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

Yang Zhao1,2,3, Hefen Sun1,2,3✉, Yuanyuan Zhao1,2, Qiqi Liu1,2, Yang Liu1, Yifeng Hou1,2 and Wei Jin1,2✉

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Alternative splicing of mRNA precursors allows cancer cells to create different protein isoforms that promote growth and survival. Compared to normal cells, cancer cells frequently exhibit a higher diversity of their transcriptomes. A comprehensive understanding of splicing regulation is required to correct the splicing alterations for the future precision oncology. A quantitative proteomic screen was performed to identify the regulators associated the metastasis in triple-negative breast cancer. Multiple in vitro and in vivo functional analyses were used to study the effects of NSrp70 on breast cancer metastasis. Next, transcriptomic sequencing (RNA-seq) and alternative splicing bioinformatics analysis was applied to screen the potential targets of NSrp70. Moreover, in vitro splicing assays, RNA pull-down, and RNA immunoprecipitation assay were used to confirm the specific binding between NSrp70 and downstream target genes. Furthermore, the prognostic value of NSrp70 was analyzed in a cohort of patients by performing IHC. We uncovered NSrp70 as a novel suppressor of breast cancer metastasis. We discovered that NSrp70 inhibited the skipped exon alternative splicing of NUMB, promoted the degradation of transforming growth factor receptor 1 through lysosome pathway, and regulated TGFβ/SMAD-mediated epithelial-mesenchymal transition phenotype in breast cancer cells. Furthermore, high NSrp70 expression correlated with a better prognosis in breast cancer patients. Our findings revealed that splicing regulator NSrp70 serves as a metastasis suppressor.

Oncogene (2022) 41:3409–3422; https://doi.org/10.1038/s41388-022-02349-z

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.

To explore splicing factors that may function in TNBC, we utilized the isobaric tags for relative and absolute quantitation (iTRAQ) technology followed by nanoscale high-performance liquid chromatography tandem mass spectrometry to analyze the different expression of nuclear proteins in TNBC cell line 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. NSrp70 was identified as a candidate protein.

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 [8]. It was first identified as a coiled-coil domain containing 55 protein and it was one of the nuclear speckles that are the storage sites of mRNA splicing regulators, including protein splicing factors such as small nuclear ribonucleoproteins, spliceosome subunits, and other...
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 Breast Cancer Gene-Expression Miner v4.7 database analysis. f TCGA and GTEx data-based box of NSrp70 expression in tumor-adjacent/tumor and healthy breast tissue. g Distant metastasis-free survival (DMFS) analysis of NSrp70 expression in breast cancer samples based on Breast Cancer Gene-Expression Miner v4.7 database analysis. h Overall survival (OS) analysis of NSrp70 expression in breast cancer samples based on Breast Cancer Gene-Expression Miner v4.7 database analysis.
non-small nuclear ribonucleoprotein. Therefore, these structures are involved in an elaborate regulation of gene expression [9]. Based on this structure, NSrp70 belongs to the SR protein family and plays an important role in alternative splicing [8]. 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. By analyzing the expression of NSrp70 in different types of breast cancer in Breast Cancer Gene-Expression Miner v4.7, we found that NSrp70 may be relatively lower in TNBC tissue than in other...
types of breast cancer. The prognosis of NSrp70 analysis indicated that lower NSrp70 expression was associated with poor prognosis in breast cancer. Furthermore, NSrp70 was downregulated in the highly metastatic breast cell line, suggesting that it was 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, we analyzed the differential expression of 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). The gene ontology analysis results of downregulated proteins in both highly metastatic cell lines and the parental cell line are listed in Supplementary Table S1. By combining the expression fold-change, their expression in tissue, prognostic values in the online database, and their function by overexpression of the candidate genes, 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 Supplementary Fig. S1a, b). 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 was generally detected in the luminal breast cancer cells (MCF-7, ZR-75-30), which are considered weakly metastatic cell lines, but it was the lower expression in basal-like breast cancer cells (BT-549, Hs-578T). In particular, in the highly metastatic cell lines (MDA-MB-231HM and MDA-MB-231BO), NSrp70 expression was weaker (Fig. 1c, d). In addition, analysis based on Breast Cancer Gene-Expression Miner v4.7 revealed that NSrp70 was downregulated in different subtypes of breast cancer and relatively lower in TNBC when compared with normal tissue according to PAM50 subtypes (Fig. 1e). By the TCGA database analysis, lower NSrp70 was found in tumor tissue when compared with normal breast tissue (Fig. 1f). Furthermore, it indicated that lower NSrp70 expression was associated with poor prognosis (Fig. 1g, h). TCGA and online Kaplan–Meier plot analysis showed expression of NSrp70 was various in different cancer and its function may depend on cancer types (Supplementary Fig. S2). 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. Furthermore, two guide RNAs targeting NSrp70 significantly reduced the expression of NSrp70 in MDA-MB-231, Hs-578T, and BT-549 cell lines (Fig. 2a, 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, d). Furthermore, 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 [13]. The cells were then delivered via tail intravenous or 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 (Supplementary Fig. S3a, 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 of CRIM1 and PAPBC1 with PSI was less than 0.1. Then the effect of NSrp70 on other 8 genes was confirmed with real-time polymerase chain reaction (RT-PCR) (Supplementary Figs. S7 and S8). The splicing variants of MTMR2 and MTMR3 were significantly increased lung metastasis in vivo (Fig. 2h, i). These observations indicated 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 [13]. The cells were then delivered via tail intravenous or 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 (Supplementary Fig. S3a, b). This indicated that NSrp70 has no significant effect on breast cancer cell proliferation.

SE was the most common alternative splicing subtype in the two NSrp70-overexpressing cell lines analyzed (Fig. 3a and Supplementary Fig. S4). The scatter plots for the other four subtypes are shown in Supplementary Fig. S5. 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 (Supplementary Fig. S6). After filtering the data by Bayes factor of ≥10 events, 10 potential genes were identified. Their alternative splicing events with a sashimi-style plot showed that the impact on alternative splicing of CRIM1 and PAPBC1 with ΔPSI was less than 10%. Then the effect of NSrp70 on other 8 genes was confirmed with real-time polymerase chain reaction (RT-PCR) (Supplementary Figs. S7 and S8). The splicing variants of MTMR2 and MTMR3 were significantly enriched when NSrp70 was knocked down or overexpressed. The splicing variants of CD97, CKL1, RHOC, and
WHSC1 showed no significant change when NSrp70 was overexpressed (Supplementary Figs. S7 and S8). Only the splicing variants of NUMBv (variants generated specifically after exon 12 skipped) and CD44 (variants generated specifically after exon 6–13 skipped) were significantly enriched after NSrp70 knockdown while they were decreased after NSrp70 overexpression (Fig. 3b and Supplementary Fig. S9a). Previous study had shown the effect of NSrp70 on CD44 [14]; therefore, NUMB might be another candidate downstream gene of NSrp70.

To further explore whether NSrp70 interacted with its target mRNA, we used 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 [14] (Fig. 3c). Pre-mRNAs of NUMB and CD44 were efficiently immunoprecipitated from NSrp70-WT cells, but not from NSrp70-M15 and NSrp70-RS1M cells (Fig. 3c and Supplementary Fig. S9b). We also performed RNA pull-down assays to confirm the physical binding of the target mRNA and NSrp70.
protein (Fig. 3d and Supplementary Fig. S9c). These observations suggested that NSrp70 could bind to specific pre-mRNAs to regulate their splicing.

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 prevent TGFβ-induced EMT**

The above experiments indicated that NSrp70 regulated alternative splicing of NUMB. To investigate 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, TGFβ 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, TGFβ signaling pathway was significantly deactivated. Depletion of NSrp70 elevated the expression of TBR1 and phosphorylation of Smad3 (p-Smad3) in MDA-MB-231 cells. By contrast, overexpression of NSrp70 in MDA-MB-231HM cells downregulated TBR1 and p-Smad3 protein levels. However, the expression of TBR2 and Smad3 was not affected by the NSrp70 status (Fig. 4b).

TGFβ stimulates the EMT in breast cancer cells [15]. To investigate whether NSrp70 mediated cell EMT by affecting the TGFβ signal pathway, we evaluated EMT marker expression in MDA-MB-231HM and Hs-578T cells with overexpressed NSrp70 (Fig. 4c, d), and in MDA-MB-231 and BT-549 cells with downregulated NSrp70 expression (Fig. 4e). Western blotting analysis revealed that overexpression of NSrp70 resulted in inhibition TGFβ signaling and EMT with higher E-cadherin expression, but lower fibronectin and vimentin expression (Fig. 4c, d). By contrast, NSrp70 deficiency had an opposite effect which induced activation of TGFβ signaling and EMT with lower E-cadherin expression but higher N-cadherin, vimentin, and fibronectin expression (Fig. 4e).

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. The relative recovery of NUMB mRNAs with NSrp70 or truncation variants (bottom and right panel).

**NSrp70 accelerates TBR1 degradation by NUMB-mediated lysosome pathway**

We observed a significant decrease in TBR1 protein levels following NSrp70 overexpression (Fig. 4b). However, the alternation of TBR1 mRNA levels was modest and not significant (Supplementary Fig. S11). These observations suggest that NSrp70 predominantly suppresses TBR1 expression on a post-transcriptional level. To check whether NSrp70 was associated with the TBR1 protein degradation pathway, MDA-MB-231HM cells were incubated with cycloheximide (CHX) which inhibited protein synthesis. TBR1 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, overexpression of NUMBv in MDA-MB-231 cells restrained TBR1 degradation after the treatment of CHX (Fig. 5b).

It has been reported that NUMB is cargo-selective endocytic adaptors and may regulate lysosome maturation [16] and recent studies showed that NUMB was a crucial regulator of lysosomal function and autophagic progression [17, 18]. To identify whether TBR1 is a new cargo of NUMB, we performed IP assay. The results showed that both NUMBv and NUMBwt interact with TBR1 (Fig. 5c).

**Interestingly, TBR1 degradation was inhibited when the cells were treated with the lysosome pathway inhibitor bfa1 in both NSrp70 and NUMBv overexpression cells (Fig. 5d, e). However, the NUMBwt did not show a significant effect under the treatment of bfa1 (Fig. 5f). We evaluated the co-localization of TBR1 with early endosome antigen 1 (EEA1), member RAS oncogene family 7**
(RAB7), and RAB11 utilizing immunofluorescence. We observed that the co-localization of TβR1 and EEA1, and RAB7 was more pronounced in NSrp70 knockdown cell lines. Meanwhile, the co-localization of TβR1 and RAB11 was not greatly affected (Fig. 5g).

Collectively, these findings suggested that depleting NSrp70 potentiated TGFβ/SMAD signaling by inhibiting the TβR1 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 TCGA and Breast Cancer Gene-Expression Miner v4.7 database and found that NSrp70 was downregulated in breast cancer (Fig. 1e, f).

To evaluate the clinical importance of NSrp70 in breast cancer, we performed an immunohistochemical analysis of tissue microarrays.
The characteristics of the cohort are summarized in Table 1. Because were included in the subsequent analysis. The clinicopathological data were missing for two patients; hence, 248 cases (TMAs) containing samples from 250 breast cancer patients.

Knockdown of NSrp70 promotes TGFβ-induced EMT. a Kyoto Encyclopedia of Genes and Genomes signaling pathway (KEGG) enrichment histogram. MDA-MB-231HM (pCDH vs. Nsrp70) and Hs-S787 (pCDH vs. Nsrp70) group analyses of transcriptome sequence data were performed. b Western blotting evaluation of the expression of TjR1/2 and key proteins of the TGFβ signaling in Nsrp70-overexpression and knockdown cell lines. c, d Expression of TGFβ signaling and EMT-related proteins in Nsrp70-overexpressing MDA-MB-231HM and Hs-S787 cells after treatment with TGFβ signaling agonist TGFβ1. e Expression of TGFβ signaling and EMT-related proteins in Nsrp70 knockdown BT-549 and MDA-MB-231 cells after treatment with GW788388. F Transwell assay to detect changes in cell invasion ability of MDA-MB-231 and Hs-S787 cells with a stably knocked down Nsrp70 after treatment with a TGFβ signaling suppressor GW788388 (magnification, ×100). g Overexpression of NUMBβ, and NUMBwt in MDA-MB-231 cells. h Transwell assay to detect changes in cell invasion ability of the Nsrp70, NUMBv, NUMBwt, Nsrp70, and NUMBv co-overexpression in MDA-MB-231 cells (magnification, ×100). i Expression of TGFβ signaling and EMT-related proteins in NUMBv, and NUMBwt-overexpressing MDA-MB-231 cell lines. j Expression of TGFβ signaling and EMT-related proteins in Nsrp70-overexpressing and NUMBv co-overexpressing MDA-MB-231HM cell lines. wt: wild type; v- alternative splicing variants.

(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 = 3.978, 95% confidence interval: 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**

Alternative splicing is the main source of protein diversity in human genes. Whereas alternative splicing is tightly regulated in normal cells, dysregulation of the splicing factors, or mutations in the cis-regulatory elements can lead to the production of proteins with aberrant functions in cancer cells [19]. Previous studies indicated that 27 types of cancers were associated with more than 15,000 cancer-specific splice variants [20]. Alternative splicing has been reported to influence breast cancer by participating in the regulation of therapeutic resistance [21]. Therefore, targeting alternative splicing will be a novel prevention and therapeutic strategy for breast cancer.

Nsrp70 belongs to the SR-related protein family and is recognized as an important splicing factor [14, 22]. Nsrp70 has been reported to play an important role in schizophrenia [10], embryonic development [11], and adult acute leukemia [12]. However, there is no evidence indicating its function in carcinoma. In this study, we found that Nsrp70 was a suppressor in TNBC in vivo and in vitro. Besides, Nsrp70 was related to a better prognosis. Therefore, a better understanding of its target genes will help to modulate the splicing and develop a new approach to cancer therapy. Our study showed that Nsrp70 regulated the SE splicing of the downstream genes NUMB and suppressed TNBC metastasis. Modulating the splicing may present a new approach to cancer therapy.

In the current study, we showed that Nsrp70 regulated the alternative splicing of CD44 and NUMB. CD44 is a cell surface protein that modulates cellular signaling by forming coreceptor complexes with various receptor tyrosine kinases [23, 24] and it has several transcriptional variants [25, 26]. It has been reported that Nsrp70 modulated CD44 alternative splicing by forming CD44v5 [14]. In this study, we identified that Nsrp70 regulated the expression of variant CD44v10, generated by exon 6–13 skipping, and found it may promote cell migration (Supplementary Fig. S9). This founding was in line with the previous reports that CD44v10 was reported to be closely associated with the cytotoxicity of leukemia hematopoietic stem cells [27] and with the histological grade and survival of renal cell carcinoma [28].

Previous studies indicated that NUMB was an adaptor protein inhibiting Notch signaling by leading to Notch degradation through ubiquitylation and lysosomal degradation [29]. NUMB positively regulated autophagic flux by regulating lysosomal function [18]. In addition, the alternative splicing of NUMB was found to be regulated by the activated Notch signaling pathway, which affected the proliferation of cervical cancer cells [30]. Besides, NUMB exon 9 inclusion was correlated with reduced levels of NUMB protein expression and activation of the Notch signaling pathway [31], whereas exon 12 inclusion promoted the proliferation and transformation of lung cancer cells [32]. Furthermore, NUMB was recently shown to be a new player in EMT. Knockdown of NUMB reduced the cell-cell adhesion, resulting in cell metastasis and proliferation [33]. Our study showed that Nsrp70 regulated NUMB alternative splicing by skipping exon 12. Furthermore, we found that NUMBwt suppressed while NUMBv enhanced cell migration. NUMBv inhibited the degradation of TβR1 via the lysosome pathway which promoted TGFβ-induced EMT ultimately.

Previous studies reported that Nsrp70 bound to splicing factor SRSF1 and SRSF2 via an arginine/serine-like region and counteracted their alternative splicing activity [14]. The RS region of Nsrp70 contains the binding sites of both SR proteins and target mRNA. In accordance with the previous study, our study showed that Nsrp70 antagonized the splicing-promoting effect of SRSF1, and both Nsrp70 and SRSF1 affected splicing in a dose-dependent manner. In addition, Nsrp70 mutation with SR deletion showed a weaker combination with NUMB pre-mRNA. In addition to SRSF1, Nsrp70 works in tandem with SRSF1 and SRSF2 [14, 22]. The complexity of alternative splicing is complicated and the mechanism needs to be further explored. Also, we noticed that the expression of 15 splicing factors also changed when Nsrp70 was overexpressed to some extent (Supplementary Table S2). The result indicated that Nsrp70 may also affect the alternative splicing of NUMB by regulating these splicing factors in addition to directly targeting the pre-mRNA of NUMB to inhibit alternative splicing. This hypothesis needs to be further demonstrated.

Taken together, our study demonstrated that Nsrp70 was a novel tumor metastatic suppressor. Mechanistically, Nsrp70 inhibited the SE splicing of the downstream gene NUMB, promoting the TβR1 degradation via lysosome pathways. This transformation inhibited TGFβ/SMAD-mediated EMT, ultimately inhibiting the metastasis of breast cancer cells. Thus, our findings are expected to provide insights into the mechanism and therapy of cancer metastasis.

**MATERIALS AND METHODS**

**Ethics statement**

The present study was conducted according to the documented standards of the Ethics Committee of Fudan University Shanghai Cancer Center
Fig. 5  *NSrp70* levels correlate with *TβR1* degradation. a, b *NSrp70*-overexpressing MDA-MB-231HM and NUMBv-overexpressing cells were incubated with 50 μM CHX for 0–24 h then the expression of *TβR1/2* was evaluated by western blotting. c Western blot analysis of immunoprecipitation assay with NUMBwt and NUMBv-overexpressing MDA-MB-231HM cells. d, e Western blot analysis of *TβR1* expression in *NSrp70*-overexpressing MDA-MB-231HM and NUMBv-overexpressing cells following incubation with 400 nM bfa1 for the indicated time periods. f 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. v: alternative splicing variants; wt: wild type. g Expression of *TβR1* and lysosome pathway-related indicators EEA1, Rab7, and Rab11 was detected by immunofluorescent in MDA-MB-231 cells with stably downregulated NSrp70 (magnification, ×600).
Furthermore, 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 high 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% CO2 incubator at 37 °C. Constructs were transfected by Polybrene (Sigma-Aldrich, St. Louis, MO) according to the instructions of the manufacturer’s instructions. Then 48 h 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–III invasive ductal carcinoma collected randomly at the Department of Breast Surgery in 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–III. The clinical data included age, menstrual status, histological grade, tumor size, lymph node status, ER, PR, HER-2 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

TMAs were obtained from archived formalin-fixed, paraffin wax-embedded carcinoma samples from patients in FDUSCC. IHC data were analyzed as previously described [34]. Patient information is reported in a previous paper from the authors’ laboratory [35]. The IHC variables were scored as described elsewhere [36]. 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 RT-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 S3.

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. NSrp70 guide

| Variables | NSrp70 | |
|-----------|--------|--------|
|           | Low expression | High expression | Total |
|           | N=117 | N=131 | N=248 |
| 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 | | | 9 | 7.6 | 3 | 2.3 | 12 | 4.8 |
| 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 |
| Unknown | 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 |

LN lymph nodes, ER estrogen receptor, PR progesterone receptor, HER-2 human epidermal growth factor receptor 2. Bold values indicates statistical significant p values (p < 0.05).
RNAs and the negative control were designed considering the CRISPR/Cas9 system [37]. The detailed primers we used are listed in the Supplementary methods.

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 primers used are shown in the Supplementary methods.

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 a medium containing FBS was added. After 24 h, the infected cells were subjected to selection with 2 μg/mL puromycin for 1 week.

**Western blot analysis**

The detailed protocol is presented in the Supplementary methods. The antibodies used in this study are listed in Supplementary Table S4.

**Transwell assays and kinetic wound-healing assay**

The detailed protocol is presented in the Supplementary methods.

**Immunofluorescence**

MDA-MB-231 stable cell lines grown on coverslips were fixed with 4% paraformaldehyde for 30 min at room temperature, permeabilized with 0.5% Triton X-100 for 5 min 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 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 μL of protein G-agarose beads (Roche). Next, the precleared supernatant was incubated with 50 μL 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-TGFβ1 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, and 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 [38]. The detailed protocol is presented in the Supplementary methods.

Two truncations of Nsrp70-M15 and RS1M were synthesized by Genewiz. The sequences of M15 and RS1M are provided in the Supplementary methods.

Primers used in RIP are listed in Supplementary methods.

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 [39]. 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 (HiScript™ T7 Quick High Tied RNA Synthesis kit EZ2050, 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% NP-40, 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 μL of streptavidin Dynabeads (Dynabeads™ M-280 streptavidin, Invitrogen, Thermo Fisher Scientific) for 20 min at 4 °C. Next, 20 μg/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 μg of biotin-labeled NUMB (full length). The mixture was then incubated for 1.5 h followed by the addition of 40 μL 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× 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. RNA sequences used are listed in the Supplementary methods.

**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 [40] 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 [41, 42]. 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 over-expression 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′-TGGCCAGAAGTAGAAGGGGAG-3′
- **Reverse primer**: 5′-CGTATGAGGAGTGATGTCG-3′

| Variables | Univariate analysis HR (95% CI) | p value | Multivariate analysis 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) | 0.03 |
| 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) | 0.012 |
| LN status | 0.388 (0.207–0.728) | 0.003 | 0.569 (0.288–1.125) | 0.105 |
| ER | 1.264 (0.676–2.263) | 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) | <0.001 |

LN lymph nodes, ER estrogen receptor, PR progesterone receptor, HER-2 human epidermal growth factor receptor 2, CI confidence interval. Bold values indicate statistically significant p values (p < 0.05).
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'-TGCTCTAGAAGTGCGGAGGGTGGTGAT-3'
Reverse primer: 5'-GGGATCCCTTAGAAGGACGAGGATC-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 × 10^5 cells labeled with GFP/fluorescein were injected into female BALB/c nude mice (N = 6) to study lung metastasis. After 4–6 weeks, the lungs were removed and metastatic foci were performed by Student’s t-test (two-tailed). The cumulative survival time was calculated by the Kaplan–Meier method (log-rank test). The univariate and multivariate analyses were based on Cox proportional hazards regression model. The results with p < 0.05 were considered statistically significant.

Quantification and statistical analysis
All data are reported as mean ± standard deviation 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 Student’s t-test (two-tailed). The cumulative survival time was calculated by the Kaplan–Meier method (log-rank test). The univariate and multivariate analyses were based on Cox proportional hazards regression model. The results with p < 0.05 were considered statistically significant.

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, Luos S, Kreshukova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008;456:467–70.
4. Venables JP. Aberrant and alternative splicing in cancer. Cancer Res. 2004;64:647–54.
5. David CJ, Manley JL. Alternative pre-mRNA splicing regulation in cancer: pathways and programs unihed. Genes Dev. 2010;24:2343–64.
6. Ladomery M. Aberrant alternative splicing is another hallmark of cancer. Int J Cell Biol. 2013;2013:66786.
7. Read A, Natrajan R. Splicing dysregulation as a driver of breast cancer. Endocr Rev. 2016;37:331–67.
8. Kim YD, Lee JY, Oh KM, Araki M, Araki K, Yamamura K, et al. NSrp70 is a novel 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;291:6169–81.
9. Imamichi T, Hitaka A, Inoue Y. The roles of TGF-beta signaling in carcinogenesis and breast cancer metastasis. Breast Cancer. 2012;19:118–24.
10. Pece S, Confolaniero S, P Pr, Di Fiore PP. NUMB-ing down cancer by more than just a NOTCH. Biochim Biophys Acta. 2011;1815:26–43.
11. Reithmeier A, Panizza E, Krumpel M, Orre LM, Branca RMM, Lehtio J, et al. 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.
12. Sun H, Liu Y, Zhang L, Shao X, Liu K, Ding Z, et al. Numb positively regulates autophagic flux via regulating lysosomal function. Biochem Biophys Res Commun. 2017;491:780–6.
13. Zhang J, Manley JL. Misregulation of pre-mRNA alternative splicing in cancer. Cancer Discov. 2013;3:1228–37.
14. Kahles A, Lehmann KV, Toussaint NC, Hüser M, Stark SG, Sachsendberg T, et al. Comprehensive analysis of alternative splicing across tumors from 8,705 patients. Cancer Cell. 2016;34:211–24.e216.
15. Lee SC, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. Nat Med. 2016;22:976–86.
16. Tripathi V, Song DY, Zong X, Shesvov SP, Hearn S, Fu XD, et al. SRSF1 regulates the assembly of pre-mRNA processing factors in nuclear speckles. Mol Biol Cell. 2012;23:3694–706.
17. Cheng C, Yaffe MB, Sharp PA. A positive feedback loop couples Rac activation and CD44 alternative splicing. Genes Dev. 2006;20:1715–20.
18. Sherman LS, Rizvi TA, Karyala S, Raper N. CD44 enhances neutregulin signaling by Schwann cells. J Cell Biol. 2000;150:1071–84.
19. Bell MV, Cowper AE, Lefranc MP, Bell J, Scroeton GR. Influence of intron length on alternative splicing of CD44. Mol Cell Biol. 1998;18:5930–41.
20. Loh TJ, Moon H, Cho S, Jung DW, Hong SE, Kim DH, et al. SC35 promotes splicing of the CS-V6-C6 isoform of CD44 pre-mRNA. Oncol Rep. 2014;31:273–9.
21. Enh U, Megapetsch AP, Gia 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.
22. Li N, Tsui M, Kanda K, Murakami Y, Kanayama H, Kagawa S. Analysis of CD44 variant v10 expression and its prognostic value in renal cell carcinoma. BJU Int. 2000;85:114–8.
23. Kandachi R, Roegiers F. Endocytosis and control of Notch signaling. Curr Opin Cell Biol. 2012;24:534–40.
24. Santoloni E, Curi P, Salcini AE, Gagliani MC, Pellicci PG, Tachetti C, et al. Numb is an endocytic protein. J Cell Biol. 2000;151:1345–52.
25. Misquitta-AlI CM, Cheng E, O’Hanlon D, Liu N, McGlade CJ, Tsao MS, et al. Global profiling and molecular characterization of alternative splicing events mis-regulated in lung cancer. Mol Cell Biol. 2011;31:138–50.
26. Dong FY, Xu F, Wei WJ, Luo YG, Heiner M, Cao LJ, et al. The RNA-binding protein QKI suppresses cancer-associated aberrant splicing. PLoS Genet. 2014;10:e1004289.
27. Wang Z, Li SS. Numb: a new player in EMT. Cell Adhes Migr. 2010;4:176–9.
28. Xiao X, Wang L, Wei P, Chi Y, Li D, Wang Q, et al. Role of MUC20 overexpression as a predictor of recurrence and poor outcome in colorectal cancer. J Transl Med. 2013;11:151.
29. Sun HF, Yang XL, Zhao Y, Tian Q, Chen MT, Zhao YY, et al. Loss of TMEM126A contributes 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;291:6169–81.
30. Chatterjee R, Stass S, Ohno S, Iversen LB, Goldberg M. Use of minigene systems to dissect alternative splicing elements. Oncol Rep. 2014;31:273–8.
31. Cooper TA. Use of minigene systems to dissect alternative splicing elements. Oncol Rep. 2014;31:273–8.
32. Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, et al. Circular intronic long noncoding RNAs. Mol Cell. 2013;51:792–806.
33. Wang Z, Li SS. Numb: a new player in EMT. Cell Adhes Migr. 2010;4:176–9.
34. Xiao X, Wang L, Wei P, Chi Y, Li D, Wang Q, et al. Role of MUC20 overexpression as a predictor of recurrence and poor outcome in colorectal cancer. J Transl Med. 2013;11:151.
35. Sun HF, Yang XL, Zhao Y, Tian Q, Chen MT, Zhao YY, et al. Loss of TMEM126A promotes extracellular matrix remodeling, epithelial-to-mesenchymal transition, and breast cancer metastasis by regulating mitochondrial retrograde signaling. Cancer Lett. 2019;440:189–201.
36. Zheng YZ, Cao ZG, Hu Y, Zhao 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;145:349–58.
37. Hu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014;157:1262–78.
38. Townley-Tilson WH, Pendergrass SA, Marzluff WF, Whitmore GF. Global analysis of mRNAs bound to the histone stem-loop binding protein. RNA. 2006;12:1853–67.
39. Katz Y, Wang ET, Airoldi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat Methods. 2010;7:1009–15.
40. Cooper TA. Use of minigene systems to dissect alternative splicing elements. Methods. 2005;37:331–40.
41. Stoss O, Stolov P, Hartmann AM, Naylor Q, Stamm S. The in vivo minigene approach to analyze tissue-specific splicing. Brain Res Brain Res Protoc. 1999;4:383–94.

ACKNOWLEDGEMENTS
This work was supported by a 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). We gratefully acknowledge the help of professor Igor Vladimirovich Reshehtov (J.M. Sechenov First
Moscow State Medical University of the Ministry of Health of the Russian Federation for the support of clinical statistical analysis.

AUTHOR CONTRIBUTIONS
WJ, Yang Z, and HS contributed to the study design. Yang Z, Yuanyuan Z, and QL contributed to the methodology. YL and YH performed statistical analysis and interpretation. Yang Z wrote the original draft. WJ and HS revised the manuscript. WJ supervised the study and provided financial support.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41388-022-02349-z.

Correspondence and requests for materials should be addressed to Hefen Sun or Wei Jin.

Reprints and permission information is available at http://www.nature.com/reprints

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