Electronic Journal of Biotechnology 82 (2026) 100718

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

journal homepage: www.elsevier.com/locate/ejbt

Research article

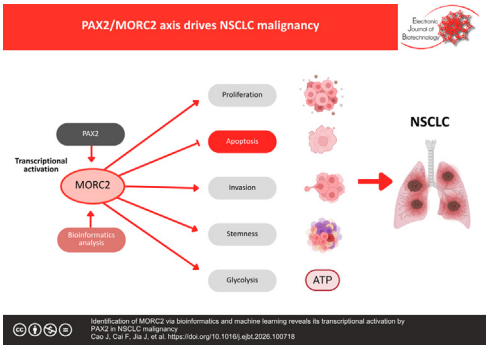
Identification of MORC2 via bioinformatics and machine learning reveals its transcriptional activation by PAX2 in NSCLC malignancy ☆

Junyan Cao, Fang Cai, Jun Jia, Jiangtao Guo, Xiu Yang *

Department of Oncology, Taiyuan Central Hospital, The Ninth Clinical Medical College of Shanxi Medical University, Taiyuan City, Shanxi, China

GRAPHICAL ABSTRACT

Identification of MORC2 via bioinformatics and machine learning reveals its transcriptional activation by PAX2 in NSCLC malignancy



ARTICLE INFO

Article history:
Received 28 December 2025
Accepted 28 April 2026
Available online 12 June 2026

Keywords:
Apoptosis
Glycolysis
Microrchidia family CW-type zinc finger 2
MORC2
Non-small cell lung cancer
Paired box 2
PAX2
Prognostic biomarker
Transcriptional regulation
Tumor invasion

ABSTRACT

Background: Non-small cell lung cancer (NSCLC) represents the top contributor to cancer-associated deaths globally, and its suboptimal clinical outcomes are linked to late-stage detection and acquired resistance to existing therapies. Microrchidia family CW-type zinc finger 2 (MORC2) has been found to exhibit oncogenic potential in NSCLC, but its transcriptional regulatory mechanisms and role in glycolytic reprogramming remain unclear.
Results: As a key gene, MORC2 was found to be overexpressed in NSCLC and linked to poor prognosis. MORC2 silencing impaired cell proliferation, stemness, and invasion, diminished glycolysis, and enhanced apoptosis. PAX2 was confirmed to directly transcriptionally activate MORC2; overexpression of MORC2 reversed the defects in malignant phenotypes induced by PAX2 knockdown. *In vivo*, depletion of PAX2 suppressed tumor growth, which was partially rescued by MORC2 overexpression.
Conclusions: The PAX2/MORC2 axis was demonstrated to drive NSCLC malignancy, supporting MORC2 as a potential prognostic marker and therapeutic target for NSCLC.

☆ Audio abstract available in Supplementary material.

Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

* Corresponding author.
E-mail address: zxyyyangxiu@163.com (X. Yang).

https://doi.org/10.1016/j.ejbt.2026.100718

Tumor progression

How to cite: Cao J, Cai F, Jia J, et al. Identification of MORC2 via bioinformatics and machine learning reveals its transcriptional activation by PAX2 in NSCLC malignancy. Electron J Biotechnol 2026;82. https://doi.org/10.1016/j.ejbt.2026.100718.

1. Introduction

Lung cancer (LC) is the most commonly diagnosed malignancy worldwide and remains the leading cause of cancer-related mortality, with approximately 2.5 million new cases and 1.8 million deaths reported globally in 2022 alone [1,2]. Based on histopathological classification, non-small cell lung cancer (NSCLC) accounts for 85–90% of all LC cases, and approximately 70% patients present with advanced-stage disease at the time of diagnosis [3,4,5]. The poor clinical outcomes of NSCLC are primarily attributed to late-stage diagnosis, high metastatic potential, and inevitable acquired resistance to current therapeutic strategies, which often emerges via target gene modification or alternative pathway activation [6,7,8,9] While advances in molecular biology have unraveled key oncogenic pathways driving NSCLC progression, the identification of critical molecular drivers and their regulatory networks, particularly those governing core malignant phenotypes such as uncontrolled proliferation, cancer stemness, and metabolic reprogramming, remains incomplete. This knowledge gap continues to hinder efforts to develop accurate diagnostic biomarkers and effective targeted treatment approaches.

Transcriptomic profiling of tumor tissues and matched adjacent normal tissues has become an indispensable tool for identifying differentially expressed genes (DEGs) associated with tumor initiation and progression, providing insights into molecular subtypes, prognostic markers, and therapeutic response predictors [10,11]. Publicly available datasets such as GSE29249, which includes paired NSCLC and adjacent normal tissues, have been widely utilized for DEG mining in LC research, serving as valuable resources for biomarker discovery [12]. However, the large volume of genomic data generated by such profiling requires robust analytical tools to prioritize biologically relevant candidates. Machine learning algorithms, including least absolute shrinkage and selection operator (LASSO), logistic regression and random forest (RF), have emerged as powerful approaches for feature selection from high-dimensional data: LASSO effectively reduces dimensionality by penalizing non-essential genes, while RF quantifies gene importance based on predictive model performance [13]. Yet their application to systematically identify key oncogenes linked to both phenotypic malignancy and metabolic alterations in NSCLC remains underexplored.

Microrchidia family CW-type zinc finger 2 (MORC2), an emerging chromatin remodeler localized mainly in the nucleus of cancer cells, has recently gained attention for its oncogenic roles [14,15]. Earlier investigations have shown that MORC2 was overexpressed in LC tissues in contrast to adjacent non-tumor tissues, and this expression pattern correlates positively with disease advancement. The underlying mechanisms include the activation of the Wnt/β-catenin signaling cascade, which serves to enhance angiogenesis and recruit TAMs [16]. Additionally, MORC2 has been implicated in glycolytic metabolism in cancer: It forms a complex with MYC-associated factor X (MAX) in breast cancer to regulate the expression levels of glycolytic enzymes, hexokinase 1 (HK1) and lactate dehydrogenase A (LDHA) included, and c-Myc has also been shown to transcriptionally regulate MORC2 to modulate LDHA expression and glucose metabolism, thereby enhancing cell proliferation and migration [17,18]. Given that metabolic reprogramming (e.g., the Warburg effect) is a core pathological feature of NSCLC that sustains rapid tumor growth, MORC2 may fulfill a conserved role in mediating the regulation of glycolytic metabolism in this disease [19]. Despite these observations, the precise mechanisms underlying MORC2 overexpression in NSCLC remain unclear; specifically, its transcriptional regulators have not been identified, and its potential role in mediating glycolytic reprogramming to drive NSCLC malignancy has not been experimentally validated.

Paired box 2 (PAX2), classified as one member of the PAX family of key regulatory transcriptional factors, is critical for embryonic development and has been increasingly linked to tumorigenesis through modulation of cell survival, proliferation, and invasion [20]. In NSCLC, PAX2 has demonstrated diagnostic utility for distinguishing NSCLC from small cell lung cancer (SCLC) [21]. However, its downstream transcriptional targets and molecular mechanisms in NSCLC pathogenesis remain poorly understood. Specifically, direct experimental evidence for a regulatory interaction between PAX2 and MORC2 is lacking. Furthermore, the functional relevance of a potential PAX2/MORC2 axis in driving key NSCLC malignant phenotypes, such as proliferation, invasion, and glycolysis, remains unexplored.

This study was designed to systematically identify and characterize key genes contributing to NSCLC pathogenesis. It integrated bioinformatic analysis of the GSE29249 dataset with machine learning approaches to screen for critical candidate genes. The research then focused on validating the expression and prognostic significance of the prioritized gene MORC2 across multiple datasets and clinical specimens. Subsequently, the impact of MORC2 on the functional phenotypes of NSCLC cells, including proliferation, metastasis, and glycolysis, was investigated. Finally, the study explored the hypothesis that PAX2 acted as an upstream transcriptional regulator of MORC2 and assessed the functional contribution of this PAX2/MORC2 axis to oncogenic processes both *in vitro* and *in vivo*.

2. Materials and methods

2.1. Bioinformatic analysis of DEGs and heatmap construction

The GSE29249 dataset, which includes 6 paired NSCLC tissues and adjacent normal tissues detected via microarray technology, was retrieved from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). Through the online platform Sangerbox (https://www.sangerbox.com/login.html), DEGs between NSCLC and normal tissues were identified by using the "limma" package in R, with filtering criteria set as adjusted p-value < 0.05 and absolute $\log_2$ fold change ($|\log_2 FC|$) > 1. The "pheatmap" package in R was employed to conduct hierarchical clustering analysis on the 10 prominently upregulated and 10 prominently downregulated protein-coding DEGs (with non-coding RNAs omitted). For sample clustering (columns), the distance metric was set to "Euclidean" (via dist.method = "euclidean") and the clustering method to "ward.D2" (via cluster rows = TRUE, cluster cols = TRUE, clustering method = "ward.D2"), which minimized within-cluster variance. For gene clustering (rows), the same distance metric and clustering method were applied. Expression values were normalized via z-score transformation

(row-wise) before plotting to ensure consistent scaling across genes.

2.2. Machine learning-based screening of key candidate genes

To identify candidate genes, two complementary machine learning-based analytical approaches were employed in the analysis of the top 10 upregulated and top 10 downregulated protein-coding DEGs. For LASSO logistic regression, the R package "glmnet 4.1.8" served to build a logistic regression model with L1 regularization (alpha = 1, since alpha = 1 specifies the LASSO penalty). A 3-fold cross-validation (K = 3) was implemented using a penalized logistic regression framework with L1 regularization (LASSO), with the response variable set as "tissue type" (NSCLC vs. normal) and predictors as the expression matrix of DEGs. λ.min (the λ value achieving the lowest cross-validation error) was set as the optimal penalty parameter, and genes displaying non-null coefficients at λ.min were identified as candidate features. For RF analysis, the R package "randomForest 4.7.1.2" was applied. A 3-fold cross-validation was performed to optimize the number of trees (ntree): tree counts were tested in the range of 100–1000 with a step size of 100 (i.e., 100, 200, . . ., 1000 trees), and the average test-set error rate was computed for each ntree. The optimal ntree was defined as the smallest tree count where the test-set error rate stabilized (no further significant decrease with increasing ntree). The "mean decrease in Gini index", a measure that gauges how greatly each gene lowers node impurity, was used to evaluate variable importance; genes with a mean decrease in Gini index > 5 (a predefined threshold based on model stability) were regarded as high-importance candidates. Via the "VennDiagram" package in R, genes overlapping between the two algorithms were designated as key candidates. Further screening of the literature showed that MORC2 can facilitate malignant progression in NSCLC, with further investigations into MORC2 accordingly initiated [16,22].

2.3. Bioinformatic validation of MORC2 expression and prognostic significance

Multiple public databases were utilized to confirm MORC2′s expression and its prognostic significance in NSCLC. The Cancer Genome Atlas (TCGA) Lung Adenocarcinoma (LUAD) dataset (515 tumor tissues vs. 59 adjacent normal tissues) and TNMPLOT database were analyzed to compare MORC2 mRNA expression between NSCLC and normal tissues. The Kaplan-Meier Plotter database (https://kmplot.com/analysis/) was used to evaluate prognosis: Patients were stratified into groups characterized by high versus low MORC2 expression levels; overall survival curves were generated accordingly, and significance levels were evaluated through the log-rank test. Additionally, the GSE29249 dataset was re-examined by employing the "limma" package within R software to confirm MORC2 expression differences.

2.4. Quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was isolated from 43 paired clinical specimens of NSCLC (tumor and corresponding adjacent normal tissues) and *in vitro* cultured cells with TRIzol reagent (Invitrogen, Carlsbad, CA, USA), strictly adhering to the manufacturer's protocol. Using 1 μg of total RNA, complementary DNA (cDNA) was produced by means of the PrimeScript RT Reagent Kit with gDNA Eraser (TaKaRa, Dalian, China). qRT-PCR was performed on a StepOnePlus Real-Time PCR System using SYBR Premix Ex Taq II (TaKaRa), with β-actin as the internal reference. Relative gene expression (MORC2 and PAX2) was calculated via the $2^{-\Delta\Delta Ct}$ method. Primers used are listed in Table 1.

Table 1
Primers sequences used for qPCR.

Name		Primers for PCR (5′-3′)
MORC2	Forward	TTCCCACTCGTATTTGGGGC
	Reverse	AGAGAGAGCAGGATGGCAGA
PAX2	Forward	CGGCAGTCCTGAAGTTGAGT
	Reverse	CTCGGAGCCCCTCTTTGTTT
β-actin	Forward	CTTCGCGGGCGACGAT
	Reverse	CCACATAGGAATCCTTCTGACC

2.5. Western blot (WB)

From clinical tissues, cultured cells, or xenograft tumor tissues, total protein was extracted with RIPA lysis buffer (Solarbio, Beijing, China) to prevent proteolysis. Protein concentration was assessed by means of the Bicinchoninic Acid (BCA) Protein Assay Kit (Solarbio). Equal amounts of protein (30 μg per lane) were separated by 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to polyvinylidene fluoride (PVDF) membranes (Millipore, Billerica, MA, USA). Membranes were blocked with 5% non-fat milk in Tris-buffered saline with Tween 20 (TBST) for 1 h at room temperature and then incubated overnight at 4°C with primary antibodies diluted in the same blocking solution: anti-MORC2 (ab315903, 1:1000, Abcam, Cambridge, UK), anti-PAX2 (PA5-118142, 1:1000, Invitrogen), and anti-β-actin (ab8227, 1:5000, Abcam). After TBST washes, membranes were incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies (goat anti-rabbit IgG (ab205718, 1:5000, Abcam); goat anti-mouse IgG (ab205719, 1:5000, Abcam)) for 1 h at room temperature. Visualization of protein bands was achieved using the Millipore Enhanced Chemiluminescence (ECL) Detection Kit, and subsequent quantification was conducted via ImageJ.

2.6. Cell culture and transfection

Normal human bronchial epithelial (16HBE) cells, together with the NSCLC cell lines Calu-6, A549, and H1299, were acquired from Procell Life Science & Technology Co., Ltd. located in Wuhan, China. All cells were cultured in RPMI-1640 medium (Gibco, Grand Island, NY, USA) supplemented with 10% fetal bovine serum (FBS; HyClone, Logan, UT, USA) and 1% penicillin–streptomycin (Solarbio) at 37°C in a 5% CO_2 humidified incubator. For gene manipulation: Small interfering RNAs (siRNAs) targeting MORC2 (si-MORC2), PAX2 (si-PAX2), and non-targeting control (si-NC) were obtained from GenePharma Co., Ltd. (Shanghai, China); overexpression plasmids pcDNA3.1-MORC2 (OE-MORC2) and empty vector (pcDNA3.1-NC) were constructed by GenScript Biotech Corporation (Nanjing, China). Lipofectamine 3000 reagent (Invitrogen) was utilized for transfection procedures in accordance with the manufacturer's specified guidelines. For single-gene silencing of MORC2 or PAX2, cells were transfected with si-MORC2/si-PAX2 or si-NC, while for rescue experiments, they were co-transfected with si-PAX2 plus OE-MORC2 (or pcDNA3.1-NC). All functional assays were conducted 48 h post-transfection.

2.7. Cell counting kit-8 (CCK-8)

At a seeding density of 3×10^3 cells per well, cells were inoculated into 96-well plates and cultured for 24, 48, or 72 h to allow for proliferation. CCK-8 assay reagent (10 μL, Solarbio) was added to experimental wells and blank wells (medium alone) at each experimental time point, followed by 2 h incubation at 37°C. The absorbance at 450 nm was assayed by means of a microplate reader. Cell viability was calculated relative to the si-NC group.

2.8. Annexin V-fluorescein isothiocyanate/propidium iodide (Annexin V-FITC/PI) staining

After being harvested, cells were washed twice with cold phosphate-buffered saline (PBS, Solarbio) and resuspended in 1× binding buffer, resulting in a final concentration of 1×10^5 – 1×10^6 cells/mL (to ensure sufficient cell count for statistics and avoid aggregation). Apoptosis was measured using the Annexin V-FITC/PI Apoptosis Detection Kit (BD Biosciences, Franklin Lakes, NJ, USA): Incorporation of 5 μL each of Annexin V-FITC conjugate and PI into the cell suspension was performed, followed by a 15-min incubation in darkness at room temperature to facilitate staining. Following incubation, 400 μL of 1× binding buffer was used to dilute the sample (to match the flow cytometer's sample volume requirement), and the suspension was filtered through a 300-mesh cell strainer (Solarbio) to remove cell clumps (preventing flow cytometer nozzle blockage). The BD FACSCanto II flow cytometer was used to analyze stained cells. For accurate gating, three controls were set: blank control (unstained cells, to adjust forward/side scatter (FSC/SSC) gates), Annexin V-FITC single-stained control (to adjust FITC channel compensation), PI single-stained control (to adjust PI channel compensation). Data were processed with FlowJo. The percentage of apoptotic cells (Annexin V^+/PI^- and Annexin V^+/PI^+) was computed.

2.9. Sphere formation assay

Cells were plated at a density of 5×10^2 cells/well and subsequently cultured in serum-free RPMI-1640 medium, and 2% B27 supplement (Gibco) at 37°C with 5% CO_2. After 7 d, spheres > 50 μm in diameter were counted under an inverted microscope. The efficiency of sphere formation was defined as the ratio of the number of spheres generated to the total number of seeded cells, multiplied by 100% to obtain a percentage value.

2.10. Transwell invasion assay

Transwell chambers (8 μm pore size, Corning) were coated in advance with 50 μL of Matrigel, diluted 1:8 in serum-free RPMI-1640 medium (Corning), followed by incubation at 37°C for 30 min to facilitate polymerization. Cells were resuspended using serum-free RPMI-1640 medium and inoculated into the upper chamber at 5×10^4 cells per well. 10% FBS-supplemented RPMI-1640 medium (600 μL) was dispensed into the lower chamber, serving as the chemoattractant to trigger cell migration. Cells that failed to invade the lower chamber were eliminated from the upper surface with a sterile cotton swab after 24 h of incubation at 37°C under a humidified 5% CO_2 environment. Cells invading the lower surface underwent fixation with 4% paraformaldehyde (Solarbio) for 15 min, followed by staining with 0.1% crystal violet (Solarbio) for 20 min and gentle rinsing with sterile PBS. Invaded cells were counted manually with ImageJ, subsequent to capturing five randomly chosen fields per Transwell insert through an inverted microscope.

2.11. Glycolysis-related parameter detection

Cell suspensions were dispensed into 6-well plates (2×10^5 cells/well). For glucose consumption and lactate production, culture medium was collected; glucose concentration was measured via Glucose Assay Kit (BioVision, Milpitas, CA, USA), and lactate concentration via Lactate Assay Kit (BioVision) per manufacturers' protocols. Values were normalized to cell number. For the ATP/ADP ratio, cells were lysed, and ATP/ADP levels were detected using the ATP/ADP Ratio Assay Kit (BioVision). The ratio was calculated as ATP concentration/ADP concentration.

2.12. Prediction of transcription factor PAX2 binding to MORC2 promoter

Potential transcription factors regulating the human MORC2 gene are identified via the GeneCards database (https://www.genecards.org/), which lists candidate binding sites for the MORC2 promoter, including PAX2. Among these predicted factors, only PAX2 is associated with the promotion of NSCLC progression, thus being selected as the target transcription factor for subsequent binding site prediction.

Potential binding sites of PAX2 on the MORC2 promoter are identified through TFBS prediction utilizing the JASPAR database (https://jaspar.genereg.net/analysis). JASPAR is chosen for its provision of a free, open-access collection of experimentally validated vertebrate TFBS motifs, which offers comprehensive annotation of human transcription factors and quantitative assessment of binding affinity via relative scores, thereby supporting precise selection of high-confidence binding sites for functional experiments.

The step-by-step workflow for JASPAR analysis is as follows: the 1000 bp DNA sequence upstream of the MORC2 transcription start site (TSS) is input into the JASPAR online analysis tool. The JASPAR CORE (Vertebrates) database is selected, and the human PAX2 transcription factor motif (MA0067.2) is specified as the target. Potential PAX2 binding sites are generated, and the site with the highest relative score (0.8311, corresponding to the sequence CCCCCGGAGCGGGGCGG) is chosen as the core binding region for subsequent mutagenesis assays. Based on these predictive results, it is hypothesized that PAX2 can induce transcriptional regulation of MORC2 by binding to its promoter region.

2.13. Luciferase reporter assay

Wild-type (WT) MORC2 promoter (containing the predicted PAX2 binding site) and mutant (MUT) MORC2 promoter (binding site mutated) were synthesized and cloned into the pGL3-Basic vector (Promega, Madison, WI, USA) by GenScript. 24-well cell culture plates were inoculated with A549 and H1299 cells (1×10^5 cells/well), followed by overnight incubation. A549 and H1299 cells were inoculated into 24-well plates at 1×10^5 cells per well and cultured overnight. Cells were co-transfected with 500 ng pGL3-WT/MUT-MORC2, 50 ng pRL-TK (Renilla luciferase control, Promega), and 50 nM si-PAX2/si-NC using Lipofectamine 3000. Luciferase activity was assayed using the Promega Dual-Luciferase Reporter Assay System and a GloMax 20/20 Luminometer (Promega). Relative luciferase activity was derived from the ratio of firefly luciferase activity to Renilla luciferase activity, ensuring data standardization. The MORC2 promoter was amplified by PCR from human genomic DNA. Primers were designed using Primer5 software. The wild-type (WT) MORC2 promoter fragment was cloned into the pGL3-basic vector. Site-directed mutagenesis was performed to mutate the PAX2 binding site in the MORC2 promoter, using the following primers: sense: 5′-AGAG CATGCGTCTTGAGC-3′; anti-sense: 5′-GGGTGTTGGTTGGTAG GAGG-3′.

2.14. Chromatin immunoprecipitation (ChIP) assay

The ChIP Assay Kit (Abcam) was used to perform the ChIP assay following the manufacturer's instructions. H1299 cells were subjected to cross-linking with 1% formaldehyde for 10 min at room temperature, followed by termination of the reaction using 0.125 M glycine (Solarbio). After lysis of the cells, chromatin was disrupted into 200–500 bp fragments through the use of an ultrasonic disruptor. Overnight incubation of sheared chromatin was performed at 4°C with 5 μg anti-PAX2 antibody (PA5-118142, 1:100, Invitrogen) or normal rabbit IgG (negative control,

ab171870, 1:100, Abcam). Protein A/G agarose beads (Abcam) were used to capture the antibody-chromatin complex, which was washed, eluted, and de-cross-linked. DNA was purified via phenol–chloroform extraction, and qRT-PCR was performed to detect MORC2 promoter enrichment. Enrichment was calculated relative to IgG control. The primers of the si-PAX2 were: sense: 5′-GACTCGCCTAAGGTCACTCC-3′; anti-sense: 5′-GGAGC CATCTTGTCTCGTCGC-3′.

2.15. In vivo nude mouse xenograft model establishment and staining

Female 6-week-old BALB/c nude mice (Beijing Vital River Laboratory Animal Technology Co., Ltd., Beijing, China) were reared in a controlled specific pathogen-free (SPF) environmental setting. Random allocation of mice to individual experimental cohorts was performed before conducting the study. By an investigator masked to group assignment, tumor volume assessment was executed with strict adherence to experimental blinding guidelines, maintaining measurement impartiality. H1299 cells were stably transduced with sh-NC (negative control), sh-PAX2 (PAX2 knockdown), or sh-PAX2 + OE-MORC2 (PAX2 knockdown + MORC2 overexpression) to generate three cell lines: sh-NC, sh-PAX2, and sh-PAX2 + OE-MORC2. For tumor xenograft establishment, 5×10^6 cells suspended in 100 μL of a 1:1 PBS/Matrigel mixture (Corning) were subcutaneously injected into each mouse's right flank (n = 6 per group). Using a vernier caliper, tumor volume was measured at 5-d intervals, with the calculation based on [Equation 1]:

$$\text{Volume} = \frac{\text{Length} \times \text{Width2}}{2} \quad (1)$$

The mice were humanely euthanized 28 d after injection, and the resulting tumors were excised, weighed, and imaged immediately thereafter to preserve experimental samples. Tumor tissue protein was extracted for WB analysis to confirm PAX2 and MORC2 expression.

Tissue morphology was examined via Hematoxylin-eosin (HE) staining. Briefly, formalin-fixed paraffin-embedded (FFPE) tissue slices underwent deparaffinization with xylene (Solarbio) and subsequent rehydration via a graded ethanol series, culminating in distilled water. The sections were then stained with hematoxylin (Solarbio) for 8 min, followed by differentiation in 1% acid alcohol and bluing in 0.5% ammonia water. Thereafter, eosin (Solarbio) was used for counterstaining the tissue sections for 3 min. After the staining process, the sections were processed for dehydration through a graded ethanol series, cleared with xylene, and affixed in neutral balsam (Solarbio). For immunohistochemistry (IHC), sections were deparaffinized, rehydrated, and antigen-retrieved in citrate buffer (pH 6.0). After blocking endogenous peroxidase (3% H_2O_2), sections were incubated overnight at 4℃ in a solution containing the respective primary antibodies: anti-MORC2 (ab315903, 1:200, Abcam), anti-Ki67 (ab15580, 1:300, Abcam), and anti-PAX2 (PA5-118142, 1:200, Abcam). Sections were then incubated with HRP-conjugated secondary antibody (goat anti-rabbit IgG (H+L), ab205718, 1:500, Abcam) for 1 h, developed with DAB (Solarbio), and counterstained with hematoxylin. Images of the stained sections were obtained through the use of a light microscope, and positive cells were quantified using ImageJ. Taiyuan Central Hospital, The Ninth Clinical Medical College of Shanxi Medical University's Animal Ethics Committee reviewed and approved all animal study protocols, which were executed in compliance with the Guide for the Care and Use of Laboratory Animals.

2.16. Statistical analysis

Experimental data, collected from no fewer than three independent biological replicates, were reported as mean ± standard deviation (SD) and statistically analyzed via GraphPad Prism and R software. Two-group comparisons used unpaired Student's *t*-test; For comparisons among more than two groups, one-way ANOVA was performed followed by Tukey's post-hoc test for pairwise comparisons between all groups, or by Dunnett's post-hoc test when comparing multiple treatment groups to a single control group; survival analysis (Kaplan-Meier plots) used log-rank test; PAX2-MORC2 expression correlation used Pearson correlation coefficient; DEG screening (GSE29249) used adjusted $p < 0.05$ and $|\log_2 FC|>1$. Two-tailed $p < 0.05$ was significant.

3. Results

3.1. Bioinformatics screening identifies transcriptional alterations in NSCLC

To uncover genes with potential roles in NSCLC pathogenesis, an initial transcriptomic analysis was conducted using the publicly available GSE29249 dataset. Differential expression profiling between NSCLC and matched adjacent normal tissues revealed widespread gene dysregulation. A substantial proportion of genes met the significance thresholds ($p < 0.05$ and $|\log_2 FC| > 1$), leading to the identification of 2464 DEGs: 1018 upregulated and 1446 downregulated transcripts, which were visualized via the volcano plot (Fig. 1A). The plot displayed a marked asymmetry, with a prominent cluster of significantly upregulated genes distributed on the right side, indicating a strong signature of transcriptional activation in tumor tissues.

For further prioritization of biologically relevant candidates, a hierarchical clustering heatmap (Fig. 1B) was constructed, displaying the top 10 upregulated and top 10 downregulated protein-coding DEGs. The resulting heatmap demonstrated a clear segregation of samples into two major clusters corresponding perfectly to their histological classification: NSCLC and normal. The color gradient, representing expression levels from low (light orange) to high (dark brown), further illustrated distinct gene expression profiles between the two conditions, with one large gene cluster showing uniformly higher expression in tumor samples. This integrated bioinformatic approach successfully delineated the global transcriptional landscape of NSCLC, establishing a reliable foundation for subsequent focused investigation.

3.2. Machine learning integration identifies MORC2 as a key candidate in NSCLC

Following the identification of a broad set of differentially expressed genes, a refined analysis was necessary to pinpoint the most promising candidates driving NSCLC pathogenesis. To this end, two distinct machine learning algorithms—LASSO logistic regression and RF—were employed for robust feature selection. For the LASSO analysis (Fig. 2A), the coefficient path plot illustrated how gene coefficients shrink with increasing log (λ), while the partial likelihood deviance plot identified the optimal λ values: λ.min (−5.3921) and λ.1se (−4.8339). At λ.min, five genes (MORC2, FAM115B, PTGER1, MS4A14, and CCDC88A) retained non-zero coefficients, indicating their potential discriminatory power.

To complement this approach and assess feature importance based on an alternative algorithm, an RF analysis was conducted (Fig. 2B). The model's error rate stabilized at a low level as the number of trees increased, confirming its predictive stability. The analysis of variable importance, measured by the mean decrease in Gini index, identified MORC2 and PTGER1 as the two genes contributing most significantly to the model's accuracy. Refinement of the candidate set via Venn diagram analysis of LASSO and RF outputs confirmed MORC2 and PTGER1 as overlapping key candidates, consistent with the variable importance results from RF (Fig. 2C).

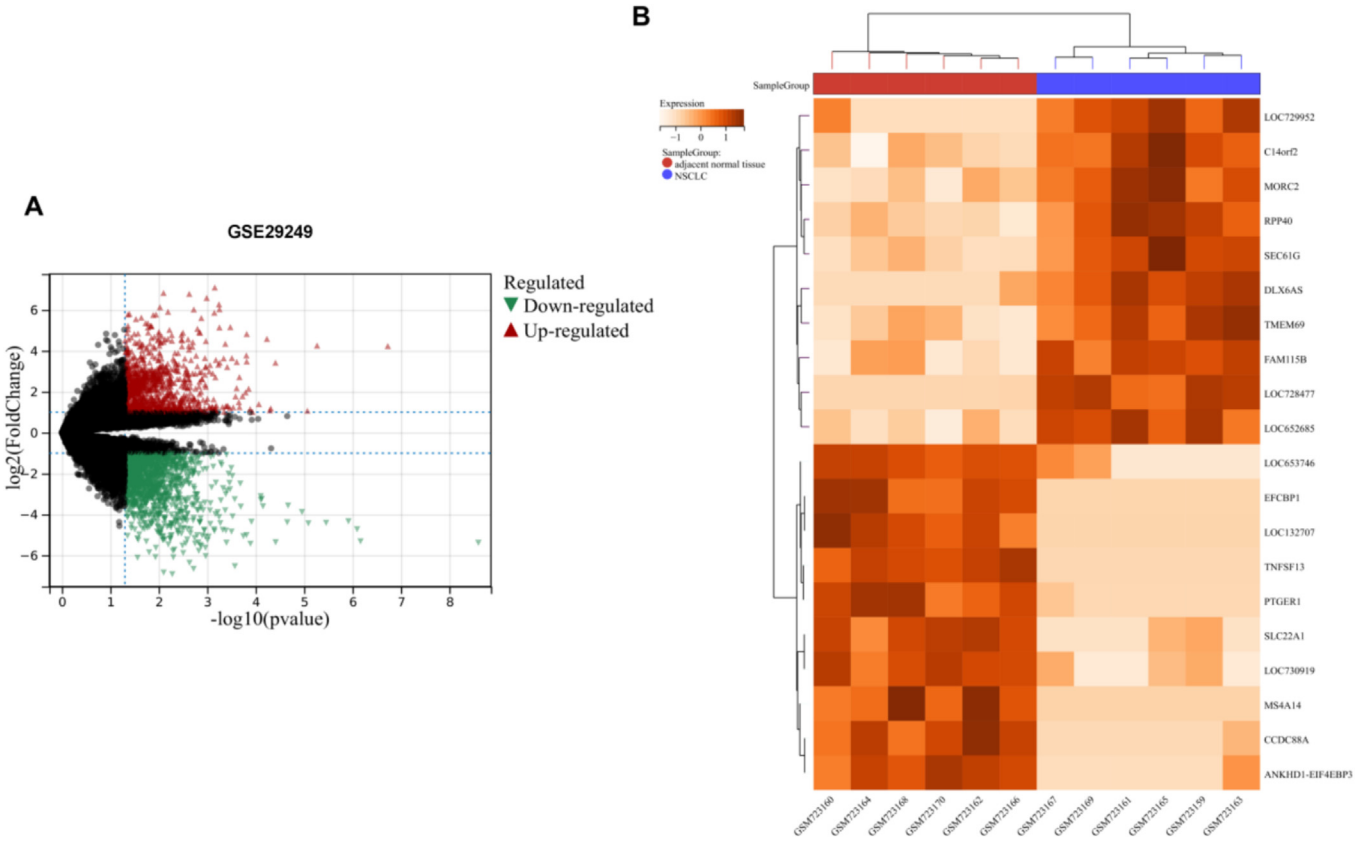


Fig. 1. Transcriptomic analysis of the GSE29249 **dataset reveals differential gene expression in NSCLC. (A)** Volcano plot illustrating genome-wide expression differences between NSCLC and adjacent normal tissues. Each point represents a single gene; red upward triangles denote significantly upregulated genes, green downward triangles denote significantly downregulated genes ($p < 0.05$, $|\log_2 FC| > 1$), and black dots represent non-significant genes. Dotted lines indicate the applied thresholds for statistical significance ($-\log10(p\text{-value})$) and $\log_2(FC)$. **(B)** Heatmap with hierarchical clustering of the differentially expressed genes across all samples in the dataset. Rows correspond to individual genes, and columns represent individual tissue samples. Expression levels are color-coded from low (light orange) to high (dark brown), as shown in the sidebar. The dendrogram and color bar at the top indicate sample clustering and group assignment (NSCLC in blue, adjacent normal tissue in red), respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

By integrating LASSO and RF algorithms, MORC2 and PTGER1 were efficiently prioritized as key candidate genes. Guided by prior reports on MORC2′s role in NSCLC progression [16,22], the research focused on dissecting MORC2′s functional and mechanistic contributions to the disease.

3.3. Elevated expression of MORC2 in NSCLC correlates with poor prognosis

After identifying MORC2 as a candidate via machine learning and previous studies, its expression patterns and clinical relevance were systematically validated across datasets, clinical specimens, and cell models. To evaluate MORC2 expression in large-scale cohorts, the TCGA-LUAD dataset was analyzed first. In contrast to 59 adjacent normal tissue counterparts, the expression of MORC2 transcripts was markedly increased in primary tumors (n = 515) (***$p < 0.001$, Fig. 3A). This observation was corroborated in an independent cohort, where scatter plots showed significantly higher MORC2 gene expression in tumor samples relative to normal counterparts (p = 7.48e-31, Fig. 3B). Kaplan-Meier survival analysis was conducted to explore the clinical implications of MORC2 expression. The results showed that elevated MORC2 expression correlated with significantly shorter overall survival in NSCLC patients (hazard ratio = 1.59, log-rank $P = 0.0053$, Fig. 3C).

Next, upon re-analyzing the GSE29249 dataset, it was observed that MORC2 mRNA levels were robustly elevated in tumor tissues compared to normal tissue controls (***$p < 0.001$, Fig. 3D). These findings were further validated using qRT-PCR on 43 paired clinical specimens, where tumor tissues exhibited marked MORC2 mRNA upregulation relative to adjacent non-tumor tissues (***$p < 0.001$, Fig. 3E). Integrating molecular and clinical data presented a survival curve stratified by MORC2 expression (low vs. high, $p = 0.0384$), which exhibited a significant elevation in tumor tissues relative to normal tissues (***$p < 0.001$, Fig. 3F). WB analysis of MORC2 across a panel of NSCLC tumor specimens (N1-T7) consistently indicated upregulation in malignant tissues (Fig. 3G).

Finally, MORC2 expression was investigated in cell models. The levels of mRNA (Fig. 3H) and protein (Fig. 3I) were quantified in normal bronchial epithelial cells (16HBE) in contrast to NSCLC cell lines (Calu-6, A549, and H1299). Both analyses revealed significantly higher MORC2 expression in tumor cell lines compared to 16HBE, reinforcing its oncogenic association. Across public databases, clinical specimens, and cell models, these cross-modal validations consistently demonstrated that MORC2 was robustly upregulated in NSCLC and correlated with adverse prognosis, providing a mechanistic foundation for subsequent functional investigations.

3.4. Knockdown of MORC2 impairs malignant phenotypes and alters glycolytic metabolism in NSCLC cells

Having established the clinical and prognostic relevance of elevated MORC2 expression, this study next aimed to elucidate its functional role in NSCLC pathogenesis. To investigate the conse-

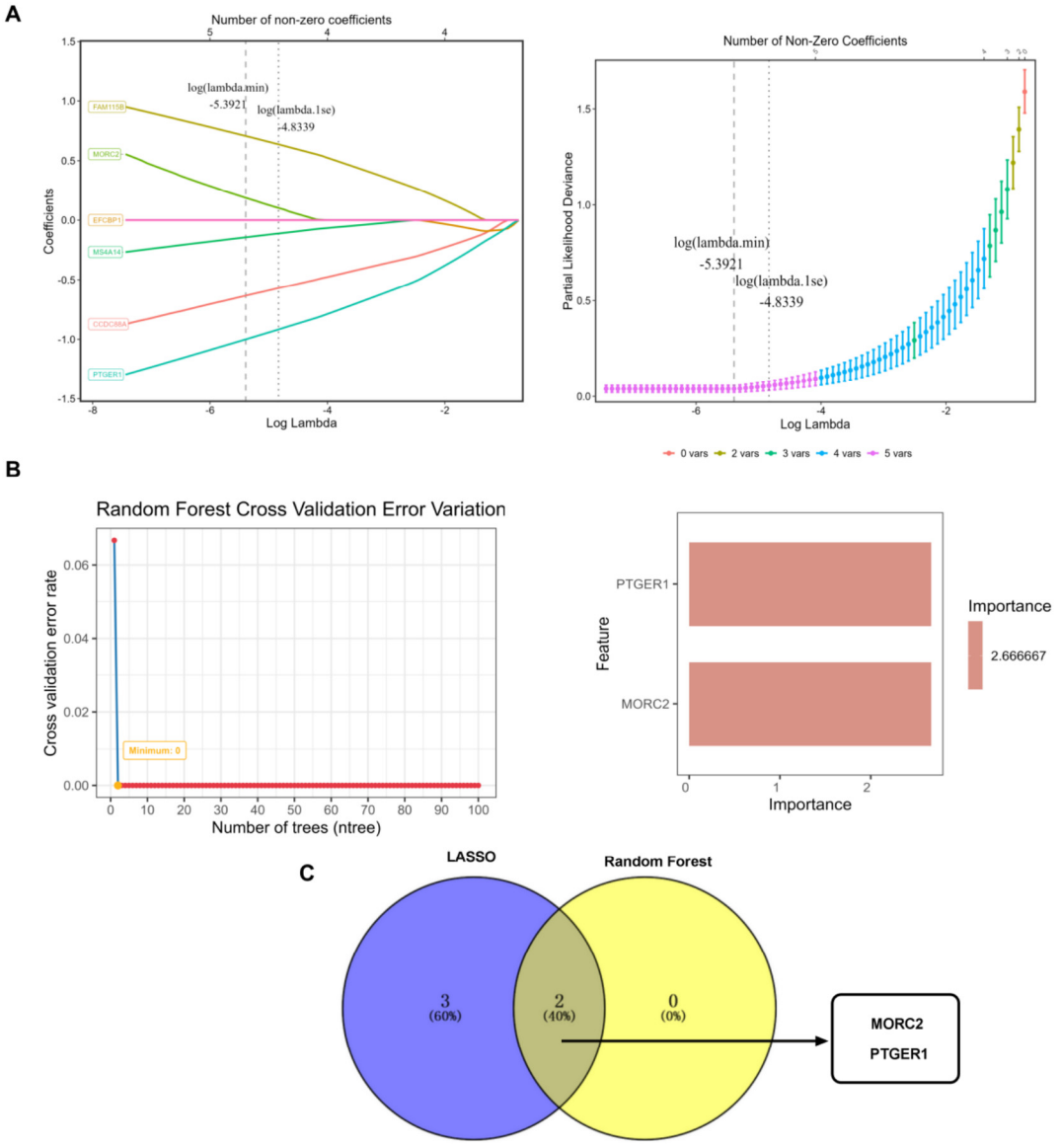


Fig. 2. Machine learning-based screening of key genes in NSCLC. (A) Feature selection using LASSO logistic regression, displaying the coefficient paths of candidate genes across penalty parameters (log (λ)) and the associated 3-fold cross-validation error curve. The dotted vertical line marks the optimal lambda (λ.min) yielding the minimum cross-validation error. **(B)** Feature selection using the RF algorithm, showing the relationship between model error and the number of trees, as well as a bar plot of the mean decrease Gini index (a measure of variable importance) for the top two genes. **(C)** Venn diagram depicting the overlap between key candidate genes identified by the LASSO regression and RF analyses.

quences of MORC2 loss-of-function, its expression was stably silenced in A549 and H1299 NSCLC cell lines using specific si-MORC2, with a si-NC serving as controls. Protein-level validation of knockdown efficiency was performed initially; WB-based analysis confirmed a notable downregulation of MORC2 expression in both si-MORC2-transfected cell lines as opposed to the respective si-NC-transfected control groups (***$p < 0.001$, Fig. 4A).

Following the confirmation of efficient MORC2 knockdown, a series of functional assays was pursued. For the purpose of evaluating the effect on cellular proliferation, CCK-8 assays were conducted over 72 h. Silencing MORC2 led to a significant decrease in the viability of both A549 and H1299 cells compared to the si-NC groups (***$p < 0.001$, Fig. 4B). Given the observed suppression of proliferation, the effect of MORC2 knockdown on cell death was thereafter investigated. Flow cytometric analysis of Annexin V/PI staining revealed that the depletion of MORC2 significantly increased the percentage of apoptotic cells in both cell lines (***$p < 0.001$, Fig. 4C), indicating that MORC2 contributes to cell survival. Beyond core survival mechanisms, the role of MORC2 in cancer stemness and metastatic potential was further explored, which are critical for tumor progression and recurrence. Using a sphere-formation assay, it was found that cells with reduced MORC2 expression formed significantly fewer and smaller tumorspheres compared to si-NC groups, suggesting a compromised self-renewal capacity (**$p < 0.01$, ***$p < 0.001$, Fig. 4D). Moreover, invasion, a hallmark of metastatic potential, was analyzed using Transwell assays. A substantial decrease in the number of invasive cells in si-MORC2 groups relative to si-NC (**$p < 0.01$, ***$p < 0.001$, Fig. 4E), highlighting diminished invasiveness.

Since metabolic reprogramming, particularly enhanced glycolysis, was a hallmark of aggressive cancers, whether MORC2 influences this pathway was then investigated. Analysis of glycolytic parameters showed that silencing MORC2 significantly reduced glucose consumption (Fig. 4F) and lactate production (Fig. 4G) in

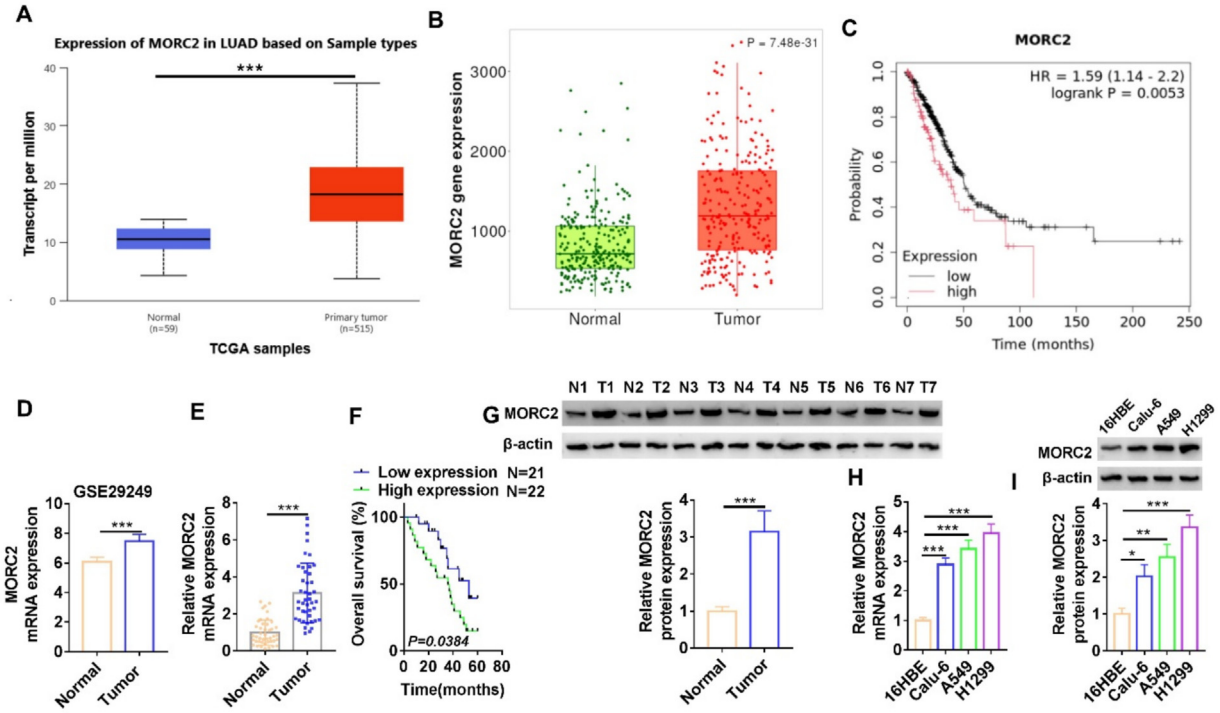


Fig. 3. Multi-modal validation of MORC2 expression and prognostic impact in NSCLC. (A) MORC2 transcript level in TCGA LUAD samples (normal, n = 59; primary tumor, n = 515) was analyzed by RNA-seq. (B) Scatter plot of MORC2 gene expression in an independent cohort (normal vs. tumor). (C) Kaplan-Meier survival curve stratified by MORC2 expression. (D) MORC2 mRNA in GSE29249 (normal vs. tumor). E qRT-PCR of MORC2 in 43 paired clinical specimens (normal vs. tumor). F Survival curve of patients with low/high MORC2. (G) WB of MORC2 in NSCLC tumor specimens (N1-T7, β-actin as a loading control). (H-I) MORC2 mRNA and protein in normal bronchial epithelial cells (16HBE) and NSCLC cell lines (Calu-6, A549, and H1299). Data are presented as the mean ± SD from three independent experiments. Statistical analyses: Student's *t*-test; one-way ANOVA with Dunnett's post-hoc test; log-rank test (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

the culture medium. Concomitantly, the cellular ATP/ADP ratio, an indicator of energy status, was also decreased upon MORC2 knockdown (Fig. 4H), linking its function to the maintenance of glycolytic flux and energy homeostasis in NSCLC cells. In summary, functional interrogation through genetic silencing demonstrated that MORC2 was essential for maintaining the proliferative, anti-apoptotic, stem-like, invasive, and glycolytic phenotypes of NSCLC cells, positioning it as a multifaceted promoter of tumor malignancy.

3.5. Transcriptional regulation of MORC2 by PAX2 in NSCLC

To dissect the upstream mechanisms driving MORC2 overexpression in NSCLC, the transcriptional regulation of MORC2 by PAX2 was probed. A bioinformatic pipeline, informed by Genecards and JASPAR analyses, first flagged PAX2 as a putative transcription factor for MORC2 (Fig. 5A). The PAX2 binding site within the MORC2 promoter was predicted and subsequently corroborated by consensus sequence alignment and motif visualization (Fig. 5B). Leveraging clinical datasets, PAX2 expression was assessed next. In TCGA NSCLC samples, PAX2 mRNA levels were notably higher relative to their adjacent normal tissue counterparts (***$p < 0.001$, Fig. 5C), aligning with its predicted regulatory function. A robust positive correlation between PAX2 and MORC2 mRNA expression was observed in NSCLC specimens ($r = 0.8407$, ***$p < 0.001$, Fig. 5D), implying a potential regulatory interplay.

siRNA-mediated silencing of PAX2 in A549 and H1299 cells was performed to validate this interaction. WB confirmed efficient PAX2 protein depletion in si-PAX2-transfected cells (***$p < 0.001$, Fig. 5E). Luciferase reporter assays were conducted using constructs containing either the WT or MUT MORC2 promoter. In both A549 (**$p < 0.01$, Fig. 5F) and H1299 (***$p < 0.001$, Fig. 5G) cells, si-PAX2 transfection attenuated the luciferase activity of the WT reporter while it had no such effect on the MUT variant, confirming specific PAX2-mediated transactivation. ChIP assays further revealed PAX2 binding to the MORC2 promoter in H1299 cells, with enrichment significantly higher than IgG controls (***$p < 0.001$, Fig. 5H). Subsequent qRT-PCR and WB showed that PAX2 knockdown substantially reduced MORC2 mRNA (***$p < 0.001$, Fig. 5I) and protein levels (**$p < 0.01$, ***$p < 0.001$, Fig. 5J) in both cell lines, cementing PAX2 as a transcriptional activator of MORC2. The integration of bioinformatic prediction and experimental validation enabled the identification of PAX2 as a direct transcriptional activator of MORC2 in NSCLC, revealing a critical regulatory axis underlying its overexpression.

3.6. Rescue experiments elucidate MORC2-mediated oncogenic functions of PAX2 in NSCLC

Building on the identification of PAX2 as a transcriptional activator of MORC2, the study next performed rescue experiments to dissect whether PAX2 exerts its oncogenic functions through MORC2. WB analysis confirmed that PAX2 knockdown significantly reduced MORC2 protein levels, and this reduction was effectively rescued by the concurrent overexpression of MORC2 (Fig. 6A), establishing the experimental model for functional interrogation.

Taking into account PAX2′s documented role in enhancing cellular proliferation, its impact on cell viability was first assessed. CCK-8 assays demonstrated that silencing PAX2 markedly inhibited the growth of both A549 and H1299 cells over 72 h. Notably, the impaired cell viability induced by PAX2 deficiency was substantially, though not completely, restored by re-expressing MORC2 (Fig. 6B). This finding prompted an examination of cell survival. Annexin V-FITC/PI staining and flow cytometry showed that

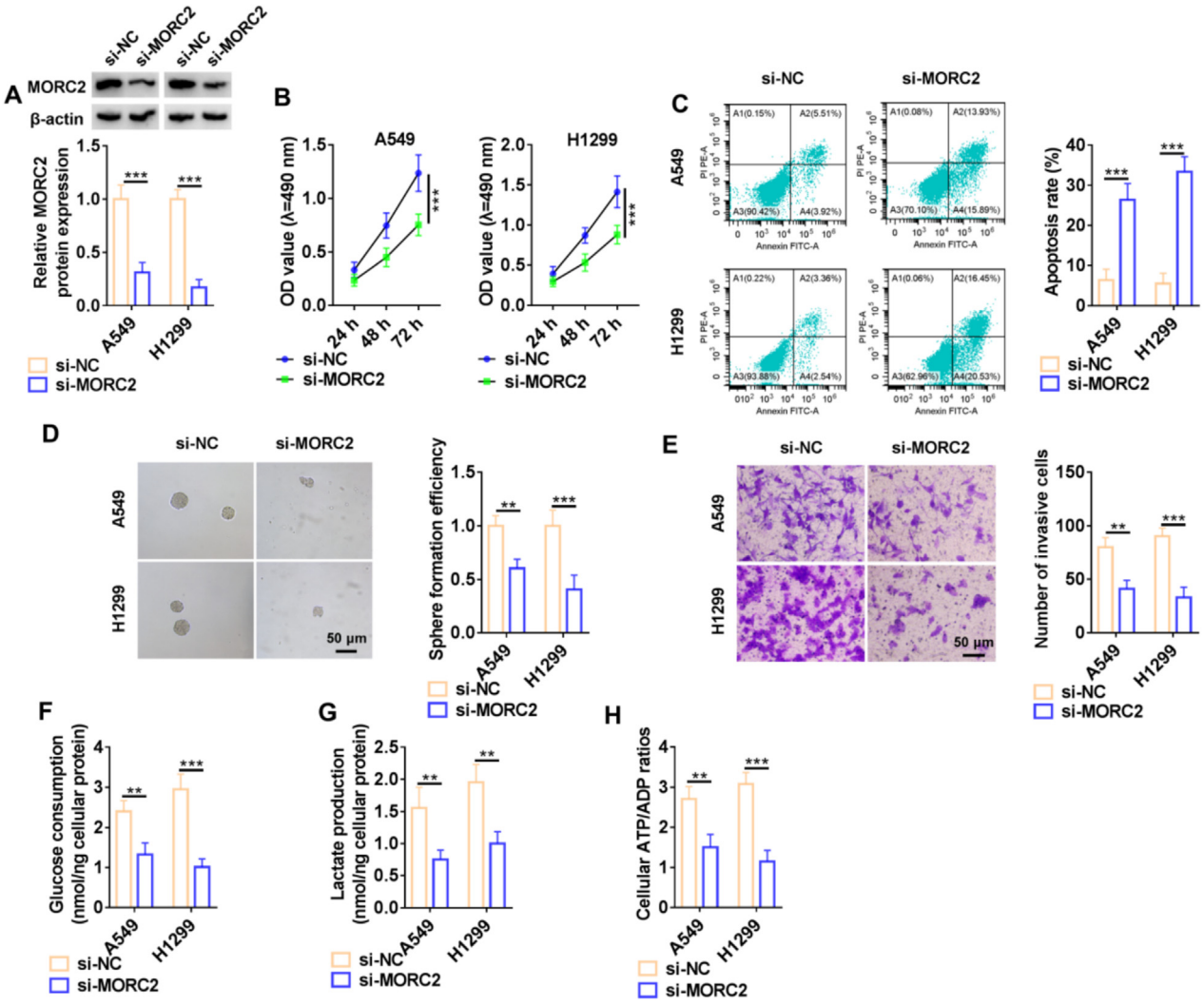


Fig. 4. Functional effects of MORC2 knockdown in NSCLC cells. A549 and H1299 cells were transfected with control siRNA (si-NC) or MORC2-targeting siRNA (si-MORC2). (A) WB analysis of MORC2 protein levels in A549 and H1299 cells transfected with si-NC or si-MORC2 (β-actin as a loading control). (B) CCK-8 assay measuring proliferation of si-NC-transfected or si-MORC2-transfected A549 and H1299 cells at 24, 48, and 72 h. (C) Flow cytometry analysis of apoptosis via Annexin V-FITC/PI staining in transfected cells. X-axis (Annexin V-FITC): phosphatidylserine externalization (early apoptosis); Y-axis (PI): membrane integrity (PI labels damaged membranes, late apoptosis/necrosis). Quadrants: Q3 (Annexin V^-/PI^-, viable); Q4 (Annexin V^+/PI^-, early apoptosis); Q2 (Annexin V^+/PI^+, late apoptosis); Q1 (Annexin V^-/PI^+, necrosis). Bar graphs present total apoptosis (Q2 + Q4). (D) Sphere formation assay evaluating cancer stem cell properties in transfected cells (scale bar = 50 μm). (E) Transwell invasion assay assessing the invasive capacity of transfected cells (scale bar = 50 μm). (F-H) Glycolytic activity analysis: glucose consumption, lactate production, and ATP/ADP ratios in transfected cells. Data are presented as mean ± SD (n = 3 independent experiments). Statistical analyses: Student's *t*-test (**$p < 0.01$, ***$p < 0.001$).

the increased apoptosis induced by si-PAX2 was attenuated in the si-PAX2 + OE-MORC2 group (Fig. 6C), suggesting MORC2 mediates PAX2′s anti-apoptotic effects.

Whether the PAX2/MORC2 axis influences more aggressive cancer cell properties was examined through an evaluation of its effects on stemness and invasive capacity. Tumorsphere formation assays showed that PAX2 silencing severely compromised the self-renewal capability of NSCLC cells, an effect that was again partially reversed by MORC2 overexpression (Fig. 6D). Similarly, Transwell invasion assays indicated that the invasive potential suppressed by PAX2 knockdown was significantly rescued by the reconstitution of MORC2 expression (Fig. 6E).

Therefore, to further characterize the underlying metabolic alterations, key parameters of glycolysis were measured. Silencing PAX2 induced a substantial decline in glucose consumption (Fig. 6F) and lactate production (Fig. 6G), alongside a reduced cellular ATP/ADP ratio (Fig. 6H). Crucially, the forced expression of MORC2 in PAX2-deficient cells effectively restored glycolytic flux and energy homeostasis. Taken together, these comprehensive rescue experiments demonstrated that MORC2 was a critical downstream executor of PAX2, mediating its oncogenic functions in promoting proliferation, survival, stemness, invasion, and glycolytic metabolism in NSCLC cells.

3.7. PAX2 depletion inhibits tumor growth in vivo, partially rescued by MORC2 overexpression

To translate *in vitro* observations of PAX2′s oncogenic potential into an *in vivo* context, a xenograft tumor model was established. WB was first performed to confirm PAX2 silencing in xenograft tumors (Fig. 7A). Relative to the sh-NC group, PAX2 protein expression was dramatically reduced in the sh-PAX2 group, verifying successful knockdown at the protein level.

Following tumor inoculation, tumor growth patterns were longitudinally followed over a 28-day interval (Fig. 7B). Tumor growth in the sh-PAX2 group was significantly inhibited compared to the sh-NC group. Notably, forced MORC2 expression in the sh-PAX2 + OE-MORC2 group partially reversed this inhibition: the res-

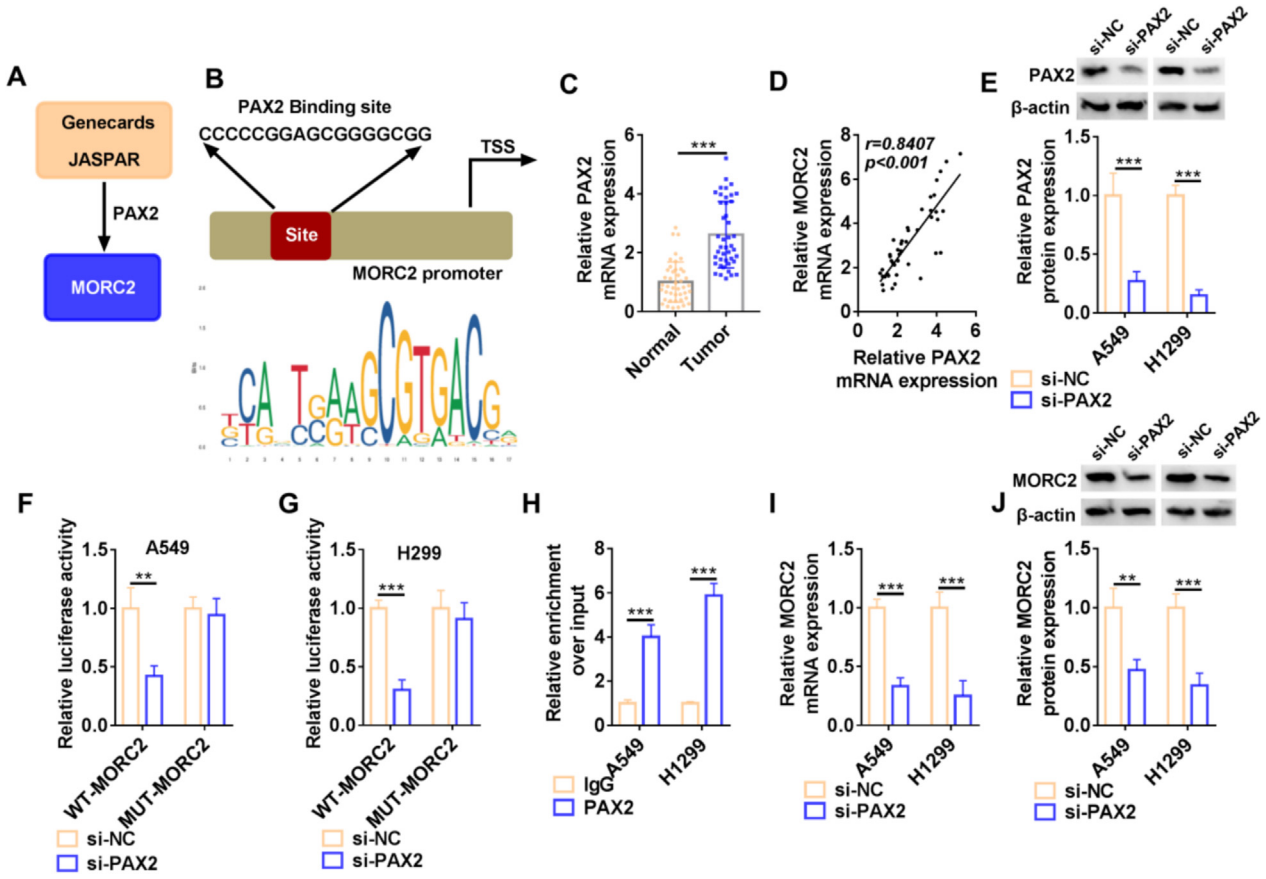


Fig. 5. The transcription factor PAX2 transcriptionally activates MORC2 expression. (A) Bioinformatic workflow using Genecards and JASPAR to predict PAX2 as a transcription factor for MORC2. (B) Predicted PAX2 binding site in the MORC2 promoter, including consensus sequence and motif visualization. (C) PAX2 mRNA expression in TCGA NSCLC samples (normal vs. tumor). (D) Pearson correlation analysis between PAX2 and MORC2 mRNA levels in NSCLC samples. (E) WB of PAX2 protein in A549 and H1299 cells transfected with si-NC or si-PAX2 (β-actin as a loading control). (F-G) Luciferase reporter assays in A549 and H1299 cells transfected with si-NC/si-PAX2 and WT/MUT MORC2 promoter constructs. (H) ChIP assay detecting PAX2 enrichment at the MORC2 promoter in H1299 cells (IgG as a control). (I-J) MORC2 mRNA and protein levels in si-NC/si-PAX2-transfected A549 and H1299 cells. Statistical analyses: Student's *t*-test, Pearson correlation (**$p < 0.01$, ***$p < 0.001$).

cue group displayed a faster tumor growth rate than the sh-PAX2 group, which indicated a partial rescue effect. At the study endpoint, tumor weights were measured to corroborate growth differences (Fig. 7C). The sh-PAX2 group displayed markedly lower tumor weights than the sh-NC group. Importantly, MORC2 overexpression in the sh-PAX2 background significantly increased tumor weight relative to the sh-PAX2 group alone, further supporting MORC2′s role in reversing PAX2 depletion-induced growth inhibition.

Successful MORC2 re-expression in the rescue group was confirmed by WB analysis (Fig. 7D). Versus the non-targeting sh-NC group, MORC2 protein expression was decreased in the sh-PAX2 group, while forced MORC2 overexpression ameliorated this issue in the sh-PAX2 + OE-MORC2 group, verifying the effectiveness of the rescue strategy. Notably, MORC2 protein expression in xenograft tumor tissues was specifically detected (Fig. 7D). To explicitly demonstrate the *in vivo* knockdown efficiency of PAX2-shRNA, PAX2 protein expression in the same set of tumor tissues was analyzed. Relative to the sh-NC group, PAX2 protein expression was significantly reduced in the sh-PAX2 group, confirming successful PAX2 silencing at the protein level; no significant difference in PAX2 expression was observed between the sh-PAX2 group and the sh-PAX2 + OE-MORC2 group, which indicated that MORC2 overexpression did not affect PAX2 protein levels (**Fig. S1**).

Histological and immunohistochemical staining were conducted to assess tumor morphology, proliferation, and MORC2 expression (Fig. 7E). HE staining revealed a more disorganized and less dense tumor architecture in the sh-PAX2 group, which was partially normalized by MORC2 overexpression. Ki67 immunostaining (a proliferation marker) showed fewer positive cells in the sh-PAX2 group, with numbers increasing in the sh-PAX2 + OE-MORC2 group. Consistent with WB, MORC2 IHC demonstrated reduced staining in the sh-PAX2 group and recovery in the rescue group. These *in vivo* data demonstrated that PAX2 depletion suppressed tumor growth, and this effect was partially mediated by MORC2, as reintroducing MORC2 alleviated the growth inhibition, suggesting a functional relationship between PAX2 and MORC2 in tumor progression.

4. Discussion

NSCLC remains a leading cause of cancer-related mortality globally, largely due to late-stage diagnosis, high metastatic potential, and the limited efficacy of existing targeted therapies [7,8]. Identifying key molecular targets and their regulatory mechanisms is critical for developing precise diagnostics and therapies. This study integrated bioinformatics, machine learning, and *in vitro*/*in vivo* experiments to explore MORC2′s role in NSCLC, revealing that MORC2 was upregulated in NSCLC, correlated with poor prognosis, promoted malignant phenotypes and glycolysis, and was transcriptionally activated by PAX2—forming a PAX2/MORC2 axis driving NSCLC progression. These experimental observations deliver perspectives concerning the pathogenesis of NSCLC and prospective therapeutic targets for clinical application.

Transcriptomic analysis of GSE29249 identified 2464 DEGs (1018 upregulated, 1446 downregulated) between NSCLC and nor-

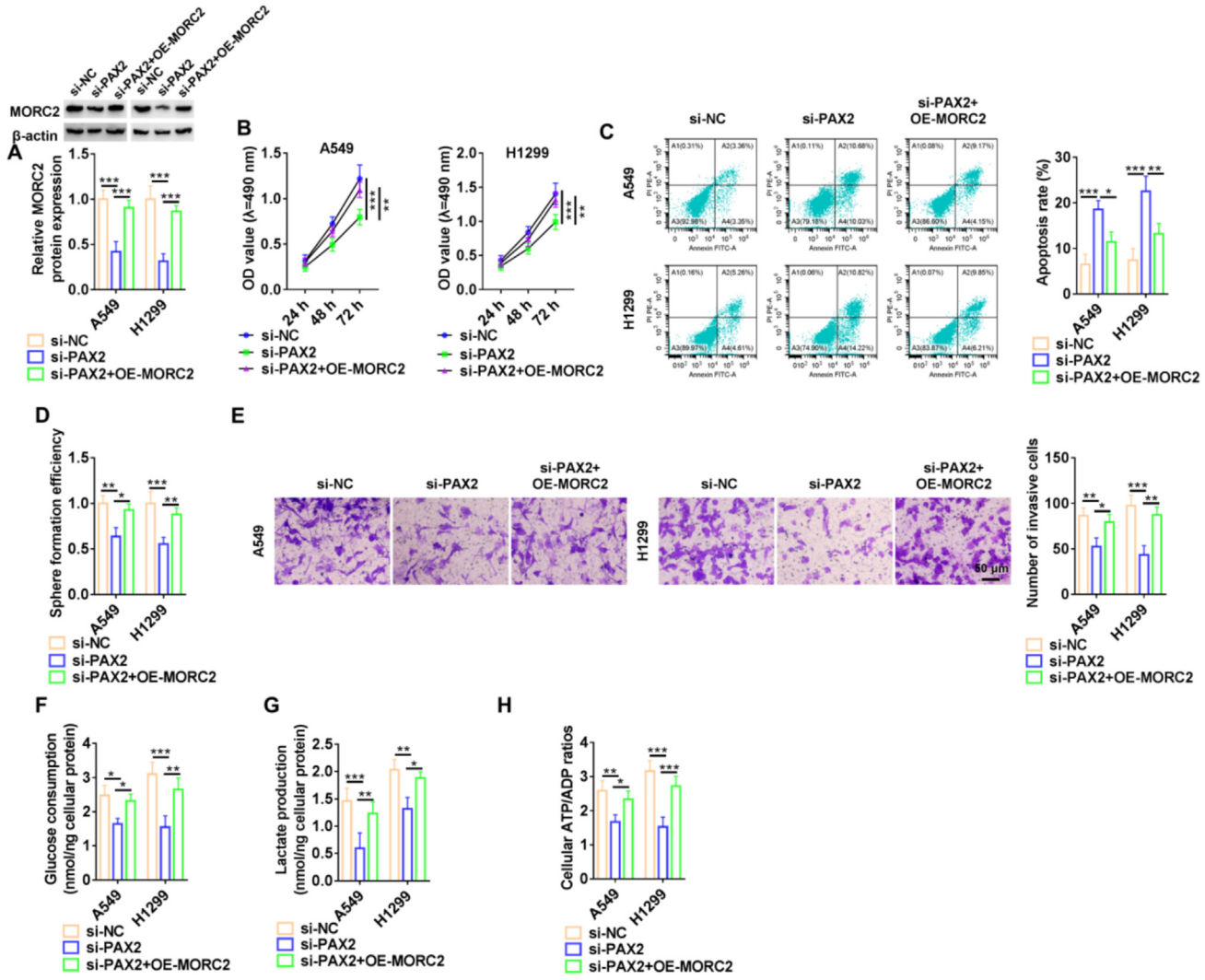


Fig. 6. MORC2 restoration rescues the oncogenic phenotypes impaired by PAX2 knockdown in NSCLC cells. A549 and H1299 cells were subjected to three treatments: control siRNA (si-NC), PAX2-targeting siRNA (si-PAX2), or si-PAX2 combined with MORC2 overexpression (si-PAX2 + OE-MORC2). (A) WB analysis of MORC2 protein expression in A549 and H1299 cells. β-Actin served as a loading control. (B) Cell viability assessed by the CCK-8 assay in A549 and H1299 cells under the indicated treatment conditions at 24, 48, and 72 h post-transfection. (C) Annexin V-FITC/PI double staining analyzed apoptosis in A549 and H1299 cells by flow cytometry. (D) Quantitative analysis of tumorsphere formation assays performed with A549 and H1299 cells under the specified treatment conditions. (E) Representative images and quantification of Matrigel invasion assays for A549 and H1299 cells under the indicated treatments. The scale bar represents 50 μm. The bar graph shows the number of invaded cells per field. (F-H) Measurement of glycolytic parameters in treated A549 and H1299 cells: Glucose consumption in the culture medium over 24 h, Lactate production in the culture medium over 24 h, and (H) Intracellular ATP/ADP ratio. Data are shown as mean ± SD (n = 3 independent experiments). Statistical significance was analyzed by one-way ANOVA with Tukey's post-hoc test (*p < 0.05, **p < 0.01, ***p < 0.001).

mal tissues. This pattern of widespread transcriptional dysregulation is consistent with previous observations in NSCLC, where malignant transformation is accompanied by extensive changes in gene expression [23]. Hierarchical clustering confirmed distinct tumor-normal expression profiles, validating DEG reliability. To filter candidates, two complementary machine learning tools, including LASSO logistic regression and RF, were used. This strategy is widely recognized in cancer biomarker research for its ability to reduce data dimensionality and improve the robustness of candidate selection [24]. LASSO identified five genes (MORC2, FAM115B, PTGER1, MS4A14, and CCDC88A) at optimal λ.min, while RF (via mean decrease in Gini index) highlighted MORC2 and PTGER1. Their intersection confirmed MORC2/PTGER1 as core candidates, consistent with an NSCLC exosome-related gene study that also applied LASSO and RF for key gene screening [25]. MORC2 was prioritized for further study due to prior evidence: It promotes LC angiogenesis via Wnt/β-catenin [16] and regulates ovarian cancer glycolysis [22], supporting its oncogenic role in NSCLC.

Multi-dimensional validation confirmed MORC2 upregulation in NSCLC. TCGA-LUAD analysis showed higher MORC2 in 515 tumors vs. 59 normal tissues, corroborated by GSE29249 reanalysis [12]. qRT-PCR/WB of 43 paired specimens confirmed MORC2 overexpression at mRNA/protein levels, extending previous reports of MORC2 upregulation in NSCLC [16]. Kaplan-Meier analysis linked high MORC2 to shorter overall survival (HR = 1.59, log-rank P = 0.0053), consistent with Liu et al. [26], who associated MORC2 with lymph node metastasis and poor prognosis. Liao et al. [27] also validated MORC2 as a prognostic marker, noting its potential to complement traditional biomarkers. Additionally, MORC2 was overexpressed in NSCLC cell lines (Calu-

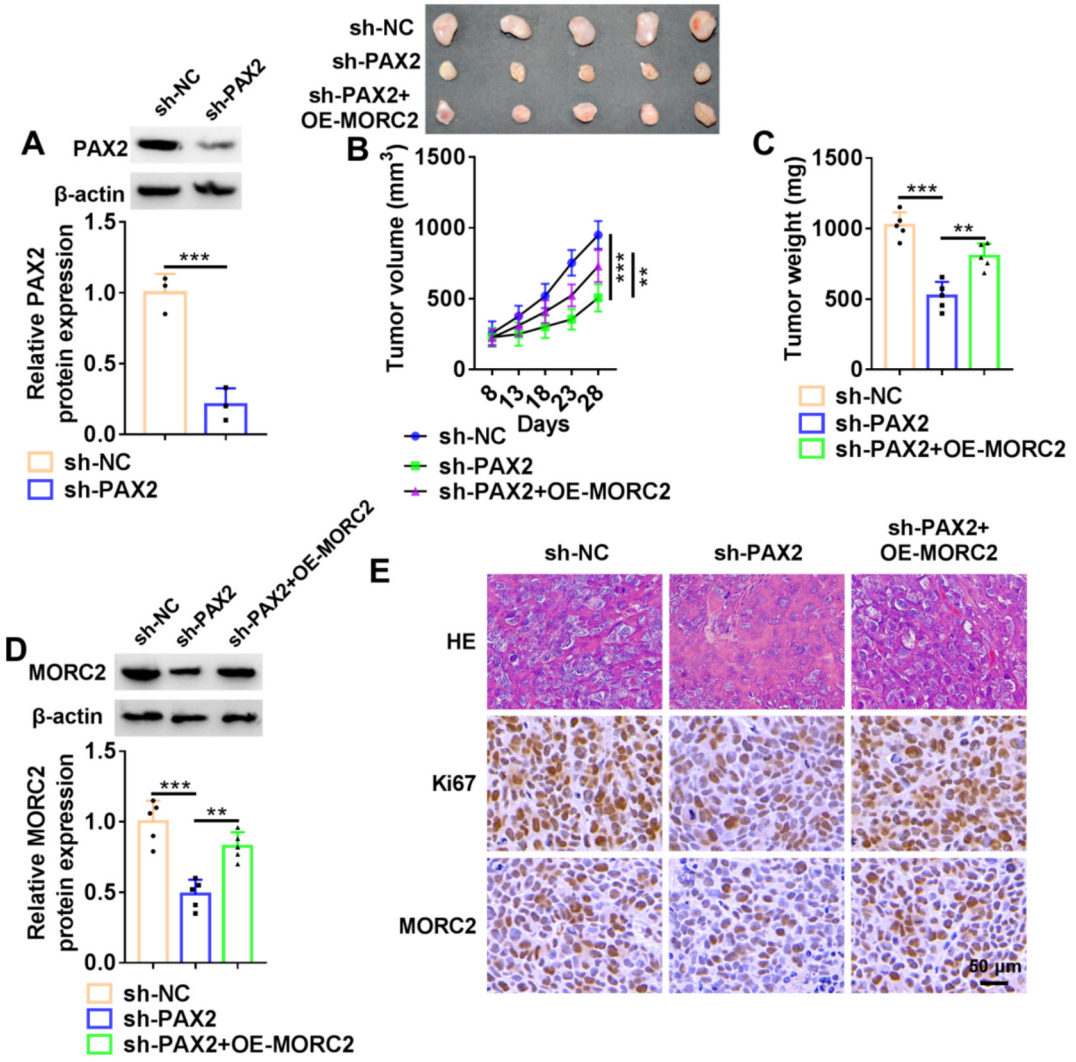


Fig. 7. PAX2 depletion suppresses tumor growth *in vivo*, partially reversed by MORC2 overexpression. Tumor xenografts were established by inoculating nude mice with cancer cells stably transduced with sh-NC (negative control), sh-PAX2 (PAX2 knockdown), or sh-PAX2 + OE-MORC2 (PAX2 knockdown + MORC2 overexpression). (A) WB analyzed PAX2 protein levels in tumor tissues. (B) Tumor volumes were measured every 5 days for 28 days. (C) Final tumor weights were recorded at study termination. (D) WB assessed MORC2 protein expression in tumor tissues. (E) Histological analysis via HE staining and IHC staining for Ki67 (proliferation marker) and MORC2, with scale bar = 50 μm. Data are presented as mean ± SD from at least three independent experiments. Statistical significance was determined by Student's *t*-test or one-way ANOVA with Tukey's post-hoc test (**$p < 0.01$, ***$p < 0.001$).

6, A549, and H1299) vs. normal 16HBE cells, supporting its oncogenic function and providing a cellular model for mechanistic studies.

siRNA-mediated MORC2 knockdown inhibited NSCLC cell proliferation (CCK-8) and induced apoptosis (Annexin V/PI), consistent with MORC2′s pro-survival role in other cancers [28,29]. MORC2 depletion also reduced tumorsphere formation (stemness) and invasion (Transwell), aligning with Liu *et al.*, who linked stemness regulators to NSCLC metastasis [30], and extended prior observations of MORC2-mediated lung cancer growth [16] by linking MORC2 to stemness and invasion. Notably, this study identified MORC2′s novel role in glycolysis, a key cancer metabolic hallmark [31]. MORC2 knockdown reduced glucose consumption, lactate production, and ATP/ADP ratio, mirroring Dong et al. [22], who reported MORC2-mediated glycolysis in ovarian cancer, and Guddeti et al. [18], who demonstrated that MORC2 modulates LDHA expression and glucose metabolism to drive breast cancer progression. This cross-cancer conservation suggests MORC2 may universally regulate glycolysis (e.g., via glycolytic enzymes [17], expanding its oncogenic functions to metabolic reprogramming.

Bioinformatics (Genecards/JASPAR) identified PAX2 as a putative MORC2 transcription factor, with a conserved binding site in the MORC2 promoter. PAX2, a known NSCLC oncogene [21,32], was upregulated in TCGA NSCLC tumors, and its mRNA strongly correlated with MORC2 (r = 0.8407). Experimental validation confirmed PAX2′s direct regulation: siRNA-mediated PAX2 knockdown reduced MORC2 mRNA/protein; luciferase assays showed PAX2 depletion inhibited wild-type (but not mutated) MORC2 promoter activity; ChIP demonstrated PAX2 binding to the MORC2 promoter. These findings extend PAX2′s regulatory network, while PAX2 regulates VEGF in renal cancer [33], this study identifies MORC2 as a lung cancer target, highlighting tissue-specific programs.

H1299 xenograft experiments validated the PAX2/MORC2 axis. PAX2 knockdown inhibited tumor growth (reduced volume/weight) and altered architecture, consistent with *in vitro* results. Critically, MORC2 overexpression partially rescued PAX2 depletion's effects, restoring tumor growth and proliferation (Ki67 staining), establishing causality. This rescue experiment is critical as it establishes causality, confirming that PAX2 exerts its oncogenic

effects at least in part through MORC2. Wu et al. [34] used a similar *in vivo* rescue strategy to validate the miR-486-3p/CRABP2 axis in NSCLC, and the current study's consistent approach strengthens the translational relevance of the PAX2/MORC2 axis. Additionally, immunohistochemical staining confirmed reduced MORC2 expression in PAX2-knockdown tumors and recovery in the rescue group, further supporting the regulatory relationship between PAX2 and MORC2. These *in vivo* data complement *in vitro* findings, providing a comprehensive view of the PAX2/MORC2 axis in NSCLC progression. Inevitably, the clinical cohort size (43 paired specimens) is relatively small, which may limit the generalizability of the prognostic findings; incomplete exploration of MORC2′s downstream pathways, and no testing of PAX2/MORC2-targeted therapies.

This study demonstrates the PAX2/MORC2 axis's critical role in NSCLC: PAX2 transcriptionally activates MORC2, whose overexpression promotes proliferation, survival, stemness, invasion, and glycolysis, and correlates with poor prognosis. These findings expand NSCLC molecular understanding and identify MORC2 as a potential prognostic biomarker/therapeutic target. Future studies should validate the axis in larger cohorts, explore MORC2′s downstream pathways, and develop targeted therapies to improve NSCLC outcomes.

CRediT authorship contribution statement

Junyan Cao: Writing – original draft, Validation, Software, Resources, Methodology, Conceptualization. **Fang Cai:** Validation, Supervision, Resources, Formal analysis, Data curation. **Jun Jia:** Validation, Software, Methodology, Formal analysis. **Jiangtao Guo:** Software, Investigation, Formal analysis. **Xiu Yang:** Writing – review & editing, Validation, Software, Investigation.

Financial support

None.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary material

https://doi.org/10.1016/j.ejbt.2026.100718.

Data availability

No data was used for the research described in the article.

References

[1] Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229–63. https://doi.org/10.3322/caac.21834. PMid: 38572751.

[2] Liu Y, Han Z, Li Y, et al. Global research trends on ALK-TKIs in non-small cell lung cancer: A bibliometric analysis. Front Pharmacol 2025;16:1665174. https://doi.org/10.3389/fphar.2025.1665174. PMid: 40894217.

[3] Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/American thoracic society/European respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol 2011;6(2):244–85. https://doi.org/10.1097/JTO.0b013e318206a221. PMid: 21252716.

[4] Molina JR, Yang P, Cassivi SD, et al. Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc 2008;83(5):584–94. https://doi.org/10.4065/83.5.584. PMid: 18452692.

[5] Relli V, Trerotola M, Guerra E, et al. Abandoning the notion of non-small cell lung cancer. Trends Mol Med 2019;25(7):585–94. https://doi.org/10.1016/j.molmed.2019.04.012. PMid: 31155338.

[6] Li L, Li Y, Ma X, et al. A pomalidomide-based gefitinib PROTAC degrader effectively inhibits lung cancer progression in EGFR-TKIs-acquired resistant models by targeting EGFR degradation and ETFA-mediated ATP generation. Bioorg Chem 2025;164:108864. https://doi.org/10.1016/j.bioorg.2025.108864. PMid: 40815992.

[7] Orooji N, Babaei S, Fadaee M, et al. Novel therapeutic approaches for non-small cell lung cancer: An updated view. J Drug Target 2025;33(8):1306–21. https://doi.org/10.1080/1061186x.2025.2489986. PMid: 40186594.

[8] Deshpand R, Chandra M, Rauthan A. Evolving trends in lung cancer: Epidemiology, diagnosis, and management. Indian J Cancer 2022;59(Suppl. 1):S90–S105. https://doi.org/10.4103/ijc.IJC_52_21. PMid: 35343194.

[9] Kandemir O, Demir F. An investigation of the relationship between ^{18}F-FDG PET/CT parameters of primary tumors and lymph node metastasis in resectable non-small cell lung cancer. Curr Radiopharm 2024;17(1):111–6. https://doi.org/10.2174/1874471016666230829100703. PMid: 37644744.

[10] Jin Y, Wang Z, He D, et al. Identification of novel subtypes based on ssGSEA in immune-related prognostic signature for tongue squamous cell carcinoma. Cancer Med 2021;10(23):8693–707. https://doi.org/10.1002/cam4.4341. PMid: 34668665.

[11] Jia ZC, Li YQ, Zhou BW, et al. Transcriptomic profiling of human granulosa cells between women with advanced maternal age with different ovarian reserve. J Assist Reprod Genet 2023;40(10):2427–37. https://doi.org/10.1007/s10815-023-02915-8. PMid: 37589858.

[12] Herrera LRM. In silico approach in designing a novel multi-epitope vaccine candidate against non-small cell lung cancer with overexpressed G protein-coupled receptor 56. Asian Pac J Cancer Prev 2020;21(8):2297–306. https://doi.org/10.31557/apjcp.2020.21.8.2297. PMid: 32856858.

[13] Zhou LX, Zhou Q, Gao TM, et al. Machine learning predicts acute respiratory failure in pancreatitis patients: A retrospective study. Int J Med Inf 2024;192:105629. https://doi.org/10.1016/j.ijmedinf.2024.105629. PMid: 39321493.

[14] Jacquier A, Roubille S, Lomonte P, et al. Microrchidia CW-type zinc finger 2, a chromatin modifier in a spectrum of peripheral neuropathies. Front Cell Neurosci 2022;16:896854. https://doi.org/10.3389/fncel.2022.896854. PMid: 35722617.

[15] Zhang J, Yang Y, Dong Y, et al. Microrchidia family CW–type zinc finger 2 promotes the proliferation, invasion, migration and epithelial–mesenchymal transition of glioma by regulating PTEN/PI3K/AKT signaling via binding to N–myc downstream regulated gene 1 promoter. Int J Mol Med 2022;49(2):16. https://doi.org/10.3892/ijmm.2021.5071. PMid: 34913078.

[16] Liu M, Sun X, Shi S. MORC2 enhances tumor growth by promoting angiogenesis and tumor-associated macrophage recruitment via Wnt/β-catenin in lung cancer. Cell Physiol Biochem 2018;51(4):1679–94. https://doi.org/10.1159/000495673. PMid: 30504718.

[17] Guddeti RK, Pacharla H, Yellapu NK, et al. MORC2 and MAX contributes to the expression of glycolytic enzymes, breast cancer cell proliferation and migration. Med Oncol 2023;40(3):102. https://doi.org/10.1007/s12032-023-01974-2. PMid: 36802305.

[18] Guddeti RK, Thomas L, Kannan A, et al. The chromatin modifier MORC2 affects glucose metabolism by regulating the expression of lactate dehydrogenase A through a feed forward loop with c-Myc. FEBS Lett 2021;595(9):1289–302. https://doi.org/10.1002/1873-3468.14062. PMid: 33626175.

[19] Cai H, Zhang F, Xu F, et al. Metabolic reprogramming and therapeutic targeting in non-small cell lung cancer: Emerging insights beyond the Warburg effect. Front Oncol 2025;15:1564226. https://doi.org/10.3389/fonc.2025.1564226. PMid: 40469185.

[20] Li L, Hossain SM, Eccles MR. The role of the PAX genes in renal cell carcinoma. Int J Mol Sci 2024;25(12):6730. https://doi.org/10.3390/ijms25126730. PMid: 38928435.

[21] Ren Y, Hou J, Xu A, et al. Diagnostic utility of PAX2 and PAX5 in distinguishing non-small cell lung cancer from small cell lung cancer. Int J Clin Exp Pathol 2015;8(11):14709–16. PMid: 26823795.

[22] Dong N, Lei P, Cai H, et al. The impact of MORC2 on glycolysis and the responsiveness of paclitaxel-resistant ovarian cancer cells. Cytotechnology 2025;77(6):192. https://doi.org/10.1007/s10616-025-00864-x. PMid: 41216303.

[23] Zhang T, Qiu L, Cao J, et al. ZFP36 loss-mediated BARX1 stabilization promotes malignant phenotypes by transactivating master oncogenes in NSCLC. Cell Death Dis 2023;14(8):527. https://doi.org/10.1038/s41419-023-06044-z. PMid: 37587140.

[24] Cantor E, Guauque-Olarte S, León R, et al. Knowledge-slanted random forest method for high-dimensional data and small sample size with a feature selection application for gene expression data. Biodata Min 2024;17(1):34. https://doi.org/10.1186/s13040-024-00388-8. PMid: 39256872.

[25] Sun Z, Du T, Yang G, et al. Identification of exosome-related genes in NSCLC via integrated bioinformatics and machine learning analysis. Sci Rep 2025;15(1):22962. https://doi.org/10.1038/s41598-025-04485-4. PMid: 40595797.

[26] Liu J, Shao Y, He Y, et al. MORC2 promotes development of an aggressive colorectal cancer phenotype through inhibition of NDRG1. Cancer Sci 2019;110(1):135–46. https://doi.org/10.1111/cas.13863. PMid: 30407715.

[27] Liao X, Liu C, Ding Z, et al. High expression of MORC2 predicts worse neoadjuvant chemotherapy efficacy in triple negative breast cancer. Medicine

(Baltimore) 2023;102(25):e34164. https://doi.org/10.1097/md.0000000000034164. PMid: 37352040.
[28] Zhang FL, Yang SY, Liao L, et al. Dynamic SUMOylation of MORC2 orchestrates chromatin remodelling and DNA repair in response to DNA damage and drives chemoresistance in breast cancer. Theranostics 2023;13(3):973–90. https://doi.org/10.7150/thno.79688. PMid: 36793866.
[29] Liao G, Liu X, Wu D, et al. MORC2 promotes cell growth and metastasis in human cholangiocarcinoma and is negatively regulated by miR-186-5p. Aging 2019;11(11):3639–49. https://doi.org/10.18632/aging.102003. PMid: 31180332.
[30] Liu L, Tao T, Liu S, et al. An RFC4/Notch1 signaling feedback loop promotes NSCLC metastasis and stemness. Nat Commun 2021;12(1):2693. https://doi.org/10.1038/s41467-021-22971-x. PMid: 33976158.
[31] Ahamed F, Van Criekinge M, Wang ZJ, et al. Modeling hyperpolarized lactate signal dynamics in cells, patient-derived tissue slice cultures and murine models. NMR Med 2021;34(3):e4467. https://doi.org/10.1002/nbm.4467. PMid: 33415771.
[32] Li Z, Huang M, Wang J, et al. Betulin, an active component from Chinese herb birch bark, suppresses tumor angiogenesis and tumor growth by inhibiting the PAX2/VEGF-A/VEGR2 signaling pathway in non-small cell lung cancer. Phytomedicine 2025;148:157445. https://doi.org/10.1016/j.phymed.2025.157445. PMid: 41166909.
[33] Maliszewska-Olejniczak K, Brodaczewska KK, Bielecka ZF, et al. Three-dimensional cell culture model utilization in renal carcinoma cancer stem cell research. In: Barata M, editor. Epithelial cell culture: Methods and protocols methods in molecular biology, vol 1817. New York: Humana Press; 2018. p. 47–66.
[34] Wu Y, Xie J, Wang H, et al. Circular RNA hsa_circ_0011298 enhances taxol resistance of non-small cell lung cancer by regulating miR-486-3p/CRABP2 axis. J Clin Lab Anal 2022;36(5):e24408. https://doi.org/10.1002/jcla.24408. PMid: 35396749.