



OPEN A novel lncRNA AC112721.1 promotes the progression of triple-negative breast cancer by directly binding to THBS1 and regulating miR-491-5p/C2CD2L axis

Ping Ma^{1,3}, Shiyao Kang^{1,3}, Huimin Li¹, Ming Li¹, Yuan Zhao², Hongjun Yuan¹, Jianyu Pang¹, Wenru Tang¹✉ & Miaomiao Sheng¹✉

Triple-negative breast cancer (TNBC) seriously threatens women's health, and long noncoding RNAs (lncRNAs) are emerging as critical regulators of gene expression and play fundamental roles in TNBC. This study aimed to identify lncRNAs that represent effective targets for the early diagnosis or treatment of TNBC. Here, we utilized the TCGA database to analyze differentially expressed genes, and survival analysis and ROC curve analysis were also performed. Notably, we identified a novel lncRNA, AC112721.1, that is significantly overexpressed in TNBC and is associated with poor overall survival in TNBC patients. Loss- and gain-of-function experiments revealed that AC112721.1 significantly promoted cell proliferation and migration, suppressed cell apoptosis in vitro and inhibited tumorigenesis in vivo. Further study of the mechanisms underlying these effects revealed that AC112721.1 regulates the Ras pathway by directly binding to THBS1 protein and functions as a ceRNA by sponging miR-491-5p to increase the expression of C2CD2L, thereby influencing the progression of TNBC. Our findings indicate that AC112721.1 may represent a new biomarker for evaluating TNBC prognosis and treating TNBC.

Keywords Triple-negative breast cancer, AC112721.1, THBS1, miR-491-5p, C2CD2L, Progression

The global cancer statistics for 2020 revealed that there were as many as 2.26 million new cases of breast cancer that year, surpassing the 2.2 million cases of lung cancer and making breast cancer the leading cancer worldwide¹. In China, as in many other countries, breast cancer has become the most common cancer among women, with a high incidence that ranks it as the number one threat to women's health².

Triple-negative breast cancer (TNBC) accounts for approximately 10–20% of all breast cancer cases. TNBC is characterized by the absence of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) expression³. TNBC tends to occur at a younger age than do other types of breast cancer; it is poorly differentiated, has high invasiveness, and is prone to metastasize to internal organs⁴. It has a high tumor mutation frequency, making it challenging to study the mechanisms of the disease. The effectiveness of TNBC treatment is poor, and TNBC is more likely than other types of breast cancer to develop drug resistance; patients with TNBC have a five-year survival rate of less than 80%. Additionally, conventional hormone therapy and targeted therapy methods are ineffective for TNBC. Currently, comprehensive treatment, including surgical resection, radiotherapy, and chemotherapy, is typically used⁵. Chemotherapy is the main treatment modality for TNBC, but it has significant side effects⁶. To date, the specific pathogenesis and mechanism of TNBC have not been fully clarified. Therefore, elucidating the complex molecular mechanism underlying TNBC progression and identifying the key molecules that regulate its progression are crucial for the treatment of TNBC.

Long noncoding RNAs (lncRNAs), a class of RNA molecules that are greater than 200 nucleotides in length, have various functions within organisms⁷. First, they serve as messenger molecules, transmitting information between the nucleus and the cytoplasm and regulating gene expression⁸. Second, lncRNAs bind to DNA

¹Laboratory of Molecular Genetics of Aging & Tumor, Medical School, Kunming University of Science and Technology, Chengong Campus, 727 South Jingming Road, Kunming 650500, Yunnan, China. ²Department of Oncology, Puer People's Hospital, Puer 665000, Yunnan, China. ³Ping Ma and Shiyao Kang contributed equally to this work. ✉email: twr@sina.com; shengmm@kust.edu.cn

sequences and thereby act as transcriptional regulators, influencing chromatin structure and modifications and thereby affecting gene transcription⁹. In addition, lncRNAs regulate the posttranscriptional processing of RNA. Finally, lncRNAs act as competitive endogenous RNAs, influencing the regulatory effects of miRNAs¹⁰. In addition to these functions, accumulating evidence has demonstrated that lncRNAs regulate cellular differentiation, proliferation, apoptosis, and other cellular processes and that abnormal lncRNA expression and overactivation are often involved in cancer initiation and progression.

Recently, it has been suggested that lncRNA dysregulation may play an essential role in the development and recurrence of TNBC¹¹. For example, Liu et al. reported that JAG1 enhances angiogenesis in TNBC by promoting secretion of the exosomal lncRNA MALAT1¹². Our previous studies also revealed that downregulation of HAGLROS inhibits the proliferation and migration of MDA-MB-231 cells¹³. LINC01234 was shown to promote TNBC progression by regulating the miR-429/SYNY1 axis¹⁴. The results of these studies suggest that lncRNAs may be novel and promising biomarkers for the treatment of TNBC, and identifying additional specific lncRNAs and elucidating their detailed mechanism of action in TNBC pathogenesis is highly important.

In this study, we used the TCGA database to predict molecules that are differentially expressed in TNBC and are strongly associated with TNBC prognosis. Through preliminary screening, we identified a novel lncRNA, AC112721.1, that is overexpressed in TNBC. By conducting *in vivo* and *in vitro* experiments, we investigated the role of AC112721.1 in TNBC progression. Furthermore, through transcriptome sequencing, RNA pull-down-MS, and database analysis, we elucidated the potential target genes and signaling molecules regulated by AC112721.1. Taken together, our data suggest that AC112721.1 is a potential biomarker and therapeutic target for TNBC.

Results

Identification of a novel lncRNA, AC112721.1, that is highly expressed in TNBC and associated with poor prognosis

To identify novel lncRNAs in TNBC, we analyzed differential gene expression in four subtypes of breast cancer using the TCGA database (Fig. 1A). We also analyzed the correlation between all of the identified differentially expressed genes and overall survival and found that 159 genes, including 14 novel lncRNAs (Supplementary Table 4), were significantly associated with survival time in TNBC patients (Fig. 1B). Among them, AC112721.1 expression was markedly higher in TNBC tissues than in normal tissues, and high AC112721.1 expression correlated with poor survival (Fig. 1C,D). Next, we generated receiver operating characteristic (ROC) curves of AC112721.1 expression and measured the area under the ROC curve (AUC) (0.889) (Fig. 1E). Finally, we evaluated the expression of AC112721.1 in clinical samples and found that AC112721.1 was highly expressed in TNBC tissues (Fig. 1F). Taken together, these data suggest that AC112721.1 is highly expressed in breast tumors and is specifically correlated with poor survival in TNBC patients. AC112721.1 can also be used as a biomarker to diagnose TNBC with high sensitivity and specificity.

AC112721.1 promotes TNBC cell growth and migration

To investigate the possible role of AC112721.1 in the pathogenesis of TNBC, we first measured the levels of AC112721.1 in four breast cancer cell lines and in an immortalized human mammary cell line (MCF-10 A) and found that AC112721.1 was significantly more highly expressed in TNBC cells (MDA-MD-468 and MDA-MB-231) (Fig. 2A).

We next used shRNA to knock down AC112721.1 expression in MDA-MB-231 cells. The efficiency of AC112721.1 knockdown was confirmed via qRT-PCR (Fig. 2B). CCK-8, flow cytometry and migration assays were performed on the cell lines. The results revealed that knockdown of AC112721.1 significantly decreased cell proliferation, promoted cell apoptosis, and inhibited cell migration (Fig. 2C–E). In contrast, overexpression of AC112721.1 increased cell proliferation and migration (Fig. 2F–H). Similar results were obtained in MDA-MB-468 cells (Supplementary Fig. 1A–D).

To confirm the function of AC112721.1 in the tumorigenicity of TNBC, stable sh-AC112721.1 or sh-negative (control) MDA-MB-231 cell lines were injected *in situ* into nude immunodeficient mice. The results revealed that AC112721.1 knockdown significantly inhibited tumor growth, as shown by decreased tumor volume and weight (Fig. 2I–K). These findings indicate that AC112721.1 expression may contribute to the development and progression of TNBC.

AC112721.1 interacts directly with THBS1 and promotes the progression of TNBC by upregulating THBS1

To determine how AC112721.1 functions in TNBC, we conducted transcriptome sequencing and compared the gene expression profiles of AC112721.1-knockdown MDA-MB-231 cells with those of cells transfected with scrambled DNA. The results revealed that there were 290 differentially expressed genes in the si-AC112721.1 group, 242 of which have names (Fig. 3A). Next, we performed an RNA pull-down assay using biotinylated AC112721.1 probes to determine the potential binding proteins in MDA-MB-231 cells. Via mass spectrometry analysis (Fig. 3B), 77 proteins were found to interact with AC112721.1 (Supplementary Table 5); comparison of these with the 242 differentially expressed genes identified in the previous data yielded THBS1 (Fig. 3C and Supplementary Fig. 2A). The interaction between AC112721.1 and THBS1 was further confirmed via RIP assays (Fig. 3D). Additionally, we found that expression of THBS1 decreased or increased significantly when AC112721.1 was knocked down or overexpressed, respectively (Fig. 3E and F and Supplementary Fig. 2C), consistent with the sequencing results (Supplementary Fig. 2B). Therefore, we focused on THBS1.

We investigated whether THBS1 promotes the progression of TNBC. First, we examined its expression in MDA-MB-231 cells and found it to be significantly upregulated compared with that in MCF-10 A cells (Fig. 3G). Moreover, knockdown of THBS1 significantly inhibited cell proliferation and migration and promoted apoptosis

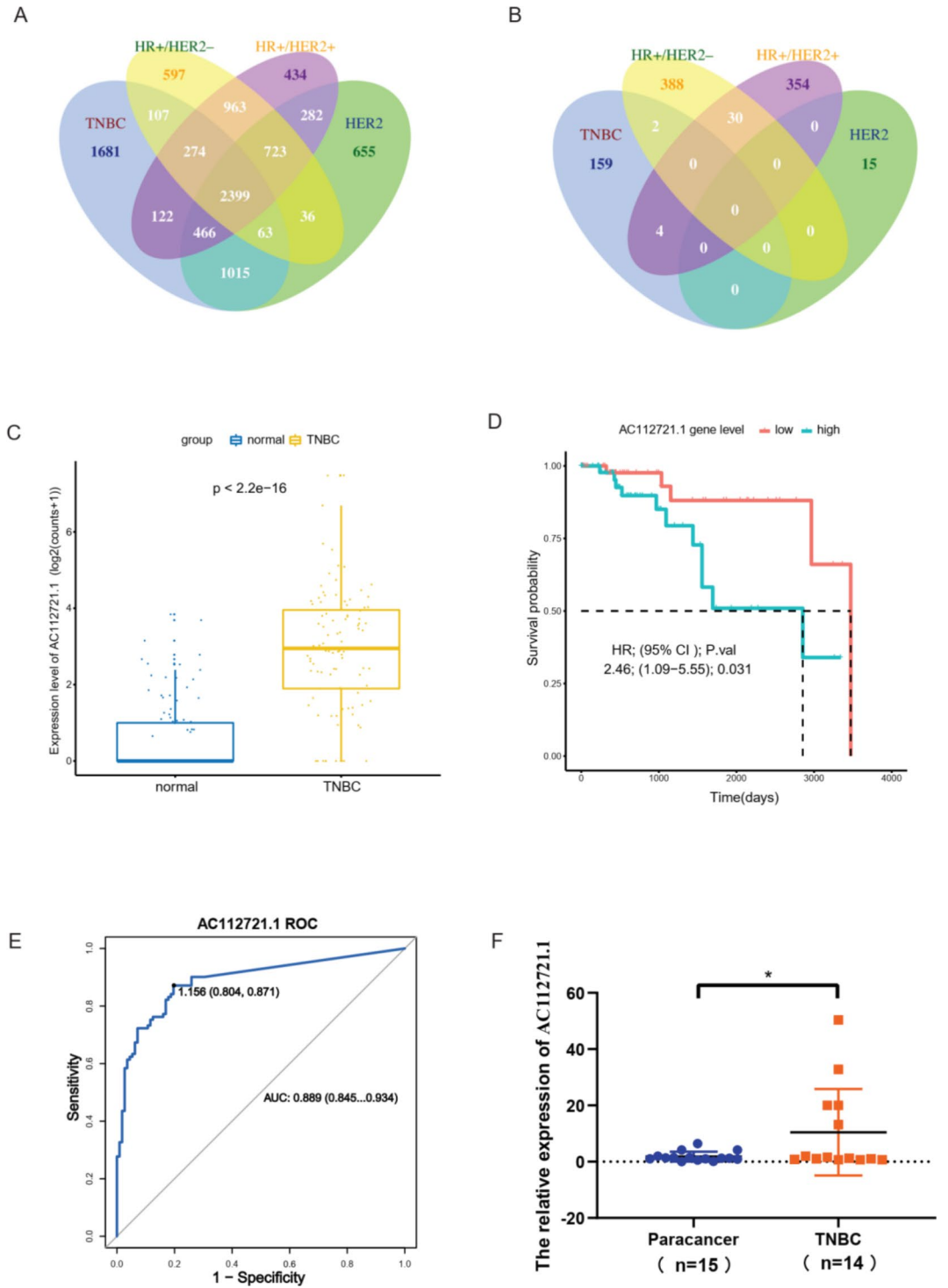


Fig. 1. AC112721.1 was highly expressed in TNBC and associated with poor prognosis. (A) Venn Diagram showed the number of genes that were significantly differentially expressed in four subtypes of breast cancer. (B) Venn Diagram showed the number of genes that were significantly associated with survival time in four subtypes of breast cancer. (C) The expression levels of AC112721.1 in TNBC tissues ($N=163$) and normal tissues ($N=61$) in the TCGA database. (D) Kaplan-Meier survival analysis of TNBC patients stratified by AC112721.1 expression level. (E) ROC curves were used to determine the diagnostic value of AC112721.1 expression in TNBC based on TCGA data. (F) The relative expression of AC112721.1 was detected by qRT-PCR in TNBC tissues ($N=14$) and paracancer tissues ($N=15$) (Wilcoxon test). N, number. * $p < 0.05$.

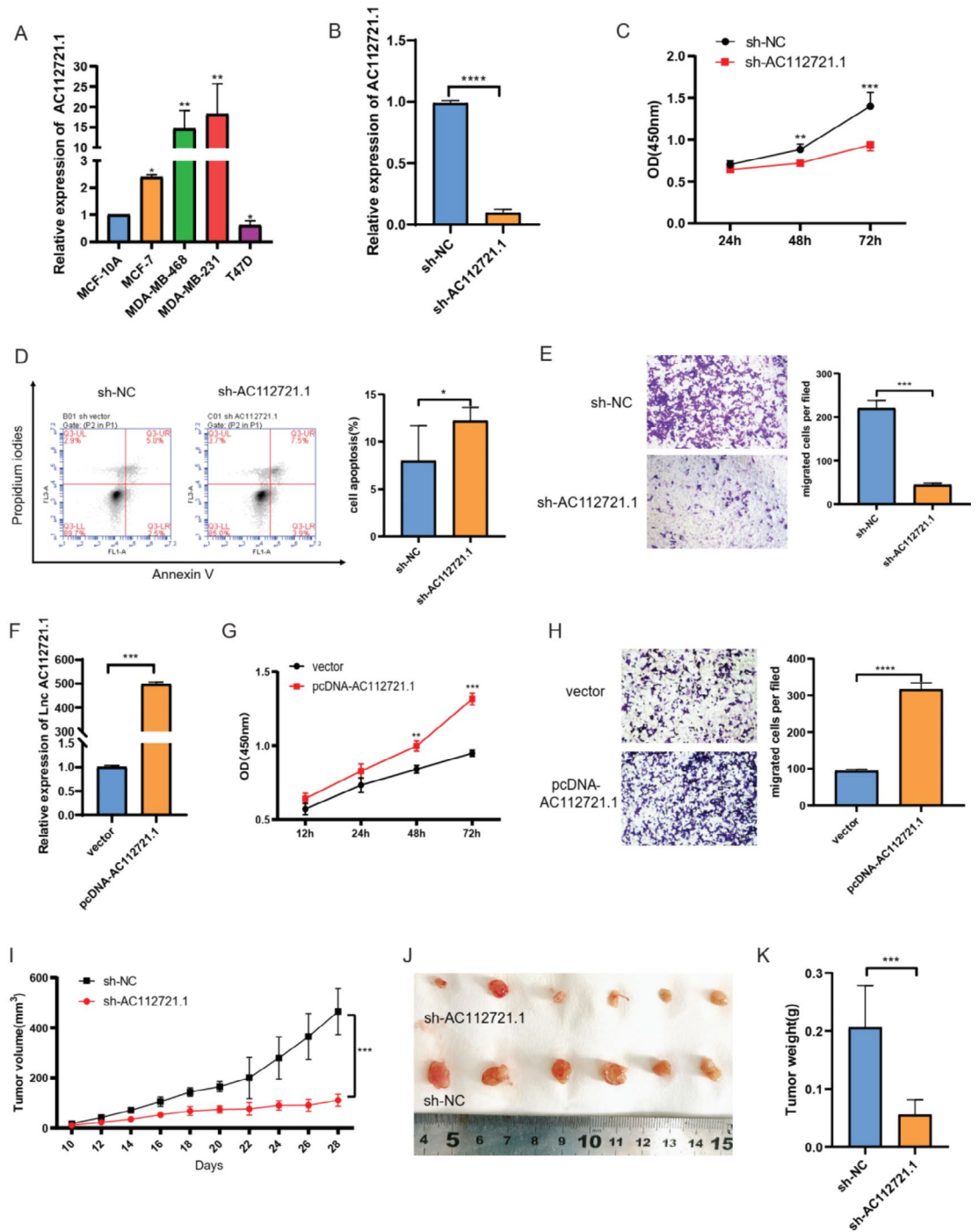


Fig. 2. AC112721.1 promoted TNBC cell growth and migration in vitro and in vivo. (A) The relative expression of AC112721.1 in four breast cancer cell lines and one immortalized human mammary cell line (MCF-10 A). (B) The expression efficiency of plasmid-mediated AC112721.1 knockdown was detected in MDA-MB-231 cell line using qRT-PCR. (C) CCK-8 assays showing the growth of MDA-MB-231 cells treated with sh-AC112721.1. (D) FACS was performed to determine the apoptosis rate of MDA-MB-231 cells after knockdown of AC112721.1. (E) Transwell assays in MDA-MB-231 cells treated with sh-AC112721.1. (F) The expression efficiency of plasmid-mediated AC112721.1 overexpression was detected in MDA-MB-231 cell line using qRT-PCR. (G) CCK-8 assays assessed proliferation of MDA-MB-231 cells transfected with pcDNA-AC112721.1. (H) Cell migration ability of MDA-MB-231 cells was determined after infection with pcDNA-AC112721.1. (I) Tumor volume was measured every 2 days. (J) Images of tumors of each group. (K) Tumor weights of each group. NC negative control. FACS, fluorescence activated cell sorting. ANOVA was used for statistical difference analysis. All experiments were performed at least thrice with triplicate samples. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

(Fig. 3H–K and Supplementary Fig. 2D); we obtained similar results in MDA-MB-468 cells (Supplementary Fig. 2E–G).

On the basis of the sequencing results, we performed a Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis. This revealed that the MAPK signaling pathway was the most downregulated pathway in AC112721.1-silenced cells (Fig. 3L). Further analysis revealed that the protein levels of three molecules involved in the MAPK signaling pathway (RAS, phosphorylated MEK and phosphorylated ERK) were obviously increased in cells with AC112721.1 overexpression and decreased after AC112721.1 knockdown (Fig. 3M and Supplementary Fig. 2H). However, the levels of these three proteins were downregulated when THBS1 was knocked down (Fig. 3N). These results suggest that AC112721.1 may regulate the MAPK signaling pathway by targeting the THBS1 protein, thus affecting the progression of TNBC.

AC112721.1 functions as a sponge for miR-491-5p in MDA-MB-231 cells

Studies have shown that lncRNAs can also act as ceRNAs to bind miRNAs and thereby play important regulatory roles in the occurrence and development of tumors. To explore the miRNAs regulated by AC112721.1 in TNBC, the 6 miRNAs that show the highest correlations with AC112721.1, namely, hsa-miR-24-1-5p, hsa-miR-3926-1, hsa-miR-483-5p, hsa-miR-489-5p, hsa-miR-491-5p, and hsa-miR-6788-5p, were predicted via correlation analysis and the cis-trans regulation method (Supplementary Table 6). Further RT-PCR revealed that only hsa-miR-491-5p expression was negatively correlated with AC112721.1 expression (Supplementary Fig. 3A,B). Compared with that in MCF-10 A cells, the expression of miR-491-5p was lower in MDA-MB-231 cells (Supplementary Fig. 3C). A dual-luciferase reporter assay demonstrated that the miR-491-5p mimic significantly reduced the luciferase activity present in wild-type AC112721.1 cells but not that present in mutant AC112721.1 cells (Fig. 4A), suggesting that miR-491-5p binds directly to AC112721.1.

To investigate the biological functions of miR-491-5p in TNBC cells, we overexpressed or knocked down miR-491-5p in MDA-MB-231 cells. The results revealed that cell proliferation and migration were decreased and cell apoptosis was increased when miR-491-5p was overexpressed (Fig. 4B–E); similar results were observed in MDA-MB-468 cells (Supplementary Fig. 3D–F). In contrast, opposite results were observed when miR-491-5p was knocked down (Fig. 4F–I). We then cotransfected MDA-MB-231 cells with the miR-491-5p mimic and AC112721.1; we found that upregulation of miR-491-5p weakened the ability of AC112721.1 overexpression to promote cell proliferation and migration (Fig. 4J,K).

AC112721.1 regulates the progression of TNBC through the miR-491-5p/C2CD2L axis

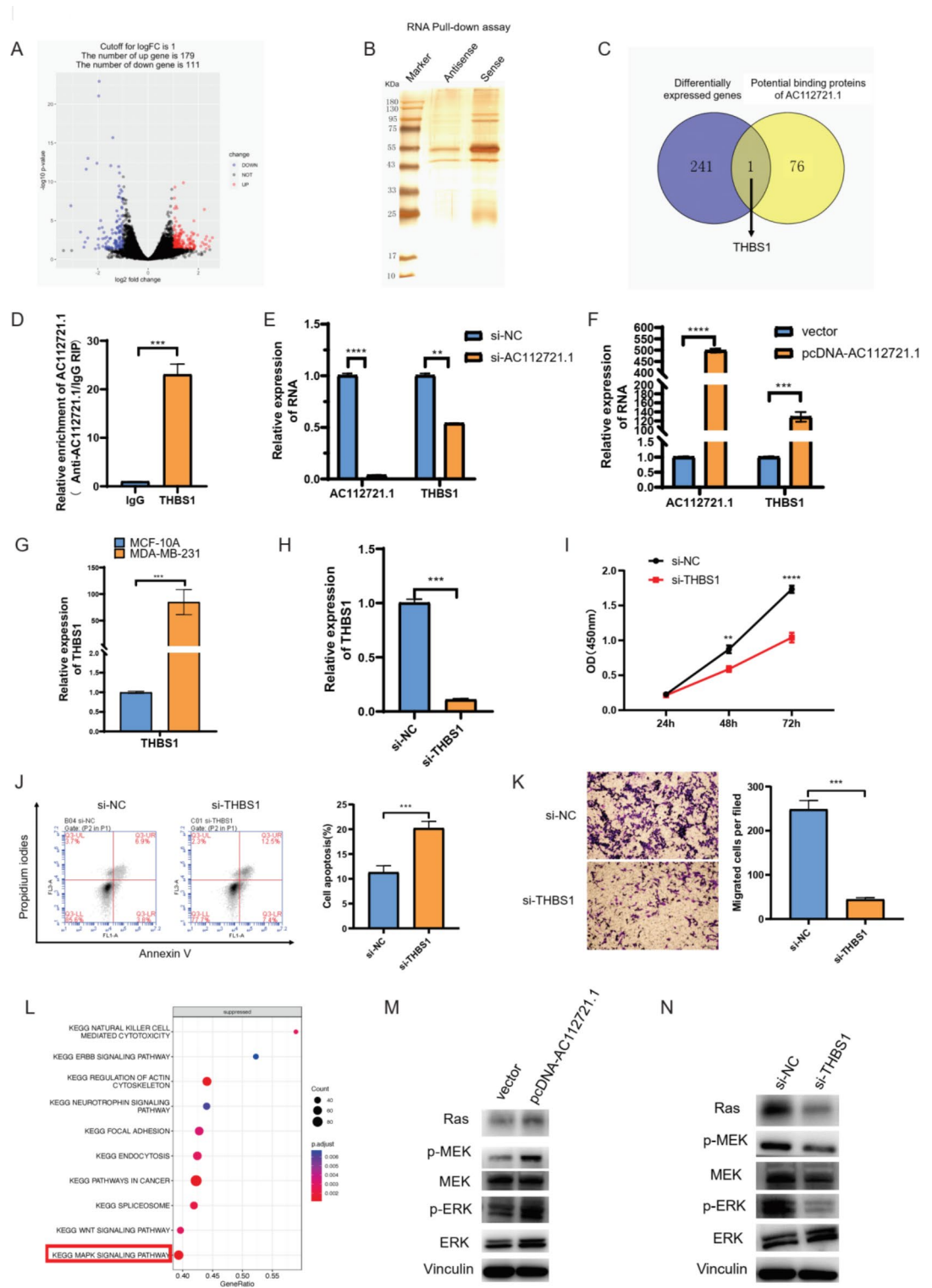
To further clarify the downstream target genes of miR-491-5p, TargetScan, RNA22, PolymiRT, and MiRBase were used for prediction (Supplementary Table 7). In this way, 45 potential target genes were identified (Fig. 5A). On the basis of the differential expression of target genes and information obtained from literature mining, C2CD2L was preliminarily identified. qRT-PCR and Western blot experiments revealed a positive correlation between the expression levels of C2CD2L and AC112721.1 in MDA-MB-231 cells (Fig. 5B and Supplementary Fig. 4A) as well as a negative correlation between the expression levels of C2CD2L and miR-491-5p (Fig. 5C); moreover, the increase in the mRNA levels of C2CD2L induced by AC112721.1 overexpression was reduced by miR-491-5p upregulation (Fig. 5D). A luciferase assay further revealed that transfection with the miR-491-5p mimic reduced the luciferase activity of the C2CD2L 3'UTR luciferase construct but not that of the mutant construct (Fig. 5E). These results indicate that C2CD2L is the target gene of miR-491-5p and that AC112721.1 antagonizes the inhibitory effect of miR-491-5p on C2CD2L expression.

Given the above results, we investigated whether C2CD2L regulates the progression of TNBC. We found that C2CD2L was significantly upregulated in MDA-MB-231 cells (Supplementary Fig. 4B) and TNBC tumors (Supplementary Fig. 4C) and that patients with high C2CD2L expression exhibited poorer overall survival than those with low C2CD2L expression (Supplementary Fig. 4D). These findings suggest that C2CD2L may play an important role in TNBC progression. We then knocked down C2CD2L in MDA-MB-231 cells (Fig. 5F); such knockdown significantly reduced cell proliferation and migration and increased cell apoptosis (Fig. 5G–I). Similar results were observed in MDA-MB-468 cells (Supplementary Fig. 4E–G). Collectively, these findings demonstrate that AC112721.1 is able to regulate the progression of TNBC through the miR-491-5p/C2CD2L axis.

Discussion

TNBC is the most rapidly progressing type of malignant breast cancer and the most difficult subtype of breast cancer to diagnose and treat in clinical practice. Compared with patients with other subtypes of breast cancer, TNBC patients have characteristics such as short overall survival, high malignancy, strong invasive potential, and high recurrence rates⁶. Because TNBC is often diagnosed at an advanced stage, its clinical treatment poses a great challenge. Because TNBC does not express ER, PR, or HER2, patients with TNBC often do not benefit from endocrine or targeted therapies¹⁵. Conventional chemotherapy remains the main treatment option for early and advanced TNBC patients. In recent years, neoadjuvant chemotherapy with platinum and ruthenium-containing drugs has been shown to have positive therapeutic effects in some TNBC patients who are sensitive to chemotherapy¹⁵. However, more than 50% of TNBC patients with stage I–III tumors still experience recurrence, and more than 37% of patients die within 5 years⁵. Therefore, identifying new tumor markers and elucidating their pathogenic mechanisms are highly important for promoting TNBC diagnosis and improving treatment efficacy.

LncRNAs are defined as transcripts of more than 200 nucleotides. They often interact with DNA, RNA, and proteins through their complex three-dimensional structures and thereby participate in a wide range of biological activities, including regulation of gene expression, chromatin inactivation, pluripotency and differentiation, as well as regulation of alternative splicing^{9,16}. Recent studies have shown that lncRNAs are



important regulatory factors in cancer-related cellular pathways^{17,18}. Dysregulation of lncRNAs is associated with cell proliferation, invasion, migration, chemotherapy resistance, and prognosis in TNBC^{19,20}. Owing to these unique characteristics, lncRNAs have the potential to serve as novel diagnostic and prognostic biomarkers for TNBC treatment.

The aim of this study was to identify lncRNAs that can provide effective targets for the early diagnosis or treatment of TNBC. On this basis, we first obtained transcriptome sequencing data from TNBC patients and clinical information on these patients from the TCGA database website. Differential expression analysis of TNBC and normal samples was performed using three methods: limma, edgeR, and DESeq2. The differentially expressed genes were then subjected to survival analysis and ROC curve analysis. Among these genes, AC112721.1 expression was markedly higher in TNBC tissues than in normal tissues, and high AC112721.1 expression correlated with poor survival. Moreover, the ROC curve analysis results also demonstrated that this

◀ **Fig. 3.** AC112721.1 directly interacted with THBS1 and promoted the progression of TNBC by upregulating THBS1. (A) Volcano plot showed the differentially expressed genes (si-AC112721.1 vs. si-NC, $|\log FC| > 1$, $p < 0.05$). (B) Sliver staining for RNA pull-down assay with the specific biotin-labeled AC112721.1 probe in MDA-MB-231 lysates. (C) Venn plot showing the overlap of AC112721.1 binding proteins and differentially expressed genes. (D) RIP assays were performed using THBS1 antibody. qRT-PCR was used to detect AC112721.1 enrichment, using IgG antibody as the control. (E) The relative expression level of THBS1 after AC112721.1 knockdown in MDA-MB-231 cells. (F) The relative expression level of THBS1 after AC112721.1 overexpression in MDA-MB-231 cells. (G) The relative expression level of THBS1 in different cell lines. (H) qRT-PCR analysis of THBS1 expression in MDA-MB-231 cells transfected with si-NC or si-THBS1. (I) Cell proliferation of MDA-MB-231 cells was detected by CCK-8 assay with the indicated treatment. (J) FACS was performed to determine the apoptosis rate of MDA-MB-231 cells after knockdown of THBS1. (K) Cell migration ability of MDA-MB-231 cells was determined after transfection with si-THBS1. (L) Top 10 KEGG enrichment pathway^{29–31}. (M) The impact of overexpressed AC112721.1 on the Ras pathway in MDA-MB-231 cells was measured using Western blot. (N) The impact of knocking down THBS1 on the Ras pathway in MDA-MB-231 cells was measured using Western blot. NC negative control. FACS, fluorescence activated cell sorting. ANOVA was used for statistical difference analysis. All experiments were performed at least thrice with triplicate samples. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

molecule has high sensitivity and specificity in predicting TNBC. Therefore, we selected AC112721.1 as the focus of our study.

AC112721.1 is a novel lncRNA, and no information on its role or mechanism of action in TNBC is currently available. Therefore, we conducted an initial investigation of the role of AC112721.1 in TNBC. The results revealed that AC112721.1 is significantly upregulated in TNBC cells (MDA-MB-231) compared with normal breast cells (MCF-10 A) and with other types of breast cancer cells. Knockdown of AC112721.1 suppressed the proliferation and migration of MDA-MB-231 cells while promoting apoptosis, whereas overexpression of AC112721.1 yielded the opposite results. In vivo experiments also demonstrated that AC112721.1 knockdown inhibited tumor growth. These findings suggest that AC112721.1 can promote the development of TNBC.

To further elucidate the mechanism of action of AC112721.1 in TNBC, we performed transcriptome sequencing and proteomic analysis. This led to the identification of thrombospondin 1 (THBS1) as a protein that binds specifically to AC112721.1. THBS1 is a major endogenous antiangiogenic protein that is secreted by various cell types. Research has shown that THBS1 is overexpressed in late-stage tumors, such as prostate cancer, pancreatic cancer, glioblastoma, breast cancer, gallbladder cancer, and melanoma^{21–23}. Furthermore, THBS1 has been found to regulate the immune response and the metastasis of tumor cells²⁴. In breast cancer, high expression of THBS1 is associated with poor prognosis²⁵. However, the mechanism of action of THBS1 in TNBC remains unclear. In the present study, THBS1 expression correlated positively with the expression of AC112721.1. Additionally, THBS1 knockdown significantly inhibited the proliferation and migration of MDA-MB-231 cells, promoted their apoptosis, and suppressed the Ras/MAPK pathway. These data indicate that AC112721.1 may regulate the proliferation, migration, and apoptosis of TNBC cells by binding to THBS1.

Studies have shown that lncRNAs can also act as ceRNAs, regulating the expression of other mRNAs by competitively binding to miRNAs. Such RNAs participate in the regulation of gene expression, form complex regulatory networks, and play important roles in cellular functions and disease-related processes. Here, through bioinformatics analysis and relevant experimental validation, we revealed that AC112721.1 can also upregulate the expression of C2CD2L by sponging miR-491-5p, promoting the proliferation and migration of TNBC cells while inhibiting apoptosis. C2CD2L is a membrane protein that is anchored to the endoplasmic reticulum²⁶. It mediates ER-plasma membrane contacts, activates the AKT-mTOR pathway, and regulates cell growth and protein synthesis. C2CD2L is highly expressed in the brain and in pancreatic islets²⁷. Loss of C2CD2L in insulin-secreting cells led to defects in insulin release²⁸. However, the mechanism of action of C2CD2L in TNBC remains unclear. Our study preliminarily confirmed that C2CD2L plays a role in promoting TNBC progression.

Although our study has produced several new findings, it has several limitations that should be mentioned. First, further investigations are needed to determine the biological functions of THBS1 overexpression. Second, rescue experiments and in vivo experiments should be performed to clarify the regulatory role of the AC112721.1/miR-491-5p/C2CD2L axis in TNBC. Third, it will also be necessary to test the level of expression of AC112721.1 and its downstream target genes in clinical samples and to determine their correlation with various pathological indicators and prognosis in patients with TNBC.

In conclusion, our study identified a novel lncRNA, AC112721.1, that acts as a regulator in TNBC. We demonstrated for the first time that AC112721.1 promotes cellular proliferation and migration and inhibits cellular apoptosis both in vitro and in vivo. Mechanistically, AC112721.1 regulates the Ras pathway by binding directly to the THBS1 protein and functions as a ceRNA by sponging miR-491-5p to increase the expression of C2CD2L, thereby influencing the progression of TNBC (Fig. 5J). Our findings provide new insight into TNBC pathogenesis and will be useful as a basis for exploring candidates for therapeutic targets in TNBC in the future.

Experimental procedures

Clinical specimens

Twenty triple-negative breast cancer samples and 20 paracancerous tissues were collected from 30 patients who underwent surgery at the First People's Hospital of Yunnan Province. All of the obtained tissue samples were histologically confirmed by a pathologist. The samples were immediately frozen in liquid nitrogen. The protocol was approved by the Ethical Committee of First People's Hospital of Yunnan Province (No. KMUST-MEC-016).

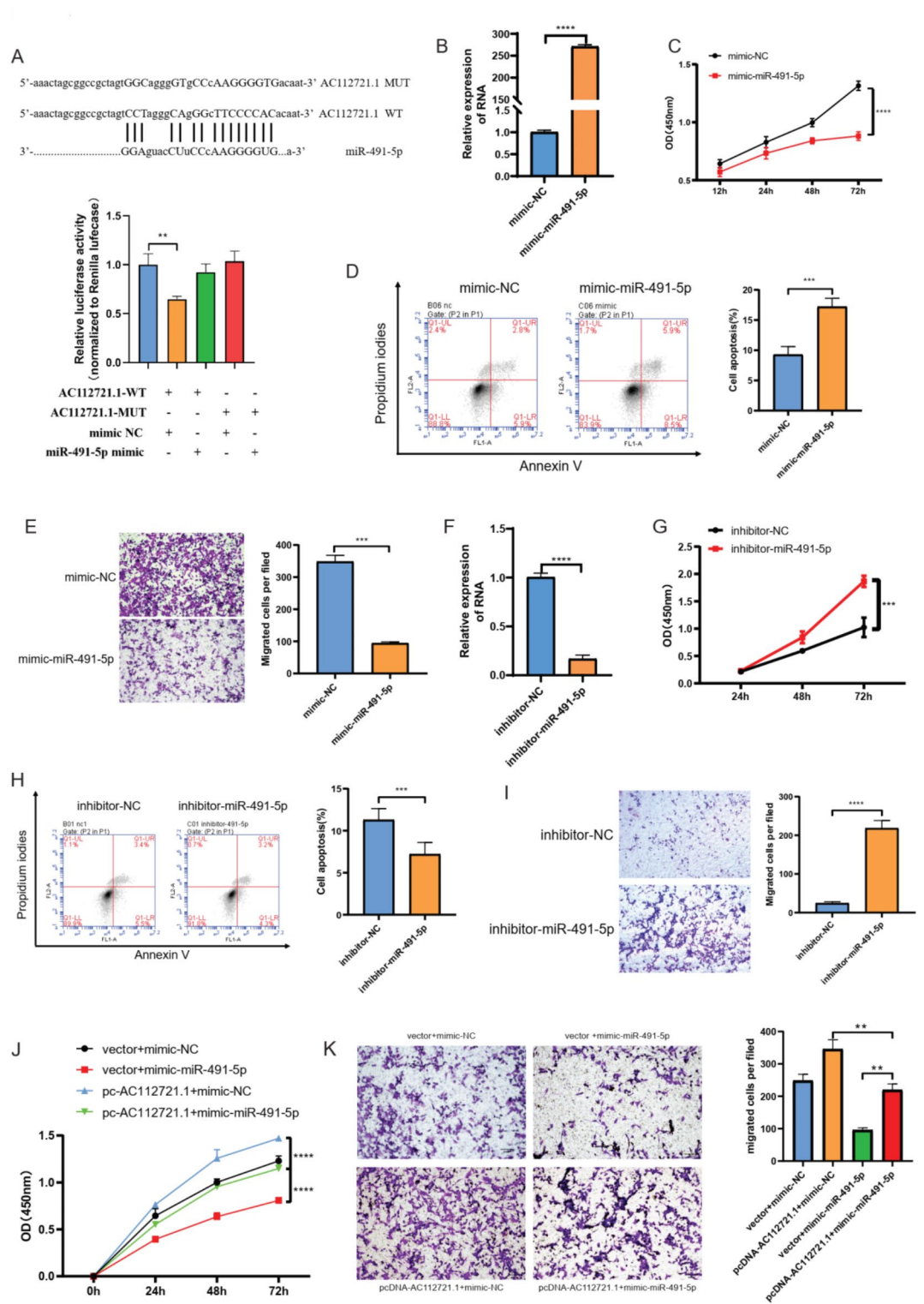


Fig. 4. AC112721.1 functioned as a miR-491-5p sponge in MDA-MB-231 cells. **(A)** The relative luciferase activities were detected in 293T cells co-transfected with miR-429 mimic or NC mimic and AC112721.1 WT or AC112721.1 MUT. **(B)** The expression efficiency of mimic-miR-491-5p was assessed using qRT-PCR. **(C)** Overexpression of miR-491-5p inhibited cell proliferation. **(D)** Overexpression of miR-491-5p promoted cell apoptosis. **(E)** Overexpression of miR-491-5p inhibited cell migration. **(F)** The knockdown efficiency of inhibitor-miR-491-5p was assessed using qRT-PCR. **(G)** Knockdown of miR-491-5p promoted cell proliferation. **(H)** Knockdown of miR-491-5p inhibited cell apoptosis. **(I)** Knockdown of miR-491-5p promoted cell migration. **(J)** The cell viability of MDA-MB-231 cells treated with pcDNA-AC112721.1 and mimic-miR-491-5p. **(K)** The cell migration ability of MDA-MB-231 cells treated with pcDNA-AC112721.1 and mimic-miR-491-5p. ANOVA was used for statistical difference analysis. All experiments were performed at least thrice with triplicate samples. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

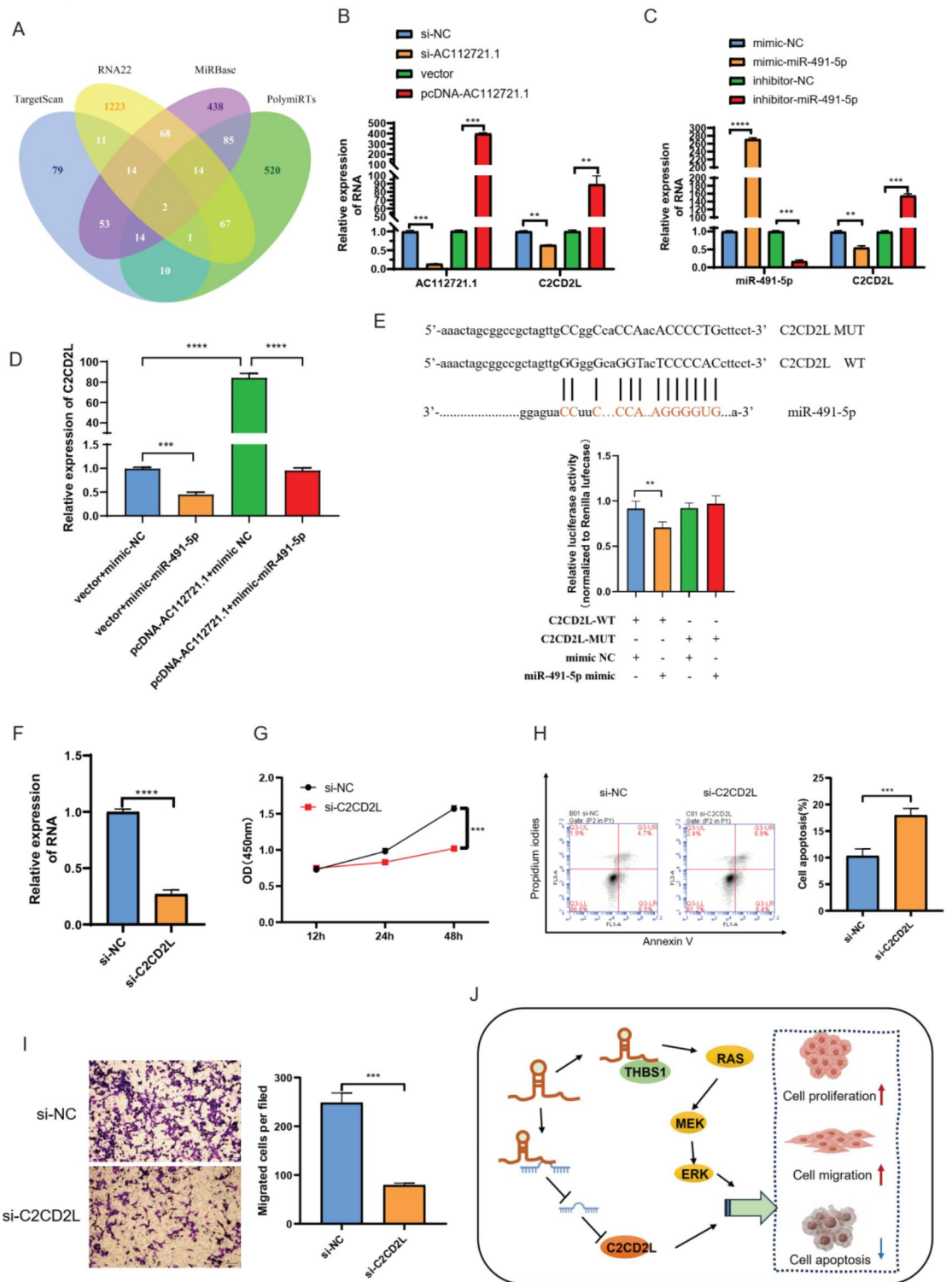


Fig. 5. AC112721.1 regulated the progression of TNBC through the miR-491-5p/C2CD2L axis. **(A)** Venn Diagram showed the target genes of miR-491-5p. **(B)** qRT-PCR was used to detect the expression level of C2CD2L after knockdown or overexpression of AC112721.1. **(C)** qRT-PCR was used to detect the expression level of C2CD2L after knockdown or overexpression of miR-491-5p. **(D)** The increase in the mRNA levels of C2CD2L induced by AC112721.1 overexpression was reduced by miR-491-5p upregulation. **(E)** The relative luciferase activities were detected in 293T cells co-transfected with miR-491-5p mimic or NC mimic and C2CD2L WT or C2CD2L MUT. **(F)** The knockdown efficiency of C2CD2L was assessed using qRT-PCR. **(G)** Knockdown of C2CD2L inhibited cell proliferation. **(H)** Knockdown of C2CD2L promoted cell apoptosis. **(I)** Knockdown of C2CD2L inhibited cell migration. **(J)** The mechanistic scheme of AC112721.1 promoting the progress of TNBC. ANOVA was used for statistical difference analysis. All experiments were performed at least thrice with triplicate samples. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Informed consent was obtained from all of the human subjects. Research on human subjects adheres to the Declaration of Helsinki principles.

The clinical information on the patients is shown in Supplementary Table 1.

TCGA data collection

The RNA-Seq datasets of 1,083 breast cancer samples and 104 normal samples were obtained from the TCGA database (<http://cancergenome.nih.gov/>). The clinical information on the samples in the TCGA cohort was matched with the RNA data. The differentially expressed genes (DEGs) in TNBC tissues compared with normal tissues were analyzed using three R language packages: DEGseq2, edgeR and limma. In this study, an absolute log₂-fold change > 1 and $p < 0.05$ were set as the cutoff criteria for the definition of DEGs. For survival analysis, we first extracted relevant variables such as patient survival status and survival time from the clinical data. We then grouped patients on the basis of gene expression levels and finally used the “survival” package to plot Kaplan–Meier survival curves. We used the pROC package in R software to assess the diagnostic value of AC112721.1. The receiver operating characteristic (ROC) curve is a tool that can be used to evaluate the performance of classification models. A curve is plotted by comparing the actual categories and the model’s predicted probabilities. According to the gene expression level, thresholds are set one by one, and the false positive rate (FPR) and the true positive rate (TPR) are calculated. The x-axis represents the FPR, and the y-axis represents the TPR, that is, the ability of the model to identify truly positive samples. The larger the area under the curve (AUC) is, the greater is the sensitivity and specificity of the test.

Cell culture

Four human breast cancer cell lines (MDA-MB-231, MDA-MB-468, MCF-7, and T47D) and the immortalized mammary epithelial-like cell line MCF-10 A were purchased from the Peking Union Medical College Cell Culture Center (Beijing, China) and cultured at 37 °C in a humidified incubator with 5% CO₂. MDA-MB-468, MDA-MB-231, and MCF-7 cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM; C11995500BT, Thermo Fisher Scientific, Waltham, USA) supplemented with 10% fetal bovine serum (FBS, 10099141, Thermo Fisher Scientific); T47D cells were maintained in Roswell Park Memorial Institute-1640 medium (C11875500BT, Thermo Fisher Scientific) supplemented with 10% FBS. MCF-10 A cells were cultured in DMEM/F12 (12500-062, Thermo Fisher Scientific) supplemented with 10% horse serum (HyClone, Utah, USA), 1 mg/mL epidermal growth factor (Merck KGaA, Darmstadt, Germany), 1 mg/mL cholera toxin (Merck KGaA), 20 mg/mL insulin (Merck KGaA), and 1 mg/mL hydrocortisone (Merck KGaA). All cell lines used were free of mycoplasma contamination.

Cell transfection

A lentivirus-based small hairpin RNA (shRNA) targeting AC112721.1 was constructed by General Biosystems (Anhui, China) and cloned and inserted into the pGreenPuro vector. The lentiviral vector was cotransfected with the packing vectors psPAX2 and pMD2.G into 293T cells for lentivirus production. To establish stable cell lines, MDA-MB-231 cells were transduced with the above lentivirus particles (2 mL) in the presence of polybrene (10 µg). After infection for 12 h, the culture medium containing the lentivirus particles was removed, and 4 mL of fresh medium containing puromycin was added to continue the culture. After 2 weeks, qRT-PCR was performed to detect AC112721.1 expression and confirm the stability of the cell lines.

For transient transfection, MDA-MB-231 cells were transfected with small interfering RNAs (siRNAs), pcDNA3.1-AC112721.1 plasmids, miRNA mimics or miRNA inhibitors using TransIntro EL transfection reagent (TransGen Biotech, Beijing, China) according to the manufacturer’s instructions. AC112721.1 siRNA, THBS1 siRNA, C2CD2L siRNA, the miR-491-5p mimic, the miR-491-5p inhibitor and matched negative controls were purchased from RiboBio (Guangzhou, China). The sequences of these RNAs are listed in Supplementary Table 2.

RNA extraction and quantitative RT-PCR

Total RNA was extracted from tissues and cell lines using TRIzol Reagent (T9424, Merck KGaA) according to the manufacturer’s instructions. cDNA was synthesized from RNA via the GoScript Reverse Transcription System (Promega, Wisconsin, USA). For qPCR, a specific forward primer was designed for each miRNA, and reverse primers were targeted to the stem-loop sequence. The expression levels of the miRNAs/mRNAs were analyzed via qPCR using SYBR Green Master Mix (Roche, Basel, Switzerland) with U6/GAPDH as an internal control. The expression level of each gene was calculated and normalized via the $2^{-\Delta\Delta Ct}$ method. An ABI 7300 instrument (Life Technology, USA) was used to conduct the qRT-PCR and collect the data. The primers used and their sequences are shown in Supplementary Table 2.

Cell proliferation assays

Cell proliferation was measured via a Cell Counting Kit-8 (CCK-8) (RiboBio, Guangzhou, China). The cells were seeded in 24-well plates and transfected at various time points; 50 µL of CCK-8 solution was then added to each well, and the cells were incubated for 2040 min at 37 °C under 5% CO₂. Finally, the absorbance was measured using a microplate reader (BioTek) at 450 nm.

Apoptosis analysis

An annexin V-fluorescein isothiocyanate apoptosis detection kit (11988549001, Roche) was used to measure cell apoptosis according to the manufacturer’s instructions. The cells were digested with trypsin and washed twice with PBS. The harvested cells were incubated in buffer supplemented with 1 µL of annexin V and 2 µL of propidium iodide for 30 min at 4 °C and analyzed via flow cytometry (Becton Dickinson, Franklin Lakes, USA).

Western blot analysis

Protein lysates were separated by electrophoresis on sodium dodecyl sulfate- polyacrylamide gels, and the separated proteins were transferred to polyvinylidene fluoride membranes (Merck KGaA). The membranes were blocked with 5% BSA at room temperature for 1 h. The corresponding primary antibodies, including vinculin (Bioss, bsm541448, 1:5000), Ras (CST, 8832, 1:1000), MEK (CST, 4694, 1:1000), p-MEK (CST, 9154, 1:1000), ERK (CST, 4695, 1:1000), and p-ERK (CST, 4370, 1:1000), were incubated with the samples overnight at 4 °C. The membranes were washed three times with Tris-buffered saline containing Tween-20 for 10 min and incubated with secondary antibodies for 2 h at room temperature. The protein bands were examined after incubation of the membranes in Immobilon Western Chemiluminescent reagent (Merck KGaA), and images were captured with a chemiluminescence imaging system (Tanon 5200, Shanghai, China).

RNA immunoprecipitation (RIP)

RIP assays were performed using a Magna RIP Kit (17-700, Millipore Corporation, USA) according to the manufacturer's instructions. Briefly, MDA-MB-231 cells were harvested and lysed with RIP lysis buffer. The magnetic beads were washed three times with RIP wash buffer, followed by incubation with 5 µg of anti-THBS1 (CST, 37879 S) or anti-IgG (CST, 2729) antibodies for 30 min at room temperature. The cell lysates were then incubated with magnetic beads at 4 °C overnight. The incubated samples were subsequently washed 6 times with RIP wash buffer. Finally, the immunoprecipitated RNAs were eluted and analyzed via RT-PCR.

Transwell assays

After transfection for 48 h, cells were seeded in 6.5-mm transwell chambers with 8-µm pores (3422, Corning Incorporated, New York, USA). The upper chambers were filled with 200 µL of a suspension of transfected cells (8×10^4 cells) in medium supplemented with 0.1% FBS, and the lower chambers were filled with 500 µL of complete medium. After 24 h of incubation, the nonmigratory cells that remained on the upper surface were removed with a cotton swab. The migrated cells on the lower surface of the membrane insert were fixed in 4% paraformaldehyde in PBS and stained with 0.1% crystal violet. The numbers of cells in the lower chambers were counted via light microscopy.

Luciferase activity analyses

The wild-type (AC112721.1-WT) and mutant (AC112721.1-MUT) fragments and the 3' untranslated region (UTR) of C2CD2L (C2CD2L-WT and C2CD2L-MUT, respectively) were cloned and inserted into the pmel and XbaI sites of the PmirGLO vector (Promega). Both constructs were verified by sequencing. 293T cells were cultured in 12-well plates and cotransfected with 50 nM control mimic or miR-491-5p mimic or with 2 µg of either the wild-type vector or the mutant vector using Lipofectamine 2000 (Invitrogen). At 48 h posttransfection, the relative luciferase activity was calculated by normalizing the firefly luminescence to the *Renilla* luminescence using the Dual-Luciferase Reporter Assay (Promega) according to the manufacturer's instructions. The values represent the means \pm standard deviations of the values obtained in three independent assays performed in three experiments.

Animal experiments

All protocols involving the use of animals were approved by the Animal Ethics Committee of Kunming University of Science and Technology (PZWH (dian) K2024-0001). Four-week-old female null mice were purchased from Hunan SJA Laboratory Animal Co., Ltd. Each animal was subcutaneously injected with a total of 1×10^6 MDA-MB-231 cells transfected with lentivirus (sh-NC or sh-AC112721.1). The tumor was measured every 2 days, and its volume was calculated via the following formula: $V = 0.5 \times \text{length} \times \text{width}^2$. After 4 weeks, the mice were sacrificed by dislocation, and the tumors were excised and weighed. All methods were carried out in accordance with relevant guidelines and regulations. All methods are reported in accordance with ARRIVE guidelines.

RNA sequencing

Total RNA was extracted from MDA-MB-231 cells transfected with si-AC112721.1 or si-NC. The integrity of the RNA was tested using an Agilent 2100 bioanalyzer, and RNA libraries were constructed using the NEBNext Ultra Directional RNA Library Prep Kit (Illumina, USA). Each library was sequenced on an Illumina NovaSeq 6000 platform according to the protocol recommended by Novogene Technology Co., Ltd. (Beijing, China).

The steps in differential expression gene (DEG) analysis were as follows: First, the raw gene expression data were preprocessed; preprocessing included data cleaning, normalization, and filtering. The samples were then grouped according to the experimental conditions used, and statistical methods were used to compare the levels of gene expression in the different groups. To minimize the false discovery rate caused by multiple hypothesis testing, the p values were usually corrected. Finally, for the identified DEGs, biological interpretation (gene function annotation, pathway analysis²⁹⁻³¹, gene ontology analysis, and analysis of other parameters) was conducted.

RNA pull-down assays

Biotinylated probes targeting AC112721.1 were synthesized by GuangZhou Lab Biotechnology Co., Ltd. MDA-MB-231 cells were lysed in RIP buffer, and the supernatant of the lysate was incubated with biotinylated probes at 4 °C overnight. Streptavidin-linked magnetic beads were added to the mixture, and the mixture was incubated at room temperature for 1 h. The captured proteins were eluted and analyzed by mass spectrometry.

Mass spectrometry (MS)

For mass spectrometry, the proteins were subjected to SDS-PAGE, and the resulting bands were cut and identified via mass spectrometry. Proteome Discoverer 1.4 (Version 1.4.0.288, Thermo Fisher) was used to convert the RAW files into MGF format MS files. The MGF format mass spectrometry files and the protein search library were input into ProteinPilot Software 4.5 (version 1656, AB Sciex) for mass spectrometry analysis. The built-in parameter settings in ProteinPilot Software 4.5 were as follows:

Item	Parameters
Detected protein threshold [unused ProtScore (Conf)] >	0.05 (10.0%)
Paragon algorithm	4.5.0.0, 1654
Annotations retrieved from uniProt	Yes
Sample type	Identification
Cys. Alkylation	MMTS
Digestion	Trypsin
Instrument	Orbi MS, Orbi MS/MS
ID Focus	Biological modifications
Search effort	Thorough
FDR analysis	Yes
User-modified parameter files	No

The search database used in this experiment was *Homo sapiens*: <https://www.uniprot.org/proteomes/UP000005640>. After the control group data were subtracted from the identified protein data of the experimental group, the remaining differentially expressed proteins were considered target binding proteins.

The exact sequence of AC112721.1 used for RNA pull-down is shown in Supplementary Table 3.

Prediction of AC112721.1 target miRNAs

First, the R language Corplot R package was used to analyze the correlations between the DEGs and AC112721.1; in this way, the miRNAs associated with HAGLROS were identified ($|R2| > 0.3$ & $p < 0.05$). Second, AC112721.1-related miRNAs were predicted via the cis-trans regulation method¹³. Third, using the cor.test function in R, we calculated the correlation between AC112721.1 and all related miRNAs via the Pearson correlation coefficient. In this way, the 6 miRNAs with the highest correlations with AC112721.1 were identified; they were hsa-miR-24-1-5p, hsa-miR-3926-1, hsa-miR-483-5p, hsa-miR-489-5p, hsa-miR-491-5p, and hsa-miR-6788-5p.

Potential targets of miR-491-5p

We used four online databases (TargetScan, RNA22, PolymiRTs and MiRBase) to predict the potential targets of miR-491-5p. The target genes we screened were simultaneously predicted via at least three databases.

Statistical analyses

The data are presented as the means \pm standard deviations (SDs). All the statistical analyses were conducted using SPSS version 20 software (IBM Corp., Armonk, USA), GraphPad Prism 7 (GraphPad Software, version 7.0, La Jolla, USA), and R language (version 3.3.1; <http://www.r-project.org/>; R Foundation). Values of $p < 0.05$ were considered to indicate statistical significance. All experiments were performed at least three times in triplicate.

Data availability

The original contributions presented in the study are included in the article/supplementary files. Further inquiries can be directed to the corresponding author.

Received: 25 July 2024; Accepted: 16 December 2024

Published online: 30 December 2024

References

- Sung, H. et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249. <https://doi.org/10.3322/caac.21660> (2021).
- Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2021. *CA Cancer J. Clin.* **71**, 7–33. <https://doi.org/10.3322/caac.21654> (2021).
- So, J. Y., Ohm, J., Lipkowitz, S. & Yang, L. Triple negative breast cancer (TNBC): non-genetic tumor heterogeneity and immune microenvironment: emerging treatment options. *Pharmacol. Ther.* **237**, 108253. <https://doi.org/10.1016/j.pharmthera.2022.108253> (2022).
- Anders, C. K. et al. The prognostic contribution of clinical breast cancer subtype, age, and race among patients with breast cancer brain metastases. *Cancer* **117**, 1602–1611. <https://doi.org/10.1002/cncr.25746> (2011).
- Li, Y. et al. Recent advances in therapeutic strategies for triple-negative breast cancer. *J. Hematol. Oncol.* **15**, 121. <https://doi.org/10.1186/s13045-022-01341-0> (2022).
- Yin, L., Duan, J. J., Bian, X. W. & Yu, S. C. Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res.* **22**, 61. <https://doi.org/10.1186/s13058-020-01296-5> (2020).
- Bridges, M. C., Daulagala, A. C. & Kourtidis, A. LNCcation: lncRNA localization and function. *J. Cell. Biol.* **220** <https://doi.org/10.1083/jcb.202009045> (2021).
- Skalska, L., Beltran-Nebot, M., Ule, J. & Jenner, R. G. Regulatory feedback from nascent RNA to chromatin and transcription. *Nat. Rev. Mol. Cell. Biol.* **18**, 331–337. <https://doi.org/10.1038/nrm.2017.12> (2017).

9. Herman, A. B., Tsitsipatis, D. & Gorospe, M. Integrated lncRNA function upon genomic and epigenomic regulation. *Mol. Cell.* **82**, 2252–2266. <https://doi.org/10.1016/j.molcel.2022.05.027> (2022).
10. Salmena, L., Poliseno, L., Tay, Y., Kats, L. & Pandolfi, P. P. A ceRNA hypothesis: the rosetta stone of a hidden RNA language? *Cell* **146**, 353–358. <https://doi.org/10.1016/j.cell.2011.07.014> (2011).
11. Singh, D. D., Lee, H. J. & Yadav, D. K. Recent clinical advances on long non-coding RNAs in triple-negative breast cancer. *Cells* **12** <https://doi.org/10.3390/cells12040674> (2023).
12. Liu, J. et al. JAG1 enhances angiogenesis in triple-negative breast cancer through promoting the secretion of exosomal lncRNA MALAT1. *Genes Dis.* **10**, 2167–2178. <https://doi.org/10.1016/j.gendis.2022.07.006> (2023).
13. Zheng, L. et al. Clinical significance and mechanism of long noncoding RNA HAGLROS in triple negative breast cancer. *Pathol. Res. Pract.* **231**, 153810. <https://doi.org/10.1016/j.prp.2022.153810> (2022).
14. Bi, M. et al. LncRNA LINC01234 promotes triple-negative breast cancer progression through regulating the miR-429/SYNY1 axis. *Am. J. Transl. Res.* **13**, 11399–11412 (2021).
15. Morrison, L. & Okines, A. Systemic therapy for metastatic triple negative breast cancer: current treatments and future directions. *Cancers (Basel)* **15** <https://doi.org/10.3390/cancers15153801> (2023).
16. Yip, C. W., Sivaraman, D. M., Prabhu, A. V. & Shin, J. W. Functional annotation of lncRNA in high-throughput screening. *Essays Biochem.* **65**, 761–773. <https://doi.org/10.1042/EBC20200061> (2021).
17. Wang, L. et al. Long noncoding RNA (lncRNA)-Mediated competing endogenous RNA networks provide novel potential biomarkers and therapeutic targets for colorectal cancer. *Int. J. Mol. Sci.* **20** <https://doi.org/10.3390/ijms20225758> (2019).
18. McCabe, E. M. & Rasmussen, T. P. lncRNA involvement in cancer stem cell function and epithelial-mesenchymal transitions. *Semin. Cancer Biol.* **75**, 38–48. <https://doi.org/10.1016/j.semcancer.2020.12.012> (2021).
19. Banerjee, M. & Rajeswari, V. D. Critical review on the different roles of exosomes in TNBC and exosomal-mediated delivery of microRNA/siRNA/lncRNA and drug targeting signalling pathways in triple-negative breast cancer. *Molecules* **28** <https://doi.org/10.3390/molecules28041802> (2023).
20. Zhang, W., Guan, X. & Tang, J. The long non-coding RNA landscape in triple-negative breast cancer. *Cell. Prolif.* **54**, e12966. <https://doi.org/10.1111/cpr.12966> (2021).
21. Davis, P. J., Sudha, T., Lin, H. Y. & Mousa, S. A. Thyroid hormone, hormone analogs, and angiogenesis. *Compr. Physiol.* **6**, 353–362. <https://doi.org/10.1002/cphy.c150011> (2015).
22. Kale, A., Rogers, N. M. & Ghimire, K. Thrombospondin-1 CD47 signalling: from mechanisms to medicine. *Int. J. Mol. Sci.* **22** <https://doi.org/10.3390/ijms22084062> (2021).
23. Zhang, X., Huang, T., Li, Y. & Qiu, H. Upregulation of THBS1 is related to immunity and chemotherapy resistance in gastric cancer. *Int. J. Gen. Med.* **14**, 4945–4957. <https://doi.org/10.2147/IJGM.S329208> (2021).
24. Kaur, S. et al. Functions of thrombospondin-1 in the tumor microenvironment. *Int. J. Mol. Sci.* **22** <https://doi.org/10.3390/ijms22094570> (2021).
25. Folgueira, M. A. et al. Markers of breast cancer stromal fibroblasts in the primary tumour site associated with lymph node metastasis: a systematic review including our case series. *Biosci. Rep.* **33** <https://doi.org/10.1042/BSR20130060> (2013).
26. Xie, B. et al. The endoplasmic reticulum-plasma membrane tethering protein TMEM24 is a regulator of cellular Ca²⁺ homeostasis. *J. Cell. Sci.* **135** <https://doi.org/10.1242/jcs.259073> (2022).
27. Sun, E. W. et al. Lipid transporter TMEM24/C2CD2L is a ca(2+)-regulated component of ER-plasma membrane contacts in mammalian neurons. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 5775–5784. <https://doi.org/10.1073/pnas.1820156116> (2019).
28. Lees, J. A. et al. Lipid transport by TMEM24 at ER-plasma membrane contacts regulates pulsatile insulin secretion. *Science* **355** <https://doi.org/10.1126/science.aah6171> (2017).
29. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951. <https://doi.org/10.1002/pro.3715> (2019).
30. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51**, D587–D592. <https://doi.org/10.1093/nar/gkac963> (2023).
31. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30. <https://doi.org/10.1093/nar/28.1.27> (2000).

Author contributions

PM, WT and MS contributed to the design and conception of the study. PM, SK, HL and ML conducted experiments. YZ collected the clinical samples. HY and JP performed the statistical analysis. PM, HY and MS wrote the first draft of the manuscript. All authors read and approved the submitted version. All authors reviewed the manuscript.

Funding

This work was supported by the Yunnan High-level Personnel Training Support Program (YNWR-QN-BJ-2020-243), the Kunming University of Science and Technology & Puer People's Hospital Joint Special Project on Medical Research (grant number KUST-PE2022004Y) and the Yunnan Province Science and Technology Program, China (grant number 202401AT070372).

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

Twenty breast cancer samples and 20 paracancerous tissues were collected from 30 patients who underwent surgery at the First People's Hospital of Yunnan Province. The protocol was approved by the Ethical Committee of First People's Hospital of Yunnan Province (No. KMUST-MEC-016). Informed consent was obtained from all the human subjects. The animal experiments were approved by the Animal Ethics Committee of Kunming University of Science and Technology (PZWH (dian) K2024-0001). Research on human subjects adheres to the Declaration of Helsinki principles.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-83665-0>

[0.1038/s41598-024-83665-0](https://doi.org/10.1038/s41598-024-83665-0).

Correspondence and requests for materials should be addressed to W.T. or M.S.

Reprints and permissions information is available at www.nature.com/reprints.

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

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

© The Author(s) 2024