Oncol Lett. 2017;13(4):2784–90. https://doi.org/10.3892/ol.2017.5766.

47. Mine KL, Shulzhenko N, Yambartsev A, Rochman M, Sanson GF, Lando M, et al. Gene network reconstruction reveals cell cycle and antiviral genes as major drivers of cervical cancer. Nat Commun. 2013;4(1):1806. https://doi.org/10.1038/ncomms2693.

48. van Dam PA, van Dam PJ, Rolfo C, Giallombardo M, van Berckelaer C, Trinh XB, et al. In silico pathway analysis in cervical carcinoma reveals potential new targets for treatment. Oncotarget. 2016;7(3):2780–95. https://doi.org/10.18632/oncotarget.6667.

49. Luo Y, Wu Y, Peng Y, Liu X, Bie J, Li S. Systematic analysis to identify a key role of CDK1 in mediating gene interaction networks in cervical cancer development. Ir J Med Sci. 2016;185(1):231–9. https://doi.org/10.1007/s11845-015-1283-8.

50. Jiang L, Zeng X, Wang Z, Ji N, Zhou Y, Liu X, et al. Oral cancer overexpressed 1 (ORAOV1) regulates cell cycle and apoptosis in cervical cancer HeLa cells. Mol Cancer. 2010;9(1):20. https://doi.org/10.1186/1476-4598-9-20.

51. Mou Z, Xu X, Dong M, Xu J. MicroRNA-148b acts as a tumor suppressor in cervical Cancer by inducing G1/S-phase cell cycle arrest and apoptosis in a Caspase-3-dependent manner. Med Sci Monit. 2016;22:2809–15. https://doi.org/10.12659/MSM.896862.

52. Xie Y, Shen YT, Kapoor A, Ojo D, Wei F, De Melo J, et al. Dataset on the effects of CYB5D2 on the distribution of HeLa cervical cancer cell cycle. Data Brief. 2016;6:811–6. https://doi.org/10.1016/j.dib.2016.01.036.

53. Zhao JL, Zhang L, Guo X, Wang JH, Zhou W, Liu M, et al. miR-212/132 downregulates SMAD2 expression to suppress the G1/S phase transition of the cell cycle and the epithelial to mesenchymal transition in cervical cancer cells. IUBMB Life. 2015;67(5):380–94. https://doi.org/10.1002/iub.1381.

54. Das M, Prasad SB, Yadav SS, Modi A, Singh S, Pradhan S, et al. HPV-type-specific response of cervical cancer cells to cisplatin after silencing replication licensing factor MCM4. Tumour Biol. 2015;36(12):9987–94. https://doi.org/10.1007/s13277-015-3782-7.

55. Wang Y, Li Y, Zhang WY, Xia QJ, Li HG, Wang R, et al. mRNA expression of minichromosome maintenance 2 in colonic adenoma and adenocarcinoma. Eur J Cancer Prev. 2009;18(1):40–5. https://doi.org/10.1097/CEJ.0b013e32830c8d5a.

56. Gakiopoulou H, Korkolopoulou P, Levidou G, Thymara I, Saetta A, Piperi C, et al. Minichromosome maintenance proteins 2 and 5 in non-benign epithelial ovarian tumours: relationship with cell cycle regulators and prognostic implications. Br J Cancer. 2007;97(8):1124–34. https://doi.org/10.1038/sj.bjc.6603992.

57. Burger M, Denzinger S, Hartmann A, Wieland WF, Stoehr R, Obermann EC. Mcm2 predicts recurrence hazard in stage ta/T1 bladder cancer more accurately than CK20, Ki67 and histological grade. Br J Cancer. 2007;96(11): 1711–5. https://doi.org/10.1038/sj.bjc.6603784.

58. Zhang X, Teng Y, Yang F, Wang M, Hong X, Ye LG, et al. MCM2 is a therapeutic target of lovastatin in human non-small cell lung carcinomas. Oncol Rep. 2015;33(5):2599–605. https://doi.org/10.3892/or.2015.3822.

59. Zheng J. Diagnostic value of MCM2 immunocytochemical staining in cervical lesions and its relationship with HPV infection. Int J Clin Exp Pathol. 2015;8(1):875–80.

60. Liu D, Zhang XX, Xi BX, Wan DY, Li L, Zhou J, et al. Sine oculis homeobox homolog 1 promotes DNA replication and cell proliferation in cervical cancer. Int J Oncol. 2014;45(3):1232–40. https://doi.org/10.3892/ijo.2014.2510.

61. Niu G, Wang D, Pei Y, Sun L. Systematic identification of key genes and pathways in the development of invasive cervical cancer. Gene. 2017;618: 28–41. https://doi.org/10.1016/j.gene.2017.03.018.

62. Yan R, Shuai H, Luo X, Wang X, Guan B. The clinical and prognostic value of CXCL8 in cervical carcinoma patients: immunohistochemical analysis. Biosci Rep. 2017;37(5):BSR20171021.

63. Bai L, Yao N, Qiao G, Wu L, Ma X. CXCL5 contributes to the tumorigenicity of cervical cancer and is post-transcriptionally regulated by miR-577. Int J Clin Exp Pathol. 2020;13(12):2984–93.

64. Yu B, Chen L, Zhang W, Li Y, Zhang Y, Gao Y, et al. TOP2A and CENPF are synergistic master regulators activated in cervical cancer. BMC Med Genet. 2020;13(1):145. https://doi.org/10.1186/s12920-020-00800-2.

65. Chen H, Wang X, Jia H, Tao Y, Zhou H, Wang M, et al. Bioinformatics analysis of key genes and pathways of cervical Cancer. Onco Targets Ther. 2020;13:13275–83. https://doi.org/10.2147/OTT.S281533.

66. Ouyang F, Liu J, Xia M, Lin C, Wu X, Ye L, et al. GINS2 is a novel prognostic biomarker and promotes tumor progression in early-stage cervical cancer. Oncol Rep. 2017;37(5):2652–62. https://doi.org/10.3892/or.2017.5573.

67. Zhang W, He W, Shi Y, Gu H, Li M, Liu Z, et al. High expression of Klf20A is associated with poor overall survival and tumor progression in early-stage cervical squamous cell carcinoma. PLoS One. 2016;11(12):e0167449. https://doi.org/10.1371/journal.pone.0167449.

Publisher's Note