ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Integrative Analysis of Multi-Omic Data for the Characteristics of Endometrial Cancer

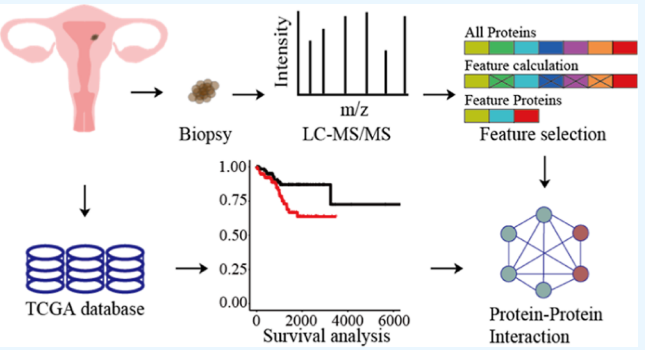
Tong Li, Zhijun Ruan, Chunli Song, Feng Yin, Tuanjie Zhang, Liyun Shi, Min Zuo, Linlin Lu, Yuhao An,* Rui Wang,* and Xiyang Ye*

Cite This: *ACS Omega* 2024, 9, 14489−14499

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Endometrial cancer (EC) is a frequently diagnosed gynecologic cancer. Identifying reliable prognostic genes for predicting EC onset is crucial for reducing patient morbidity and mortality. Here, a comprehensive strategy with transcriptomic and proteomic data was performed to measure EC's characteristics. Based on the publicly available RNA-seq data, death-associated protein kinase 3, recombination signal-binding protein for the immunoglobulin kappa J region, and myosin light chain 9 were screened out as potential biomarkers that affect the EC patients' prognosis. A linear model was further constructed by multivariate Cox regression for the prediction of the risk of being malignant. From further integrative analysis, exosomes were found to have a highly enriched role that might participate in EC occurrence. The findings were validated by qRT-polymerase chain reaction (PCR) and western blotting. Collectively, we constructed a prognostic-gene-based model for EC prediction and found that exosomes participate in EC incidents, revealing significantly promising support for the diagnosis of EC.



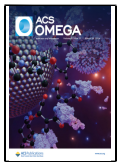
INTRODUCTION

Endometrial cancer (EC) is a commonly diagnosed gynecologic malignancy among women.[1] The precursor lesion for endometrioid adenocarcinoma of the endometrium, which accounts for the majority of endometrial carcinomas, is endometrial hyperplasia (EH). EH is a noninvasive, abnormal proliferation of the endometrial lining of the uterus and is associated with a significant risk of concurrent EC or progression to EC.[2] EH could be further classified into two subtypes during its development: endometrial hyperplasia without atypia (EHA)[3] and endometrial atypical hyperplasia (EAH). EHA is a benign disease without significant somatic genetic changes, along with a significant risk of transforming EC and persistent EH.[4] EAH is regarded as a precancerous condition leading to EC,[5] whose pathological progression is complex and exhibits a multiplicity over time and spatial distribution.[6] It is widely recognized that women with EAH have a higher risk of progressing to EC compared to those with EHA. However, the magnitude of this risk is uncertain.[7]

The rising incidence of EC and its growing population of new diagnoses underscore the severe challenge to women's health. To combat this condition, it is essential to identify reliable prognostic biomarker genes to assist in the risk assessment of malignancy and to inform clinical treatment decisions for EC. Nowadays, high-throughput technologies have made it possible to obtain large-scale information at multidimensional biological expression levels. An integrated analysis that encompasses multiple biological layers, such as transcripts, proteins, or metabolites, provides a comprehensive approach to gain a more detailed understanding of the molecular mechanisms underlying public health and disease.[8]

In this study, we collected the publicly available ribonucleic acid (RNA)-seq data, including 572 tissues (normal = 35, EC = 537). Death-associated protein kinase-3 (DAPK3) and recombination signal binding protein for immunoglobulin kappa J region (RBPJ) were identified as the potential prognostic genes to affect the patients' prognosis. Based on the above two genes, a linear model was constructed to evaluate the risk of being malignant in EC patients. Meanwhile, the biopsies (EHA = 3, EAH = 4, EH = 6, and EC = 7) were collected for proteomic data generation by liquid-chromatogram tandem mass spectrometry (LC-MS/MS). In prognostic gene validation, the DAPK3 protein was found in our proteomics analysis. Through protein−protein interaction analysis, a DAPK3 binding partner called myosin light chain

Received: January 11, 2024
Revised: February 22, 2024
Accepted: February 27, 2024
Published: March 13, 2024



9 (MYL9) was found to significantly affect the EC patients' prognosis, making us speculate that MYL9 might be closely related to EC. Four of the six genes were further validated by qRT-polymerase chain reaction (PCR) and western blotting, including three exosome-related genes (HSPE1, RAB14, and SEC31A) and one prognostic gene (MYL9). To investigate the relationships at both transcriptomic and proteomic levels, we also compared their shared genes and the biological annotations from the significantly changed genes during EC incidents, finding that most exosomal proteins might participate in tumorigenesis. EHA and EAH have similar patterns of feature proteins involved in exosomes, indicating they might carry a higher risk compared with EH. Altogether, the utilization of integrative analysis based on a multiomics strategy can be used to discover the prognostic genes for the clinical diagnosis of EC.

EXPERIMENTAL PROCEDURES

Sample Preparation for MS. After washing with ice-cold phosphate-buffered saline, approximately 50−100 mg of uterine tumors and normal tissues were cut into small pieces. Samples were then homogenized by a homogenizer (Bertin Precellys, France) at 4 °C and processed with the following settings: 8 × 20 s at 5500 rpm, break 30 s. The freeze-crushed tissue samples were precipitated with ice-cold acetone at −20 °C overnight. The protein precipitant was centrifuged at 14,000*g* for 30 min. The tissue was washed separately with 1 mL of acetone and then air-dried on ice. Tissues were softly homogenized separately in 500 μL of lysis buffer (8 M urea, 100 mM Tris, pH 8.0, 1:100 v/v MCE protease inhibitor cocktail). Lysates were precleared by centrifugation at 20,000*g* for 30 min at 4 °C and protein concentrations were determined by a Bradford assay (Thermo Fisher Scientific). Proteins were reduced with 10 mM dithiothreitol for 30 min at 37 °C and subsequently alkylated with 55 mM iodoacetamide for 45 min in the dark. Before digestion, samples were loaded into the Microcon Ultracel YM-30 filtration devices (Millipore, Billerica, MA, USA) and centrifuged at 14,000*g* for 20 min. After three washes in 50 mM ABC, trypsin (Pierce, Thermo Fisher Scientific) solution was added to the filter (enzyme-to-protein ratio 1:100 w/w), and samples were incubated at 37 °C overnight. Peptides were collected by centrifugation, followed by an additional wash with 50 mM ABC. The digestion was stopped by acidifying the solution to a final concentration of 1% (v/v) formic acid. Tryptic peptides were desalted on a C18 SPE and dried for LC-MS/MS analysis.

LC-MS/MS Analyses. All analyses were performed using an EASY-nLC 1200 system (Thermo Fisher Scientific) on an Orbitrap 480. After reconstituting peptides in 20 μL of 0.1% FA, 1 μg of the peptide mixture was injected and loaded directly onto a C18 column (25 cm/75 μm, 2 μm beads, Thermo Fisher Scientific) and separated with a 90 min gradient from 4 to 40% B at 300 nL/min in typically. Parameters are as follows in Full MS/data dependent—MS^2 TopN mode: mass analyzer over m/z range of 350−1500 with a mass resolution of 60,000 (at m/z = 200) in a data-dependent mode, 1.6 m/z isolation window. Twenty of the most intense ions are selected for MS/MS analysis at a resolution of 15,000 using the collision mode of HCD.

Data Acquisition. Tandem MS data were queried against a human database using Proteome Discoverer version 2.4 software (Thermo Fisher Scientific). The normalized abundance of a given protein was calculated from the average area of the three most intense peptide signals. For this software, proteins for which area intensities were below the minimum range or were not detected were assigned an area of zero. For the proteins that were identified by multiple UniProt[9] IDs (release 2023/12/26).

The transcriptome profiling (RNA-seq), which was preprocessed by fragments per kilobase of an exon model per million mapped fragments (FPKM), was from the EC project of the Cancer Genome Atlas (TCGA)[10] database, including 537 tumor samples and 35 normal samples. The analyses were based on the genes expressed in all samples using a cutoff value of FPKM ≥1.

Computational Analyses. Software tools used for this study were available as open source packages under the R v4.2.2 environment, including: "tidyverse"[11] v2.0.0 for basic data operations, such as data cleaning; "missForest"[12] v1.5 for empty value filling; "limma"[13] v3.54.1 for differential expression analysis; "pheatmap"[14] v1.0.12 for plotting all samples' abundance and classification; the DAVID[15] database for terms enrichment, such as KEGG[16] and GO term;[17] "Boruta"[18] v8.0.0 for feature selection; "survival"[19] v3.5-5 and "survminer"[20] v0.4.9 for drawing survival curve and Cox regression; "leaps"[21] v3.1 for best subset regression (BSR); "glmnet"[22] v4.1-7 for lasso regression with 10-fold cross validation; "pathview"[23] v1.40.0 for alterations specific pathways demonstration.

Quantitative Reverse Transcription Polymerase Chain Reaction. Transfer the accurately weighed RNA extraction sample to a liquid nitrogen precooled mortar and grind the tissue with grinding (liquid nitrogen needs to be continuously added to the mortar during the grinding process) until it is ground into powder. Then, an appropriate amount of RNA extraction reagent was added to the powder and mixed well. The above mixture was transferred to a centrifuge tube and thoroughly mixed by repeatedly pipetting and blowing. Let it stand at room temperature for 5 min, then centrifuge at 12,000 rpm at 4 °C for 5 min. Carefully aspirate the supernatant and transfer it into a new centrifuge tube before proceeding with subsequent RNA extraction operations. The process of RNA extraction was done using the Trizol reagent (Accurate Biotechnology, Changsha, China) according to the manufacturer's protocol. Total RNA was quantified by a spectrophotometer (Nano-Drop ND-2000), and 2 μg of total RNA was reverse transcribed to complementary DNA using the reverse transcriptase kit (Accurate Biotechnology, Changsha, China) according to the manufacturer's instructions. The messenger RNA (mRNA) levels of the pS2 gene were detected by RT-PCR using the SYBR Green Premix qPCR Kit (Accurate Biotechnology, Changsha, China) in the Bio-Rad CFX Connect PCR system.

The primer of HSPE1-forward (5′-3′) was CAACAGTAGTCGCTGTTGGA, and that of HSPE1-reverse (5′-3′) was CCTCCATATTCTGGGAGAAGAAC.

The primer of MYL9-forward was (5′-3′) GTCCCAGATCCAGGAGTTTAAG, and that of MYL9-reverse (5′-3′) was CATCATGCCCTCCAGGTATT.

The primer of SEC31A-forward was (5′-3′) GAAGTTGTGATTGCCCAGAATG, and that of SEC31A-reverse (5′-3′) was GCACCAGAAGCTACCAGATTAG.

The primer of MSN-forward was (5′-3′) CCACCTGGCTGAAACTCAATAA, and that of MSN-reverse (5′-3′) was GGACACATCCTCAGGGTAGAA.

https://doi.org/10.1021/acsomega.4c00375
ACS Omega 2024, 9, 14489−14499

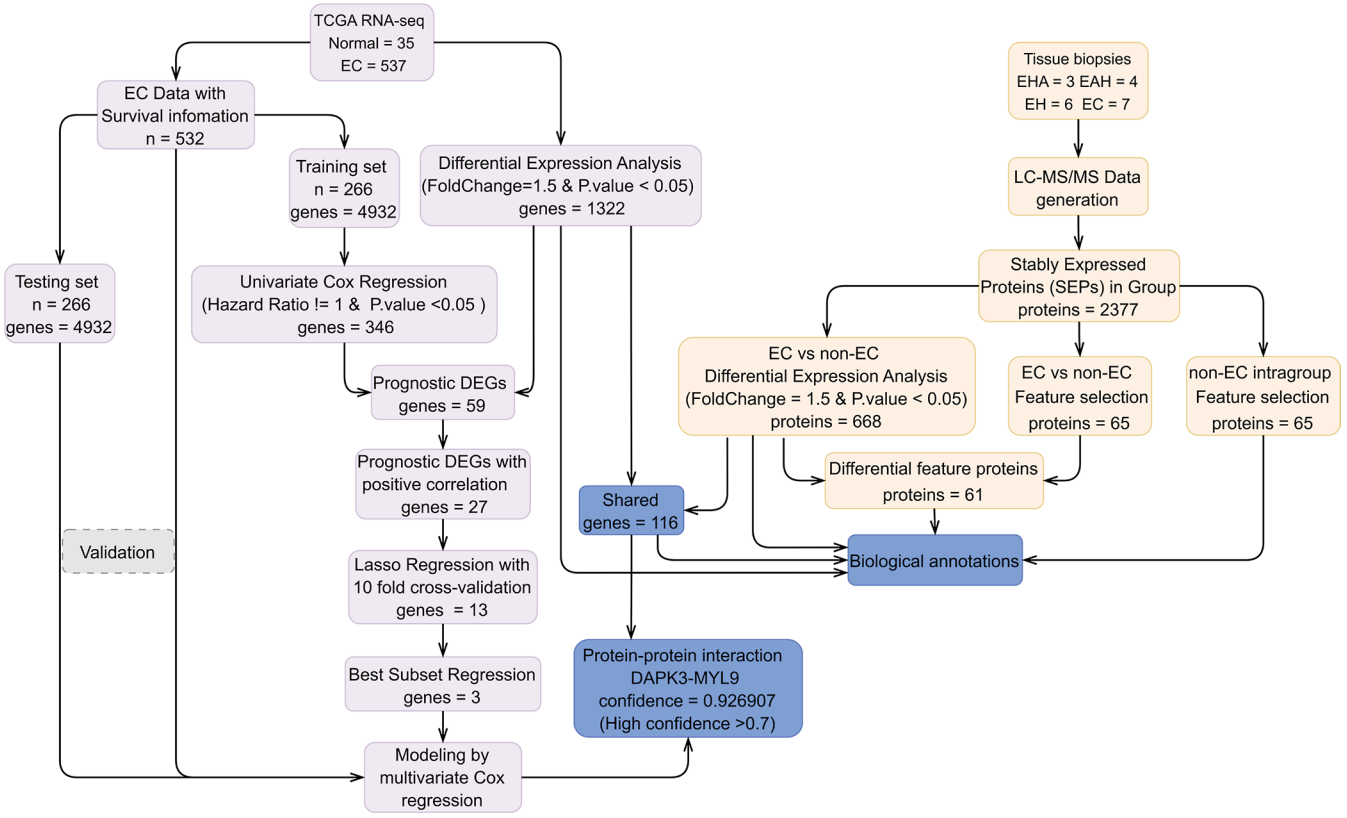


Figure 1. Flowchart of the whole study.

The primer of RAB14-forward was (5′-3′) GGAGCGATTTAGGGCTGTTA, and that of RAB14-reverse (5′-3′) was ACCAGCTGCTTAAGTGGTTAT.

The primer of DAPK3-forward was (5′-3′) CGTTCACTACCTGCACTCTAAGC, and that of DAPK3-reverse (5′-3′) was CCGAAGTCGATGAGCTTGAT.

Western Blotting Analysis. 30 μg of EVP proteins were boiled using 4 × sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer at 95 °C for 10 min, run on 15% SDS-PAGE gels (polyacrylamide gel electrophoresis), and transferred onto poly(vinylidene difluoride) membranes (Millipore, USA). Membranes were sequentially blocked with 5% nonfat milk (w/v) and incubated with primary antibodies, including an Exosome Panel Kit (Abcam, ab275018) and Anti-GAPDH (TransGen Biotech, HC301-01), overnight at 4 °C. After washing three times with 1× TBST, membranes were incubated with antimouse (Cell Signaling, 7076S) or antirabbit (Cell Signaling, 7074S) secondary antibodies for 1 h and washed again to remove unbound antibodies. Bound antibody complexes were imaged by a ChemiDoc Imager (Bio-Rad).

RESULTS AND DISCUSSION

Data Processing and Analysis Strategies. The whole flowchart of this study is shown in Figure 1. We first collected the RNA-seq data of EC from the TCGA[10] database and prepared the expression data and clinical information for the further associated analysis.[24−26] To provide enough data for the model validation, we also divided the whole RNA-seq data set into two equal subsets: the training and the testing data set. The differentially expressed genes (DEGs) between the normal and cancer samples were combined with the prognostic genes screened out by univariate Cox regression for the selection of the genes whose expression is positively correlated with the hazard ratio by the least absolute shrinkage and selection operator (LASSO) regression for biomarker discovery. BSR was performed to reduce the count of independent variables for model simplification. Finally, a linear model was constructed for the risk score prediction and validated by the whole RNA-seq data set and the testing data set.

At the proteomic level, four types of biopsies were collected and calculated by two strategies: one was that the EC and non-EC (EHA/EAH/EH) samples were compared to get the significantly changed proteins along the tumorigenesis; another strategy was that the feature proteins of the three non-EC samples were measured by the random forest algorithm. After biological annotations, some hints were found to indicate the possible conversion from nontumor to tumor, and exosomes were finally found to possibly participate in the conversion of nontumor to tumor.

Discovery of Prognosis-Related Genes. Data mining based on publicly available data plays a vital role in addressing clinical issues. In this study, we collected RNA-seq data from 35 normal individuals and 537 EC patients from the TCGA[10] database. 4932 genes with more than 1 FPKM were identified that were expressed in all samples. To optimize model generalization during the training process, the size of the test set was made equal to that of the training set. According to the screening methods and criteria discussed above, 346 prognosis-related genes (hazard ratio ≠1) were found in the training set (n = 266). To link survival information and genes' abundance, the prognosis-related genes were intersected with the DEGs. 27 genes with positive correlations between abundance and hazard ratio were screened out, including 3 up-regulated DEGs

https://doi.org/10.1021/acsomega.4c00375
ACS Omega 2024, 9, 14489−14499

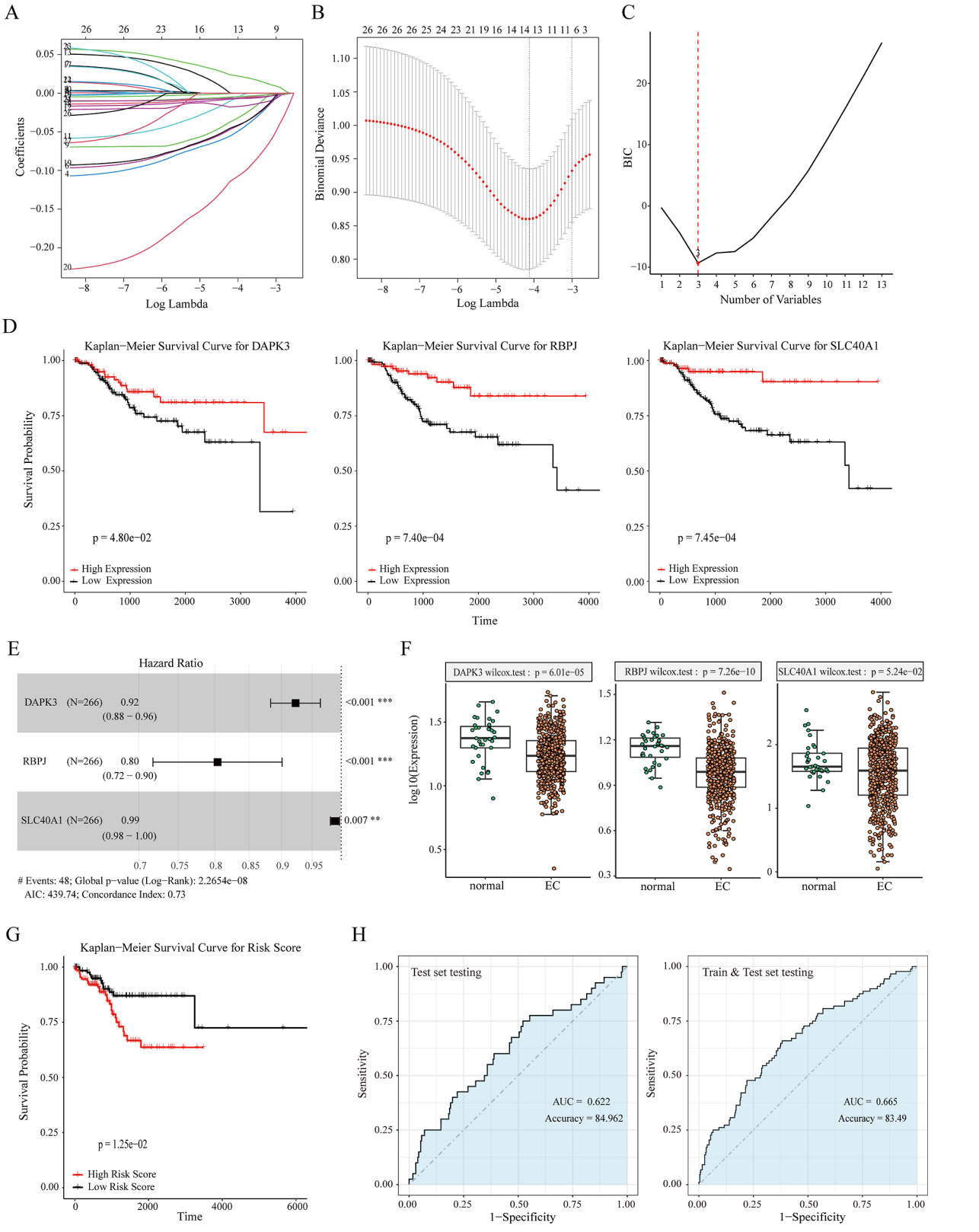


Figure 2. (A) LASSO coefficient profiles of training set genes. (B) LASSO regression with 10-fold cross-validation obtained 13 prognostic genes by minimum lambda value. (C) Best three biomarker candidates from BSR. (D) Kaplan−Meier survival analysis of DAPK3, RBPJ, and SLC40A1. (E) Forest plot of three prognostic genes from the multivariate Cox regression analysis. (F) Expression of the three genes in normal and EC tissues. (G) Kaplan−Meier survival analysis of the risk score model. (H) Model evaluation.

with HR >1 and 24 down-regulated DEGs with HR <1. LASSO regression with 10-fold cross-validation was performed based on the 27 genes to get the optimal lambda value that came from the minimum partial likelihood deviance (Lambda min = 0.015), which was related to 13 potential biomarker candidates (Figure 2A,B). For further analysis by BSR, the

https://doi.org/10.1021/acsomega.4c00375
ACS Omega 2024, 9, 14489−14499

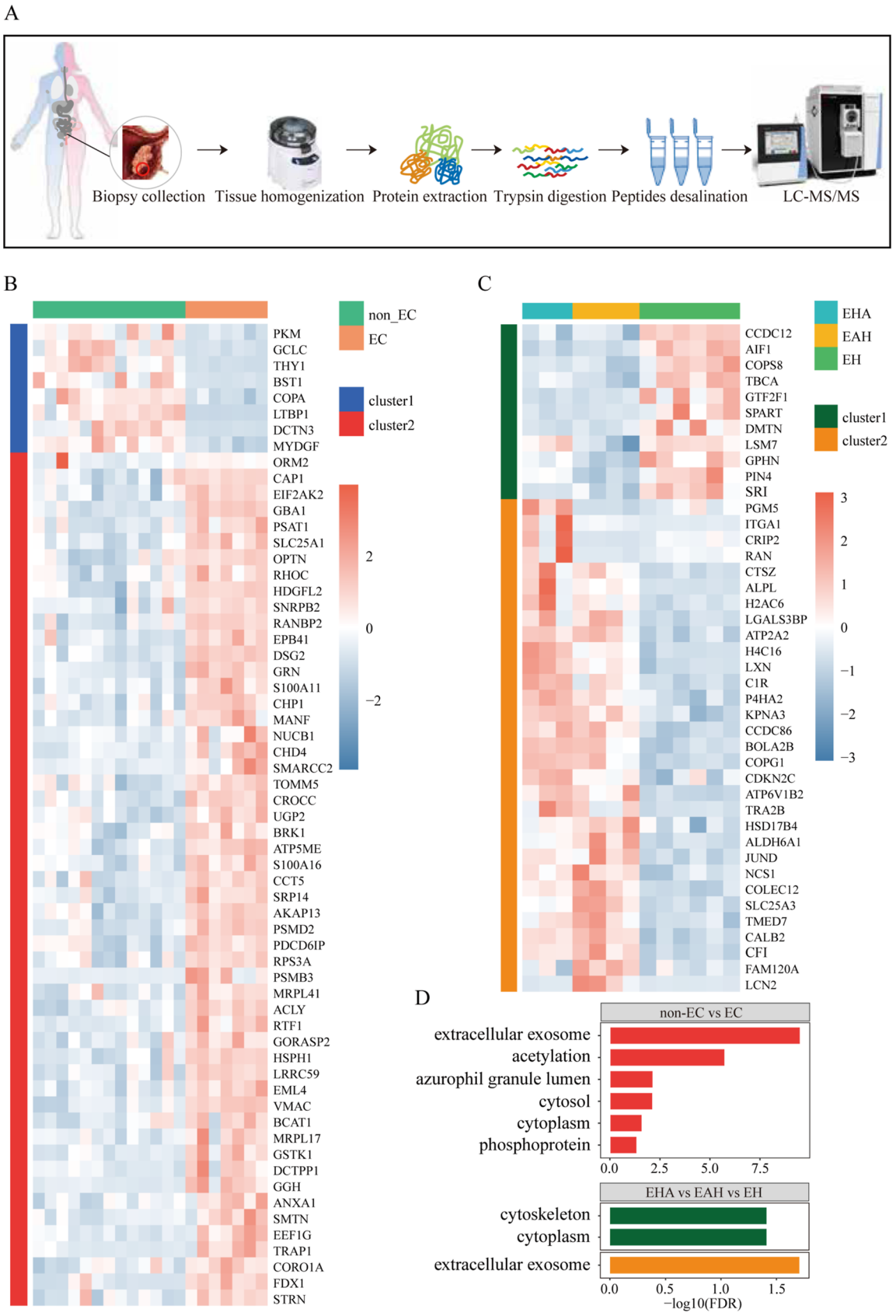


Figure 3. (A) Workflow of proteomic data generation. (B) Patterns of non-EC and EC feature proteins screened out by a random forest algorithm. (C) Patterns of three types of non-EC samples. (D) Biological annotations of the feature proteins.

optimal three prognostic genes were selected according to the lowest value of the Bayesian information criterion in the evaluation process of model performance, namely DAPK3, RBPJ, and SLC40A1 (Figure 2C). Through Kaplan−Meier

https://doi.org/10.1021/acsomega.4c00375
ACS Omega 2024, 9, 14489−14499

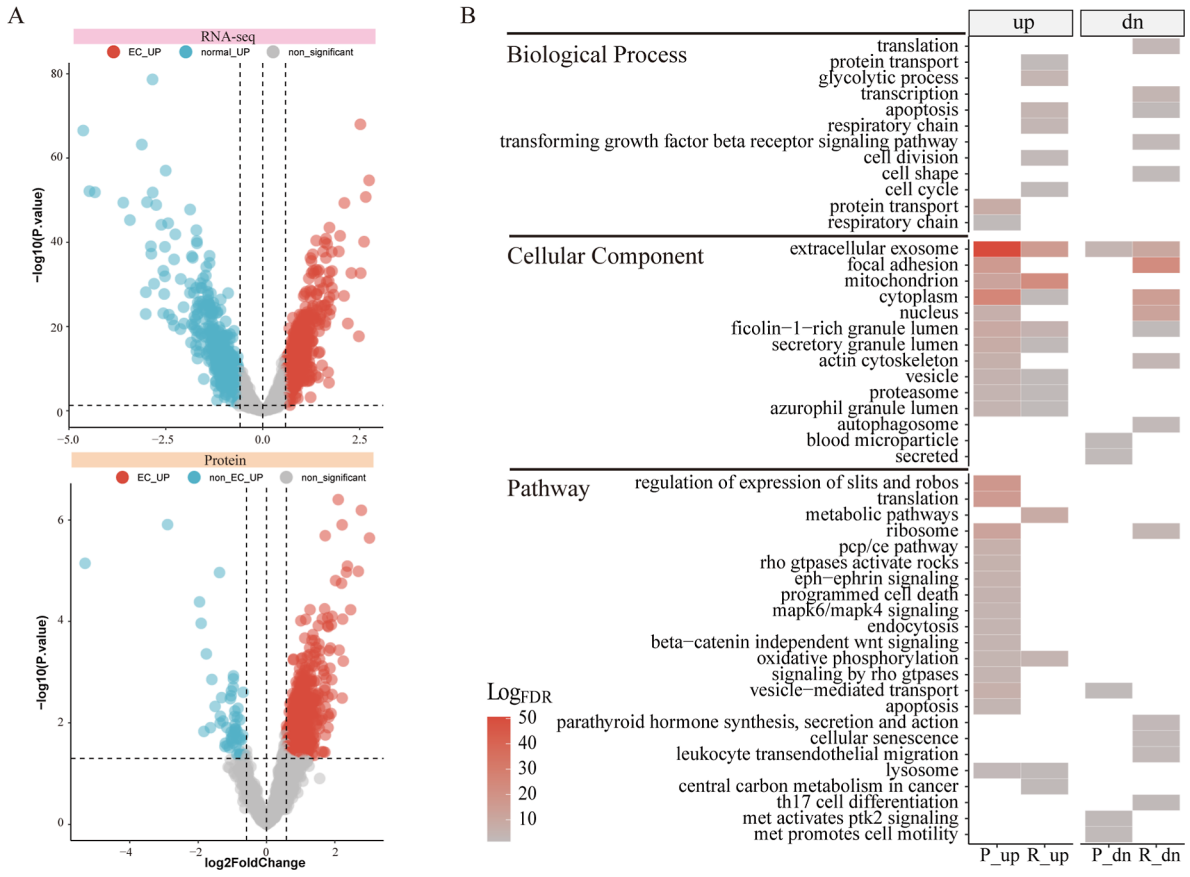


Figure 4. (A) Distributions of DEG products. (B) Biological annotations of DEGs at the transcriptomic and proteomic levels.

analysis, it was found that the expression of genes significantly impacted the patients' prognosis (Figure 2D). The forest plots (Figure 2E) and boxplots (Figure 2F) provided more details about the three prognostic genes. To ascertain the existence of a significant prognostic value for each gene for EC patients, we conducted a multivariate Cox proportional hazards regression analysis and constructed a linear model. After optimizing the model, the two gene-based models for risk score evaluation were established with respective Cox coefficients: risk score = −0.009767733 × Exp (DAPK3) − 0.021571099 × Exp (RBPJ). The risk score of each patient in the testing set was calculated to plot the survival curve, showing that patients with high scores had poorer life conditions compared to those with low scores. The result suggests that the model holds a predictive value in assessing the risk of malignancy (Figure 2G). To validate the performance of the model, both the testing set (n = 266) and the entire data set (n = 532) were used for evaluation. This comprehensive approach ensures a more accurate and reliable assessment of the model's predictive capabilities. Time-dependent ROC analysis showed that the AUC for the testing and the entire set were 0.622 and 0.665, respectively (Figure 2H). Indeed, the model was constructed using only two genes, highlighting the significant role of these two prognostic genes in predicting overall survival among EC patients. Collectively, these genes could potentially be valuable biomarkers for prognosis and might even be targeted for therapeutic interventions.

Feature Protein Selection by Machine Learning. Gene expression has been defined as the "production of an observable phenotype by a gene—usually by directing the synthesis of a protein".[27] To investigate the protein characteristics during tumorigenesis, EC and three types of non-EC biopsies were collected for proteomic data generation using LC-MS/MS (Figure 3A), which is a central analytical technique for protein research and for the study of biomolecules in general. Feature selection is a fundamental step in many machine learning pipelines, which aims to simplify the problem by removing useless features that would introduce unnecessary noise. The genes were treated as features used for the recognition of different types of samples in this study. 61 and 42 feature proteins were identified to distinguish between EC and non-EC types and to discern the intragroup differences among the three types of non-EC, respectively (Figure 3B,C). Interestingly, in the abnormal samples, such as EHA, EAH, and EH, EHA and EAH showed similar patterns compared with those of EH (Figure 3C). As biological annotation showed, the up-regulated proteins in EC were involved in extracellular exosomes and cytoplasm, containing the same functions as the annotations of non-EC feature proteins (Figure 3D). Additionally, the overexpressed proteins in EHA and EAH were the same terms as the EC feature selection, indicating that both of them might carry a higher risk of being malignant than that of EH.

Alterations in Target Pathways during EC. Genetic alterations as well as various cell-signaling pathways have been implicated in EC development and progression, including the MAPK/ERK pathway and Wnt/β-catenin signaling cascades (together with APC/β-catenin signaling). Endometrial tumor tissues have been shown to contain mutations in these signaling pathways, which are generally regarded as the primary drivers of carcinogenesis.[28] MAPK/ERK expression is critical for development, and their hyperactivation plays a

https://doi.org/10.1021/acsomega.4c00375
ACS Omega 2024, 9, 14489−14499

Figure 5. (A) Patterns of shared DEG products. (B) Biological annotations of protein-RNA shared DEGs. (C) Biological annotations of protein-RNA shared DEGs. (D) Expression trend of the shared 116 DEG products at two levels.

well-known role in cancer development and progression.[29] The Wnt/β-catenin signaling pathway is an extremely conserved pathway that is involved in a variety of cellular processes in the female genital system, including development, cell proliferation, cell survival, adhesion, and motility, as well as the regulation of the menstrual cycle.[30,31] Based on the above findings, we measured the gene products at different expression levels. Interestingly, the genes associated with these two pathways have reverse expression patterns (Figure S1). Along with EC occurrence, the proteins were highly expressed, but the RNAs were down-regulated. Although it is widely recognized that gene expression can exhibit inconsistent trends, the inverse expression in specific pathways, which is

https://doi.org/10.1021/acsomega.4c00375
ACS Omega 2024, 9, 14489−14499

Table 1. Fold Change Values of Exosome-Related Genes From Transcriptomic Data

SYMBOL	EC_vs_normal_RNAseq	EC_vs_EH_qRT_PCR
HSPE1	2.1091455	1.536972908
MSN	0.5322263	1.122774077
RAB14	0.6373773	0.251442077
SEC31A	0.6646919	0.001384277

strongly associated with cancer development, still warrants careful consideration.

Integrative Differential Expression Analysis. The advent of omics technologies has generated an ever-growing number of omics datasets that allow researchers to study gene expression at various levels. Multiomics analysis in biomedical research could help explain the complex relationships between molecular layers, improving disease prevention, early detection, and prediction. To investigate the connections between transcriptomic and proteomic levels, we measured the significantly changed genes in tumorigenesis. The biological annotations of DEGs (RNAs = 1322, proteins = 668) were integrated (Figure 4A). The biological annotations were related to cell proliferation, growth, communication, and immigration, indicating that the DEGs have a closely tight connection with cancer occurrence (Figure 4B).

The comparison of normal and EC from RNA-seq data showed similar counts of up- and down-regulated genes. However, at the protein level, overexpressed proteins make up a significant proportion of the differentially expressed proteins shared with the abnormal tissues (Figure 5A,B), which means there were many variations up-regulated along with EC incidents even though compared with the benign non-EC samples. The biological annotations (Figure 5C) were consistent with the cellular component annotations (Figure 4B), demonstrating that exosomes were critical for EC development. There were 116 DEGs shared and expressed at both the protein and RNA levels (Figure 5D). Based on the proteomic data, DAPK3, and its binding partner MYL9 were found. MYL9 had a tight connection with DAPK3 (confidence = 0.926907) and had been identified as a fibroblast-specific biomarker of a poorer prognosis for colorectal cancer.[32] Low MYL9 abundance also was reported a significant association with the development of nonsmall-cell lung cancer.[33] Thus, we speculated that MYL9 could also be a biomarker candidate for EC.

Validation of Exosome-Related and Prognostic Genes. To substantiate the insights resulting from omics data analysis, it is essential to furnish solid evidence. We selected six genes associated with exosomes and prognosis, including RAB14, SEC31A, HSPE1, MSN, DAPK3, and MYL9, and measured their expression by qRT-PCR and western blotting.

Heat shock protein family E (Hsp10) member 1 (HSPE1) is usually used together with the cochaperonin heat shock protein 60 (Hsp60) to maintain protein homeostasis.[34] Hsp60 is well-known as being related to exosomes, so we speculated that HSPE1 might participate in exosome activity and cooperate with Hsp60. Moesin (MSN) is part of the ezrin, radixin, and moesin protein families. These proteins play crucial roles in linking the plasma membrane to the actin cytoskeleton. MSN has been reported as a potential exosomal protein biomarker for bladder cancer.[35] Ras-related protein Rab-14 (RAB14) is a small GTPase involved in the regulation of intracellular membrane trafficking and vesicle transport, which plays a crucial role in vesicular transport within cells, especially during vesicular transport from the Golgi apparatus to the cell surface.[36] There were reports showing RAB14 protein expression was positively correlated with increased tumor size,[37] and its activity could regulate exosome secretions and cell growth.[38] Protein transport protein Sec31A (SEC31A) is a core component of the COPII coat complex, which is involved in the formation of transport vesicles from the endoplasmic reticulum (ER), contributing to the efficient transport of proteins from the ER to the Golgi apparatus. This gene is essential for maintaining proper protein trafficking and secretion within the cell.[39] DAPK3 was a potential prognostic gene found in RNA-seq data mining, which is a serine/threonine kinase that belongs to the DAP kinase family. Members of this family serve as crucial regulators of cell apoptosis. DAPK3 plays a pivotal role in diverse biological processes such as cell apoptosis, autophagy, cytoskeleton remodeling, and immune response.[40−42] As a DAPK3 binding partner, MYL9 was widely reported to be overexpressed in several colorectal cancer cell lines, promoting cell proliferation, invasion, migration, and angiogenesis, while silencing MYL9 exerted the opposite effects.[43]

In Table 1, the four genes related to exosomes have the same expression trend in RNA-seq data and qRT-PCR, except for MSN. Table 2 and Figure 6A show the results at the RNA level are consistent. High expression of MYL9 mRNA was associated with worse survival status in EC patients (Figure 6B). MYL9 was validated at the protein level by western blotting (Figure 6C). The result of univariate Cox regression was not significant (HR = 1, p-value = 0.2), suggesting that additional factors may synergize with MYL9 to worsen the prognosis of EC patients. Overall, the experiments supported our findings based on the omics data, even though the comparisons were normal versus EC and EH versus EC.

CONCLUSIONS

In summary of our study, DAPK3, RBPJ, and MYL9 have been identified as potential prognostic biomarkers for EC, and a linear model has been developed based on these findings to predict the patients' prognosis. To describe the characteristics of EC at different levels, differential expression analysis and machine learning algorithms were adopted, finding that extracellular exosomes might be associated with EC development. The biological pathways named the "MAPK signaling pathway" and "WNT signaling pathway" are well-known and involved in tumorigenesis, and the DEG products in the two pathways have inverse trends in this study, demonstrating that there are potential regulation networks that exist among multibiolayers depending on this way. Three exosome-related

Table 2. Fold Change Values of Potential Biomarkers at Different Levels

SYMBOL	EC_vs_normal_RNAseq	EC_vs_EH_RT_qRT_PCR	EC_vs_EH_Protein	EC_vs_EH_WB
DAPK3	0.65795354	0.007340118	1.190202	0.515297896
MYL9	0.08244865	0.411552033	2.216607	1.12696007

https://doi.org/10.1021/acsomega.4c00375
ACS Omega 2024, 9, 14489−14499

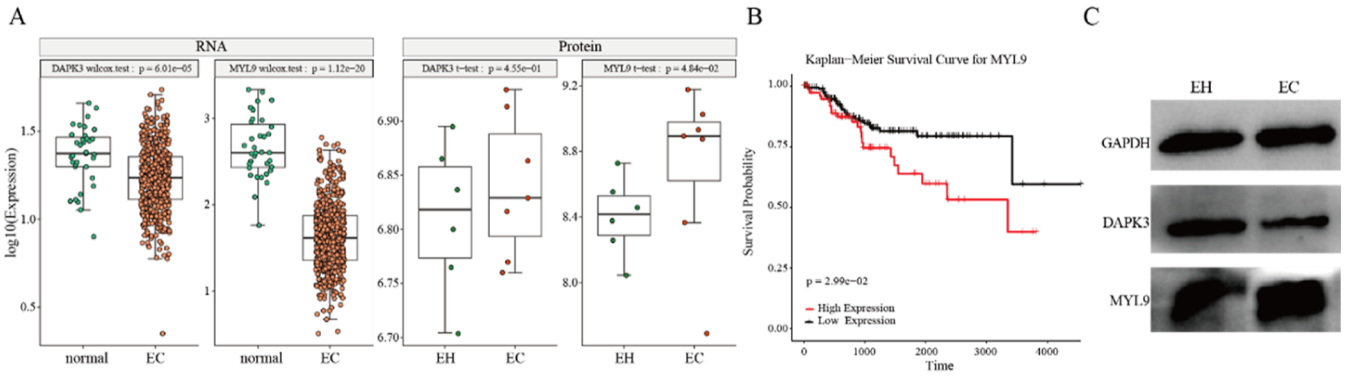


Figure 6. (A) Expression of DAPK3 and MYL9 in omic data. (B) Kaplan–Meier survival analysis of MYL9. (C) Western blotting of two potential biomarkers.

genes (HSPE1, RAB14, and SEC31A) and one prognostic gene (MYL9) were validated by wet experiments, supporting the omic findings resulting above. Altogether, based on the approach of integrating multiomics data from the transcriptome and proteome, we provided a landscape for a better understanding of EC incidents.

■ ASSOCIATED CONTENT

Data Availability Statement

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the iProX partner repository with the data set identifier PXD045217.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.4c00375.

Alterations of gene products in MAPK and WNT signaling pathways during EC incidents (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Yuhao An − *Shenzhen Bay Laboratory, Pingshan Translational Medicine Center, Shenzhen 518118, China*; orcid.org/0009-0006-0224-9709; Email: anyh@szbl.ac.cn

Rui Wang − *Shenzhen Bay Laboratory, Pingshan Translational Medicine Center, Shenzhen 518118, China*; Email: wangrui@szbl.ac.cn

Xiyang Ye − *Department of Gynecology, Shenzhen People's Hospital, Shenzhen, Guangdong 518020, China*; Email: szyxy2009@qq.com

Authors

Tong Li − *Department of Gynecology, Shenzhen People's Hospital, Shenzhen, Guangdong 518020, China*

Zhijun Ruan − *Shenzhen Bay Laboratory, Pingshan Translational Medicine Center, Shenzhen 518118, China*

Chunli Song − *School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen 518055, China*

Feng Yin − *School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen 518055, China*

Tuanjie Zhang − *Shenzhen Bay Laboratory, Pingshan Translational Medicine Center, Shenzhen 518118, China*

Liyun Shi − *Department of Gynecology, Shenzhen People's Hospital, Shenzhen, Guangdong 518020, China*

Min Zuo − *Department of Pathology, Shenzhen People's Hospital, Shenzhen, Guangdong 518020, China*

Linlin Lu − *International Institute for Translational Chinese Medicine, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong 510006, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c00375

Notes

The authors declare no competing financial interest.

Ethics Statement: All experimental protocols were approved by the Ethics Committee of Shenzhen People's Hospital and were carried out in accordance with the approved guidelines. This study adhered to Measures for Ethical Review of Life Science and Medical Research Involving Human Beings (Document Number: no. 4 [2023] of the National Health Commission), as deliberated and adopted by the National Science and Technology Ethics Committee. For studies involving human subjects, written informed consent was obtained from all participants, and all procedures were in line with the ethical standards of the Shenzhen People's Hospital Institutional Review Board and with the Helsinki Declaration. For research involving human subjects, we ensured that each participant was fully informed about the purpose, procedures, potential risks, and benefits of the study and that their participation was voluntary. Measures to protect the privacy and personal information on the subjects were strictly implemented.

■ ACKNOWLEDGMENTS

This work was supported by two grants from the Shenzhen Science and Technology Program (JCYJ20200109140406047 and JCYJ20230807151059004), two grants from the National Natural Science Foundation of China (21977010 and 22307084), and a grant from the Shenzhen High-tech Zone Development Special Plan Pingshan District Innovation Platform Construction Project (29853MKCJ202300208). This work is supported by the Proteomic Platform of the Pingshan Translational Medicine Center, Shenzhen Bay Laboratory.

■ REFERENCES

(1) Sung, H.; Ferlay, J.; Siegel, R. L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide

for 36 Cancers in 185 Countries. *CA Cancer J. Clinicians* **2021**, *71* (3), 209−249.
(2) Nees, L. K.; Heublein, S.; Steinmacher, S.; Juhasz-Böss, I.; Brucker, S.; Tempfer, C. B.; Wallwiener, M. Endometrial hyperplasia as a risk factor of endometrial cancer. *Arch. Gynecol. Obstet.* **2022**, *306* (2), 407−421.
(3) Sobczuk, K.; Sobczuk, A. New classification system of endometrial hyperplasia WHO 2014 and its clinical implications. *Menopausal Rev.* **2017**, *3* (3), 107−111.
(4) Ren, H.; Zhang, Y.; Duan, H. Recent advances in the management of postmenopausal women with non-atypical endometrial hyperplasia. *Climacteric* **2023**, *26*, 411−418.
(5) Sanderson, P. A.; Critchley, H. O. D.; Williams, A. R. W.; Arends, M. J.; Saunders, P. T. K. New concepts for an old problem: the diagnosis of endometrial hyperplasia. *Hum. Reprod. Update* **2016**, *23* (2), 232−254.
(6) Zhang, H.; Kong, W.; Han, C.; Liu, T.; Li, J.; Song, D. Correlation of Metabolic Factors with Endometrial Atypical Hyperplasia and Endometrial Cancer: Development and Assessment of a New Predictive Nomogram. *Cancer Manage. Res.* **2021**, *13*, 7937−7949.
(7) Koskas, M.; Uzan, J.; Luton, D.; Rouzier, R.; Daraï, E. Prognostic factors of oncologic and reproductive outcomes in fertility-sparing management of endometrial atypical hyperplasia and adenocarcinoma: systematic review and meta-analysis. *Fertil. Steril.* **2014**, *101* (3), 785−794.e3.
(8) Davey Smith, G.; Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*. *Int. J. Epidemiol.* **2003**, *32* (1), 1−22.
(9) Bateman, A.; Martin, M.-J.; Orchard, S.; Magrane, M.; Ahmad, S.; Alpi, E.; Bowler-Barnett, E. H.; Britto, R.; Bye-A-Jee, H.; Cukura, A.; The UniProt Consortium; et al. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **2022**, *51* (D1), D523−D531.
(10) Tomczak, K.; Czerwińska, P.; Wiznerowicz, M. Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Współczesna Onkologia* **2015**, *1A*, 68−77.
(11) Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; McGowan, L.; François, R.; Grolemund, G.; Hayes, A.; Henry, L.; Hester, J.; et al. Welcome to the Tidyverse. *J. Open Source Software* **2019**, *4* (43), 1686.
(12) Stekhoven, D. J.; Bühlmann, P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* **2012**, *28* (1), 112−118.
(13) Ritchie, M. E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C. W.; Shi, W.; Smyth, G. K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43* (7), No. e47.
(14) Kolde, R.; Kolde, M. R. *Package "pheatmap"*, R Package, 2018. *1*.
(15) Huang, D. W.; Sherman, B. T.; Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4* (1), 44−57.
(16) Kanehisa, M.; Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28* (1), 27−30.
(17) Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **2004**, *32* (90001), D258−D261.
(18) Kursa, M. B.; Rudnicki, W. R. Feature selection with the Boruta package. *J. Stat. Software* **2010**, *36*, 1−13.
(19) Therneau, T. *A Package for Survival Analysis in R*, 2024.
(20) Kassambara, A.; Kosinski, M.; Biecek, P.; Fabian, S. *Package "survminer"*, Drawing Survival Curves using 'ggplot2' (R package version 03 1), 2017. *3*.
(21) Lumley, T. *Package "leaps"*, Regression subset selection. Thomas Lumley Based on Fortran Code by Alan Miller, 2013. http://CRAN.R-project.org/package=leaps (accessed Mar 18, 2018).
(22) Friedman, J.; Hastie, T.; Tibshirani, R.; Narasimhan, B.; Tay, K.; Simon, N.; Qian, J.; Yang, J. *Package "glmnet"*, CRAN R Repositary, 2021.
(23) Luo, W.; Brouwer, C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* **2013**, *29* (14), 1830−1831.
(24) Ling, D.; Liu, A.; Sun, J.; Wang, Y.; Wang, L.; Song, X.; Zhao, X. Integration of IDPC Clustering Analysis and Interpretable Machine Learning for Survival Risk Prediction of Patients with ESCC. *Interdiscip. Sci.: Comput. Life Sci.* **2023**, *15* (3), 480−498.
(25) Xu, R.; Wang, J.; Zhu, Q.; Zou, C.; Wei, Z.; Wang, H.; Ding, Z.; Meng, M.; Wei, H.; Xia, S.; Wei, D.; Deng, L.; Zhang, S. Integrated models of blood protein and metabolite enhance the diagnostic accuracy for Non-Small Cell Lung Cancer. *Biomarker Res.* **2023**, *11* (1), 71.
(26) Sindhoo, A.; Sipy, S.; Khan, A.; Selvaraj, G.; Alshammari, A.; Casida, M. E.; Wei, D.-Q. ESOMIR: a curated database of biomarker genes and miRNAs associated with esophageal cancer. *Database* **2023**, *2023*, baad063.
(27) Buccitelli, C.; Selbach, M. mRNAs, proteins and the emerging principles of gene expression control. *Nat. Rev. Genet.* **2020**, *21* (10), 630−644.
(28) Levine, D. Integrated genomic characterization of endometrial carcinoma. *Nature* **2013**, *497* (7447), 67−73.
(29) Guo, Y.-J.; Pan, W.-W.; Liu, S.-B.; Shen, Z.-F.; Xu, Y.; Hu, L.-L. ERK/MAPK signalling pathway and tumorigenesis (Review). *Exp. Ther. Med.* **2020**, *19* (3), 1997−2007.
(30) Zhan, T.; Rindtorff, N.; Boutros, M. Wnt signaling in cancer. *Oncogene* **2017**, *36* (11), 1461−1473.
(31) Fatima, I.; Barman, S.; Rai, R.; Thiel, K. W.; Chandra, V. Targeting Wnt Signaling in Endometrial Cancer. *Cancers* **2021**, *13* (10), 2351.
(32) Zhou, Y.; Bian, S.; Zhou, X.; Cui, Y.; Wang, W.; Wen, L.; Guo, L.; Fu, W.; Tang, F. Single-Cell Multiomics Sequencing Reveals Prevalent Genomic Alteration s in Tumor Stromal Cells of Human Colorectal Cancer. *Cancer Cell* **2020**, *38* (6), 818−828.e5.
(33) Tan, X.; Chen, M. MYLK and MYL9 expression in non-small cell lung cancer identified by bioinformatics analysis of public expression data. *Tumor Biol.* **2014**, *35* (12), 12189−12200.
(34) Caruso Bavisotto, C.; Alberti, G.; Vitale, A. M.; Paladino, L.; Campanella, C.; Rappa, F.; Gorska, M.; Conway de Macario, E.; Cappello, F.; Macario, A. J. L.; Marino Gammazza, A. Hsp60 Post-translational Modifications: Functional and Pathological Consequences. *Front. Mol. Biosci.* **2020**, *7*, 95.
(35) Li, X.-X.; Yang, L.-X.; Wang, C.; Li, H.; Shi, D.-S.; Wang, J. The Roles of Exosomal Proteins: Classification, Function, and Applications. *Int. J. Mol. Sci.* **2023**, *24* (4), 3061.
(36) Junutula, J. R.; De Maziére, A. M.; Peden, A. A.; Ervin, K. E.; Advani, R. J.; van Dijk, S. M.; Klumperman, J.; Scheller, R. H. Rab14 is involved in membrane trafficking between the Golgi complex and endosomes. *Mol. Biol. Cell* **2004**, *15* (5), 2218−2229.
(37) Ramírez-Torres, A.; Gil, J.; Contreras, S.; Ramírez, G.; Valencia-González, H. A.; Salazar-Bustamante, E.; Gómez-Caudillo, L.; García-Carranca, A.; Encarnación-Guevara, S. Quantitative Proteomic Analysis of Cervical Cancer Tissues Identifies Proteins Associated With Cancer Progression. *Cancer Genomics Proteomics* **2022**, *19* (2), 241−258.
(38) Maziveyi, M.; Dong, S.; Baranwal, S.; Mehrnezhad, A.; Rathinam, R.; Huckaba, T. M.; Mercante, D. E.; Park, K.; Alahari, S. K. Exosomes from Nischarin-Expressing Cells Reduce Breast Cancer Cell Motility and Tumor Growth. *Cancer Res.* **2019**, *79* (9), 2152−2166.
(39) Stagg, S. M.; LaPointe, P.; Razvi, A.; Gürkan, C.; Potter, C. S.; Carragher, B.; Balch, W. E. Structural basis for cargo regulation of COPII coat assembly. *Cell* **2008**, *134* (3), 474−484.
(40) Bialik, S.; Kimchi, A. The death-associated protein kinases: structure, function, and beyond. *Annu. Rev. Biochem.* **2006**, *75*, 189−210.

https://doi.org/10.1021/acsomega.4c00375
ACS Omega 2024, 9, 14489−14499

(41) Deiss, L. P.; Feinstein, E.; Berissi, H.; Cohen, O.; Kimchi, A. Identification of a novel serine/threonine kinase and a novel 15-kD protein as potential mediators of the gamma interferon-induced cell death. *Genes Dev.* **1995**, *9* (1), 15−30.
(42) Shohat, G.; Spivak-Kroizman, T.; Cohen, O.; Bialik, S.; Shani, G.; Berrisi, H.; Eisenstein, M.; Kimchi, A. The pro-apoptotic function of death-associated protein kinase is controlled by a unique inhibitory autophosphorylation-based mechanism. *J. Biol. Chem.* **2001**, *276* (50), 47460−47467.
(43) Feng, M.; Dong, N.; Zhou, X.; Ma, L.; Xiang, R. Myosin light chain 9 promotes the proliferation, invasion, migration and angiogenesis of colorectal cancer cells by binding to Yes-associated protein 1 and regulating Hippo signaling. *Bioengineered* **2022**, *13* (1), 96−106.

https://doi.org/10.1021/acsomega.4c00375
ACS Omega 2024, 9, 14489−14499