

# Identification of Key Genes Associated With Radiotherapy Resistance in Cervical Cancer by Bioinformatics

**zhenhua Zhang**

Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China, 646000

**kechao Xiang**

Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China, 646000

**longjing Tan**

Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China, 646000

**xiuju Du**

Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China, 646000

**huailin He**

Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China, 646000

**dan Li**

Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China, 646000

**qinglian Wen** (✉ [wql73115@hotmail.com](mailto:wql73115@hotmail.com))

Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China, 646000

**li Li**

Affiliated Cancer Hospital of Guangxi Medical University, Guangxi, China

---

## Research Article

**Keywords:** bioinformatics, cervical squamous cell carcinoma, radiotherapy, differentially expressed genes, prognosis

**Posted Date:** February 1st, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1103416/v1>

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

[Read Full License](#)

---

# Abstract

**Background:** Cervical cancer is one of the common malignant tumors in women, among which squamous cell carcinoma of the cervix accounts for more than 90%. Currently, 30% of patients with intermediate to advanced squamous cervical cancer are still uncontrolled or recurrent after standard radical simultaneous radiotherapy; therefore, the search for key genes affecting the sensitivity of radiotherapy may lead to new strategies for treatment.

**Methods:** Differentially expressed genes were identified by GEO2R from the gene expression omnibus (GEO) website, prognosis-related genes for cervical cancer were obtained from other databases, then gene ontology (GO) was analyzed by DAVID. Meanwhile, protein-protein interaction network was constructed by STRING, By Online analysis of differentially expressed genes, prognostic genes, and CCDB(cervical cancer gene database website: <http://crdd.osdd.net/raghava/ccdb/index.php>) through the OncoLnc database, we aim to search for differentially expressed genes associated with prognosis validation by immunohistochemistry.

**Result:** 298 differentially expressed genes, 724 genes associated with prognosis were found. Eventually, 2 genes, ASPH and NKAPP1, were identified through validation as genes that affect both sensitivities to radiotherapy and survival. Then, immunohistochemical results showed that the ASPH gene was indeed highly expressed in the radiotherapy-resistant group and had lower OS and PFS.

**Conclusion:** This study aims to better understand the characteristics of cervical cancer radiation therapy resistance-related genes through bioinformatics, and provide further research ideas for finding new mechanisms and potential therapeutic targets related to cervical cancer radiation therapy.

## Introduction

Cervical cancer is one of the common malignant tumors in women, and squamous cell carcinoma of the cervix accounts for more than 90%. For patients with intermediate and advanced stages, simultaneous radiotherapy is the most effective treatment. However, it has been found that with the same clinical factors such as the same stage and the same synchronized radiotherapy regimen, some of the poor outcomes were previously attributed to individual differences, and with the flourishing of genetic research, it is now believed that individual differences (especially genes in different individuals) differ in treatment sensitivity. Currently, the main research direction in cervical cancer treatment sensitivity studies is still radiotherapy sensitivity studies. Radiotherapy sensitivity mainly includes four factors: the intrinsic sensitivity of tumor cells, the presence of oxygen-depleted cells, the proportion of oxygen-depleted clonal cells and the repair of tumor radiation damage[1]. It is clinically important to find the key molecular biological mechanisms affecting the intrinsic sensitivity of tumor cells, which can further carry out molecular typing, develop new molecular targeting drugs, and guide and improve the therapeutic efficacy. The molecular mechanisms associated with radiotherapy sensitivity in cervical cancer are very complex, such as the BCL2 family proteins BCL2 and BCL-XL[2], genes such as EGFR, HER-2, p53, p21, Ki-67, HIF,

VEGF, COX-2[3], FA/BRCA pathway of which FANCD2, RAD51, BRCA1 and BRIP1 and other genes[4], but different literature reports different genes and pathways resistant to radiotherapy, the variability of the foothold of each study, the different results obtained, may not be comprehensive, how to analyze these massive gene-related data to obtain clinically valuable genetic information, bioinformatics method[5] provides us with the research method.

The intrinsic sensitivity of tumor cells to radiotherapy is related to the intrinsic sensitivity of cells before radiotherapy and the damage repair of tumor cells after radiotherapy, etc. The intrinsic sensitivity of tumor cells is the result of the joint action of some oncogenes and oncogenes in the process of tumor development, and the damage occurs after tumor cells receive radiotherapy, while some genetic changes occur damage repair, which is a complex multi-genes and multiple genetic pathways cross each other in the whole process. The whole process is a complex biological process with multiple genes and multiple genetic pathways intersecting each other. It is difficult to analyze the radiosensitivity of tumors by analyzing the expression of one or several genes alone in a comprehensive manner, and later molecular biology techniques such as cDNA and gene microarrays were adopted to obtain informative, reproducible, and easy and reliable genetic information and applied to radiotherapy sensitivity studies in cervical cancer[4, 6, 7].

Gene expression data analysis is an important part of bioinformatics and is a hot spot and focus of current bioinformatics research. It reflects the abundance of mRNA, the gene transcription product, in the cell obtained by direct or indirect measurement, which is the data that allows analysis of which genes have been altered in expression, how they are related to each other, and how their activities are affected under different conditions.

## Materials And Methods

### 1.1 Materials

#### 1.1.1 GEO database

The Gene Expression Omnibus (Gene Expression Omnibus, GEO, <http://www.ncbi.nlm.nih.gov/geo/>)[8] is the largest public database for storing high-throughput molecular abundance data in the world, mainly gene expression data. Users can submit, store and retrieve data in multiple formats and use them for free. Using cervical cancer as the search keyword, restricting the study type to expression profiling by array, and restricting the species to homo sapiens, the data set of gene expression profiles related to radiotherapy for cervical cancer that have been reported worldwide were retrieved, and data from controlled studies that matched the sensitivity and resistance to radiotherapy for cervical cancer were selected.

Relevant data were retrieved from the GEO database until December 2015, and the selected data met the following criteria: (1) the dataset must be genome-wide expression mRNA microarray data for cervical cancer; (2) these data must be controlled studies on radiotherapy sensitivity versus radiotherapy

resistance for squamous cervical cancer; (3) the dataset case (radiotherapy resistant)-control (radiotherapy sensitive) groups must include or exceed 3 (4) Clear information on sensitivity and resistance of each sample must be given. Data sets that meet the above criteria will be included in this study. Among them, GSE19526 which was the data for the in vitro study of radiotherapy sensitivity, due to only 1 sample was not included in the analysis, and GSE6213 was not included in the analysis because of the lack of samples corresponding to radiotherapy sensitivity, only 2 data sets finally met our criteria, GSE56303[9] and GSE56363[4] (see Table 1). 2 data sets Both were different grouped gene expression datasets for evaluating the recent outcome after radiotherapy in patients with intermediate to advanced cervical cancer, except that GSE56303 had a group of some pathological specimens with adenocarcinoma and GSE56363 were all squamous carcinomas.

### 1.1.2 proteinatlas data

The proteinatlas data, published in Science by Mathias Uhlen et al[10] in 2017, are transcriptome-wide data using a systematic level approach to analyze protein-coding genes associated with clinical outcomes in 17 major cancers, including cervical cancer as well. All data were published in its <http://www.proteinatlas.org/humanproteome/pathology>, giving data for our analysis of prognosis-related indicators.

### 1.1.3 STRING database

STRING (search tool for the retrieval of interacting genes/proteins), at <https://string-db.org/>, is a database for finding known protein-protein interrelationships and predicting protein-protein interrelationships, which is currently the One of the largest information about protein interactions, contains experimental data, text mining results in pubmed, integrated data from other databases, and can also be used to predict results using bioinformatics methods, including analysis of genetic data based on microarray data. The study of protein interactions through this database helps to mine the core regulatory genes.

### 1.1.3 CCDB data

The Cervical Cancer Gene Database CCDB (Cervical Cancer Gene Database, at <http://crdd.osdd.net/raghava/ccdb/index.php>)[11], is a manually compiled, experimentally validated catalog containing genes involved in different stages of the cervical cancer formation process. Each entry contains information about genes and protein sequences, their location, structure, function, chromosomal position, accession number, gene, CDS size, etc. In addition, the database is richly cross-referenced with databases such as Unigene, HPRD, HGNC, Ensemble, and OMIM.CCDB also provides references to the relevant literature for the genes included in the database.

### 1.1.4 OncoLnc data

OncoLnc data, at <http://www.oncolnc.org/>, is a TCGA-based repurposing online analysis tool that allows survival analysis of mRNA, miRNA, or lncRNA, which also contains survival information related to cervical cancer. The relevant genes can be analyzed online at this URL to obtain whether the gene affects survival.

## 1.1.5 Immunohistochemistry

### 1.1.5.1 Experimental objects

Patients with intermediate to advanced cervical squamous carcinoma were selected, of which 87 specimens were preserved in the Department of Pathology of the Cancer Hospital of Guangxi Medical University from January 2011 to June 2015, including 44 cases in the radiotherapy-sensitive group and 43 cases in the radiotherapy-resistant group. Each pathological specimen was sectioned by the pathology teacher from paraffin block in 5 sheets of 5 microns each, attached to anti-dislodging slides, fixed in a constant temperature oven at 60°C for 2 hours, cooled and mounted in slide boxes for spare.

### 1.1.5.2 Main reagents

- 1) Mouse anti-human ASPH polyclonal antibody IgG, purchased from Bio-Swamp
- 2) General secondary antibody Beijing Lambert Biotechnology Co.
- 3) DAB staining solution Nanjing Jiancheng Biotechnology Co.
- 4) Other reagents such as: sodium citrate buffer, PBS, hydrogen peroxide, different concentrations of ethanol, different xylenes, hematoxylin, neutral gum, etc. were provided by the laboratory of the Department of Pathology, Southwest Medical University Hospital.

### 1.1.5.3 Main experimental apparatus

General microscopes, Olympus microscopes, autoclaves, refrigerators, thermostats, spikers, etc. were provided by the laboratory of the Department of Pathology, Southwest Medical University Hospital.

## 1.2 Method

### 1.2.1 Batch effect processing

In the sequencing, microarray, RNA-seq, DNA methylation, and proteomics, and other histological studies, as the samples are collected and processed in multiple batches as well as collected at different times, and also the fluorescent signals are converted into digital information, there may be some technical reasons such as analysis platform, laboratory samples, experimenters, programmers, or reagents that may produce differences during processing, called batch effects, and deviations in batch effects may lead to incorrect analysis of downstream results. Therefore, homogenization of batch effects is needed, including visualization of data, hierarchical clustering, principal component analysis, and analysis of variance.

In the actual sample processing, it is necessary to first consider whether batch correction is needed, and then how to correct and what tool to choose for correction. BatchQC[12] is an R package for the Shiny App, a tool that deals with these issues by providing interactive diagnostics, visualization, and statistical

analysis. BatchQC can also apply some existing batch correction methods so that the user can compare the advantages and disadvantages between the methods in an interactive way and ultimately select the results of the batch correction. The output of BatchQC is organized into multiple tabs, which will eventually generate a report that contains the results of summary, differential expression analysis, significance of results, median correlation, heat map, expression correlation between samples, and PCA analysis, so the tool is a good choice.

### 1.2.2 Screening for differentially expressed genes

A large amount of gene expression data obtained by microarray, gene microarray and other technologies, but these data are characterized by small sample size, high dimensionality, strong correlation, etc. There may be significantly different genes for small changes, and a statistical method needs to be selected to find out these meaningful genes with significant differences. Commonly used methods include ANOVA t-test, which usually doubles or halves the expression level as a criterion to determine whether there is an expression difference.

Differentially expressed gene screening using the limma package. limma is an R/Bioconductor package that can cover every major step and function of gene analysis, from data import, preprocessing, normalization to differential expression analysis and gene characterization. We first transformed the data set originating from GEO into an expression matrix of genes, then preprocessed, normalized with the limma package, specified the study and control groups, and finally calculated differentially expressed genes with the lmFit function. In this study, we took  $P \leq 0.05$  as the threshold value and combined with the fold change (FC) analysis method, when fold change (FC)  $\geq 2$ , the gene was considered up-regulated, and if  $FC \leq 0.5$ , the gene was considered down-regulated.

Every major step and function of gene analysis can be covered, from data import, preprocessing, normalization to differential expression analysis and gene characterization, etc., at <http://www.bioconductor.org/> web page. Firstly, the expression profile dataset downloaded from GEO was alerted to pre-processing and standardization with the algorithm of limma package to extract differentially expressed genes from the expression profile data, in which the study and control groups were set, and the thresholds of p-value and logFC were set to extract differentially expressed genes.  $p < 0.05$  was used as the criterion for screening differentially expressed genes. Combined with the ploidy analysis (fold change, FC), when  $\log FC < 0$ , gene expression was down-regulated; when  $\log FC > 0$ , it indicated up-regulation.

### 1.2.3 Gene function annotation

DAVID software (Database for Annotation, Visualization and Integrated Discovery, <https://david.ncifcrf.gov/>)[5, 7] is a bioinformatics database containing many integrated biological data and analytical tools DAVID provides a large scale list of genes or proteins that can be used to obtain systematic and comprehensive annotated information on biological functions, mainly for gene probe conversion, gene function classification, gene enrichment analysis and functional annotation clustering.

## 1.2.4 Immunohistochemistry

### 1.2.4.1 Experimental methods and procedures

#### 1) Dewaxing and hydration of slices

The prepared paraffin sections were placed in xylene I for 10 min, xylene II for 10 min, 100% ethanol for 5 min, 100% ethanol for 54 min, 95% ethanol for 5 min, 85% ethanol for 2 min, 70% ethanol for 2 min, 50% ethanol for 2 min, followed by 2 immersion washes in distilled water for 5 min each, followed by 5 min in 0.01 PBS solution in turn.

#### 2) Antigen repair

Sections were immersed in 0.01 M citrate buffer (PH 6.0), heated to boiling in an autoclave and then cooled naturally for 35 min, followed by PBS rinsing 3 times for 2 min each.

#### 3) Closure of endogenous peroxidase

Add 100  $\mu$ L of endogenous peroxidase blocker, incubate for 10 min at room temperature, and rinse 3 times with PBS for 2 min each time.

#### 4) Addition of primary antibody

After filter paper is blotted dry around the closure solution, add primary antibody (mouse anti-human ASPH polyclonal antibody, PBS as negative control) at a concentration of 1:200, and then incubate in a 37°C thermostat for 1 hour, and then rinse the PBS solution 3 times for 3 minutes each time.

#### 5) Add secondary antibody

After absorbing the excess liquid on the slide with filter paper, add universal secondary antibody on the slide sequentially, incubate for 30 minutes at 27°C, and then rinse 3 times with PBS solution for 3 minutes each time.

#### 6) DAB color development

Prepare DAB chromogenic solution, add 100  $\mu$ L of freshly prepared DAB chromogenic solution to each section and incubate for 5 min at room temperature.

#### 7) Re-staining

After rinsing the sections under running water, 100  $\mu$ L of hematoxylin solution was added for restaining, and distilled water was rinsed for 5 minutes.

#### 8) Dehydration, transparency and sealing

Place the rinsed sections into the ethanol concentration of 70%, 80%, 90%, 95%, 100%, 100% for 2 minutes each, then immerse the sections into xylene I for 2 minutes and xylene II for 2 minutes, dry the xylene around the sections, use neutral resin drops next to the tissue then cover with a coverslip, seal the number and dry naturally at room temperature.

#### 1.2.4.2 Determination of immunohistochemical results

Positive protein expression persisted in the cytoplasm or cell membrane of cervical tissues, and brownish-yellow granules were considered positive and no staining was considered negative, and were counted using the semi-quantitative integration method. Five areas of each section were randomly selected under a 400× microscopic field of view, and the average percentage of positive cells in the field of view of that section was calculated. The number of positive cells was scored 0%, 6%-30% was scored 1, 31%-50% was scored 2, 51%-75% was scored 3, and >75% was scored 4. Positive intensity is not colored 0 points, light yellow counts as 1 point, yellow counts as 2 points, and reddish-brown counts as 3 points. The percentage of positive cell count and the positive intensity score were multiplied to give a total score of 0 as -, 1 to 4 as +, 5 to 8 as ++, and 9 to 12 as +++, with + considered as low expression and +++ to ++++ as high expression. For all 87 specimens, 2 pathologists were asked to score them individually.

#### 1.2.4.3 Statistical analysis

All data were analyzed using R language 3.4.3 software (<https://CRAN.R-project.org/package=survival>), survival curves and survival rates were plotted using the Kaplan-Meier method, differences in survival rates between groups were tested using the log-rank test (log-rank method), and P-values were obtained using the chi-square test. Univariate and multifactorial analyses were performed using the ratio COX regression model with 95% confidence intervals (95% CI) for the risk ratio area, and all data were analyzed using a two-sided test, with  $P < 0.05$  indicating statistical differences.

## Results

### 2.1 Batch effect treatment results

The BatchQC package was used to remove the batch effects from the two combined expression data sets, and the results showed that the overall expression levels were similar between the samples (see Figure 1). On the one hand, to test the effect of removing the batch effect, and on the other hand, to test whether the batch effect could distinguish the radiotherapy-sensitive group from the radiotherapy-resistant group, a principal component analysis (PCA) was done on the data after removing the batch effect, and the first 3 were taken for the PCA (see Figure 2). We can see from the figure below that the batch effect is not completely removed, but the results are still acceptable, and at the same time the 2 groups are relatively well differentiated.

### 2.2 Screening results of differentially expressed genes

The data after removing the batch effect were combined together for analysis, with batch as a covariate, and using the limma package, 298 differentially expressed genes (DEGs) were calculated, of which 199 genes were expressed up-regulated and 99 genes were expressed down-regulated. Then we used these differentially expressed genes to do sample clustering, and the results are shown in Figure 3. The 298 differentially expressed genes obtained from the graphs can distinguish the radiotherapy-sensitive group from the resistant group relatively well because the results are too many to be shown completely. The above conclusion also verified that the batch effect was well removed and the obtained differentially expressed genes could well distinguish the sensitive group from the resistant group.

### 2.3 Gene function analysis

Enrichment analysis of these differentially expressed genes for functional classification was first done with DAVID, and the results are shown in Figures 4 (for visualization, the opposite of the logarithm of the base 10 was taken for the raw probabilities). Also, the differentially expressed genes were validated against the 724 genes associated with prognosis from the proteintlas data, and the correlation with the differentially expressed genes obtained in this study was tested using Fisher's exact test, and a p-value of 0.0388 was obtained, indicating that the results obtained were significant. The results showed that differentially expressed genes were significant mainly in functional pathways such as variable shear and energy metabolism.

We did GO functional enrichment of 724 prognostic genes from the obtained differentially expressed genes and proteintlas data, respectively, and the results are shown in Figure 5 to Figure 8. We found that the functions of these two batches of genes are significantly different from both the molecular functional point of view and the biological process point of view. Differentially expressed genes in molecular functions are mainly functions such as base binding related, in biological processes mainly related to DNA stability, cell and organ cell membranes. Whereas in proteintlas data in molecular functions are mainly cytokine interactions, binding of kinases, etc., and in biological processes are mainly related to the activity of leukocytes and lymphocytes.

We downloaded all human protein-protein interaction pairs from STRING, and based on this network all protein-protein interactions between differentially expressed genes will be extracted, and the results are shown in Figure 9. the center of the network is the ELAVL1 gene, which encodes the ELAV1 protein, an RNA-binding protein that is involved in the differentiation of embryonic stem cells and has been shown to be associated with prognosis of a variety of cancers. This is followed by the HSP90AA1 gene, which encodes a protein primarily involved in ATPase activity.

### 2.4 Validation of differentially expressed genes

First, because CCDB data is a set of data related to the catalog of experimental validation of genes involved in different stages of the formation process of cervical cancer, first on this database, the 724 prognosis genes (referred to as Prognosis) for which differentially expressed genes, proteintlas data were obtained in this study were obtained Wayne plots (see Figure 10), and it was seen that no

intersection was found for all 3 groups of genes. Seven genes with intersection of differential expression data with prognosis-related data and nine genes with intersection of differential expression data with prognosis-related data were found (see Table 2).

Then, the 3 data from the previous period were intersected with on OncoLnc and the results are shown in Figure 11. after the clinical survival validation for each gene, we analyzed with 50 as lower percentile and upper percentile. There was no difference in the analysis of survival curves for any of the 9 genes obtained by intersecting differentially expressed genes with CCDB. Validating the obtained differentially expressed genes with the 7 genes that had intersection with the Prognosis group, only ASPH and NKAPP1 had differences in survival curves with a statistically significant  $P < 0.05$ , as shown in Figures 12 and 13. from the survival curves, it can be seen that the low expression of ASPH and high expression of NKAPP1 may lead to increased sensitivity to radiotherapy, thus improving cervical squamous cancer patients' efficacy.

## 2.4 Immunohistochemical results

### 2.4.1 ASPH expression results and patient's general condition

All 87 patients who underwent ASPH expression analysis, combined with the retrospective data analysis collected in Chapter 1, their basic information is shown in Table 3, high expression of ASPH in the radiotherapy-resistant group accounted for 69.76% (30/43) and high expression in the radiotherapy-sensitive group accounted for 40.9% (18/26), the rate of high expression of ASPH in the radiotherapy-resistant group was higher than that in the radiotherapy-sensitive group,  $p < 0.05$ , the difference was statistically significant grouped according to ASPH expression, between the two groups, the mortality rate of those with high ASPH expression was 54.2% significantly higher than that of 28.2% in the low expression group and statistically different, and the difference between their OS and PFS was not statistically significant. There were no significant differences in age, stage, tumor diameter, lymph node metastasis, NLR, PLR, lymph node metastasis, total time of radiotherapy, number of concurrent chemotherapy sessions (TR), and number of adjuvant chemotherapy sessions between the two groups. See Table 4 and Figure 14.

### 2.4.2 ASPH expression profile and survival analysis results

At the end of follow-up, 50 of the 87 patients were alive, and there were significant differences in OS, PFS, and survival between the ASPH low expression group and the high expression group, and significantly lower OS and PFS in those with high ASPH expression, and significant differences. (See Figure 15 and Figure 16)

### 2.4.3 Analysis of clinical factors affecting OS of radiotherapy for intermediate and advanced squamous cervical cancer

Univariate analysis showed that high ASPH expression was a prognostic indicator of OS, with 2.068 times the risk of high ASPH expression compared with low expression (HR=2.06, 95% CI=1.02-4.18,

P<0.005). However, multifactorial analysis did not show statistical differences. Among other factors univariate analysis of radiotherapy sensitivity, tumor diameter, NLR, lymph node metastasis status and total duration of radiotherapy were influential factors of OS, and multifactorial analysis showed that radiotherapy sensitivity, tumor diameter and total duration of radiotherapy were independent factors affecting OS. (See Table 5)

#### 2.4.4 Analysis of clinical factors affecting PFS of radiotherapy for intermediate and advanced squamous carcinoma of the cervix

Eighty-seven patients examined for ASPH were analyzed. Univariate analysis showed that high ASPH expression was a prognostic indicator of PFS, with a 4.8-fold higher risk for high ASPH expression than for low expression (HR=2.12, 95% CI=1.05-4.30, P<0.005). However, multifactorial analysis did not show statistical differences. Among other factors, univariate analysis of radiotherapy sensitivity, tumor diameter, NLR, lymph node metastasis status and total duration of radiotherapy were influential factors of PFS, while multifactorial analysis showed that radiotherapy sensitivity, tumor diameter and total duration of radiotherapy were independent factors affecting PFS.

## Discussion

In addition to the clinicopathologic-related factors, the efficacy of radiotherapy for cervical cancer is related to the intrinsic sensitivity of tumor cells to radiotherapy, and the intrinsic resistance to radiotherapy is a key factor that needs to be unresolved, which is also the focus and hot spot of research in recent years. Starting in 2005 Gaffney, D. K. et al[13] reported that gene expression analysis using RNA amplification after pathological tissue biopsy from patients with cervical cancer was feasible, laying the foundation for future studies. Weidhaas, J.B.[14] found that gene expression was altered during radiotherapy for cervical cancer, which could predict treatment efficacy. Later many researchers started to use this technique to perform analysis of gene expression in radiotherapy-sensitive and resistant groups in the expectation of finding genes and pathways associated with radiotherapy resistance to guide treatment to improve prognosis.

The current published studies using gene microarray studies regarding the analysis of radiotherapy sensitivity in cervical cancer retrieved from pubmed are An, J.S et al. at the Cancer Hospital of Chinese Academy of Medical Sciences in 2013 at the earliest[15], Balacescu, O. et al. in 2014[4] and Fernandez-Retana, J. et al. in 2015[9], and these 3 papers, the first one failed to include the relevant data in this bioinformatics analysis because it did not upload the relevant expressed genes to public databases such as GEO, and only the latter two were included in this study for bioinformatics analysis. We used bioinformatics to obtain differentially expressed genes for these 2 data sets and analyzed them in terms of gene functional classification, molecular function, biological processes, and protein interactions, and the results showed that the differentially expressed genes obtained could well distinguish radiotherapy-sensitive from radiotherapy-resistant, in terms of gene functional classification mainly in the functional pathways mainly in variable shear and energy metabolism, and found that the proteins involved in The

two main ELAVL1 and HSP90AA1 genes involved in protein interactions were identified. The results are somewhat different from the literature published by An, J.S et al. This study showed that radiotherapy sensitivity was related in multiple pathways: DNA damage repair, apoptosis, cell cycle, MAPK signaling pathway, anaerobic glycolysis and glutathione metabolism, in the radiotherapy-sensitive group SMUG1 and CDK7 genes were downregulated in the DNA damage repair pathway, ATM genes were upregulated, in the MAPK pathway IL1R1, PDGFRA and TGFB3 genes upregulated and HRAS downregulated, IL1R1, PRKAR1A and ATM upregulated while CAST, BNIP3 and BAK1 downregulated in tumor necrosis-related genes, and TGFB3 and ATM upregulated while CDK7, RAN and HRAS downregulated in cell cycle-related genes. The same study by Balacescu O et al[4], also included in our bioinformatics analysis, considered the DNA damage repair pathway as a major pathway study for radiotherapy sensitivity, which showed differences in 17 genes on this pathway (RAD51, BRIP1, BLM, BRCA1, BRCA2, BRCC3, HLTF, FANCD2, FANCI, FANCM, FANCL, ATF1, E2F4, E2F2, SMARCA2, SMARCA4, and RFC1), with a particular focus on which overexpression of BRCA1, BRCA2, RAD51, BRIP1 (BACH1), FANCD2, BLM, and RFC in the radiotherapy-resistant group can be detected in their pathways by cell cycle arrest and homologous recombination leading to the activation of DNA repair mechanisms. In addition Kitahara, O. et al[7] studying the expression of radiotherapy sensitivity genes in cervical cancer with radiotherapy alone found 171 genes differentially between the radiotherapy sensitive and resistant groups, 121 genes were upregulated and 50 genes were downregulated in the radiotherapy sensitive group, and upregulation of genes related to MAPK pathway such as MAP3K2 and RAB5C family in the gene pathway played an important role in radiotherapy. The upregulation of MAPK pathway related genes such as MAP3K2 and RAB5C family played an important role in radiotherapy sensitivity, and the downregulation of DNA repair related genes such as XRCC5 in the radiotherapy sensitive group, and the downregulation of other pathway genes such as LDHA, this study especially focused on the high expression of XRCC5 in the radiotherapy resistant group, and the high expression of ALDH1 and RBP1 in the radiotherapy sensitive group. An, J.S et al[15] found that the PDGFRA and PRKAR gene families were consistently upregulated in the radiotherapy-sensitive group, but no overlapping consistent genes were found between these publications and the results of the current bioinformatics analysis.

In the above studies, most of them found partial overlap with radiotherapy-sensitive related gene pathways, but there are fewer studies on whether radiotherapy-sensitive genes are directly related to prognosis. We found 7 genes that were consistent with prognosis and 9 genes that were consistent with the formation process of cervical cancer by bioinformatics although we did not find genes common to the 3 data sets, and found differentially expressed genes affecting radiotherapy sensitivity in cervical cancer by gene function classification and analysis with proteomic data showed that the main function is on the pathway of energy metabolism, which is different from the pathway affecting prognosis differently, suggesting that the two may work together through synergy to accomplish prognosis-related effects. We identified two genes, ASPH and NKAPP1, as genes that affect both sensitivity to radiotherapy and survival through validation of experimental and clinically relevant data.

The NKAPP1 (NFKB activating protein pseudogene) gene is an NK- $\kappa$ B activating protein pseudogene, a segment of the DNA base sequence that is very similar to the sequence of a gene that has been in other

organisms, but is unable to perform its original function and make proteins, but studies have shown that these genes are often involved in transcriptional regulation and play a role in multiple functions in cancer pathogenesis, and subtype analysis of cancer can be used as a prognosis-related biological indicator[16, 17], and their expression in pathological tissue specimens could not be studied in this study because of the absence of protein function.

ASPH gene (aspartate beta-hydroxylase), aspartate-aspartate beta-hydroxylase is a highly conserved deoxygenase present in cells since embryonic stage in mammalian arteries[18]. ASPH gene DNA length is 2277bp and contains 27 exons, which encodes 4 of the proteins ASPH, HUMBUG Junctate and junctin, of which ASPH is the major translation product. ASPH overexpression may promote tumor cell formation, proliferation, invasion and metastasis[18], becoming an indicator of malignant expression and considered as a potential tumor marker, which is currently more studied in hepatocellular carcinoma. The positive expression rate of ASPH in cervical cancer cells is about 88.5%[19], and there are no relevant studies on its role in cervical cancer and related functional studies on whether there is a difference in ASPH expression in radiotherapy-sensitive and resistant groups.

The final results showed that in intermediate and advanced cervical squamous carcinoma specimen pathology, the high expression of ASPH was higher in the radiotherapy-resistant group than in the radiotherapy-sensitive group, and the high expression of ASPH affected the OS and PFS of patients, and those with high expression had significantly lower OS and PFS than those with low expression, and the univariate analysis of ASPH high expression showed that ASPH was one of the prognostic influencing factors affecting OS and PFS, but the multi The multifactorial analysis did not yield consistent results, which may be related to the small sample size, with only 87 cases included in the analysis, and whether increasing the number of cases may yield positive results, which needs to be confirmed by more studies. There are very few studies on ASPH in cervical cancer, one published in vitro study showed that ASPH was positively expressed in 3 specimens in cervical cancer cell lines, but the results were not very convincing because of the in vitro experiment and the very effective sample size. ASPH in cervical cancer as well as radiotherapy related studies are even less reported[19].

ASPH may become a new tumor marker. Since most studies have shown that ASPH is overexpressed in malignant tumor cells[19] and low or no expression in normal tissues, and overexpressed ASPH can be detected by releasing from tumor cells into human serum and body fluids, ASPH can be used as a novel tumor marker and therapeutic target. Zhang Jizong et al[20] showed that ASPH alone has lower sensitivity than GP73 but higher than AFP for detecting primary liver cancer, but lower specificity than both 2. However, ASPH combined with AFP and GP73 is higher than the single test, with 96% sensitivity, 98% specificity and 97% accuracy in the period. Whether both can be used as tumor markers for other tumors, including cervical cancer, needs to be explored in more studies.

ASPH may become a marker suggestive of prognosis. In studies of primary liver cancer showed that hepatocellular carcinoma cells with high expression of ASPH are more active, migratory, invasive and metastatic, and et al[21] similarly confirmed that ASPH can promote cancer cell re-interrogation

migration, in vivo metastasis and distant metastasis through relevant signaling pathways, suggesting that tumor cells are more aggressive. In vivo studies using computer-assisted synthesis of compounds that inhibit  $\beta$ -hydroxylase have shown to inhibit the North signaling pathway in hepatocellular carcinoma to produce antitumor effects. In cholangiocarcinoma[22] ASPH may promote the progression of cholangiocarcinoma by regulating RB1 phosphorylation. We showed that high expression of ASPH in cervical cancer is more likely to resist radiotherapy, suggesting that possibly those with high ASPH expression have poor sensitivity to treatment with radiotherapy.

ASPH may also be a new pathway for therapy. Studies in pancreatic cancer[23] showed that the drug SNS-622-DM1 coupled with ASPH has a better antitumor effect in vitro, and animal experiments can limit proliferation, promote apoptosis, and have a significant inhibitory effect on tumor growth and metastatic foci. In the study of hepatocellular carcinoma, the antigenicity of ASPH can be used to load ASPH on dendritic cells to induce the production of CD4+ T cells, and the antitumor effect can be achieved through this immune response[24]. Such immunotherapy can also inhibit the cytotoxicity of bile duct cancer cells and suppress the growth and metastasis of intrahepatic tumors.

Our study showed a higher rate of ASPH positive expression in the radiotherapy-resistant group in intermediate to advanced squamous cervical cancer. It is speculated that there may be higher concentrations of ASPH in the blood, and it may be possible that the sensitivity of radiotherapy may be enhanced by corresponding immunotherapy or targeted drugs, thus improving the efficacy, but of course this all needs to be verified by further studies.

## Conclusion

We identified 7 genes consistent with prognosis and 9 genes consistent with the process of cervical cancer formation, among which ELAVL1 and HSP90AA1 are involved in protein inter-righting, and identified 2 genes, ASPH and NKAPP1, as both genes affecting radiotherapy sensitivity and survival, through the analysis of the included GEO database on radiotherapy sensitivity in intermediate and advanced squamous cervical cancer, using bioinformatics methods with relevant databases. Immunohistochemical results showed that ASPH was more highly expressed in the radiotherapy-resistant group than in the radiotherapy-sensitive group in intermediate and advanced squamous cervical cancer, and those with high ASPH expression had lower OS and PFS and could be prognostic indicators in intermediate and advanced squamous cervical cancer, and ASPH could be a tumor marker, prognostic indicator and therapeutic target in squamous cervical cancer.

We hereby declare that the use of all experimental specimens was approved by the Clinical Trial and Ethics Committee of the Affiliated Cancer Hospital of Guangxi Medical University, and the experimental procedures were carried out in strict accordance with the guidelines and regulations related to clinical trials, and informed consent was obtained from the study subjects or their legal guardians.

## Declarations

## Ethical approval and consent to participate

Our trial was approved by the Ethics Committee of the Affiliated Cancer Hospital of Guangxi Medical University, and informed consent was also obtained from the subjects.

## Availability of data and materials

Our data and immunohistochemical specimens were mainly obtained from the Affiliated Cancer Hospital of Guangxi Medical University.

## Conflict of interest

There is no conflict of interest involved in our experiments.

## Author Contributions

zhenhua Zhang, qinglian Wen and li Li conceived and designed the analysis; kechao Xiang Collected the data, Performed the analysis and wrote the paper; longjing Tan Collected the data; huailin He Contributed data or analysis tools; xiuju Du Performed the analysis; All authors reviewed the manuscript. moreover, zhenhua Zhang and kechao Xiang are the first authors, qinglian Wen and li Li are corresponding authors.

## Acknowledgements

We appreciate the contributions of all authors to the manuscript. We also sincerely thank the Department of Pathology, Affiliated Cancer Hospital, Guangxi Medical University for their help.

## References

1. Yin Weibo YZ, Xu Guozhen.: **Radiation Therapy of Oncology (Fourth Edition)[M]**. Beijing, Beijing Medical University and Peking Union Medical University United Press 2008.
2. Wootipoom V, Lekhyananda N, Phunggrassami T, Boonyaphiphat P, Thongsuksai P: **Prognostic significance of Bax, Bcl-2, and p53 expressions in cervical squamous cell carcinoma treated by radiotherapy**. *Gynecol Oncol* 2004, **94**(3):636-642.
3. Petera J, Sirak I, Beranek M, Vosmik M, Drastikova M, Paulikova S, Soumarova R: **Molecular predictive factors of outcome of radiotherapy in cervical cancer**. *Neoplasma* 2011, **58**(6):469-475.
4. Balacescu O, Balacescu L, Tudoran O, Todor N, Rus M, Buiga R, Susman S, Fetica B, Pop L, Maja L *et al*: **Gene expression profiling reveals activation of the FA/BRCA pathway in advanced squamous cervical cancer with intrinsic resistance and therapy failure**. *BMC Cancer* 2014, **14**:246.
5. Moore JH: **Bioinformatics**. *Cell Physiol* 2007(213(2)):365-369.
6. Harima Y, Togashi A, Horikoshi K, Imamura M, Sougawa M, Sawada S, Tsunoda T, Nakamura Y, Katagiri T: **Prediction of outcome of advanced cervical cancer to thermoradiotherapy according to**

- expression profiles of 35 genes selected by cDNA microarray analysis.** *Int J Radiat Oncol Biol Phys* 2004, **60**(1):237-248.
7. Kitahara O, Katagiri T, Tsunoda T, Harima Y, Nakamura Y: **Classification of sensitivity or resistance of cervical cancers to ionizing radiation according to expression profiles of 62 genes selected by cDNA microarray analysis.** *Neoplasia* 2002, **4**(4):295-303.
  8. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M *et al*: **NCBI GEO: archive for functional genomics data sets—update.** *Nucleic acids research* 2013, **41**(Database issue):D991-995.
  9. Fernandez-Retana J, Lasa-Gonsebatt F, Lopez-Urrutia E, Coronel-Martinez J, Cantu De Leon D, Jacobo-Herrera N, Peralta-Zaragoza O, Perez-Montiel D, Reynoso-Noveron N, Vazquez-Romo R *et al*: **Transcript profiling distinguishes complete treatment responders with locally advanced cervical cancer.** *Transl Oncol* 2015, **8**(2):77-84.
  10. Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhorji G, Benfeitas R, Arif M, Liu Z, Edfors F *et al*: **A pathology atlas of the human cancer transcriptome.** *Science (New York, NY)* 2017, **357**(6352).
  11. Agarwal SM, Raghav D, Singh H, Raghava GP: **CCDB: a curated database of genes involved in cervix cancer.** *Nucleic acids research* 2011, **39**(Database issue):D975-979.
  12. Manimaran S, Selby HM, Okrah K, Ruberman C, Leek JT, Quackenbush J, Haibe-Kains B, Bravo HC, Johnson WE: **BatchQC: interactive software for evaluating sample and batch effects in genomic data.** *Bioinformatics* 2016, **32**(24):3836-3838.
  13. Gaffney DK, Winter K, Fuhrman C, Flinner R, Greven K, Ryu J, Forbes A, Kerlin K, Nichols RC, Zempolich K: **Feasibility of RNA collection for micro-array gene expression analysis in the treatment of cervical carcinoma: a scientific correlate of RTOG C-0128.** *Gynecol Oncol* 2005, **97**(2):607-611.
  14. Weidhaas JB, Li SX, Winter K, Ryu J, Jhingran A, Miller B, Dicker AP, Gaffney D: **Changes in gene expression predicting local control in cervical cancer: results from Radiation Therapy Oncology Group 0128.** *Clin Cancer Res* 2009, **15**(12):4199-4206.
  15. An JS, Huang MN, Song YM, Li N, Wu LY, Zhan QM: **A preliminary study of genes related to concomitant chemoradiotherapy resistance in advanced uterine cervical squamous cell carcinoma.** *Chinese medical journal* 2013, **126**(21):4109-4115.
  16. Han L, Yuan Y, Zheng S, Yang Y, Li J, Edgerton ME, Diao L, Xu Y, Verhaak RGW, Liang H: **The Pan-Cancer analysis of pseudogene expression reveals biologically and clinically relevant tumour subtypes.** *Nat Commun* 2014, **5**:3963.
  17. Lu XJ, Gao AM, Ji LJ, Xu J: **Pseudogene in cancer: real functions and promising signature.** *J Med Genet* 2015, **52**(1):17-24.
  18. Ince N, de la Monte SM, Wands JR: **Overexpression of human aspartyl (asparaginy) beta-hydroxylase is associated with malignant transformation.** *Cancer Res* 2000, **60**(5):1261-1266.
  19. Song Kai XX, Wang Wei: **Distribution and detection of ASPH in tumor cells and tumor tissues.** *Journal of Cellular and Molecular Immunology* 2010(26(02)):141-144.

20. Zhang Jizong MJ, Ding Hai: **The diagnostic significance of serum aspartate-asparagine  $\beta$  hydroxylase (ASPH) molecule combined with AFP and GP73 in the diagnosis of primary liver cancer.** *Journal of Shandong University (Medical Sciences) Edition* 2014(52(06)):78-80+84.
21. Zou Q HY, Wang H, Wang K, Xing X, Xia Y, Wan X, Li J, Jiao B, Liu J, Huang A, Wu D, Xiang H, Pawlik TM, Wang H, Lau WY, Wang Y, Shen F. : **Hydroxylase Activity of ASPH Promotes Hepatocellular Carcinoma Metastasis Through Epithelial-to-Mesenchymal Transition Pathway.** 2018 May(31):287-298.
22. Huang CK, Iwagami Y, Zou J, Casulli S, Lu S, Nagaoka K, Ji C, Ogawa K, Cao KY, Gao JS *et al*: **Aspartate beta-hydroxylase promotes cholangiocarcinoma progression by modulating RB1 phosphorylation.** *Cancer Lett* 2018, **429**:1-10.
23. Nagaoka K, Bai X, Ogawa K, Dong X, Zhang S, Zhou Y, Carlson RI, Jiang ZG, Fuller S, Lebowitz MS *et al*: **Anti-tumor activity of antibody drug conjugate targeting aspartate-beta-hydroxylase in pancreatic ductal adenocarcinoma.** *Cancer Lett* 2019, **449**:87-98.
24. Shimoda M, Tomimaru Y, Charpentier KP, Safran H, Carlson RI, Wands J: **Tumor progression-related transmembrane protein aspartate- $\beta$ -hydroxylase is a target for immunotherapy of hepatocellular carcinoma.** *Journal of hepatology* 2012, **56**(5):1129-1135.

## Tables

Table 1 Disease status of 2 sets of genome-wide expression data sets related to radiotherapy and chemotherapy sensitivity of cervical cancer

GEO serial number	Contributor	Chip platform	Sensitive group (number)	Resistance group (number)
GSE56303	Fernandez-Retana J	NimbleGen	63	22
GSE56363	Balacescu O	Agilent-014850	12	9

Table2 Genes at the intersection of 3 databases

DEGs & Prognosis	DEGs & CCDB	Prognosis & CCDB		
ASPH	ATP9A	ADAM9	LDHA	PDK2
DENND2D	CXCR2	AQP3	MCM2	PLOD2
ETFB	FGF1	BCL2	MCM3	POP5
FARSA	GTF2F2	CDA	MCM5	PTPN6
HNRNPA3	HSP90AA1	DSG2	MDM2	RARRES3
RCC2	LSM3	E2F1	MGMT	RASSF1
TARDBP	S100A10	EREG	MMP1	SPP1
	S100P	FGFR2	MMP3	TFRC
	TWIST1	GLTP	NME2	TP73
		IL1A	OSMR	VEGFA
		ITGB1	PCNA	

Table 3 General characteristics of the patient

Clinical factors	ASPH low expression group (N=39)	ASPH high expression group (N=48)	P-value
Radiochemoradiotherapy sensitivity number%			0.013
Sensitivity to chemoradiotherapy	26 (66.7)	18 (37.5)	
Chemoradiotherapy resistance	13 (33.3)	30 (62.5)	
Age (years)	54.03 (10.36)	53.69 (7.90)	0.863
Survival status number, %)			0.027
Survival	28 (71.8)	22 (45.8)	
Dead	11 (28.2)	26 (54.2)	
OSmouth	32.82 (19.62)	30.65 (19.88)	0.611
PFSmouth	31.44 (20.54)	27.54 (20.39)	0.380
stage(number%)			0.385
IIb-IIIa	14 (35.9)	12 (25.0)	
≥IIIb	25 (64.1)	36 (75.0)	
Tumor diameter (number%)			0.551
≤4cm	18 (46.2)	18 (37.5)	
>4cm	21 (53.8)	30 (62.5)	
hemoglobin (mean (SD))	106.51 (22.71)	110.29 (21.82)	0.432
NLR (mean (SD))	3.01 (1.82)	3.15 (1.66)	0.715
PLR (mean (SD))	187.85 (116.10)	186.15 (82.97)	0.937
Lymph node metastasis number%)			0.722
No	22 (56.4)	30 (62.5)	
Yes	17 (43.6)	18 (37.5)	
Total radiotherapy time (%)			0.659
TR≤8 week	19 (48.7)	20 (41.7)	
TR > 8 week	20 (51.3)	28 (58.3)	
Number of concurrent chemotherapy courses (times)	3.03 (1.77)	3.60 (1.54)	0.107
Number of adjuvant chemotherapy courses (times)	1.38 (1.55)	1.38 (1.57)	0.977

Table 4 Analysis of clinical factors affecting OS of radiotherapy and chemotherapy for intermediate and advanced cervical squamous cell carcinoma

Clinical factors	cases	Univariate analysis		Multivariate analysis	
		HR (95% CI)	P value	HR (95% CI)	P value
group			<0.001		<0.001
Chemoradiotherapy sensitive group	44	ref		ref	
Chemoradiotherapy resistance group	43	4.44 (2.09-9.43)		5.67 (2.37-13.56)	
ASPH			0.0448		0.2602
low	39	ref		ref	
high	48	2.06 (1.02-4.18)		1.60 (0.71-3.62)	
age(year)	87	0.99 (0.95-1.03)	0.564	0.97 (0.92-1.03)	0.3411
Stage			0.243		0.3155
IIb-IIIa	26	ref		ref	
≥IIIb	61	1.60 (0.73-3.53)		1.60 (0.64-4.04)	
Tumor diameter			0.0161		0.0735
≤4cm	36	ref		ref	
>4cm	51	2.33 (1.12-4.81)	0.0229	3.22 (1.29-8.04)	0.0124
hemoglobin	87	1.00 (0.98-1.01)	0.776	1.01 (0.99-1.03)	0.2566
NLR	87	1.27 (1.06-1.52)	0.0086	1.02 (0.79-1.31)	0.9067
PLR	87	1.00 (1.00-1.00)	0.203	1.00 (1.00-1.00)	0.8209
Lymph node metastasis			0.018		0.3182
No	52	ref		ref	
Yes	35	2.20 (1.15-4.22)		1.51 (0.67-3.38)	
Total duration of radiotherapy (month)					
≤8	39	ref		ref	
> 8	48	2.06 (1.03-4.12)	0.0411	2.82 (1.24-6.40)	0.0132

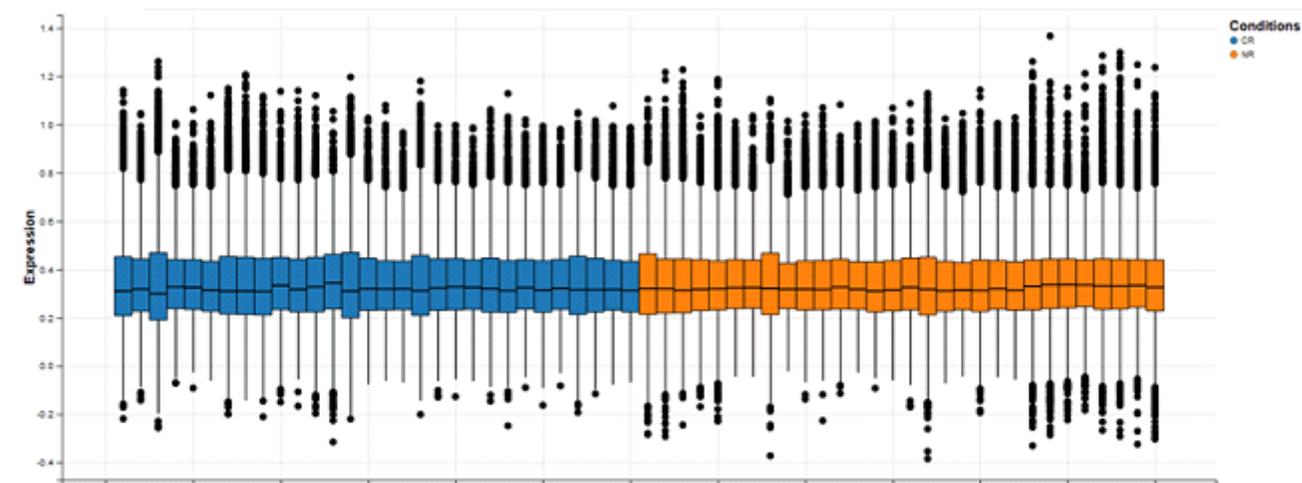
Number of concurrent chemotherapy courses (weeks)	87	0.94 (0.79-1.12)	0.469	0.82 (0.64-1.04)	0.1052
Adjuvant chemotherapy course (weeks)	87	0.99 (0.80-1.23)	0.946	0.97 (0.74-1.26)	0.7997

Table 5 Analysis of clinical factors affecting PFS of radiotherapy and chemotherapy in intermediate and advanced cervical squamous cell carcinoma

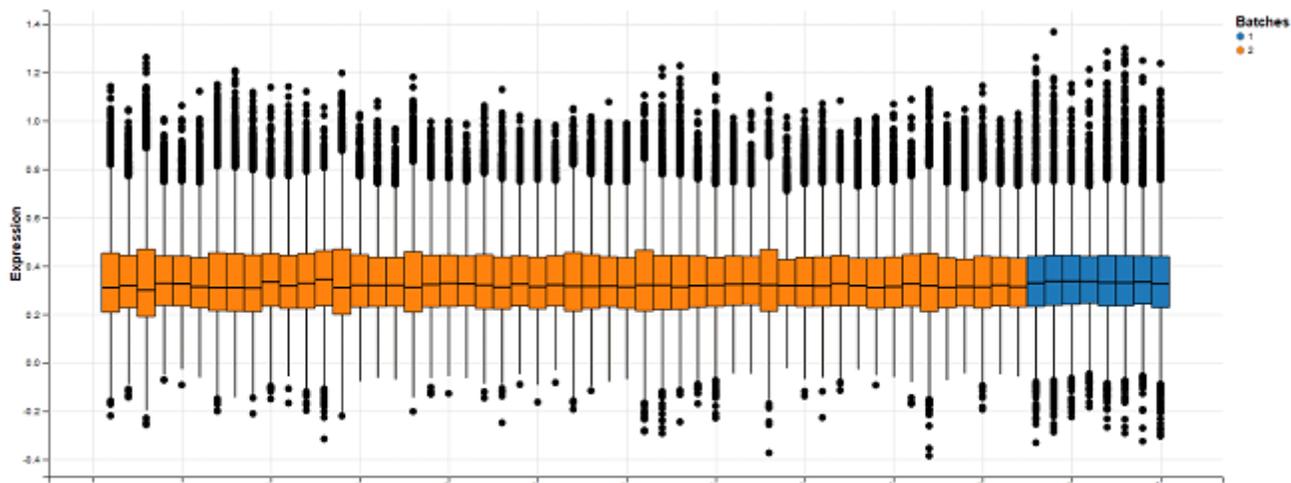
Clinical factors	Cases	Univariate analysis		Multivariate analysis	
		HR (95% CI)	P value	HR (95% CI)	P value
group			<0.001		<0.001
Chemoradiotherapy sensitive group	44	ref		ref	
Chemoradiotherapy resistance	43	4.80 (2.26-10.20)		6.10 (2.58-14.41)	
ASPH			0.0369		0.3326
low	39	ref		ref	
high	48	2.12 (1.05-4.30)		1.50 (0.66-3.39)	
Age	87	0.99 (0.95-1.03)	0.536	0.97 (0.92-1.03)	0.2989
Stage					
IIb-IIIa	26	ref		ref	
≥IIIb	61	1.65 (0.75-3.62)	0.216	1.60 (0.63-4.04)	0.3250
Tumor diameter					
≤4cm	36	ref		ref	
>4cm	51	2.35 (1.14-4.87)	0.0213	3.33 (1.31-8.45)	0.0115
hemoglobin	87	1.00 (0.98-1.01)	0.758	1.01 (0.99-1.03)	0.2463
NLR	87	1.25 (1.05-1.50)	0.0132	0.99 (0.78-1.27)	0.9595
PLR	87	1.00 (1.00-1.01)	0.186	1.00 (1.00-1.01)	0.7840
Lymph node metastasis			0.0194		0.3037
No	52	ref		ref	
Yes	35	2.18 (1.13-4.18)		1.52 (0.69-3.37)	
Total duration of radiotherapy (month)					
≤8	39	ref		ref	
> 8	48	2.04 (1.02-4.08)	0.0433	2.54 (1.14-5.66)	0.0231

Number of concurrent chemotherapy courses (weeks)	87	0.93 (0.78-1.12)	0.453	0.80 (0.63-1.02)	0.0667
Adjuvant chemotherapy course (weeks)	87	1.01 (0.81-1.25)	0.956	0.99 (0.76-1.28)	0.9193

## Figures



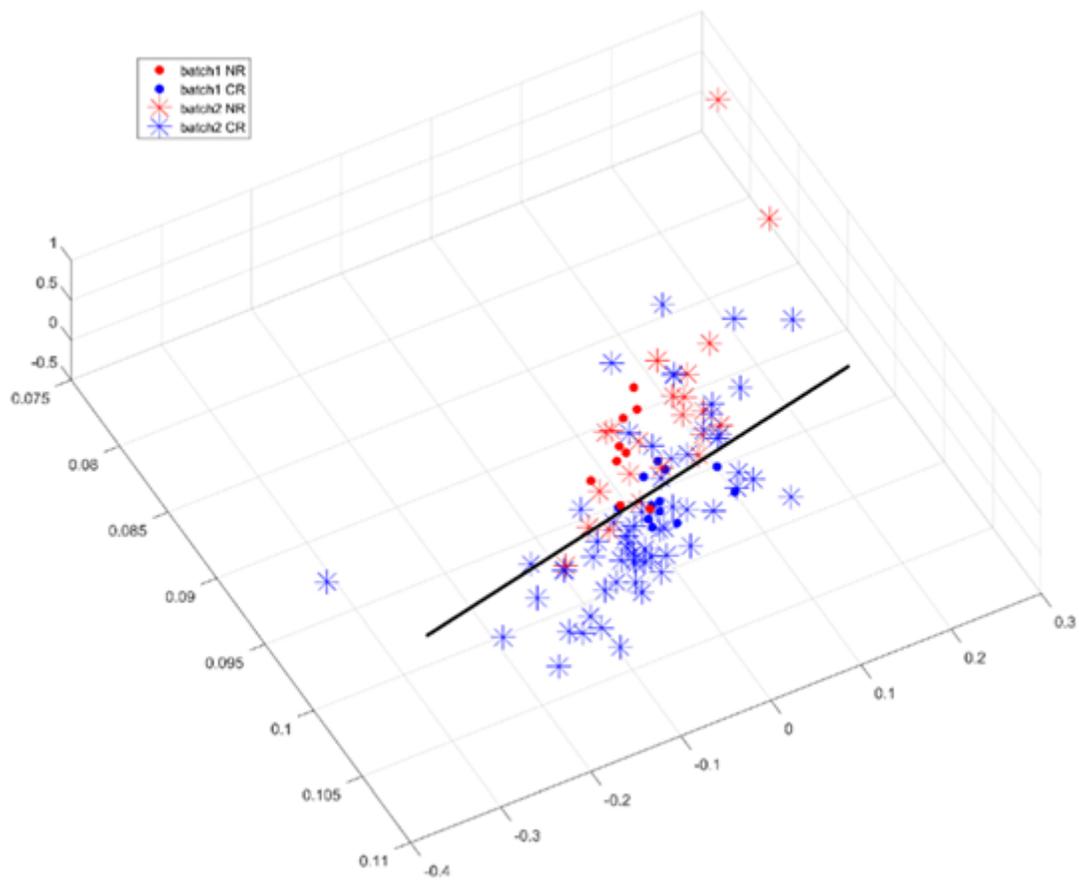
a



b

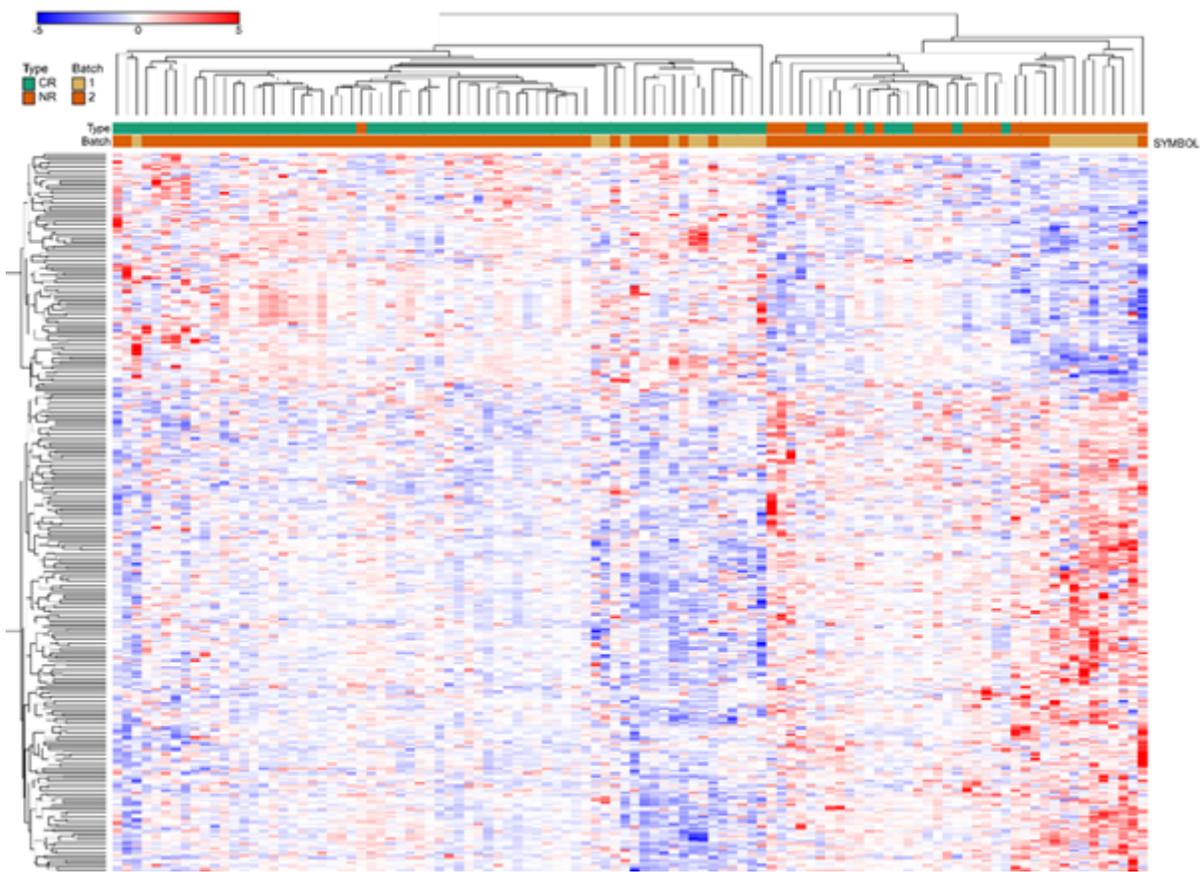
Figure 1

The whole expression level of each sample after removing the batch effect



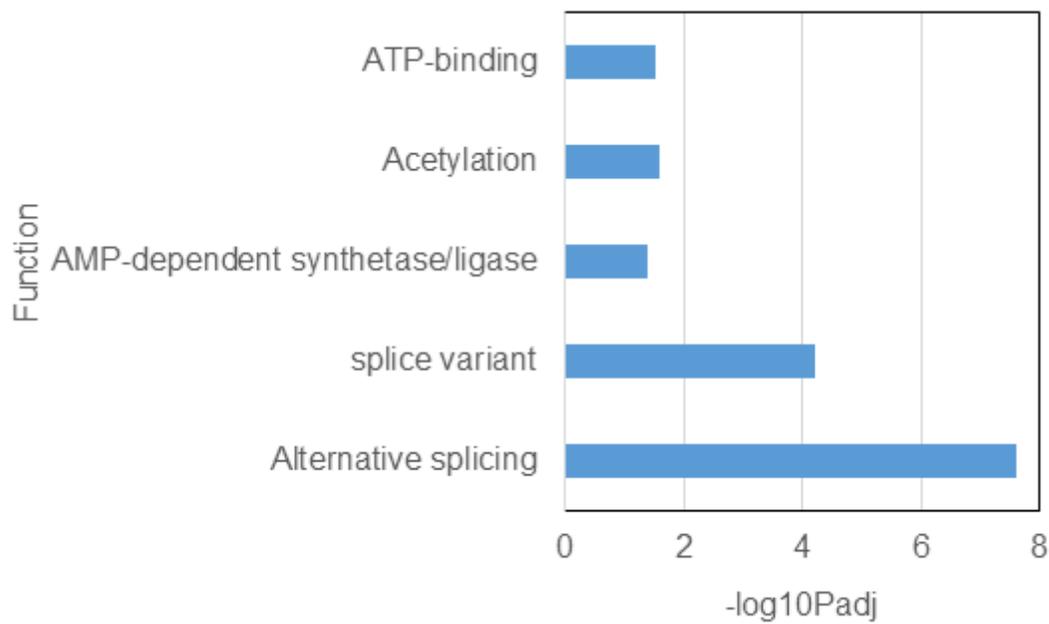
**Figure 2**

Results of principal component analysis



**Figure 3**

hierarchical clustering analysis of 298 differentially expressed genes



**Figure 4**

The gene annotation in DEGs

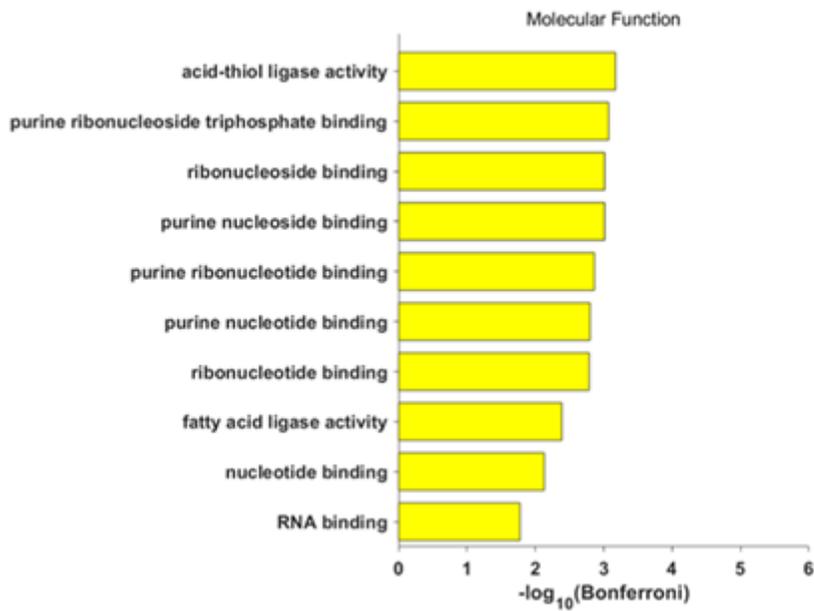


Figure 5

The cluster result of molecular function in DEGs

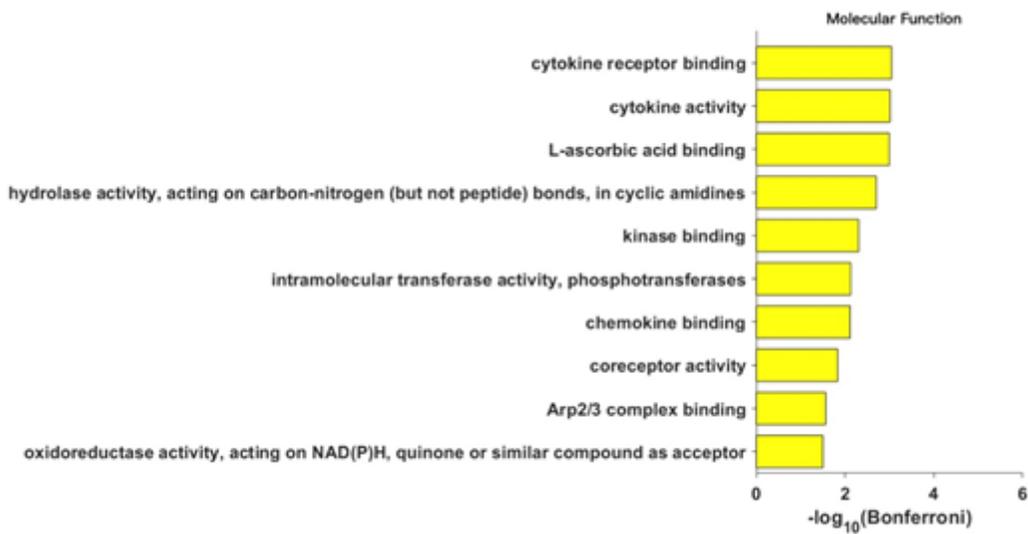


Figure 6

The cluster result of molecular function in prognosis genes

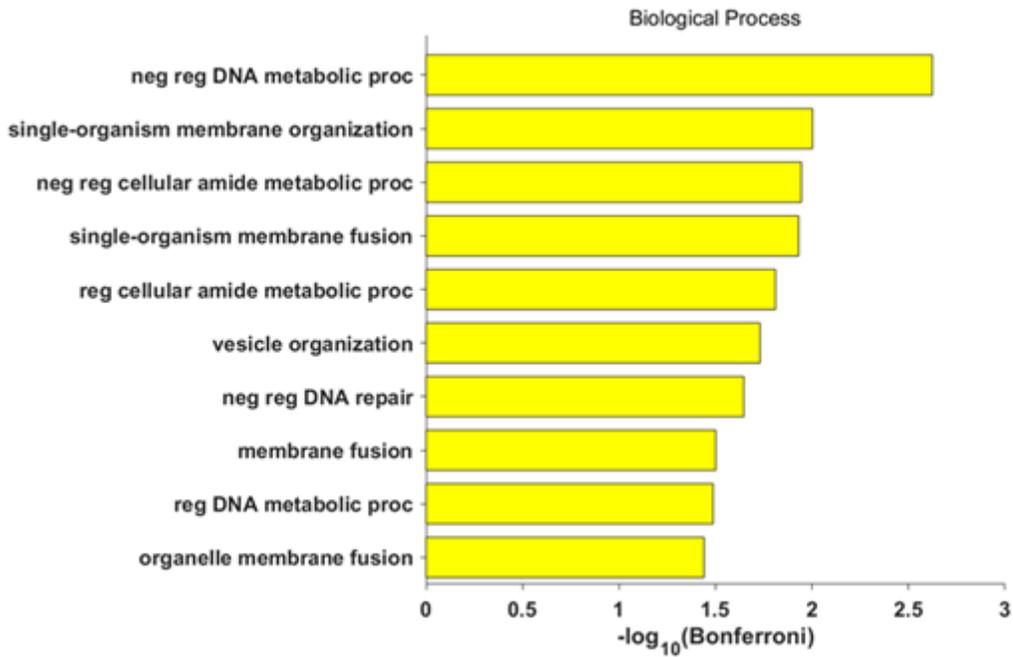
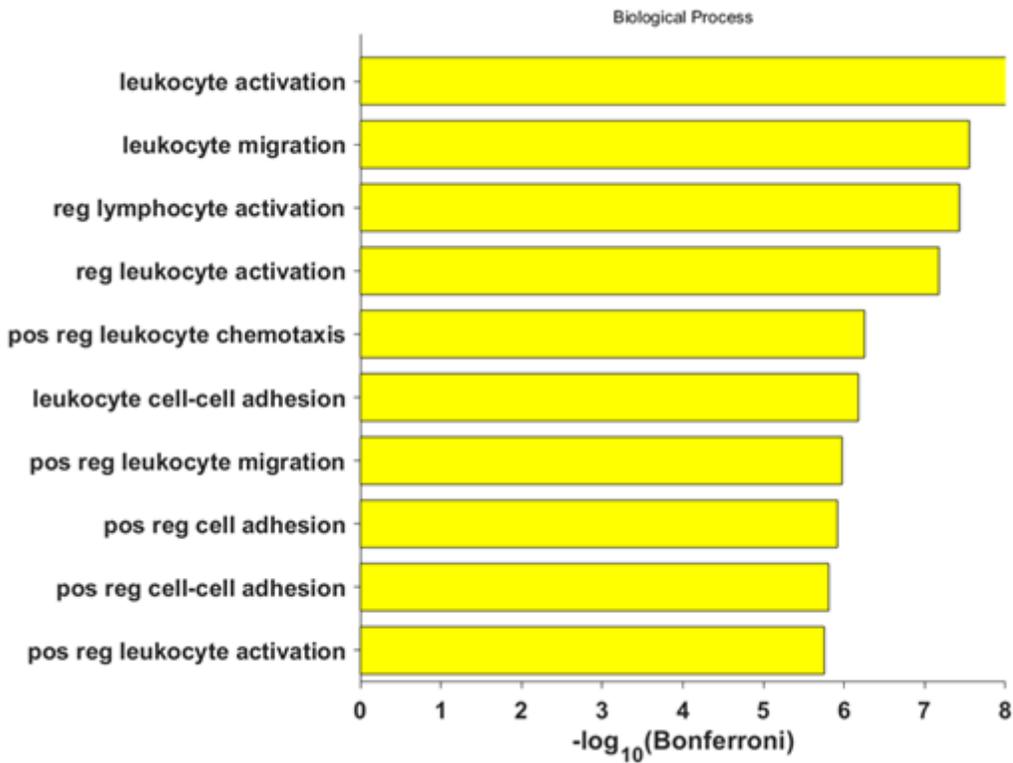


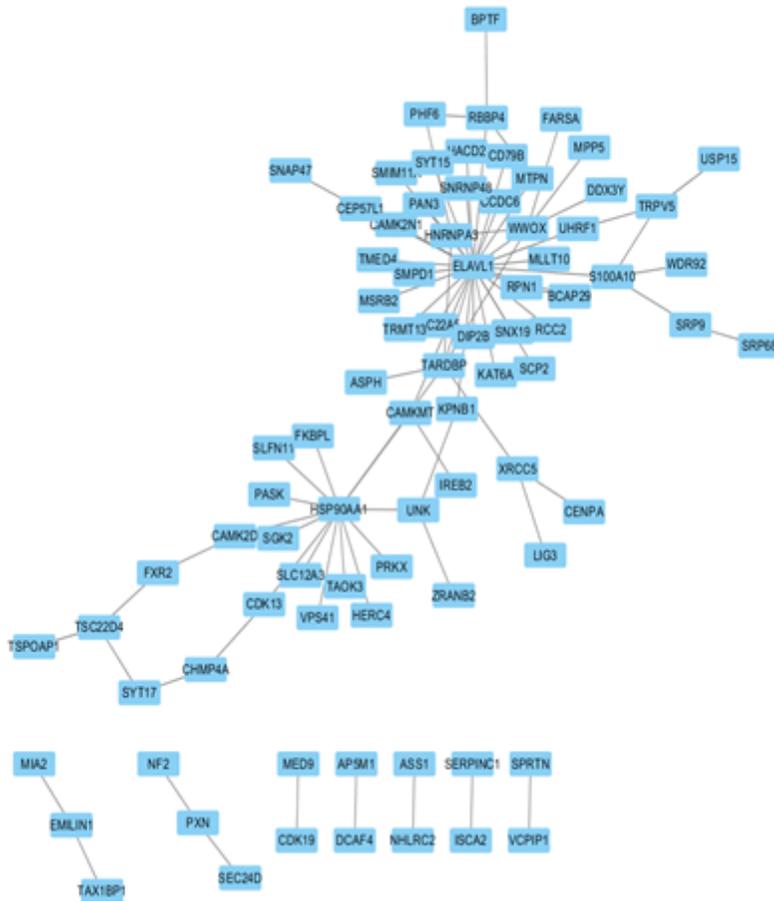
Figure 7

The cluster result of biological process in DEGs



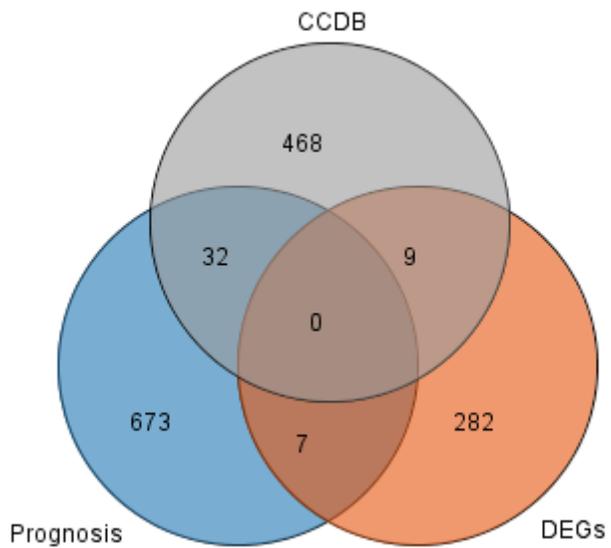
**Figure 8**

The cluster result of biological process in prognosis genes



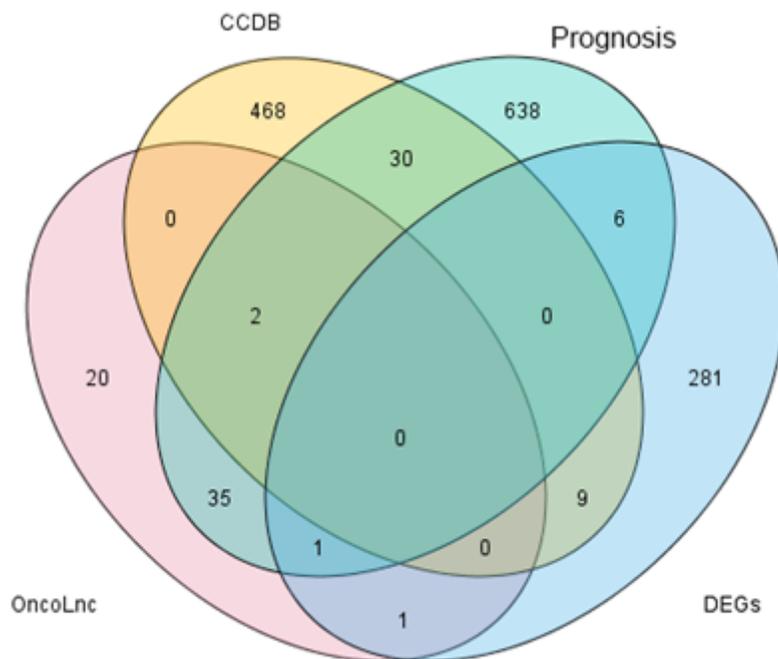
**Figure 9**

Protein interactions between DEGs



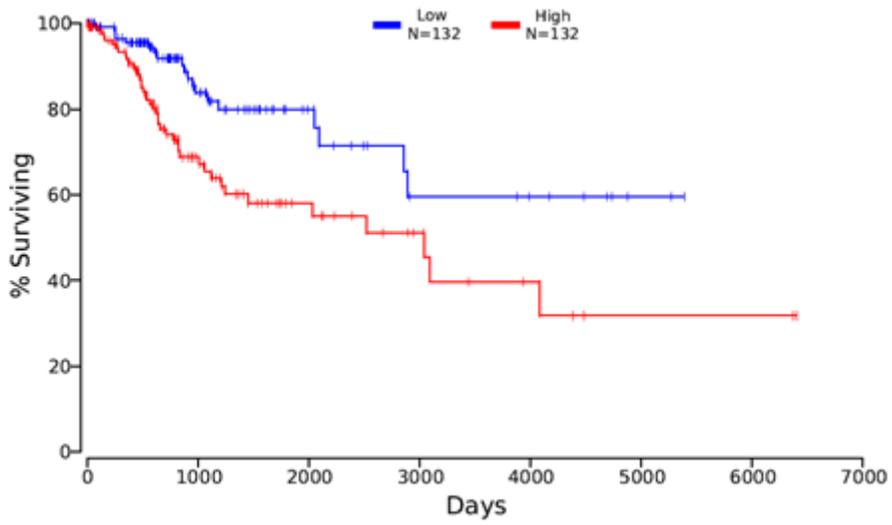
**Figure 10**

Venn diagram of three gene database sources



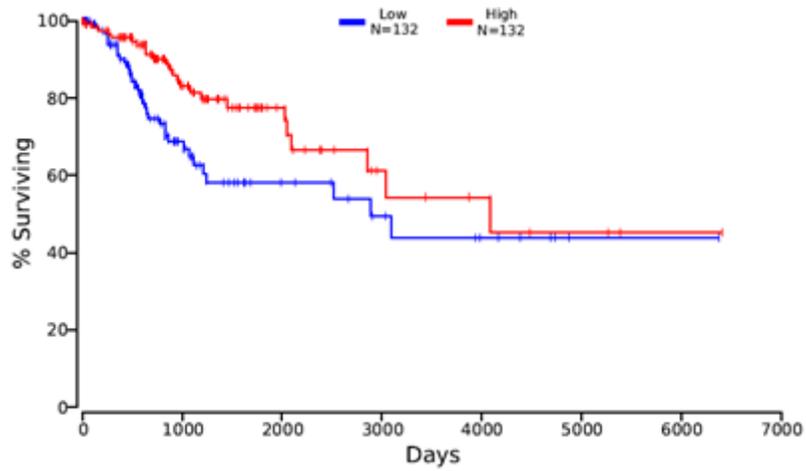
**Figure 11**

Venn diagram of three gene database sources with OncoLnc



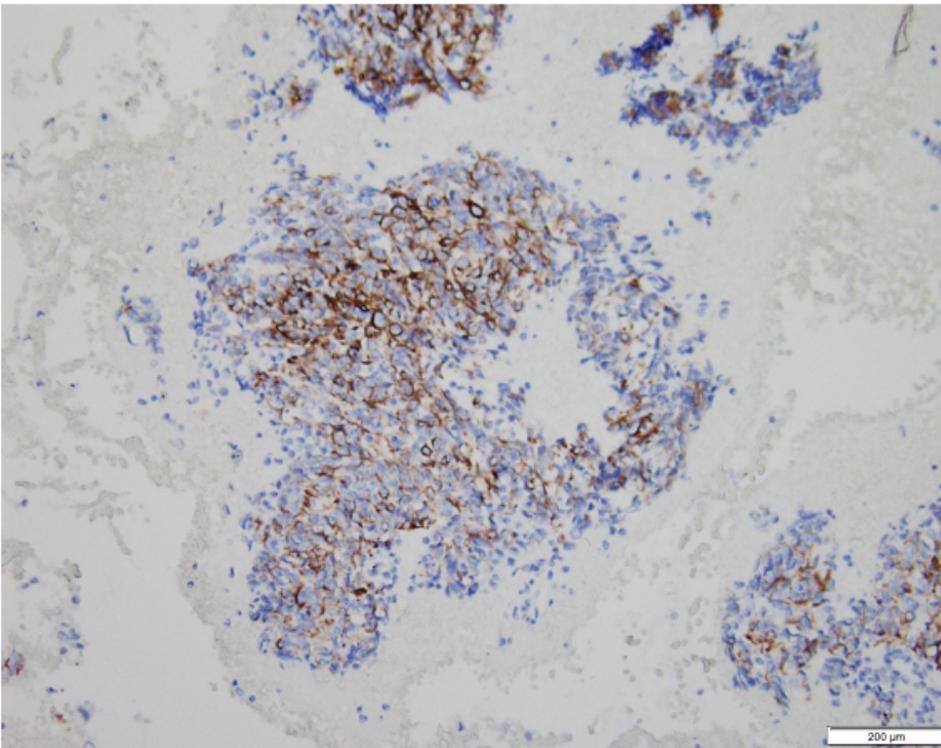
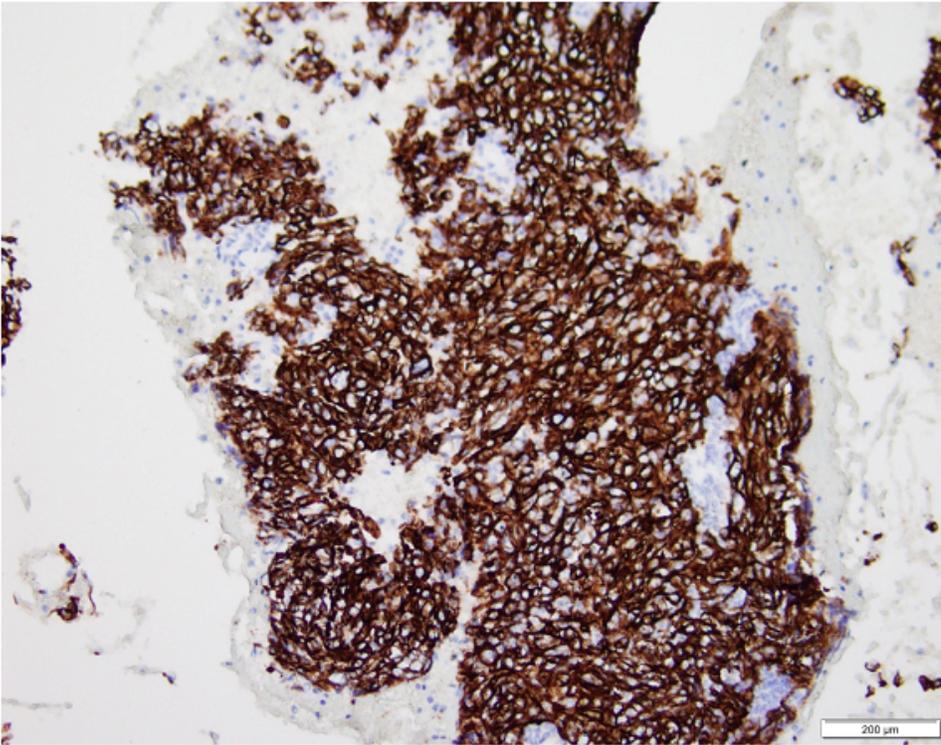
**Figure 12**

the Kaplan-Meier plot of ASPH in OncoLnc



**Figure 13**

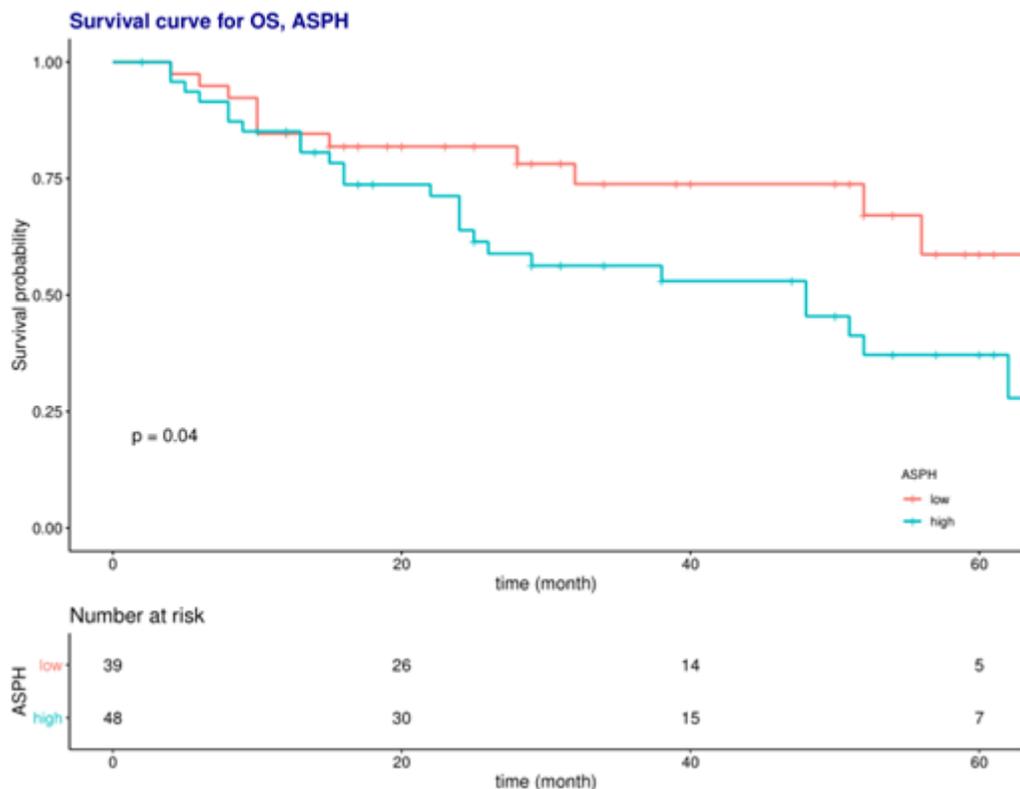
the Kaplan-Meier plot of NKAPP1 in OncoLnc



**Figure 14**

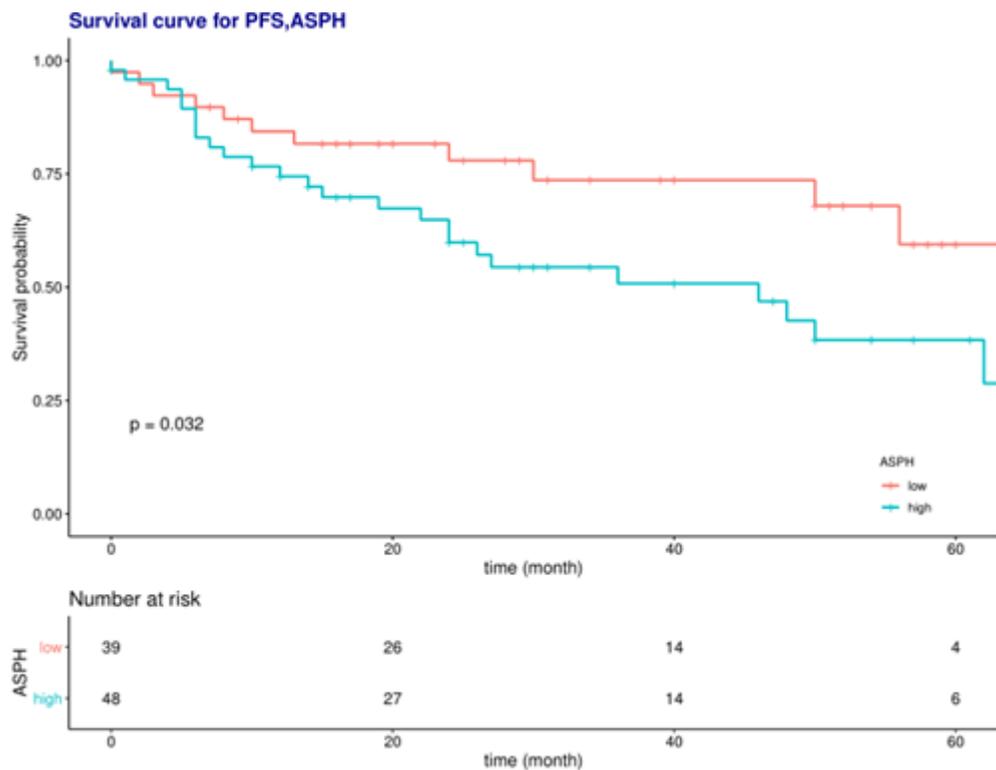
ASPH expression in cervical cancer tissues

Figure A is high expression results. Figure B is low expression results.



**Figure 15**

Survival curves of overall survival in low ASPH expression and high ASPH expression



**Figure 16**

Survival curves of progression-free survival in low ASPH expression and high ASPH