

Research Article

A Five-Genes-Based Prognostic Signature for Cervical Cancer Overall Survival Prediction

Menghuang Zhao , Wenbin Huang , Shuangwei Zou , Qi Shen , and Xueqiong Zhu 

Department of Obstetrics and Gynecology, The Second Affiliated Hospital of Wenzhou Medical University, Wenzhou 325027, China

Correspondence should be addressed to Xueqiong Zhu; zjwzzxq@163.com

Received 19 October 2019; Revised 8 January 2020; Accepted 21 February 2020; Published 26 March 2020

Academic Editor: Jacques Camonis

Copyright © 2020 Menghuang Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aims. This study is aimed at identifying a prognostic signature for cervical cancer. **Main Methods.** The gene expression data and clinical information of cervical cancer and normal cervical tissues were acquired from The Cancer Genome Atlas and from three datasets of the Gene Expression Omnibus database. DESeq2 and Limma were employed to screen differentially expressed genes (DEGs). The overlapping DEGs among all datasets were considered the final DEGs. Then, the functional enrichment analysis was performed. Moreover, the Cox proportional hazards regression was performed to establish a prognostic signature of the DEGs. The Kaplan-Meier analysis was applied to test the model. Relationships between gene expression and clinicopathological parameters in cervical cancer, including age, HPV status, histology, stage, and lymph node metastasis, were analysed by the chi-square test. The somatic mutations of these prognostic genes were assessed through cBioPortal. The robustness of the model was verified in another two independent validation cohorts. **Key Findings.** In total, 169 overlapping upregulated genes and 29 overlapping downregulated genes were identified in cervical cancer compared with normal cervical tissues. Functional enrichment analysis indicated that the DEGs were mainly enriched in DNA replication, the cell cycle, and the p53 signalling pathway. Finally, a 5-gene- (ITM2A, DSG2, SPP1, EFNA1, and MMP1) based prognostic signature was built. According to this model, each patient was given a prognostic-related risk value. The Kaplan-Meier analysis showed that a higher risk was related to worse overall survival in cervical cancer, with an area under the receiver operating characteristic curve of 0.811 for 15 years. The validity of this model in the prediction of cervical cancer outcome was verified in another two independent datasets. In addition, our study also found that the low expression of ITM2A was associated with cervical adenocarcinoma. Interestingly, DSG2 was associated with the HPV status of cervical cancer. **Significance.** Our study constructed a prognostic model in cervical cancer and discovered two novel genes, ITM2A and DSG2, associated with cervical carcinogenesis and survival.

1. Introduction

Cervical cancer ranks as the fourth most common cancer in women worldwide [1]. Even in developed countries such as the United States, there were approximately 13 170 newly diagnosed cases of cervical cancer and 4 250 deaths in 2019 [2]. Although some advances have been made in the screening technology for cervical cancer and human papillomavirus vaccine applications, the survival rate remains poor [3]. Another critical problem with cervical cancer is the heterogeneity in clinical outcomes among individuals with the same disease [4]. Consequently, there is a great need to establish

a molecular tool to help predict patient outcomes and outline individualized treatment plans.

With the impressive progress of the next-generation sequencing technology, various genetic alterations have been revealed, including the high-frequency mutations of genes and dysregulated signalling pathways in cancer [5]. Recently, numerous studies have shown the clinical importance of the messenger RNA expression in numerous cancers, including cervical cancer [6]. Messenger RNAs play pivotal roles in *diverse* physiological and pathological processes, such as differentiation, cell proliferation, development, apoptosis, and stress responses [7].

Transcriptome sequencing technologies (RNA-Seq) and microarray data provide an ideal platform for cancer genetic studies. The Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) repositories offer abundant cervical cancer sample resources, which may be helpful to explore reliable biomarkers. In the current study, gene expression in cervical cancer was profiled to identify critical genes predictive of patient outcomes. Through the differentially expressed gene (DEG) analysis and the Cox proportional hazards model, a signature based on the expression of several key genes was finally constructed as a prognostic signature for cervical cancer. This prognostic model could be used as an effective tool to predict the prognosis of patients with cervical cancer. The findings can also aid in identifying new therapeutic targets for cervical cancer.

2. Materials and Methods

2.1. Data Source. A high-throughput sequencing dataset and the clinical information of cervical cancer patients were downloaded from the TCGA website for the screening of prognostic signatures. Three microarray datasets (GSE6791 [8], GSE7803 [9], and GSE9750 [10]) were obtained from the GEO database. In total, 423 samples, including 378 cervical cancer samples and 45 normal cervical tissue samples, were collected in the present study.

2.2. Identification of DEGs. Raw count data were extracted from TCGA. The DESeq2 package [11] for R was used to normalize and identify DEGs. DEGs in cervical cancer samples compared with adjacent normal cervical tissues were obtained with the threshold of adjusted P value < 0.05 and absolute \log_2 -based fold change > 1 . The Limma package [12] for R was used for the microarray data to identify DEGs. The thresholds of absolute \log_2 -based fold change > 1 and false discovery rate (FDR) < 0.05 were employed to screen out DEGs. The overlapping DEGs among all datasets were considered the final DEGs.

2.3. Functional Enrichment. The Gene Ontology (GO) resource provides scientific knowledge about the functions of genes from many different organisms [13, 14]. The GO enrichment analysis was performed by employing the online software Database for Annotation, Visualization, and Integration Discovery (DAVID, version 6.8) [15]. GO terms were composed of three categories, molecular function (MF), cellular component (CC), and biological process (BP). Pathway enrichment was analysed by The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database [16], which consists of a collection of biochemical pathway diagrams of reaction networks and molecular interactions. $P < 0.05$ was considered a significant difference.

2.4. Prognostic Signature Construction and Evaluation. Expression profiles with fragments per kilobases per million values were downloaded from TCGA, and the \log_2 transformation was performed. The prognostic value of each DEG was analysed by the univariate Cox proportional hazards regression analysis using the survival package [17] for R. DEGs that were significantly related to overall survival

($P < 0.05$) were retained for further analysis. Next, the multivariable Cox proportional hazards stepwise regression with backward selection in SPSS (version 18.0) was used to establish a prognostic model. A prognostic signature was built based on the linear weighted combination of the gene expression values. The weight of each gene was the regression coefficient (β) identified from the multivariate Cox proportional hazards regression analysis. The prognostic risk score for predicting overall survival (OS) was as follows: Risk score = $\sum \beta_i \times \text{expGene}_i$. According to the median risk score, patients were divided into low- and high-risk groups. Then, the prognostic signature model was examined by the Kaplan-Meier analysis. To assess the predictive accuracy of the prognostic signature for cervical cancer outcome, a receiver operating characteristic (ROC) curve was plotted by the survivalROC package for R.

2.5. Analysis of Clinicopathological Parameters in Cervical Cancer. The expression levels of the prognostic genes and clinicopathological parameters of cervical cancer patients, including age, HPV status, histology, stage, and lymph node metastasis, were downloaded from TCGA. Only samples with relevant clinical information were included. The patients were classified into low and high expression groups according to the median gene expression value. The relationship between gene expression and clinicopathological parameters in cervical cancer was analysed by the chi-square test. $P < 0.05$ was considered to indicate a statistically significant difference.

2.6. Prognostic Gene Mutations. The somatic mutations of the prognostic genes in cervical cancer patients, including missense mutations, truncation mutations, fusions, amplifications, and deep deletions, were assessed through cBioPortal [18].

2.7. Validation of the Robustness of the Prognostic Signature. The robustness of the prognostic signature for predicting survival in cervical cancer patients was evaluated in the validation cohorts GSE52903 and GSE39001 [19, 20]. The differential expression levels of the corresponding genes were compared by Student's t -test. The Kaplan-Meier analysis was applied to confirm the prognostic signature of the prediction model.

2.8. Prognostic Values of Clinical Parameters and Risk Score in Cervical Cancer. Furthermore, patients with all clinical information, including age, HPV status, histology, stage, and lymph node metastasis, were collected from TCGA. Then the univariate Cox proportional hazards regression and multivariable Cox proportional hazards regression were used to analyse the prognostic values of clinical parameters and risk score in cervical cancer. $P < 0.05$ was considered to indicate a statistically significant difference.

3. Results

3.1. Identification of Differentially Expressed Genes between Cervical Cancer and Normal Cervical Tissue. In the TCGA dataset, compared with normal cervical samples, 3 769 genes

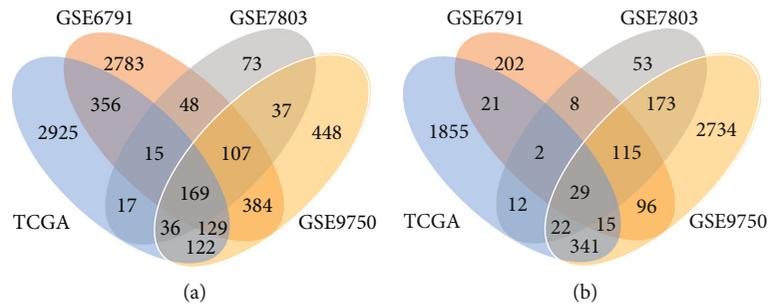


FIGURE 1: Venn diagrams showing the overlapping DEGs in the four datasets. (a) Upregulated DEGs. (b) Downregulated DEGs.

were identified to be significantly upregulated, while 2 297 genes were discovered to be significantly downregulated in cervical cancer. For the GSE6791 dataset, the differential expression analysis between cervical cancer and normal cervical samples identified 3 991 upregulated DEGs and 488 downregulated DEGs. For the GSE7803 and GSE9750 datasets, 502 upregulated and 414 downregulated DEGs and 1 432 upregulated and 3 525 downregulated DEGs were found in cervical cancer, respectively. Altogether, compared with the normal cervical samples, 169 upregulated DEGs and 29 downregulated DEGs in cervical cancer overlapped among the three microarray datasets and the TCGA dataset (Figure 1).

3.2. Functional Enrichment in Cervical Cancer. GO enrichment showed that the DEGs were primarily related to biological process functions, which included DNA replication ($P = 2.23 \times 10^{-27}$), cell division ($P = 1.12 \times 10^{-26}$), and G1/S transition of the mitotic cell cycle ($P = 2.05 \times 10^{-22}$). The GO results of the cellular compartment category revealed that the DEGs were mainly enriched in the nucleoplasm ($P = 1.04 \times 10^{-20}$), nucleus ($P = 6.42 \times 10^{-17}$), and midbody ($P = 8.89 \times 10^{-12}$). In terms of molecular functions, the DEGs were involved in protein binding ($P = 8.21 \times 10^{-14}$), ATP binding ($P = 1.54 \times 10^{-8}$), and DNA binding ($P = 1.57 \times 10^{-8}$). The top 10 GO terms are shown in Figure 2.

The KEGG pathway enrichment showed that the DEGs were significantly enriched in the DNA replication, cell cycle, and p53 signalling pathways (Figure 3).

3.3. Prognostic Signature Construction and Evaluation. The DEGs were further analysed to identify survival-related genes of cervical cancer through the univariate Cox proportional hazards analysis. The analysis identified 8 upregulated DEGs that were significantly related to poor overall survival in the TCGA cohort and 3 downregulated DEGs that were related to good overall survival ($P < 0.05$). In total, 11 survival-related DEGs were found.

Then, the multivariable Cox proportional hazards analysis was performed to build a prognostic model. Five genes, ITM2A, DSG2, SPP1, EFNA1, and MMP1, which showed a significant P value, were selected (Table 1). A linear model was calculated with the risk score = $(-0.308 \times \text{ITM2A}) + (0.341 \times \text{DSG2}) + (0.172 \times \text{SPP1}) + (0.392 \times \text{EFNA1}) + (0.116 \times \text{MMP1})$. The risk score was calculated for each patient in the TCGA dataset. According to the median risk

score, the patients were divided into high- and low-risk groups. The Kaplan-Meier analysis was conducted and showed that the high-risk patient group had a markedly shorter overall survival than the low-risk patient group (Figure 4(a)). The distributions of the risk score, survival status, and gene expression levels of the individual patients are shown in Figure 4(b). Furthermore, the ROC analysis was performed to evaluate the predictive accuracy of the prognostic model. The area under the curve (AUC) for the 10-year prediction was 0.762. The AUC for the 15-year prediction was up to 0.811 (Figure 4(c)).

3.4. Analysis of Clinicopathological Parameters in Cervical Cancer. The associations between the expression of the 5 prognostic genes and the clinicopathological parameters of patients with cervical cancer were determined. As presented in Table 2, ITM2A expression was associated with the histology of cervical cancer. Cervical cancer patients with a low expression of ITM2A were more likely to have adenocarcinoma or adenosquamous carcinoma, while cervical cancer patients with a high expression of ITM2A tended to have squamous carcinoma. However, ITM2A exhibited no association with age, HPV status, stage, or lymph node metastasis in cervical cancer.

The relationships between DSG2 expression and different clinicopathological parameters in cervical cancer are demonstrated in Table 3. Interestingly, DSG2 expression was related to HPV status. All cervical cancer patients with a high expression of DSG2 were HPV positive. However, approximately 9.1% of patients with a low expression of DSG2 were HPV negative. DSG2 expression was not correlated with other clinicopathological parameters in cervical cancer patients.

Tables 4 and 5 show that EFNA1 and MMP1 exhibited no association with any clinicopathological parameters in cervical cancer patients.

Table 6 shows the results of SPP1 in cervical cancer. SPP1 was found to be related to the age and histology of cervical cancer patients. Patients with high SPP1 expression were older than those with low SPP1 expression. The low expression group tended to have adenocarcinoma-type cervical cancer.

3.5. Prognostic Gene Mutations. The somatic mutations of the five prognostic genes were further explored in this study (Figure 5). In cervical cancer, 10% of patients showed genetic

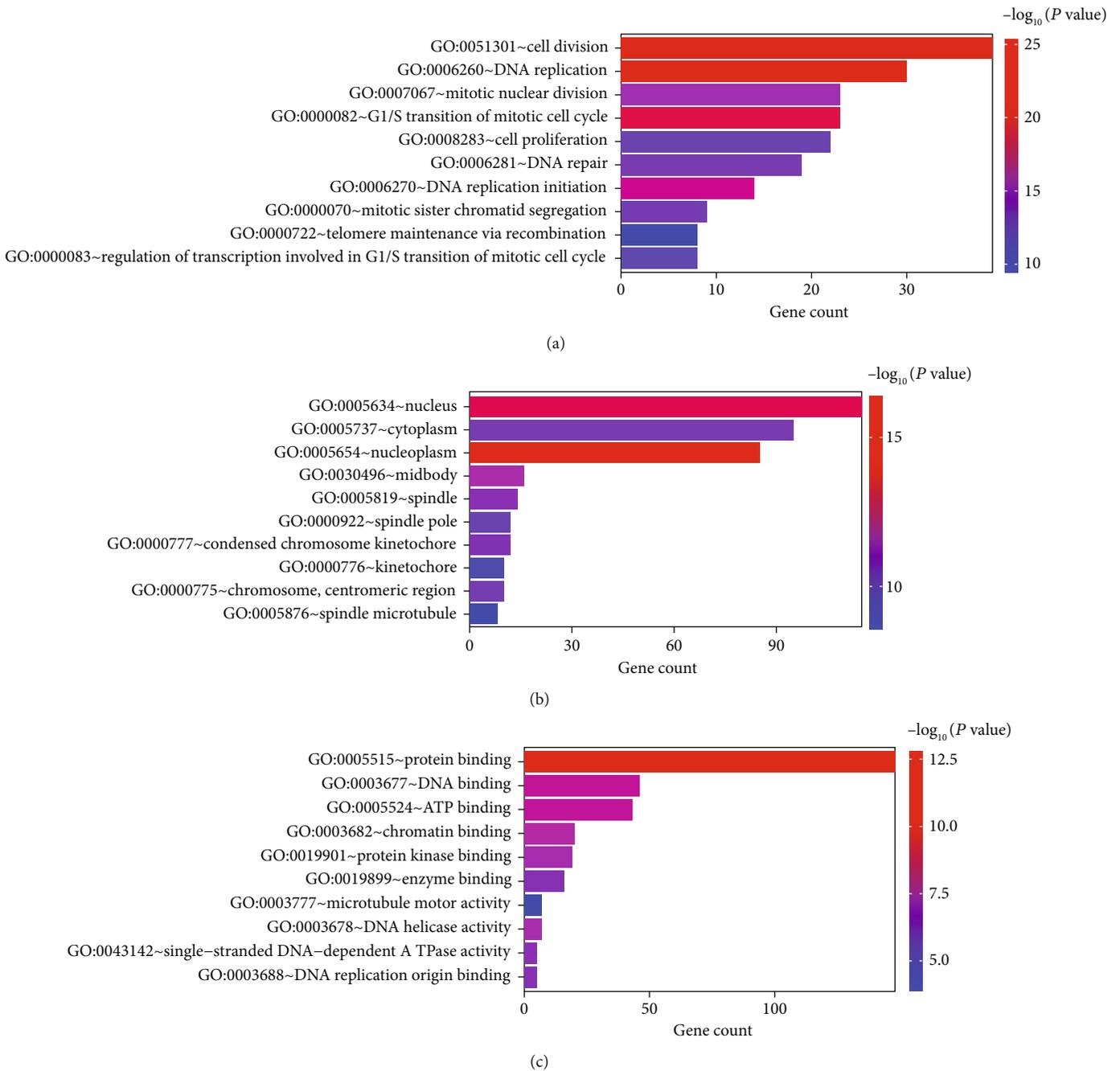


FIGURE 2: The results of the GO enrichment analysis of DEGs. (a) Biological processes. (b) Cellular components. (c) Molecular functions.

alterations in MMP1. EFNA1 mutation occurred in 2.5% of cervical cancer patients. Amplification was the most common mutation in these two genes. The mutation rates of DSG2, SPP1, and ITM2A were 2.2%, 1.8%, and 1.8%, respectively. Deletion could be seen in ITM2A. These mutations might contribute to the aberrant expression of the corresponding genes.

3.6. Validation of the Robustness of the Prognostic Signature. The robustness of the prognostic signature for predicting survival in cervical cancer patients was evaluated in the validation cohorts GSE52903 and GSE39001. The GSE52903

dataset was composed of 17 normal cervical tissues and 55 cervical cancer samples. First, the differential expression level of the five prognostic biomarkers between normal cervical tissue and cervical cancer samples was tested. ITM2A was downregulated in cervical cancer compared with normal cervical samples, and the four other genes (DSG2, SPP1, EFNA1, and MMP1) were markedly upregulated in cervical cancer, which was consistent with the results above (Figures 6(a)–6(e)). Then, the risk scores of individual cervical cancer patients were calculated according to the prognostic linear model shown above. The patients were divided into high- and low-risk groups. The Kaplan-Meier results

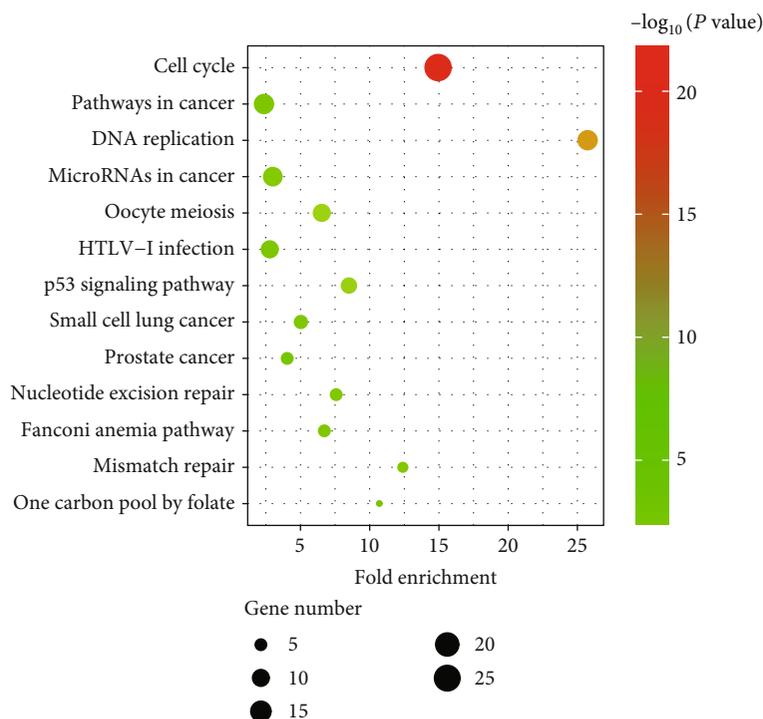


FIGURE 3: The results of the KEGG enrichment analysis of DEGs.

demonstrated that a high-risk score was associated with shorter overall survival, which confirmed the results discovered above (Figure 6(f)).

The GSE39001 dataset consisted of 12 normal cervical tissues and 43 cervical cancer samples. The above results were further verified in the GSE39001 dataset (Figure 7).

3.7. Prognostic Values of Clinical Parameters and Risk Score in Cervical Cancer. Furthermore, the prognostic values of clinical parameters of patients with cervical cancer were also analysed (Table 7). The univariate Cox proportional hazards regression analysis showed that lymph node metastasis and risk score were associated with overall survival. However, the multivariable Cox proportional hazards regression revealed that only the risk score was an independent prognostic indicator of cervical cancer survival.

4. Discussion

Cervical cancer is one of most lethal malignancies in women, causing a heavy burden to families and the world [21]. The accurate prediction of cervical cancer outcomes is of great importance for therapy method selection and prognosis improvement. In the present study, four independent datasets were used to identify DEGs between cervical cancer and normal cervical tissue. In total, 198 DEGs, including 29 downregulated and 169 upregulated genes, were identified. To better understand the mechanism of cervical tumourigenesis, enrichment analysis of the DEGs was performed. In terms of GO enrichment, the DEGs were mainly enriched in DNA replication, cell division, and G1/S transition of the mitotic cell cycle. The KEGG pathway results showed that

TABLE 1: The multivariate Cox regression analysis of overall survival.

Gene symbol	Coefficient	Hazard ratio (HR)	P value
ITM2A	-0.308	0.735	0.042
DSG2	0.341	1.407	0.018
SPP1	0.172	1.187	0.011
EFNA1	0.392	1.479	0.012
MMP1	0.116	1.123	0.015

the DEGs were mainly involved in DNA replication and the cell cycle. These findings were consistent with the fact that DNA replication and cell cycle regulation play important roles in the development of cancer [22].

To identify survival-associated DEGs, univariate Cox proportional hazards regression was performed. In total, 11 survival-related DEGs were found. Then, by stepwise selection, 5 independent prognostic mRNAs (SPP1, EFNA1, MMP1, ITM2A, and DSG2) were selected and combined to construct a prognostic signature. The prognostic model based on these five genes showed high efficiency in distinguishing good versus poor outcomes for cervical cancer patients. According to the Hosmer-Lemeshow classification method, areas under the ROC curve of 0.80 to 0.90 are considered to be excellent [23]. The areas under the ROC curve of this model reached 0.811 for the 15-year prediction. The robustness of the prognostic signature was confirmed in two independent cohorts. In the clinic, this prognostic signature might serve as a predictive indicator of cervical cancer survival.

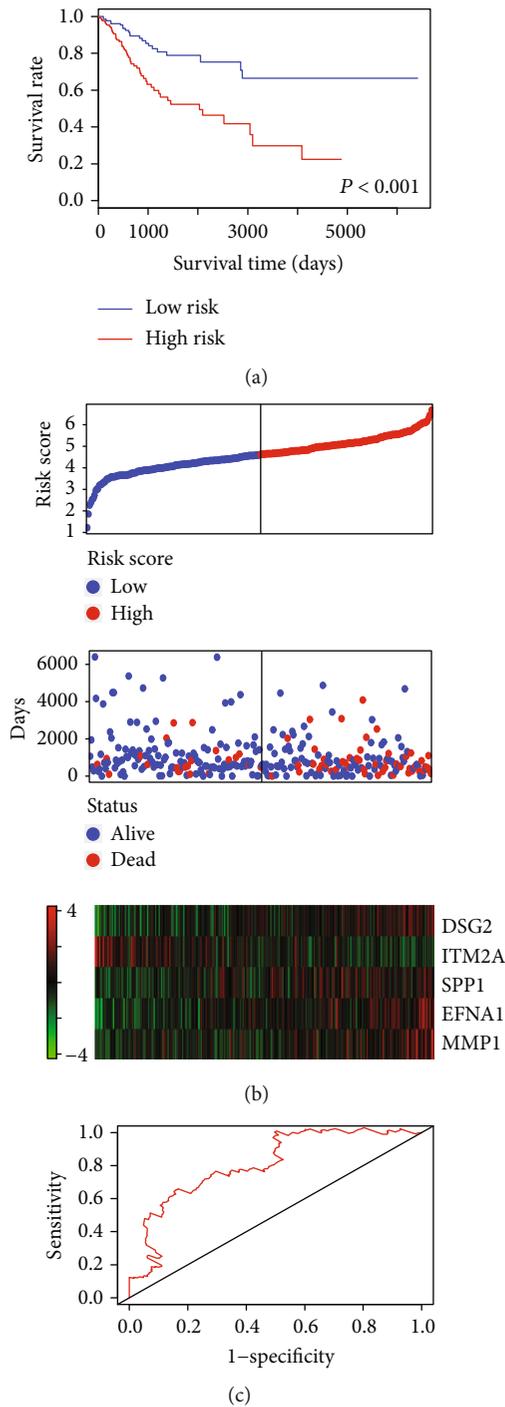


FIGURE 4: Prognostic evaluation of the 5-gene signature in cervical cancer patients. (a) The Kaplan-Meier risk survival curve analysis of overall survival in cervical cancer patients. (b) The distributions of the risk score, survival status, and gene expression levels of the patients. (c) ROC curve analysis of the 5-gene signature.

SPP1 is a secreted glycoprophosphoprotein of the SIBLING family that plays a critical role in physiological and pathophysiological processes [24]. Elevated expression of SPP1 has been observed in multiple cancers, including colon cancer, lung cancer, prostate cancer, breast cancer, ovarian cancer, multiple myeloma, acute myeloid leukaemia, and

chronic myeloid leukaemia [25–27]. SPP1 expression was also higher in cervical cancer [28]. In the study by Cho, SPP1 had a 50.6% sensitivity and 95.0% specificity as a diagnostic biomarker for cervical cancer [29]. Higher expression of SPP1 was strongly related to worse disease-free survival and overall survival in patients with cervical cancer [29]. The present findings were consistent with those of previous studies.

EFNA1, a glycosyl-phosphatidylinositol-anchored ligand, binds to its receptor EPHA2 to promote autophosphorylation [30, 31]. Autophosphorylation then triggers downstream signalling to regulate cell migration and growth [30]. EFNA1, a tumour-related gene, was significantly upregulated in numerous tumours, including cervical cancer [32–36]. EFNA1 expression was reported to be related to deep invasion, parametrial invasion, tumour size, and outcome in cervical carcinoma [36, 37]. Our study also observed the high expression of EFNA1 in cervical cancer, and its expression was related to poor overall survival.

Matrix metalloproteinase-1 (MMP1) functions as an enzyme in the degradation of the extracellular matrix in both normal physiological processes and disease processes [38]. Growing evidence has shown that MMP1 plays critical roles in tumourigenesis and cancer metastasis [39]. Some studies have observed the upregulation of MMP1 in cervical cancer [40]. Knockdown of MMP1 could inhibit cervical cancer cell invasion, migration, and proliferation via epithelial-mesenchymal transition [41]. MMP1 was correlated with lymph node metastasis and indicated unfavourable survival in cervical cancer [41]. The present study supported previous reports by showing that the high expression of MMP1 was significantly related to poor overall survival in cervical cancer.

ITM2A is a type II integral membrane protein containing a BRICHOS domain [42, 43]. Some studies supported that ITM2A was associated with cell differentiation, including chondrogenesis, odontogenesis, and myogenesis stages [44–46]. A previous study found that ITM2A was markedly downregulated in invasive ovarian carcinomas compared with normal ovarian tissues [47]. In addition, the loss of ITM2A was significantly correlated with higher FIGO stage, recurrence, chemoresistance, and poorer prognosis in ovarian cancer [47]. The expression of ITM2A induced G2/M cell cycle arrest and inhibited ovarian cancer cell growth by decreasing the expression of cyclin B1, p-CDC2, CDC2, and CDC25C [47]. All of the above results indicated that ITM2A might serve as a tumour suppressor in ovarian cancer. However, there has been no paper about the expression and prognostic value of ITM2A in cervical cancer. In this study, ITM2A was shown to be downregulated and related to good outcomes in cervical cancer. Different expression levels of ITM2A were also associated with the specific histological phenotype of cervical cancer. Cervical cancer patients with a low expression of ITM2A were more likely have adenocarcinoma or adenosquamous carcinoma, while cervical cancer patients with a high expression of ITM2A tended have squamous carcinoma. Compared with squamous cell carcinoma, cervical adenocarcinoma has a worse outcome [48, 49]. Our study found that the low expression of ITM2A was associated with poor prognosis and cervical

TABLE 2: Association between ITM2A expression and the clinicopathological parameters of cervical cancer patients.

Clinicopathological parameters		Low expression	High expression	P value
Age	<50 years	91	89	0.708
	>50 years	59	63	
HPV status	Negative	6	3	0.430
	Positive	80	89	
Histology	Adenocarcinoma	20	11	0.016*
	Adenosquamous	3	0	
	Squamous	63	81	
Stage	I-II	63	78	0.138
	III-IV	20	14	
Lymph node metastasis	Positive	15	18	0.572
	Negative	35	53	

* $P < 0.05$.

TABLE 3: Association between DSG2 expression and the clinicopathological parameters of cervical cancer patients.

Clinicopathological parameters		Low expression	High expression	P value
Age	<50 years	82	98	0.061
	>50 years	69	53	
HPV status	Negative	9	0	0.016*
	Positive	90	79	
Histology	Adenocarcinoma	18	13	0.373
	Adenosquamous	3	0	
	Squamous	78	66	
Stage	I-II	78	63	0.803
	III-IV	18	16	
Lymph node metastasis	Positive	15	18	0.054
	Negative	57	31	

* $P < 0.05$.

TABLE 4: Association between EFNA1 expression and the clinicopathological parameters of cervical cancer patients.

Clinicopathological parameters		Low expression	High expression	P value
Age	<50 years	88	92	0.543
	>50 years	64	58	
HPV status	Negative	7	2	0.171
	Positive	82	87	
Histology	Adenocarcinoma	15	16	0.331
	Adenosquamous	3	0	
	Squamous	71	73	
Stage	I-II	70	71	0.730
	III-IV	18	16	
Lymph node metastasis	Positive	18	15	0.738
	Negative	45	43	

TABLE 5: Association between MMP1 expression and the clinicopathological parameters of cervical cancer patients.

Clinicopathological parameters		Low expression	High expression	<i>P</i> value
Age	<50 years	83	97	0.075
	>50 years	69	53	
HPV status	Negative	6	3	0.391
	Positive	78	91	
Histology	Adenocarcinoma	18	13	0.375
	Adenosquamous	2	1	
	Squamous	64	80	
Stage	I-II	68	73	0.460
	III-IV	14	20	
Lymph node metastasis	Positive	12	21	0.095
	Negative	47	41	

TABLE 6: Association between SPP1 expression and the clinicopathological parameters of cervical cancer patients.

Clinicopathological parameters		Low expression	High expression	<i>P</i> value
Age	<50 years	101	79	0.010*
	>50 years	50	72	
HPV status	Negative	3	6	0.494
	Positive	86	83	
Histology	Adenocarcinoma	23	8	0.002*
	Adenosquamous	0	3	
	Squamous	66	78	
Stage	I-II	74	67	0.237
	III-IV	14	20	
Lymph node metastasis	Positive	13	20	0.17
	Negative	47	41	

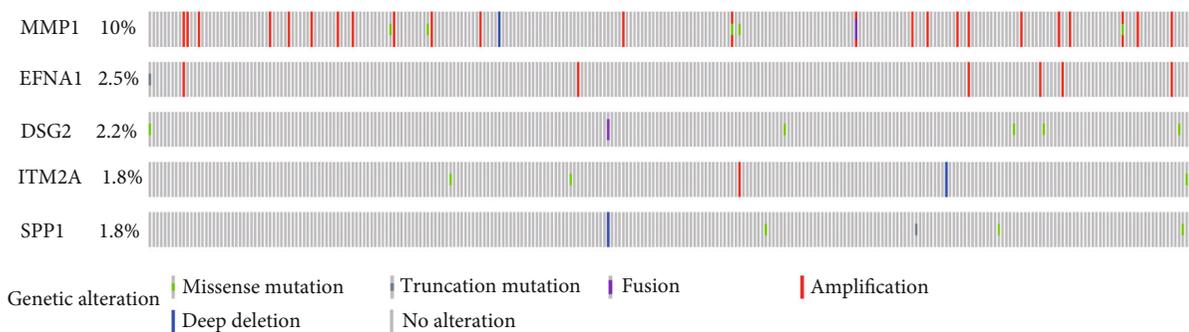
**P* < 0.05.

FIGURE 5: Somatic mutations of the five genes contained in the prognostic signature.

adenocarcinoma. In clinical practice, ITM2A might be used as an indicator for the classification of cervical cancer, and cervical cancer patients with low ITM2A expression could be treated more aggressively. Of course, this finding still needs further investigation.

Intercellular adhesion plays an important role in the maintenance of multicellular structures and cell-cell signal transmission [50]. Desmosomes are intercellular connections that provide strong adhesion strength, and the dysregulation of desmosome components may lead to cancer progression

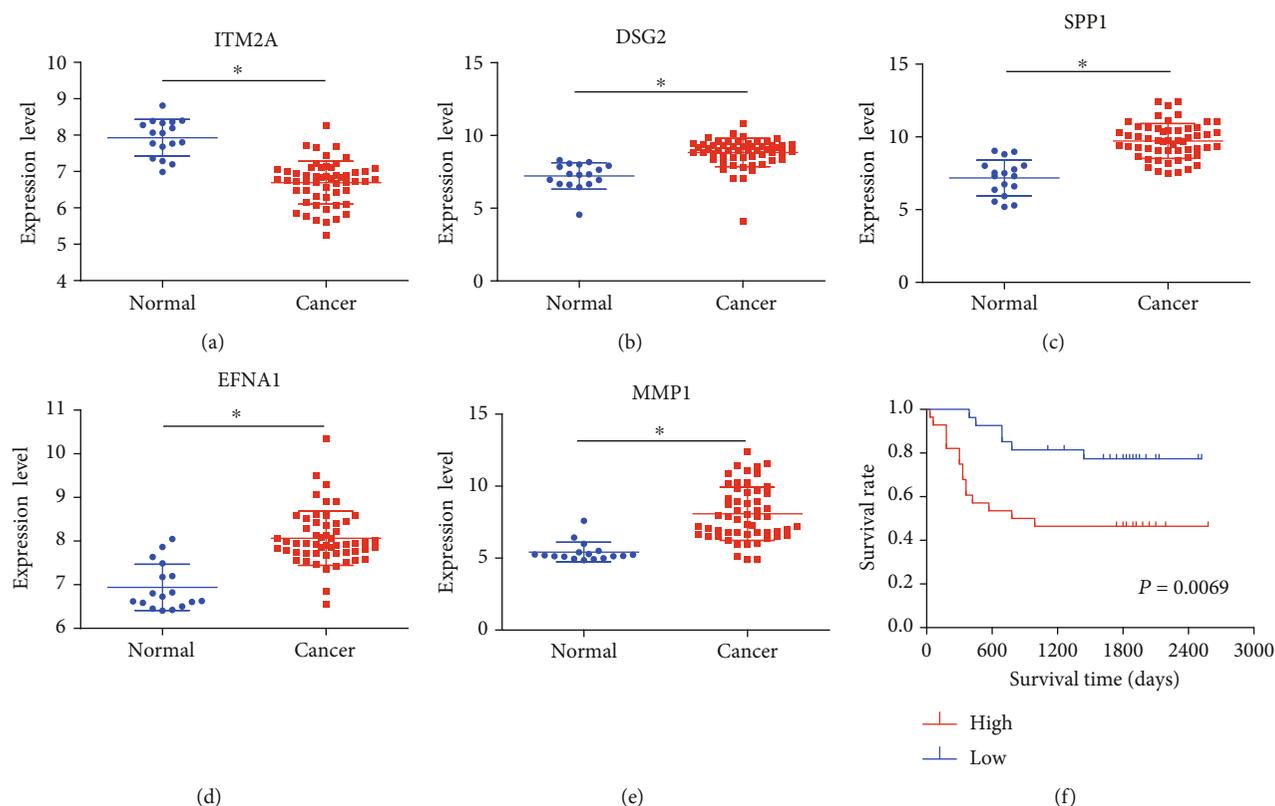


FIGURE 6: The expression level and prognostic value of the five genes in the GSE52903 validation dataset. Differential expression levels of the five genes between cervical cancer and normal cervical tissues: (a) ITM2A, (b) DSG2, (c) SPP1, (d) EFNA1, and (e) MMP1. (f) The Kaplan-Meier risk survival curve analysis of overall survival in cervical cancer patients.

by altering cellular signalling pathways [51]. DSG2 is a transmembrane desmosomal cadherin [52]. DSG2 has different functions in various cancers and can act as either an oncogene or a tumour suppressor. Several investigators reported high DSG2 expression in melanomas, human epithelial squamous cell cancer, basal cell cancer, non-small cell lung cancer, and colonic adenocarcinoma [52–55]. However, DSG2 was downregulated in prostate carcinoma, pancreatic tumours, and diffuse-type gastric cancer [56–58]. DSG2 was reported to be involved in regulating cell proliferation and tumorigenesis in cancers. Cai et al. [54] revealed that the knockdown of DSG2 expression led to the growth inhibition of non-small cell lung cancer cells and G1 phase arrest. Kamekura et al. [55] found that the loss of DSG2 inhibited cell proliferation through EGFR signalling in colon cancer. Vessel formation was attenuated in the DSG2 loss strain of mice [59]. Tan et al. [53] found that DSG2 might play crucial roles in regulating vasculogenic mimicry activity in human melanoma, which is another important method of blood supply in tumours. DSG2 also served as a poor prognostic indicator in melanoma [53]. There have been no reports about DSG2 in cervical cancer. In this study, the upregulation of DSG2 was observed in cervical cancer and indicated a poor outcome. Notably, DSG2 was associated with the HPV status of cervical cancer. These results suggested that DSG2 may be involved in HPV-induced cervical cancer. However, the involvement of the synergistic effects of DSG2 and HPV on cervical carcinogenesis needs further study.

5. Conclusion

In conclusion, the present study combined 5 microarray datasets with an RNA-Seq dataset and constructed and validated a 5-gene mRNA expression-based signature, which may serve as an independent indicator of cervical cancer survival. The prognostic values of SPP1, EFNA1, and MMP-1 in cervical cancer have been investigated in previous studies. Similar results were also found in our study. It is worth noting that ITM2A and DSG2 are two novel genes that have never been studied in cervical cancer. Furthermore, ITM2A was related to different histological phenotypes of cervical cancer. DSG2 was associated with the HPV status of cervical cancer. These findings lead to a better understanding of the underlying mechanisms of cervical cancer and help lay a foundation for making precise individual clinical treatment decisions. However, the present findings need further verification in future studies. Knockdown and overexpression of these five genes in cell and animal models could be performed to verify the carcinogenic functions of these genes in cervical cancer.

Data Availability

Expression data and clinical information of cervical cancer and normal cervical samples analyzed in this study may be acquired from the TCGA (<http://cancergenome.nih.gov/>) and GEO (<http://www.ncbi.nlm.nih.gov/geo/>).

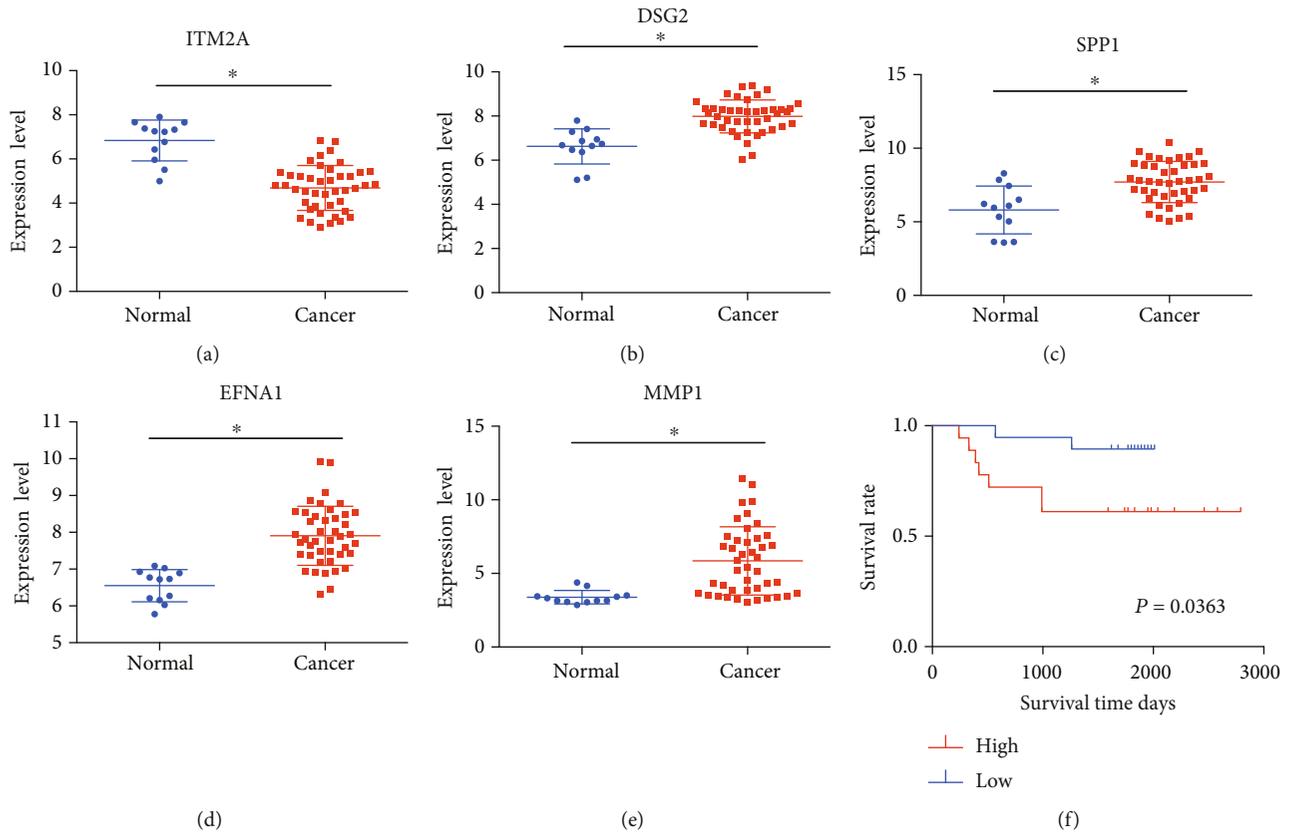


FIGURE 7: The expression level and prognostic value of the five genes in the GSE39001 validation dataset. Differential expression levels of the five genes between cervical cancer and normal cervical tissues: (a) ITM2A, (b) DSG2, (c) SPP1, (d) EFNA1, and (e) MMP1. (f) The Kaplan-Meier risk survival curve analysis of the overall survival in cervical cancer patients.

TABLE 7: Prognostic values of clinical parameters and risk score in cervical cancer.

Variables		Cases	Univariate Cox HR	Univariate Cox P value	Multivariable Cox HR	Multivariable Cox P value
Age	<50 years	72				
	>50 years	48	0.794	0.643		
HPV status	Positive	7				
	Negative	113	0.044	0.460		
Histology	Squamous	94				
	Adenocarcinoma	23	1.277	0.670		
	Adenosquamous	3	<0.001	0.985		
Stage	I-II	102				
	III-IV	18	0.442	0.429		
Lymph node metastasis	Positive	87				
	Negative	33	0.312	0.020*	0.819	0.735
Risk score			4.138	<0.001*	3.845	0.001*

* $P < 0.05$, HR: hazards ratio.

Ethical Approval

All data in this study were obtained from publicly available datasets. No additional patients were enrolled in present study.

Conflicts of Interest

The authors have declared that no competing interests exist.

Acknowledgments

This work was supported by grants from the Science and Technology Planning Project of Wenzhou City (No. ZS2017006) and Center for Uterine Cancer Diagnosis & Therapy Research of Zhejiang Province. The sponsor had no involvement in the collection, analysis, and interpretation of data or in the writing of the manuscript.

Supplementary Materials

Table SI: differentially expressed genes. (*Supplementary Materials*)

References

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.
- [2] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA: a Cancer Journal for Clinicians*, vol. 69, no. 1, pp. 7–34, 2019.
- [3] J. Zhu, S. Ji, Q. Hu et al., "Concurrent weekly single cisplatin vs triweekly cisplatin alone with radiotherapy for treatment of locally advanced cervical cancer: a meta-analysis," *Cancer Management and Research*, vol. 2018, pp. 1975–1985, 2018.
- [4] C. Gao, C. Zhou, J. Zhuang et al., "MicroRNA expression in cervical cancer: novel diagnostic and prognostic biomarkers," *Journal of Cellular Biochemistry*, vol. 119, no. 8, pp. 7080–7090, 2018.
- [5] S. Serrati, S. De Summa, B. Pilato et al., "Next-generation sequencing: advances and applications in cancer diagnosis," *OncoTargets and Therapy*, vol. 9, pp. 7355–7365, 2016.
- [6] X. Q. Cao, H. S. Lu, L. Zhang, L. L. Chen, and M. F. Gan, "MEKK3 and survivin expression in cervical cancer: association with clinicopathological factors and prognosis," *Asian Pacific Journal of Cancer Prevention*, vol. 15, no. 13, pp. 5271–5276, 2014.
- [7] M. Rashel, N. Alston, and S. Ghazizadeh, "Protein kinase D1 has a key role in wound healing and skin carcinogenesis," *The Journal of Investigative Dermatology*, vol. 134, no. 4, pp. 902–909, 2014.
- [8] D. Pyeon, M. A. Newton, P. F. Lambert et al., "Fundamental differences in cell cycle deregulation in human papillomavirus-positive and human papillomavirus-negative head/neck and cervical cancers," *Cancer Research*, vol. 67, no. 10, pp. 4605–4619, 2007.
- [9] Y. Zhai, R. Kuick, B. Nan et al., "Gene expression analysis of preinvasive and invasive cervical squamous cell carcinomas identifies HOXC10 as a key mediator of invasion," *Cancer Research*, vol. 67, no. 21, pp. 10163–10172, 2007.
- [10] L. Scotto, G. Narayan, S. V. Nandula et al., "Identification of copy number gain and overexpressed genes on chromosome arm 20q by an integrative genomic approach in cervical cancer: potential role in progression," *Genes, Chromosomes & Cancer*, vol. 47, no. 9, pp. 755–765, 2008.
- [11] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2," *Genome Biology*, vol. 15, no. 12, p. 550, 2014.
- [12] M. E. Ritchie, B. Phipson, D. Wu et al., "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic acids research*, vol. 43, article e47, 2015.
- [13] M. Ashburner, C. A. Ball, J. A. Blake et al., "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.
- [14] The Gene Ontology Consortium, "The Gene Ontology Resource: 20 years and still GOing strong," *Nucleic Acids Research*, vol. 47, no. D1, pp. D330–D338, 2019.
- [15] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources," *Nature Protocols*, vol. 4, no. 1, pp. 44–57, 2009.
- [16] H. Ogata, S. Goto, K. Sato, W. Fujibuchi, H. Bono, and M. Kanehisa, "KEGG: Kyoto Encyclopedia of Genes and Genomes," *Nucleic Acids Research*, vol. 27, no. 1, pp. 29–34, 1999.
- [17] O. O. Aalen, "A linear regression model for the analysis of life times," *Statistics in Medicine*, vol. 8, no. 8, pp. 907–925, 1989.
- [18] E. Cerami, J. Gao, U. Dogrusoz et al., "The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data," *Cancer Discovery*, vol. 2, no. 5, pp. 401–404, 2012.
- [19] I. Medina-Martinez, V. Barron, E. Roman-Bassaure et al., "Impact of gene dosage on gene expression, biological processes and survival in cervical cancer: a genome-wide follow-up study," *PLoS One*, vol. 9, no. 5, article e97842, 2014.
- [20] A. M. Espinosa, A. Alfaro, E. Roman-Basaure et al., "Mitosis is a source of potential markers for screening and survival and therapeutic targets in cervical cancer," *PLoS One*, vol. 8, no. 2, article e55975, 2013.
- [21] L. M. Seng, A. N. Rosman, A. Khan et al., "Awareness of cervical cancer among women in Malaysia," *International Journal of Health Sciences*, vol. 12, no. 4, pp. 42–48, 2018.
- [22] G. H. Williams and K. Stoeber, "The cell cycle and cancer," *The Journal of Pathology*, vol. 226, no. 2, pp. 352–364, 2012.
- [23] D. W. Hosmer and S. Lemeshow, *Applied logistic regression*, John Wiley and Sons, New York, NY, 2nd edition, 2000.
- [24] D. N. Haylock and S. K. Nilsson, "Osteopontin: a bridge between bone and blood," *British Journal of Haematology*, vol. 134, no. 5, pp. 467–474, 2006.
- [25] N. S. Fedarko, A. Jain, A. Karadag, M. R. Van Eman, and L. W. Fisher, "Elevated serum bone sialoprotein and osteopontin in colon, breast, prostate, and lung cancer," *Clinical Cancer Research*, vol. 7, no. 12, pp. 4060–4066, 2001.
- [26] B. Zeng, M. Zhou, H. Wu, and Z. Xiong, "SPP1 promotes ovarian cancer progression via integrin β 1/FAK/AKT signaling pathway," *OncoTargets and therapy*, vol. 11, pp. 1333–1343, 2018.

- [27] R. Liersch, J. Gerss, C. Schliemann et al., "Osteopontin is a prognostic factor for survival of acute myeloid leukemia patients," *Blood*, vol. 119, no. 22, pp. 5215–5220, 2012.
- [28] A. Thomas, U. Mahantshetty, S. Kannan et al., "Expression profiling of cervical cancers in Indian women at different stages to identify gene signatures during progression of the disease," *Cancer Medicine*, vol. 2, no. 6, pp. 836–848, 2013.
- [29] H. Cho, S. W. Hong, Y. J. Oh et al., "Clinical significance of osteopontin expression in cervical cancer," *Journal of Cancer Research and Clinical Oncology*, vol. 134, no. 8, pp. 909–917, 2008.
- [30] H. Wada, H. Yamamoto, C. Kim et al., "Association between ephrin-A1 mRNA expression and poor prognosis after hepatectomy to treat hepatocellular carcinoma," *International Journal of Oncology*, vol. 45, no. 3, pp. 1051–1058, 2014.
- [31] Y. Y. Mao, F. Y. Jing, M. J. Jin et al., "rs12904 polymorphism in the 3'UTR of EFNA1 is associated with colorectal cancer susceptibility in a Chinese population," *Asian Pacific Journal of Cancer Prevention : APJCP*, vol. 14, no. 9, pp. 5037–5041, 2013.
- [32] X. D. Cui, M. J. Lee, G. R. Yu et al., "EFNA1 ligand and its receptor EphA2: potential biomarkers for hepatocellular carcinoma," *International Journal of Cancer*, vol. 126, no. 4, pp. 940–949, 2010.
- [33] R. Nakamura, H. Kataoka, N. Sato et al., "EPHA2/EFNA1 expression in human gastric cancer," *Cancer Science*, vol. 96, no. 1, pp. 42–47, 2005.
- [34] S. Abraham, D. W. Knapp, L. Cheng et al., "Expression of EphA2 and Ephrin A-1 in carcinoma of the urinary bladder," *Clinical Cancer Research*, vol. 12, no. 2, pp. 353–360, 2006.
- [35] S. E. Larkin, S. Holmes, I. A. Cree et al., "Identification of markers of prostate cancer progression using candidate gene expression," *British Journal of Cancer*, vol. 106, no. 1, pp. 157–165, 2012.
- [36] R. Holm, G. V. de Putte, Z. Suo, A. K. Lie, and G. B. Kristensen, "Expressions of EphA2 and EphrinA-1 in early squamous cell cervical carcinomas and their relation to prognosis," *International Journal of Medical Sciences*, vol. 5, no. 3, pp. 121–126, 2008.
- [37] D. Wu, Z. Suo, G. B. Kristensen et al., "Prognostic value of EphA2 and EphrinA-1 in squamous cell cervical carcinoma," *Gynecologic Oncology*, vol. 94, no. 2, pp. 312–319, 2004.
- [38] C. M. Overall and R. A. Dean, "Degradomics: systems biology of the protease web. Pleiotropic roles of MMPs in cancer," *Cancer Metastasis Reviews*, vol. 25, no. 1, pp. 69–75, 2006.
- [39] L. Lu, Y. Sun, Y. Li, and P. Wan, "The polymorphism *MMP1* – 1607 (1G>2G) is associated with a significantly increased risk of cancers from a meta-analysis," *Tumour Biology*, vol. 36, no. 3, pp. 1685–1693, 2015.
- [40] T. Rajkumar, K. Sabitha, N. Vijayalakshmi et al., "Identification and validation of genes involved in cervical tumourigenesis," *BMC Cancer*, vol. 11, no. 1, p. 80, 2011.
- [41] R. Tian, X. Li, Y. Gao, Y. Li, P. Yang, and K. Wang, "Identification and validation of the role of matrix metalloproteinase-1 in cervical cancer," *International Journal of Oncology*, vol. 52, no. 4, pp. 1198–1208, 2018.
- [42] K. Pittois, J. Wauters, P. Bossuyt, W. Deleersnijder, and J. Merregaert, "Genomic organization and chromosomal localization of the *Itm2a* gene," *Mammalian Genome*, vol. 10, no. 1, pp. 54–56, 1999.
- [43] L. Sanchez-Pulido, D. Devos, and A. Valencia, "BRICHOS: a conserved domain in proteins associated with dementia, respiratory distress and cancer," *Trends in Biochemical Sciences*, vol. 27, no. 7, pp. 329–332, 2002.
- [44] D. Van den Plas and J. Merregaert, "In vitro studies on *Itm2a* reveal its involvement in early stages of the chondrogenic differentiation pathway," *Biology of the Cell*, vol. 96, no. 6, pp. 463–470, 2004.
- [45] D. Van den Plas and J. Merregaert, "Constitutive overexpression of the integral membrane protein *Itm2A* enhances myogenic differentiation of C2C12 cells," *Cell Biology International*, vol. 28, no. 3, pp. 199–207, 2004.
- [46] M. Kihara, T. Kiyoshima, K. Nagata et al., "Itm2a expression in the developing mouse first lower molar, and the subcellular localization of *Itm2a* in mouse dental epithelial cells," *PLoS One*, vol. 9, no. 7, article e103928, 2014.
- [47] T. M. Nguyen, I. W. Shin, T. J. Lee et al., "Loss of *ITM2A*, a novel tumor suppressor of ovarian cancer through G2/M cell cycle arrest, is a poor prognostic factor of epithelial ovarian cancer," *Gynecologic Oncology*, vol. 140, no. 3, pp. 545–553, 2016.
- [48] J. Jonska-Gmyrek, L. Gmyrek, A. Zolciak-Siwinska, M. Kowalska, and B. Kotowicz, "Adenocarcinoma histology is a poor prognostic factor in locally advanced cervical cancer," *Current Medical Research and Opinion*, vol. 35, no. 4, pp. 595–601, 2019.
- [49] S. Mabuchi, M. Okazawa, K. Matsuo et al., "Impact of histological subtype on survival of patients with surgically-treated stage IA2-IIB cervical cancer: adenocarcinoma versus squamous cell carcinoma," *Gynecologic Oncology*, vol. 127, no. 1, pp. 114–120, 2012.
- [50] A. Gupta, D. Nitou, D. Brennan-Crispi et al., "Cell cycle- and cancer-associated gene networks activated by *Dsg2*: evidence of cystatin A deregulation and a potential role in cell-cell adhesion," *PLoS One*, vol. 10, no. 3, article e0120091, 2015.
- [51] K. Hutz, J. Zeiler, L. Sachs, S. Ormanns, and V. Spindler, "Loss of desmoglein 2 promotes tumorigenic behavior in pancreatic cancer cells," *Molecular Carcinogenesis*, vol. 56, no. 8, pp. 1884–1895, 2017.
- [52] D. Brennan and M. G. Mahoney, "Increased expression of *Dsg2* in malignant skin carcinomas: a tissue-microarray based study," *Cell Adhesion & Migration*, vol. 3, no. 2, pp. 148–154, 2009.
- [53] L. Y. Tan, C. Mintoff, M. Z. Johan et al., "Desmoglein 2 promotes vasculogenic mimicry in melanoma and is associated with poor clinical outcome," *Oncotarget*, vol. 7, no. 29, pp. 46492–46508, 2016.
- [54] F. Cai, Q. Zhu, Y. Miao, S. Shen, X. Su, and Y. Shi, "Desmoglein-2 is overexpressed in non-small cell lung cancer tissues and its knockdown suppresses NSCLC growth by regulation of p27 and CDK2," *Journal of Cancer Research and Clinical Oncology*, vol. 143, no. 1, pp. 59–69, 2017.
- [55] R. Kamekura, K. N. Kolegraff, P. Nava et al., "Loss of the desmosomal cadherin desmoglein-2 suppresses colon cancer cell proliferation through EGFR signaling," *Oncogene*, vol. 33, no. 36, pp. 4531–4536, 2014.
- [56] A. G. Barber, M. Castillo-Martin, D. M. Bonal, B. A. Rybicki, A. M. Christiano, and C. Cordon-Cardo, "Characterization of desmoglein expression in the normal prostatic gland. Desmoglein 2 is an independent prognostic factor for aggressive prostate cancer," *PLoS One*, vol. 9, no. 6, article e98786, 2014.

- [57] V. C. Ramani, L. Hennings, and R. S. Haun, "Desmoglein 2 is a substrate of kallikrein 7 in pancreatic cancer," *BMC Cancer*, vol. 8, no. 1, article 373, 2008.
- [58] M. Yashiro, N. Nishioka, and K. Hirakawa, "Decreased expression of the adhesion molecule desmoglein-2 is associated with diffuse-type gastric carcinoma," *European Journal of Cancer*, vol. 42, no. 14, pp. 2397–2403, 2006.
- [59] L. M. Ebert, L. Y. Tan, M. Z. Johan et al., "A non-canonical role for desmoglein-2 in endothelial cells: implications for neoangiogenesis," *Angiogenesis*, vol. 19, no. 4, pp. 463–486, 2016.