

# NSrp70 suppresses metastasis in triple-negative breast cancer by modulating Numb/T $\beta$ R1/EMT axis

**Yang Zhao**

Fudan University Shanghai Cancer Center

**Hefen Sun**

Department of Breast Surgery, Fudan University Shanghai Cancer Center

**Yuanyuan Zhao**

Fudan University Shanghai Cancer Center

**Qiqi Liu**

Shandong Provincial QianFoShan Hospital

**Yang Liu**

Fudan University Shanghai Cancer Center

**Yifeng Hou**

Fudan University Shanghai Cancer Center

**Wei Jin** (✉ [weijin7207@126.com](mailto:weijin7207@126.com))

Department of Breast Surgery, Fudan University Shanghai Cancer Center

---

## Article

**Keywords:** NSrp70, alternative splicing, breast cancer, metastasis, TGF $\beta$  signaling

**Posted Date:** March 18th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-311222/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**NSrp70 suppresses metastasis in  
triple-negative breast cancer by modulating Numb/TβR1/EMT axis**

Yang Zhao<sup>1,2#</sup>, Hefen Sun<sup>1,2#</sup>, Yuanyuan Zhao<sup>1,2</sup>, Qiqi Liu<sup>1,2</sup>, Yang Liu<sup>1</sup>, Yifeng Hou<sup>1,2</sup>,  
Wei Jin<sup>1,2,\*</sup>

<sup>1</sup> Department of Breast Surgery, Key Laboratory of Breast Cancer in Shanghai, Fudan University Shanghai Cancer Center, Shanghai 200032, China

<sup>2</sup> Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China

\*Correspondence: Dr. Wei Jin, Department of Breast Surgery, Key Laboratory of Breast Cancer in Shanghai, Fudan University Shanghai Cancer Center, Shanghai, 200032, China. Tel: +86-21-64175590-3423; Fax: +86-21-64031696; E-mail: [jinwei7207@163.com](mailto:jinwei7207@163.com)

#These authors have contributed equally to this work.

## ABSTRACT

Metastasis is a major cause of death in individuals suffering from triple-negative breast cancer. Alternative splicing of mRNA precursor allows cancer cells to create different protein isoforms which may promote metastasis. Quantitative proteomic analysis of primary and metastatic breast cancer cells revealed that nuclear speckle-related protein 70 (NSrp70) was significantly downregulated in highly metastatic cells. Downregulation of NSrp70 promoted the migration and invasion of breast cancer cells *in vitro* and *in vivo*. Mechanistically, we found that NSrp70 inhibited the skipped exon alternative splicing of NUMB,

promoted the degradation of TGF receptor 1(T $\beta$ R1) through lysosome pathway, and regulated TGF $\beta$ /SMAD-mediated epithelial-mesenchymal transition (EMT) phenotype in breast cancer cells. Furthermore, high NSrp70 expression correlated with better prognosis in breast cancer patients. Our findings revealed that splicing regulator NSrp70 may serve as a metastasis suppressor.

**Keywords:** NSrp70, alternative splicing, breast cancer, metastasis, TGF $\beta$  signaling

## INTRODUCTION

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer-related death among women worldwide[1]. Breast cancer metastasis has become the leading cause of breast cancer death. Triple-negative breast cancer (TNBC) is a highly invasive and aggressive cancer subtype characterized by the lack of expression of the estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (Her-2) and the percentage of TNBC is approximately 15-20% of all breast cancer patients. No prognostic biomarkers or therapeutic targets have been identified for the prognostic or treatment for TNBC. Thus, understanding the molecular mechanisms underpinning TNBC cancer metastasis is crucial for the development of effective strategies for treating this deadly disease.

Alternative splicing is critical at the post-transcriptional stage of gene expression, during which exons from a single gene are assembled in different ways to produce several protein isoforms in eukaryotic organisms[2]. Over 90% of human genes produce transcripts that are alternatively spliced, and 60% of splice variants encode distinct protein isoforms[3]. Aberrant splicing is common in cancer[4]. Cancer cells often benefit from this flexibility by producing proteins that promote tumor survival[5]. Dysregulation of alternative splicing is widely considered a new hallmark of cancer and alternative splicing products are acknowledged as potentially useful biomarkers[6]. TNBC is a heterogeneous disease driven by a large repertoire of molecular abnormalities[7]. Additional molecular signatures of TNBC need to be identified for improved diagnosis and treatment. Since cell-specific splicing patterns exist in different cell types, we hypothesized that TNBC may also exhibit characteristic splicing signatures that could be exploited for the development of new strategies for TNBC treatment.

Nsrp70 was first identified as a coiled-coil domain containing 55 (CCDC55) protein. Nuclear speckles are located in the inter-chromatin regions of the mammalian cell nucleoplasm and are the storage sites of mRNA splicing regulators, including

protein splicing factors such as small nuclear ribonucleoproteins, spliceosome subunits, and other non-small nuclear ribonucleoprotein. Therefore, these structures are involved in an elaborate regulation of gene expression[8]. NSrp70 is an important splicing factor containing an N-terminal RNA recognition motif, a C-terminal arginine/serine (RS)-like region, and two coiled-coil domains at each terminus[9]. Based on this structure, NSrp70 belongs to the SR protein family and plays an important role in alternative splicing[9]. To date, functional studies into NSrp70 have focused on schizophrenia[10], embryonic development[11], and adult acute leukemia[12]. However, the function of NSrp70 in carcinoma remains unknown.

We utilized the isobaric tags for relative and absolute quantitation (iTRAQ) technology followed by nanoscale high-performance liquid chromatography (nano-HPLC) tandem mass spectrometry (MS/MS) to analyze MDA-MB-231HM (higher lung-metastatic potential cells), MDA-MB-231BO (higher bone-metastatic potential cells), and MDA-MB-231 cells (parent cell). Because of their similar genetic background, these cells provide a unique model for identifying candidate metastasis-associated biomarkers and potential therapeutic targets for TNBC. We identified the candidate gene NSrp70 in the nuclear component. NSrp70 was downregulated in the highly metastatic breast cell line, suggesting that it is associated with metastasis of TNBC. These observations indicate that NSrp70 is a suppressor of TNBC metastasis.

## RESULTS

### **NSrp70 expression is downregulated in highly metastatic breast cancer cells**

Using the iTRAQ-labeling method in a previously described model system[13], we analyzed the differential expression nuclear located proteins in human breast cancer cell line MDA-MB-231, and its highly metastatic sublines MDA-MB-231HM and MDA-MB-231BO (Fig. 1a). We focused on the candidate protein NSrp70, which was significantly downregulated in MDA-MB-231HM and MDA-MB-231BO cells compared with the parental MDA-MB-231 cells (Fig. 1b, and Fig. S1a and S1b). The protein and mRNA-level expression of NSrp70 was also evaluated in a panel of breast cancer cell lines. In agreement with the iTRAQ data, NSrp70 expression were generally detected in the luminal breast cancer cells (MCF-7, ZR-75-30, and T47D), which are considered weakly metastatic cell lines, but it was lower expression in basal-like breast cancer cells (BT549, Hs-578T). In particular, in the highly metastatic cell lines (MDA-MB-231HM and MDA-MB-231BO), NSrp70 expression was weaker (Fig. 1c and d). In addition, Oncomine expression analysis revealed that NSrp70 was downregulated in different subtypes of breast cancer when compared with normal breast tissue and lower NSrp70 was correlated with higher recurrence and metastasis risk in breast cancer patient (Fig. 1e-h). Further, online Kaplan-Meier plot analysis indicated that downregulation of NSrp70 was associated with poor prognosis (Fig. 1i-k). Taken together, these observations suggested that NSrp70 might be a tumor suppressor in breast cancer.

### **Loss of NSrp70 promotes breast cancer cell metastasis *in vitro* and *in vivo***

To further investigate the role of NSrp70 in tumor migration and invasion, we

constructed stable NSrp70 knockdown and overexpressing cell lines. Western blotting analysis demonstrated that NSrp70 was successfully overexpressed in Hs-578T and MDA-MB-231HM cell lines. Further, two guide RNAs targeting NSrp70 significantly reduced the expression of NSrp70 in MDA-MB-231, Hs-578T, and BT-549 cell lines (Fig. 2a and b). We then evaluated the role of NSrp70 in the malignant phenotype of the breast cancer cells *in vitro*. The analysis indicated that NSrp70 knockdown in MDA-MB-231 and BT-549 cells significantly promoted cell migration and invasion. In addition, overexpression of NSrp70 in Hs-578T and MDA-MB-231HM cells significantly decreased cell migration and invasion ability (Fig. 2c and d). Further, NSrp70 knockdown in MDA-MB-231 and BT-549 cells significantly promoted the cell healing ability (Fig. 2e-g). These observations suggested that NSrp70 overexpression could suppress breast cancer cell migration and invasion *in vitro*.

To explore the ability of NSrp70 to influence cell migration *in vivo*, we labeled NSrp70 knockdown MDA-MB-231HM cells and NSrp70 stably overexpressing MDA-MB-231 cells with a retroviral construct expressing a green fluorescent protein (GFP)/luciferase fusion protein[14]. The cells were then delivered intravenously or via the mammary fat pad injection into the nude or NOD/scid mice. Tumor cell colonization and outgrowth in the lungs were monitored by noninvasive bioluminescent imaging (BLI) 6-8 weeks after treatment. We observed that low expression of NSrp70 in MDA-MB-231 cells significantly increased lung metastasis *in vivo*. By contrast, overexpression of NSrp70 in MDA-MB-231HM cells decreased lung metastasis *in vivo* (Fig. 2h, i). These observations indicated that low expression of NSrp70 promotes breast cancer cell metastasis both *in vitro* and *in vivo*.

In addition, we assessed the role of NSrp70 in breast cancer cell proliferation in MDA-MB-231HM and MDA-MB-231 cells. Neither upregulation nor downregulation affected the cell proliferation ability (Fig. S2a, b). This indicated that NSrp70 has no significant effect on breast cancer cell proliferation.

### **NSrp70 inhibits skipped exon (SE) alternative splicing of target pre-mRNA**

To determine the mechanism of NSrp70-mediated metastasis, we performed RNA-sequencing in NSrp70-overexpressing MDA-MB-231HM and Hs-578T cells, combined with bioinformatics alternative splicing analysis to survey the potential target genes of NSrp70. The analysis revealed that NSrp70 modulated alternative splicing via SE, alternative 3'-splice sites (A3SS), alternative 5'-splice sites (A5SS), intron retention (RI), and mutually exclusive exons (MXE) (Fig. 3a and Fig. S 3).

SE was the most common alternative splicing subtype in the two NSrp70-overexpressing cell lines analyzed (Fig. 3a and Fig. S3). The scatter plots for the other four subtypes are shown in Fig. S4. We took the intersection of different subtypes of alternative splicing events of the two NSrp70-overexpressing cell lines. The SE subtype intersection contained 13,462 splicing events, which was the largest number among the five subtypes analyzed (Fig. S5). After filtering the data by Bayes factor of  $\geq 10$  events, 10 potential genes were identified. Real-time polymerase chain reaction (Real Time-PCR) analysis revealed that splicing variants of NUMBv (exon 12 skipped) were

enriched after NSrp70 knockdown (Fig. 3b).

To further explore whether NSrp70 interacted with its target mRNA, we used an RNA immunoprecipitation (RIP) assay to analyze MDA-MB-231 cells transfected with constructs encoding NSrp70 wild type (NSrp70-WT), and its truncated variants NSrp70-M15 and NSrp70-RS1M. Pre-mRNAs of *NUMB* and *CD44* were efficiently immunoprecipitated from NSrp70-WT cells, but not from NSrp70-M15 and NSrp70-RS1M cells (Fig. 3c). We also performed RNA pull-down assays to confirm the physical binding of the target mRNA and NSrp70 protein (Fig. 3d). These observations suggested that NSrp70 provides a binding site for target mRNA.

By performing IP-mass, we found that NSrp70 interacts with SRSF1 (Fig. 3d, e). To explore the roles of these proteins in the inclusion or exclusion of target mRNA exons, we tested the effects of NSrp70 and SRSF1 on the *NUMB* exon 12 minigene. NSrp70 increased *NUMB* exon 12 inclusion in a transfection concentration-dependent manner (Fig. 3f). To further investigate whether SRSF1 affected the alternative site selection on the *NUMB* exon 12, MDA-MB-231HM cells were transfected with increasing or decreasing amounts of vectors encoding NSrp70 and SRSF1. PCR analysis revealed that SRSF1 increased *NUMB* exon 12 exclusion in a concentration-dependent manner, an effect that was opposite to that of NSrp70 (Fig. 3g).

### **NSrp70 modulates NUMB alternative splicing to promotes transforming growth factor $\beta$ -induced epithelial-mesenchymal transition**

The above experiments indicated that NSrp70 regulated alternative splicing of *NUMB*. To identify the mechanism underlying NSrp70-mediated metastasis, we performed KEGG pathway analysis of the RNA-sequence data. The analysis revealed enrichment of several pathways, such as gap junction, TGF $\beta$  signaling, cell adhesion molecule, and Ras signaling pathways (Fig. 4a). We next used western blotting to evaluate whether these pathways were related to NSrp70. Among these pathways, TGF $\beta$  signaling pathway was significantly activated. Depletion of NSrp70 elevated the expression of T $\beta$ R1 and phosphorylation of Smad3 (p-Smad3) in MDA-MB-231 cells. By contrast, overexpression of NSrp70 in MDA-MB-231HM cells downregulated TGF $\beta$  receptor type 1 (T $\beta$ R1) and p-Smad3 protein levels. However, the expression of T $\beta$ R2 and Smad3 was not affected by the NSrp70 status (Fig. 4b).

TGF $\beta$  stimulates the EMT in breast cancer cells[15]. To investigate whether NSrp70 mediated cell EMT by affecting the TGF $\beta$  signal pathway, we evaluated EMT marker expression in MDA-MB-231, BT549 and Hs-578T cells with downregulated NSrp70 expression, and in MDA-MB-231HM and Hs-578T cells with overexpressed NSrp70 (Fig. 4b-e). Western blotting analysis revealed that NSrp70 deficiency resulted in low E-cadherin expression, and high N-cadherin and fibronectin expression. By contrast, overexpression of NSrp70 in MDA-MB-231HM cells had an adverse effect (Fig. 4b). In addition, TGF $\beta$  or T $\beta$ R1/2 inhibitor (GW788388) treatment of NSrp70-overexpressing or -knockout cells confirmed the effect of NSrp70 on TGF $\beta$  signal pathway (Fig. 4c-e). Further, GW788388 treatment weakened the promotion of metastasis of Hs-578T and MDA-MB-231 cells elicited by downregulation of NSrp70 (Fig. 4f). Further, immunofluorescence analysis of NSrp70 downregulated MDA-MB-

231 cells confirmed the trends in the expression of key proteins. The expression of E-cadherin and N-cadherin was downregulated after treatment with pathway inhibitor (Fig. S6).

To further explore whether NSrp70 affected the TGF $\beta$  signaling by impacting splicing of NUMB, we constructed NUMBwt and NUMBv stable overexpression cell lines, and the NUMBv referred to variants generated specific after alternative splicing. The transwell assay indicated that NUMBv rather than NUMBwt promoted breast cancer cell metastasis (Fig4. g). We then evaluated expression of TGF $\beta$  signaling and EMT-related proteins in the cells. When NSrp70 was co-overexpressed with NUMBv, the expression of target proteins was rescued, in addition to activating TGF $\beta$  signaling and EMT, which were inhibited in NSrp70-overexpressing cells. (Fig. 4h). NUMB variants rescued the metastatic ability of breast cancer cells that had been reduced by NSrp70 (Fig. 4g).

### **NSrp70 accelerates T $\beta$ R1 degradation by NUMB mediated lysosome pathway**

We observed a significant increase in T $\beta$ R1 protein levels following NSrp70 knockdown (Fig. 4b). However, the alternation of T $\beta$ R1 mRNA levels was modest and not significant (Fig. S7). These observations suggested that NSrp70 predominantly suppresses T $\beta$ R1 expression on a post-transcriptional level.

To check whether NSrp70 was associated with the T $\beta$ R1 protein degradation pathway, MDA-MB-231HM cells were incubated with cycloheximide (CHX). T $\beta$ R1 was degraded more rapidly and became less detectable within 6 h of CHX treatment in the NSrp70-overexpressing cells than in the control cell line (Fig. 5a). In addition, treatment of CHX in MDA-MB-231 cells restrained T $\beta$ R1 degradation after the overexpression of NUMBv (Fig5.b).

It has been reported that NUMB is cargo-selective endocytic adaptors and may regulate lysosome maturation[16] and recent studies showed that NUMB is a crucial regulator of lysosomal function and autophagic progression[17, 18]. To identify whether T $\beta$ R1 is a new cargo of NUMB, we performed IP assay. The results showed that both NUMBv and NUMBwt interaction with T $\beta$ R1(Fig. 5c).

Interestingly, T $\beta$ R1 degradation was inhibited when the cells were treated with the lysosome pathway inhibitor NH<sub>4</sub>Cl in both NSrp70 and NUMBv overexpression cells (Fig. 5d, e). However, the NUMBwt did not show significant effect under the treatment of other lysosome pathway inhibitor bfa1 (Fig. 5f). We evaluated the co-localization of T $\beta$ R1 with early endosome antigen 1 (EEA1), member RAS oncogene family 7 (RAB7), and RAB11 utilizing immunofluorescence. We observed that the co-localization of T $\beta$ R1 and EEA1, and RAB7 was more pronounced in NSrp70 knockdown cell lines. Meanwhile, the co-localization of T $\beta$ R1 and RAB11 was not greatly affected (Fig. 5g). Collectively, these findings suggested that depleting NSrp70 potentiates TGF $\beta$ /SMAD signaling by inhibiting the T $\beta$ R1 degradation induced by lysosome pathway.

### **Low NSrp70 Expression Correlates with Poor Patient Prognosis**

To determine the clinical relevance of the above findings in advanced human cancers, we first analyzed the data in the Oncomine database and found that NSrp70 was

downregulated in breast cancer (Fig. 1e). To evaluate the clinical importance of NSrp70 in breast cancer, we performed an immunohistochemical analysis of tissue microarrays (TMAs) containing samples from 250 breast cancer patients. Follow-up data were missing for two patients; hence, 248 cases were included in the subsequent analysis. The clinicopathological characteristics of the cohort are summarized in Table 1. Because NSrp70 expression varied across the breast tumor samples, we scored the NSrp70 expression and categorized it into low intensity and high intensity groups, according to the semi-quantitative evaluation, as described in the Supplementary Methods (Fig. 6a). The Kaplan-Meier survival curve analysis revealed that low NSrp70 expression correlated with decreased survival in the whole data set and patients with TNBC tumors. On the other hand, in patients with non-TNBC tumors, NSrp70 expression did not correlate with disease-free survival probability (Fig. 6b). This indicated that low NSrp70 expression served as an independent predictor of poor disease-free survival in TNBC patients. Furthermore, univariate analysis indicated that low NSrp70 expression at diagnosis was associated with an increased risk of disease relapse [hazard ratio (HR) = 3.978, 95% confidence interval (CI) 1.983-7.977;  $p < 0.001$ ]. Further multivariate Cox analysis revealed a similar trend as that indicated by the univariate analysis (Table 2).

## DISCUSSION

NSrp70 is a newly discovered protein mainly localized in the nuclear speckle[9]. Nuclear speckles are subnuclear organelles that serve as storage sites for splicing components[19]. NSrp70 belongs to the SR-related protein family and is recognized as an important splicing factor[20, 21].

In the current study, we showed that NSrp70, a splicing factor, inhibits the alternative splicing of target mRNA. We identified CD44 and NUMB as its target genes. Studies of different CD44 variants have gradually progressed in recent years. CD44 is a cell surface protein that modulates cellular signaling by forming coreceptor complexes with various receptor tyrosine kinases[22, 23]. Expression of CD44 is an excellent example of alternative splicing, further study identified different CD44 variants and CD44 exons v5 and v6 have been investigated extensively[24, 25]. It has been reported that NSrp70 modulated CD44 alternative splicing by forming CD44v5[21]. CD44v6 is an independent prognosis factor suggesting poor outcomes of colorectal cancer[26]. In the current study, we identified the variant CD44v10, generated by exon 6-13 skipping and found it may be according to improving metastatic proficiency (Fig. S8). Indeed, CD44v10 is closely associated with the niche embedment and cytotoxicity of leukemia hematopoietic stem cells[27]. Further, CD44v10 is associated with the histological grade and survival of renal cell carcinoma[28]. These observations are in agreement with those of the current study.

Previous studies indicated that NUMB was an adaptor protein inhibiting Notch signaling by leading to Notch degradation through ubiquitylation and lysosomal degradation[29-31]. And a recent research showed NUMB positively regulates autophagic flux by regulating lysosomal function[18]. In addition, the alternative splicing of NUMB was also found to be regulated by activated Notch signaling pathway,

which affects the proliferation of the cervical cancer cells[31]. Our study showed that NUMB was a new target gene of NSrp70 and NSrp70 modulated the skip of NUMB exon12. This variant was decreased in tumor cell, which was contribute to the viability of cervical carcinoma cells. Besides, the long isoform containing exon 12 was suggested to promote cell proliferation during tumorigenesis[32], whereas the short isoform inhibits cell growth[33]. Further, NUMB was recently shown to be a new player in EMT. Knockdown of NUMB reduces the cell-cell adhesion, resulting in cell metastasis and proliferation[34]. In the current study, we showed that NUMBv inhibit the degradation of T $\beta$ R1 via the lysosome pathway and promote TGF $\beta$ -induced EMT.

According to recent studies, in addition to alternative splicing affecting the activation of the TGF $\beta$  signaling pathway discussed in the current study, the TGF $\beta$  signaling pathway can itself adversely affect the occurrence of alternative splicing[35]. TGF $\beta$  is a major inducer of EMT, which is associated with cancer cell metastasis, and resistance to chemotherapy and targeted drugs, through both transcriptional and non-transcriptional mechanisms. We here found that in tumor cells, an increase in the mitotic signaling pathway prompted TGF $\beta$  to activate Smad3 and bind to PCBP1, thereby regulating EMT, cytoskeletal rearrangement, and selective splicing of cell adhesion-associated proteins.

The RS region of NSrp70 contains the binding sites of both SR proteins and target mRNA. Depletion of the RS region could severely affect the location of splicing factors and target mRNA in the nuclear speckle. Further, the RS region of NSrp70 may play important roles in the formation of protein framework and the assembly of related proteins[36]. The findings of the current study indicated antagonistic regulation of splicing by NSrp70 and SRSF1. SRSF1 promotes splicing of downstream genes, such as *CD44*, by interfering with the inhibition of heterogeneous nuclear ribonucleoprotein (hnRNP)[37, 38]. Here, we showed that NSrp70 rescues the splicing-promoting effect of SRSF1, and that NSrp70 and SRSF1 both affect splicing in a concentration-dependent manner. In addition to SRSF1, NSrp70 works in tandem with SRSF1 and SRSF2[20, 21]. SR proteins, such as SRSF1-4, generally make effect in the form of compounds[39-41].

Alternative splicing has been the main source of protein diversity in human genes and it has been considered a hallmark of cancer. Previous studies indicated that 27 types of cancers were associated with more than 15000 cancer-specific splice variants[42]. Alternative splicing has been reported to influence breast cancer by participating the regulation of therapeutic resistance[43]. Therefore, develop therapeutic approaches to specific splicing is promising. Currently, several tools have been researched to target with splice site, splicing regulatory elements or prevent the binding of transacting splicing factor such as antisense oligonucleotides (ASOs) and splicing-switching antisense oligonucleotides (SSOs) [44, 45]. Our study showed that NSrp70 inhibits the SE splicing of the downstream genes *NUMB* and suppresses TNBC metastasis. Therefore, modulating the splicing may present a new approach for cancer therapy.

Taken together, our study identified NSrp70 as a novel tumor metastatic suppressor. Mechanistically, NSrp70 inhibits the SE splicing of the downstream genes *NUMB*, promoting the T $\beta$ R1 degradation via lysosome pathways (Fig. 6c). This inhibits EMT

mediated by the TGF $\beta$  signaling, ultimately inhibiting the metastasis of breast cancer cells. Further, the presented data indicate that NSrp70 plays a key role in TNBC invasion and metastasis, providing insights into the mechanisms and strategies for the therapeutic intervention in the TGF $\beta$  pathway and cancer metastasis.

## **METHODS**

### **Ethics Statement**

The present study was conducted according to the documented standards of the Ethics Committee of FDUSCC. Further, the authors' research group members attended educational and training courses on the appropriate care and use of experimental animals.

### **Cell Culture and Transfection**

The cell lines were obtained from the Shanghai Cell Bank, Type Culture Collection Committee of Chinese Academy of Science (Shanghai, China). The highly lung metastatic cell line MDA-MB-231HM was developed by four cycles of tail vein injections and we have a current patent application for this cell line (patent number:200910174455.4). The other high metastatic cell line used in our research was MDA-MB-231BO, which has higher bone and lung metastatic characteristics. This cell line was gift from Dr. Toshiyuki Yoneda (The University of Texas, Houston, TX). All cell lines were maintained under conditions specified by the provider, and were cultured in a 5% CO<sub>2</sub> incubator at 37°C. Constructs were transfected by Polybrene (Sigma-Aldrich, St. Louis, MO) according to the instructions of the manufacturer instructions. 48h after transfection, cells were collected for RNA and protein or image analysis.

### **Study Population**

We used 250 primary breast cancer samples from patients with stages I to III invasive ductal carcinoma collected randomly at the Department of Breast Surgery in Fudan University Shanghai Cancer Center (FDUSCC, Shanghai, P.R. China) between August 2002 and March 2006. All cases had complete pathological diagnosis and follow-up records, and the histological grade was I to III. The clinical data included age, menstrual status, histological grade, tumor size, lymph node status, ER, PR, HER2 status, recurrence or metastasis time, and total survival time. The deadline for follow-up information was August 2013, with a median follow-up of 98 months. The patients were treated according to the standards used during surgery and those who were not found to be fit for surgery received adjuvant chemotherapy with different regimens for four to six cycles and/or hormone therapy (if required), according to the standard therapy during surgery. This study was approved by the Ethics Committee of FDUSCC, and each participant provided signed informed consent.

### **IHC Analysis**

Tissue microarrays (TMAs) were obtained from archived formalin-fixed, paraffin wax-embedded carcinoma samples from patients in FDUSCC. IHC data were analyzed as

previously described[46]. Patient information is reported in a previous paper from the authors' laboratory[13]. The IHC variables were scored as described elsewhere[47]. The details of TMAs construction, IHC staining, and analysis of the IHC variables are presented in the Supplementary Methods.

### **Quantitative Real-time PCR**

The total RNA was extracted using TRIzol reagent (Invitrogen Corporation) and reverse transcribed using the PrimeScript RT Reagent Kit (TaKaRa Biotechnology). The real-time polymerase chain reaction (PCR) was performed with SYBR Premix Ex Taq (TaKaRa Biotechnology) using ABI Prism 7900 (Applied Biosystems). The sequence of primers used in this study is listed in Supplementary Table.

### **Plasmids and Guide RNAs for CRISPR/Cas9**

Human NSrp70 cDNA was subcloned from the breast cancer cell line MDA-MB-231 into the pCDH-CMV-MCS-EF1-Puro lentiviral vector. The following primers were used:

Forward primer: 5'-CCGgaattc GCCACC ATG GAC TAC AAG GAC GAT GAT GAC AAG CTC GAT GGA GGA ATGGCGATTCCGGGCAGGCA-3'

Reverse primer: 5'-CGCggatccTCAATCATCTTCTTTCTCAA-3'.

NSrp70 guide RNAs and the negative control were designed considering the CRISPR/Cas9 system[48]. The following primers were used:

Crispr#1: Forward primer: 5'-CACCGAATGCAGCTTTCGTGAAGCC-3'

Reverse primer: 5'-AAACGGCTTCACGAAAGCTGCATTC-3'

Crispr#2: Forward primer: 5'-CACCGTAAGAAGCAGGCCATGAAAC-3'

Reverse primer: 5'-AAACTAAGAAGCAGGCCATGAAACC-3'.

Similar to NSrp70 cloning, human SRSF1 cDNA was subcloned from the breast cancer cell line MDA-MB-231 into the pCDH-CMV-MCS-EF1-Puro lentiviral vector with the help of EcoRI and BamHI (NEW ENGLAND BioLabs). The following primers were used:

Forward primer: 5'-TGCTCTAGAATGTCCGGGAGGTGGTGTGAT-3'

Reverse primer: 5'-CGGGATCCTTATGTACGAGAGCGAGATC-3'.

NUMB variants (NM\_001005744.2) and wild type (NM\_001005743.2) synthesis was purchased from GENEWIS and then subcloned to pCDH (NUMBv-pCDH and NUMBwt-pCDH).

### **Lentivirus Packaging and Infection**

Briefly, 293T cells were co-transfected with the lentiviral vectors pCDH (or pGIPZ) and the packaging vectors psPAX2 and pMD2G. After 48 h, viral culture supernatants were collected, filtered, and added to the cells. The cells were incubated with virus and polybrene (Sigma-Aldrich) added at a working concentration of 8 µg/mL for 12 h, and then medium containing FBS was added. After 24 h, the infected cells were subjected to selection with 2 µg/mL puromycin for one week.

### **Western Blot Analysis**

Whole-cell lysates were obtained using Pierce T-PER (Tissue Protein Extraction Reagent; Thermo Fisher Scientific Inc.) with protease inhibitor cocktail tablets and phosphatase inhibitors (Roche). The concentration of proteins was determined using the BCA Protein Assay Kit (Solarbio). In total, 30 mg of cell lysates was resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred on to polyvinylidene fluoride membranes (Millipore). The membranes were blocked with 5% milk or 5% BSA, and then incubated with primary antibodies, followed by HRP-conjugated secondary antibodies. Immunoreactive bands were identified using chemiluminescent substance, according to the manufacturer's instructions, and quantified by densitometry. The antibodies used in this study are listed in Supplementary Table S2.

### **Transwell Assays and Kinetic Wound-healing Assay**

Cells ( $5 \times 10^4$  cells/well for the migration assay and  $1 \times 10^5$  cells/well for the invasion assay) were plated in the top chambers of noncoated membranes or Matrigel-coated Transwell chambers (BD Biosciences) in 200  $\mu$ L of medium without FBS. Medium supplemented with 20% serum was used as a chemoattractant in the bottom chamber. After incubation for 15-20 h, the cells were fixed with 40% methanol and stained with 0.25% crystal violet. The cells that did not migrate through the pores were removed and the migrating cells were imaged and counted.

Cells were plated on 96-well plates (Essen ImageLock, Essen Instruments) at a concentration of  $3.6 \times 10^4/100$   $\mu$ L of medium, and a wound was scratched with a wound scratch maker (Essen Instruments). Wound confluence was monitored with a Live-Cell Imaging System and software (Essen Instruments) every 4 h for 48 h by comparing the mean relative wound density of at least three biological replicates for each experiment.

### **Immunofluorescence**

MDA-MB-231 stable cell lines grown on coverslips were fixed with 4% paraformaldehyde for 30 minutes at room temperature, permeabilized with 0.5% Triton X-100 for 5 minutes at 4°C, and incubated with primary antibodies overnight at 4°C. The slides were then incubated with Alexa 594-conjugated (red, Invitrogen) secondary antibodies for 30 min at room temperature, and then incubated with 4',6-diamidino-2-phenylindole (DAPI) for 10 min. Images were captured with a confocal laser microscope (Leica TCS SP5 II). At least 100 cells were analyzed per group.

### **Immunoprecipitation**

Stable cell lines ( $10^7$  cells) were harvested and suspended in immunoprecipitation buffer [50 mM HEPES (pH 7.6), 250 mM NaCl, 5 mM EDTA (pH 8.0), 0.1% NP-40, 1 mM phenylmethylsulfonyl fluoride, and protease inhibitor cocktail] and sonicated. After centrifugation at 13,000 rpm for 15 min at 4 °C, the supernatant was transferred into a new tube and precleared with 50  $\mu$ L of protein G-agarose beads (Roche). Next, the precleared supernatant was incubated with 50  $\mu$ L of protein G-agarose beads with

antibodies against NSrp70 (ALTAS) or IgG (Sigma), overnight at 4 °C, and washed with immunoprecipitation buffer. For the ubiquitin experiment *in vitro*, anti-TGFβR1 antibody was used (Abcam). To harvest the protein complex, 50 μL of 1× SDS loading buffer (62.4 mM Tris, pH 6.8, 10% glycerol, 2% SDS, and 0.0012% bromophenol blue) was added, the mixture incubated for 10 min at 95 °C, and analyzed by western blotting.

### **RNA-Binding Protein Immunoprecipitation**

RIP experiments were performed using an EZ-Magna RIP™ RNA-Binding Protein Immunoprecipitation kit 17-701 (Millipore), according to the manufacturer's instructions, using a modified version of a method described previously[49]. Briefly, MDA-MB-231HM cells ( $1 \times 10^7$ ) transfected with the indicated constructs were lysed on ice in complete RIP lysis buffer (Millipore). Anti-NSrp70 antibody or IgG control antibody was pre-coated onto magnetic bead protein A/G mixture for 30 min at room temperature. Antibody-coated beads were washed twice with RIP wash buffer (Millipore) using a magnetic separator. Protein lysates were then incubated with rabbit anti-NSrp70- or rabbit control IgG-coated beads overnight at 4 °C. The immunoprecipitated samples were then centrifuged and washed with ice-cold RIP wash buffer six times using a magnetic separator. After the final washing step, the RNA-protein complexes were dissociated from the beads by incubating in proteinase K buffer (Millipore) for 30 min at 55 °C. The supernatants were collected using a magnetic separator. RNA extraction from the supernatant was performed using the phenol:chloroform: isoamyl alcohol (125:24:1) (Sigma-Aldrich) method and precipitated using ethanol supplemented with salt solution I, salt solution II, and precipitation enhancer overnight. RNA pellets were then washed with and suspended in RNase-free water. The cDNA was transcribed from total RNA using the PrimeScript RT Reagent kit (TaKaRa Biotechnology).

Two truncations of NSrp70 M15 and RS1M were synthesized by Genewiz. The sequences of M15 and RS1M are provided in the Supplementary methods.

PCR was performed using the following primers:

NUMB-RIP: Forward primer: 5'- AAGCAGTGCTACCACCAGTC -3'

Reverse primer: 5'- GGGAGGGATTAGTACGCTGC -3'.

RT-PCR was performed using SYBR Premix Ex Taq (TaKaRa Biotechnology) and ABI Prism 7900 (Applied Biosystems).

### **Biotin-labeled RNA Pull-down Assay**

Biotinylated RNA pull-down assays were performed as described elsewhere[50]. Full-length DNA fragments of *NUMB* with a 5'-terminal T7 promoter were *in vitro* transcribed using the biotin RNA-labeling mix (Roche) and T7 transcription kit (HiScribe™ T7 Quick High Yield RNA Synthesis kit E2050S, BioLabs). The biotin-labeled RNA was purified by using Direct-zol RNA Miniprep R2052 kit (ZYMO RESEARCH). MDA-MB-231HM cells ( $10^7$  cells) were collected and suspended in 1 mL of RIPA buffer (25 mM Tris, pH 7.5, 150 mM KCl, 0.5 mM DTT, 0.5% NP40, 1 mM phenylmethylsulfonyl fluoride, 2 mM VRC, and protease inhibitor cocktail), followed by sonication. After centrifugation at 13,000 rpm for 15 min at 4 °C, the

supernatant was transferred into a new tube and precleared with 40  $\mu$ L of streptavidin Dynabeads (Dynabeads<sup>TM</sup> M-280 streptavidin, Invitrogen, Thermo Fisher Scientific) for 20 min at 4 °C. Next, 20  $\mu$ g/mL yeast tRNA was added to block unspecific binding and incubated for 20 min at 4 °C. The precleared lysate was supplemented with 2.5  $\mu$ g of biotin-labeled NUMB (full length). The mixture was then incubated for 1.5 h followed by addition of 40  $\mu$ L of streptavidin Dynabeads, and overnight incubation at 4 °C. Beads were washed four times for 5 min with RIPA buffer containing 0.5% sodium deoxycholate and boiled in 1 $\times$  SDS loading buffer for 10 min at 100 °C. The retrieved proteins were analyzed by western blotting with anti-NSrp70 (ATLAS) and anti-SRSF1 (Proteintech) antibodies.

The RNA sequences were as follows:

NUMB-pull-down:

Forward primer: 5'- TAATACGACTCACTATAGGGAGAAT  
GAACAAATTACGGCAAAG -3'

Reverse primer: 5'- TTAAAGTTCAATTTCAAACG -3';

### **mRNA-Sequencing and Alternative Splicing**

The total RNA from the cells was purified using TRIzol reagent (Invitrogen). The RNA quality was evaluated using an Agilent 2100 bioanalyzer. MISO software[51] was utilized to quantify the expression level of alternatively spliced genes from the mapped read data and to identify differentially regulated isoforms or exons across samples. The human genome (hg19) references of five known alternative splicing events were downloaded from the MISO website and prepared according to MISO manuals. The results were filtered based on both percent spliced-in (PSI) differences and PSI distribution plots generated by Sashimi plot analysis.

### ***In Vivo* Splicing Assays**

*In vivo* splicing assays were performed essentially as described previously[52, 53]. Briefly, a splicing reporter minigene was used to co-transfect MDA-MB-231HM cells with an increasing or decreasing amount of NSrp70-pCDH and SRSF1-pCDH. Empty plasmids were added to ensure that the same amount of DNA was used in the transfection reactions. Then, 48 h after transfection, total RNA was extracted from the cells. The cDNA was transcribed using the PrimeScript RT Reagent kit (TaKaRa Biotechnology). The *NUMB* minigene that included the exon 11-13 was synthesized by Genewiz. The sequence of *NUMB* minigene is provided in the Supplementary Methods. The *NUMB* minigene overexpression system was established based on double-digestion of the pCDH-CMV-MCV-EF1-Puro vector. The following primers were used for PCR amplification of *NUMB*:

Forward primer: 5'- TGCCAGAAGTAGAAGGGGAG -3'

Reverse primer: 5'- CGTAGTGAGGGAATGTCTGC -3'.

The overexpression construct for SRSF1 was generated in an analogous manner, using the enzymes EcoRI and BamHI (NEW ENGLAND BioLabs). The following primers were used for the construction of SRSF1-pCDH:

Forward primer: 5'- TGCTCTAGAATGTCGGGAGGTGGTGTGAT -3'

Reverse primer: 5'- CGGGATCCTTATGTACGAGAGCGAGATC -3'.

The PCR products were analyzed on a 1–1.5% agarose gel, and the splicing pattern was quantified using ImageJ (National Institutes of Health).

### **Metastasis Assays in Mice**

All animal experiments were performed in accordance with the guidelines of the Institutional Animal Care and Use Committee of Fudan University under approved protocols. For cell metastasis analysis,  $3 \times 10^5$  cells labeled with GFP/luciferase were washed in serum-free DMEM and injected intravenously into female BALB/c nude mice (N=6) to study lung metastasis. After 4-6 weeks, noninvasive bioluminescence imaging (BLI) was performed to quantify metastasis using an in vivo imaging system. To establish orthotopic transplantation tumor,  $2 \times 10^6$  cells/100  $\mu$ L were injected into the mammary fat pad of 8-week-old female NOD-scid mice. After 6-8 weeks, the tumors were removed and noninvasive bioluminescence imaging (BLI) was performed to quantify metastasis using an in vivo imaging system.

### **Quantification and Statistical Analysis**

All data are reported as mean  $\pm$  standard deviation (SD) as indicated in the figure legends. The data were analyzed using PRISM 5.0 (GraphPad Software Inc.) and SPSS 18.0 software (SPSS). Comparisons between two groups were performed by the student's t-test (two-tailed). The cumulative survival time was calculated by Kaplan-Meier method (log-rank test). The univariate and multivariate analyses were based on Cox proportional hazards regression model. The results with P-value  $< 0.05$  were considered statistically significant.

## **Abbreviations**

A3SS	Alternative 3'-splice sites
A5SS	Alternative 5'-splice sites
DFS	Disease-free survival
DNA	Deoxyribonucleic acid
EEA1	Early endosome antigen 1
EMT	Epithelial-mesenchymal Transition
HER-2	Human epidermal growth factor receptor 2
IHC	Immunohistochemistry
IP	Immunoprecipitation
MXE	Mutually exclusive exons
NSrp70	Nuclear Speckle-Related Protein 70
OD	Optical density
PCR	Polymerase chain reaction
RAB7	RAS oncogene family 7
RI	Intron retention
RIP	RNA-binding protein immunoprecipitation
sgRNA	Single guided RNA
SRSF1	Serine and arginine rich splicing factor1
TCGA	The cancer genome atlas
TGF	Transforming growth factor
TMA	Tissue microarrays

## **Declarations**

### **Ethics approval and consent to participate**

The present study was approved by the Medical Ethics Committee of Fudan University Shanghai Cancer Center and Experimental animal ethics committee of Fudan University.

### **Availability of data and materials**

Datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

### **Acknowledgements**

This work was supported by the grant from National Natural Science Foundation of China (81972727) and the Municipal Human Resources Development Program for Outstanding Leaders in Medical Disciplines in Shanghai (2017BR028).

### **Author contributions**

Conceptualization, Wei Jin and Hefen Sun; Methodology, Yang Zhao and Hefen Sun; Investigation, Yuanyuan Zhao; Writing-Original Draft, Yang Zhao; Writing-Review & Editing, Qiqi Liu, Hefen Sun and Yang Liu; Funding Acquisition, Wei Jin; Resources, Yifeng Hou; Supervision, Wei Jin and Hefen Sun.

### **Competing interests**

The authors declare that they have no competing interests.

### **Footnotes**

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

Yang Zhao and Hefen Sun contributed equally to this work.

## References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A: Global cancer statistics, 2012. *CA Cancer J Clin* 2015, 65:87-108.
2. Lee Y, Rio DC: Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu Rev Biochem* 2015, 84:291-323.
3. Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB: Alternative isoform regulation in human tissue transcriptomes. *Nature* 2008, 456:470-476.
4. Venables JP: Aberrant and alternative splicing in cancer. *Cancer Res* 2004, 64:7647-7654.
5. David CJ, Manley JL: Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev* 2010, 24:2343-2364.
6. Lodomery M: Aberrant alternative splicing is another hallmark of cancer. *Int J Cell Biol* 2013, 2013:463786.
7. Read A, Natrajan R: Splicing dysregulation as a driver of breast cancer. *Endocr Relat Cancer* 2018, 25:R467-r478.
8. Spector DL, Lamond AI: Nuclear speckles. *Cold Spring Harb Perspect Biol* 2011, 3.
9. Kim YD, Lee JY, Oh KM, Araki M, Araki K, Yamamura K, Jun CD: NSrp70 is a novel nuclear speckle-related protein that modulates alternative pre-mRNA splicing in vivo. *Nucleic Acids Res* 2011, 39:4300-4314.
10. Xie J, Gizatullin R, Vukojevic V, Leopardi R: The CCDC55 couples cannabinoid receptor CNR1 to a putative DISC1 schizophrenia pathway. *Neuroscience* 2015, 310:723-730.
11. Lee SH, Kim C, Lee HK, Kim YK, Ismail T, Jeong Y, Park K, Park MJ, Park DS, Lee HS: NSrp70 is significant for embryonic growth and development, being a crucial factor for gastrulation and mesoderm induction. *Biochem Biophys Res Commun* 2016, 479:238-244.
12. Choi DB, Park MR, Kim HR, Jun CD, Kim HJ, Shim H, Kim YD, Choi C, Choi KH, Yun KJ, et al: Aberrant proteomic expression of NSRP70 and its clinical implications and connection to the transcriptional level in adult acute leukemia. *Leuk Res* 2014, 38:1252-1259.
13. Jiang HL, Sun HF, Gao SP, Li LD, Huang S, Hu X, Liu S, Wu J, Shao ZM, Jin W: SSBP1 Suppresses TGFbeta-Driven Epithelial-to-Mesenchymal Transition and Metastasis in Triple-Negative Breast Cancer by Regulating Mitochondrial Retrograde Signaling. *Cancer Res* 2016, 76:952-964.
14. Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, Viale A, Olshen AB, Gerald WL, Massague J: Genes that mediate breast cancer metastasis to lung. *Nature* 2005, 436:518-524.
15. Imamura T, Hikita A, Inoue Y: The roles of TGF-beta signaling in carcinogenesis and breast cancer metastasis. *Breast Cancer* 2012, 19:118-124.
16. Pece S, Confalonieri S, P RR, Di Fiore PP: NUMB-ing down cancer by more than just a NOTCH. *Biochim Biophys Acta* 2011, 1815:26-43.
17. Reithmeier A, Panizza E, Krumpel M, Orre LM, Branca RMM, Lehtio J, Ek-

- Rylander B, Andersson G: Tartrate-resistant acid phosphatase (TRAP/ACP5) promotes metastasis-related properties via TGFbeta2/TbetaR and CD44 in MDA-MB-231 breast cancer cells. *BMC Cancer* 2017, 17:650.
18. Sun H, Liu Y, Zhang L, Shao X, Liu K, Ding Z, Liu X, Jiang C, Li H, Li H: Numb positively regulates autophagic flux via regulating lysosomal function. *Biochem Biophys Res Commun* 2017, 491:780-786.
  19. Lamond AI, Spector DL: Nuclear speckles: a model for nuclear organelles. *Nat Rev Mol Cell Biol* 2003, 4:605-612.
  20. Tripathi V, Song DY, Zong X, Shevtsov SP, Hearn S, Fu XD, Dundr M, Prasanth KV: SRSF1 regulates the assembly of pre-mRNA processing factors in nuclear speckles. *Mol Biol Cell* 2012, 23:3694-3706.
  21. Kim CH, Kim YD, Choi EK, Kim HR, Na BR, Im SH, Jun CD: Nuclear Speckle-related Protein 70 Binds to Serine/Arginine-rich Splicing Factors 1 and 2 via an Arginine/Serine-like Region and Counteracts Their Alternative Splicing Activity. *J Biol Chem* 2016, 291:6169-6181.
  22. Cheng C, Yaffe MB, Sharp PA: A positive feedback loop couples Ras activation and CD44 alternative splicing. *Genes Dev* 2006, 20:1715-1720.
  23. Sherman LS, Rizvi TA, Karyala S, Ratner N: CD44 enhances neuregulin signaling by Schwann cells. *J Cell Biol* 2000, 150:1071-1084.
  24. Bell MV, Cowper AE, Lefranc MP, Bell JI, Sreaton GR: Influence of intron length on alternative splicing of CD44. *Mol Cell Biol* 1998, 18:5930-5941.
  25. Loh TJ, Moon H, Cho S, Jung DW, Hong SE, Kim DH, Green MR, Zheng X, Zhou J, Shen H: SC35 promotes splicing of the C5-V6-C6 isoform of CD44 pre-mRNA. *Oncol Rep* 2014, 31:273-279.
  26. Sakuma K, Sasaki E, Kimura K, Komori K, Shimizu Y, Yatabe Y, Aoki M: HNRNPLL, a newly identified colorectal cancer metastasis suppressor, modulates alternative splicing of CD44 during epithelial-mesenchymal transition. *Gut* 2018, 67:1103-1111.
  27. Erb U, Megaptche AP, Gu X, Buchler MW, Zoller M: CD44 standard and CD44v10 isoform expression on leukemia cells distinctly influences niche embedding of hematopoietic stem cells. *J Hematol Oncol* 2014, 7:29.
  28. Li N, Tsuji M, Kanda K, Murakami Y, Kanayama H, Kagawa S: Analysis of CD44 isoform v10 expression and its prognostic value in renal cell carcinoma. *BJU Int* 2000, 85:514-518.
  29. Kandachar V, Roegiers F: Endocytosis and control of Notch signaling. *Curr Opin Cell Biol* 2012, 24:534-540.
  30. McGill MA, Dho SE, Weinmaster G, McGlade CJ: Numb regulates post-endocytic trafficking and degradation of Notch1. *J Biol Chem* 2009, 284:26427-26438.
  31. Santolini E, Puri C, Salcini AE, Gagliani MC, Pelicci PG, Tacchetti C, Di Fiore PP: Numb is an endocytic protein. *J Cell Biol* 2000, 151:1345-1352.
  32. Misquitta-Ali CM, Cheng E, O'Hanlon D, Liu N, McGlade CJ, Tsao MS, Blencowe BJ: Global profiling and molecular characterization of alternative splicing events misregulated in lung cancer. *Mol Cell Biol* 2011, 31:138-150.

33. Zong FY, Fu X, Wei WJ, Luo YG, Heiner M, Cao LJ, Fang Z, Fang R, Lu D, Ji H, Hui J: The RNA-binding protein QKI suppresses cancer-associated aberrant splicing. *PLoS Genet* 2014, 10:e1004289.
34. Wang Z, Li SS: Numb: A new player in EMT. *Cell Adh Migr* 2010, 4:176-179.
35. Tripathi V, Shin JH, Stuelten CH, Zhang YE: TGF-beta-induced alternative splicing of TAK1 promotes EMT and drug resistance. *Oncogene* 2019, 38:3185-3200.
36. Shen H, Kan JL, Green MR: Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. *Mol Cell* 2004, 13:367-376.
37. Matter N, Marx M, Weg-Remers S, Ponta H, Herrlich P, König H: Heterogeneous ribonucleoprotein A1 is part of an exon-specific splice-silencing complex controlled by oncogenic signaling pathways. *J Biol Chem* 2000, 275:35353-35360.
38. Kashima T, Manley JL: A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. *Nat Genet* 2003, 34:460-463.
39. Pandit S, Zhou Y, Shiue L, Coutinho-Mansfield G, Li H, Qiu J, Huang J, Yeo GW, Ares M, Jr., Fu XD: Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. *Mol Cell* 2013, 50:223-235.
40. Anko ML, Muller-McNicoll M, Brandl H, Curk T, Gorup C, Henry I, Ule J, Neugebauer KM: The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes. *Genome Biol* 2012, 13:R17.
41. Bradley T, Cook ME, Blanchette M: SR proteins control a complex network of RNA-processing events. *Rna* 2015, 21:75-92.
42. Kahles A, Lehmann KV, Toussaint NC, Hüser M, Stark SG, Sachsenberg T, Stegle O, Kohlbacher O, Sander C, Rätsch G: Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell* 2018, 34:211-224.e216.
43. Lee SC, Abdel-Wahab O: Therapeutic targeting of splicing in cancer. *Nat Med* 2016, 22:976-986.
44. DeVos SL, Miller TM: Antisense oligonucleotides: treating neurodegeneration at the level of RNA. *Neurotherapeutics* 2013, 10:486-497.
45. Havens MA, Hastings ML: Splice-switching antisense oligonucleotides as therapeutic drugs. *Nucleic Acids Res* 2016, 44:6549-6563.
46. Xiao X, Wang L, Wei P, Chi Y, Li D, Wang Q, Ni S, Tan C, Sheng W, Sun M, et al: Role of MUC20 overexpression as a predictor of recurrence and poor outcome in colorectal cancer. *J Transl Med* 2013, 11:151.
47. Zheng YZ, Cao ZG, Hu X, Shao ZM: The endoplasmic reticulum stress markers GRP78 and CHOP predict disease-free survival and responsiveness to chemotherapy in breast cancer. *Breast Cancer Res Treat* 2014, 145:349-358.
48. Hsu PD, Lander ES, Zhang F: Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 2014, 157:1262-1278.
49. Townley-Tilson WH, Pendergrass SA, Marzluff WF, Whitfield ML: Genome-

- wide analysis of mRNAs bound to the histone stem-loop binding protein. *Rna* 2006, 12:1853-1867.
50. Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, Zhu S, Yang L, Chen LL: Circular intronic long noncoding RNAs. *Mol Cell* 2013, 51:792-806.
  51. Katz Y, Wang ET, Airoidi EM, Burge CB: Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods* 2010, 7:1009-1015.
  52. Cooper TA: Use of minigene systems to dissect alternative splicing elements. *Methods* 2005, 37:331-340.
  53. Stoss O, Stoilov P, Hartmann AM, Nayler O, Stamm S: The in vivo minigene approach to analyze tissue-specific splicing. *Brain Res Brain Res Protoc* 1999, 4:383-394.

## Figure Legends

### **Fig 1. NSrp70 Expression is Downregulated in Highly Metastatic Breast Cancer Cells**

**a** Schematic overview of quantitative protein analysis approaches. Three biological replicates were analyzed per group. **b** Scatter plot of the relative expression of different genes in MDA-MB-231, MDA-MB-231HM, and MDA-MB-231BO cells. **c** Relative expression of NSrp70 in breast cancer cell lines, assessed by western blotting. **d** Relative expression of NSrp70 mRNA in breast cancer cell lines, evaluated by real-time polymerase chain reaction (RT-PCR). **e** Relative expression of NSrp70 in different subtypes of breast cancer, based on TCGA database analysis. **f-g** Oncomine-based box of NSrp70 expression in primary tumors and metastasis sites based on Chin Breast2 (n=219) and Loi breast (n=87) analysis. **h** Oncomine-based box of NSrp70 expression in primary tumors and recurrent sites based on Loi breast3 analysis (n=77). **i-j** Oncomine-based box of NSrp70 expression in “alive and deceased at 5 years” based on Esserman breast (n=54) and Chin breast (n=169) analysis. **k** Distant metastasis-free survival analysis of NSrp70 expression in breast cancer samples using an online Kaplan-Meier plotter.

### **Fig 2. Loss of NSrp70 Promotes Breast Cancer Cell Migration and Metastasis *In Vitro* and *In Vivo***

**a** Stable knockdown of NSrp70 in MDA-MB-231, Hs-578T, and BT549 cells. **b** NSrp70 overexpression in MDA-MB-231HM and Hs-578T cells. **c-d** Migration and invasion in cell lines upon NSrp70 overexpression and knockdown, evaluated by a transwell assay. **c** Images of representative fields (magnification, ×100) of invasive cells. **d** Histograms depicting the transwell assay results analysis. Statistical analysis was performed by using Student's *t*-test (n = 3). Error bars, SD. \**p* < 0.05 vs. control. **e-g** Effect of NSrp70 on the migration of MDA-MB-231 and BT-549 cells, evaluated in a wound-healing assay. The red lines indicate the initial scratch wound location and the yellow area indicated the scratch wound mask. Images were captured every 4 h after

wounding. **h** The effect of NSrp70 on lung metastasis in mouse. BLI of three representative mice from each group 4-6 weeks after treatment are shown. The top two panels show animals after injection of NSrp70-overexpressing MDA-MB-231HM cells or NSrp70 knockdown MDA-MB-231 cells. The lower panel shows an animal after injection of NSrp70-overexpressing MDA-MB-231HM cells into the mammary fat pad. **i** Maximum BLI signals of each group. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

### **Fig 3. NSrp70 Inhibits SE Alternative Splicing of Target Pre-mRNA**

**a** Statistical histogram for the number of five alternative splicing events occurring in NSrp70-overexpressing cell lines Hs-578T and MDA-MB-231HM. Scatter plot of SE alternative splicing subtypes in the NSrp70-overexpressing cell lines MDA-MB-231HM and Hs-578T. **b** NSrp70 inhibits *NUMB* exon 12 SE splicing. The left panels show alternative splicing isoforms and a Sashimi-style plot. The right panels show RT-PCR verification of changes in the target gene splicing after NSrp70 overexpression or downregulation. \*\*\* $p < 0.001$ . **c** Construction pattern diagram of two truncated variants (M15 and RS1M) of NSrp70. RIP assay evaluation of the binding of NSrp70 or truncation variants to *NUMB* mRNAs. **d** RNA pull-down from *NUMB*-overexpressing MDA-MB-231 cell lines. Western blot analysis of immunoprecipitation assay with NSrp70-overexpressing MDA-MB-231HM cells. **e** Silver-staining of immunoprecipitates from overexpressed NSrp70 MDA-MB-231HM cells. **f-g** *In vivo* splicing assays. **f** NSrp70 inhibits *NUMB* exon 12 exclusion. MDA-MB-231HM cells were transfected with decreasing amounts of a vector encoding NSrp70 (2-0  $\mu\text{g}$ ). A parental vector (EV, 0-2  $\mu\text{g}$ ) was added to ensure that similar amount of cDNA was used in all transfection experiments. The increased or reduced protein levels of NSrp70 were confirmed by western blotting with the indicate antibodies (top panel). Exon inclusion or exclusion was evaluated by RT-PCR (center panel). The ratios of exclusion or inclusion of *NUMB* exon12 are shown as bar graphs (bottom panel). **g** Promotion of *NUMB* minigene exon12 exclusion by SRSF1. MDA-MB-231HM cells were co-transfected with increasing amounts of plasmids encoding SRSF1 (0-4  $\mu\text{g}$ ), decreasing amounts of plasmids encoding NSrp70 (2-0  $\mu\text{g}$ ), and the *NUMB* minigene (1  $\mu\text{g}$ ).

### **Fig 4. Knockdown of NSrp70 Promotes TGF $\beta$ -induced EMT**

**a** Kyoto Encyclopedia of Genes and Genomes signaling pathway enrichment histogram. MDA-MB-231HM (pCDH vs. NSrp70) and Hs-578T (pCDH vs. NSrp70) group analyses of transcriptome sequence data were performed. **b** Western blotting evaluation of the expression of T $\beta$ R1/2 and key proteins of the TGF $\beta$  signaling in NSrp70-overexpressing and knockdown cell lines. **c,d** Expression of TGF $\beta$  signaling and EMT-related proteins in NSrp70 knockdown MDA-MB-231 cells and NSrp70-overexpressing Hs-578T cells after treatment with a TGF $\beta$  signaling activator. **e** Expression of TGF $\beta$  signaling and EMT-related proteins in NSrp70 knockdown BT549 and Hs-578T cells after treatment with GW788388. **f** Transwell assay to detect changes in cell invasion ability of MDA-MB-231 and Hs-578T cells with a stably knocked down NSrp70 after treatment with a TGF $\beta$  signaling suppressor (magnification,  $\times 100$ ). **g**

Overexpression of NUMBv, and NUMBwt in MDA-MB-231 cells. Transwell assay to detect changes in cell invasion ability of the NUMBv-, and NUMBwt-overexpressing cell lines (magnification, ×100). **h** Expression of TGFβ signaling and EMT-related proteins in NUMBv-, and NUMBwt-overexpressing MDA-MB-231 cell lines, and NSrp70-overexpressing and NUMBv co-overexpressing MDA-MB-231HM cell lines. V: alternative splicing variants.

**Fig 5. NSrp70 Levels Correlate with TβR1 Degradation**

**a** Western blot analysis of immunoprecipitation assay with NUMBwt and NUMBv-overexpressing MDA-MB-231HM cells. **b,c** After incubation with of NSrp70-overexpressing MDA-MB-231HM and NUMBv-overexpressing cells with 50 μM CHX for 0-24 h, the expression of TβR1/2 was evaluated by western blotting. **d** Western blot analysis of TβR1/2 expression in NSrp70-overexpressing MDA-MB-231HM cells following incubation with 10 μM NH<sub>4</sub>Cl for the indicated time periods. **e** Western blot analysis of TβR1 expression in NUMBv or NUMBwt-overexpressing MDA-MB-231HM cells following incubation with 400 nM bfa1 for the indicated time periods. **f** Western blot analysis of TβR1 expression in NUMBv-overexpressing MDA-MB-231 cells following incubation with 10 μM NH<sub>4</sub>Cl for the indicated time periods. v: alternative splicing variants; wt: wild type. **g** Immunofluorescently-detected TβR1 and lysosome pathway-related indicators EEA1, Rab7, and Rab11 in MDA-MB-231 cells with stably downregulated NSrp70 (magnification, ×600).

**Fig 6. Low NSrp70 Expression Correlates with Poor Patient Prognosis**

**a** Representative image of high and low expression of NSrp70 detected by immunohistochemistry (IHC). Top image, ×100 magnification; bottom image, ×400 magnification. **b** Kaplan-Meier curves showing that the disease-free survival of individuals with low NSrp70 expression is significantly lower than that of individuals with a high NSrp70 expression ( $p < 0.001$ , log-rank test). Cumulative DFS curves of patients with TNBC ( $p < 0.0049$ , log-rank test). Cumulative DFS curves of patients with other non-TNBC (Luminal A, Luminal B, and HER-2 positive subtype) showing no significant difference. **c** Model of the mechanism of NSrp70 (→: promotion; ⇨: inhibition).

## Tables

### Table1

Alteration of NSrp70 protein expression levels in relation to clinicopathological characteristics in breast cancer patients.

Variables		NSrp70						Total	p
		Low expression n=117		High expression n=131		n=248			
		No.	%	No.	%	No.	%		
age(years)	10-49	52	44.4	66	50.4	118	47.6	0.35	
	50-85	65	55.6	65	49.6	130	52.4		
Marital status	Married	115	98	125	95.4	236	95	0.44	
	Not married	1	1	3	2.3	4	2.5		
	Unknown	1	1	3	2.3	4	2.5		
Grade	I	8	6.8	4	3	12	5	0.319	
	II	53	45.2	66	50.4	119	48		
	III	42	36	51	39	93	37.5		
	Ungraded	14	12	10	7.6	24	9.5		
Tumor size(cm)	≤2	48	41	64	49	112	45.3	0.221	
	>2 and ≤5	58	50	63	48	121	48.7		
	>5	9	7.6	3	2.3	12	4.8		
	Unkonwn	2	1.4	1	0.7	3	1.2		
LN status	Negative	55	47	95	73	150	60	<0.001	
	Positive	62	53	36	27	98	40		
ER	Negative	66	56	75	57	141	57	0.569	
	Positive	50	43	56	43	106	42.6		
	Unknown	1	1	0	0	1	0.4		
PR	Negative	87	74	95	73	182	73	0.743	
	Positive	28	24	35	26.6	63	25		
	Unknown	2	2	1	0.4	3	2		
HER-2	Negative	61	52	84	64	145	58.5	0.105	
	Positive	55	47	47	36	102	41.1		
	Unknown	1	1	0	0	1	0.4		

Notes: LN: lymph nodes; ER: estrogen receptor; PR: progesterone receptor; Her-2: human epidermal growth factor receptor 2.

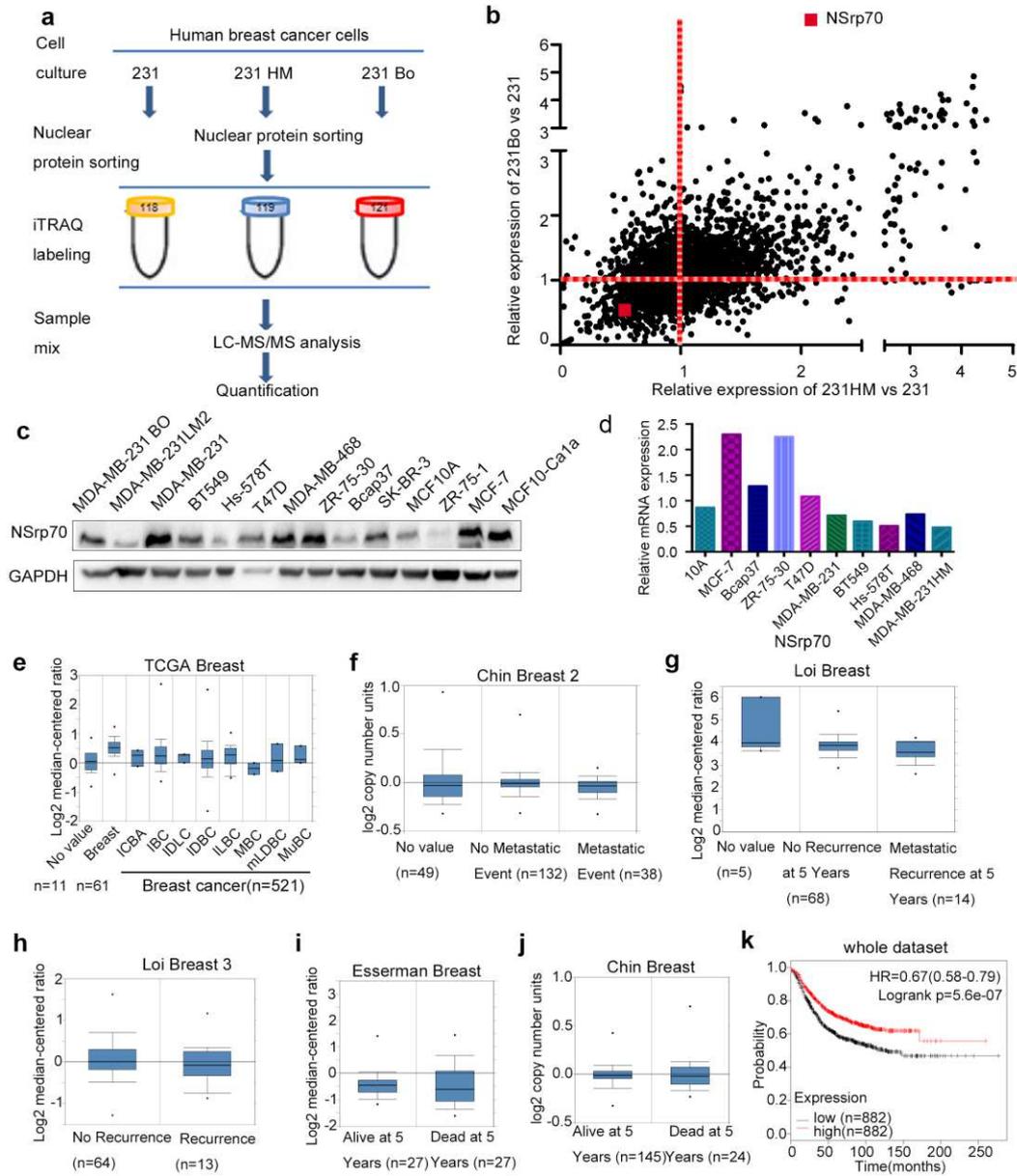
**Table2**

Univariate and multivariate survival analysis of NSrp70 expression in breast cancer patients.

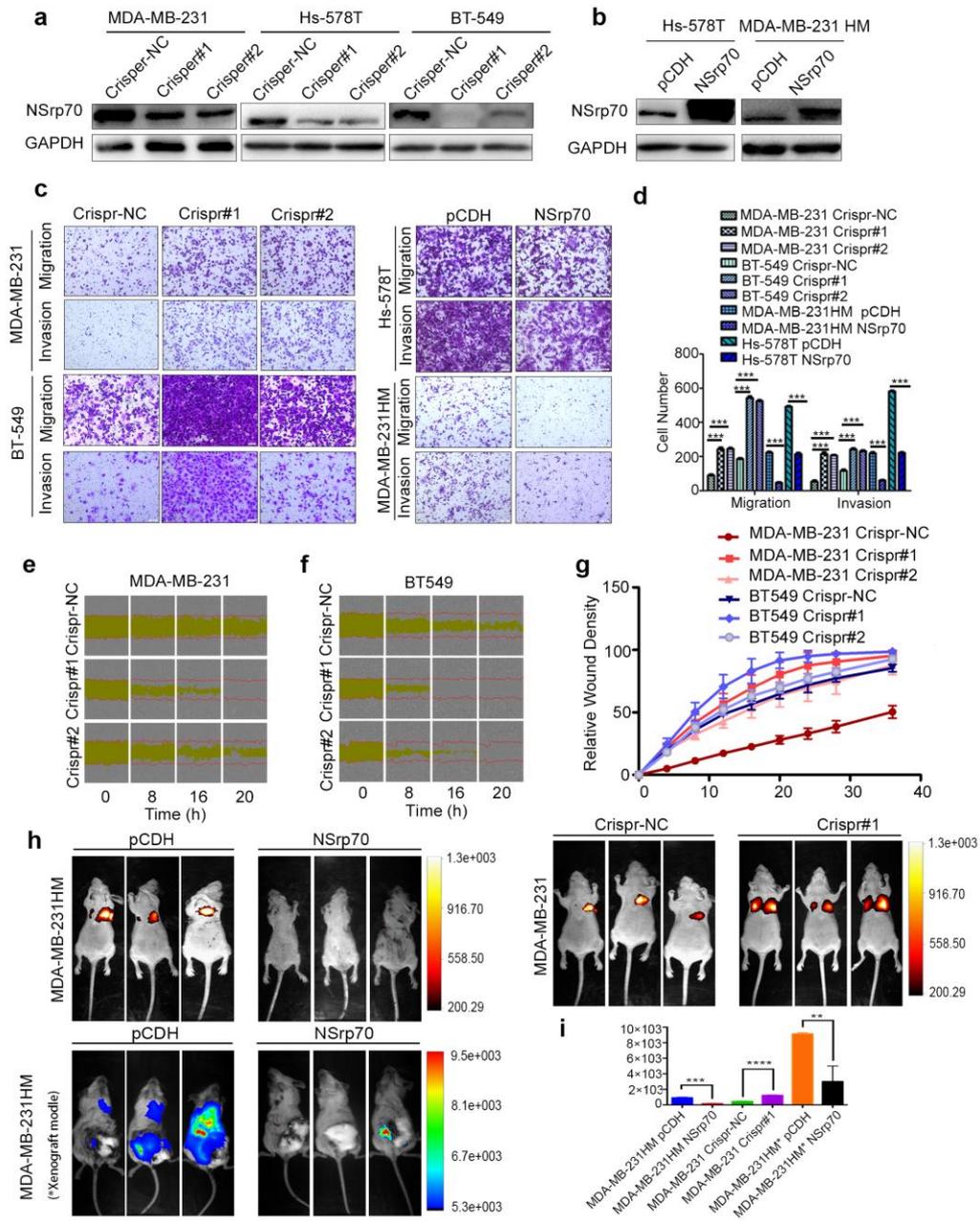
Variables	Univariate analysis		Multivariate analysis	
	HR (95%CI)	p-value	HR (95%CI)	p-value
Age(years)	0.592 (0.309-1.133)	0.114	0.549(0.281-1.076)	0.081
Menstrual status	10.984 (1.389-86.838)	0.023	12.434 (1.268-121.915)	<b>0.03</b>
Grade	0.613 (0.329-1.144)	0.124	0.701 (0.356-1.381)	0.304
Tumor size(cm)	0.213 (0.094-0.483)	<0.001	0.321 (0.132-0.783)	<b>0.012</b>
LN status	0.388 (0.207-0.728)	0.003	0.569 (0.288-1.125)	0.105
ER	1.264 (0.676-2.363)	0.463	0.796(0.353-1.795)	0.583
PR	2.16 (0.954-4.889)	0.065	2.402 (0.855-6.748)	0.096
HER-2	1.191 (0.635-2.233)	0.586	1.791 (0.867-3.7)	0.115
NSrp70	3.978 (1.983-7.977)	<0.001	3.927(1.856-8.31)	<b>&lt;0.001</b>

Notes: LN: lymph nodes; ER: estrogen receptor; PR: progesterone receptor; Her-2: human epidermal growth factor receptor 2; CI: confidence interval.

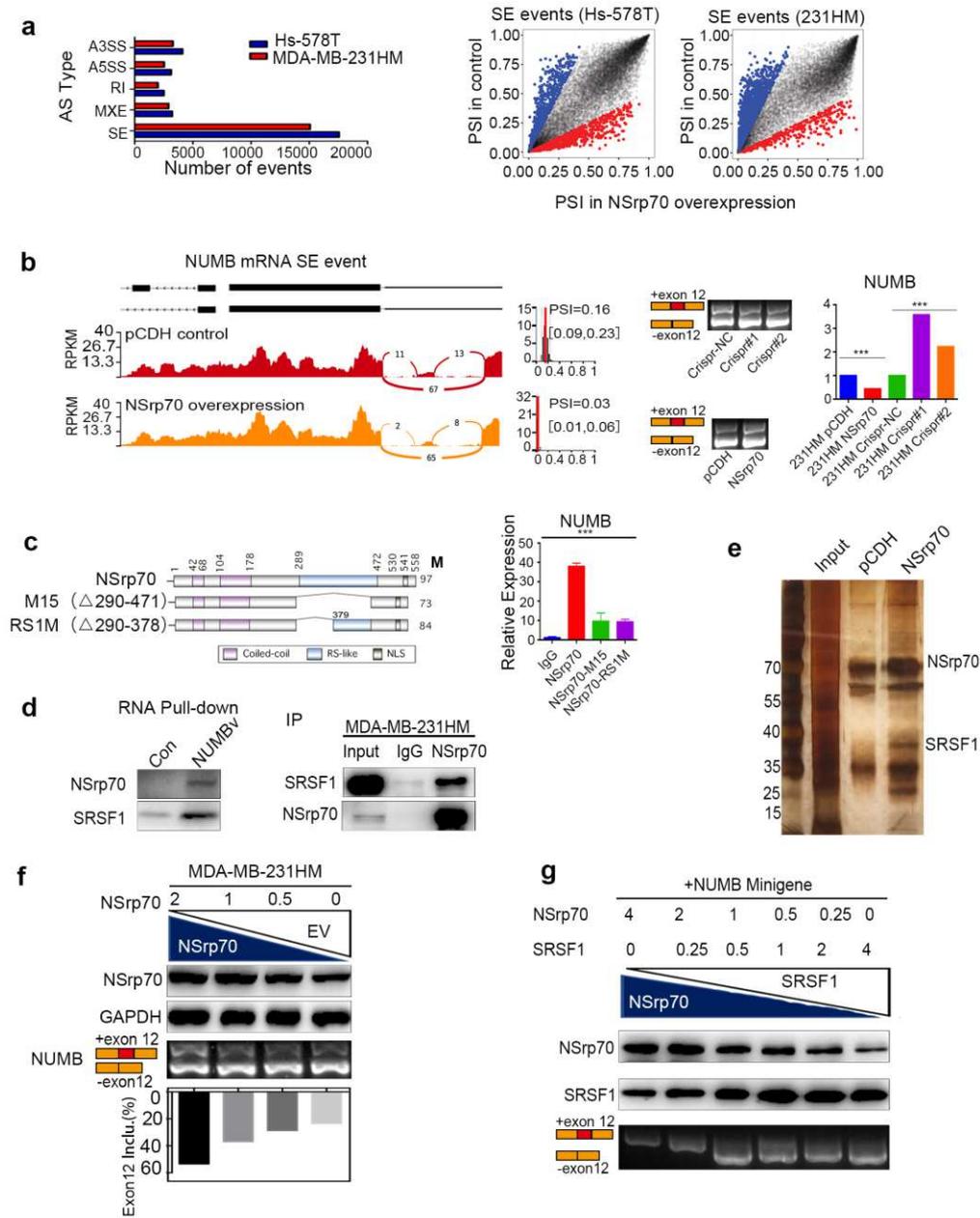
**Fig. 1**



**Fig. 2**

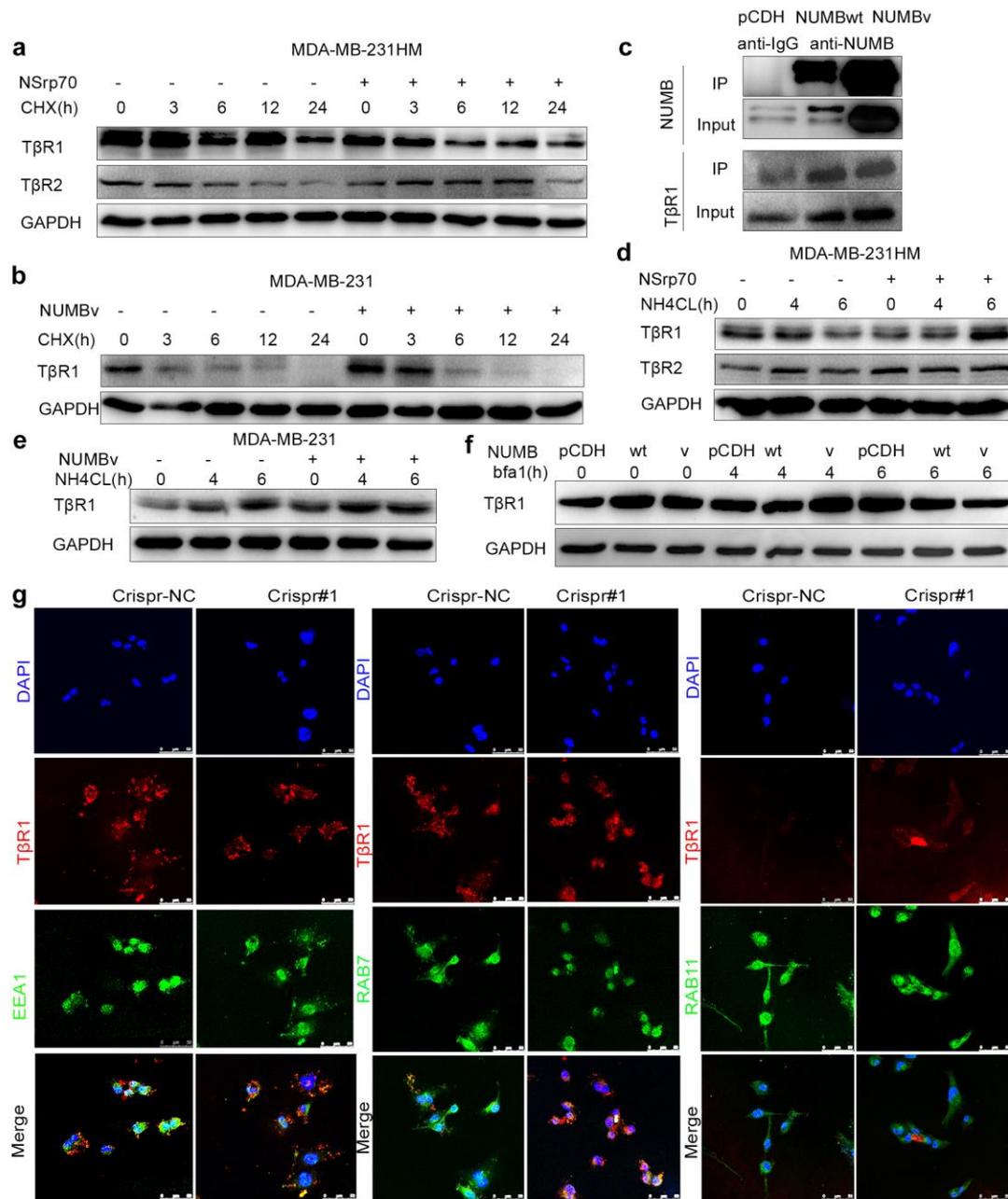


**Fig. 3**

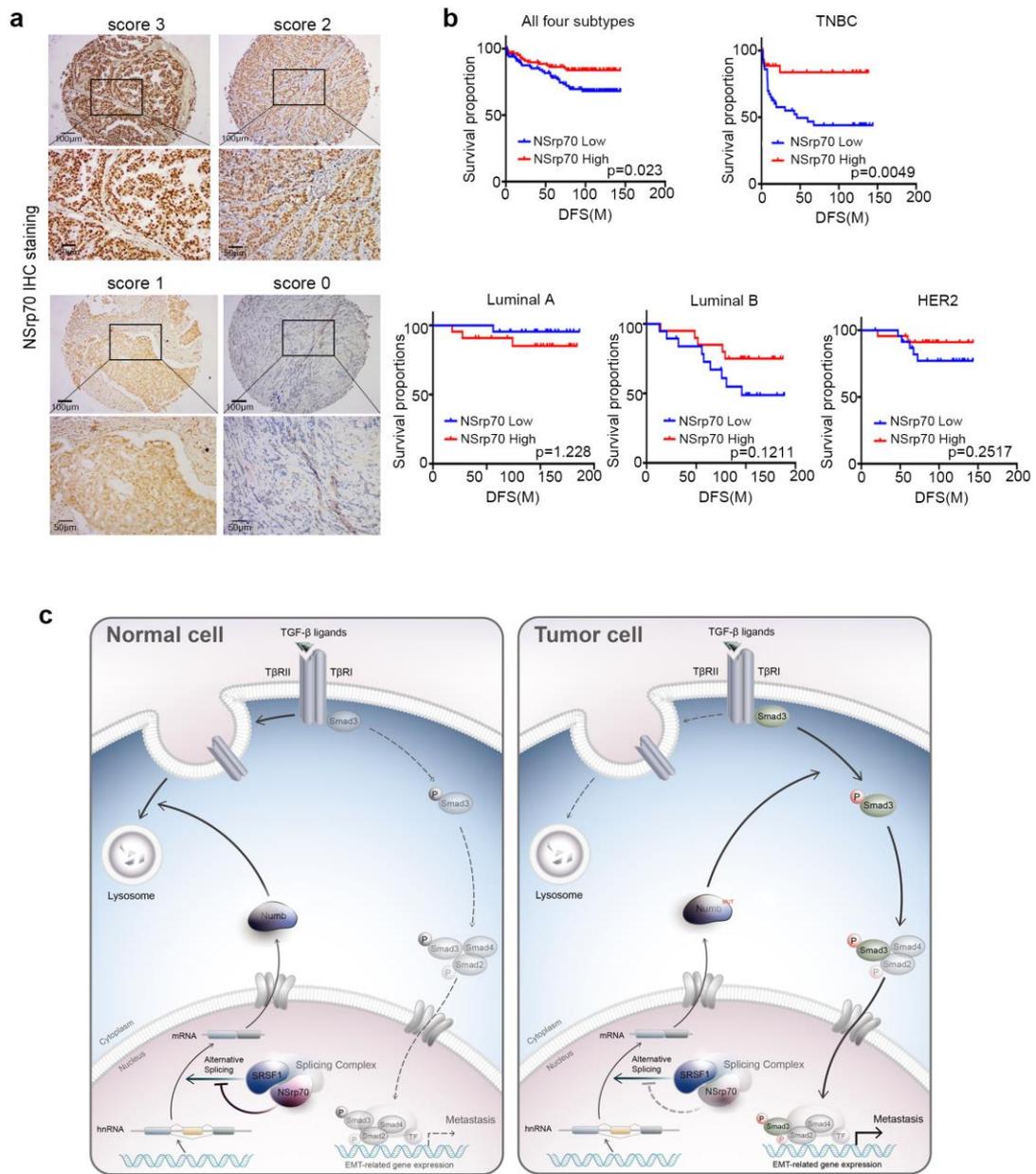




**Fig. 5**



**Fig. 6**



# Figures

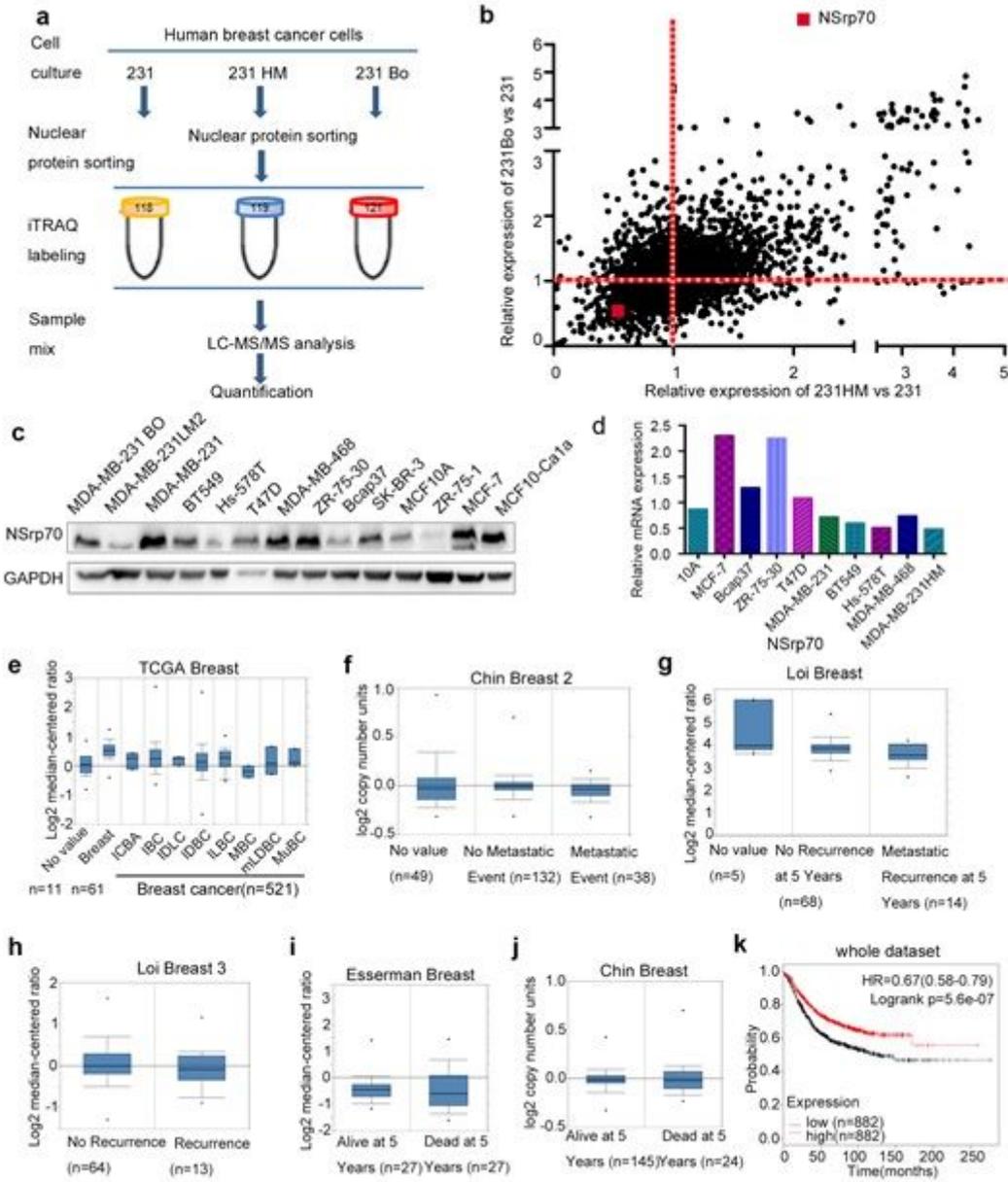
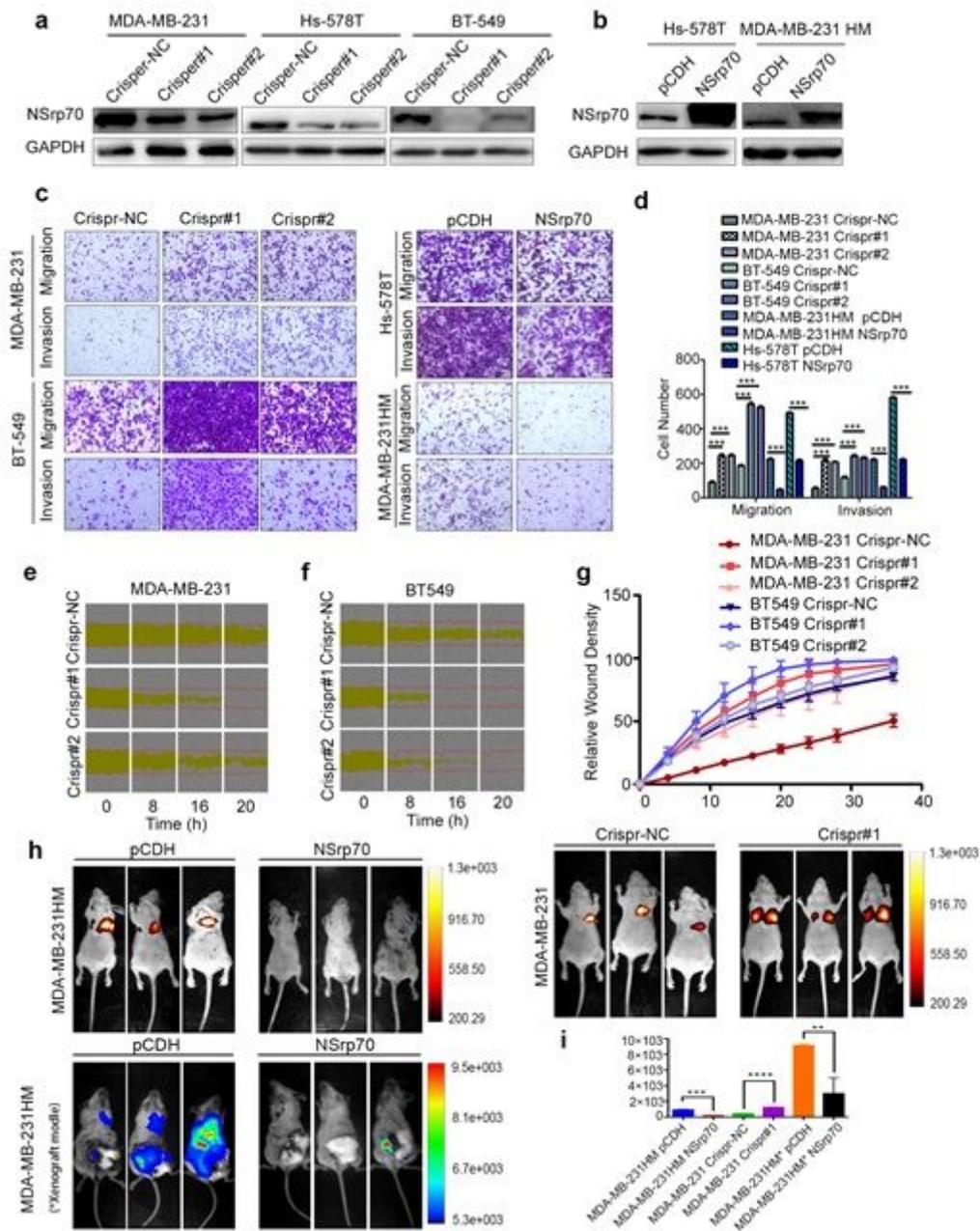


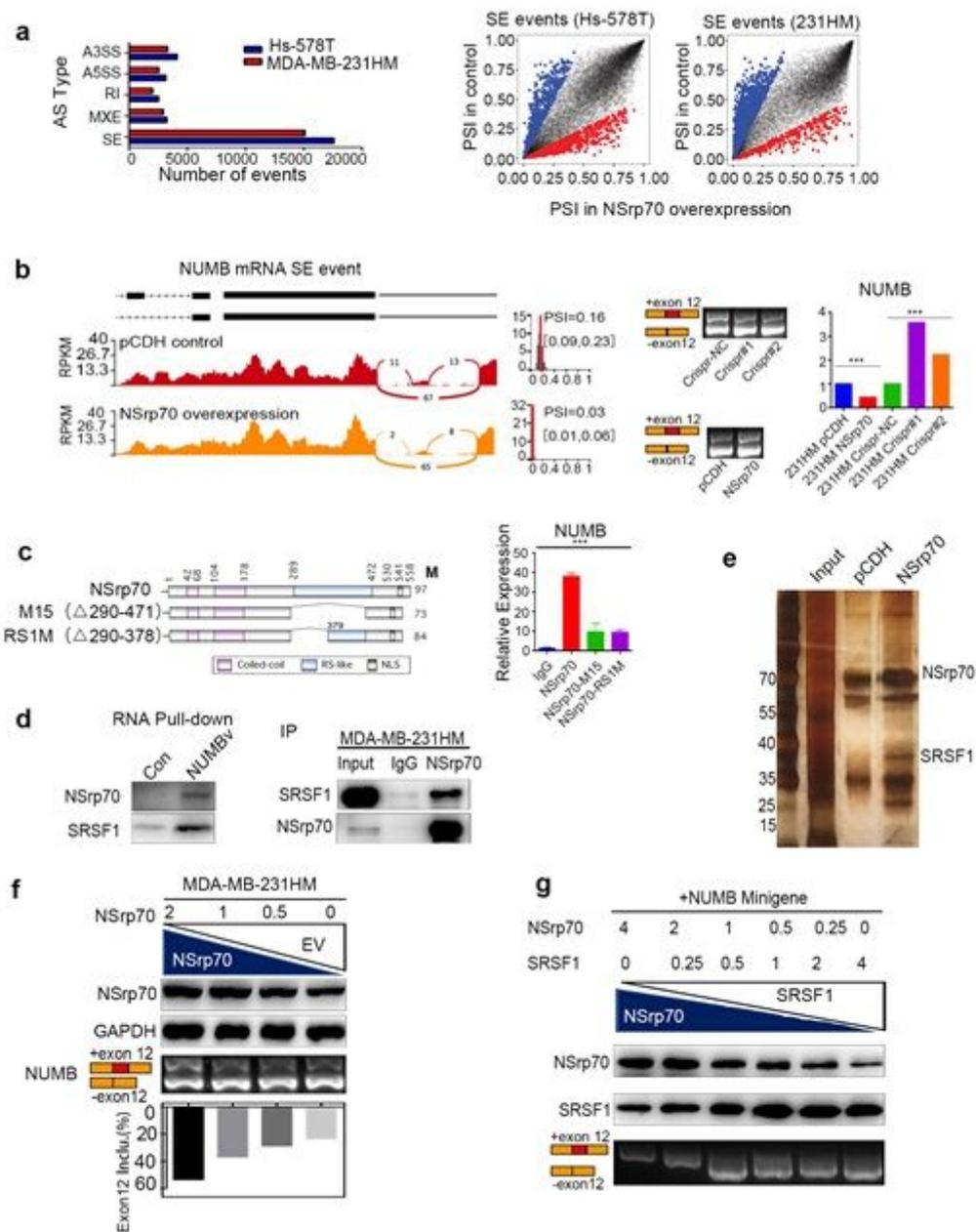
Figure 1

NSrp70 Expression is Downregulated in Highly Metastatic Breast Cancer Cells



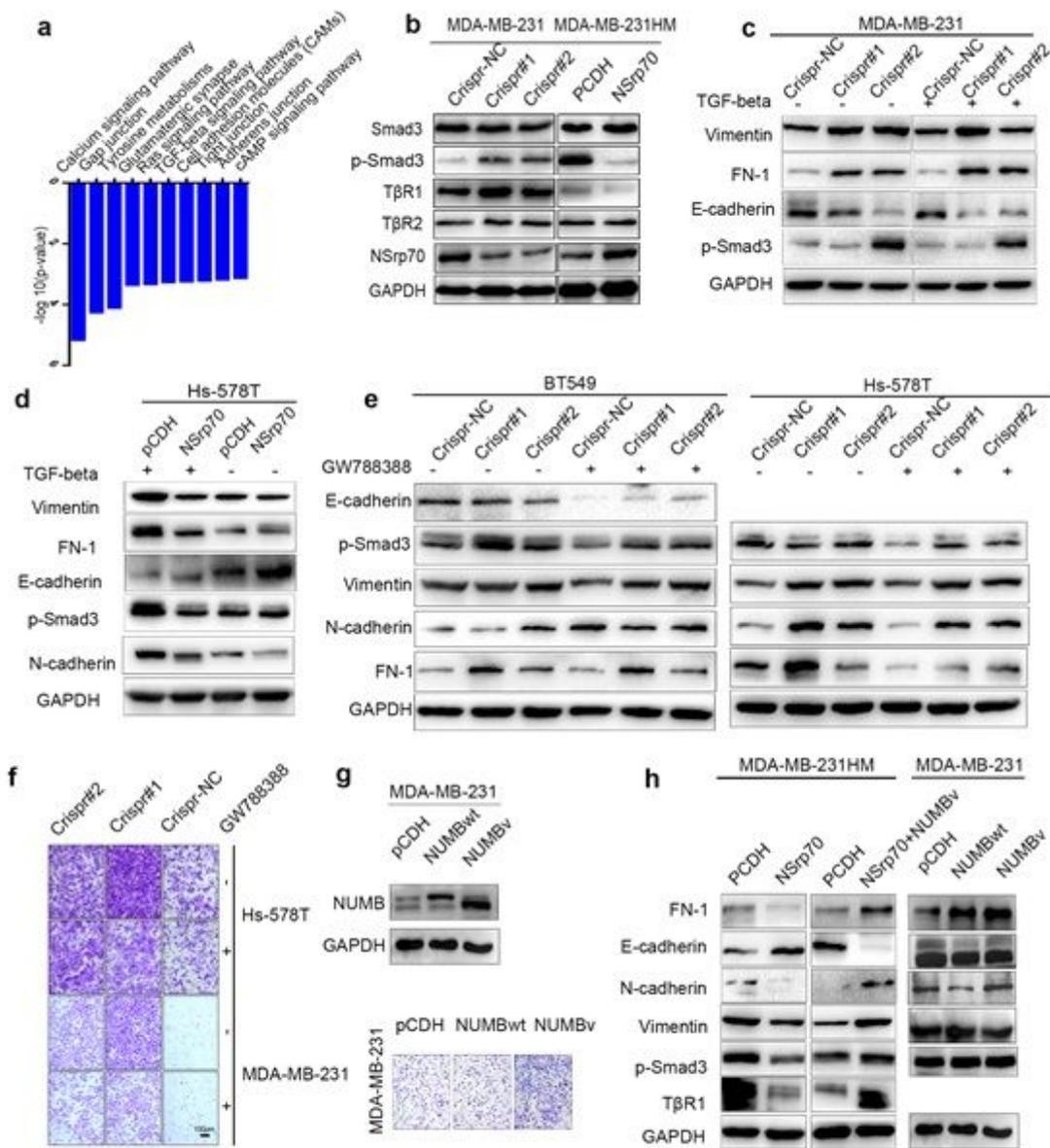
**Figure 2**

Loss of NSrp70 Promotes Breast Cancer Cell Migration and Metastasis In Vitro and In Vivo



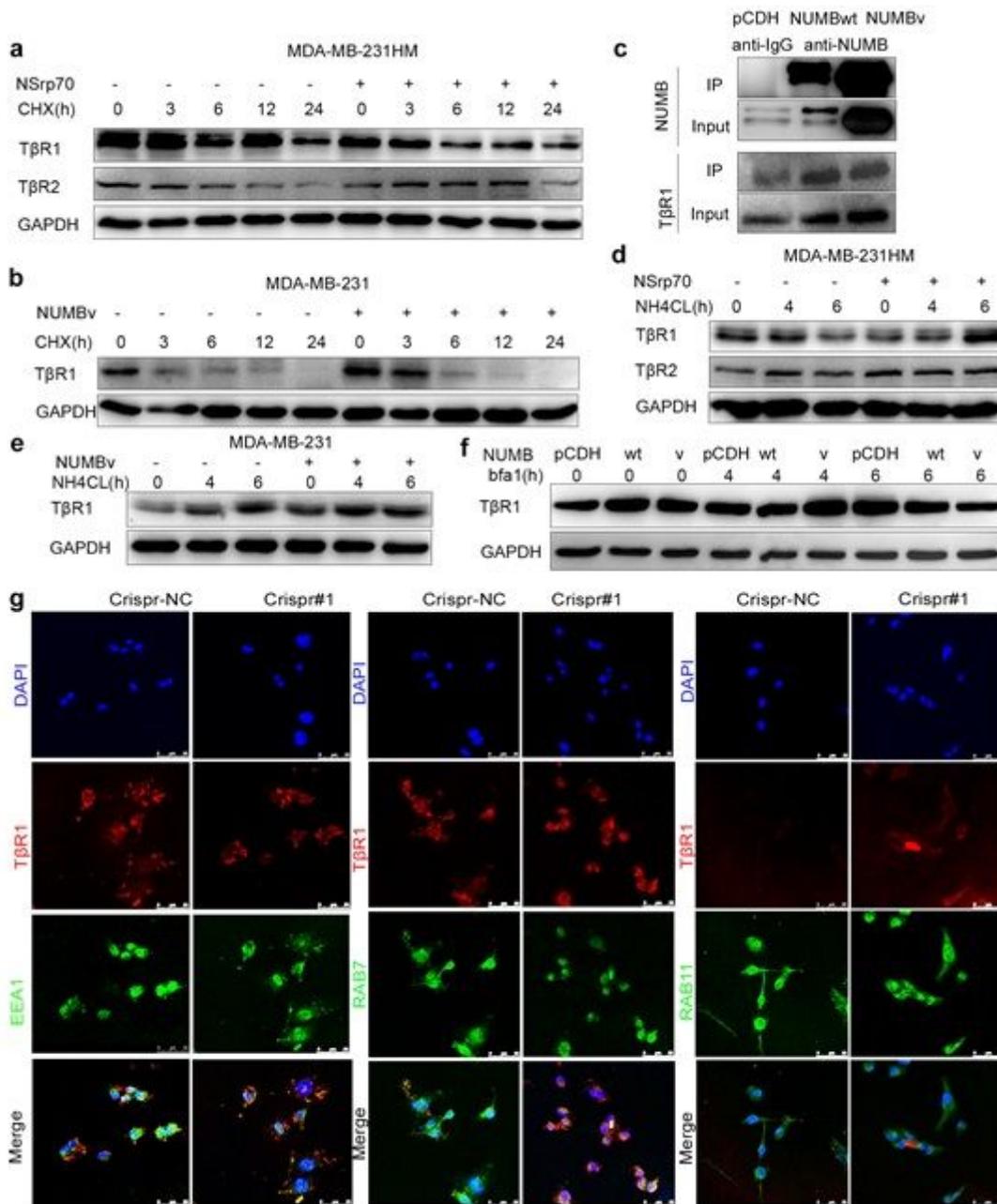
**Figure 3**

NSrp70 Inhibits SE Alternative Splicing of Target Pre-mRNA



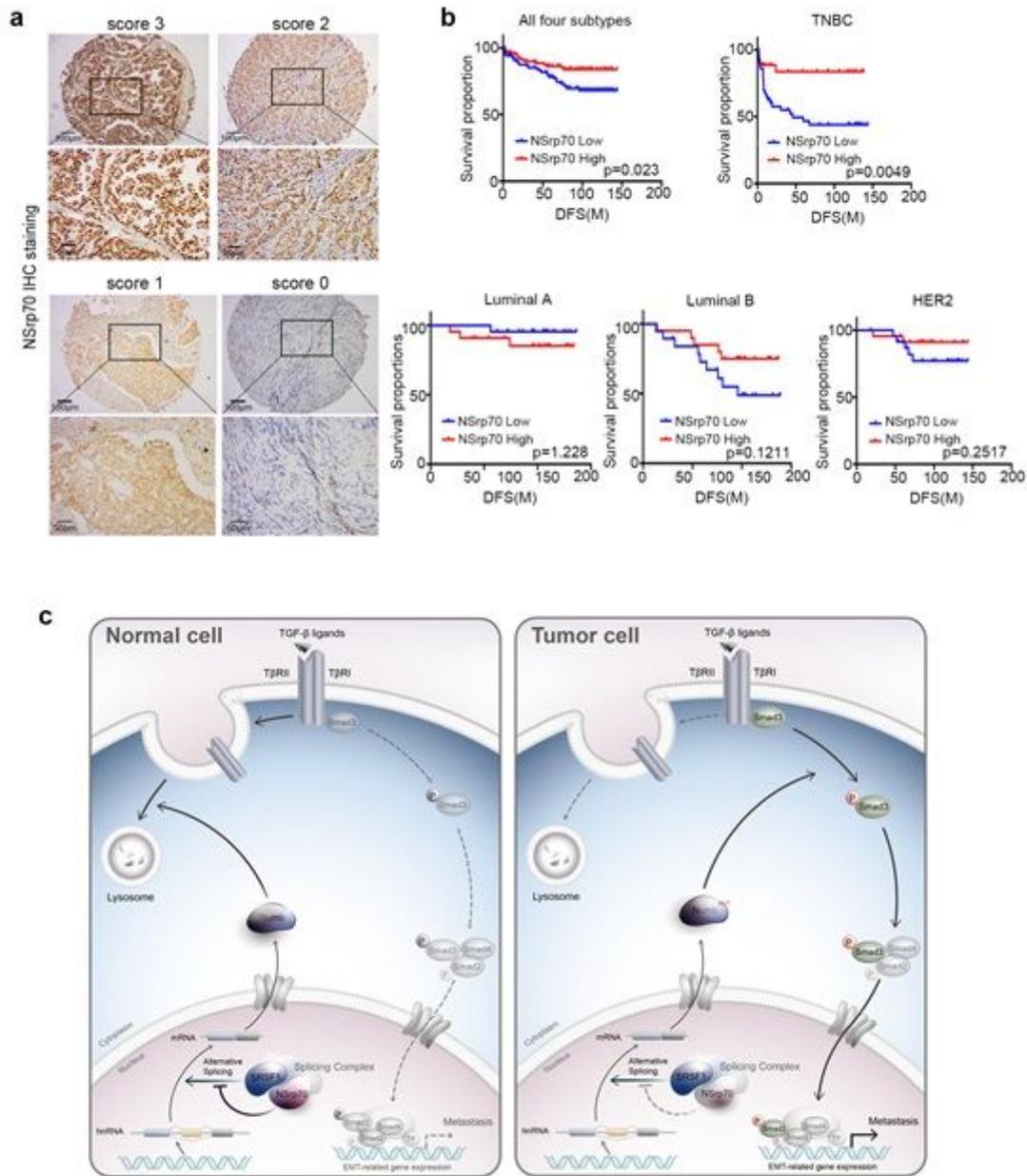
**Figure 4**

Knockdown of NSrp70 Promotes TGFβ-induced EMT



**Figure 5**

NSrp70 Levels Correlate with TBR1 Degradation



**Figure 6**

Low NSrp70 Expression Correlates with Poor Patient Prognosis

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFiles.docx](#)