Identification of therapeutic targets and prognostic biomarkers from the hnRNP family in invasive breast carcinoma

Jiawei Zhou1,2, Yugang Guo1, Zheng Huo1, Yuxin Xing1,2, Jintao Fang1, Guohui Ma1, Qinghui Han1,2, Mengzhen Wang1,2, Qian Xu1

1Henan Provincial Engineering Laboratory of Insects Bio-reactor, Nanyang Normal University, NanYang 473000, China
2School of Chemistry and Pharmaceutical Engineering, Nanyang Normal University, NanYang 473000, China

Correspondence to: Qian Xu; email: xuqian7666@nynu.edu.cn
Keywords: heterogeneous nuclear ribonucleoproteins, invasive breast carcinoma, metastasis, prognostic markers
Received: March 24, 2020 Accepted: November 13, 2020 Published: January 20, 2021

ABSTRACT

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are RNA-binding proteins that are reported to play a crucial role in the pathogenic process of multiple malignancies. However, their expression patterns, clinical application significance and prognostic values in invasive breast carcinoma (BRCA) remain unknown. In this study, we investigated hnRNP family members in BRCA using accumulated data from Oncomine 4.5, UALCAN portal and other available databases. We explored the expression and prognostic value level of hnRNPs in BRCA. We further analyzed their association with the clinicopathological features of BRCA patients. Subsequently, we calculated the alteration frequency of hnRNPs, constructed the interaction network of hnRNPs, and examined the potential coexpression genes of hnRNPs, revealing that HNRNPU and SYNCRIPT are the core molecular genes requiring further investigation for BRCA. We validated the immunohistochemistry (IHC) pattern to simulate clinical applications based on pathology. Cell function experiments conducted in vitro indicated that HNRNPU can promote epithelial-mesenchymal transition, functionally stimulating the invasion capacity and inhibiting the viability of invasive BRCA cells. In summary, our systematic analysis demonstrated that HNRNPU was the key molecule that played a fundamental role in BRCA metastasis, which may facilitate the development of new diagnostic and prognostic markers for the analysis of BRCA progression.

INTRODUCTION

Breast carcinoma (BRCA) is the leading cause of cancer-associated mortality in women worldwide [1]. The pathogenesis of BRCA is extremely complex, reflecting the diversity in clinical subcategories at the cellular level [1, 2]. The most common subtype of BRCA is invasive ductal carcinoma (IDC), accounting for approximately 60%–75% of all BRCA cases. On the other hand, invasive lobular carcinoma (ILC) accounts for 5%–15% of all BRCA cases and is the second most common type of invasive BRCA [3]. Despite recent advances in BRCA therapy, tumor metastasis remains the leading cause of mortality in patients with BRCA. Therefore, it is imperative to identify prognostic markers and potential drug targets as well as understand the mechanisms underlying BRCA development in order to improve prognosis and promote individualized treatment.

Heterogeneous nuclear ribonucleoproteins (hnRNPs) represent a large family of RNA-binding proteins with various key cellular functions. The hnRNP family members are attracting increasing attention with respect to their association with cancer occurrence and progression [4, 5]. Cellular functions are dysregulated during tumorigenesis, including alternative splicing and translational and RNA processing [5–8]. Accumulating...
experimental evidence indicates that the hnRNP family members play pivotal roles in multiple cancers. For example, the increased expression of HNRNPA2B1 leads to tumor suppression in multiple cancers, including pancreatic ductal adenocarcinoma, ovarian cancer, hepatocellular carcinoma, lung cancer, melanoma, glioblastoma, and prostate cancer [9–15]. HNRNPC is negatively associated with the overall survival of patients with advanced gastric cancer treated with 5FU-based drugs [16]. HNRNPLL has been reported to modulate the alternative splicing of CD44 during epithelial–mesenchymal transition (EMT) in colorectal cancer metastasis [17]. HNRNPM is frequently associated with high-grade human breast tumor and promotes BRCA metastasis [18]. However, the different expression levels and the specific functions of individual hnRNPs members in BRCA, and association between hnRNPs and the clinicopathological features of BRCA remain unelucidated [19]. In addition, how and when hnRNPs are activated to interact and functionally complement each other remain unknown. Therefore, an in-depth investigation and analysis into the potential roles of hnRNPs in BRCA are warranted.

In the present study, we attempted to explore and analyze the expression patterns and clinical significance of hnRNPs in BRCA using integrated bioinformatics. In addition, we sought to determine whether hnRNPs can predict the progression and prognosis of patients with BRCA. We hope that this study will serve as a meaningful reference for understanding the roles of hnRNPs in BRCA as well as for improving the effectiveness of treatment and prognostic accuracy of patients with BRCA.

RESULTS

Expression levels of the hnRNP family members in patients with BRCA

The mRNA expression patterns of the hnRNPs were determined using the Oncomine and The Cancer Genome Atlas (TCGA) databases. Oncomine analysis revealed that there were significant changes in hnRNPs expression at the transcriptional level between different types of cancer and normal tissues following the threshold search criteria (P value < 0.05; fold change > 2, and top 10% gene rank). The significant unique analyses between cancer and normal tissues that met our selection criteria in pan-cancer analysis are shown in Figure 1A and Supplementary Table 1. HNRNPA0, HNRNPA1, HNRNPA2B1, PCBP1, HNRNPF, HNRNPK, HNRNPL, HNRNPM, SYNCRIP, and HNRNPU were significantly upregulated in the two main subtypes of BRCA (IDC and ILC) compared with in normal tissues in the datasets by Gluck et al [20], Radvanyi et al [21], Ma et al [22] and Turashvili et al [23]. As shown in Figure 1B, in the TCGA cohort, HNRNPA0, HNRNPA2B1, HNRNPC, HNRNPD, PCBP1, HNRNPF, HNRNPK, HNRNPL, HNRNPM, PTBP1, SYNCRIP, HNRNPR, and HNRNPU were found to be upregulated in 1,093 invasive BRCA tissue specimens compared with in 112 normal tissue specimens from the TCGA cohort (P < 0.05). The results of Radvanyi et al showed that the mRNA expression levels of HNRNPA1, HNRNPL, and HNRNPR were notably increased in invasive mixed BRCA tissues than in normal tissues. Radvanyi et al and Ma et al revealed that HNRNPA, PCBP1, HNRNPL, and SYNCRIP were upregulated in ductal BRCA tissues in situ than in normal tissues. However, Finak et al [24] reported significantly lower expression levels of HNRNPA0, HNRNPA1, HNRNPA2B1, HNRNPC, HNRNPD, PCBP1, HNRNPF, RBMX, HNRNPK, and HNRNPU in BRCA-associated stroma than in normal tissues. The results of Curtis from the European Genome-phenome Archive demonstrated that HNRNPU was downregulated in invasive BRCA tissues than in normal tissues. These findings suggest that the expression of hnRNP family members is associated with patients in different areas or different histological types of BRCA.

Correlation of hnRNP mRNA expression with individual stages of cancer, lymph node metastases, and pathological subtypes in TCGA-BRCA

We evaluated the relationship between the mRNA expression levels of hnRNPs and the different clinicopathological subgroups of patients in the TCGA-BRCA cohort. As shown in Figure 2, significant differences between the groups were assessed to generate P values and then illustrated using heat maps. These data indicated that the upregulated mRNA expression levels of HNRNPA0, HNRNPA2B1, HNRNPC, HNRNPD, PCBP1, HNRNPF, PTBP1, HNRNPK, HNRNPL, HNRNPM, SYNCRIP, HNRNPR, and HNRNPU were significantly associated with high cancer stages and lymph node metastases (LNM) with few exceptions. However, the mRNA expression level of HNRNPA1 was not associated with cancer stages and LNM in patients with BRCA (P > 0.05; Figure 2A, 2B). Except for HNRNPA1 and RBMX, the expression levels of the remaining members were higher in IDC tissues than in normal tissues (P < 0.0001; Figure 2C). In addition, except for HNRNPA1, HNRNPD, and RBMX, the expression levels of the remaining 12 members were higher in ILC tissues than in normal tissues (P < 0.05; Figure 2C). These results suggest that the mRNA expression levels of HNRNPA0, HNRNPA1, HNRNPA2B1, HNRNPC, HNRNPD, HNRNPF, HNRNPK, HNRNPL,
SYNCRIP, HNRNPR, and HNRNPU are markedly higher in IDC tissues than in ILC tissues ($P < 0.001$; Figure 2C), indicating that hnRNPs play significant roles in the tumorigenesis and progression of BRCA, particularly IDC and ILC. However, owing to the limited sample size of some histological groups (mixed, other, mucinous, medullary, or metaplastic), we did not present an in-depth analysis of their subgroup comparison data.

**Prognostic value of hnRNPs in BRCA**

To date, the potential prognostic value of only few members of the hnRNP family have been investigated...
Figure 2. The relationship between the mRNA expression of hnRNPs and tumor stage, nodal metastasis status, and pathological subtypes of patients with TCGA-BRCA. (A) Heat map displaying the correlation between the expression of hnRNPs and tumor stage. The boxplot shows the correlation between HNRNP expression and tumor stages. (B) Heat map displaying the association between the expression of hnRNP and nodal metastasis status. The boxplot shows the association between HNRNP expression and the nodal metastasis status. (C) Heat map displaying the correlation between the expression of hnRNPs and tumor pathological subtypes. The boxplot shows the association between HNRNP expression and tumor pathological subtypes. Gray represents *P < 0.05, blue represents **P < 0.01, green represents ***P < 0.001, red represents ****P < 0.0001, and white represents no significant difference.
in certain cohorts. Therefore, to analyze the independent prognostic value of the hnRNP family members, we analyzed the prognostic role of 15 hnRNP family members in patients with BRCA using the Kaplan–Meier (KM) method. As shown in Figure 3A, 3B, 10 of the 15 tested genes significantly correlated with OS ($P < 0.05$). Among these genes, HNRNPD, HNRNPL, and SYNCRIP were risky genes, with a hazard ratio (HR) of >1 and log-rank $P$ value of <0.05. In contrast, HNRNPA0, HNRNPA2B1, HNRNPC, PCBP1, HNRNPF, RBMX, and HNRNPR were protective genes, with an HR of <1 and log-rank $P$ value of <0.05. In addition, the mRNA expression levels of HNRNPA0, HNRNPA2B1, HNRNPC, PCBP1, RBMX, PTBP1, HNRNPK, HNRNPL, SYNCRIP, HNRNPR, and HNRNPU were associated with recurrence-free survival (RFS) in patients with BRCA ($P < 0.05$; Figure 3C). In particular, among these 11 genes, increased mRNA expression levels of HNRNPA0, PTBP1, HNRNPK, and SYNCRIP were associated with poor RFS in patients with BRCA under the selected best cutoff value ($P < 0.05$; Figure 3D).

Despite recent advances in BRCA therapy, metastasis remains the leading cause of mortality in patients with BRCA. Subgroup analysis revealed that the increased expression of HNRNPU significantly indicated a poor OS in patients with BRCA with LNM (HR, 2.11; $P = 0.029$).

![Figure 3. Correlation between hnRNP mRNA expression and clinical prognosis in patients with BRCA. (A) Forest plots showing the correlation between the expression of hnRNPs and patient OS, as examined via Kaplan–Meier analysis. (B) High hnRNP D, hnRNP L, or SYNCRIP expression was significantly associated with poor OS. (C) Forest plots showing the association between the mRNA expression levels of hnRNPs and RFS. (D) High expression of HNRNPA0, PTBP1, HNRNPK, and SYNCRIP was significantly associated with poor RFS.](image-url)
However, the increased expression of HNRNPA0, HNRNPC, HNRNPF, and HNRNPK predicted improved survival in patients with BRCA with LNM (HR, <1; P < 0.05). RFS data analysis revealed that the increased expression levels of HNRNPA0, HNRNPA2B1, HNRNPC, HNRNPD, HNRNPF, HNRNPK, HNRNPR, and HNRNPU were associated with improved RFS in patients with BRCA with LNM (HR, <1; P < 0.05). In contrast, the increased expression of PTBP1 (HR, 1.25; P = 0.032), HNRNPL (HR, 1.73; P < .0001), HNRNPM (HR, 1.42; P = 0.01), and SYNCRIP (HR, 1.84; P < .0001) were associated with poor RFS in patients with BRCA with LNM (Supplementary Table 2).

Protein expression levels of hnRNPs in BRCA and their association with the clinicopathological features of patients with BRCA

Because HNRNPA1 and RBMX were not altered in 1,093 patients with invasive BRCA compared with in 112 normal tissues from the TCGA cohort, the mRNA expression level of PCBP1 is not associated with LNM in BRCA. Therefore, we further identified the protein expression levels of the other 12 genes in patients with BRCA using the Clinical Proteomic Tumor Analysis Consortium (CPTAC) database. Our results indicated that HNRNPA2B1, HNRNPC, HNRNPD, HNRNPF, PTBP1, HNRNPK, HNRNPL, HNRNPM, SYNCRIP, and HNRNPU were upregulated in BRCA tissues compared with in normal tissues (Supplementary Figure 1), which corroborated the mRNA expression results detailed in the previous section of this manuscript. We analyzed the association between the protein expression profile and clinicopathological data and found that increased protein expression levels of HNRNPA2B1, HNRNPC, HNRNPD, HNRNPF, PTBP1, HNRNPK, HNRNPL, HNRNPM, and HNRNPU were associated with high cancer stages (Figure 4A). Further, we found that the protein expression levels of HNRNPA2B1, HNRNPC, HNRNPD, HNRNPF, PTBP1, HNRNPK, HNRNPL, HNRNPM, SYNCRIP, and HNRNPU were

![Figure 4. Correlation between hnRNP protein expression, tumor stage, and pathological subtypes.](image)

(A) Heat map displaying the correlation between hnRNP protein expression and tumor stage. The boxplot shows the association between HNRNPU protein expression and tumor stage. (B) Heat map displaying the correlation between hnRNP protein expression and tumor pathological subtypes. The boxplot shows the association between HNRNPU protein expression and tumor pathological subtypes. Gray represents *P < 0.05, blue represents **P < 0.01, green represents ***P < 0.001, red represents ****P < 0.0001, and white represents no significant difference.
higher in IDC tissues than in normal tissues and that the protein expression levels of HNRNPC, HNRNPF, PTBP1, HNRNPL, and HNRNPM were higher in ILC tissues than in normal tissues. In addition, the expression levels of HNRNPC, HNRNPF, HNRNPL, HNRNPM, SYNCRIP, HNRNPR, and HNRNPU were significantly higher in IDC tissues than in ILC tissues (Figure 4B; P < 0.05).

Genetic mutation, interaction, and coexpression analyses of hnRNPs

The different mutation spectra seen in the geographical and ethnic populations can be used to identify the environmental exposure contributing toward BRCA development. We considered their possible roles in BRCA risk and their association with the type of cancer developed. Therefore, we calculated the alteration frequency of hnRNPs across the TCGA-BRCA cohort using the eBioPortal online tool. hnRNPs were altered in 595 samples from 1065 patients (60%) in the TCGA-BRCA cohort (Figure 5A). Notably, HNRNPU, HNRNPL, and HNRNPA2B1 were the three most frequently altered genes (14%, 2.9%, and 2.5%, respectively), and amplification, mutation, and deep deletion were the three most common types of mutations. We also confirmed the alteration frequency of hnRNPs in the two other datasets (Metastatic Breast Cancer Project Provisional and Breast Cancer; METABRIC) (Supplementary Figure 2A, 2B). Similar to the results of the TCGA-BRCA cohort, HNRNPU was the most frequently altered gene in these two datasets (13% and 23%, respectively). Next, we explored the potential coexpression genes of hnRNPs. Pearson’s correlation analysis revealed the following significant correlations: HNRNPA0 with PTBP1 and HNRNPM; HNRNPA2B1 with HNRNPC, HNRNPD, HNRNPL, and HNRNPU; HNRNPC with HNRNPF, HNRNPK, HNRNPL, and HNRNPU; HNRNPD with HNRNPL; HNRNPF with HNRNPK; PTBP1 with HNRNPL and HNRNPM; and HNRNPK with HNRNPR and HNRNPU (R > 0.4; N = 1093; Figure 5B). These results indicate that several hnRNPs interact with each other during the pathogenesis of BRCA. We conducted a protein–protein interaction (PPI) network analysis of hnRNPs to present the interaction among the 12 genes using STRING. As shown in Figure 5C, the purple connecting line shows the correlation among the genes when experimentally determined. We then constructed another network for hnRNPs and the most frequently altered coexpressed genes. The results showed that 43 genes, including BRCA2, TP53, ZNF532, CYLC1, COL12A1, SEMA5A, PLXNB3, and PCM1, were closely related with alterations in hnRNPs. Using Cytoscape, we found that of the 12 hnRNPs, HNRNPU, SYNCRIP, HNRNPL, and PTBP1 were the hub genes in the network (Figure 5D). Next, we studied the potential biological processes of the top 50 genes that positively and negatively correlated with HNRNPU mutation in BRCA using Linkomics (Supplementary Figure 2C–2E). The Gene Ontology (GO) term functional enrichment analyses of these genes are summarized (P < 0.01) in Supplementary Figure 2F–2H. Among the GO terms for biological processes, the top five terms were “protein binding,” “ion binding,” “molecular transducer activity,” “hydrolase activity,” and “transporter activity.”

Immunohistochemistry (IHC)-predicted patterns for potential clinical applications

Because HNRNPU and SYNCRIP were identified as the core molecular genes and SYNCRIP positively correlated with OS and HNRNPU positively correlated with tumor stage and metastasis, we further validated the IHC pattern from the Human Protein Atlas database to simulate clinical applications on the basis of pathology. Several IHC images of hnRNP proteins displayed a trend toward differential expression in BRCA subtype samples (IDC or ILC) compared with normal tissue samples. These results were in accordance with the transcriptional level of hnRNPs. SYNCRIP staining showed negative or low staining in the cytoplasm and cell membranes of normal tissues, and HNRNPU staining showed negative staining in the cytoplasm and cell membrane. In BRCA (IDC and ILC) tissues, strong and moderate staining patterns of SYNCRIP were clearly located in the nucleus, cytoplasm, and cell membrane. For HNRNPU, moderate and strong patterns were mainly located in the nucleus, cytoplasm, and cell membrane. This unique IHC staining of BRCA could help distinguish cancer tissues from normal tissues as well as predict clinical outcomes (Figure 6A, 6B). The relative staining intensity of SYNCRIP was negative (two cases) for normal tissues; weak (three cases), moderate (four cases), and strong (one case) for IDC tissues; and weak (one case) and moderate (two cases) for ILC tissues (Figure 6C). For HNRNPU staining, both cases were weak for normal tissue; moderate (two cases) and strong (six cases) for IDC tissues; and moderate (one case) and strong (three cases) for ILC tissues (Figure 6D).

HNRNPU promoted breast cancer cell proliferation, migration, and invasion

Because HNRNPU was discovered as a key gene in BRCA among the hnRNP family members and was associated with nodal metastasis in BRCA, we speculated that HNRNPU is involved in the regulation
of BRCA metastasis. To test this hypothesis, we first investigated the expression of HNRNPU in 32 breast IDC and adjacent normal tissues. The expression level of HNRNPU was higher in IDC tissues than in normal tissues (Figure 7A). We then examined the expression level of HNRNPU in a series of BRCA cell lines (MCF-7, ZR-75-1, MDA-MB-330, and MDA231-LM2) and a normal breast epithelial cell line (MCF-10A). The

![Figure 5](image_url)

**Figure 5. Genetic mutation, interaction, and coexpression analyses of hnRNPs in patients with BRCA.** (A) Summary of the alteration frequencies of nodal metastasis-related hnRNPs in patients with TCGA-BRCA (Breast Invasive Carcinoma, TCGA). The number of patients for each BRCA subtype is indicated at the top of each column and each type of alteration, including deletions (blue), amplification (red), multiple alterations (gray), or mutations (green), is shown using the cBioPortal. (B) Correlation plot based on Spearman’s correlation test results to show the correlation of gene expression among the 12 hnRNP family members in BRCA. (C) PPI among hnRNPs was derived from STRING. The purple connection indicates the protein interaction determined by the experiment and the blue connection indicates that the proteins are homologous. (D) The 12 hnRNP family members and coexpressed gene networks were screened using Cytoscape. Red and yellow represent the top 10 hub genes.
Figure 6. The IHC expression pattern of SYNCRIP and HNRNPU in normal breast tissues and BRCA. (A, B) Representative images are shown for strong, moderate, weak and negative expression of SYNCRIP and HNRNPU in normal breast tissues and BRCA (IDC and ILC). The black rectangle indicates a higher magnification of the indicated area in the image. (C) Bar chart of the IHC staining intensity of SYNCRIP for BRCA (a total of 13 cases). (D) Bar chart of the IHC staining intensity of hnRNPU for BRCA (a total of 14 cases).

Figure 7. HNRNPU expression is upregulated in BRCA cell lines and clinical tumor samples. Altering expression of HNRNPU proteins in clinical tumor samples (A) and four human BRCA cell lines (B), as determined via western blot analysis. Glyceraldehyde 3-phosphate dehydrogenase was used as the control. ZR-75-1 and MDA-MB-330 cells were transfected with shRNA and the mRNA (C), and protein (D) expression levels of HNRNPU were measured via reverse transcription quantitative real-time polymerase chain reaction and western blotting. *P < 0.05; **P < 0.01; ***P < 0.001.
expression level of HNRPNU was higher in the BRCA cell lines than in MCF-10A cells (Figure 7B). To investigate the functional roles of HNRPNU in invasive BRCA progression, HNRPNU expression was knocked down using shRNAs in ZR-75-1 and MDA-MB-330 cells (Figure 7C, 7D). MTT assay revealed that HNRPNU knockdown substantially reduced the proliferation ability of ZR-75-1 and MDA-MB-330 cells (Figure 7A, 8A). Moreover, the wound-healing and Matrigel invasion assays revealed that HNRPNU knockdown significantly inhibited the migration and invasion abilities of ZR-75-1 and MDA-MB-330 cells ($P < 0.05$, Figure 8C–8H). Further analysis of HNRPNU knockdown in ZR-75-1 and MDA-MB-330 cells revealed that HNRPNU depletion can increase the expression level of the epithelial marker E-cadherin and decrease the expression levels of the mesenchymal markers vimentin, N-cadherin, and matrix metalloproteinase 9, suggesting that HNRPNU is important for the EMT phenotype (Figure 8I–8L). Taken together, our results suggest that HNRPNU knockdown significantly inhibits the proliferation, migration, and invasion of BRCA cells in vitro. Therefore, we reached the tentative conclusion that HNRPNU is the key molecule that plays a fundamental role in BRCA metastasis.

**DISCUSSION**

BRCA is one of the most common malignancies in women, with 1.20 million cases reported annually in women worldwide [1]. Moreover, approximately one in eight women in the United States develops BRCA during the course of her lifetime. Its incidence is expected to increase globally [25]. Although several molecular factors have emerged as predictive markers of BRCA, it remains challenging to clarify the prognostic significance of those biomarkers.

Accumulated evidence indicates that hnRNPs play a variety of potential roles in the inhibition of apoptosis, angiogenesis, cell invasion, and EMT. The significant association among the hnRNP family members and clinical pathological features and survival outcomes of patients with cancer has been confirmed, indicating that hnRNPs are novel and promising cancer therapeutic targets and predictive biomarkers for treatment response and prognostic evaluation [16, 26–31]. In BRCA, the hnRNP family members, including HNRNPA2B1, HNRNPC, HNRNPH1, HNRNPK, HNRNP and PTBP1, function as oncogenes in BRCA progression [7, 8, 18, 32–36]. Liu et al reported that HNRNPA2B1 is a negative regulator of BRCA metastasis [36]. Hu et al...
revealed that HNRNPA2B1 knockdown inhibits cell proliferation, induces apoptosis, and prolongs the S phase of the cell cycle in MCF-7 and MDA-MB-231 cells via the STAT3 and ERK1/2 signaling pathways [37]. Wu et al showed that HNRNPC repression inhibits cell proliferation and tumor growth in BRCA cells [38]. Yang et al reported that HNRNPM promotes BRCA progression by regulating the axin/β-catenin signaling pathway [34]. In addition, HNRNP/H1 has been reported to result in the production of HER2 splice variants in BRCA [7]. PTBP1 has been shown to promote the growth of BRCA cells via the PTEN/Akt pathway and autophagy [8]. As there are few studies on other members of the hnRNP family, understanding the initial molecular events leading to BRCA development may provide opportunities for prophylactic intervention, improve our ability to predict the risk of BRCA, and offer advanced strategies for early detection. Therefore, in our study, we provided the first systemic analysis of the hnRNP family members in BRCA as therapeutic targets for the diagnosis, prognosis, and treatment of BRCA.

The results of our study showed that the mRNA expression levels of HNRNPA0, HNRNPA2B1, HNRNPC, HNRNPD, PCBP1, HNRNPF, HNRNPK, HNRNPL, HNRNPM, PTBP1, SYNCRIP, HNRNPR, and HNRNP were upregulated in patients with BRCA, whereas those of HNRNPA1 and RBMX were not altered in BRCA tissues compared with in normal tissues based on the TCGA dataset. Clinical correlation analysis revealed that increased mRNA and protein expression levels of HNRNPA2B1, HNRNPC, PCBP1, HNRNPF, HNRNPK, HNRNPL, HNRNPM, SYNCRIP, and HNRNP were observed in subgroups with high pathological stage and LNM. To validate the prognostic indicators, we found that increased mRNA expression levels of HNRNPD, HNRNP, and SYNCRIP were associated with poor OS rates and that increased mRNA expression levels of HNRNPA0, PCBP1, HNRNPK, and SYNCRIP were associated with poor RFS rates. Furthermore, subgroup analysis revealed that high expression of HNRNP in 32 patients with BRCA cancer and found that its expression was considerably higher in cancer tissues than in adjacent tissues. In addition, we found that HNRNP suppressed human BRCA metastasis by promoting EMT development, functionally stimulating the invasion ability and inhibiting the viability of invasive BRCA cell lines. These results suggest that HNRNP could be used as a potential biomarker to predict the prognosis of patients with IDC.

The progression of the invasion and migration of cancer cells is propelled by different functions and interaction of multiple genes at multiple steps. In recent years, EMT has been believed to be a key step in the invasion and metastasis of cancer cells [39]. Previous studies have shown that several hnRNP family members are implicated in the regulation of EMT in several human cancer types and are regarded as cancer metastasis suppressors [17, 18, 20, 40, 41]. We investigated the function of HNRNPU in regulating EMT. In our study, we first detected the expression of HNRNPU in 32 patients with BRCA cancer and found that its expression was considerably higher in cancer tissues than in adjacent tissues. In addition, we found that HNRNPU suppressed human BRCA metastasis by promoting EMT development, functionally stimulating the invasion ability and inhibiting the viability of invasive BRCA cell lines. These results suggest that HNRNPU could be used as a potential biomarker to predict the prognosis of patients with IDC.

Accumulating evidence suggests that RNA epigenetics has received increased attention in the past few years [4]. The hnRNP family members not only play an important role in alternative splicing but also participate in the regulation of microRNAs (miRNAs) or long non-coding RNAs [4, 42]. Studies have reported that HNRNPA2/B1 binds to m6A markers in a part of the mRNA transcripts and promotes primary mRNA processing. Half of the miRNAs regulated by m6A depend on HNRNPA2/B1 [42, 43]. Vascular endothelial growth factor A transcript expression in hypoxic conditions is regulated by HNRNP in competing with miRNA binding in the 3'-untranslated region of vascular endothelial growth factor A [44]. Similarly, PTBP1 binds to let-7 miRNA and human Argonaute 2, thereby altering their association with human mRNAs [45]. Many hnRNPs appear in the same complex, indicating that multiple hnRNPs have the same structure and function. Studies on the regulation of alternative splicing in cells may begin with the direction of the interaction mechanism between the low-complexity regions of hnRNPs and m6A-modified RNA as development progresses.

In conclusion, in this study, we aimed to understand the clinical significance of the hnRNP family in BRCA and the molecular mechanism based on big data. We hope that our findings offer new perspectives in future research and clinical applications for patients with BRCA. However, some potential limitations of this study should be discussed. Although we confirmed that hnRNPs play crucial roles in the pathogenesis of BRCA, additional experiments are warranted to confirm the association between hnRNPs and BRCA as new indicators in large patient cohorts.
MATERIALS AND METHODS

Ethics statement

The study has been conducted in accordance with the ethical standards and the Declaration of Helsinki as well as according to national and international guidelines. All procedures were approved by the Ethics Committee of the Institute of Clinical Pharmacology. Informed consent has been obtained.

Bioinformatics analysis

During the validation process, the transcription levels of hnRNP in different types of cancer were analyzed using the Oncomine 4.5 (https://www.oncomine.org) gene expression array dataset [46] and TIMER (https://cistrome.shinyapps.io/timer/) [47]. The mutation frequency of hnRNPs in a variety of cancers was analyzed based on TCGA, and a detailed diagram was constructed for hnRNPs in BRCA samples based on mutations. CNA data were analyzed using the online database cBioPortal (https://www.cbioportal.org/) [48]. To intuitively analyze the relationship between the mRNA and protein expression levels of hnRNPs and clinicopathology, we compared the patterns of the expression levels of hnRNPs in major BRCA types and downloaded a boxplot based on these patterns from the UALCAN web portal (http://ualcan.path.uab.edu/), which is an interactive Web portal that provides gene expression analyses based on TCGA and MET500 cohort data as well as protein expression analysis based on CPTAC data [49]. The comparison between hnRNPs expression and TCGA-BRCA tumor pathological subtypes, lymph metastasis status, and major subclasses were selected to create heat maps. Additionally, LinkedOmics (http://www.linkes.org/login.php) was used to analyze the mRNA expression levels of hnRNPs coexpression genes in the TCGA-BRCA cohort based on Pearson’s correlation coefficient [50]. All results were graphically presented in a heat map. To analyze OS and RFS, we divided the patients into two groups based on the best-performing threshold using KM plotter (https://www.kmplot.com). Next, the distribution of HR with 95% confidence intervals of OS and RFS across the hnRNPs was selected to create forest plots. Further, to examine the differences in the protein expression levels of hnRNPs between IDC, ILC, and normal tissues, we obtained the immunohistochemistry data of SYNCRIP and HNRNPU from the HPA database (https://www.proteinatlas.org/) [51].

Patient samples

All samples (IDC and adjacent normal tissues) were obtained from Nanyang Second People's Hospital (Nanyang, Henan, China) between October 2019 and June 2020 (n = 32). All pathological features were confirmed by experienced pathologists, and none of the patients received preoperative anticancer treatment.

Cell culture

The human breast cell lines MCF-10A, ZR-75-1, MCF-7, MDA231-LM2, and MDA-MB-330 were obtained from the Type Culture Collection of the Chinese Academy of Sciences (Shanghai, China). Cells were cultured in Dulbecco’s modified Eagle medium (Invitrogen, Shanghai, China) supplemented with 10% fetal bovine serum (Gibco, Logan, UT) and 100 U of penicillin and streptomycin at 37°C in a humidified atmosphere containing 5% CO2.

Cell transfection

The shRNA for the specific inhibition of HNRNPU expression and a negative control shRNA were synthesized by GenePharma Co., Ltd. (Shanghai, China). Exponentially growing untreated cells were plated 24 h before transfection. The plated cells were transfected with HNRNPU shRNA using 5 μL of Lipofectamine MAX following the manufacturer’s instructions (Invitrogen, Carlsbad, CA, USA). After treatment, the cells were processed for further analysis.

RNA extraction and quantitative real-time polymerase chain reaction

Total RNA was extracted from the cultured cells or tissue samples using the TRIzol reagent (Invitrogen) following the manufacturer’s instructions. First-strand cDNA synthesis was performed using the Prime Script RT Master Mix (TaKaRa, Dalian, China) and reverse transcription quantitative real-time polymerase chain reaction was performed using the SYBR Green PCR kit (TaKaRa) according to the manufacturer’s instructions. Glyceraldehyde 3-phosphate dehydrogenase was used as an endogenous control. The gene expression was calculated using the 2−ΔΔCt method. All data represent the average of three replicates.

Cell viability assays

One day after transfection, cells were re-seeded into 96-well plates. Cell viability was measured using a Cell Counting Kit 8 assay (Dojindo Laboratories, Kumamoto, Japan), according to the manufacturer's instructions. The absorbance value of each well was measured at 450 nm. Five replicates were set for each group and the whole experiment was repeated three times.
**Wound-healing assay**

Cells were seeded in 6-well plates and were transfected with the shRNA after 24 h. After culturing the cells for an appropriate amount of time, artificial wounds were made using a micropipette tip. The wounded monolayers were washed with phosphate-buffered saline (PBS) to remove cell debris. The distance between the two edges of the wound was calculated and imaged at 0 and 24 h under a microscope.

**Transwell invasion assays**

For the invasion assays, cells were seeded in a 24-well Corning FluoroBlok chamber precoated with Matrigel (BD Biosciences, Billerica, MA). A medium containing 20% FBS was added to the lower chamber in a 5% CO₂ incubator at 37°C C for 24 h. Then, the cells remaining on the lower side of the membrane were stained with 4% paraformaldehyde for 15 min, stained with 0.1% cresyl violet, washed three times with PBS, and air-dried. Five fields were randomly selected for counting the number of migrated cells and images (200× magnification) were captured under a phase-contrast microscope.

**Western blot analysis**

Standard Western blotting protocols were followed [18]. Whole-cell lysates were probed with antibodies specific to GAPDH (Santa Cruz, CA, USA), HNRNPU, MMP9, E-Cadherin, Vimentin, and N Cadherin (Cell Signaling, UK).

**Statistical analysis**

Data are expressed as the mean ± standard error of the mean for at least three independent experiments. Analysis of variance was used to distinguish the differences in each group. P < 0.05 was considered statistically significant. All analyses were performed using GraphPad Prism version 8.0 (Graph Pad Software, CA, USA).

**AUTHOR CONTRIBUTIONS**

QX conceived and designed the study. JW-Z, YG-G, JT-F, and ZH analyzed and interpreted the data. GH-M, MZ-W, QH-H, and YX-X conducted the experiments. GH-M, JW-Z, and YG-G analyzed the data. QX and JW-Z prepared Figures and wrote the manuscript. All authors have read and approved the final version.

**CONFLICTS OF INTEREST**

The authors declare no conflicts of interest.

**FUNDING**

We thank Qiming Jane Wang for providing helpful feedback during the preparation of this manuscript. This work was supported by the National Natural Science Foundation of China (#81701185), High School Key Research Project of Henan Province (#18A320003), Henan Science and Technology Innovation Team of Insect Bio-reactor (#C20140047), and Special Funded Projects of Nanyang Normal University (#17307).

**REFERENCES**

8. Wang X, Li Y, Fan Y, Yu X, Mao X, Jin F. PTBP1 promotes the growth of breast cancer cells through the PTEN/Akt


PMID:18438415

PMID:12636465

PMID:24626777

PMID:31355266

PMID:31819776

PMID:31828152

PMID:31492158

PMID:12871781

PMID:11121407

PMID:11510693

PMID:30021377

PMID:23585894

PMID:31901866

PMID:28351333

PMID:30158112

PMID:12189386
https://doi.org/10.1186/s12935-016-0368-4 
PMID: 28077929

https://doi.org/10.18632/aging.103333 
PMID: 32526707

https://doi.org/10.1016/j.cell.2015.08.011 
PMID: 26321680

https://doi.org/10.1038/ncomms3980 
PMID: 24356509

https://doi.org/10.1038/emboj.2011.38 
PMID: 21343907

https://doi.org/10.1371/journal.pone.0033144 
PMID: 22427970

https://doi.org/10.1593/neo.07112 
PMID: 17356713

https://doi.org/10.1093/nar/gkaa407 
PMID: 32442275

https://doi.org/10.1126/scisignal.2004088 
PMID: 23550210

https://doi.org/10.1016/j.neo.2017.05.002 
PMID: 28732212

https://doi.org/10.1093/nar/gkx1090 
PMID: 29136207

https://doi.org/10.1126/science.1260419 
PMID: 25613900
Supplementary Figure 1. Boxplot showing the relative protein expression of hnRNPs in cancer and normal tissues. Data are mean ± SE. *P < 0.05; **P < 0.01; ***P < 0.001.
Supplementary Figure 2. Genetic mutation of HNRNPU, associated with gene expression, and enrichment analysis of differentially expressed genes in patients with TCGA-BRCA. (A) Verification of the top three mutations of the differentially expressed hnRNPs in patients with BRCA (The Metastatic Breast Cancer Project Provisional, February 2020, and Metastatic Breast Cancer) (INSERM, PLoS Med 2016). (B) Verification of the top three mutations in differentially expressed hnRNPs in BRCA (Breast Cancer METABRIC, Nature 2012 and Nat Commun 2016). (C) The t-test was used to analyze the effect of HNRNPU mutation on its gene expression in patients with BRCA. (D, E) Heat maps showing the genes positively and negatively correlated with HNRNPU in BRCA (top 50). The significantly enriched GO annotations of HNRNPU coexpression gene (top 100) in BRCA. (F) Cellular components. (G) Biological processes. (H) Molecular functions.
Supplementary Tables

Please browse Full Text version to see the data of Supplementary Table 1.

Supplementary Table 1. The mRNA levels of hnRNPs upregulated in different types of breast cancer tissues and normal tissues at transcriptome level (ONCOMINE).

Supplementary Table 2. Correlation of hnRNPs mRNA expression and clinical prognosis in BRCA with LNM by kaplan-meyer plotter.

<table>
<thead>
<tr>
<th>Gene</th>
<th>N</th>
<th>OS</th>
<th>RFS</th>
<th>Log rank P</th>
<th>N</th>
<th>OS</th>
<th>RFS</th>
<th>Log rank P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
<td>HR</td>
<td>Log rank P</td>
<td>Low</td>
<td>High</td>
<td>HR</td>
<td>Log rank P</td>
</tr>
<tr>
<td>HNRNPA0</td>
<td>106</td>
<td>207</td>
<td>0.65(0.44-0.97)</td>
<td>0.033</td>
<td>834</td>
<td>299</td>
<td>0.58(0.45-0.75)</td>
<td>2.00E-05</td>
</tr>
<tr>
<td>HNRNPA2B1</td>
<td>127</td>
<td>50</td>
<td>0.72(0.40-1.29)</td>
<td>0.26</td>
<td>358</td>
<td>366</td>
<td>0.72(0.56-0.92)</td>
<td>0.0091</td>
</tr>
<tr>
<td>HNRNPC</td>
<td>55</td>
<td>122</td>
<td>0.34(0.20-0.58)</td>
<td>3.60E-05</td>
<td>187</td>
<td>537</td>
<td>0.52(0.40-0.68)</td>
<td>8.60E-07</td>
</tr>
<tr>
<td>HNRNPD</td>
<td>135</td>
<td>178</td>
<td>0.78(0.53-1.16)</td>
<td>0.22</td>
<td>391</td>
<td>742</td>
<td>0.81(0.66-0.99)</td>
<td>0.039</td>
</tr>
<tr>
<td>PCBP1</td>
<td>80</td>
<td>233</td>
<td>0.71(0.47-1.08)</td>
<td>0.11</td>
<td>351</td>
<td>782</td>
<td>0.90(0.73-1.11)</td>
<td>0.31</td>
</tr>
<tr>
<td>HNRNPF</td>
<td>80</td>
<td>233</td>
<td>0.65(0.43-0.98)</td>
<td>0.038</td>
<td>282</td>
<td>851</td>
<td>0.80(0.64-0.99)</td>
<td>0.041</td>
</tr>
<tr>
<td>PTBP1</td>
<td>81</td>
<td>232</td>
<td>0.66(0.43-1.01)</td>
<td>0.056</td>
<td>791</td>
<td>342</td>
<td>1.25(1.02-1.54)</td>
<td>0.032</td>
</tr>
<tr>
<td>HNRNPK</td>
<td>78</td>
<td>235</td>
<td>0.57(0.38-0.85)</td>
<td>0.0054</td>
<td>771</td>
<td>362</td>
<td>0.71(0.57-0.89)</td>
<td>0.0027</td>
</tr>
<tr>
<td>HNRNPL</td>
<td>223</td>
<td>90</td>
<td>1.37(0.92-2.05)</td>
<td>0.12</td>
<td>295</td>
<td>838</td>
<td>1.73(1.33-2.23)</td>
<td>2.50E-05</td>
</tr>
<tr>
<td>HNRNPM</td>
<td>52</td>
<td>125</td>
<td>1.62(0.79-3.31)</td>
<td>0.19</td>
<td>536</td>
<td>188</td>
<td>1.42(1.08-1.86)</td>
<td>0.011</td>
</tr>
<tr>
<td>SYNCRIP</td>
<td>159</td>
<td>154</td>
<td>1.32(0.89-1.96)</td>
<td>0.16</td>
<td>702</td>
<td>431</td>
<td>1.84(1.50-2.24)</td>
<td>1.20E-09</td>
</tr>
<tr>
<td>HNRNPR</td>
<td>83</td>
<td>94</td>
<td>1.55(0.89-2.70)</td>
<td>0.12</td>
<td>497</td>
<td>227</td>
<td>0.59(0.44-0.80)</td>
<td>0.00056</td>
</tr>
<tr>
<td>HNRNPU</td>
<td>59</td>
<td>118</td>
<td>2.11(1.06-4.19)</td>
<td>0.029</td>
<td>180</td>
<td>544</td>
<td>0.71(0.54-0.94)</td>
<td>0.015</td>
</tr>
</tbody>
</table>