RESEARCH ARTICLE

# Integrative bioinformatics and experiments identify *RIBC2* as a key regulator in the esophageal cancer

Xuan Zheng[1], Yishuang Cui[2], Xuemin Yao[2], Yanan Wu[1], Yanlei Ge[2,3], Ye Jin[4,5], Junqing Gan[2,6], Weinan Yao[2,3], Yanna Bi[2,3], Guogui Sun (ID)[2,5]*

**1** School of Public Health, North China University of Science and Technology, Tangshan, Hebei, China, **2** Affiliated Hospital of North China University of Science and Technology, Tangshan, Hebei, China, **3** Hebei Province Innovation Center for Biological Cell Function Development and Precise Detection Technology, Tangshan, Hebei, China, **4** Clinical Medicine School, North China University of Science and Technology, Tangshan, Hebei, China, **5** Department of Hebei Key Laboratory of Medical-Industrial Integration Precision Medicine, Tangshan, Hebei, China **6** Sun Guogui Innovation Studio, Tangshan, Hebei, China

* guogui_sun2021@sina.com

## Abstract

Early detection of esophageal cancer (EC) remains a major challenge due to the limited understanding of its initial molecular alterations. Therefore, this study aimed to identify the key molecular drivers involved in EC carcinogenesis. Human normal esophageal epithelial cells were subjected to chronic malignant transformation, followed by assessment of their morphological changes, proliferative capacity, clonogenic potential, migration, and invasion abilities. To elucidate the molecular mechanisms underlying tumorigenesis, transcriptome sequencing was performed and integrated with clinical datasets from two independent EC cohorts. Machine learning algorithms were then applied to pinpoint diagnostic and prognostic gene signatures, which were further validated through comprehensive in vitro and in vivo experiments. Differential expression analysis and machine learning identified RIB43A domain with coiled-coils 2 (RIBC2) as a strong diagnostic and prognostic biomarker for EC. RIBC2 expression was markedly upregulated in chronically transformed epithelial cells, established EC cell lines, and clinical tumor specimens, and its elevation was associated with unfavorable clinicopathological characteristics. Functional studies revealed that silencing RIBC2 significantly inhibited cell proliferation, migration, and invasion in both transformed and EC cells. Moreover, immune profiling indicated that high RIBC2 expression was linked to an immune-excluded tumor microenvironment, implying a potential role in modulating responsiveness to immunotherapy. These findings reveal RIBC2 as a novel driver of EC initiation and progression, highlighting its potential as a biomarker for early diagnosis and as a promising target for therapeutic intervention.

## Introduction

Esophageal cancer (EC) is the eighth most prevalent malignancy globally and the sixth leading cause of cancer-related deaths (https://gco.iarc.who.int/media/globocan/factsheets/populations/900-world-fact-sheet.pdf). Notably, nearly 52.7% of all EC-related fatalities worldwide occur in China [1]. Despite advances in first-line treatment options—including surgery, radiotherapy, chemotherapy, targeted therapies, and novel immunotherapeutic approaches, the overall prognosis for EC patients remains dismal. This is primarily attributed to challenges in early detection and the high recurrence rate after treatment [2]. These challenges underscore an urgent need to identify novel and robust biomarkers capable of facilitating early diagnosis and timely therapeutic intervention, ultimately improving patient outcomes in EC.

With the advent of high-throughput sequencing and integrative multi-omics technologies, bioinformatics has emerged as a powerful tool for cancer biomarker discovery [3,4]. Through the systematic analysis of large-scale datasets, bioinformatic approaches enable the identification of critical molecular alterations and signaling pathways involved in tumor initiation and progression. In the case of EC, computational analyses have proven valuable for prioritizing candidate biomarkers with diagnostic, prognostic, and therapeutic potential. For instance, Ren et al. utilized transcriptomic data from the TCGA-EC cohort to identify *ANGPTL7*, *CSRP1*, *EXPH5*, *F2RL2*, *KCNMA1*, *MAGEC3*, *MAMDC2*, and *SLC4A9* as fibroblast-associated gene signatures linked to EC prognosis [5]. Similarly, Yuan et al. analyzed TCGA gene expression profiles and identified miR-205-3p, miR-452-3p, and miR-6499-3p as promising biomarkers for EC staging [6]. By high-dimensional single-cell proteomics, Han et al. built a CCR4/CCR6 chemokine-based model to stratify EC patients with different response to neoadjuvant chemoradiotherapy combined with immunotherapy [7]. The microbiomes of EC patients showed that intratumoral Streptococcus signatures could predict the response of neoadjuvant chemoradiotherapy [8]. Additionally, integrative multi-omics, combining transcriptomics, proteomics, metabolomics and microbiomics may provide complementary signals and better understanding of molecular mechanisms [9]. In EC, for example, PGK1 is found to reprogram glucose metabolism and contributes to EC progression by analyzing data of genomics, proteomics and phosphoproteomics from EC patients covering 9 histopathological stages and 3 phases [10]. Integrated analyses of genomics, epigenomics, transcriptomics and proteomics could classify EC patients into four molecular subtypes and identify 28 features in the prediction of anti-PD1 therapy response [11].

Machine learning provides practical tools for feature selection and model construction when analyzing multi-omics data, and may improve the diagnostic accuracy and prognostic prediction in EC. For example, HSPD1 and MAP1LC3B were identified as key prognostic genes by LASSO regression, random forest and XGBoost algorithms [12]. Li et al. established a clinic-radiomics nomogram by machine learning to predict the overall survival after definitive chemotherapy of EC patients [13]. Sun et al. employed 10 machine learning algorithms to generate 101 algorithm combinations and identified key neoantigen-related genes. These genes could stratify EC patients into groups with differential infiltration levels of dendritic cells, macrophages and B

cells [14]. However, the strength of bioinformatics predictions is often constrained by the lack of clinical and experimental validation. Therefore, transforming in silico findings into clinically meaningful insights demands rigorous analytical frameworks coupled with comprehensive biological verification.

To overcome the aforementioned limitations, we developed a comprehensive experimental and computational framework to uncover key molecular drivers of EC. Specifically, we established an in vitro EC transformation model and generated corresponding gene expression profiles to capture the molecular alterations associated with malignant progression. These experimental datasets were subsequently integrated with transcriptomic data from the TCGA-EC cohort to enhance the robustness and clinical relevance of biomarker discovery. Furthermore, in vitro and in vivo assays were conducted to experimentally validate the functional role of the candidate markers in EC progression. Collectively, this study provides a systematic and experimentally grounded strategy for elucidating the molecular mechanisms underlying esophageal tumorigenesis. The insights gained from this work may pave the way for the development of novel biomarkers and therapeutic targets, ultimately improving the early detection, prognosis, and clinical management of EC.

## Materials and methods

### Tissue samples and cell lines

A total of 50 paired tissue samples, comprising EC tissues and their corresponding adjacent non-tumorous tissues, were collected from the Pathology Department of Tangshan People's Hospital. The adjacent tissues, located at least 5.0 cm from the tumor margin, were verified as histologically normal through hematoxylin and eosin (H&E) staining. All experiments involving human tissues and clinical information were approved by the Ethics Committee of Tangshan People's Hospital (Approval No.: RMYY-LLKS-2024042). All experiments were performed in accordance with the Declaration of Helsinki. Informed consents were obtained from all participants in the current study.

The human normal esophageal epithelial cell line Het-1A (RRID: CVCL_3702, male) was obtained from IMMOCELLBio (Xiamen, China) and maintained in Endothelial Cell Medium (Cat. No. 1001, ScienCell). The human EC cell lines KYSE-410 (esophageal squamous cell carcinoma, ESCC; RRID: CVCL_1352, male), TE-10 (ESCC; RRID: CVCL_1760, male), and OE33 (esophageal adenocarcinoma, EAC; RRID: CVCL_0471, female) were purchased from ProcellBio (Wuhan, China). These EC cell lines were cultured in RPMI-1640 medium (Cat. No. C11875500BT, Gibco) supplemented with 10% fetal bovine serum (FBS; Cat. No. 10099−141, Gibco) and 1% penicillin–streptomycin (Cat. No. 15140−122, Gibco). For cell passaging or subculturing, adherent cells were detached using 0.25% trypsin–EDTA (Cat. No. 25200−056, Gibco). All cells were maintained at 37°C in a humidified incubator containing 5% $CO_2$. The identity of all cell lines was verified by short tandem repeat (STR) profiling within the past three years, and all experiments were carried out using mycoplasma-free cultures.

### N-methyl-N′-nitro-N-nitrosoguanidine (MNNG) exposure and malignant transformation

After the Het-1A cells had fully adhered, they were divided into two experimental groups: Het-1A-T (treated with MNNG) and Het-1A-N (treated with solvent control). MNNG was first dissolved in DMSO to prepare a 1 mol/L stock solution, which was subsequently diluted with culture medium to achieve a final concentration of 2 μM. When the cells reached 60–70% confluency, they were exposed to 2 μM MNNG or the corresponding solvent control for 24 hours per passage, and this treatment cycle was repeated for 50 consecutive generations. The selected environmental exposure dose of MNNG was based on previously published studies.

### Data collection and processing

Two independent cohorts from the TCGA and GEO databases were used in this study to identify key genes associated with the diagnosis and prognosis of EC. The TCGA-EC cohort comprised 161 EC tissues and 11 normal esophageal tissues, while the GEO-EC dataset (GSE53625) included 179 EC samples and 179 matched normal tissues. The

corresponding transcriptomic profiles and clinical data from both cohorts were retrieved for comprehensive integrative analysis.

In parallel, total RNA was extracted from three Het-1A-T samples (MNNG-treated) and three Het-1A-N samples (solvent-treated controls) using the mirVana™ miRNA Isolation Kit (Cat. No. AM1561, Ambion) according to the manufacturer's instructions. RNA integrity was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), and only samples with an RNA Integrity Number (RIN) ≥ 7 were included for subsequent analyses. RNA sequencing libraries were prepared using the TruSeq Stranded Total RNA with Ribo-Zero Gold Kit (Illumina) following standard protocols, and sequencing was performed on the Illumina HiSeq™ 2500 platform (or equivalent), generating 125 bp/150 bp paired-end reads. The raw sequencing data were obtained in FASTQ format and underwent quality control filtering to remove low-quality reads and adapter sequences. The resulting high-quality clean reads were then aligned to the human reference genome for downstream transcriptomic analysis.

All sequencing data were normalized and expressed as fragments per kilobase of transcript per million mapped reads (FPKM). Differentially expressed genes (DEGs) were identified using the DESeq2 package, with comparisons made between EC and normal tissues (DEGs1) and between Het-1A-T and Het-1A-N groups (DEGs2). Genes meeting the criteria of $|\log_2$ fold change$| > 1$ and an adjusted p-value $< 0.05$ were considered statistically significant.

## Bioinformatics analyses

The overlapping DEGs derived from DEGs1 and DEGs2 were first subjected to univariate Cox regression analysis to identify genes significantly associated with EC prognosis. The functional enrichment of these key DEGs was performed using ClueGO to elucidate their potential biological roles. Subsequently, the key DEGs were analyzed using four machine learning algorithms, CatBoost, Random Forest, Support Vector Machine (SVM), and Logistic Regression to construct a robust predictive model for EC occurrence. In parallel, Least Absolute Shrinkage and Selection Operator (LASSO) regression was applied to select feature genes from the key DEGs, which were then used to calculate a risk score. The association between the risk score and patient survival was evaluated through Kaplan–Meier (K–M) survival analysis and receiver operating characteristic (ROC) curve assessment. Furthermore, the biological functions and immune landscape associated with the identified feature genes were explored by comparing the high-risk and low-risk groups, providing insights into their roles in EC progression and tumor immune microenvironment modulation. The basic information of EC patients in low- and high-risk groups were shown in S1 Table.

In addition, to investigate the role of RIBC2 in tumor immune microenvironment and its potential molecular mechanisms in regulating EC, we (1) estimated the immune infiltration by IOBR package [15], (2) performed GSEA by GseaVis package [16], (3) identified the DEGs by DESeq2 using $|\log_2$ fold change$| > 1$ and an adjusted p-value $< 0.05$, and performed functional enrichment and protein-protein interaction network in low- and high-RIBC2 expression groups.

## RNA extraction and reverse transcription quantitative polymerase chain reaction (RT-qPCR)

Total RNA was isolated from paraffin-embedded tissue samples and cell cultures using the Paraffin-Embedded Tissue Total RNA Extraction Kit (Cat. No. R310, GeneBetter) and the Universal Total RNA Kit (Cat. No. R013, GeneBetter), respectively, in accordance with the manufacturers' protocols. The concentration and purity of the extracted RNA were quantified using a SpectraMax QuickDrop spectrophotometer (Molecular Devices, USA). Subsequently, complementary DNA (cDNA) was synthesized from the RNA samples using the PrimeScript™ RT Master Mix (Cat. No. RR036A, TaKaRa) following the manufacturer's instructions. Quantitative real-time PCR (qPCR) was performed on a QuantStudio™ 3 System (Thermo Fisher Scientific, USA) employing TB Green® Premix Ex Taq™ II (Cat. No. RR820A, TaKaRa). For the detection of RIBC2 expression, qPCR was carried out under the following cycling conditions: initial denaturation at 95°C for 30 seconds, followed by 40 cycles of denaturation at 95°C for 5 seconds and annealing/extension at 60°C for 34 seconds.

Relative gene expression levels were calculated using the $2^{-\Delta\Delta Cq}$ method, with GAPDH serving as the internal control. The primer sequences used in the current study were as follows: RIBC2 forward 1, 5'-GACATGAAACCTTTGCTGCTGAA-3' and reverse 1, 5'-CCGAACATCATTATCTGACTGC-3'; RIBC2 forward 2, 5'-AGGCCCTCTACACAGAGACAAG-3' and reverse 2, 5'-GCTTTTTCCTTTCCACTGACTC-3'; GAPDH forward, 5'-ACCCACTCCTCCACCTTTGA-3' and reverse, 5'-CCACCCTGTTGCTGTAGCCA-3'. In the case of RIBC2, the primers forward 1 and reverse 1, which generated a 215 bp product, were used in the initial experiments using 50 paired FFPE tumour and adjacent non-tumour tissues. To further reduce the potential influence of RNA integrity and to assess the robustness of the expression pattern, another RIBC2 primers (forward 2 and reverse 2) yielding a 147 bp product was used, and RT-qPCR was repeated in a subset of 20 paired FFPE tumour and adjacent non-tumour tissues from the same 50-pair cohort.

## Western blot assay

Cells were lysed on ice for 10 minutes using RIPA lysis buffer (Cat. No. P0013, Beyotime) supplemented with 1% protease inhibitor cocktail (Cat. No. ST506, Beyotime). The resulting lysates were centrifuged at 14,000×g for 5 minutes at 4°C, and the supernatant containing total protein was collected. Protein concentrations were determined using a Bicinchoninic Acid (BCA) Protein Assay Kit (Cat. No. PC0020, Solarbio). Equal amounts of protein were separated on a 10% SDS–polyacrylamide gel (SDS-PAGE) and transferred onto polyvinylidene fluoride (PVDF) membranes. The membranes were blocked with 5% skim milk in TBST (Tris-buffered saline containing 0.1% Tween-20) for 1 hour at room temperature (RT) to prevent nonspecific binding. This was followed by an incubation period at 4°C overnight with diluted primary antibodies, which included RIBC2 Polyclonal antibody (15272–1-AP, 1:1000 dilution, Proteintech), and GAPDH Polyclonal antibody (10494–1-AP, 1:5000 dilution, Proteintech). On the following day, after three washes with TBST (10 minutes each at RT), the membranes were incubated with the corresponding HRP-conjugated secondary antibodies (1:5000 dilution) for 1 hour at RT. The membranes were then washed three additional times with TBST (10 minutes per wash). Protein bands were visualized using an Enhanced Chemiluminescence (ECL) detection kit, and the signal was captured with a gel imaging system (Model C300, Azure Biosystems).

## Cell transfection

The short hairpin RNAs (shRNAs) plasmid targeting RIBC2 and a negative control (NC) shRNA were custom-designed and synthesized by GenCefeBio (Wuxi, China). The target sequences for the shRNAs were as follows: sh-RIBC2–1, GGCTATCAATGACTTCCAACA; sh-RIBC2–2, GCTGCAGTTTGACGAGACAGC; sh-RIBC2–3, GCTGGAGCAGATCCGCCTAGT. Het-1A-T, TE-10, and OE33 cells were seeded in 6-well plates and allowed to reach approximately 80% confluency within 24 hours. Transfection was carried out using the Lipofectamine™ 3000 Transfection Kit (Cat. No. L3000008, Invitrogen) according to the manufacturer's instructions. Specifically, a transfection mixture containing 2.5 μg of shRNA plasmid, 5 μL of P3000 reagent, 7.5 μL of Lipofectamine™ 3000, and 250 μL of Opti-MEM (Cat. No. 31985−070, Gibco) was gently prepared and incubated at room temperature for 15 minutes before being added dropwise to the cell cultures. After transfection, the cells were incubated at 37°C for 48 hours, after which they were harvested for subsequent experimental analyses.

## Cell doubling time

Het-1A-N and Het-1A-T cells were seeded separately into 12-well plates at a density of $1 \times 10^4$ cells per well and cultured in standard growth medium. Cell proliferation was monitored over a 7-day period, with cell counts from three parallel wells recorded daily using a CytoSMART™ Cell Counter (Corning, USA). The population doubling time (TD) was calculated using the following formula: $TD = t \times \log 2 / (\log N_t - \log N_i)$; where t is duration, $N_t$ is cell number on Day t, $N_i$ is initial cell number.

## Soft agar colony formation assay

A base layer was first prepared by mixing 1.2% sterile low–melting point agarose with 2×DMEM supplemented with FBS in a 1:1 ratio, and 1 mL of this mixture was added to each well of a 6-well plate. For the upper layer, cells were suspended in a mixture of 0.7% low–melting point agarose and 2×DMEM containing FBS, which was then gently overlaid onto the solidified base layer. After 14 days of incubation, cell colonies were visualized and documented using a light microscope (TS2, Nikon, Japan).

## Cell counting kit-8 (CCK-8) assay

Transfected Het-1A-T, TE-10, and OE33 cells (as described above) were seeded into 96-well plates, and cell viability was evaluated using the CCK-8 assay (Cat. No. MF128−01, Mer5bio) following the manufacturer's protocol. At designated time points, CCK-8 reagent was added to each well and incubated at 37°C for 2 hours. The absorbance was then measured at 450 nm using a microplate reader (Multiskan FC, Thermo Fisher Scientific, USA) to determine cell viability.

## Colony formation assay

Het-1A-N and Het-1A-T cells, as well as transfected Het-1A-T, TE-10, and OE33 cells (as described previously), were seeded into 6-well plates and cultured for 7–14 days to allow colony formation. After incubation, the cells were fixed with methanol for 15 minutes and subsequently stained with 0.1% (v/v) crystal violet (Cat. No. G1063, Solarbio) for another 15 minutes at RT. Colonies containing more than 50 cells were manually counted, and the average colony number was used to quantify the colony-forming ability of the cells, reflecting their proliferative potential.

## Wound healing assay

Het-1A-N and Het-1A-T cells, as well as transfected Het-1A-T, TE-10, and OE33 cell lines (as described previously), were seeded into 6-well plates to form a confluent monolayer. A uniform scratch was then created using a 200 μL pipette tip to generate a wound gap in each well. The wells were subsequently washed three times with 1×PBS to remove detached cells and incubated in serum-free medium at 37°C. Cell migration into the wound area was observed and photographed under a light microscope (Model TS2, Nikon, Japan) at 0 and 24 hours. The extent of wound closure was quantified using ImageJ software (1.51a, National Institutes of Health).

## Transwell invasion assay

A Transwell chamber system (Cat. No. 3422, Corning) was employed to evaluate the invasive capacity of cells. Briefly, Matrigel (Cat. No. 354234, Corning) was thawed at 4°C, and 50 μL of a Matrigel–medium mixture (ratio 1:8) was added to pre-chilled Transwell inserts and allowed to solidify at 37°C for 1 hour. Subsequently, Het-1A-N and Het-1A-T cells, or transfected Het-1A-T, TE-10, and OE33 cells (as described earlier), were seeded into the upper chamber containing serum-free Opti-MEM (Cat. No. 31985−070, Gibco). The lower chamber was filled with complete growth medium supplemented with 20% FBS, 100 U/mL penicillin, and 0.1 mg/mL streptomycin to serve as a chemoattractant. After 24 hours of incubation at 37°C in a 5% $CO_2$ atmosphere, non-invading cells remaining on the upper surface of the membrane were gently removed with a cotton swab. The inserts were then fixed with methanol for 15 minutes and stained with 1% (v/v) crystal violet (Cat. No. G1062, Solarbio) for an additional 15 minutes at RT. Representative fields were imaged under a light microscope (Model DM4B, Leica, Germany), and the number of invaded cells per field was manually counted to quantify cell invasion.

## In vivo xenograft assay

Four-week-old female BALB/c nude mice were obtained from HuafukangBio (Beijing, China) and housed under specific pathogen-free (SPF) conditions. Het-1A-N, Het-1A-T, KYSE-450 and OE33 cells were cultured, resuspended in sterile

PBS, and subcutaneously injected into the mice at a concentration of $1 \times 10^7$ cells per mouse. Tumor growth was monitored, and tumor volume was measured on day 28 post-injection, calculated using the formula: volume = length×(width$^2$)/2. At the end of the experiment, mice were euthanized by cervical dislocation under deep anesthesia induced with sodium pentobarbital (40 mg·kg$^{-1}$, i.p.). Anesthetic depth was verified by the absence of pedal and palpebral reflexes. Death was confirmed by the absence of heartbeat and corneal reflex for ≥2 min. Tumors were then excised and weighed for further analysis. All animal experiments were carried out with approved protocols, relevant guidelines and regulations of the Laboratory Animal Ethics Committee of North China University of Technology (protocol code: 2023-SY-070). All methods were reported in accordance with ARRIVE guidelines.

### H&E staining

Het-1A-N and Het-1A-T cells were fixed in 95% ethanol for 10 minutes. For tissue sections, samples were first incubated at 60°C for 1 hour, followed by dewaxing in xylene and rehydration through a graded series of ethanol solutions. Subsequently, both cells and tissue sections were stained with hematoxylin for 3 minutes, subjected to brief differentiation in 1% hydrochloric acid alcohol, and then counterstained with eosin for 2 minutes. After washing with PBS, the samples were mounted using neutral gum. All staining procedures were performed at room temperature (RT). After staining, the cells and tissue sections were examined and imaged using a light microscope (DM4B, Leica, Germany).

### Immunohistochemistry (IHC)

Paraffin-embedded tissue sections were first incubated at 60°C for 1 hour to melt excess paraffin, followed by deparaffinization in xylene I and II for 10 minutes each. The sections were then rehydrated through a graded series of ethanol solutions (100%, 90%, 80%, and 70%, each for 5 minutes) and rinsed three times with PBS at RT for 5 minutes per wash. To quench endogenous peroxidase activity, the sections were treated with 3% hydrogen peroxide ($H_2O_2$) for 8 minutes, followed by antigen retrieval using a sodium citrate buffer. After retrieval, nonspecific binding was blocked with 5% bovine serum albumin (BSA) (Cat. No. AR0004, Bosterbio) at 37°C for 20 minutes. Subsequently, the tissue sections were subjected to overnight incubation at 4°C with the following primary antibodies: KI67 Polyclonal antibody (27309–1-AP, 1:10000, Proteintech) and pan-keratin Polyclonal antibody (26411–1-AP, 1:4000, Proteintech). After washing, sections were incubated with HRP-conjugated secondary antibodies (Cat. No. PV-6001, ZSGBbio) for 40 minutes at RT, followed by color development using 3,3'-diaminobenzidine (DAB) and counterstaining with hematoxylin. Finally, the slides were mounted with neutral gum, and the stained sections were examined and imaged under a light microscope (DM4B, Leica, Germany).

### Statistical analysis

All statistical analyses were conducted using R software (version 4.2.2). Comparisons between two groups were performed using either the Wilcoxon rank-sum test or the Student's t-test, depending on data distribution. The chi-squared test was applied to assess differences in categorical clinical variables. Kaplan–Meier survival curves and ROC analyses were generated to evaluate the prognostic and diagnostic significance of candidate biomarkers. All biological experiments were carried out independently in triplicate, and a p-value < 0.05 was considered statistically significant unless otherwise indicated.

## Results

### Establishment of a chronic in vitro model of esophageal epithelial malignant transformation

To mimic the malignant transformation of esophageal epithelial cells, an in vitro model was established by exposing non-tumorigenic Het-1A cells to chronic low-dose carcinogenic stimulation. After 50 consecutive passages, the

resulting transformed cells (Het-1A-T) displayed distinct morphological and phenotypic changes compared with the non-transformed controls (Het-1A-N). Specifically, Het-1A-T cells exhibited enlarged and irregular morphology, enhanced intercellular adhesion, cytoplasmic vacuolization, and the appearance of multinucleated giant cells compared to Het-1A-N cells (Fig 1A). Cell proliferation analysis revealed that Het-1A-T cells had a significantly higher growth rate (Fig 1B) and a shorter doubling time (Fig 1C) than Het-1A-N cells. Functional assays including soft agar colony formation, plate colony formation, wound healing, and Transwell invasion further confirmed that Het-1A-T cells possessed markedly enhanced migratory and invasive abilities compared to their non-transformed counterparts (Fig 1D–1G). In in vivo tumorigenicity assays using nude mice, Het-1A-T cells, similar to the positive control ESCC (KYSE-450) and EAC (OE33) cells formed subcutaneous tumors, while Het-1A-N cells failed to do so (Fig 1H–1I). Het-1A-T cells generated tumours with malignant histology and high Ki-67 and pan-CK expression, confirming that transformed squamous epithelial cells acquired tumourigenic capacity in vivo (Fig 1H–K). Collectively, these findings confirm the successful establishment of a malignant transformation model, providing a robust system for subsequent mechanistic and transcriptomic analyses.

## Identification of candidate diagnostic biomarkers for early EC

To identify potential biomarkers for the early detection of EC, we first examined transcriptomic alterations associated with malignant transformation in vitro. A total of 1,580 DEGs were identified between Het-1A-T and Het-1A-N cells, comprising 791 upregulated and 789 downregulated genes (S1A and S1B Fig). To enhance the clinical relevance of our findings, transcriptomic data from the TCGA-EC cohort were integrated, yielding 3,510 DEGs between tumor and normal esophageal tissues (S1C and S1D Fig). Cross-comparison of the in vitro and in vivo datasets revealed 481 overlapping DEGs (S1E Fig). Univariate Cox regression analysis of these overlapping genes identified 22 DEGs significantly associated with patient prognosis, including *HSPA1B*, *STK32A*, *GLTPD2*, *ABCC6*, *ESM1*, *COX6B2*, *APLN*, *BTG2*, *CACNG8*, *DUSP2*, *RIPPLY3*, *SLIT2*, *SLIT3*, *RIBC2*, *CXCL8*, *PRR18*, *HSPA1A*, *NOL4*, *SLC24A4*, *PHYHIP*, *ACP7*, and *LMF1* (S2 Table). Functional enrichment analysis indicated that these genes were predominantly involved in signal transduction, cell proliferation, and cytokine-mediated signaling pathways (S1F Fig). Using these 22 prognostic DEGs, we constructed and evaluated four machine learning models for EC diagnosis. Among them, the logistic regression model exhibited the best performance, achieving an AUC of 1.0 in the TCGA training cohort and 0.84 in the GEO validation cohort (Fig 2A–2D). Feature importance analysis revealed that *ESM1*, *ABCC6*, *CXCL8*, *RIBC2*, and *APLN* were the top five contributors to the model (Fig 2E and 2F). Further SHAP dependency analysis demonstrated that elevated expression of *ESM1*, *CXCL8*, *RIBC2*, and *APLN* corresponded to higher SHAP values, indicating a positive association with EC diagnosis, while *ABCC6* expression showed a negative correlation, suggesting a protective effect (Fig 2G). Overall, these findings identify *ESM1*, *ABCC6*, *CXCL8*, *RIBC2*, and *APLN* as robust diagnostic biomarkers with strong potential for early-stage EC detection.

## Identification of candidate prognostic biomarkers for EC

To identify robust prognostic biomarkers for EC, the initial set of 22 prognosis-related DEGs was further refined using LASSO regression analysis, resulting in the identification of 17 key prognostic genes: *HSPA1B*, *STK32A*, *ABCC6*, *ESM1*, *COX6B2*, *APLN*, *BTG2*, *CACNG8*, *RIPPLY3*, *SLIT2*, *SLIT3*, *RIBC2*, *CXCL8*, *HSPA1A*, *SLC24A4*, *ACP7*, and *LMF1* (Fig 3A). A risk score was computed for each EC patient based on the expression levels and LASSO-derived coefficients of these 17 genes, and patients were subsequently categorized into high-risk (risk score^high) and low-risk (risk score^low) groups (Fig 3B). In the TCGA-EC cohort, Kaplan–Meier survival analysis revealed that patients in the low-risk group had significantly longer overall survival compared to those in the high-risk group (Fig 3C). Given the different biological backgrounds of ESCC and EAC, we further performed Kaplan-Meier analyses stratified by histological subtype, and it was also found that low-risk group had better survival than high-risk group in both TCGA-ESCC and TCGA-EAC cohorts (S2 Fig). The predictive accuracy of this prognostic model was further validated using time-dependent ROC curves, yielding AUC values of 0.83, 0.84, and 0.79 for 1-, 2-, and 3-year survival, respectively (Fig 3D). Notably, patients with advanced
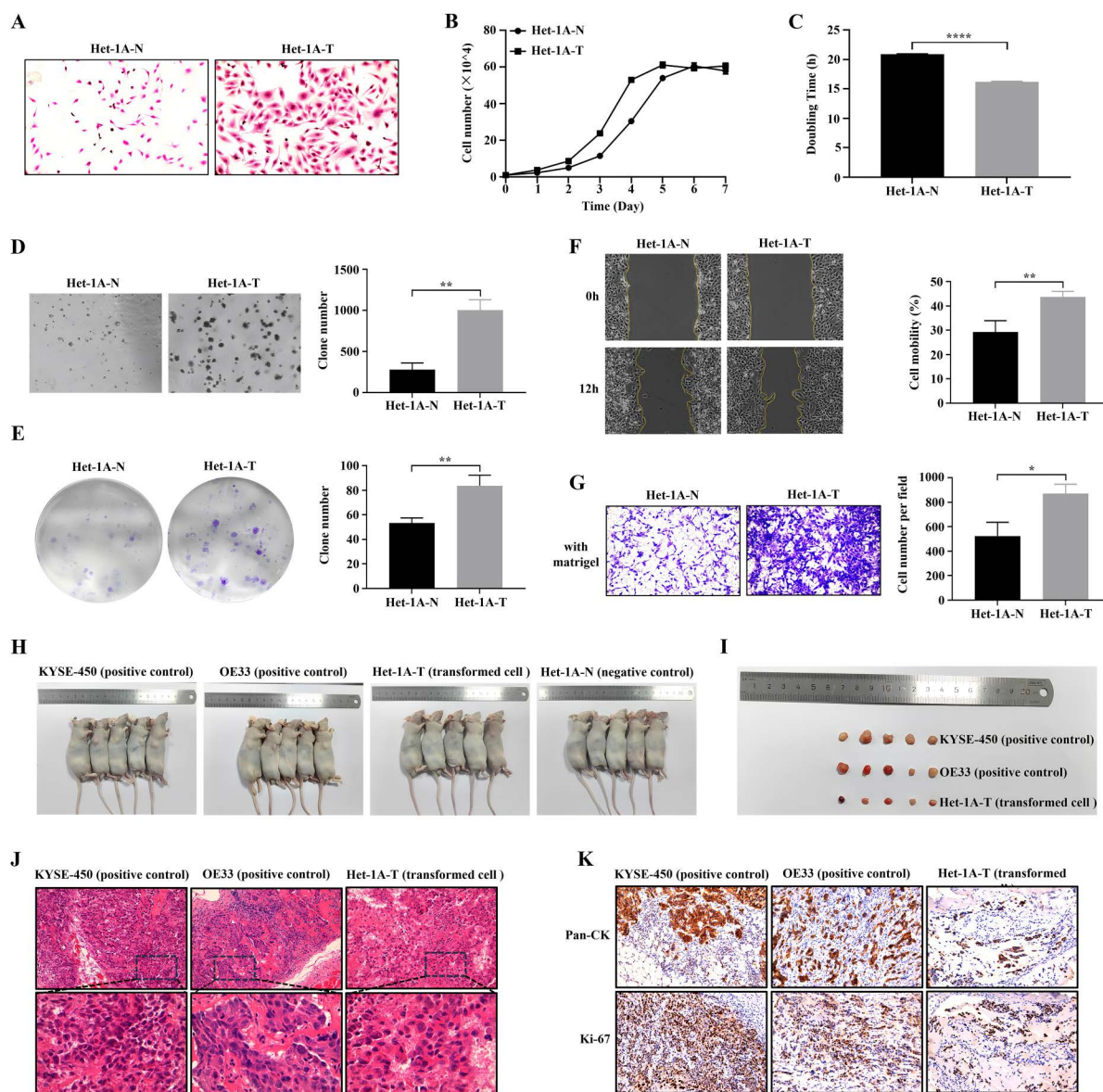
**Fig 1. Establishment of an in vitro model of esophageal epithelial malignant transformation.** (A) Representative H&E-stained images showing distinct morphological differences between non-transformed Het-1A-N and transformed Het-1A-T cells (magnification, ×100). (B, C) Cell growth curves (B) and doubling time (C) of Het-1A-N and Het-1A-T cells. (D-G) Evaluation of malignant characteristics in Het-1A-N and Het-1A-T cells using soft agar colony formation (D, ×40), colony formation (E), wound healing (F, ×40), and Transwell invasion (G, ×100) assays. (H) Images of nude mice bearing subcutaneous xenografts generated from KYSE-450 (ESCC positive control), OE33 (EAC positive control), Het-1A-T (transformed squamous epithelial cells) and Het-1A-N (negative control). (I) Photographs of resected xenograft tumours from KYSE-450, OE33 and Het-1A-T groups; no palpable tumours were formed by Het-1A-N cells. (J) Representative H&E staining of xenograft tumours derived from KYSE-450, OE33 and Het-1A-T cells; higher-magnification views of the boxed areas are shown in the lower panels. (K) Representative immunohistochemical staining of xenograft tumors showing Pan-CK and Ki-67 expression (magnification, ×200). *$P < 0.05$, **$P < 0.01$ and ****$P < 0.0001$.

nodal (N1–N3) or distant metastatic (M1) disease exhibited significantly higher risk scores than those with N0 or M0 status ([Fig 3E](#) and [3F](#)), suggesting that the prognostic gene signature is reflective of tumor aggressiveness and metastatic potential. Further immune profiling revealed distinct tumor microenvironment characteristics between the high- and
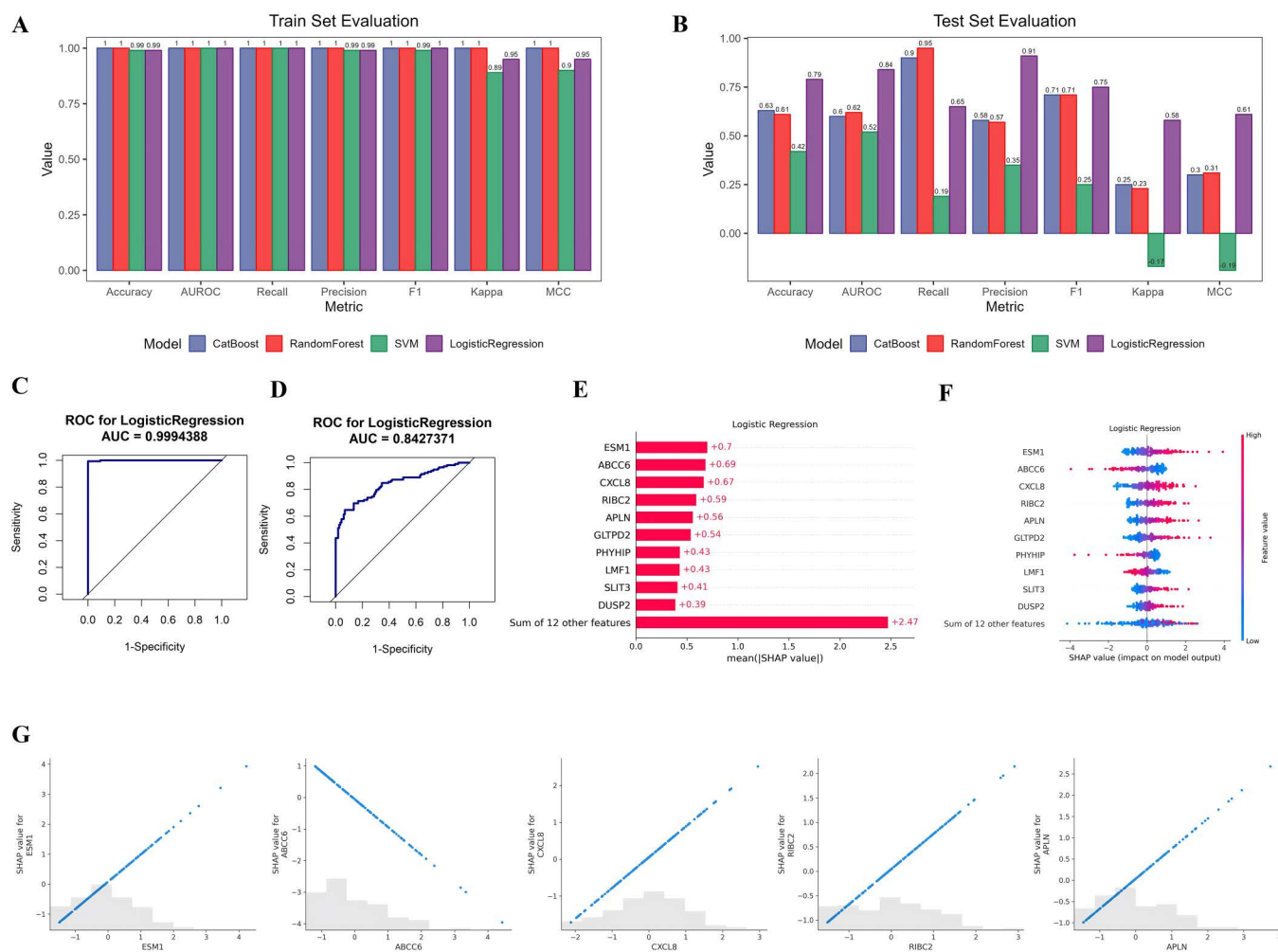
**Fig 2. Construction and interpretation of a diagnostic model for EC based on machine learning.** (A, B) Comparative performance assessment of four machine learning algorithms (CatBoost, Random Forest, Support Vector Machine and Logistic Regression) in the training (A) and testing (B) cohorts. (C, D) ROC curves and corresponding AUC values for the logistic regression model in the training (C) and testing (D) datasets. (E) Mean SHAP value plot showing the top 10 most influential features contributing to the logistic regression model. (F) SHAP summary plot illustrating the distribution and magnitude of SHAP values for all features, ranked according to their impact on model predictions. (G) SHAP dependence plots for the top five diagnostic genes, demonstrating the relationship between gene expression levels and their influence on EC classification. Positive SHAP values represent a stronger contribution toward EC prediction.

https://doi.org/10.1371/journal.pone.0340850.g002

low-risk groups, underscoring the immunological relevance of this prognostic model. The low-risk group (risk score^low) displayed significant enrichment of immune-related functions, including type I/II interferon (IFN) responses, immune checkpoint activity, para-inflammation, and increased HLA gene expression (Fig 3G). In addition, the infiltration of both innate and adaptive immune cells such as immature dendritic cells (iDCs), macrophages, T helper cells, and natural killer (NK) cells was markedly higher in the low-risk group (Fig 3H). To elucidate the molecular mechanisms underlying the distinct phenotypes between risk groups, GSEA was performed. The high-risk group (risk score^high) was characterized by the enrichment of metabolic pathways, including lipid absorption and sterol homeostasis (Fig 3I), whereas the low-risk group showed predominant enrichment of immune-related pathways, such as immune response-regulating signaling pathways, leukocyte-mediated immunity, and the T cell receptor complex (Fig 3J). Together, these results suggest
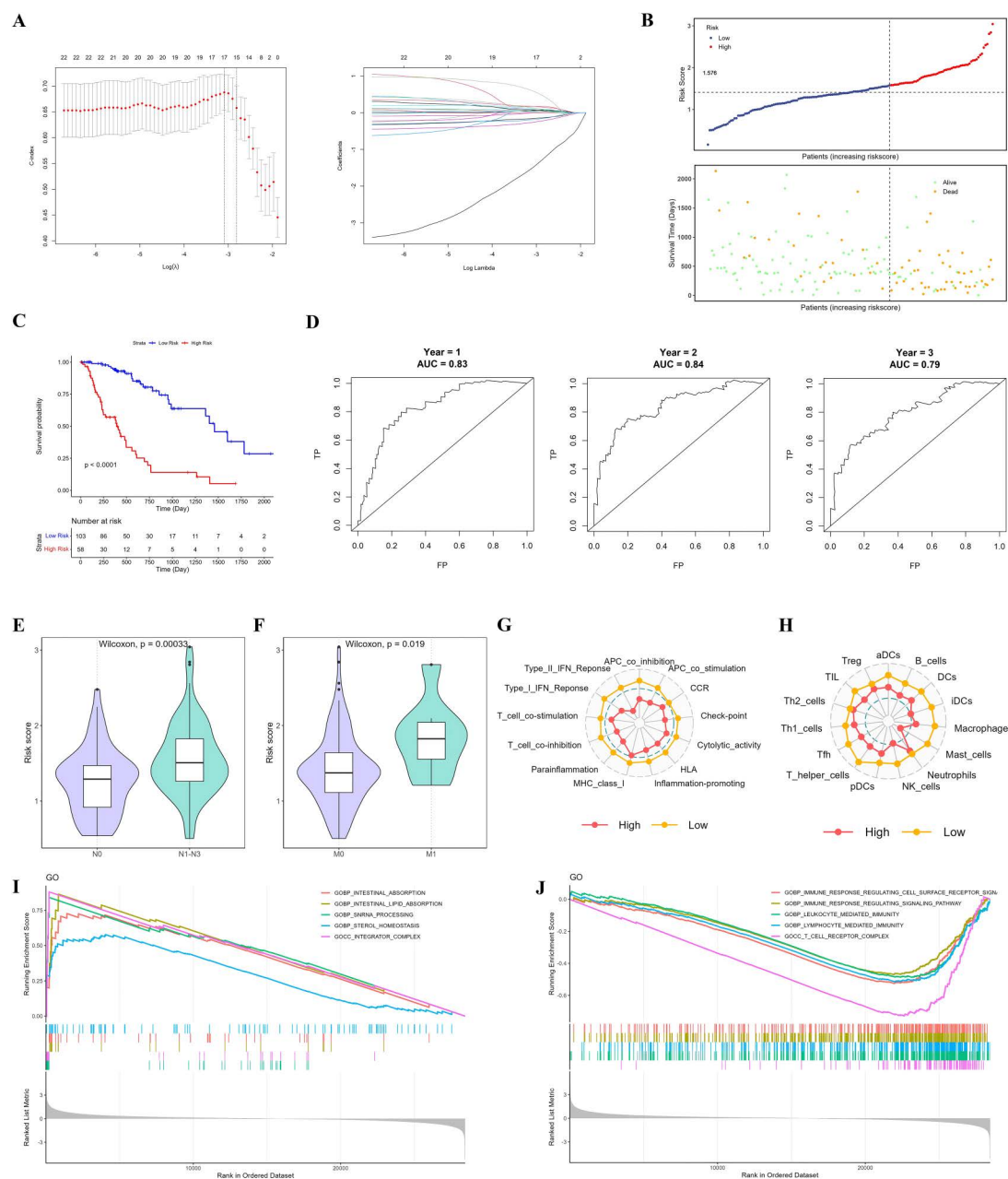
**Fig 3. Construction and evaluation of a prognostic risk model for EC.** (A) LASSO Cox regression analysis of 22 prognosis-associated DEGs, with the optimal lambda value determined through 10-fold cross-validation. (B) Distribution of survival status among patients in the low- and high-risk groups. (C) Kaplan–Meier survival analysis indicating significantly improved overall survival in the low-risk group compared to the high-risk group. (D) Time-dependent ROC curves demonstrating the predictive accuracy of the risk model for 1-, 2-, and 3-year overall survival. (E, F) Box plots comparing risk scores across N (E) and M (F) staging categories. (G, H) Radar plots illustrating differences in immune function enrichment (G) and immune cell infiltration (H) between the low- and high-risk groups. (I, J) GSEA showing significant enrichment of metabolism-related pathways in the high-risk group (I) and immune-related pathways in the low-risk group (J).

https://doi.org/10.1371/journal.pone.0340850.g003

a strong association between metabolic reprogramming, immune exclusion, and poor prognosis in EC, indicating that disrupted immune regulation and altered metabolism may jointly drive disease progression.

## Identification of *RIBC2* as a potential diagnostic and prognostic biomarker

Among the diagnostic and prognostic biomarkers identified, *RIBC2* expression was found to be significantly upregulated in both transformed Het-1A-T cells (Fig 4A) and esophageal tumor tissues (Fig 4B). Kaplan–Meier survival analysis demonstrated that high *RIBC2* expression was strongly associated with poorer overall survival in EC patients (Fig 4C). Additionally, it was observed that the adverse prognostic impact of RIBC2 high expression was significant in EAC, whereas in ESCC the Kaplan-Meier curves did not show a statistically significant difference between the low- and high-RIBC2 expression groups (S3 Fig). Furthermore, ROC curve analysis confirmed the diagnostic utility of *RIBC2* in EC, yielding an AUC of 0.793 (Fig 4D). Specifically, RIBC2 showed diagnostic potential in EC with AUCs of 0.772 for EAC and 0.812 for ESCC, respectively (S4 Fig). Immune correlation analysis revealed that patients with elevated *RIBC2* expression exhibited higher levels of immune checkpoint molecules such as *CD276* and *NT5E*, while showing reduced expression of *IL6R* (Fig 4E). In the context of the cancer–immunity cycle, high *RIBC2* expression was linked to enhanced activity in the recognition of cancer cells by T cells phase, but decreased activity in the immune cell infiltration into tumors phase (Fig 4F). Additionally, these patients demonstrated significantly higher TIDE exclusion scores, indicative of an immune-excluded phenotype (Fig
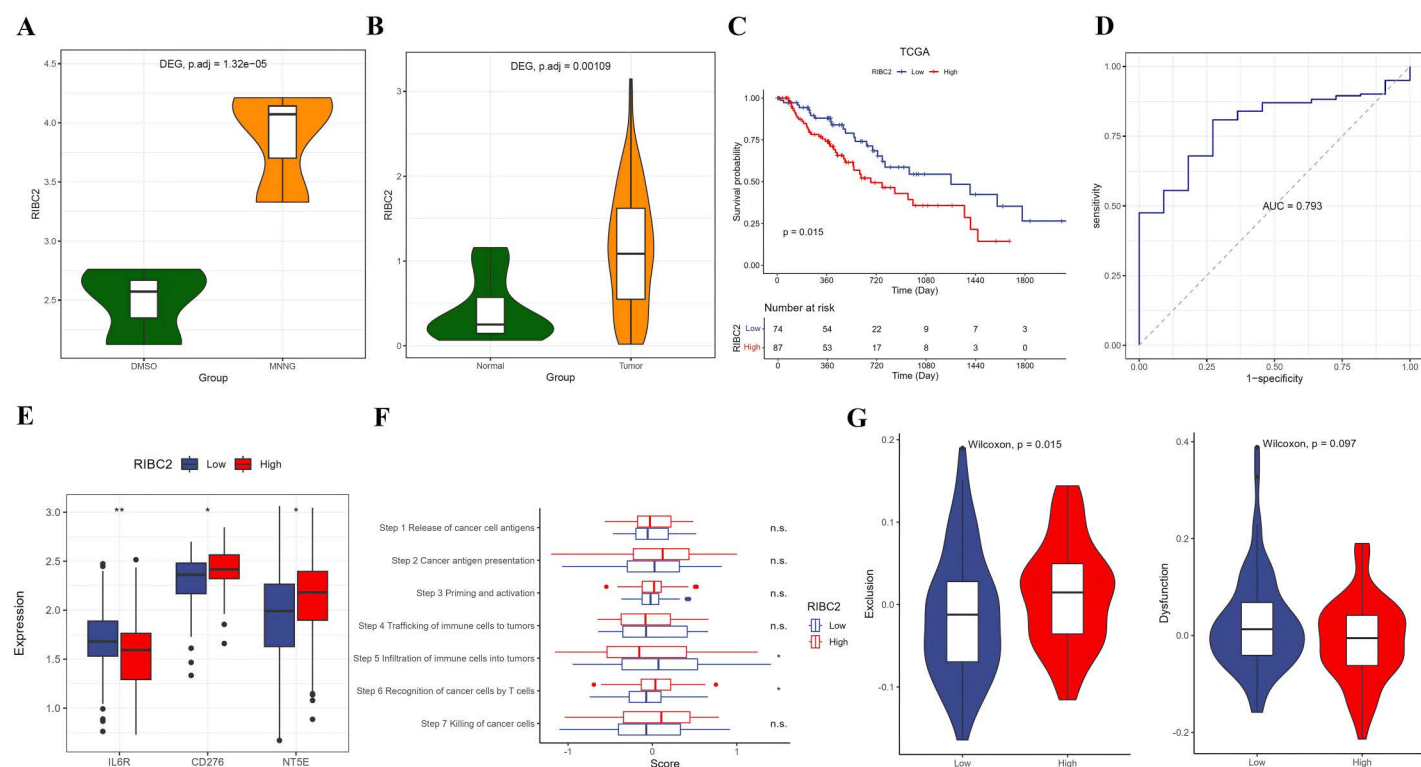


**Fig 4. RIBC2 in diagnosis, prognosis and immunity of EC.** (A, B) Differential expression of RIBC2 between Het-1A-T and Het-1A-N cells (A) and between tumor and normal tissues in the TCGA-EC cohort (B). (C) Kaplan-Meier survival curves demonstrating that high RIBC2 expression correlates with poorer overall survival in EC patients. (D) ROC curve illustrating the diagnostic accuracy of RIBC2 expression for distinguishing EC from normal tissues. (E) Comparison of immune checkpoint gene expression between high- and low-RIBC2 expression groups. (F) Correlation analysis between RIBC2 expression and the cancer–immunity cycle. (G) Comparison of TIDE exclusion and dysfunction scores across high- and low-RIBC2 expression groups, reflecting differences in potential immunotherapy responsiveness.

https://doi.org/10.1371/journal.pone.0340850.g004

4G). We also found that the infiltration of CD4 Tem and CD8 T cells were significantly higher in the low RIBC2 expression group, while the abundance of Th2 cells were much higher in the high RIBC2 expression group (S5A Fig). Further GSEA analysis showed that pyrimidine metabolism, cell cycle, mismatch repair and DNA replication were activated, while O-glycan biosynthesis and sphingolipid metabolism were suppressed in the high RIBC2 expression group (S5B Fig). Moreover, we identified a total of 1,255 DEGs between low- and high-RIBC2 expression groups (S5C Fig). These DEGs were significantly enriched into KEGG pathways of neuroactive ligand-receptor interaction, cornified envelope formation, nicotine addition, biological processes of immune response and peptidase activity (S5D Fig). Moreover, by PPI analysis, we found that RIBC2 may interact with TEKT2, C5orf49 and CFAP299 (S5E Fig). Collectively, these findings suggest that *RIBC2* serves as both a diagnostic and prognostic biomarker for EC and may also play a crucial role in modulating the tumor immune microenvironment, potentially contributing to immune evasion and reduced responsiveness to immunotherapy.

We subsequently validated the expression pattern and clinical relevance of *RIBC2* using both in vitro models and clinical samples. As shown in Fig 5A and 5B, *RIBC2* expression was significantly higher in Het-1A-T cells compared with Het-1A-N cells. Similarly, elevated *RIBC2* levels were observed in multiple EC cell lines including KYSE-410, TE-10, and
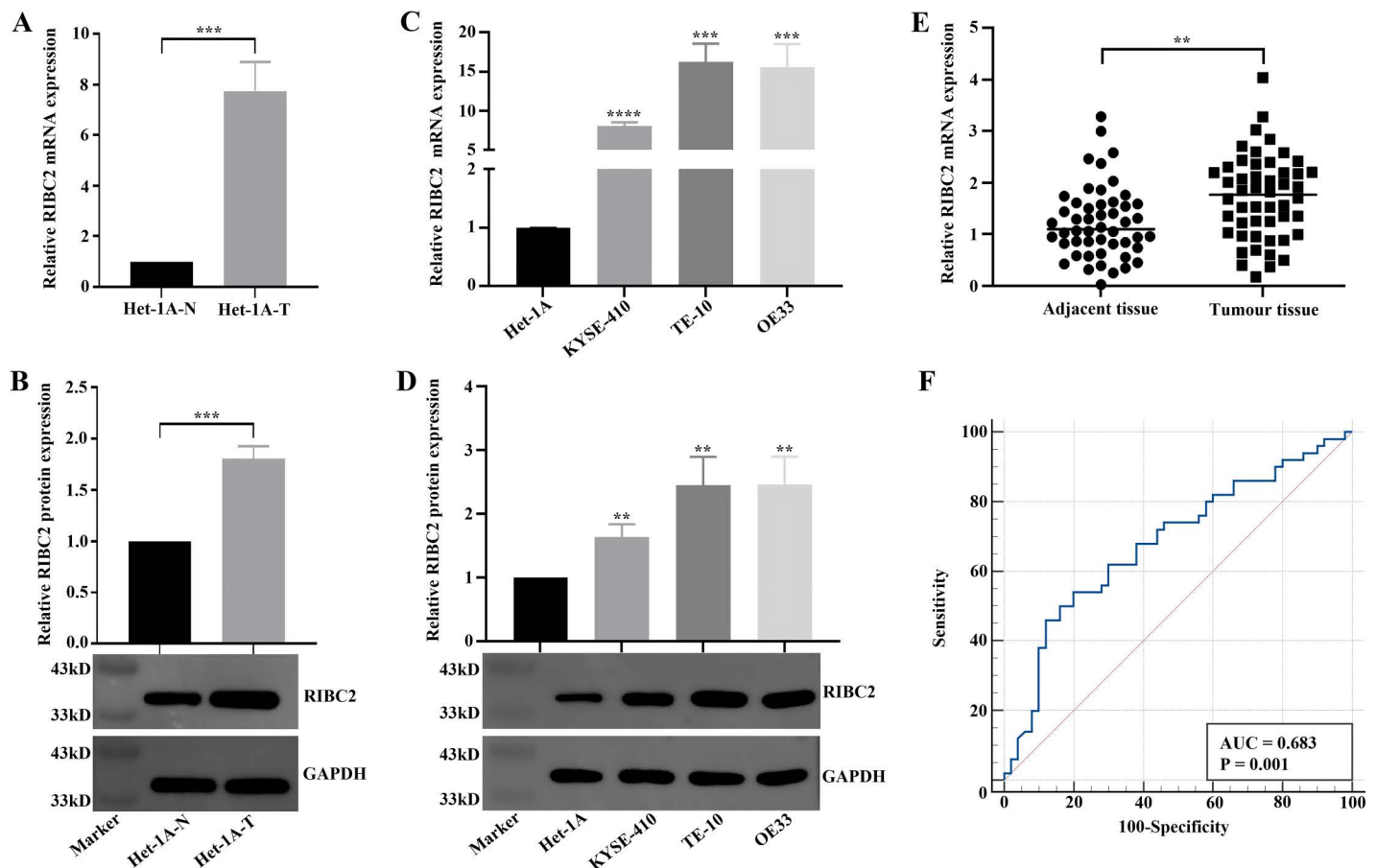


**Fig 5. Validation of RIBC2 expression and its clinical relevance in EC.** (A, B) Quantification of RIBC2 mRNA (A) and protein (B) expression levels in Het-1A-N and Het-1A-T cells. (C, D) Comparison of RIBC2 mRNA (C) and protein (D) expression levels between the normal esophageal epithelial cell line (Het-1A) and EC cell lines (KYSE-410, TE-10, and OE33). (E) Analysis of RIBC2 mRNA expression in 50 paired EC tissues and their corresponding adjacent normal tissues. of EC. (F) ROC curve illustrating the diagnostic performance of tissue RIBC2 expression. **P < 0.01, ***P < 0.001 and ****P < 0.0001.

https://doi.org/10.1371/journal.pone.0340850.g005

OE33 relative to the non-tumorigenic esophageal epithelial cell line Het-1A (Fig 5C and 5D). To validate these findings in vivo, RT-qPCR analysis was conducted on paired tumor and adjacent non-tumor tissues from EC patients. Consistent with the cell line data, *RIBC2* expression was markedly upregulated in tumor tissues (Figs 5E and S6). Clinicopathological correlation analysis revealed that high *RIBC2* expression was significantly associated with older patient age and advanced tumor stage (S3 Table). Furthermore, ROC curve analysis demonstrated that *RIBC2* effectively distinguished EC tissues from adjacent normal tissues, yielding an AUC of 0.683 (Fig 5F). Collectively, these findings confirm that *RIBC2* is consistently overexpressed in EC and is correlated with unfavorable clinical characteristics, highlighting its potential as a clinically relevant biomarker.

## Promotion of malignant phenotypes in Het-1A-T cells by RIBC2

Since *RIBC2* was found to be significantly upregulated in transformed cells, we next examined its functional role in malignant Het-1A-T cells, using TE-10 and OE33 EC cell lines as positive controls. Three short hairpin RNAs (shRNAs) targeting *RIBC2*, sh-*RIBC2*–1, sh-*RIBC2*–2, and sh-*RIBC2*–3 were designed, all of which effectively reduced *RIBC2* mRNA expression (Fig 6A). Among them, sh-*RIBC2*–2 (hereafter referred to as sh-*RIBC2*) exhibited the most efficient knockdown and was selected for further analysis. Western blotting confirmed a marked decrease in RIBC2 protein levels following transfection (Fig 6B). Functionally, *RIBC2* silencing significantly suppressed cell proliferation, as demonstrated by CCK-8 and colony formation assays (Fig 6C and 6D). Moreover, wound healing and Transwell invasion assays revealed that RIBC2 knockdown notably reduced the migratory and invasive abilities of Het-1A-T, TE-10, and OE33 cells (Fig 6E and 6F). Taken together, these results demonstrate that *RIBC2* promotes the proliferative, migratory, and invasive behavior of transformed esophageal epithelial cells, underscoring its potential role as a key driver of malignant transformation in EC.

## Discussion

EC continues to be a malignancy with a poor prognosis, primarily due to the lack of reliable early diagnostic approaches and an incomplete understanding of the molecular mechanisms underlying tumor initiation. In this study, we utilized a well-established in vitro malignant transformation model of esophageal epithelial cells to examine the molecular processes involved in EC tumorigenesis, a system that has been extensively applied in both in vitro and in vivo research [17–19]. The transformed cells demonstrated markedly increased proliferation, migration, invasion, and tumor-forming ability in nude mice, consistent with previous findings. This model thus provides a robust experimental platform for transcriptomic profiling and the identification of key molecular drivers contributing to EC development.

Transcriptome sequencing of normal and transformed esophageal epithelial cells, combined with publicly available transcriptomic datasets from EC tissues, facilitated the identification of key candidate biomarkers using a machine learning–based diagnostic framework. Among the identified genes, the top five contributors to diagnostic accuracy were *ESM1*, *ABCC6*, *CXCL8*, *RIBC2*, and *APLN*. Notably, *ESM1* has been reported to be highly expressed across various cancer types including ovarian, gastric, breast, esophageal, and cervical cancers and is known to promote tumor growth and progression primarily through its roles in angiogenesis and fatty acid metabolism [20–24]. In EC, *ESM1* expression is markedly upregulated in tumor tissues and shows a positive correlation with advanced tumor stage and poorer survival outcomes. Functional studies have demonstrated that *ESM1* knockdown can suppress EC cell proliferation, migration, and invasion [23]. *ABCC6*, a member of the ATP-binding cassette (ABC) transporter family, plays diverse roles in human diseases. For instance, *ABCC6* deficiency leads to pseudoxanthoma elasticum, a disorder characterized by ectopic calcification [25]. In hepatocellular carcinoma, *ABCC6* silencing has been shown to enhance cell proliferation and suppress apoptosis through activation of the PPAR signaling pathway [26]. Moreover, *ABCC6* has been implicated in the regulation of lipid and cholesterol metabolism [27,28]. However, its specific function in EC remains poorly understood. A recent study by Zhang et al. suggested that *ABCC6* may serve as a potential target of cranberry proanthocyanidins, which exhibit inhibitory effects on EC growth [29]. The significance of *CXCL8* (also known as interleukin-8) has been well established in
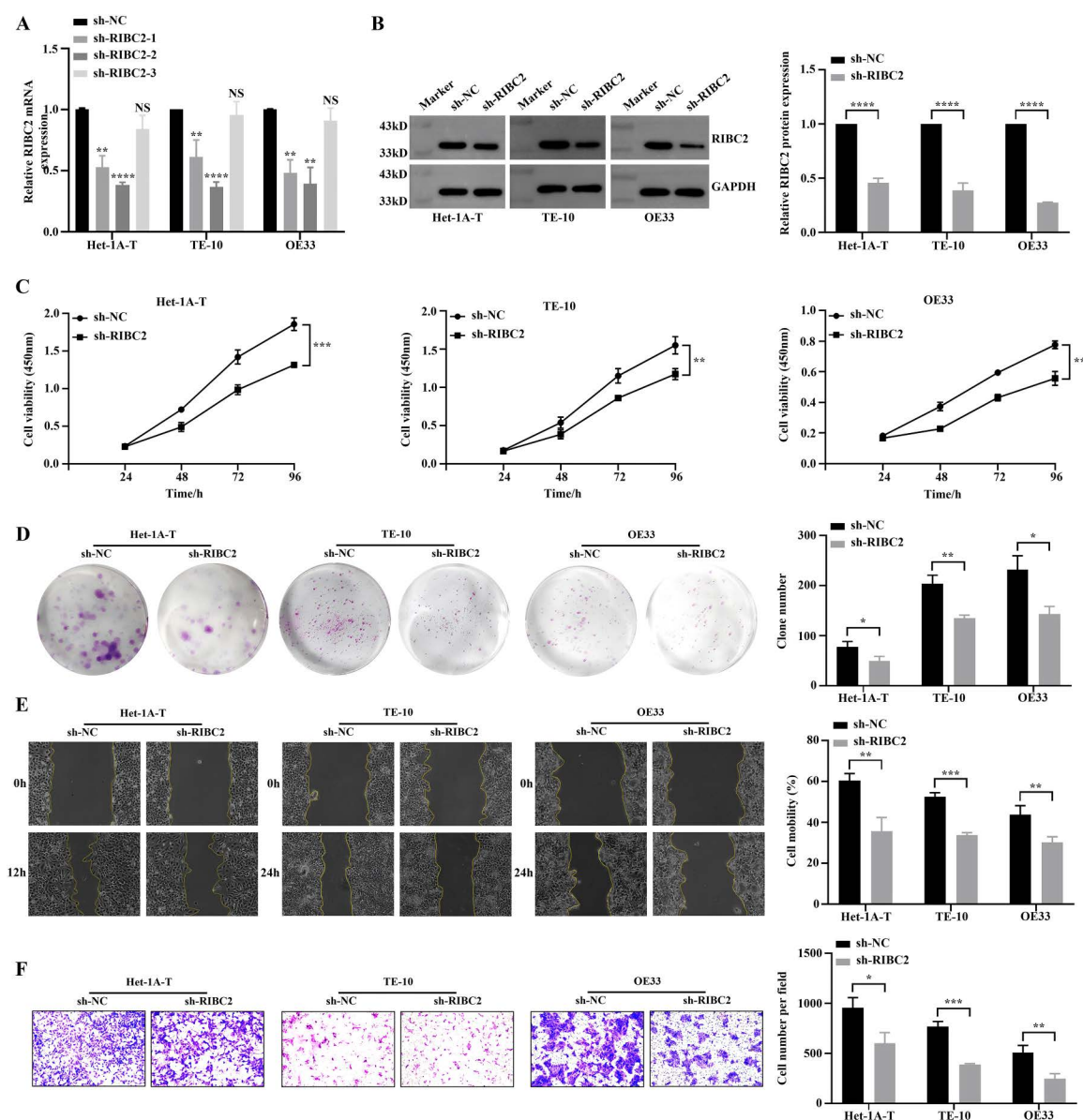
**Fig 6. RIBC2 promotes malignant transformation of human esophageal epithelial cells.** (A, B) Quantification of RIBC2 mRNA (A) and protein (B) expression levels following transfection with RIBC2-specific shRNA or negative control shRNA. (C-F) Functional assessment of cell proliferation (C), colony formation (D), migration (magnification, 40×) (E), and invasion (magnification, 100×) (F) in RIBC2 knockdown and control groups. $^{NS}P > 0.05$, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$.

both inflammatory disorders and cancer biology. It is a cytokine secreted by immune, stromal, and tumor cells, playing a pivotal role in tumor initiation, progression, and therapeutic response through interactions with its specific receptors [30]. The molecular functions of *CXCL8* in cancer development and therapy have been extensively characterized. *CXCL8* promotes cell proliferation, invasion, and survival, while also contributing to tumor stemness, angiogenesis, and remodeling of the tumor microenvironment through activation of the PI3K/Akt, MAPK, JAK/STAT3, and FAK/Src signaling pathways, as comprehensively reviewed by Xiong et al [31]. APLN binds to its receptor APLNR and participates in the pathogenesis of

cancers [32], heart diseases [33], reproduction dysfunction [34] and diabetes [35]. In EC, *APLN* has been shown to influence cell proliferation, apoptosis, migration, and stemness, with several studies exploring the underlying signaling mechanisms [36,37]. Regarding *RIBC2*, it has been identified as prognostic markers in liver and cervical cancers by bioinformatic analysis [38,39]. Higher expressions of RIBC2 were associated with better survival in cervical cancer, but worse outcome in liver cancer. However, its biological function and molecular mechanisms in cancers remain unclear. It has been reported that RIBC2 is expressed in motile cilia and plays a crucial role in ciliary movement by interacting with YbX2 [40], which could modulate mRNA stability, stemness and chemoresistance in endometrial cancer [41,42].

However, to date, the biological role and mechanistic significance of RIBC2 in EC remain largely unexplored, representing an important gap in current research. Accordingly, we conducted bioinformatic, clinical, and in vitro investigations to elucidate the role of *RIBC2* in EC. Transcriptomic analyses revealed that *RIBC2* expression was markedly elevated in both EC tissues and MNNG-induced malignantly transformed esophageal epithelial cells, with its expression levels closely correlated with EC onset and progression. These observations were subsequently validated using clinical specimens, confirming the upregulation and clinical relevance of *RIBC2* in EC. Functional experiments further demonstrated that *RIBC2* modulates key malignant phenotypes, including cell proliferation, migration, and invasion, in transformed esophageal cells and in both histological subtypes (ESCC and EAC cell lines). These results suggest that MNNG may contribute to esophageal carcinogenesis at least in part via *RIBC2* upregulation. We note that ESCC and EAC differ in cell of origin, risk factors and molecular architectures. Although RIBC2 knockdown produced similar phenotypes across ESCC (KYSE-410, TE-10) and EAC (OE33) cells, the underlying pathways may differ by histology. Future work using subtype-matched 3D organoids or in vivo models, together with patient-level analyses stratified by histology will be important to define subtype-specific mechanisms and clinical relevance of *RIBC2*. In addition, to validate the molecular mechanisms of RIBC2, we will confirm the putative interaction network by co-immunoprecipitation (co-IP) for candidate partners (TEKT2, C5orf49, CFAP299) in ESCC (KYSE-410, TE-10) and EAC (OE33) cells, followed by subcellular localization. Second, guided by our enrichment results, we will determine pathway-related proteins (e.g., keratinization/cornified-envelope markers such as KRT5/KRT14; neuroactive ligand–receptor nodes including nAChR subunits with downstream p-ERK/p-AKT) to evaluate whether pathway modulation attenuates RIBC2-dependent proliferation/migration/invasion. Furthermore, patients with low versus high RIBC2 expression exhibit differences in the immune microenvironment and in activated/ suppressed pathways, suggesting that therapeutic strategies may need to be tailored by RIBC2 status. Multi-omics, including genomics, epigenomics, transcriptomics, proteomics/ phospho-proteomics, metabolomics, and radiomics, provides a useful tool to translate laboratory findings into therapeutically relevant biomarkers [43,44]. Briefly, integrating multi-omics with network pharmacology can identify druggable pathways and network nodes associated with RIBC2, and develop subtype-specific strategies for patients with high versus low RIBC2 expression. Thus, in the future work, we will stratify ESCC and EAC cohorts by RIBC2 expression and perform multi-omics profiling to refine pathway signatures in the high- versus low-RIBC2 groups. Using these signatures, we will select candidate drugs via network pharmacology, followed by molecular docking against potential targets and experimental validation.

Collectively, these findings establish *RIBC2* as a previously unrecognized diagnostic and prognostic biomarker in EC, potentially contributing to tumor initiation and progression by enhancing malignant cellular behaviors. Its consistent overexpression across both experimental models and clinical samples, along with its association with an immune-excluded tumor microenvironment, underscores its potential as a biomarker for early detection and a novel therapeutic target. For example, RIBC2 can serve as a biopsy marker to aid pathologists in evaluating indeterminate dysplasia or very early lesions. Beyond tissue, minimally invasive cytology (esophageal brush/balloon) could quantify RIBC2 mRNA as part of a small RT-qPCR panel to improve detection in at-risk patients. In addition, cfRNA/EV-RNA assays may offer a nonendoscopic option for high-risk screening and post-treatment surveillance. It is worth noting that the accurate diagnostic value of RIBC2 needs validation in the real clinical world across different cohorts. Further mechanistic and translational studies will be essential to elucidate the precise role of *RIBC2* in EC pathogenesis and to determine whether therapeutic targeting of *RIBC2* could help suppress tumor initiation or enhance patient responsiveness to existing therapies.

## Supporting information

**S1 Table. The basic information of EC patients in low- and high-risk groups.**
(DOCX)

**S2 Table. List of 22 genes significantly associated with the prognosis of EC patients.**
(DOCX)

**S3 Table. Correlation analysis between RIBC2 expression levels and various clinicopathological characteristics of EC patients.**
(DOCX)

**S1 Fig. Integrated transcriptomic analysis revealing common DEGs shared between the in vitro malignant transformation model and clinical EC tissue datasets.** (A, B) Volcano plots and heatmap showing DEGs between Het-1A-T and Het-1A-N cells. (C, D) Volcano plots and heatmap showing DEGs between tumor and normal tissues in the TCGA-EC cohort. (E) Venn diagram showing the overlap between DEGs identified in the in vitro model and in vivo tissues. (F) Functional enrichment analysis of the 22 prognosis-related DEGs.
(TIF)

**S2 Fig. Kaplan–Meier survival analysis of the low- and high-risk groups in the TCGA-EAC and TCGA-ESCC cohorts.**
(JPG)

**S3 Fig. Kaplan–Meier survival analysis of the low- and high-RIBC2 expression groups in the TCGA-EAC and TCGA-ESCC cohorts.**
(JPG)

**S4 Fig. ROC curves of RIBC2 in the diagnosis of TCGA-EAC and TCGA-ESCC patients.**
(JPG)

**S5 Fig. RIBC2-associated immune context, differential expression, pathway enrichment and PPI network in EC.**
(A) Comparison of immune cell infiltration between low- and high-RIBC2 expression groups. (B) GSEA for KEGG pathways using genes ranked by the differential signal between high- and low-RIBC2 expression groups. (C) Volcano plot showing DEGs between high- and low-RIBC2 expression groups. (D) Functional enrichment analysis of DEGs between high- and low-RIBC2 expression groups. (E) PPI network highlighting RIBC2 and its candidate partners derived from STRING. Edges indicate curated or predicted interactions.
(JPG)

**S6 Fig. Validation of RIBC2 by RT-qPCR using RIBC2 primer 2.** Analysis of RIBC2 mRNA expression in 20 paired EC tissues and their corresponding adjacent normal tissues.
(JPG)

**S1 File. The original blots.**
(PDF)

## Author contributions

**Conceptualization:** Guogui Sun.

**Data curation:** Xuan Zheng, Xuemin Yao.

**Formal analysis:** Xuan Zheng, Yishuang Cui, Yanan Wu.

**Investigation:** Weinan Yao.

**Methodology:** Xuan Zheng, Yishuang Cui.

**Project administration:** Guogui Sun.

**Resources:** Yanna Bi.

**Software:** Xuan Zheng, Yishuang Cui, Xuemin Yao, Junqing Gan.

**Supervision:** Guogui Sun.

**Validation:** Yanlei Ge, Ye Jin.

**Visualization:** Xuan Zheng.

**Writing – original draft:** Xuan Zheng.

**Writing – review & editing:** Guogui Sun.

## References

1. Li J, Xu J, Zheng Y, Gao Y, He S, Li H, et al. Esophageal cancer: Epidemiology, risk factors and screening. Chin J Cancer Res. 2021;33(5):535–47. https://doi.org/10.21147/j.issn.1000-9604.2021.05.01 PMID: 34815628

2. Shoji Y, Koyanagi K, Kanamori K, Tajima K, Ogimi M, Ninomiya Y, et al. Immunotherapy for esophageal cancer: Where are we now and where can we go. World J Gastroenterol. 2024;30(19):2496–501. https://doi.org/10.3748/wjg.v30.i19.2496 PMID: 38817664

3. Hu J, Liu F, Zhang J, Yin L, Cao W, Xu W, et al. Spatially resolved transcriptomic analysis of the adult human prostate. Nat Genet. 2025;57(4):922–33. https://doi.org/10.1038/s41588-025-02139-9 PMID: 40169792

4. Zhou R, Tang X, Wang Y. Emerging strategies to investigate the biology of early cancer. Nat Rev Cancer. 2024;24(12):850–66. https://doi.org/10.1038/s41568-024-00754-y PMID: 39433978

5. Ren Q, Zhang P, Zhang X, Feng Y, Li L, Lin H, et al. A fibroblast-associated signature predicts prognosis and immunotherapy in esophageal squamous cell cancer. Front Immunol. 2023;14:1199040. https://doi.org/10.3389/fimmu.2023.1199040 PMID: 37313409

6. Yuan P, Gao X, Xu M, Qiu L, Xiong Z, Shen J, et al. Novel miRNA markers and their mechanism of esophageal squamous cell carcinoma (ESCC) based on TCGA. Sci Rep. 2024;14(1):27261. https://doi.org/10.1038/s41598-024-76321-0 PMID: 39516222

7. Han D, Han Y, Guo W, Wei W, Yang S, Xiang J, et al. High-dimensional single-cell proteomics analysis of esophageal squamous cell carcinoma reveals dynamic alterations of the tumor immune microenvironment after neoadjuvant therapy. J Immunother Cancer. 2023;11(11):e007847. https://doi.org/10.1136/jitc-2023-007847 PMID: 38016720

8. Wu H, Leng X, Liu Q, Mao T, Jiang T, Liu Y, et al. Intratumoral Microbiota Composition Regulates Chemoimmunotherapy Response in Esophageal Squamous Cell Carcinoma. Cancer Res. 2023;83(18):3131–44. https://doi.org/10.1158/0008-5472.CAN-22-2593 PMID: 37433041

9. Khaghani F, Hemmati M, Ebrahimi M, Salmaninejad A. Emerging Multi-omic Approaches to the Molecular Diagnosis of Mitochondrial Disease and Available Strategies for Treatment and Prevention. Curr Genomics. 2024;25(5):358–79. https://doi.org/10.2174/0113892029308327240612110334 PMID: 39323625

10. Li L, Jiang D, Zhang Q, Liu H, Xu F, Guo C, et al. Integrative proteogenomic characterization of early esophageal cancer. Nat Commun. 2023;14(1):1666. https://doi.org/10.1038/s41467-023-37440-w PMID: 36966136

11. Liu Z, Zhao Y, Kong P, Liu Y, Huang J, Xu E, et al. Integrated multi-omics profiling yields a clinically relevant molecular classification for esophageal squamous cell carcinoma. Cancer Cell. 2023;41(1):181-195.e9. https://doi.org/10.1016/j.ccell.2022.12.004 PMID: 36584672

12. Tian F, He X, Wang S, Liang Y, Wang Z, Hu M, et al. Integrating single-cell sequencing and machine learning to uncover the role of mitophagy in subtyping and prognosis of esophageal cancer. Apoptosis. 2025;30(3–4):1021–41. https://doi.org/10.1007/s10495-024-02061-1 PMID: 39948301

13. Li L, Qin Z, Bo J, Hu J, Zhang Y, Qian L, et al. Machine learning-based radiomics prognostic model for patients with proximal esophageal cancer after definitive chemoradiotherapy. Insights Imaging. 2024;15(1):284. https://doi.org/10.1186/s13244-024-01853-y PMID: 39613907

14. Sun Y, Tang Y, Qi Q, Pang J, Chen Y, Wang H, et al. 101 Machine Learning Algorithms for Mining Esophageal Squamous Cell Carcinoma Neoantigen Prognostic Models in Single-Cell Data. Int J Mol Sci. 2025;26(7):3373. https://doi.org/10.3390/ijms26073373 PMID: 40244296

15. Fang Y, Kong Y, Rong G, Luo Q, Liao W, Zeng D. Systematic Investigation of Tumor Microenvironment and Antitumor Immunity With IOBR. Med Research. 2025;1(1):136–40. https://doi.org/10.1002/mdr2.70001

16. Zhang J, Li H, Tao W, Zhou J. GseaVis: An R Package for Enhanced Visualization of Gene Set Enrichment Analysis in Biomedicine. Med Research. 2025;1(1):131–5. https://doi.org/10.1002/mdr2.70000

17. Chen Y, Feng H, Zhang H, Li X. High expression of DNA methyltransferase 1 in Kazakh esophageal epithelial cells may promote malignant transformation induced by N-methyl-N'-nitro-N-nitrosoguanidine. Hum Exp Toxicol. 2019;38(9):1060–8. https://doi.org/10.1177/0960327119851254 PMID: 31117802

18. Li SY, Guo YS, Yi L, Chen Y. The role of LncRNA TUG1 in DNA damage and malignant transformation induced by Helicobacter pylori and N-methyl-N'-nitro-N-nitrosoguanidine on human esophageal epithelial cells. Food Chem Toxicol. 2024;192:114928.

19. Wang J, Zhang W, Zhang R, Yang H, Li Y, Wang J, et al. MiR-101-3p Promotes Tumor Cell Proliferation and Migration via the Wnt Signal Pathway in MNNG-Induced Esophageal Squamous Cell Carcinoma. Toxics. 2024;12(11):824. https://doi.org/10.3390/toxics12110824 PMID: 39591002

20. Li D, Su X, Xue S, Yao L, Yu D, Tang X, et al. Targeting ESM1/ VEGFα signaling axis: a promising therapeutic avenue for angiogenesis in cervical squamous cell carcinoma. J Cancer. 2023;14(10):1725–35. https://doi.org/10.7150/jca.84654 PMID: 37476182

21. Liu WT, Yang Y, He BC, Ma F, Sun F, Guo M. ESM1 promotes triple-negative breast cancer cell proliferation through activating AKT/NF-κB/Cyclin D1 pathway. Ann Transl Med. 2021;12(5):104.

22. Yang J, Shu G, Chen T, Dong A, Dong C, Li W, et al. ESM1 Interacts with c-Met to Promote Gastric Cancer Peritoneal Metastasis by Inducing Angiogenesis. Cancers (Basel). 2023;16(1):194. https://doi.org/10.3390/cancers16010194 PMID: 38201620

23. Li J, Yang D, Zhang C, Wei S, Zhao R, Dai S, et al. ESM1 Is a Promising Therapeutic Target and Prognostic Indicator for Esophageal Carcinogenesis/Esophageal Squamous Cell Carcinoma. Biomed Res Int. 2022;2022:5328192. https://doi.org/10.1155/2022/5328192 PMID: 35937390

24. Zhang J, Ouyang F, Gao A, Zeng T, Li M, Li H, et al. ESM1 enhances fatty acid synthesis and vascular mimicry in ovarian cancer by utilizing the PKM2-dependent warburg effect within the hypoxic tumor microenvironment. Mol Cancer. 2024;23(1):94. https://doi.org/10.1186/s12943-024-02009-8 PMID: 38720298

25. Favre G, Laurain A, Aranyi T, Szeri F, Fulop K, Le Saux O, et al. The ABCC6 Transporter: A New Player in Biomineralization. Int J Mol Sci. 2017;18(9):1941. https://doi.org/10.3390/ijms18091941 PMID: 28891970

26. Zhao Z, Zhao Z, Wang J, Zhang H, Xi Z, Xia Q. ABCC6 Knockdown Fuels Cell Proliferation by Regulating PPARα in Hepatocellular Carcinoma. Front Oncol. 2022;12:840287. https://doi.org/10.3389/fonc.2022.840287 PMID: 35280774

27. Ibold B, Tiemann J, Faust I, Ceglarek U, Dittrich J, Gorgels TGMF, et al. Genetic deletion of Abcc6 disturbs cholesterol homeostasis in mice. Sci Rep. 2021;11(1):2137. https://doi.org/10.1038/s41598-021-81573-1 PMID: 33483533

28. Plümers R, Osterhage MR, Lindenkamp C, Knabbe C, Hendig D. Targeting ABCC6 in Mesenchymal Stem Cells: Impairment of Mature Adipocyte Lipid Homeostasis. Int J Mol Sci. 2022;23(16):9218. https://doi.org/10.3390/ijms23169218 PMID: 36012482

29. Zhang Y, Weh KM, Tripp BA, Clarke JL, Howard CL, Sunilkumar S, et al. Cranberry Proanthocyanidins Mitigate Reflux-Induced Transporter Dysregulation in an Esophageal Adenocarcinoma Model. Pharmaceuticals (Basel). 2023;16(12):1697. https://doi.org/10.3390/ph16121697 PMID: 38139823

30. Liu Q, Li A, Tian Y, Wu JD, Liu Y, Li T, et al. The CXCL8-CXCR1/2 pathways in cancer. Cytokine Growth Factor Rev. 2016;31:61–71. https://doi.org/10.1016/j.cytogfr.2016.08.002 PMID: 27578214

31. Xiong X, Liao X, Qiu S, Xu H, Zhang S, Wang S, et al. CXCL8 in Tumor Biology and Its Implications for Clinical Translation. Front Mol Biosci. 2022;9:723846. https://doi.org/10.3389/fmolb.2022.723846 PMID: 35372515

32. Wang Q, Wang B, Zhang W, Zhang T, Liu Q, Jiao X, et al. APLN promotes the proliferation, migration, and glycolysis of cervical cancer through the PI3K/AKT/mTOR pathway. Arch Biochem Biophys. 2024;755:109983. https://doi.org/10.1016/j.abb.2024.109983 PMID: 38561035

33. Wang W, Shen M, Fischer C, Basu R, Hazra S, Couvineau P, et al. Apelin protects against abdominal aortic aneurysm and the therapeutic role of neutral endopeptidase resistant apelin analogs. Proc Natl Acad Sci U S A. 2019;116(26):13006–15. https://doi.org/10.1073/pnas.1900152116 PMID: 31189595

34. Song K, Yang X, An G, Xia X, Zhao J, Xu X, et al. Targeting APLN/APJ restores blood-testis barrier and improves spermatogenesis in murine and human diabetic models. Nat Commun. 2022;13(1):7335. https://doi.org/10.1038/s41467-022-34990-3 PMID: 36443325

35. Li C, Cheng H, Adhikari BK, Wang S, Yang N, Liu W, et al. The Role of Apelin-APJ System in Diabetes and Obesity. Front Endocrinol (Lausanne). 2022;13:820002. https://doi.org/10.3389/fendo.2022.820002 PMID: 35355561

36. Wang Y, Wang G, Liu X, Yun D, Cui Q, Wu X, et al. Inhibition of APLN suppresses cell proliferation and migration and promotes cell apoptosis in esophageal cancer cells *in vitro*, through activating PI3K/mTOR signaling pathway. Eur J Histochem. 2022;66(3):3336. https://doi.org/10.4081/ejh.2022.3336 PMID: 35920446

37. Zhou Y, Xu R, Luo J, Li X, Zhong Y, Sun Z. Dysregulation of miR-204-5p/APLN axis affects malignant progression and cell stemness of esophageal cancer. Mutat Res. 2022;825:111791. https://doi.org/10.1016/j.mrfmmm.2022.111791 PMID: 35930907

38. Feng N-N, Du X-Y, Zhang Y-S, Jiao Z-K, Wu X-H, Yang B-M. Overweight/obesity-related transcriptomic signature as a correlate of clinical outcome, immune microenvironment, and treatment response in hepatocellular carcinoma. Front Endocrinol (Lausanne). 2023;13:1061091. https://doi.org/10.3389/fendo.2022.1061091 PMID: 36714595

39. Li N, Yu K, Lin Z, Zeng D. Identifying a cervical cancer survival signature based on mRNA expression and genome-wide copy number variations. Exp Biol Med (Maywood). 2022;247(3):207–20. https://doi.org/10.1177/15353702211053580 PMID: 34674573

40. Kwon KY, Jeong H, Jang DG, Kwon T, Park TJ. Ckb and Ybx2 interact with Ribc2 and are necessary for the ciliary beating of multi-cilia. Genes Genomics. 2023;45(2):157–67. https://doi.org/10.1007/s13258-022-01350-w PMID: 36508087

41. Cai Y, Li N, Li H. YBX2 modulates mRNA stability via interaction with YTHDF2 in endometrial cancer cells. Exp Cell Res. 2023;427(1):113586. https://doi.org/10.1016/j.yexcr.2023.113586 PMID: 37030331

42. Suzuki I, Yoshida S, Tabu K, Kusunoki S, Matsumura Y, Izumi H, et al. YBX2 and cancer testis antigen 45 contribute to stemness, chemoresistance and a high degree of malignancy in human endometrial cancer. Sci Rep. 2021;11(1):4220. https://doi.org/10.1038/s41598-021-83200-5 PMID: 33602962

43. Liu Y, Zhang S, Liu K, Hu X, Gu X. Advances in drug discovery based on network pharmacology and omics technology. Current Pharmaceutical Analysis. 2024;21(1):33–43. https://doi.org/10.1016/j.cpan.2024.12.002

44. Wang Z, Zhao Y, Zhang L. Emerging trends and hot topics in the application of multi-omics in drug discovery: A bibliometric and visualized study. Current Pharmaceutical Analysis. 2024;21(1):20–32. https://doi.org/10.1016/j.cpan.2024.12.001