

RESEARCH

Open Access



SIM2, associated with clinicopathologic features, promotes the malignant biological behaviors of endometrial carcinoma cells

Hua Nie^{1,2} and Yu Chen^{1,2*}

Abstract

Background Endometrial carcinoma (EC) poses a significant threat to women's health. Identifying effective prognostic biomarkers and therapeutic targets is essential for improving survival rates in EC patients. This study aimed to identify key regulators involved in EC progression and investigate the biological functions of SIM bHLH transcription factor 2 (SIM2) in EC.

Methods Gene expression profiles and clinical data from EC and control samples were retrieved from the TCGA and GEO databases. Differential expression analysis and weighted gene co-expression network analysis (WGCNA) were used to identify genes associated with EC tumorigenesis and progression. The least absolute shrinkage and selection operator (LASSO) method was applied to further screen prognostic genes and construct a prognostic risk model. The expression and biological function of SIM2 were analyzed using the GEPIA, HPA, and LinkedOmics databases. SIM2 knockdown and overexpression models were established in EC cell lines, and their function was validated through qRT-PCR, CCK-8, flow cytometry, and western blot. Additionally, an in vivo lung/liver metastasis model was employed to further validate the cancer-promoting properties of SIM2 in EC.

Results WGCNA identified 343 EC-related genes. Cox regression analysis and LASSO were further applied to identify 13 prognostic genes, leading to the development of a robust prognostic risk model that effectively predicted EC patients' clinical outcomes. Significant differences in the tumor immune microenvironment were observed between the high- and low-risk groups. Among these 13 genes, SIM2 was significantly overexpressed in EC tissues, and its high expression was associated with poor prognosis in EC patients. SIM2 depletion inhibited EC cell viability, induced cell cycle arrest, and promoted apoptosis. Additionally, SIM2 knockdown increased the expression of cleaved caspase-3 and reduced the levels of Cyclin D1 and CDK4 proteins, while SIM2 overexpression showed the opposite effects. In vivo, silencing SIM2 notably suppressed the metastatic potential of EC cells.

Conclusion SIM2 serves as both a biomarker and a therapeutic target for EC diagnosis and prognosis prediction, which positively modulates the malignant phenotypes of EC cells.

Keywords Endometrial carcinoma, Prognosis, WGCNA, LASSO, SIM2

*Correspondence:

Yu Chen

hustchenyu11@163.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Endometrial carcinoma (EC) is one of the most prevalent gynecological malignancies, with both its global incidence and mortality rates on the rise [1], and its age of onset gradually decreasing [2]. According to GLOBOCAN 2020, there were 417,000 new cases of EC in 2020, and 97,000 women died from this disease [3]. Several risk factors contribute to the development of EC, including obesity, diabetes, estrogen use, and polycystic ovary syndrome [4–6]. Elucidating the molecular mechanisms behind EC tumorigenesis and progression, as well as identifying effective diagnostic markers and therapeutic targets, is crucial for improving the prognosis of EC patients.

SIM bHLH transcription factor 2 (SIM2) is a transcription factor that is highly expressed in neurons and plays a key biological role in stress response and cell development [7, 8]. Deficient SIM2 is considered to be associated with Down syndrome [7]. It is also reported that the polymorphism of *SIM2* is associated with the incidence of tumors in patients with Down syndrome, suggesting the role of SIM2 in cancer biology [8]. In recent years, increasing studies have reported that SIM2 is related to tumorigenesis. However, its role in tumor progression varies across different cancer types. In esophageal cancer, SIM2 may increase the sensitivity of tumor cells to chemoradiotherapy by upregulating CD24 and cytokeratin 4 [9]. In cervical cancer, high SIM2 expression is considered to be an indicator of good prognosis [10]. However, short splice variant of SIM2 is significantly overexpressed in glioblastoma and prostate cancer and promotes the aggressiveness of tumor cells [11, 12]. However, the expression characteristics, clinical significance and biological function of SIM2 in endometrial carcinoma remain unclear.

With the rapid advancements in microarray and sequencing technologies, a growing number of cancer-related biomarkers and therapeutic targets have been identified [13]. Weighted gene co-expression network analysis (WGCNA) is an important method for exploring the relationships between genes and clinical phenotypes [14]. Additionally, Cox regression analysis and the least absolute shrinkage and selection operator (LASSO) are widely used to identify characteristic genes associated with cancer prognosis [15–17]. In the present work, it was hypothesized that SIM2 was involved in EC progression, and had the potential as a prognostic biomarker and therapeutic target. This study was performed to investigate the expression characteristics, clinical significance, and biological functions of SIM2 in EC, and preliminarily explore the downstream molecular mechanisms. With the bioinformatics approaches mentioned above, SIM2 was identified as a crucial regulator of EC progression,

and the biological function of SIM2 in regulating the malignant biological behaviors of EC cells were also verified with *in vitro* and *in vivo* assays.

Materials and methods

The overall design of the study

In this study, differentially expressed genes (DEGs) in EC tissues were identified from EC-related datasets in the Gene Expression Omnibus (GEO) and the Cancer Genome Atlas (TCGA) Program-Uterine Corpus Endometrial Carcinoma (UCEC) cohort. WGCNA was then employed to further identify genes associated with EC tumorigenesis and progression. A prognostic risk model was constructed based on the intersection of these genes using Cox regression analysis and LASSO. Subsequently, the expression pattern and biological function of SIM2, a key gene in the prognostic risk model, were analyzed. The study design is illustrated in Fig. 1.

Data download and processing

Gene expression datasets, including GSE17025, GSE63678, and GSE106191 (Table 1), were downloaded from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). The Sangerbox online platform was used to convert gene IDs into Gene Symbols, and the datasets GSE17025 and GSE63678 were merged [18]. The batch effect between the datasets was corrected using the COMBAT method.

Screening of DEGs

DEGs were screened using the “limma” package in R, with thresholds set at $P < 0.05$ and $|\log_2\text{FoldChange}| > 1$. The TCGA-UCEC RNA-seq data were obtained from the UCSC Xena database (<https://xenabrowser.net/datapages/>). The “DESeq2” package in R was employed to identify DEGs in the TCGA-UCEC dataset, with criteria set at $\text{padj} < 0.05$ and $|\log_2\text{FoldChange}| > 2$.

WGCNA

First, the median absolute deviation (MAD) of each gene was calculated, and the top 50% of genes with the smallest MAD values were excluded. Next, outlier genes and samples were removed using the `goodSamplesGenes` method in the WGCNA package in R. The WGCNA package was then used to construct a scale-free co-expression network. A soft threshold of 3 was set, after which an adjacency matrix was created, and the gene distribution was matched to a scale-free network based on the degree of connectivity. The adjacency matrix was subsequently transformed into a topological overlap matrix (TOM). Hierarchical clustering was performed to generate a dendrogram and assess the correlation between module characteristic genes and clinical labels.

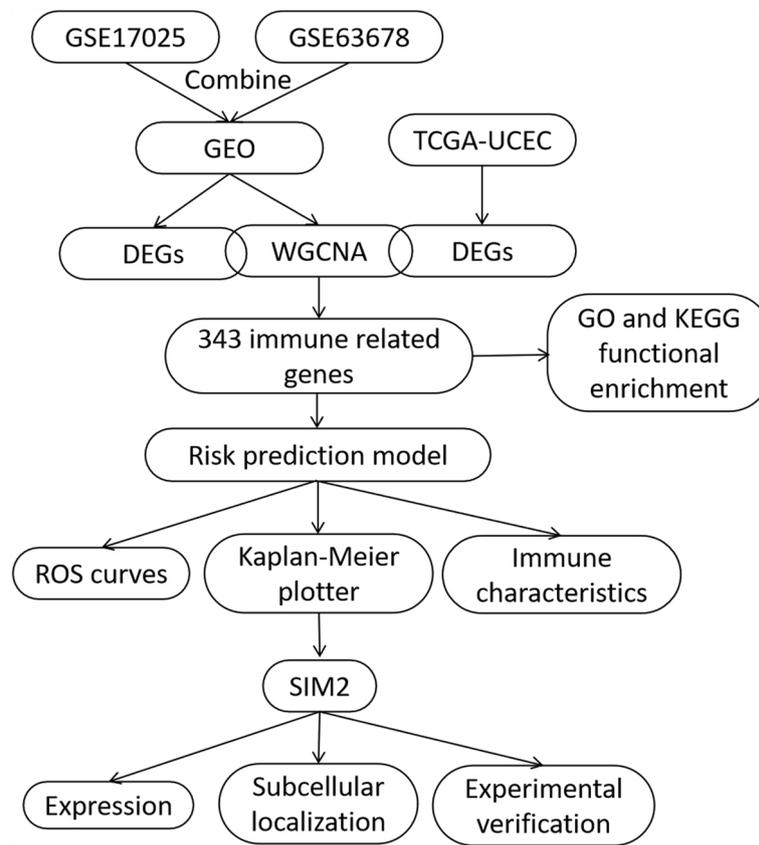


Fig. 1 Workflow of the work

Table 1 Information about the GEO datasets

Datasets	Control group	EC group	Affymetrix platform	Species
GSE17025	12	91	GPL570	Homo sapiens
GSE63678	5	7	GPL571	Homo sapiens
GSE106191	31	66	GPL570	Homo sapiens

Construction of prognostic risk model

TCGA-UCEC samples were randomly split into a training set and a test set in a 7:3 ratio. Univariate Cox regression was initially used to identify genes associated with prognosis, and these genes were further analyzed using LASSO and multivariate Cox regression to construct the prognostic risk model. A ten-fold cross-validation was conducted to determine the optimal penalty parameter (λ) for the model. Based on the average risk score, the samples were classified into high-risk and low-risk subgroups. The Kaplan–Meier method and log-rank test were used to compare the overall survival (OS) between the two groups. The accuracy of the risk model in

predicting 1-, 3-, and 5-year survival rates was evaluated using receiver operating characteristic (ROC) curves and nomograms.

Immunoinfiltration analysis

After dividing the TCGA-UCEC cohort into high-risk and low-risk groups, the levels of immune cell infiltration in each group were calculated using the CIBERSORT algorithm [19], and the results were compared between the groups. Additionally, single-sample gene set enrichment analysis (GSEA) (ssGSEA) was employed to assess the degree of immune cell infiltration in different samples using the R package “GSVA” [20]. The immune microenvironment of the two groups was further analyzed using the Estimation of Stromal and Immune cells in Malignant Tumors using Expression data (ESTIMATE) algorithm, available in the R package “ESTIMATE” [21]. The ESTIMATE, stromal, and immune scores were compared between the two groups. Furthermore, the expression levels of immune checkpoint markers, including CD28, CD274, CTLA4, LAG3, and TIGIT, were compared between the samples from the high-risk and low-risk groups.

Gene expression, subcellular localization and biological function analysis of SIM2

Gene expression data for SIM2 in UCEC and normal tissues were obtained from the Gene Expression Profiling Interactive Analysis (GEPIA) database (<http://gepia2.cancer-pku.cn/#index>) using TCGA data [22]. Additionally, the expression characteristics of SIM2 in EC tissues were validated using the GSE106191 dataset. The subcellular localization and immunohistochemical analysis of SIM2 in normal endometrial tissues and EC tissues were performed using data from the Human Protein Atlas (HPA) database [23]. The LinkedOmics database (<https://www.linkedomics.org/admin.php>) was used to identify genes correlated with SIM2 in the TCGA-UCEC cohort. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed using the Sangerbox platform [18]. GO analysis includes three categories: biological process (BP), cellular component (CC), and molecular function (MF). A *p*-value < 0.05 was considered significant for enrichment.

Cell culture and transfection

The immortalized endometrial epithelium cell line hEM15A and EC cell lines hEM15A, HEC-1-A, AN3CA, KLE and Ishikawa were obtained from the American Type Culture Collection (ATCC, Rockville, MD, USA). The cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, Carlsbad, CA, USA) enriched with 10% fetal bovine serum (FBS, ThermoFisher Scientific, MA, USA), 100 U/mL penicillin, and 0.1 mg/mL streptomycin (ThermoFisher Scientific, MA, USA) in an incubator at 37 °C with 5% CO₂. Lipofectamine[®] 3000 (ThermoFisher Scientific, MA, USA) was used for transfection of shRNA (sh-NC, sh-SIM2-1, sh-SIM2-2) and plasmids (oe-NC and oe-SIM2) into HEC-1-A and KLE cells. The target sequences for sh-SIM2-1 and sh-SIM2-2 are 5'-GCTCACGGGCAACAGTATATTA-3' and 5'-TGCATCGTGAGTGTC AATTAT-3', respectively.

Real-time quantitative polymerase chain reaction (qPCR)

Total RNA from HEC-1-A and KLE cells was extracted using TRIzol reagent (Invitrogen, Shanghai, China). The RNA was reverse-transcribed into cDNA using the HiScript IV All-in-One Ultra RT SuperMix for qPCR kit (R433-01, Vazyme, Nanjing, China). The PCR mixture was prepared using the ChamQ Blue Universal SYBR qPCR Master Mix kit (Q312-02, Vazyme, Nanjing, China) and primers. qPCR was performed on a StepOnePlus real-time PCR system (Applied Biosystems, San Francisco, CA, USA). The primer sequences used were as follows: SIM2, forward: 5'-AAGTCCAAGAATGCGGCCAA-3', and reverse: 5'-TTGTCCAGCTGCGAAGTGTGAT-3'. β-actin, forward: 5'-CATGTACGTTGCTAT

CCAGGC-3', and reverse: 5'-CTCCTTAATGTCACGCACGAT-3'.

Cell Counting Kit-8 (CCK-8) assay

Cell viability was assessed using a CCK-8 kit (Beyotime, Shanghai, China). After transfection, HEC-1-A and KLE cells were seeded in 96-well plates (2 × 10³ cells /well) and cultured in an incubator. At 0, 24, 48, and 72 h post-transfection, 10 μL of CCK-8 reagent was added to each well. After incubating for 1 h, the absorbance at 450 nm was measured using a spectrophotometer (Bio-Rad, CA, USA).

Flow cytometry

Apoptosis in HEC-1-A and KLE cells was detected using an Annexin V-FITC apoptosis detection kit (Beyotime, Shanghai, China). Briefly, the collected cells were washed with ice-cold phosphate-buffered saline (PBS) and resuspended in a binding buffer. Annexin V-FITC and propidium iodide were added, and the mixture was incubated at room temperature for 10 min, protected from light. A flow cytometer (FACScan; BD Biosciences, USA) was used to analyze the cells. For cell cycle distribution analysis, cells were washed twice with PBS and fixed with 75% ethanol. After fixation, cells were stained with propidium iodide (PI) for 10 min, followed by flow cytometry.

Western blotting

Total protein from EC cells in each group was extracted and separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), then transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore, Bedford, MA, USA). The membrane was incubated with primary and secondary antibodies (Abcam, Shanghai, China). Protein bands on the PVDF membrane were visualized using a supersensitive ECL chemiluminescence kit (Beyotime, Shanghai, China). ImageJ software (NIH, Bethesda, MD, USA) was used to analyze the gray values of each band. The antibodies used in this study were: anti-SIM2 (ab131161, 1:1000, Abcam, Shanghai, China), anti-cleaved caspase-3 (ab2302, 1:1000, Abcam, Shanghai, China), anti-Cyclin D1 (ab16663, 1:1000, Abcam, Shanghai, China), anti-CDK4 (ab108357, 1:1000, Abcam, Shanghai, China), anti-GAPDH (ab181602, 1:5000, Abcam, Shanghai, China), and secondary antibodies (1:5000, Proteintech, Wuhan, China).

Animal model

Animal experiments were approved by the Ethics Committee of Wuhan Children's Hospital. Female BALB/c nude mice (6 weeks old) were obtained from Shulaibao Co., Ltd. (Wuhan, China). HEC-1-A cells (2 × 10⁶ per mouse), with or without SIM2 depletion, were injected

into the tail vein of each mouse (6 mice per group). Four weeks after injection, the mice were euthanized. The mice were sacrificed by CO₂ asphyxiation followed by cervical dislocation. The lungs and livers were harvested, washed, fixed, and embedded in paraffin. Tissue sections were then prepared, and hematoxylin–eosin staining was performed. Metastatic tumor nodules in the lungs and livers were observed and evaluated under a microscope.

Statistical analysis

Survival curves were generated using the Kaplan–Meier method and analyzed with the log-rank test. For normally distributed continuous variables, comparisons were made using the Student's *t*-test. For non-normally distributed variables, the Wilcoxon test was used to compare the mRNA expression levels of a single gene between the high-risk and low-risk groups, or compare the data of animal assays. The experimental results were analyzed and plotted using GraphPad Prism software. A *P* value < 0.05 was deemed statistically significant.

Results

Screening of DEGs in EC and identification of EC-related genes with WGCNA

After merging the GSE17025 and GSE63678 datasets and correcting for batch effects (Fig. 2A–C), a total of 986 DEGs were identified, with 458 genes significantly upregulated and 528 genes significantly downregulated in EC (Fig. 2D). The heatmap presents the expression profiles of the top 20 genes that were most significantly upregulated and downregulated (Fig. 2E). Subsequently, WGCNA was used to identify gene modules associated with cancerous and non-cancerous tissue groupings. A soft threshold of $\beta = 3$ ($R^2 = 0.91$) was applied to construct a scale-free network (Supplementary Fig. 1A–C). Modules with a distance less than 0.8 were merged, resulting in four co-expression modules. Among these, the lightgreen module, which contained 960 genes, was significantly positively correlated with EC ($r = 0.70$, $P = 5.3e-18$), while the royablue module, containing 707 genes, was significantly negatively correlated with EC ($r = -0.56$, $P = 9.2e-11$) (Supplementary Fig. 1D–E). Additionally, genes in these two modules showed significant correlations with StromalScore and ESTIMATEScore (Supplementary Fig. 1D–E). In the TCGA-UCEC cohort, 4,762 DEGs were identified, with 3,241 genes upregulated and 1,521 genes downregulated in EC (Fig. 3A). There were 343 common genes found in the intersection of DEGs from the TCGA dataset, DEGs from the GEO dataset, and genes from the lightgreen and royablue modules (Fig. 3B). GO analysis revealed that these 343 genes were enriched in BP, such as cell division and the cell cycle (Fig. 3C), CC, such as

chromosomes, microtubule cytoskeleton, and extracellular matrix (Fig. 3D), and MF, such as glycosaminoglycan binding, extracellular matrix structural constituent, and signaling receptor binding (Fig. 3E). KEGG enrichment analysis indicated that these genes were involved in the cell cycle, oocyte meiosis, p53 signaling pathway, and IL-17 signaling pathway (Fig. 3F).

Establishment of a prognostic risk model for EC patients via LASSO

In the TCGA-UCEC cohort, a univariate Cox regression analysis was performed on the 343 genes identified previously to screen for prognostic genes. A prognostic risk model was then constructed using LASSO and multivariate Cox analysis (Fig. 4A–B). The model included 13 genes. The risk score for each patient was calculated using the following formula: risk score = (0.186 × TSPYL5 expression value) + (0.467 × NAP1L2 expression value) + (0.327 × HAPLN1 expression value) + (-0.781 × RNASE4 expression value) + (-0.329 × KLF2 expression value) + (0.422 × CKMT1B expression value) + (-0.260 × PAMR1 expression) + (1.061 × PLCL1 expression) + (0.220 × SIM2 expression) + (0.374 × ATP1B2 expression) + (0.420 × MEOX2 expression) + (0.166 × GALNT14 expression) + (0.141 × NTS expression). This prognostic risk model was evaluated using the training set, validation set, and TCGA-UCEC dataset. The ROC curve analysis revealed that the area under the curve (AUC) values for the risk score were consistently high in all three groups of EC patients (>0.7), indicating the risk score's strong diagnostic ability to predict 1-year, 3-year, and 5-year survival outcomes (Fig. 4C). Kaplan–Meier survival analysis showed that the survival rate of patients in the high-risk group was notably lower across all three groups of EC patients (Fig. 4D). CIBERSORT analysis revealed a decrease in CD8 T cell infiltration in the high-risk group (Supplementary Fig. 2A). ssGSEA showed significant differences in the infiltration levels of most immune cell types between the high- and low-risk groups (Supplementary Fig. 2B). The ESTIMATE algorithm indicated that the StromalScore, ImmuneScore, and ESTIMATEScore were significantly lower in the high-risk group compared to the low-risk group, suggesting that tumors with a high-risk score tend to be “cold tumors” with immunosuppression (Supplementary Fig. 2C). Furthermore, the expression of immune checkpoint markers CD28, CTLA4, and TIGIT was significantly higher in the low-risk group (Supplementary Fig. 2D).

Expression characteristics of genes in the prognostic risk model

Univariate Cox analysis of the 13 genes identified in the prognostic risk model revealed that RNASE4, KLF2, and PAMR1 acted as protective factors (hazard ratio < 1),

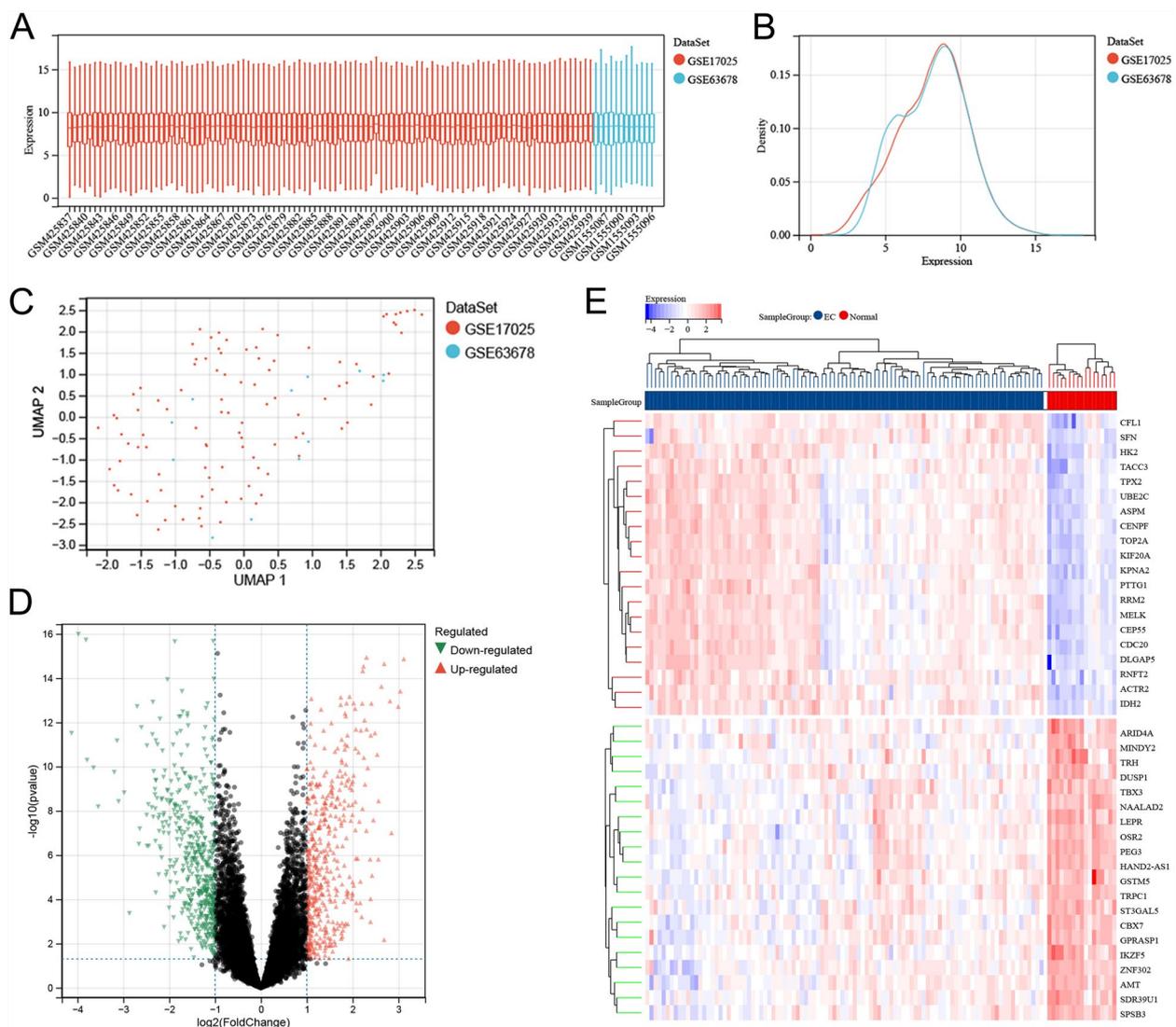


Fig. 2 Merging of GEO datasets and DEGs analysis. **A-C** Data distribution (**A**), density distribution (**B**) and UMAP distribution (**C**) after merging the data of GSE17025 and GSE63678, and removal of batch effect. **D** The volcano map shows the DEGs in EC tissues and non-cancerous tissues, with green representing significantly down-regulated genes, red representing significantly up-regulated expression genes, and black representing genes with insignificant expression differences. **E** The heat map shows the expression profiles of the top 20 up-regulated and top 20 down-regulated DEGs in EC and non-cancerous tissues. The red group represents normal tissue samples and the blue group represents EC samples. In the heat map, the deeper the red, the higher the expression level, and the deeper the blue, the lower the expression level

while the remaining genes served as risk factors (hazard ratio > 1) (Fig. 5A). The distribution of risk scores and survival status in EC patients indicated that the survival rate in the high-risk group was notably lower than that in the low-risk group (Fig. 5B). Heatmaps and box plots further revealed that RNASE4, KLF2, and PAMR1 were significantly underexpressed in the high-risk group, while the expression of the other genes was notably higher in the low-risk group in the TCGA-UCEC cohort (Fig. 5B and C). These findings suggest that RNASE4, KLF2, and PAMR1 function as tumor suppressors in EC, while the

other 10 genes likely act as oncogenes. Additionally, a nomogram was constructed that combined the risk score with clinicopathological parameters, including age, tumor grade, and tumor stage. In the nomogram, the risk score contributed the most to the overall score (Fig. 5D). The calibration curve demonstrated that the nomogram accurately predicted 1-, 3-, and 5-year survival, with high consistency between the predicted and actual survival outcomes (Fig. 5E). Furthermore, a nomogram model was developed based on the expression levels of the 13 genes, and all of these genes significantly contributed

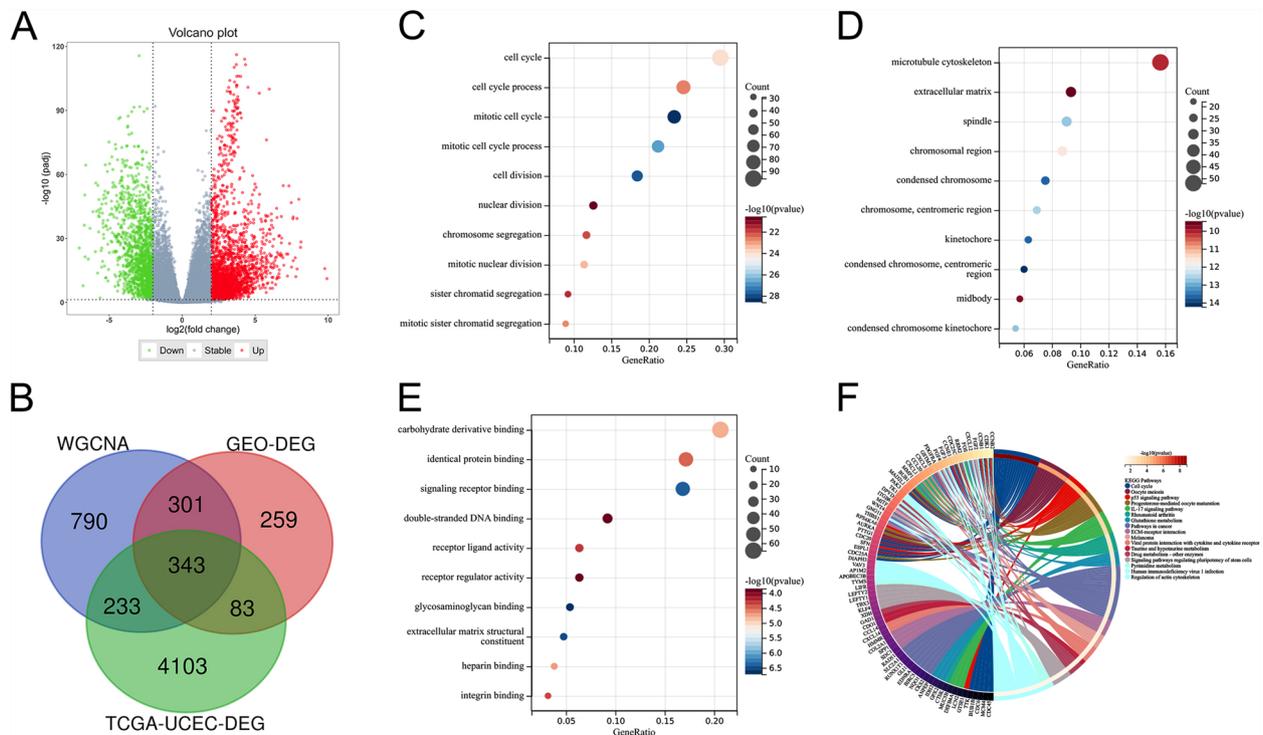


Fig. 3 Screening of EC-related genes and functional enrichment analysis. **A** The volcano map shows the DEGs between EC tissues and normal tissues in the TCGA-UCEC cohort. Green represents down-regulated genes and red represents up-regulated genes. **B** The common genes of DEGs in GEO dataset, genes in the crucial modules in WGCNA, and DEGs in TCGA-UCEC cohort were obtained by a Venn diagram. **C-E** The bubble diagram shows the results of GO enrichment analysis, including BP (**C**), CC (**D**) and MF (**E**). The size of bubble indicate indicates the gene count, the color of the bubble indicates the *P* value. **F** The string diagram shows the results of KEGG enrichment analysis. BP: biological process; CC: cellular component; MF: molecular function

to survival prediction (Fig. 5F). The corresponding calibration curves also showed strong agreement between the survival predictions and the actual observed values (Fig. 5G).

SIM2 identified as a potential oncogenic factor in EC from the 13 genes in the prognostic risk model

In the GEO dataset, CKMT1B, GALNT14, HAPLN1, NTS, and SIM2 were notably overexpressed in EC tissues compared to non-cancerous endometrial tissues (Supplementary Fig. 3A). Similarly, in the TCGA-UCEC cohort, CKMT1B, GALNT14, HAPLN1, and SIM2 were significantly overexpressed in EC tissues, while NTS was under-expressed (Supplementary Fig. 3B). Next, patients in the TCGA-UCEC cohort were divided into high-expression and low-expression groups based on the median expression value of each gene. Notably, EC patients with high expression of NAP1L2 and SIM2 exhibited significantly worse prognosis, while those with high expression of PAMR1 had a better prognosis (Supplementary Fig. 3C). No significant association was found between the expression of the other 10 genes and patient prognosis (data not shown). Given that SIM2 was highly expressed in

patients with high-risk scores, as well as in EC tissues compared to non-cancerous tissues, and its association with poor prognosis, it was hypothesized that SIM2 acts as a cancer-promoting factor in EC. Consistently, in the GSE106191 dataset, SIM2 was significantly more highly expressed in EC tissues compared to non-cancerous tissues (Supplementary Fig. 4).

In-silico analysis of expression and biological function of SIM2 in EC

In the HPA database, SIM2 was significantly overexpressed in EC tissues. Among the EC samples, 3 cases showed high expression, 6 cases showed medium expression, 1 case had low expression, and 2 cases had no detectable expression. However, SIM2 was not detected in endometrial stroma or glandular cells (Fig. 6A). Subcellular localization analysis revealed that SIM2 was predominantly localized in the nucleus (Fig. 6B). The BP and signaling pathways associated with SIM2 in EC were analyzed using the LinkedOmics database. The results showed that SIM2-related genes were mainly involved in positively regulating BP, such as DNA conformation change, chromosome

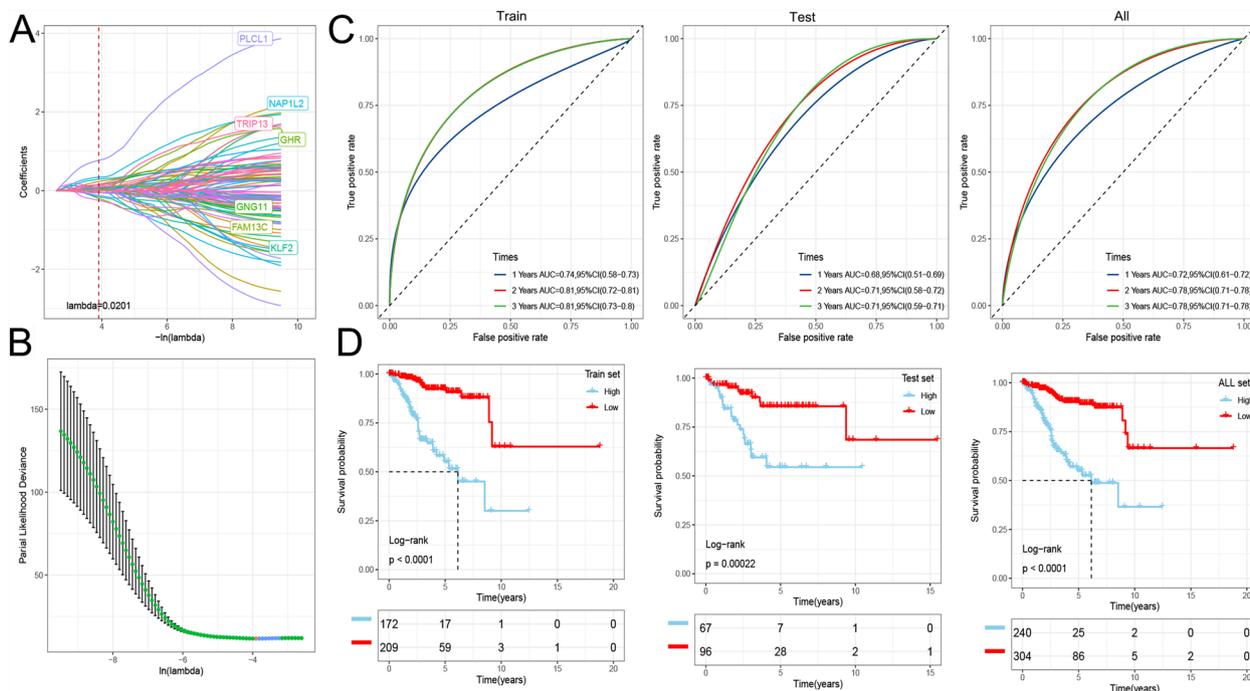


Fig. 4 Establishment of the prognostic risk model for EC. **A** Ten-fold cross-validation of adjusted parameters using LASSO-Cox regression model. **B** Locus of EC-related genes. One line represents a gene, and the horizontal axis represents the logarithm of the gene λ , and the vertical axis shows the coefficient of the gene. **C** ROC curves were applied to evaluate the predictive value of risk model in training set, validation set, and TCGA-UCEC cohort. **D** Kaplan–Meier survival curves were used to analyze the overall survival of patients in the high and low risk groups in the training set, validation set and TCGA-UCEC cohort, and log-rank test was performed

segregation, DNA strand elongation, and DNA replication. Conversely, SIM2-related genes were found to negatively regulate pathways such as adrenergic receptor signaling, muscle system processes, the cellular response to vascular endothelial growth factor (VEGF) stimuli, etc. (Fig. 6C). KEGG enrichment analysis revealed that SIM2-related genes were primarily involved in pathways associated with tumorigenesis and cancer progression, including the cell cycle, fructose and mannose metabolism, p53 signaling pathway, cGMP-PKG signaling pathway, etc. (Fig. 6D). These findings suggest that SIM2 contributes to EC tumorigenesis and progression by promoting DNA replication, cell cycle progression, and the growth of cancer cells.

SIM2 positively regulates the malignant biological behaviors of EC cells

Next, the expression level of SIM2 in different EC cell lines were detected by Western blot, and the result showed that compared with that in immortalized endometrial epithelium cell line hEM15A, the expression level of SIM2 in EC cell lines was significantly increased

(Fig. 7A). To investigate the role of SIM2 in EC, SIM2 knockdown and overexpression models were established in HEC-1-A and KLE cells (Fig. 7B and C). Among the shRNAs tested, sh-SIM2-1 exhibited the most effective inhibition of SIM2 and was therefore used for subsequent experiments. The viability of HEC-1-A and KLE cells with SIM2 knockdown was notably reduced compared to the sh-NC group. Conversely, EC cells with SIM2 overexpression showed a significant increase in cell viability compared to the oe-NC group (Fig. 7D). The apoptosis rate of EC cells in the sh-SIM2-1 group was significantly higher than that in the control group (sh-NC), while the apoptosis rate in the oe-SIM2 group was significantly lower than in the oe-NC group (Fig. 7E). Cell cycle analysis revealed that the proportion of EC cells in the G1 phase was notably higher in the sh-SIM2-1 group compared to the sh-NC group, and the proportion of cells in the S phase was significantly lower. In contrast, the proportion of EC cells in the G1 phase was significantly reduced in the oe-SIM2 group, while the proportions in the S and G2 phases were notably increased (Fig. 8A). Western blot analysis was conducted to examine the expression levels of apoptosis-related and

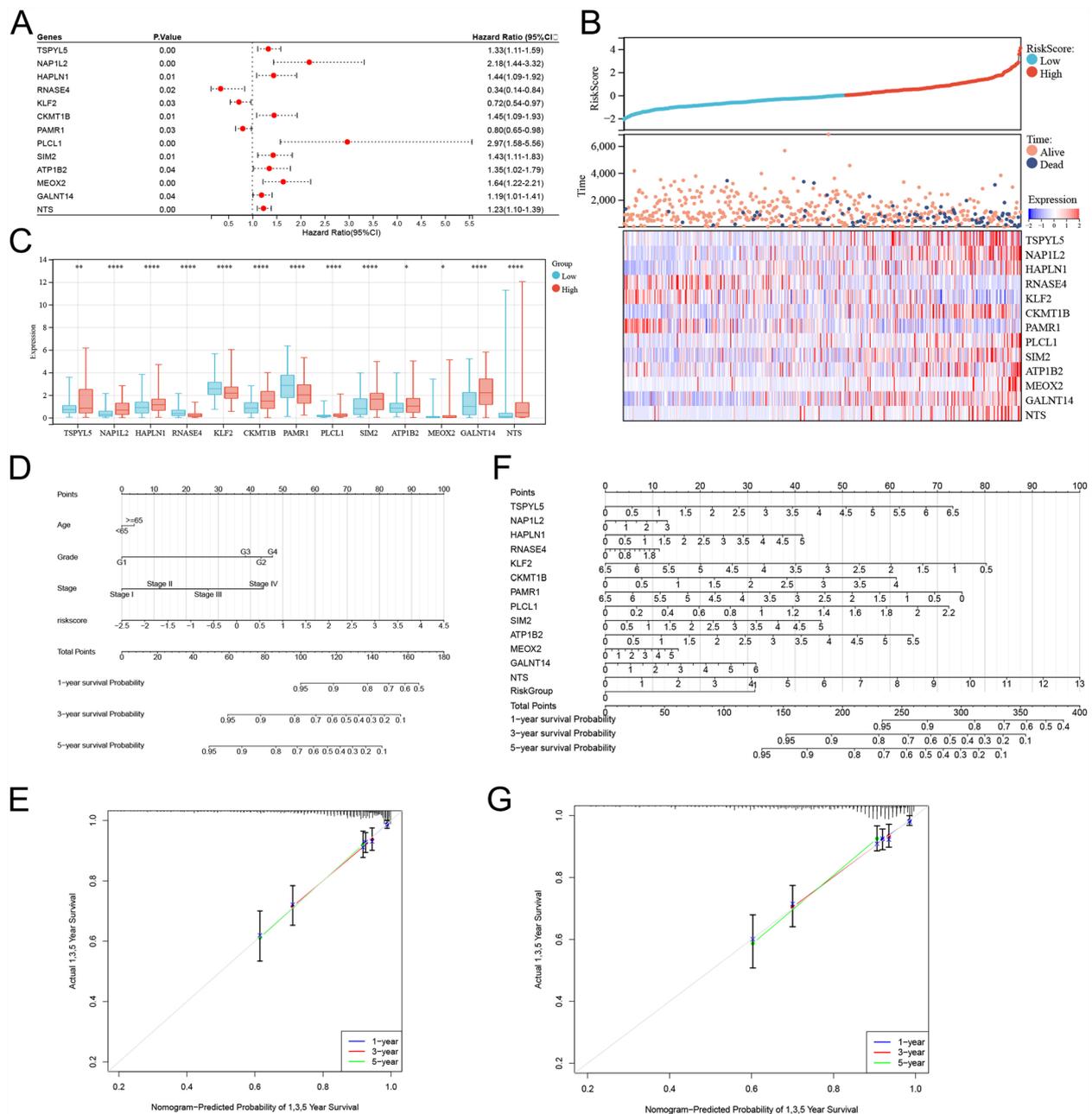


Fig. 5 Expression characteristics of the genes in the prognostic risk model. **A** The forest map shows the results of univariate Cox analysis of the 13 genes in the prognostic risk model. **B** Risk score distribution, survival status and 13 gene expression profiles of the patients with high or low risk scores in the TCGA-UCEC dataset. **C** The box plots show the expression of 13 genes in the high-low risk group in the TCGA-UCEC dataset. **D** and **E** The nomogram model including clinicopathologic features and risk score was constructed, and the calibration curves of 1 year, 3 years and 5 years were used to evaluate the prediction accuracy of nomogram. **F** and **G** The nomogram model including the expression of 13 genes was constructed, and the calibration curves of 1 year, 3 years and 5 years were used to evaluate the prediction accuracy of nomogram. Wilcoxon test was used to compare the data between two groups. *** $P < 0.001$

cell cycle-related proteins. The results showed that the protein expression level of cleaved caspase-3 was significantly higher following SIM2 knockdown, while the levels of Cyclin D1 and CDK4 were significantly reduced.

Conversely, the opposite results were observed in the oe-SIM2 groups (Fig. 8B). To further validate the biological function of SIM2 in EC progression, a nude mouse model with lung/liver metastasis was established. In this model,

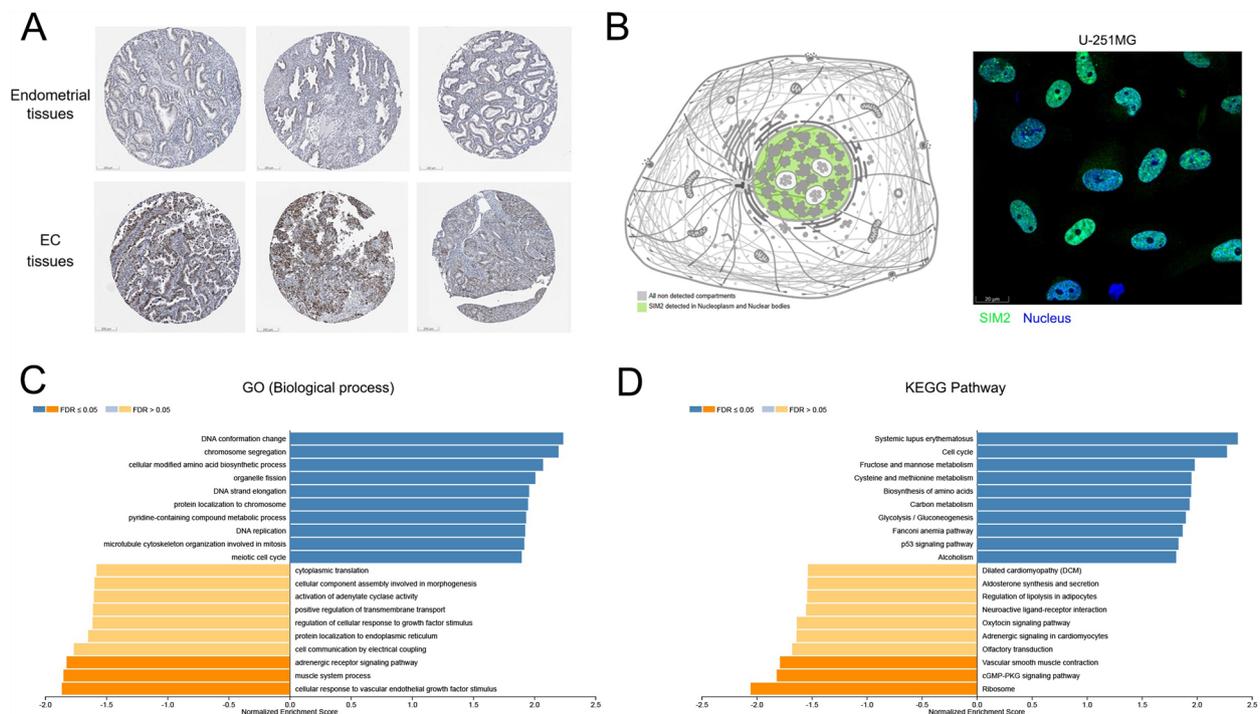


Fig. 6 Expression characteristics and biological function analysis of SIM2 in EC. **A** Representative immunohistochemical images of SIM2 in EC tissues and normal tissues in HPA database. Scale bar = 200 μ m. **B** Representative images of subcellular localization immunofluorescence of SIM2 in cells in HPA database. Scale bar = 20 μ m. **C–D** Functional enrichment analysis of SIM2-related genes, including GO analysis (biological process, C) and KEGG enrichment analysis (D), were performed using LinkedOmics database. Benjamini–Hochberg false discovery rate was applied to evaluate the enrichment

silencing SIM2 notably reduced the ability of HEC-1-A cells to form metastatic nodules in the lungs and livers of the mice (Fig. 8C and D).

Discussion

The p53 pathway plays a critical role in repressing cell cycle progression and inducing apoptosis in EC cells by inhibiting cyclin-dependent kinases [24, 25]. IL-17, a pro-inflammatory cytokine secreted by various cells within the tumor microenvironment, including T helper cells, CD8 T cells, and natural killer cells [26], has been implicated in promoting EC progression [27]. In this study, 343 genes were identified through differential expression analysis combined with WGCNA. Bioinformatics analysis suggested that these genes were primarily involved in processes such as cell division, cell cycle regulation, the p53 signaling pathway, and the IL-17 signaling pathway. Our data further support the critical roles of the p53 and IL-17 pathways in EC tumorigenesis and progression. Subsequent univariate Cox regression analysis and the LASSO identified 13 key EC-related genes, which were used to construct a prognostic risk model. This model demonstrated strong predictive efficacy for forecasting the prognosis of EC patients. Based on this risk model,

patients in the TCGA cohort were classified into high-risk and low-risk groups. In recent years, immunotherapy has significantly improved the prognosis of EC patients, particularly for those with metastatic disease [28]. CD8 T cells are known to play an essential anti-tumor role, and their recruitment to tumors can enhance the effectiveness of immune checkpoint inhibitors [29]. In this study, the level of CD8 T cells was found to be notably higher in the low-risk group compared to the high-risk group, suggesting that patients in the low-risk group may have a better response to immunotherapy. Additionally, the expression levels of immune checkpoint proteins such as CD28 (PD-1), CTLA4, and TIGIT were evidently lower in the high-risk group, indicating T cell exhaustion within the tumor microenvironment of these patients. Collectively, this prognostic risk model may prove useful for immunotherapy stratification in EC patients. Future studies should enroll more patients from diverse medical centers to validate the efficacy of this risk model. In particular, further investigation into the relationship between risk scores and patient responses to immunotherapy, as well as the status of microsatellite instability, is warranted to better understand its potential in clinical applications.

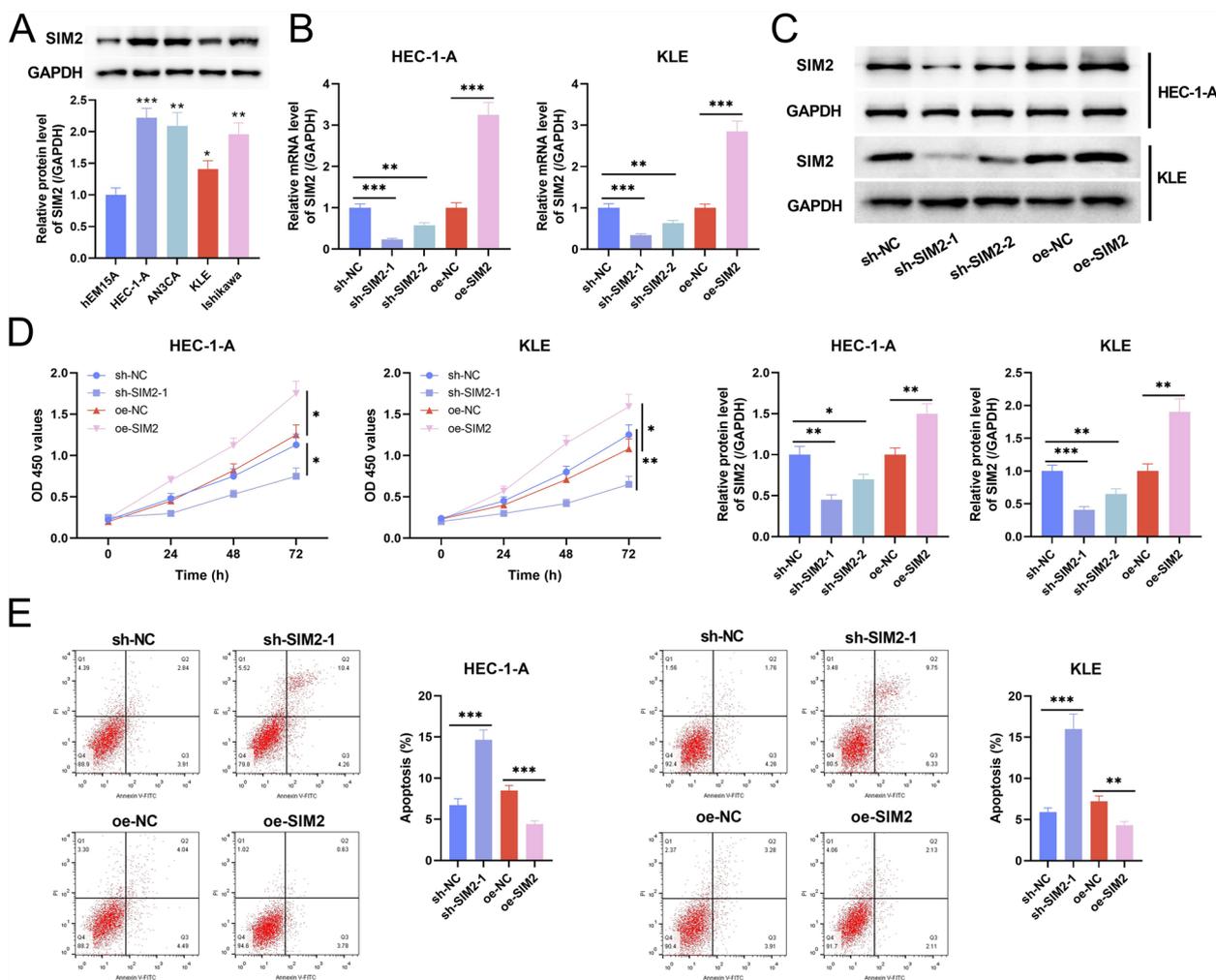


Fig. 7 SIM2 promotes the viability and represses the apoptosis of EC cells. **A** Western blot was used to detect SIM2 expression levels in different cell lines including hEM15A, HEC-1-A, AN3CA, KLE and Ishikawa. **B** qRT-PCR was used to detect SIM2 expression levels in HEC-1-A and KLE cell lines with SIM2 knockdown or overexpression. **C** Western blot was used to detect SIM2 expression levels in HEC-1-A and KLE cell lines with SIM2 knockdown or overexpression. **D** The effect of SIM2 expression on EC cell viability was detected by CCK-8 assay. **E** The effect of SIM2 expression on EC cell apoptosis was detected by flow cytometry. Student's t test was used to compare the data between two groups. * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$

In this study, SIM2, one of the 13 genes identified, was selected for further investigation. SIM2 encodes a transcription factor that plays a crucial role in neurogenesis and angiogenesis. It undergoes ubiquitination by RING-IBR-RING-type E3 ubiquitin ligases, including the parkin RBR E3 ubiquitin protein ligase [30]. Recent studies have also highlighted the involvement of SIM2 in the progression of cervical cancer, prostate cancer, and pancreatic cancer [10, 31, 32]. Our findings revealed that SIM2 was significantly overexpressed in EC tissues, and patients with high SIM2 expression had notably poorer clinical outcomes. Furthermore, SIM2 was found to positively regulate the malignant biological behaviors of EC

cells, consistent with the results of a recent study [33]. Additionally, SIM2 was predominantly localized in the nucleus, and genes associated with SIM2 were linked to BP, such as DNA conformation changes, chromosome segregation, DNA strand elongation, and DNA replication. Cyclin D primarily binds to cyclin-dependent kinase 4/6 (CDK4/CDK6), facilitating the transition from the G1 to the S phase of the cell cycle [34]. In this study, SIM2 knockdown resulted in a significant reduction in the expression levels of Cyclin D1 and CDK4 proteins, which provides preliminary insight into the mechanism by which SIM2 modulates the phenotypes of cancer cells.

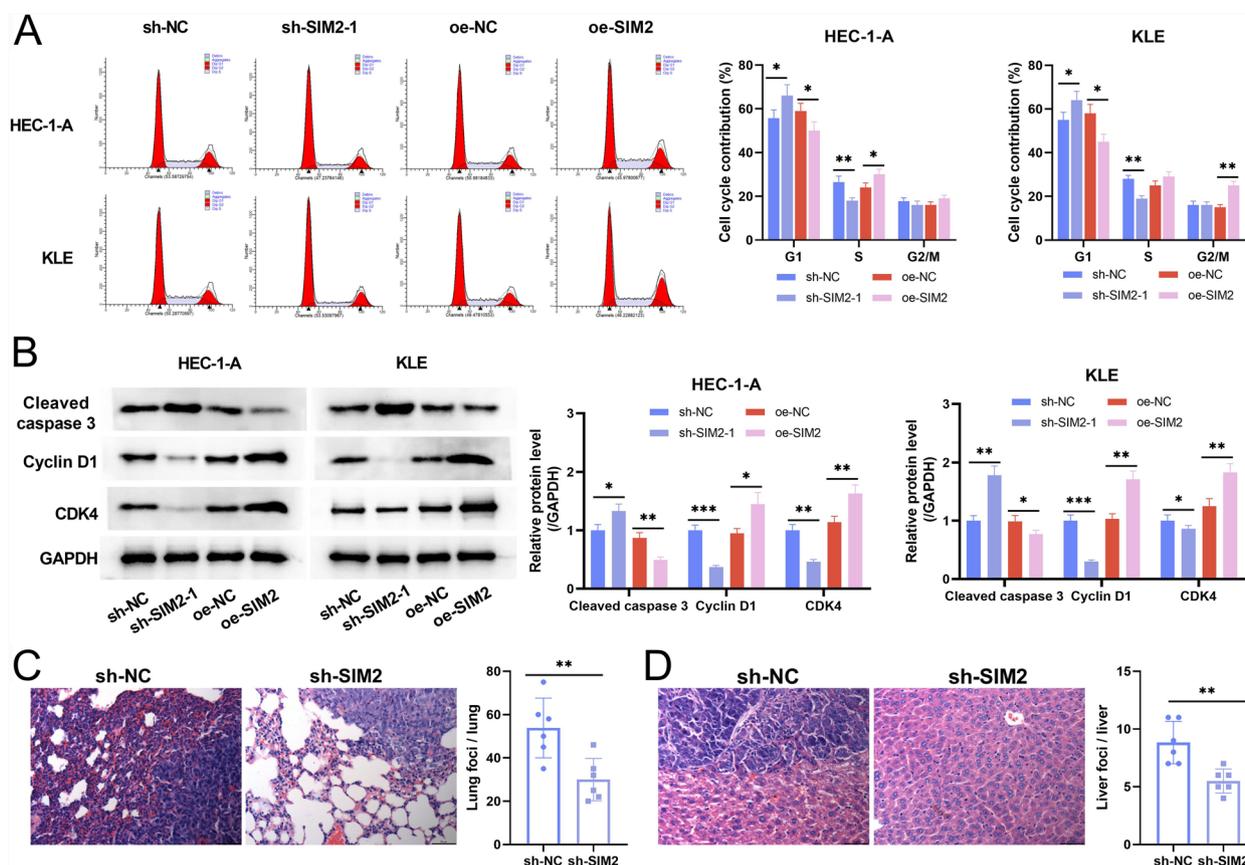


Fig. 8 SIM2 knockdown represses cell cycle progression of EC cells in vitro and metastasis of EC cells in vivo. **A** The effect of SIM2 expression on EC cell cycle progression was detected by flow cytometry. **B** The expression levels of cleaved caspase 3, Cyclin D1 and CDK4 in HEC-1-A and KLE cell lines with SIM2 knockdown or overexpression, were detected by western blot. **C** and **D** HEC-1-A cells were injected into nude mice via tail vein, and H&E staining was used to examine lung (**C**) and liver (**D**) metastases in mice, and the severity of metastasis in two groups was compared. For the in vitro assays, Student's t test was used to compare the data between two groups. For the in vivo assays, Wilcoxon test was used to compare the data between two groups. * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$

It is important to note that the other 12 genes in the prognostic model may also play a role in regulating the phenotypes of EC cells. These genes are closely associated with cancer biology. For example, HAPLN1 (hyaluronan and proteoglycan link protein-1) is highly expressed in cancer-associated fibroblasts [35], where it can promote tumor cell plasticity. It has been shown to enhance the aggressiveness of tumor cells in pancreatic and gastric cancers [36, 37]. KLF2 (Kruppel-like factor 2), a member of the transcription factor family with conserved zinc finger domains, is involved in various biological processes [38–40]. Due to its low expression in several cancer types, KLF2 is considered a tumor suppressor [41, 42]. PLCL1 (phospholipase C-like 1) encodes a protein that activates phospholipase C activity. In renal carcinoma, PLCL1 has been shown to inhibit cancer progression by regulating lipid metabolism [43]. A proteomic study has reported that PLCL1 is differentially expressed in EC and normal

tissues [44]. MEOX2 (mesenchyme homeobox-2) also plays a cancer-promoting role in various cancers [45, 46]. Future studies should aim to further clarify the expression characteristics, biological functions, and underlying mechanisms of these genes in EC.

This study has some limitations. First, the specific molecular mechanism by which SIM2 is involved in EC progression has not been fully elucidated. As a transcription factor, SIM2's abnormally high expression may induce the transcription of downstream oncogenes, thereby promoting the malignant biological behaviors of tumor cells. Additionally, the regulatory effects of SIM2 on other malignant phenotypes of tumor cells, such as angiogenesis and drug resistance, require further investigation. These hypotheses need to be validated through additional cell and animal models. Finally, the potential of SIM2 as a clinical prognostic biomarker must be confirmed in a larger cohort of clinical samples.

Conclusion

The prognostic risk model, developed using bulk RNA sequencing data, provides a useful tool for predicting the prognosis of EC patients and may offer valuable insights into the patients' responsiveness to immunotherapy. Furthermore, SIM2 promotes the malignant biological behaviors of EC cells and has the potential to serve as a therapeutic target for EC.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12885-025-14077-0>.

Supplementary Material 1.

Supplementary Material 2.

Authors' contributions

Both authors (NIE Hua & Chen Yu) participated in the experiment design, performing the experiments, preparing the manuscript. Both authors (NIE Hua & Chen Yu) read and approved the final manuscript.

Funding

This study is financially supported by the author themselves.

Data availability

The data used to support the findings of this study are available from the corresponding author upon request.

Declarations

Ethics approval and consent to participate

In the present work, the animal experiments were approved by the Ethics Committee of Wuhan Children's Hospital.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Reproductive Medicine Center, Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital) Tongji Medical College Huazhong University of Science and Technology Reproductive Medicine Center, Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital), Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430015, China. ²Department of Obstetrics and Gynecology, Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital), Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430015, China.

Received: 13 September 2024 Accepted: 3 April 2025

Published online: 11 April 2025

References

- Crosbie EJ, Kitson SJ, McAlpine JN, et al. Endometrial cancer. *Lancet* (London, England). 2022;399(10333):1412–28.
- Wang H, Xiao Y, Cai Y, et al. The Clinicopathological Features and Prognoses of Lower Uterine Segment Cancer: A Retrospective Single-Center Cohort Study. *Int J Women's Health*. 2024;16:1401–11.
- Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71(3):209–49.
- Makker V, MacKay H, Ray-Coquard I, et al. Endometrial cancer. *Nat Rev Dis Primers*. 2021;7(1):88.
- Adekanmbi V, Guo F, Hsu CD, et al. Temporal Trends in Treatment and Outcomes of Endometrial Carcinoma in the United States, 2005–2020. *Cancers*. 2024;16(7):1282.
- Sud S, Holmes J, Eblan M, Chen R, Jones E. Clinical characteristics associated with racial disparities in endometrial cancer outcomes: A surveillance, epidemiology and end results analysis. *Gynecol Oncol*. 2018;148(2):349–56.
- Button EL, Rossi JJ, McDougal DP, et al. Characterization of functionally deficient SIM2 variants found in patients with neurological phenotypes. *Biochem J*. 2022;479(13):1441–54.
- Chatterjee A, Dutta S, Mukherjee S, et al. Potential contribution of SIM2 and ETS2 functional polymorphisms in Down syndrome associated malignancies. *BMC Med Genet*. 2013;14:12.
- Takashima K, Fujii S, Komatsuzaki R, et al. CD24 and CK4 are upregulated by SIM2, and are predictive biomarkers for chemoradiotherapy and surgery in esophageal cancer. *Int J Oncol*. 2020;56(3):835–47.
- Nakamura K, Komatsu M, Chiwaki F, et al. SIM2 attenuates resistance to hypoxia and tumor growth by transcriptional suppression of HIF1A in uterine cervical squamous cell carcinoma. *Sci Rep*. 2017;7(1):14574.
- He Q, Li G, Su Y, et al. Single minded 2-s (SIM2-s) gene is expressed in human GBM cells and involved in GBM invasion. *Cancer Biol Ther*. 2010;9(6):430–6.
- Halvorsen OJ, Rostad K, Øyan AM, et al. Increased expression of SIM2-s protein is a novel marker of aggressive prostate cancer. *Clin Cancer Res*. 2007;13(3):892–7.
- Talijanovic M, Lopacinska-Jørgensen J, Høgdall EV. Gene Expression Profiles in Ovarian Cancer Tissues as a Potential Tool to Predict Platinum-based Chemotherapy Resistance. *Anticancer Res*. 2025;45(1):11–26.
- Miao Y, Konno Y, Wang B, et al. Integrated multi-omics analyses and functional validation reveal TTK as a novel EMT activator for endometrial cancer. *J Transl Med*. 2023;21(1):151.
- Tibshirani R, Friedman J. A Pliable Lasso, *Journal of computational and graphical statistics : a joint publication of American Statistical Association, Institute of Mathematical Statistics, Interface Foundation of North America*. 2020;29(1):215–25.
- Liu X, Wang W, Zhang X, et al. Metabolism pathway-based subtyping in endometrial cancer: An integrated study by multi-omics analysis and machine learning algorithms. *Molecular therapy Nucleic acids*. 2024;35(2):102155.
- Lin S, Wei C, Wei Y, Fan J. Construction and verification of an endoplasmic reticulum stress-related prognostic model for endometrial cancer based on WGCNA and machine learning algorithms. *Front Oncol*. 2024;14:1362891.
- Shen W, Song Z, Zhong X, et al. Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *Imeta*. 2022;1(3):e36.
- Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol Biol*. 2018;1711:243–59.
- Yi M, Nissley DV, McCormick F, et al. ssGSEA score-based Ras dependency indexes derived from gene expression data reveal potential Ras addiction mechanisms with possible clinical implications. *Sci Rep*. 2020;10(1):10258.
- Yoshihara K, Shahmoradgoli M, Martínez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612.
- Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45(W1):W98–102.
- M. Uhlén, L. Fagerberg L, B.M. Hallström, et al. Proteomics. Tissue-based map of the human proteome. *Science*. 2015;347(6220):1260419.
- Kukita A, Sone K, Kaneko S, et al. The Histone Methyltransferase SETD8 Regulates the Expression of Tumor Suppressor Genes via H4K20 Methylation and the p53 Signaling Pathway in Endometrial Cancer Cells. *Cancers*. 2022;14(21):5367.
- Li W, Qin Y, Chen X, Wang X. Cell division cycle associated 8 promotes the growth and inhibits the apoptosis of endometrial cancer cells by regulating cell cycle and P53/Rb signaling pathway. *Ame J Transl Res*. 2023;15(6):3864–81.
- Koh CH, Kim BS, Kang CY, et al. IL-17 and IL-21: Their Immunobiology and Therapeutic Potentials. *Immune Netw*. 2024;24(1):e2.

27. Cheng R, Xue X, Liu X. Expression of IL17A in endometrial carcinoma and effects of IL17A on biological behaviour in Ishikawa cells. *J Int Med Res.* 2020;48(9):300060520950563.
28. Villacampa C, Eminowicz G, Navarro V, et al. Immunotherapy and PARP inhibitors as first-line treatment in endometrial cancer: A systematic review and network meta-analysis. *Eur J Cancer.* 2025;220:115329.
29. Freeman P, Bellomo G, Ireland L, et al. Inhibition of insulin-like growth factors increases production of CXCL9/10 by macrophages and fibroblasts and facilitates CD8(+) cytotoxic T cell recruitment to pancreatic tumours. *Front Immunol.* 2024;15:1382538.
30. Jin X, Liu G, Zhang X, Du N. Long noncoding RNA TMEM75 promotes colorectal cancer progression by activation of SIM2. *Gene.* 2018;675:80–7.
31. Lu B, Asara JM, Sanda MG, Arredouani MS. The role of the transcription factor SIM2 in prostate cancer. *PLoS ONE.* 2011;6(12):e28837.
32. Hu Q, Wen N, Li F, et al. Single-minded homolog 2 as a potential prognostic signature and assessment of its correlation with immune cell infiltration in pancreatic cancer. *Polish J Pathol.* 2023;74(4):232–47.
33. Wei Y, Zhao X, Tang H, et al. SIM2: Its Prognostic Significance and Oncogenic Role in Endometrial Carcinoma. *Onco Targets Ther.* 2024;17:45–61.
34. Nardone V, Barbarino M, Angrisani A, et al. CDK4, CDK6/cyclin-D1 Complex Inhibition and Radiotherapy for Cancer Control: A Role for Autophagy. *Int J Mol Sci.* 2021;22(16):8391.
35. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun.* 2020;11(1):5120.
36. Zhang T, Li X, He Y, et al. Cancer-associated fibroblasts-derived HAPLN1 promotes tumour invasion through extracellular matrix remodeling in gastric cancer. *Gastric Cancer.* 2022;25(2):346–59.
37. Wiedmann L, De Angelis Rigotti F, Vaquero-Siguero N, et al. HAPLN1 potentiates peritoneal metastasis in pancreatic cancer. *Nat Commun.* 2023;14(1):2353.
38. Laha D, Deb M, Das H. KLF2 (kruppel-like factor 2 [lung]) regulates osteoclastogenesis by modulating autophagy. *Autophagy.* 2019;15(12):2063–75.
39. Tang X, Wang P, Zhang R, et al. KLF2 regulates neutrophil activation and thrombosis in cardiac hypertrophy and heart failure progression. *J Clin Investig.* 2022;132(3).
40. Wittner J, Schuh W. Krüppel-like Factor 2 (KLF2) in Immune Cell Migration. *Vaccines.* 2021;9(10):1171.
41. Lu Y, Qin H, Jiang B, et al. KLF2 inhibits cancer cell migration and invasion by regulating ferroptosis through GPX4 in clear cell renal cell carcinoma. *Cancer Lett.* 2021;522:1–13.
42. Taghehchian N, Maharati A, Akhlaghipour I, Zangouei AS, Moghbeli M. PRC2 mediated KLF2 down regulation: a therapeutic and diagnostic axis during tumor progression. *Cancer Cell Int.* 2023;23(1):233.
43. Pan Z, Huang J, Song H, et al. PLCL1 suppresses tumour progression by regulating AMPK/mTOR-mediated autophagy in renal cell carcinoma. *Aging.* 2023;15(19):10407–27.
44. Taylor AH, Konje JC, Ayakannu T. Identification of Potentially Novel Molecular Targets of Endometrial Cancer Using a Non-Biased Proteomic Approach. *Cancers.* 2023;15(18):4665.
45. Proserpio C, Galardi S, Desimio MG, et al. MEOX2 Regulates the Growth and Survival of Glioblastoma Stem Cells by Modulating Genes of the Glycolytic Pathway and Response to Hypoxia. *Cancers.* 2022;14(9):2304.
46. Peralta-Arrieta I, Trejo-Villegas OA, Armas-López L, et al. Failure to EGFR-TKI-based therapy and tumoural progression are promoted by MEOX2/GLI1-mediated epigenetic regulation of EGFR in the human lung cancer. *Eur J Cancer (Oxford England).* 1990;160(2022):189–205.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.