Predictive Proteomic Signature for the Prognosis of Cervical Cancer

Xiaoyu Ji
Affiliated Hospital of Qingdao University
Guangdi Chu
Affiliated Hospital of Qingdao University
Jinwen Jiao
Affiliated Hospital of Qingdao University
Teng Lv
Affiliated Hospital of Qingdao University
Yulong Chen
Affiliated Hospital of Qingdao University
Qin Yao (dr_yaoqin@126.com)
Affiliated Hospital of Qingdao University https://orcid.org/0000-0003-1546-1048

Research Article

Keywords: cervical cancer, proteomic panel, TCPA, TCGA, prognosis

Posted Date: August 24th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-327396/v1

Abstract

Objective: Cervical cancer (CC) is one of the most common types of malignant female cancer, and its incidence and mortality are not optimistic. Protein panels can be a powerful prognostic factor for many types of cancer. The purpose of our study was to investigate a proteomic panel to predict survival of patients with common CC.

Methods and results: The protein expression and clinicopathological data of CC were downloaded from The Cancer Proteome Atlas (TCPA) and The Cancer Genome Atlas (TCGA) database, respectively. We selected the prognosis-related proteins (PRPs) by univariate Cox regression analysis and found that the results of functional enrichment analysis were mainly related to apoptosis. We used Kaplan–Meier(K-M) analysis and multivariable Cox regression analysis further to screen PRPs to establish a prognostic model, including BCL2, SMAD3, and 4EBP1-pT70. The signature was verified to be independent predictors of OS by Cox regression analysis and the Area Under Curves. Nomogram and subgroup classification were established based on the signature to verify its clinical application. Furthermore, we looked for the co-expressed proteins of three-protein panel as potential prognostic proteins.

Conclusion: A proteomic signature independently predicted OS of CC patients, and the predictive ability was better than the clinicopathological characteristics. This signature can help improve prediction for clinical outcome and provides new targets for CC treatment.

1 Introduction

Cervical cancer is a severe public health challenge around the world. The 2018 global cancer statistics show 569,847 women develop CC and 311,365 women die from it each year[1]. Squamous cell carcinoma (SCC) and adenocarcinoma (AD) account for more than 95% of all cervical cancers[2]. Improved prognosis assessment is beneficial to guide early treatment decisions in disease management better. The International Federation of Gynecology and Obstetrics (FIGO) staging system[3] and histological type are common indicators for evaluating prognosis and determining therapeutic interventions for gynecologic tumors. FIGO staging mainly describes the degree of tumor invasion and ignores other important prognostic variables, particularly genomic or proteomic differences. FIGO/TNM staging is also difficult to explain individual differences in patients with the same stage. Therefore, it is necessary to establish a genome-based or proteomic based prognostic score system to more accurately predict the clinical prognosis of individual patients [4].

For the last century, our understanding of molecular diagnosis and treatment of cancer-based on a single biomolecular method has been limited. In recent years, high-throughput data have burgeoned, and public databases have been fully utilized[5]. A growing number of biomarkers are being discovered that may be useful for diagnosing cancer, assessing prognosis, providing promising therapeutic targets and personalizing treatments[5]. However, most of these molecules are based on levels of genomes[6], transcriptional profiling, microRNAs[7] or long non-coding RNAs[8], and they cannot represent all levels of

biological complexity. The level and function of proteins depend not only on transcription and translation but also on modification, which affects the stability and activity of proteins. So protein is a direct factor affecting cell function[9]. Reverse-phase protein array (RPPA) can quantitatively evaluate abundant protein markers by a cost-effective, precise and sensitive approach[10]. Thus we obtain the proteome data from TCPA that is an integrated protein-centric analytic module that contains RPPA data that can be visualized and analyzed in the rich context of TCGA[11]. Single protein biomarkers lack the accuracy of accurate detection of cancer, and their clinical effectiveness is minimal. A signature of proteins would be a more accurate predictor compared with a single protein[12].

Here, we have screened PRPs for constructing a model using protein expression profiles from TCPA along with patient outcomes from TCGA. We have evaluated whether these multi-panel protein prediction models without clinical indicators can be applied as a reliable biomarker and have analyzed the co-expressed proteins of the panel to identify potential prognostic biomarkers.

2 Material And Methods

2.1 Data sources

The protein expression data of 171 CC patients in the TCPA database (https://www.tcpaportal.org/) and the clinical and data of 286 common CC patients from the TCGA database (http://cancergenome.nih.gov/) were downloaded. We used level 4 protein values of CC in TCPA and used R package (impute) complete the missing values in RPPA data. A total of 164 patients were present in both databases, and the Perl (http://www.perl.org/) script was used to combine these files to obtain a matrix file of complete protein expression and survival information for the 164 CC patients (Table S1). The patients of SCC and AD were 142 and 22, respectively. A total of 217 protein markers were assayed. Further screening was performed, and 136 patients had complete clinical data, including age, histological type, histological grade, and TNM stage. The workflow of the entire study was shown in Fig. 1.

2.2 Constructing the multi-protein prognostic model and risk score assessment

2.2.1 Screen and analysis for PRPs

PRPs were screened by univariate Cox regression analysis. At the same time, the p-value and hazard ratio (HR) of each protein were recorded. We found the gene corresponding to each protein in the TCPA database and carried out Gene Ontology (GO) term functional enrichment and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses to determine the major biological properties. Protein-protein interaction (PPI) information was evaluated via the Search Tool for the Retrieval of Interacting Genes Database (STRING) (https://string-db.org/).

2.2.2 Construct the prognostic proteins model model

We divided the patients into high expression group and low expression group based on the median of the expression of protein screened by univariate Cox analysis as the cut-off point. We exerted K-M survival analysis and multivariate Cox regression analysis to capture the proteins, established the optimal multi-protein prognostic signature, and obtained the regression coefficients of each protein. The risk score was defined as the sum of the expression levels of each protein multiplied by the corresponding coefficients.

2.2.3 Prediction model assessment

Based on the signature, the risk score of each patient was calculated, and the median of the score was used as a critical value to assign patients into the high-risk group or low-risk group. K-M survival analysis was used to compare whether there were differences in survival outcomes between the two risk grades. To prove the synergistic effect of the signature on the prediction of OS, time-dependent ROC curves were used to test that the signature had higher sensitivity and specificity than a single protein. The area under the curve (AUC) > 0.60 indicates an acceptable predictive value, and AUC > 0.75 indicates a good predictive value.

To verify that our risk score could be independent of other clinicopathologic features as prognostic factors, we excluded the missing clinical data from the remaining 136 cases (Table S2). Risk score was confirmed as independent prognostic factors by univariate and multivariate Cox analysis. Receiver Operating Characteristic (ROC) analysis was used to state the prediction accuracy of the prediction model. A composite nomogram was constructed according to the important clinical parameters related to prognosis and the risk score measured by the protein signature. Calibration plots were used to evaluate the difference between the actual value and the predicted value, and time-dependent ROC curves were again used to test the sensitivity and specificity of nomogram.

2.2.4 The Molecular Subtypes of Cervical Cancer

Unsupervised class discovery is an effective method to classify individuals with similar biological characteristics. With the help of R packet ConsensusClusterPlus, consensus clustering can provide sufficient evidence to identify unsupervised classes in data sets. In this study, we identify the subsets of CC by ConsensusClusterPlus according to the expression level of the protein in the signature. The maximum number of subgroups was set as 10, and 1000 permutations were performed to ensure the stability of the classification. K-M analysis with the log-rank test was performed in these subgroups, and $p < 0.05$ was the cutoff value.

2.3 Statistical analysis

The methods mentioned above would not be repeated, and the following statistical methods were also applied to our research. We used the chi-square test to analyze the relationship between risk level and clinicopathologic features. The linear correlation analysis was used to find the co-expressed proteins of the signature, and the correlation coefficients were examined by t-test. The correlation coefficient greater than 0.3 was taken as the filtration condition. In all tests, we considered P-values of < 0.05 (two-sided) to

indicate statistical significance. Data processing, model construction and assessment were performed using R software (v 3.6.2).

3 Result

3.1 Identification and selection of PRPs

We have used univariable Cox regression analysis in our cohort to identify 29 PRPs ($p < 0.05$), including 22 high risk proteins (CYCLIND1, SMAD3, 4EBP1-pT70, RAD51, HER3-pY1289, MIG6, RAB11, 1433β, NCADHERIN, BCLXL, CRAF-pS338, BID, ASNS, FASN, MRE11, PAI1, ACC1, YB1, RAPTOR, TAZ, MEK1-pS217S221, HSP70) and 7 low risk proteins (BCL2, CHK1-pS345, KU80, SRC-pY416, ERα-pS118, JAK2, CD31). According to the p-value and HR, the results are displayed in the form of volcano plot (Fig. 2b) to distinguish high-risk and low-risk proteins visually. The PPI network better demonstrated the interaction between prognostic proteins. The network included 29 nodes and 204 intersections (coincidence interval ≥ 0.4, Fig. 2c). All the corresponding genes were visualized, and *CCND1*, *ESR1* and *SRC* could be regarded as hub genes (Fig. 2d).

3.2 Functional enrichment analysis of PRPs

Functional enrichment analysis of the 29 corresponding genes of PRPs provided a biological comprehension about these proteins, and the corresponding genes for PRPs was shown in Table I. The results of GO term functional enrichment and the KEGG pathway enrichment analyses were presented in Fig. 3.

Table I.The corresponding gene to the PRPs and the ID of the gene

protein	p-value(unCOX)	p-value(K-M)	HR	gene	geneID
NCADHERIN	0.014	0.000	2.599	*CDH2*	1000
RAB11	0.007	0.002	42.658	*RAB11A\RAB11B*	8766\9230
YB1	0.041	0.002	5.319	*YBX1*	4904
BCL2	0.002	0.019	0.169	*BCL2*	596
CHK1-pS345	0.013	0.023	0.099	*CHEK1*	1111
SMAD3	<0.001	0.023	20.309	*SMAD3*	4088
4EBP1-pT70	0.001	0.026	80.523	*EIF4EBP1*	1978
JAK2	0.046	0.031	0.408	*JAK2*	3717
MEK1-pS217S221	0.045	0.033	2.283	*MAP2K1*	5604
BCLXL	0.016	0.041	5.935	*BCL2L1*	598
CRAF-pS338	0.018	0.043	3.561	*RAF1*	5894
CYCLIND1	<0.001	0.065	11.877	*CCND1*	595
RAPTOR	0.043	0.066	7.590	*RPTOR*	57521
SRC-pY416	0.028	0.083	0.662	*SRC*	6714
MRE11	0.023	0.095	3.034	*MRE11*	4361
KU80	0.020	0.099	0.477	*XRCC5*	7520
RAD51	0.002	0.130	2.321	*RAD51*	5888
ERALPHA-pS118	0.043	0.140	0.247	*ESR1*	2099
TAZ	0.044	0.158	3.342	*TAZ*	6901
HSP70	0.049	0.182	2.050	*HSPA4*	3308
HER3-pY1289	0.004	0.206	10.541	*ERBB3*	2065
ASNS	0.022	0.211	1.899	*ASNS*	440
MIG6	0.005	0.311	26.596	*ERRFI1*	54206
BID	0.022	0.334	5.262	*BID*	637
PAI1	0.025	0.435	1.428	*SERPINE1*	5054
CD31	0.049	0.443	0.181	*PECAM1*	5175
1433BETA	0.013	0.454	10.026	*YWHAB*	7529

ACC1	0.031	0.741	2.276	*ACACA*	31
FASN	0.023	0.923	1.996	*FASN*	2194

Details of all gene enrichment analysis results were used as supplementary materials (Table S3). In biological processes, the top three enriched GO terms were extrinsic apoptotic signaling pathway, gland development and regulation of apoptotic signaling pathway. Cellular components included chromosome, telomeric region, mitochondrial outer membrane and site of DNA damage. From a molecular function point of view, the terms of ubiquitin and ubiquitin-like protein ligase binding and protein C-terminus binding included enriched genes. In the KEGG pathway enrichment analysis, these corresponding genes proved to be significantly correlated with EGFR tyrosine kinase inhibitor resistance, cellular senescence, p53 signaling pathway and PI3K-Akt signaling pathway.

3.3 Identification of a signature for the prognosis of CC

The 29 PRPs selected by univariate regression analysis, and then 11 key proteins were screened by K-M survival analysis. The 11 key proteins included NCADHERIN, RAB11, YB1, BCL2, CHK1-pS345, SMAD3, 4EBP1-pT70, JAK2, MEK1-pS217S221, BCLXL and CRAF-pS338 (Table I). Subsequently, we used multivariate Cox analysis to select BCL2, SMAD3 and 4EBP1-pT70 as signatures from 11 PRPs, and received expression coefficients of them. The Kaplan–Meier survival curves of these three proteins were shown in Figs. 4a-c. The OS was lower in the high expression group of SMAD3 and 4EBP1-pT70, while the OS was higher in the high expression group of BCL2. A prognostic model for common CC based on the three proteins was conducted by the following formula: risk score = (2.413 ×expression level of SMAD3) + (3.437 × expression level of 4EBP1-pT70) + (-1.514×expression level of BCL2). Then we calculated the median value of the risk score as 0.98, which was set as a threshold to classify the patients into the high-risk group or low-risk group, and the K-M curve depending on risk score were displayed in Fig. 4d. The AUC values of 1-year, 3-year and 5-year signature of three proteins were higher than those of any single protein(Figs. 4e-g), which indicated that the three proteins had a synergistic effect on predicting prognosis.

The heat map and the risk status were shown in one diagram (Fig. 5). The risk score increases from left to right on the risk curve. Figure 5A represents the result of grouping by the median value. In Fig. 5B, there was no significant difference in survival time, but we can know that the number of deaths increased with the increase of risk value. As shown in Fig. 5C, the expression of SMAD3 and 4EBP1-pT70 increased with the increase in risk. These two proteins are high-risk proteins, while BCL2 is low-risk protein. The images of immunohistochemical staining for PRPs in the signature were obtained from the HPA database, and we found that SMAD3 and 4EBP1-pT70(the gene encoding EIF4EBP1) had strong staining in tumor tissue. BCL2 showed weak staining in tumor tissue (Fig. 6a). The expression differences of three corresponding genes in tumor tissues and normal tissues were analyzed by GEPIA2 online tool using TCGA data. The gene expression level of SMAD3 and EIF4EBP1 were higher in tumor tissues, while that of BCL2 was higher in normal tissues (Fig. 6b).

3.4 An independent and superior predictor compared with clinical characteristics

Cases with incomplete clinical features were excluded, remained 136 cases. The independent predictive value of the risk score for common cervical cancer OS was verified by univariate and multivariate Cox regression analysis (Figs. 7a, b). Only N stage (uncox: HR 2.069, 95%CI 1.188–3.604; multcox: HR 2.079, 95% CI 1.031–4.193) and risk score (uncox: HR 1.342, 95%CI 1.184–1.521; multcox: HR 1.398, 95% CI 1.194–1.636) were the factors associated with prognosis in either univariate or multivariate analysis, and risk score was the most significant indicator($p < 0.001$). The AUC values of the risk score was 0.765, which were apparently higher than clinical and pathological factors, including age, histological grade, histological type, and TNM staging. The AUC values indicated that the prognostic model predictive outcome had good consistency with the recoded prognosis.

In order to build a more suitable and accurate clinical practice tool, a nomogram including risk score and N stage was established (Fig. 7c). It can be seen from the calibration curve that there was little deviation between the predicted results and the actual results (Fig. 7d), nomogram 1-year, 3-year, 5-year AUC values were 0.845, 0.814, 0.786, respectively, which were higher than those of the signature alone (Fig. 7e).

3.5 Identification of PRP-based subgroups and survival analysis

The protein signature was proven to be a robust tool for predicting the prognosis of CC patients. To further explore the importance of this protein signature in CC, consensus clustering analysis was performed to identify subgroups based on protein expression. The optimal number of classification was limited by many factors. Firstly, the area of the cumulative distribution function curve needed to be stable. Secondly, the correlation between categories should not be too strong. When the clustering index “k” increased from 2 to 7, $k = 3$ was demonstrated to be the optimal point to obtain the largest differences between clusters (Figs. 8a-c). Then, we conducted Kaplan-Meier survival analysis. The results showed that there were significant differences among these subgroups($p < 0.05$), and cluster 3 had the worst prognosis (Fig. 8d).

3.6 Screening for proteins co-expressed with our protein signature

To better predict survival, we screened three-protein predictor co-expressed proteins from all the proteins measured in the TCPA database. Co-expressed proteins with the absolute value of correlation coefficients greater than 0.3 were listed in Table S4. We only selected related proteomics with a correlation coefficient greater than 0.4 to be shown in Fig. 9. Sankey diagram can show all the coexpression results more intuitively (Fig. 10).

4 Discussion

With the spread of vaccination and screening programmes, the mortality associated with cervical cancer has fallen sharply in developed countries in recent years but remains high in developing regions, where nearly 90 percent of cervical cancer deaths occur[1, 13]. If the prognosis of patients can be known in time, it will sound the alarm for treatment and follow-up. As we mentioned above, advances in high-throughput sequencing and large-scale databases have facilitated to excavate large amounts of molecular data on diseases, thus providing a better understanding of the relationship between biomarkers and cancer. We presented the details of building a protein prediction model and analyzed the results in this study. Computationally, survival prediction was often attributed to regression of patient survival time. Cox regression model is the most commonly used method[14, 15].We identified 29 PRPs and further screened 11 as key proteins. Finally, we established an optimized signature consisting of BCL2, SMAD3 and 4EBP1-pT70 as a prediction signature for the prognosis of common cervical cancer. The model contains a small number of molecules that contribute to practical clinical applications.

The GO analysis of 29 PRPs revealed a high correlation with apoptosis in terms of biological processes and related to chromosome, DNA damage and some membrane structures in the terms of cellular components. Apoptosis is a manifestation of programmed cell death, a normal biological process, and its disorder is a hallmark of cancer. The dysregulation of apoptosis is not only related to the occurrence and development of tumor, but also related to drug resistance[16, 17]. The molecular regulatory mechanism of apoptosis is very complex and can be divided into the extrinsic pathway and intrinsic pathway. In the intrinsic pathway, mitochondrial outer membrane permeabilization is a trigger of apoptosis regulated by the B-cell lymphoma 2 (BCL-2) family [18, 19], which will be discussed in detail later. Under the set of molecular functions, the ubiquitin-protein ligase binding plays a regulatory role in extrinsic apoptotic signaling pathway[20, 21]. In addition, in KEGG analysis, enrichment of important pathways is not only closely related to apoptosis, but also cell senescence and autophagy.

SMAD3 is known to act as a signaling intermediate for transforming growth factor β (TGF-β), and high levels of TGF-β often predict poor prognosis[22]. Biological processes such as cell proliferation, apoptosis, differentiation, and motility have been shown to be regulated by the TGF-β/SMAD3 signaling pathway[23]. Some biomarkers, LncRNA PVT1 and Twist, have been proved to promote proliferation, invasion, and metastasis by enhancing the expression of SMAD3 in the cell trial of cervical cancer[24, 25]. These biomarkers are potential therapeutic targets. Enhanced phosphorylation of SMAD3 inhibits cell invasions, while inhibition of phosphorylation of SMAD3 promotes cancer cell invasions[26, 27].

4EBP1 has been shown to promote tumorigenesis, which is encoded by *EIF4EBP1* and affects the mTOR signaling pathway by inhibiting eukaryotic translation initiation factor 4e (EIF4E)[28]. In general, phosphorylated 4EBP1 is believed to indicate a poor tumor prognosis[29]. It is worth emphasizing that 4EBP1-pT70, a high-risk protein, was included in our panel. Interestingly, E6 and E7 of human papillomavirus were found to induce the expression and enhance the activity of eIF4E[30, 31],which indicated that 4EBP1 may be more expressed in cervical cancer associated with HPV infection. Drugs that act on the Akt /mTOR/4EBP1 pathway were being discovered. Britannin indirectly suppressed the expression of PD-L1 through this pathway, thus stabilizing the activity of T cells and inhibiting

proliferation and angiogenesis to complete the anti-tumor effect[32]. Some synthetic drugs[33, 34] that could inhibit mTOR/4EBP1 pathway showed superior anti-tumor activity, which suggested that 4EBP1 was a new anticancer drug target for specific anticancer therapy and a promising method to overcome drug resistance and enhance the anti-tumor efficacy of chemotherapy.

BCL2 is not only an effective anti-apoptotic molecule but also a carcinogenic protein[35], which can enhance the sensitivity to carcinogenesis via DNA replication stress[36]. Counter-intuitively, low expression of the BCL2 indicated worse prognosis. Of course, we were not alone in having counterintuitive results [37], and the research of Yang F et al. might provide us with ideas. They believe that miR-7-5p reduces the expression level of BCL2 in platinum-resistant cells, promotes autophagy, and increases energy generation to maintain the survival of platinum-resistant cells[38]. In addition, we speculate that BCL2 is an effective target for the treatment of common CC. In several studies, *BCL2* has been reported as a target gene for some miRNAs to inhibit tumor cells, such as miR-449a, miR-136 and miR-365[39–41]. Therefore, we believed that the elusive molecular mechanism was responsible for the better prognosis of patients with high expression, and the phenomenon needs to be explained in the context of specific diseases. It could be a therapeutic opportunity for patients with high BCL2 expression.

Given the complexity of molecular mechanisms, we should realize the necessity of establishing a multi-molecular signature. Due to the complex role of protein signaling pathways, many proteins play a bifunctional role in tumors. Understanding these specific patterns can help screen high-risk biomarkers for potential therapeutic targets. For this purpose, we also screened the co-expressed proteins related to the establishment of the signature in the database as potential research targets. However, we were unable to examine a causal relationship between risk score and prognostic status in patients with CC. And more prospective studies are needed to validate the predictive performance of our model.

In summary, the expression of 29 PRPs, and the enrichment of their functions was mainly related to apoptosis. Based on the optimized model of a three-protein signature, a promising method for predicting the prognosis of patients with cervical cancer was proposed, and several potential therapeutic targets for cervical cancer were pointed out. These results could provide additional reflection and improvement for treatment decision making and further research.

Declarations

Author Contribution

Xiaoyu Ji: Conceptualization, Writing- original draft, Writing-review & editing. Guangdi Chu: Writing-original draft, Methodology, Software.Jinwen Jiao: Software, Validation. Teng Lv: Methodology, Visualization. Yulong Chen: Visualization, Investigation. Qin Yao: Writing-Review & Editing, Project administration. All authors read and approved the final manuscript.

Competing interests

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

All authors have seen the manuscript and approved to submit to your journal. Neither the entire paper nor any part of its content has been published or has been accepted elsewhere.

Acknowledgements

We acknowledge the contributions of the TCPA Database and the TCGA Database gratefully.

Funding

This work was supported by the Shandong Province Nature Science Foundation (NO. ZR2016HM08).

Availability of data and materials

We obtained the data sets from TCPA (https://www.tcpaportal.org/) and TCGA (http://cancergenome.nih.gov/).

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68(6):394–424
2. Nogueira-Rodrigues A, Ferreira CG, Bergmann A, de Aguiar SS, Thuler LC (2014) Comparison of adenocarcinoma (ACA) and squamous cell carcinoma (SCC) of the uterine cervix in a sub-optimally screened cohort: a population-based epidemiologic study of 51,842 women in Brazil. *Gynecol Oncol* 135(2):292–296

3. Bhatla N, Aoki D, Sharma DN, Sankaranarayanan R (2018) Cancer of the cervix uteri. *Int J Gynaecol Obstet* 143 Suppl 2:22–36
4. Dasari S, Wudayagiri R, Valluru L (2015) Cervical cancer: Biomarkers for diagnosis and treatment. *Clin Chim Acta* 445:7–11
5. Shukla HD (2017) Comprehensive Analysis of Cancer-Proteogenome to Identify Biomarkers for the Early Diagnosis and Prognosis of Cancer. *Proteomes* 5(4)
6. Hirose S, Murakami N, Takahashi K et al (2020) Genomic alterations in STK11 can predict clinical outcomes in cervical cancer patients. *Gynecol Oncol* 156(1):203–210
7. Hu X, Schwarz JK, Lewis JS Jr et al (2010) A microRNA expression signature for cervical cancer prognosis. *Cancer Res* 70(4):1441–1448
8. Mao X, Qin X, Li L et al (2018) A 15-long non-coding RNA signature to improve prognosis prediction of cervical squamous cell carcinoma. *Gynecol Oncol* 149(1):181–187
9. Choi CH, Chung JY, Kang JH et al (2020) Chemoradiotherapy response prediction model by proteomic expressional profiling in patients with locally advanced cervical cancer. *Gynecol Oncol*
10. Masuda M, Yamada T (2019) Utility of Reverse-Phase Protein Array for Refining Precision Oncology. *Adv Exp Med Biol* 1188:239–249
11. Chen MM, Li J, Wang Y et al (2019) TCPA v3.0: An Integrative Platform to Explore the Pan-Cancer Analysis of Functional Proteomic Data. *Mol Cell Proteomics* 18(8 suppl 1):S15-s25
12. Tan GS, Lim KH, Tan HT et al (2014) Novel proteomic biomarker panel for prediction of aggressive metastatic hepatocellular carcinoma relapse in surgically resectable patients. *J Proteome Res* 13(11):4833–4846
13. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A (2015) Global cancer statistics, 2012. *CA Cancer J Clin* 65(2):87–108
14. Shi X, Li Y, Sun Y et al (2020) Genome-wide analysis of lncRNAs, miRNAs, and mRNAs forming a prognostic scoring system in esophageal squamous cell carcinoma. *PeerJ* 8:e8368
15. Hou Z, Yang J, Wang H, Liu D, Zhang H (2019) A Potential Prognostic Gene Signature for Predicting Survival for Glioblastoma Patients. 2019:9506461
16. Fulda S (2009) Tumor resistance to apoptosis. *Int J Cancer* 124(3):511–515
17. Ngoi NYL, Choong C, Lee J et al (2020) Targeting Mitochondrial Apoptosis to Overcome Treatment Resistance in Cancer. *Cancers (Basel)* 12(3)
18. Leverson JD, Cojocari D (2018) Hematologic Tumor Cell Resistance to the BCL-2 Inhibitor Venetoclax: A Product of Its Microenvironment? *Front Oncol* 8:458
19. Kalkavan H, Green DR (2018) MOMP, cell suicide as a BCL-2 family business. *Cell Death Differ* 25(1):46–55
20. Woo SM, Kwon TK (2019) E3 ubiquitin ligases and deubiquitinases as modulators of TRAIL-mediated extrinsic apoptotic signaling pathway. *BMB Rep* 52(2):119–126

21. Ni T, Li W, Zou F (2005) The ubiquitin ligase ability of IAPs regulates apoptosis. *IUBMB Life* 57(12):779–785
22. Millet C, Zhang YE (2007) Roles of Smad3 in TGF-beta signaling during carcinogenesis. *Crit Rev Eukaryot Gene Expr* 17(4):281–293
23. Padua D, Massague J (2009) Roles of TGFbeta in metastasis. *Cell Res* 19(1):89–102
24. Chang QQ, Chen CY, Chen Z, Chang S (2019) LncRNA PVT1 promotes proliferation and invasion through enhancing Smad3 expression by sponging miR-140-5p in cervical cancer. *Radiol Oncol* 53(4):443–452
25. Fan Q, Qiu MT, Zhu Z et al (2015) Twist induces epithelial-mesenchymal transition in cervical carcinogenesis by regulating the TGF-beta/Smad3 signaling pathway. *Oncol Rep* 34(4):1787–1794
26. Xia Q, Li C, Bian P, Wang J, Dong S (2014) Targeting SMAD3 for inhibiting prostate cancer metastasis. *Tumour Biol* 35(9):8537–8541
27. Chen N, Balasenthil S, Reuther J et al (2013) DEAR1 is a chromosome 1p35 tumor suppressor and master regulator of TGF-beta-driven epithelial-mesenchymal transition. *Cancer Discov* 3(10):1172–1189
28. Karlsson E, Perez-Tenorio G, Amin R et al (2013) The mTOR effectors 4EBP1 and S6K2 are frequently coexpressed, and associated with a poor prognosis and endocrine resistance in breast cancer: a retrospective study including patients from the randomised Stockholm tamoxifen trials. *Breast Cancer Res* 15(5):R96
29. Zhu J, Wang M, Hu D (2020) Development of an autophagy-related gene prognostic signature in lung adenocarcinoma and lung squamous cell carcinoma. *PeerJ* 8:e8288
30. Wang S, Pang T, Gao M et al (2013) HPV E6 induces eIF4E transcription to promote the proliferation and migration of cervical cancer. *FEBS Lett* 587(6):690–697
31. Pang T, Wang S, Gao M et al (2015) HPV18 E7 induces the over-transcription of eIF4E gene in cervical cancer. *Iran J Basic Med Sci* 18(7):684–690
32. Zhang YF, Zhang ZH, Li MY et al (2021) Britannin stabilizes T cell activity and inhibits proliferation and angiogenesis by targeting PD-L1 via abrogation of the crosstalk between Myc and HIF-1α in cancer. *Phytomedicine* 81:153425
33. Wang L, Guo C, Li X et al (2019) Design, synthesis and biological evaluation of bromophenol-thiazolylhydrazone hybrids inhibiting the interaction of translation initiation factors eIF4E/eIF4G as multifunctional agents for cancer treatment. *Eur J Med Chem* 177:153–170
34. Xiao D, Lu Z, Wang Z et al (2020) Synthesis, biological evaluation and anti-proliferative mechanism of fluorine-containing proguanil derivatives. *Bioorg Med Chem* 28(2):115258
35. Linette GP, Hess JL, Sentman CL, Korsmeyer SJ (1995) Peripheral T-cell lymphoma in lckpr-bcl-2 transgenic mice. *Blood* 86(4):1255–1260
36. Xie M, Park D, Sica GL, Deng X (2020) Bcl2-induced DNA replication stress promotes lung carcinogenesis in response to space radiation. *Carcinogenesis*

37. Guttà C, Rahman A, Aura C et al (2020) Low expression of pro-apoptotic proteins Bax, Bak and Smac indicates prolonged progression-free survival in chemotherapy-treated metastatic melanoma. *Cell Death Dis* 11(2):124
38. Yang F, Guo L, Cao Y, Li S, Li J, Liu M (2018) MicroRNA-7-5p Promotes Cisplatin Resistance of Cervical Cancer Cells and Modulation of Cellular Energy Homeostasis by Regulating the Expression of the PARP-1 and BCL2 Genes. *Med Sci Monit* 24:6506–6516
39. Hu J, Fang Y, Cao Y, Qin R, Chen Q (2014) miR-449a Regulates proliferation and chemosensitivity to cisplatin by targeting cyclin D1 and BCL2 in SGC7901 cells. *Dig Dis Sci* 59(2):336–345
40. Yu L, Zhou GQ, Li DC (2018) MiR-136 triggers apoptosis in human gastric cancer cells by targeting AEG-1 and BCL2. *Eur Rev Med Pharmacol Sci* 22(21):7251–7256
41. Zhu Y, Wen X, Zhao P (2018) MicroRNA-365 Inhibits Cell Growth and Promotes Apoptosis in Melanoma by Targeting BCL2 and Cyclin D1 (CCND1). *Med Sci Monit* 24:3679–3692

Figures

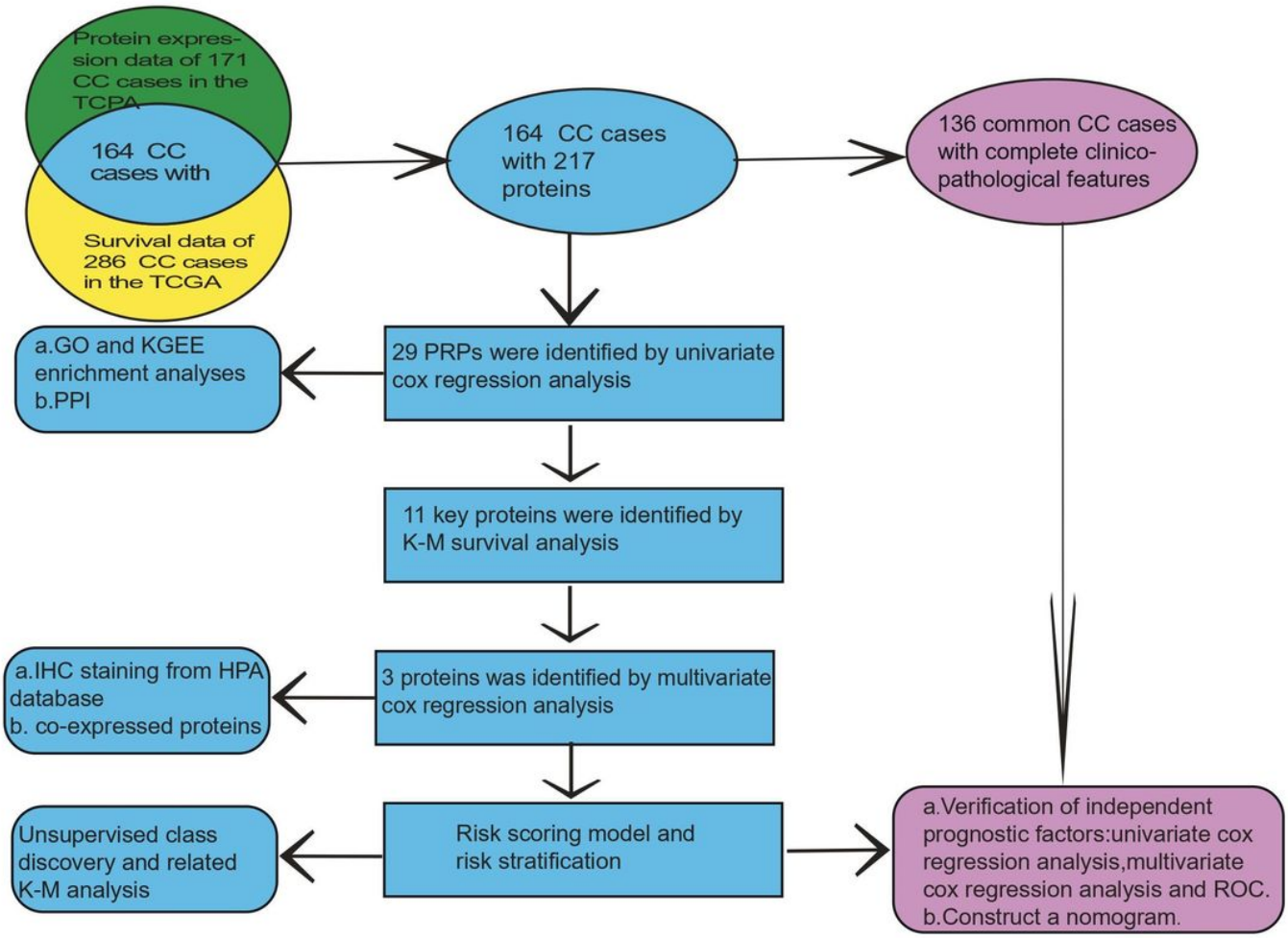

Figure 1

The workflow of this study.

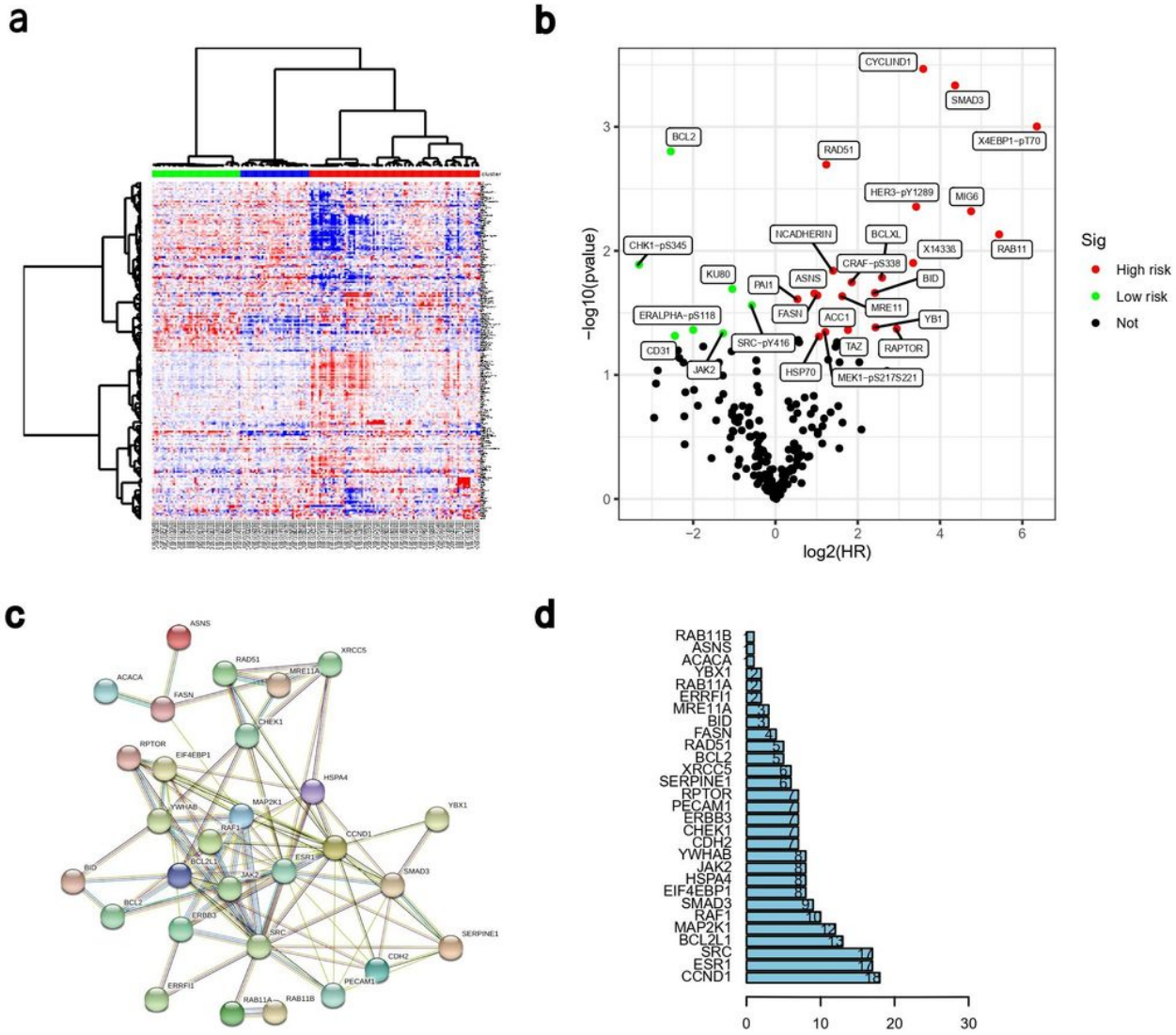

Figure 2

The PRPs in CC and the PPI network. (A) Heatmap of protein level in 171 CC cases;(B) Volcano plot of prognostic proteins. A total of 29 prognostic proteins were identified. Yellow dots represent high risk proteins, blue dots represent low risk proteins, and black dots represent non-statistically significant proteins.HR: hazard ratio. (C) The PPI network of corresponding genes for 29 PRPs. The color of the edges represents different evidence of protein-protein associations. Sources of known interactions include from curated databases (sky blue) and experimentally determined(purple), and predicted

interactions and others are represented by residual colors. (D) All genes were sequenced by the number of genes associated with them.

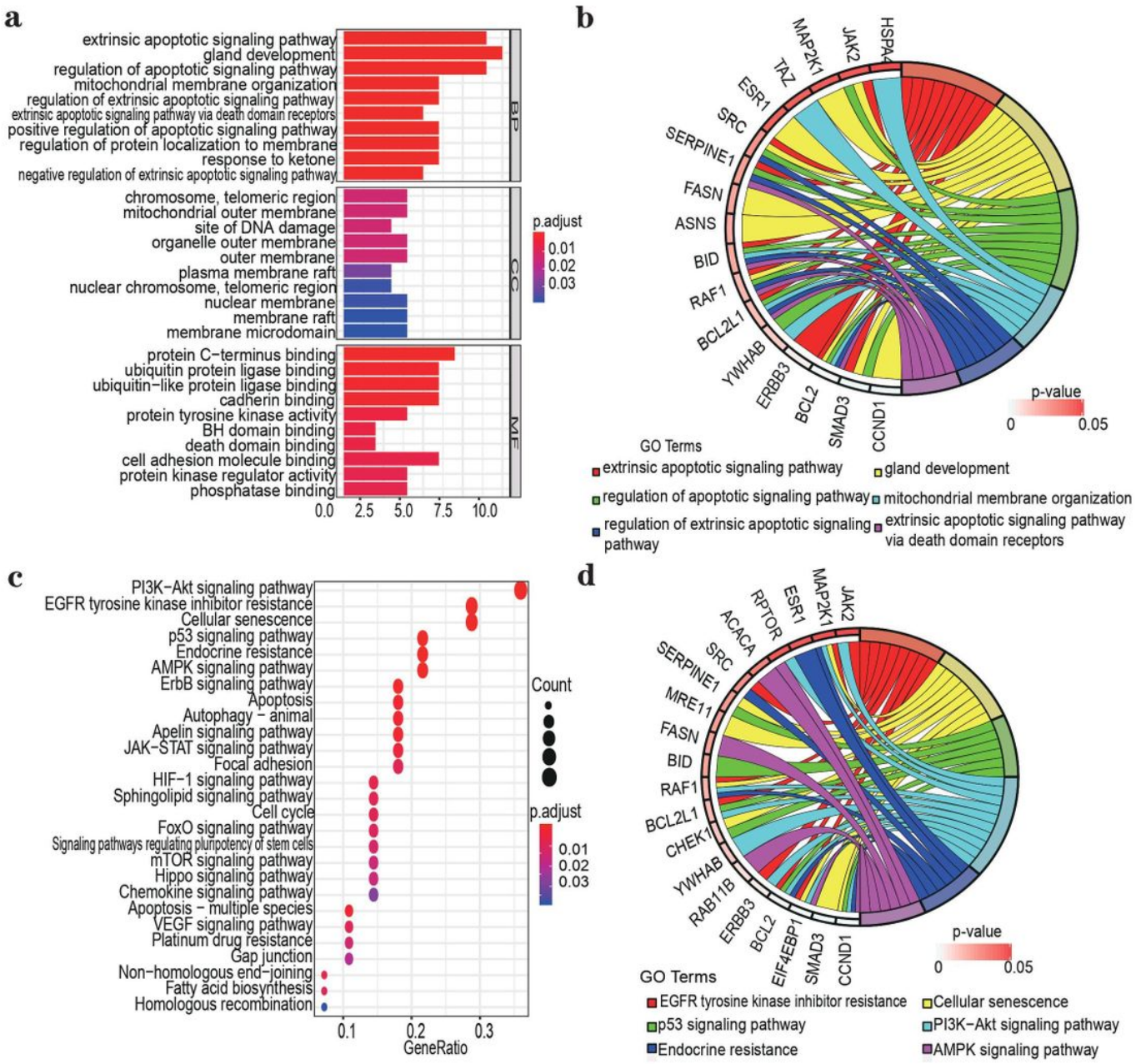

Figure 3

GO and KEGG functional enrichment analyses. The bar chart of the top 10 GO terms enriched in BP, CC and MF (A) and the upregulated and downregulated genes in the top 6 most enriched GO terms(B). The bubble chart of the 27 KEGG terms enriched (C) and the enrichment of top 6 most enriched KEGG terms on different genes (D). BP, biological processes; CC, cellular components; MF, molecular function.

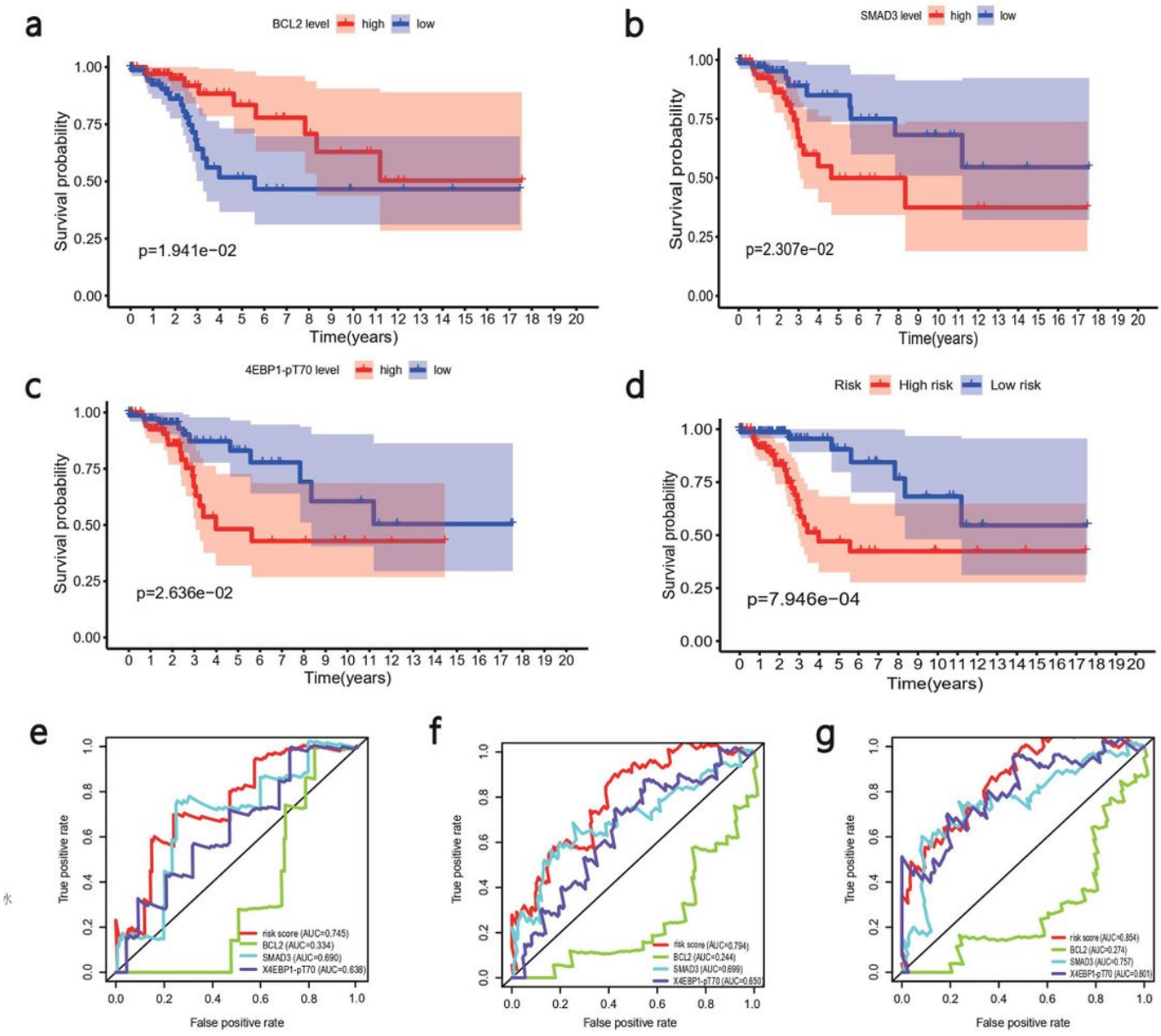

Figure 4

The Kaplan–Meier survival curves of three proteins involved in the signature and risk grade. The K-M curve reveals the relationship between the expression level of BCL2 (A), SMAD3 (B) ,4EBP1-pT70 (C) and OS, using median separation; The K-M curve reveals the relationship between the risk scores (D) and OS, using median separation.1-year (E), 3-year (F), 5-year (G)AUC of BCL2, SMAD3, 4EBP1-pT70 and signature.

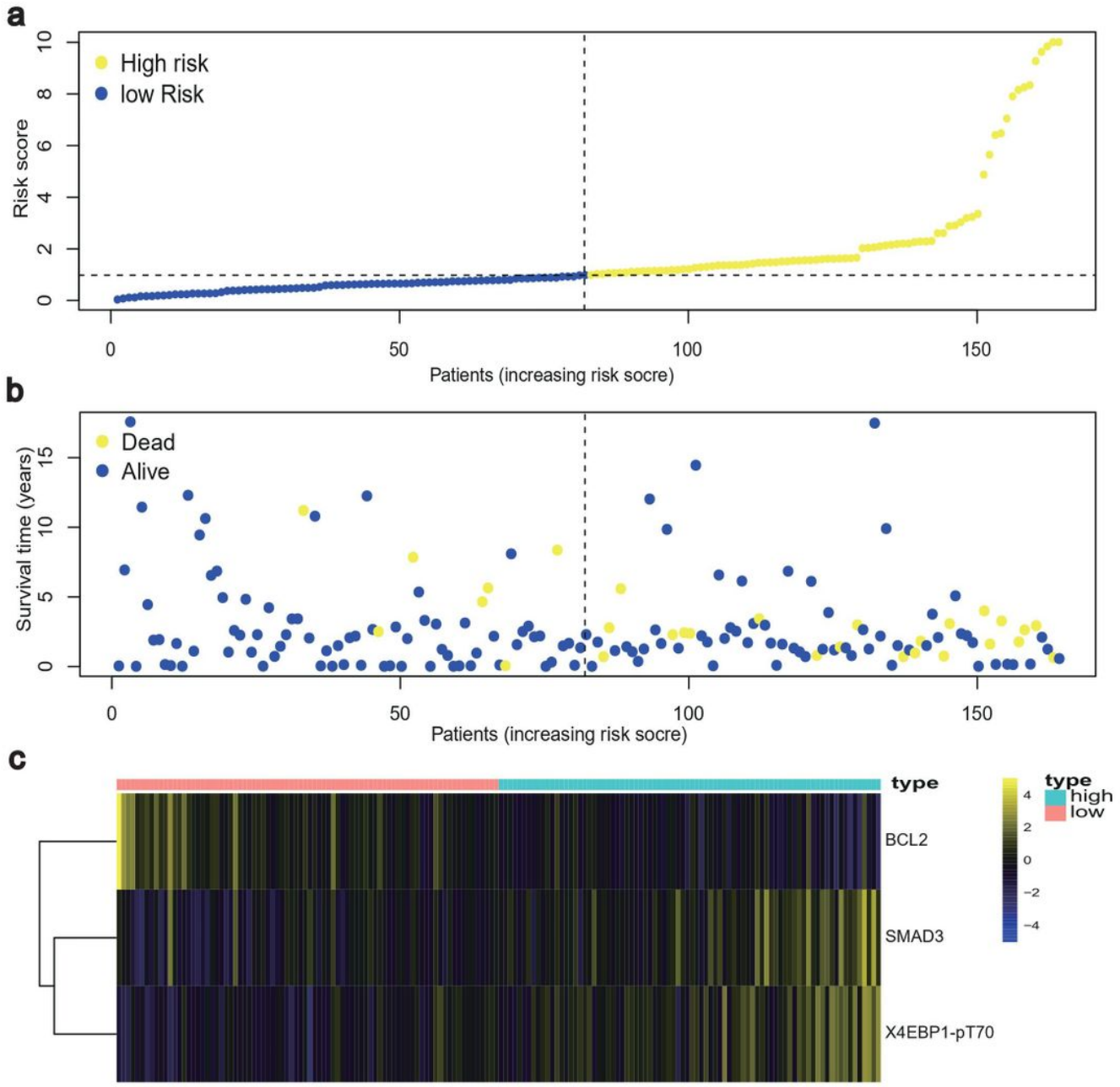

Figure 5

Three-protein risk score analysis. (A) the number of patients in different risk group; (B) the OS of patients in our cohort; (C) the heatmap of the three signature proteins expression profiles.

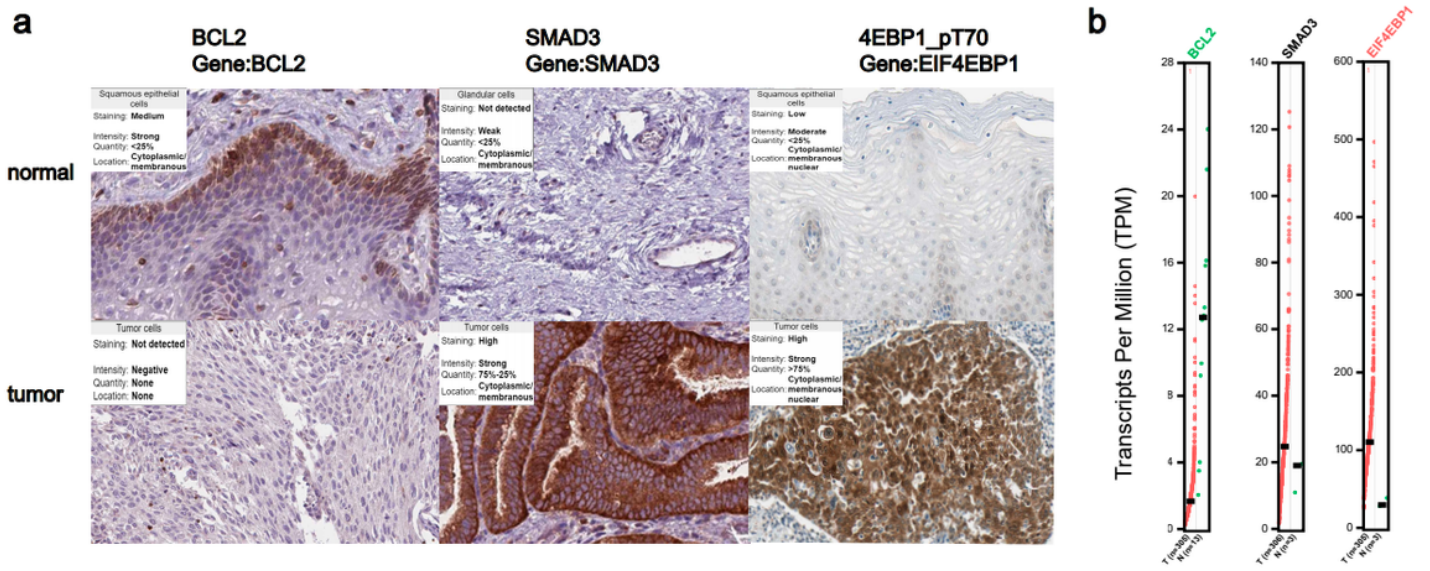

Figure 6

The expression of the three factors in normal tissues and tumor tissues were compared at the protein and gene levels. (A) The immunohistochemical images of proteins associated with risk models. (B) Comparison of expression of risk model corresponding genes in normal and tumor tissues.

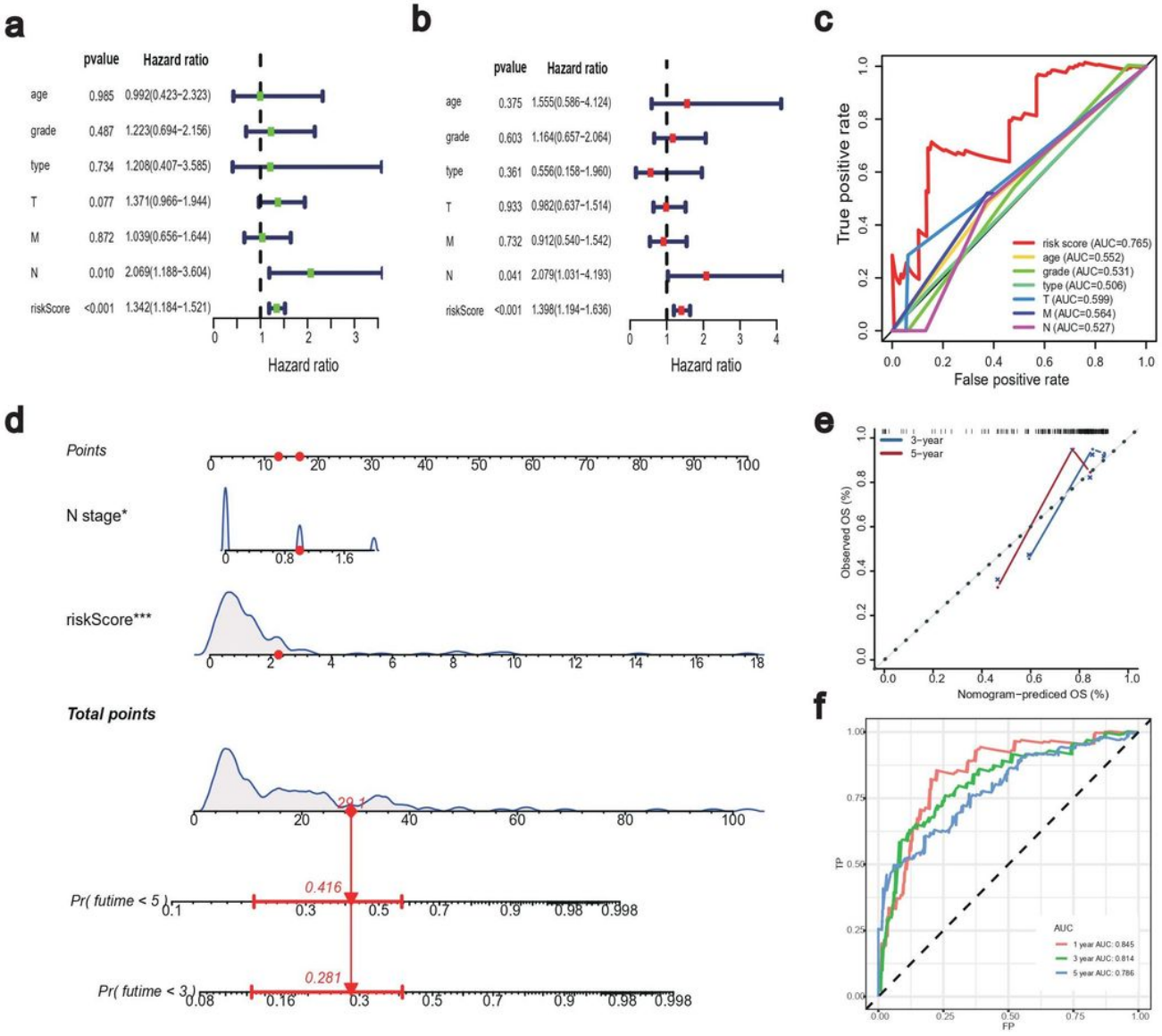

Figure 7

The independent prognostic factors of common CC and construction and validation of the nomogram. (A) The univariate Cox regression analysis, (B) the multivariate Cox regression analysis, (C) the ROC analysis of OS for the risk score and the clinicopathologic parameters. (D) Nomogram integrated protein signature-based PRPscore. (E) The calibration curve of the nomogram in 3-year and 5-year. (F) The ROC curve of the nomogram in 1-year, 3-year, and 5-year.

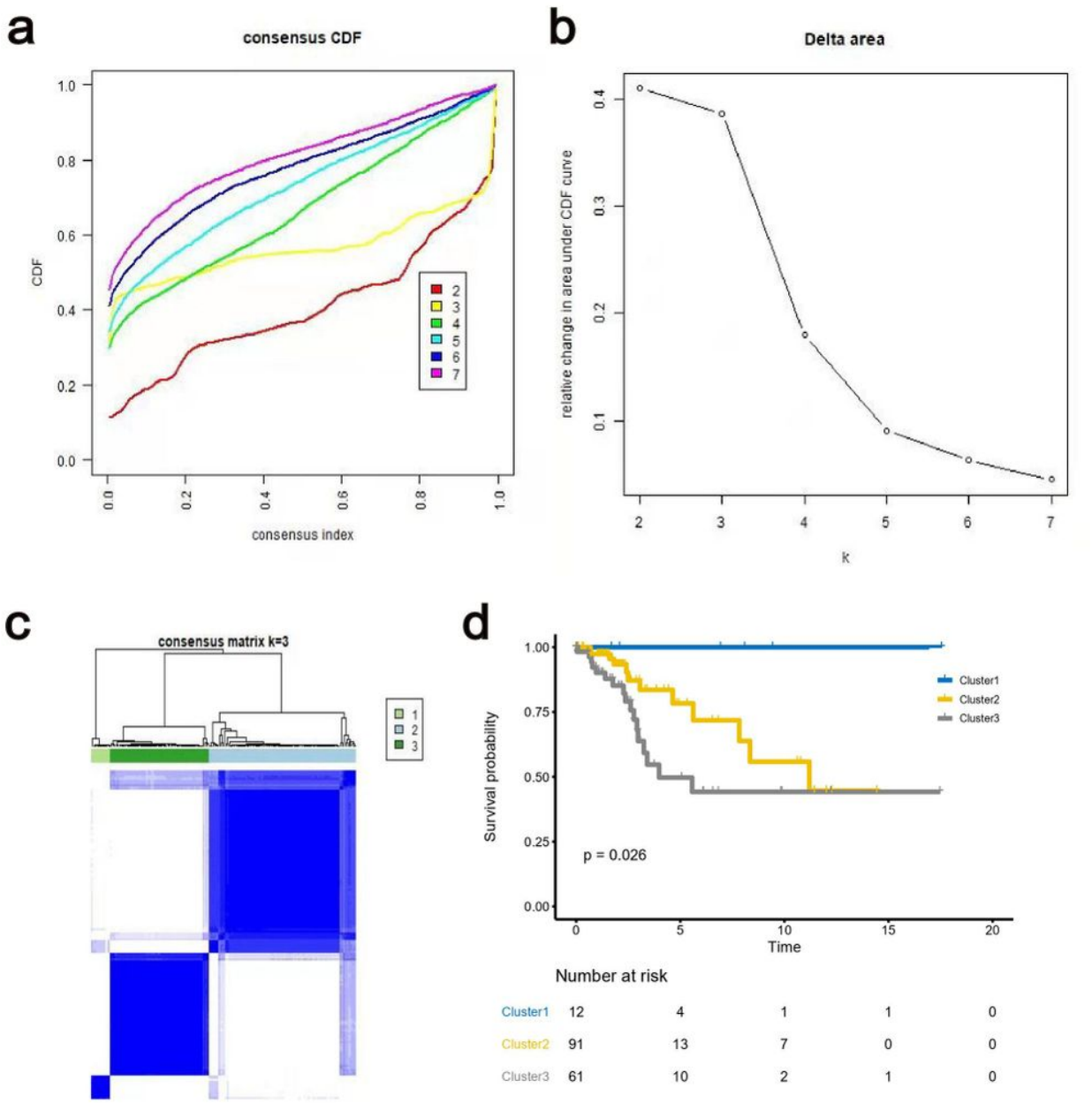

Figure 8

Identification of potential cancer subgroups. (A) Consensus Cumulative Distribution Function Plot. This graph indicated the cumulative distribution functions of the matrix, which had been clustered, for each k. (B) Delta Area Plot. It helped people to see the relative change of area under the curve between k and k-1 to determine the number of k. (C) The Consensus Matrices, which showed the distribution of patients when the cluster number was 6. (D) K–M survival analysis of clusters from consensus clustering analysis, P<0.05 was the cut-off value.

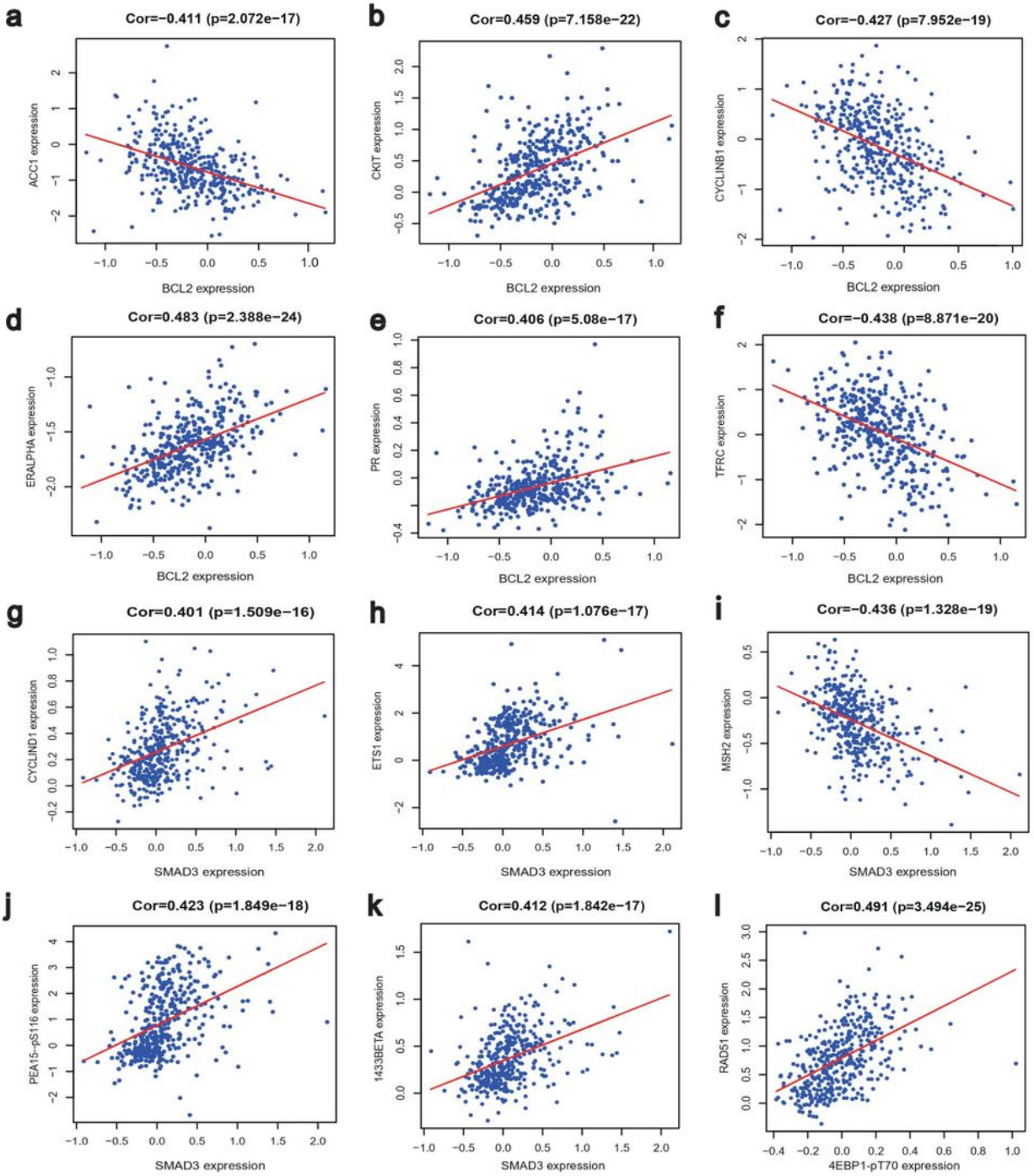

Figure 9

Co-expressed protein of signature proteins. (A-F) Co-expression of ACC1, CKIT, CYCLINB1, ER, PR, TFRC and BCL2, respectively;(G-K) Co-expression of CYCLIND1, ETS1, MSH2, PEA15-pS116, 1433β and SMAD3, respectively;(L) Co-expression of RAD51and 4EBP1-pT70. The direction of the red line to the upper right indicates positive correlation, and to the lower right indicates negative correlation. The absolute value of the correlation coefficient is greater than 0.4.

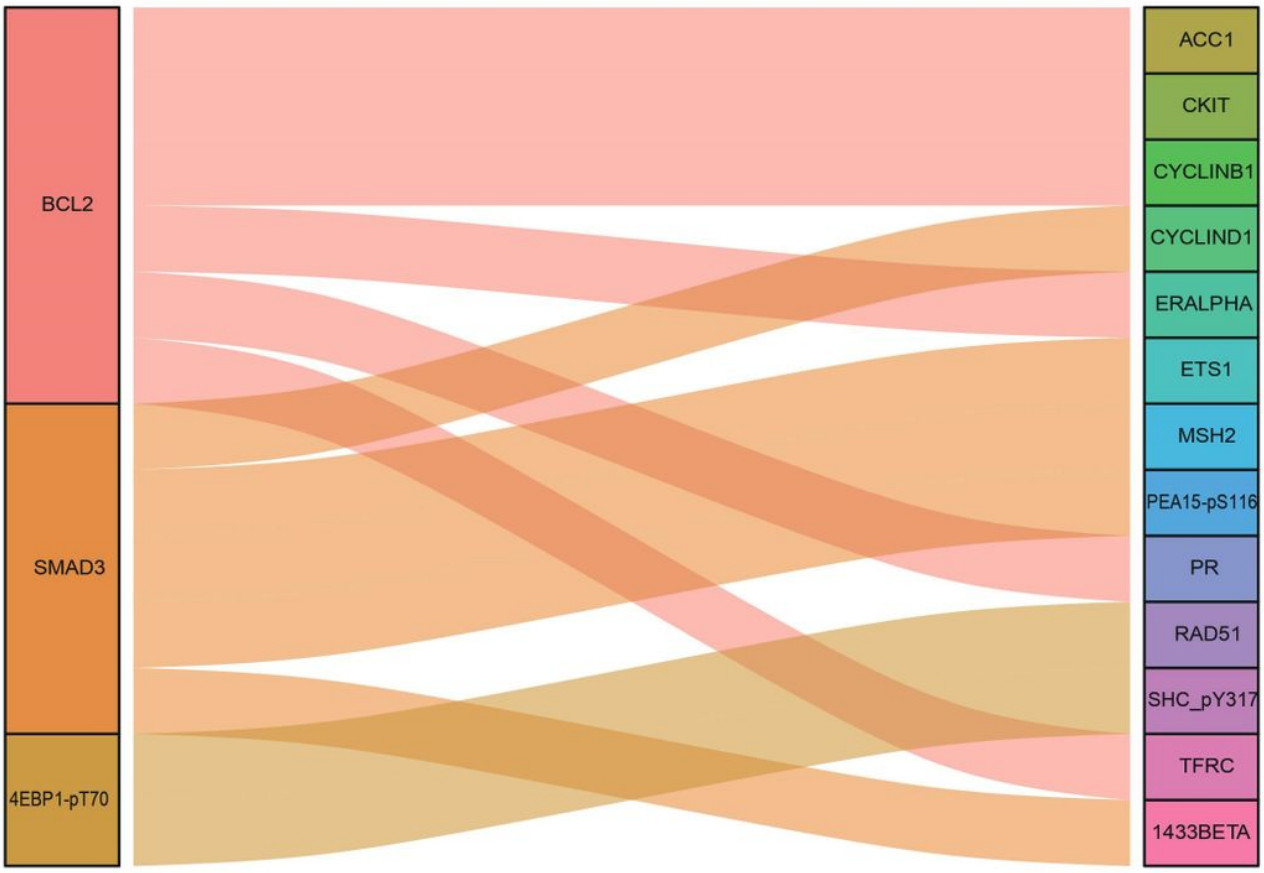

Figure 10

Sankey diagram. Co-expressed protein of signature proteins can show in Sankey diagram. The color of the lines is consistent with the co-expressed proteins in right column.

Supplementary Files

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

- SupplementaryTable1.xlsx
- SupplementaryTable2.xlsx
- SupplementaryTable3.xlsx
- SupplementaryTable4.xlsx