



Identification of a Three-RNA Binding Proteins (RBPs) Signature Predicting Prognosis for Breast Cancer

Yang Liu^{1,2}, Hefen Sun^{1,2}, Xuan Li^{1,2}, Qiqi Liu^{1,2}, Yuanyuan Zhao^{1,2}, Liangdong Li^{1,2}, Baojin Xu^{1,2}, Yifeng Hou^{1,2*} and Wei Jin^{1,2*}

¹ Department of Breast Surgery, Key Laboratory of Breast Cancer in Shanghai, Fudan University Shanghai Cancer Center, Shanghai, China, ² Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China

Background: To date, breast cancer remains the primary cause of tumor-related death among women, even though some leap-type developments of oncology have been done to slash the mortality. Considering the tumor heterogeneity and individual variation, the more reliable biomarkers are required to be identified for supporting the development of precision medicine in breast cancer.

Methods: Based on the TCGA-BRCA and METABRIC databases, the differently expressed RNA binding proteins (RBPs) between tumor and normal tissues were investigated. In this study, we focused on the communal differently expressed RBPs in four subtypes of breast cancer. Lasso-penalized Cox analysis, Stepwise-multivariate Cox analysis and Kaplan–Meier survival curve were performed to identify the hub RBP-coding genes in predicting prognosis of breast cancer, and a prognostic model was established. The efficiency of this model was further validated in other independent GSE20685, GSE4922 and FUSCC-TNBC cohorts by calculating the risk score and performing survival analysis, ROC and nomogram. Moreover, pathologic functions of the candidate RBPs in breast cancer were explored using some routine experiments *in vitro*, and the potential compounds targeting these RBPs were predicted by reviewing the Comparative Toxicogenomics Database.

Results: Here, we identified 62 RBPs which were differently expressed between the tumor and normal tissues. Thereinto, three RBPs (MRPL12, MRPL13 and POP1) acted as independent risk factors, and their expression pattern also correlated with poor prognosis of patients. A prognostic model, built with these 3-RBPs, possessed statistical significance to predict the survival probability of patients with breast cancer. Furthermore, experimental validations showed that down-regulating the expression of endogenous MRPL12, MRPL13 or POP1 could dramatically suppress the cellular viability and migration of breast cancer cells *in vitro*. Besides, some compounds (such as the Acetaminophen, Urethane and Tunicamycin) were predicted for curing breast cancer *via* targeting MRPL12, MRPL13 and POP1 simultaneously.

Conclusion: This study identified and established a 3-RBPs-based signature and nomogram for predicting the survival probability of patients with breast cancer.

OPEN ACCESS

Edited by:

Maria Felice Brizzi,
University of Turin, Italy

Reviewed by:

Lin Liu,
Nankai University, China
Gaosong Wu,
Wuhan University, China

*Correspondence:

Wei Jin
weijin7207@126.com
Yifeng Hou
newmailhou@163.com

Specialty section:

This article was submitted to
Cancer Molecular Targets
and Therapeutics,
a section of the journal
Frontiers in Oncology

Received: 03 February 2021

Accepted: 19 May 2021

Published: 12 July 2021

Citation:

Liu Y, Sun H, Li X, Liu Q, Zhao Y, Li L,
Xu B, Hou Y and Jin W (2021)
Identification of a Three-RNA Binding
Proteins (RBPs) Signature Predicting
Prognosis for Breast Cancer.
Front. Oncol. 11:663556.
doi: 10.3389/fonc.2021.663556

MRPL12, MRPL13 and POP1 might act as oncogenes in maintaining cellular viability and accelerating metastasis of breast cancer cells, implying the possibility of which to be designed as biomarkers and/or therapeutic targets for breast cancer.

Keywords: breast cancer, RNA-binding protein, prognostic model, bioinformatics analysis, MRPL12, Mrpl13, POP1

INTRODUCTION

Breast cancer is an extraordinarily heterogeneous set of neoplasms with sophisticated pathogenesis and alarming incidence, especially being a formidable threat to women. Only in 2018, almost 2.1 million of breast cancer were diagnosed worldwide, accounting for 24.2% of all the new cancer cases in females (1). Although advances in the early diagnosis and clinical intervention slashed the mortality of breast cancer significantly, breast cancer remains a feared, refractory and primary cause of tumor-related death among women. Following a curative intent while avoiding being overtreated, the novel personalized and precision medicine that constructed upon the set of reliable tumor biomarkers and targets are emerging and ongoing. Therefore, the more reliable and available tumor biomarkers would be identified, the more eligible and prescriptible therapeutic strategies would be adopted in clinical treatment of breast cancer.

RNA-binding proteins (RBPs) are ubiquitously expressed in cells and play important roles in regulating the cellular physiological and pathological processes (2). Conventionally, RBPs participate in the formation of ribonucleoprotein complexes *via* binding to the modular sequence and/or structural motifs of target RNAs, and then dynamically alter the fate or function of RNAs, such as regulating its stability, localization, modification and translation (3–5). Moreover, some proteins that are lacking the canonical RNA-binding domains also moonlight as RBPs under specific cellular context (6). For instance, the 3-hydroxyacyl-CoA dehydrogenase type 2 (HSD17B10) not only behaves as a mitochondrial enzyme which is involved in the oxidation of isoleucine and branched-chain fatty acids, but also binds to the 5'-end of most mitochondrial tRNAs that recruits RNase P to regulate the stability and activity of target mitochondrial tRNAs (7–9).

Accumulating evidence revealed an inextricable causality between the dysregulated RBPs and various human diseases, including malignant tumor (10). Intrinsically, throughout the tumorigenesis and progression, RBPs act as participants and/or coordinators to regulate the proliferation, apoptosis, senescence, angiogenesis, metastasis and invasion of cancer cells (11, 12). For instance, aberrant over-expressed IGF2BP1 in hepatocarcinoma could accelerate cell proliferation *via* motivating the expression of its target mRNAs, including *c-Myc*, *Ki67* and β -*TrCP1* (13). DND1 might competitively interact with *miR-221* against *BIM* and stabilized *BIM* mRNA, thus haploinsufficiency of DND1 assisted the breast cancer cells in escaping apoptosis *via* decreasing the expression of *BIM* (14). Moreover, ESRP1/2 might contribute to the epithelial–mesenchymal transition of cancer cells by regulating the alternative splicing of *FGFR2* and *CD44*, and in turn accelerated metastasis (15, 16).

Given that RBPs are critical regulators and effectors of cancer, they also should be regarded as hallmarks and potential targets for early diagnosis, prognosis and clinical intervention of cancer. NOP14/MRPS23/POP1 exhibit prognostic value in colorectal cancer, while APOBEC3C/DCAF13/EIF4E3/EZR possess significance in predicting prognosis of breast cancer (17, 18). Moreover, the Ribavirin, an antiviral guanosine analog, is able to slacken tumorigenesis by competitively interacting with the 5'-Cap of *eIF4E* mRNA, subsequently impeding the translation and transposition of eIF4E-regulated *Cyclin D1* (19). Nevertheless, the physiological functions and pathological features of most RBPs remain unknown, especially for the unconventional RBPs.

The benefit of integrating the high-throughput genomic analysis of clinical specimens and relative epidemiological research, systematic studies of RBPs in one specific type or subtype of cancer will be more helpful to identify the exclusive biomarkers and potential therapeutic targets in one restrictive type of cancer, such as the ER-signaling pathway has been applied as the major therapeutic target in Luminal-like breast cancer, but is inefficient in basal-like breast cancer. Therefore, in this study, the molecular classification of breast cancer was employed as a precondition to identify the extra biomarkers for breast cancer. A series of stepwise bioinformatics analysis were carried out in five transcriptional profile datasets (TCGA-BRCA, METABRIC, FUSCC-TNBC, GSE20685 and GSE4922), and identified a number of RBPs which might be involved in breast cancer. Furthermore, *in vitro* experimental validation indicated that some of these candidate RBPs would be used as generic prognostic biomarkers and potential therapeutic targets for breast cancer.

MATERIALS AND METHODS

Acquisition of Data Sources and Identification of Differently Expressed RBPs

The RNA-sequencing information of 1,001 cases of breast cancer samples and 113 cases of normal breast tissue samples with clinicopathological information were obtained from the TCGA-BRCA database (<http://portal.gdc.cancer.gov/>), wherein 653 cases of Luminal A, 150 cases of Luminal B, 39 cases of HER2-riched as well as 159 cases of TNBC samples were included. In this study, we referred to the human RBP catalog reported by Gerstberger et al. (2), which consisted of 1,542 RBPs but only 1,538 RBP-coding genes are recorded and available in the TCGA-BRCA database (**Supplementary Table 1**). To identify the differently expressed RBP-coding genes between the breast tumor and normal tissues, the “edgeR”, “pheatmap” and “ggplot2” packages were employed and performed in each

molecular subtype with normal tissues as control, in view of $|\log_2(\text{fold change})| \geq 1$ and $\text{FDR} < 0.05$. Then, the “VennDiagram” package was used to identify the communal differently expressed RBP-coding genes between the four subtypes of breast cancer. Moreover, in order to increase the credibility of bioinformatic prediction, other datasets were introduced with detailed RNA-sequencing data and complete clinicopathological information. The METABRIC dataset ($n = 1\text{m}893$) were obtained from cBioPortal (<http://www.cbioportal.org/>), the GSE20685 cohort ($n = 327$) and the GSE4922 cohort ($n = 249$), as well as our local FUSCC-TNBC dataset (GSE118527) which contains 360 cases of TNBC samples and 88 cases of paired para-carcinoma samples were obtained from GEO (<http://www.ncbi.nlm.nih.gov/>). Data from TCGA-BRCA and METABRIC databases were normalized using the “edgeR” package, while data from GEO databases were normalized using the “limma” package.

Construction of Protein–Protein Interaction (PPI) and Pathways Interaction Network

The STRING database (*version 11.0*, <http://www.string-db.org/>) was used to evaluate the protein–protein interaction information and biological functions of all the differently expressed RBPs. The PPI network was visualized by *Cytoscape 3.8.0*, and the interaction modes of biological processes along with molecular functions were built by ClueGO in *Cytoscape*.

Identification of Hub Genes

In the TCGA-BRCA dataset, Lasso-penalized Cox regression analysis and stepwise multivariate Cox regression analysis were conducted in turn to screen hub genes from the above differently expressed RBPs, that might be prognostic signature for overall survival (OS) and/or disease-free survival (DFS) in breast cancer patients. Then, the candidate genes derived from Cox regression analysis were submitted to assess their prognostic significance for survival of patients, who were grouped by the Youden index of candidate mRNA expression pattern into high- or low-expression set. Only if the candidate is significantly in both the TCGA-BRCA and METABRIC datasets, would be considered as a hub gene for further study.

Establishment and Validation of the Prognostic Model

A prognostic model for breast cancer was built based on the signature of these hub genes predicted as above, which was constructed by a linear combination of the Cox regression coefficients (β) multiplied with its mRNA expression pattern. In brief, the risk score of each patient was calculated by the formula that risk score = $(\beta_{\text{Gene1}} \times \text{expression}_{\text{Gene1}}) + (\beta_{\text{Gene2}} \times \text{expression}_{\text{Gene2}}) + \dots + (\beta_{\text{GeneN}} \times \text{expression}_{\text{GeneN}})$. Patients were grouped into high- or low-risk set according to the Youden index of their risk score, and the survival difference between these two sets was analyzed by the Kaplan–Meier estimation combined with a log-rank test in the testing datasets (TCGA-BRCA and METABRIC). Moreover, the predictive value of this model was further applied and validated into other datasets, such as the public GSE20685 cohort, GSE4922 cohort and our local FUSCC-TNBC dataset (GSE118527).

Independent Prognostic Analysis

To investigate whether this prognostic model was independent of other clinical parameters, such as age, ER status, PR status, HER2 status and TMN stage, univariate and multivariate analyses were carried out in the TCGA-BRCA dataset using Cox regression model method with forwarding stepwise procedure where $p < 0.05$ was considered as statistically significant.

Construction and Validation of a Predictive Nomogram

For application of the prognostic model, a predictive nomogram was constructed using all the independent prognostic factors derived from the multivariate Cox regression analysis. Discrimination of the nomogram was assessed by the concordance index (C-index) using a bootstrap method with 1,000 resamples, and prediction probabilities of the nomogram against actual survival of patients were evaluated by plotting the calibration curve.

Acquisition and Processing of Samples

For experimental validation, altogether 88 cases of tumor tissues and paired para-carcinoma tissues were obtained from patients with TNBC, who had undergone radical resection surgical of tumor between 2007 and 2014 with a median duration of follow-up of 45.8 months at Fudan University Shanghai Cancer Center, China. The dissected samples were stored in liquid nitrogen immediately until extraction of total RNA. Meanwhile, a series of immortalized breast cancer cell lines (such as MDA-MB-231, bone metastatic 231-BO, lung metastatic 231-HM), the mammary epithelial cell line MCF10A as well as its derived pre-malignant cell line MCF10^{DCIS}, malignant cell line MCF10^{Ca1h} and MCF10^{Ca1a} were employed. According to the recommendations from the American Type Culture Collection (ATCC), the MDA-MB-231, 231-BO and 231-HM cells were routinely maintained in *Leibovitz's L-15* basal medium supplemented with 10% fetal bovine serum and 1% penicillin–streptomycin, while the MCF-10A, MCF-10^{DCIS}, MCF-10^{Ca1h} and MCF-10^{Ca1a} cells were grown in *DEME/F-12* medium supplemented with 5% Chelex-treated horse serum, 0.01 mg/ml human insulin, 20 ng/ml epidermal growth factor, 100 ng/ml cholera toxin and 500 ng/ml hydrocortisone.

Realtime-qPCR Analysis

For evaluating the transcriptional expression pattern of the hub genes, the total RNA was isolated from tissues and cells using TRIzol reagent (*Invitrogen#15596026*), and cDNA was synthesized using the PrimeScriptTM RT reagent kit (*TaKaRa#RR037*) under the guidance of the manufacturer's instructions. Realtime-qPCR was performed using the TB Green[®] Premix Ex TaqTM II kit (*TaKaRa#RR82LR*), and the sequence of primers were listed as below: MRPL12-F 5'-ATCCCCATAGCGAAAGAACGG-3'; MRPL12-R 5'-GGACGAGGTTGATGCTTGG-3'; MRPL13-F 5'-ACATAAACCTGTGTACCATGCAC-3'; MRPL13-R 5'-GGTAGCCAGTATGCGAAGAGT-3'; POP1-F 5'-AGACAGCTGCTTCTGGAGGGT-3'; POP1-R 5'-TGACCCAGGCACCAGCAATTC-3'; Besides, GAPDH was used as internal control for normalization, whose sequence was

designed as GAPDH-F 5'-GGAGCGAGATCCCTCCAAAAT-3' and GAPDH-R 5'-GGCTGTTGTCATACTTCTCATGG-3'.

Down-Regulating the Expression of the Candidate Endogenous RBP-Coding Genes

For each candidate RBP, two shRNA oligos against different sites of CDS (coding sequence) were designed and synthesized (**Supplementary Table 2**). The pLKO.1-puro plasmids (*Addgene#8453*) carried the specific shRNA oligos for *MRPL12*, *MRPL13* or *POPI* were constructed, and be co-transfected into HEK-293T cells with the pVSVg (*Addgene#8454*) and psPAX2 (*Addgene#12260*) for packaging specific lentivirus. Then, the specific lentivirus was harvested and be co-cultured with target cells. Realtime-qPCR analysis was applied to assess the interference effect of these shRNAs on the expression of target genes.

Cellular Migration and Viability Assay

For assessing the influence of candidate RBPs on the migration of breast cancer cells, transwell assay was performed. Cells in logarithmic growth phase were harvested and suspended in DMEM-H medium, and seeded into the upper chambers (0.8 μm , *BD#353097*) at 5×10^4 cells/well. While the lower chambers were filled with 600 μl of DMEM-H medium supplemented with 20% FBS. After 14 h incubation, cells were fixed and stained with crystal violet staining solution, while the non-invading cells could be removed using swab. Then, the invaded cells were counted under microscope in six random fields at a magnification of 100 \times . Cloning efficiency was detected for evaluating the influence of these candidate RBPs on cellular viability. Cells in logarithmic growth phase were harvested and suspended in complete growth medium, and seeded into 6-well clusters at 5×10^2 cells/well. About 14 d later, cell colonies were fixed and stained with crystal violet staining solution. Then, the colony-forming efficiency was calculated by comparing the number of colonies between the gene-edited cells and its parental cells.

Target-Drug Interaction Network Analysis

For constructing a target-drug interaction network, the online database CTD (Comparative Toxicogenomics Database, <http://ctdbase.org/>) was introduced. By retrieving the CTD database, some potential compounds that might be involved in regulating these prognostic biomarkers were identified, and visualized by using *Cytoscape 3.8.0*.

Statistical Analysis

R v3.5.0 and *SPSS v23.0.0.0* were employed for statistical analysis, and the *GraphPad Prism v8.4.3* was used for scientific drawing. False discovery rate (FDR) ≤ 0.05 was designed as threshold for filtering the significantly difference between tumor and normal tissues in bioinformatics analysis, while a Student's t-test was used for comparing the difference between the test and control groups in experimental validations. Meanwhile, Kaplan-Meier curves and Cox regression analysis with log-rank *p* test were

applied into survival analysis, in which $p < 0.05$ was considered as statistically significant.

RESULTS

62 RBP-Coding Genes Were Differently Expressed in Breast Tumor Specimens

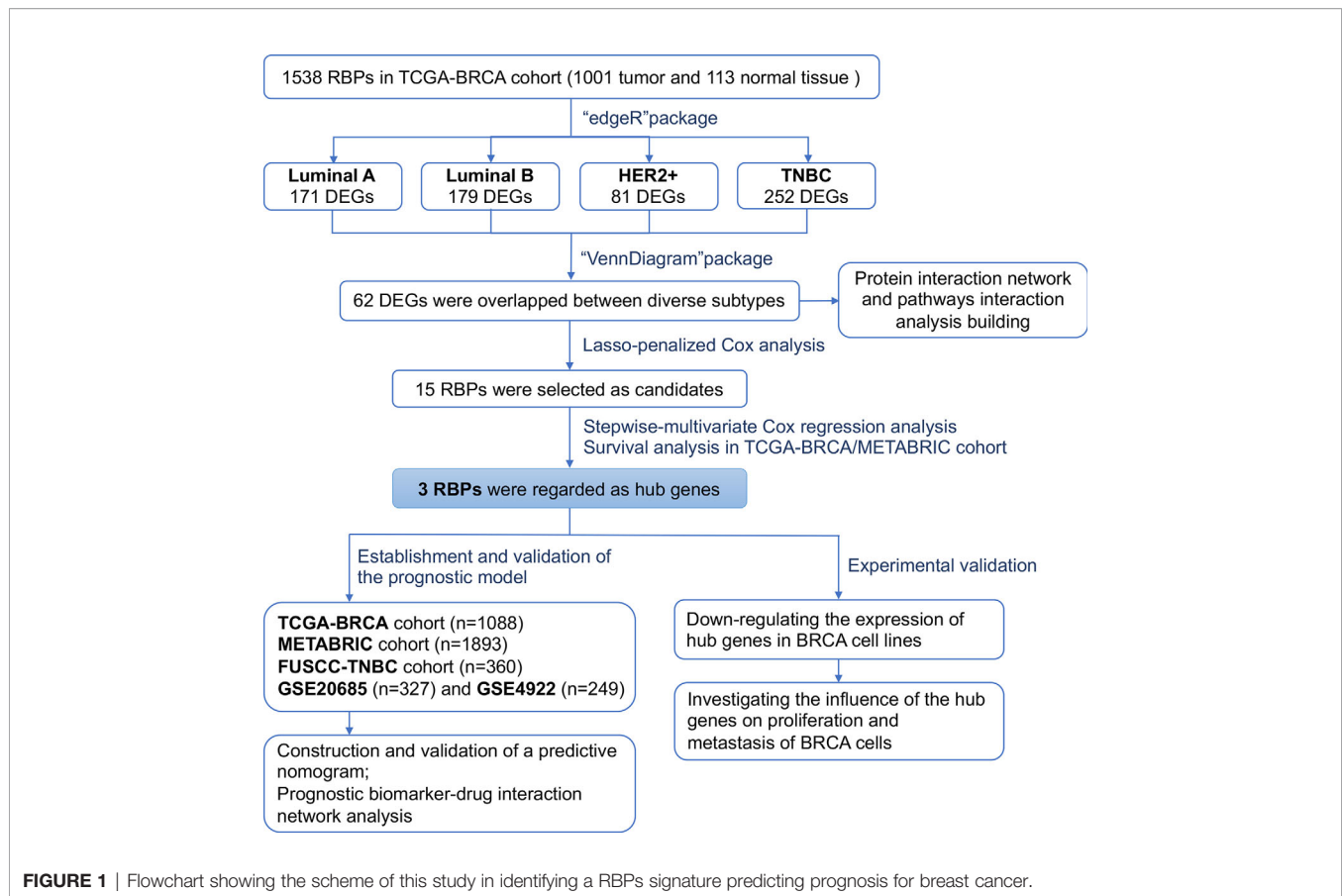
Based on the available RNA sequencing data from TCGA-BRCA database, we conducted a deep analysis of 1,538 RBP-coding genes and identified a total of 323 genes that be differently expressed between tumor and adjacent normal tissues (**Figure 1** and **Supplementary Tables 3, 4**). Given that breast cancer is a heterogeneous and polygenic disease, which exhibits biological distinctness and behavioral difference, we further employed the recognized four molecular subtyping of breast cancer into analysis. Compared to normal tissues, there were 171 RBPs differently expressed in tumor tissues of the Luminal A subtype, 179 RBPs differently expressed in the Luminal B subtype, 81 RBPs differently expressed in the HER2-riched subtype while 252 RBPs differently expressed in TNBC samples (**Figure 2A** and **Supplementary Table 4**). Notably, 62 differently expressed RBPs were communal in four subtypes of breast cancer, wherein 24 RBPs were up-regulated but 38 RBPs were down-regulated in tumor tissues (**Figure 2B** and **Supplementary Table 4**).

Functional and Pathway Interrelation Analysis of the Differently Expressed RBPs

To determine the function of these differently expressed RBPs in oncogenesis or even deterioration of breast cancer, we constructed the protein-protein interaction (PPI) network and analyzed the pathways interrelation. According to the STRING 11.0 database, 62 PPI nodes and 56 edges with a *p*-value of $2E-12$ were plot (**Figure 2C**). Notably, ELVAL2 showed the highest relationship with other candidate RBPs, but it had been winnowed out in further Cox regression analysis, it might be due to its negligible contribution to predict prognosis of breast cancer. For GO analysis, the biological process of these 62 RBPs was mainly enriched in RNA phosphodiester bond hydrolysis, DNA modification and regulation of RNA splicing, while these 62 RBPs were also played important roles in regulating the ribonuclease activity, mRNA 3'-UTR binding, mRNA binding as well as poly-A binding (**Figure 3**).

Establishment and Validation of a 3-RBPs-Signature Based Prognostic Model

Take advantage of Lasso-penalized Cox analysis, 15 RBP-coding genes which might be prognosis signature for survival were screened from the aforementioned 62 RBPs (**Figure 2A**). Subsequently, 6 RBP-coding genes with *p*-value < 0.05 were further identified as candidates by Stepwise-multivariate Cox regression analysis (**Supplementary Figures 1, 2** and **Supplementary Table 4**). As shown in **Figure 4**, patients with highly expressed *MRPL12*, *MRPL13* or *POPI* always showed poorer disease-free survival (DFS, $p < 0.05$) in the TCGA-BRCA



dataset and poorer overall survival (OS, $p < 0.05$) in the METABRIC dataset. Although *MRPL12*, *MRPL13* or *POPI* were over-expressed in tumor tissues, there were also some slight differences in expression pattern between each subtype of breast cancer, wherein TNBC cases tolerated the highest expression of these three RBPs (**Figure 4B** and **Supplementary Table 5**). In addition, the expression pattern of *MRPL12*, *MRPL13* or *POPI* were significantly associated with tumor stage (**Figure 4D** and **Supplementary Table 6**), suggesting that these three RBP-coding genes could be regarded as potential biomarkers for breast cancer prognosis.

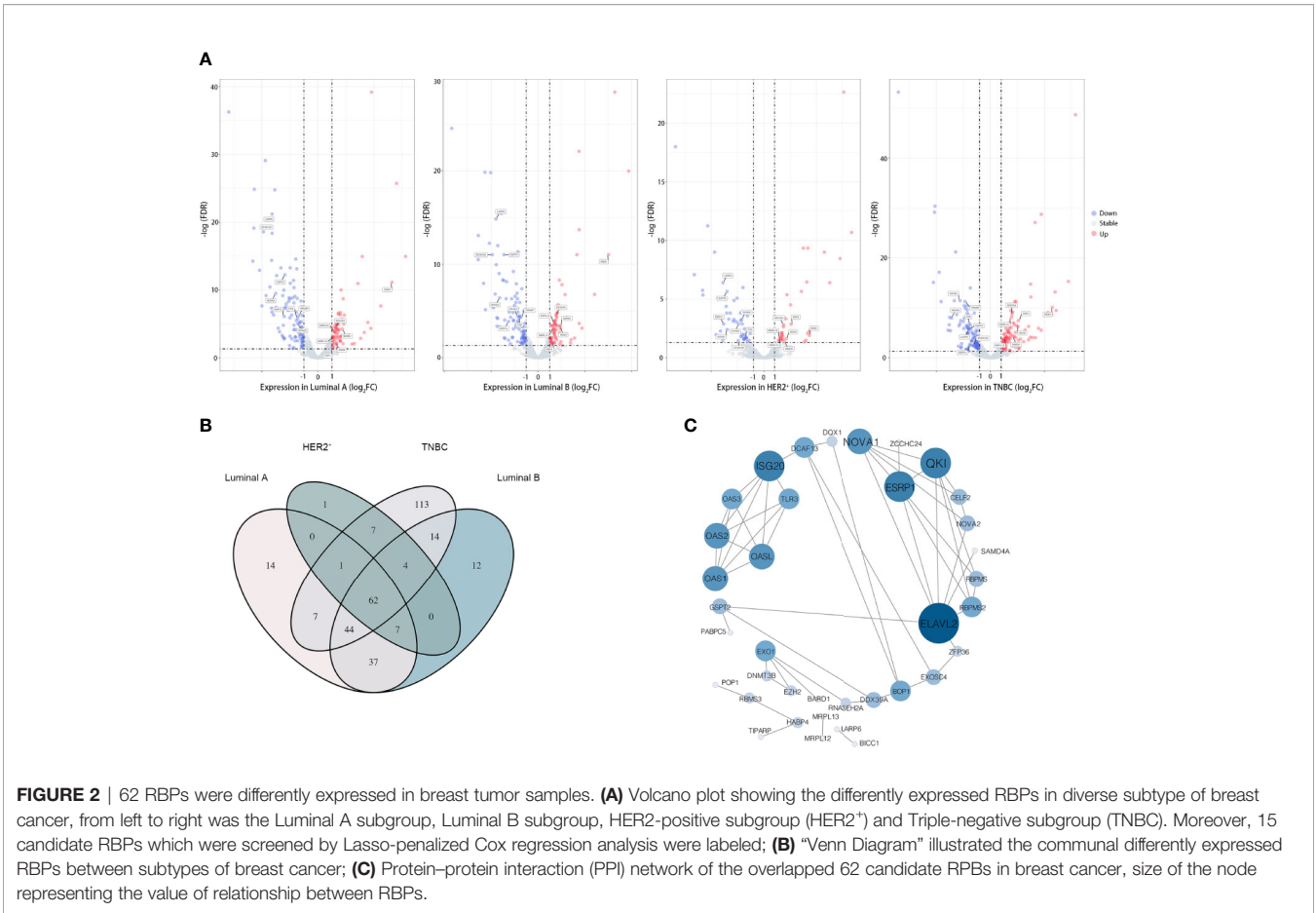
Furthermore, a prognostic model based on these three RBP-coding genes was constructed by a linear combination of the regression coefficient and its mRNA expression level. In this model, the risk score of each patient was calculated by the equation that risk score = $0.282 * \text{Expression}_{MRPL13} - 0.186 * \text{Expression}_{MRPL12} + 0.288 * \text{Expression}_{POPI}$. Patients were grouped into low- or high-risk set according to the Youden index of their risk score, and the Kaplan–Meier survival curve combined with a long-rank test was employed to assess whether these three RBP-coding genes can be used as a gene signature for predicting the survival probability of patients with breast cancer. Either within the TCGA-BRCA or METABRIC datasets, patients of high-risk were with poorer OS than those of low-risk (**Figure 5A**). However, the area under curve (AUC) value of ROC curves of this model was only

0.619 for predicting 3 years survival and 0.605 for predicting 5 years survival (**Figure 5B**). Moreover, the FUSCC-TNBC dataset ($n = 360$), GSE4922 cohort ($n = 249$) and GSE20685 cohort ($n = 327$) were introduced for external validation, which indicated a consistent result that patients in the high-risk group always showed significantly poorer survival probability than the low-risk group (**Figure 5C**), indicating a good performance of this three-gene signature for survival prediction of breast cancer patients.

Building and Validation of a Predictive Nomogram

Some 687 patients from the TCGA-BRCA dataset with complete clinicopathological information including age, ER status, PR status, HER2 status and TMN stage were submitted for further analysis. As shown in **Figure 6**, univariate and multivariate Cox regression analysis indicated that age, TNM stage and the risk score calculated from the three-gene signature were independent prognostic factors for OS of breast cancer patients.

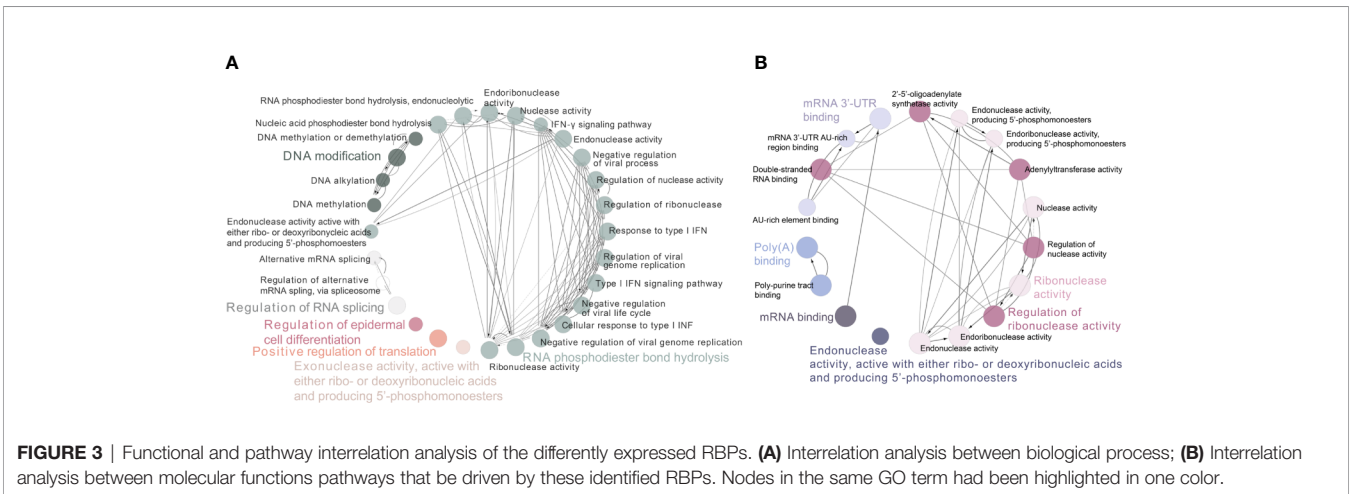
Subsequently, these three independent prognostic factors including age, TNM stage and risk score were used to build a nomogram to predict 3-year and 5-year OS of patients with breast cancer. As the calibration plots shown in **Figure 7**, this nomogram built with the combined model might under-estimate the mortality which would be the ideal nomogram for predicting short-term survival for patients with breast cancer.



MRPL12, MRPL13 and POP1 Were Functioned as Oncogenes in Breast Cancer

In order to reveal the pathological function of these three RBPs in breast cancer, some routine experimental methods were employed. In line with the above statistical analysis, real-time-

qPCR assay showed that either *MRPL12*, *MRPL13* or *POP1* were over-expressed in the clinical TNBC tissues rather than in the para-carcinoma tissues (**Figure 8A**). Similarly, *MRPL12*, *MRPL13* and *POP1* were also over-expressed in the pre-malignant cell line MCF10^{DCIS} as well as the malignant cell line MCF10^{Ca1h} and MCF10^{Ca1a} rather than in the mammary epithelial cell line MCF10A (**Figure 8B**). And the expression



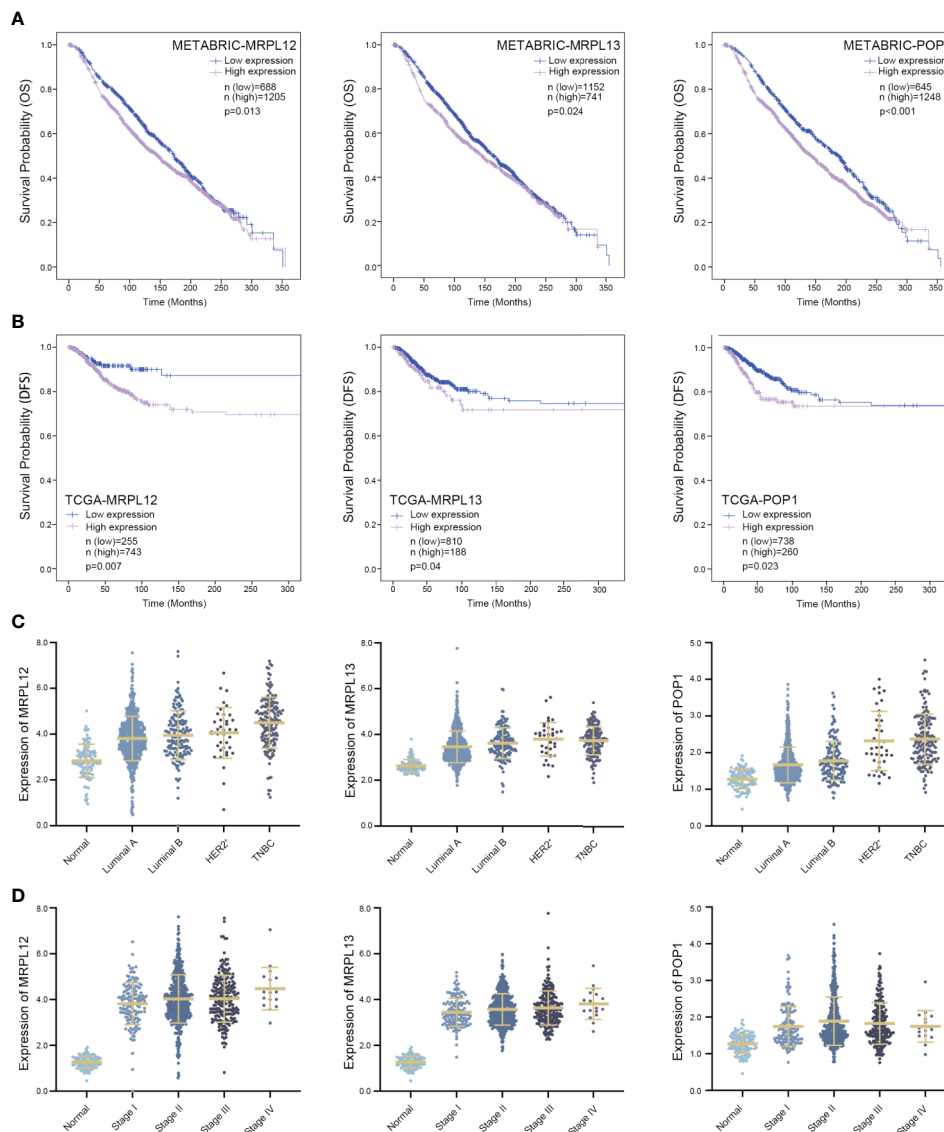


FIGURE 4 | Three RBPs might be designed as prognostic biomarkers for breast cancer. **(A)** Survival Kaplan–Meier estimate of hub RBPs in the TCGA-BRCA database ($n = 1,088$), clinical samples were grouped by Youden index of mRNA expression level; **(B)** Kaplan–Meier plotter of these hub RBPs in the METABRIC database ($n = 1,893$), in which samples were grouped by Youden index of mRNA expression level; **(C)** Comparing the expression pattern of these hub RBPs in four subtypes of breast cancer versus normal tissue; **(D)** Correlation between the expression pattern of these hub RBPs and the tumor stage in patients suffering from breast cancer. DFS, disease-free survival; OS, overall survival and $p < 0.05$ was considered as statistically significant.

pattern of these three RBPs were more abundant in the metastatic 231-HM and 231-BO than in their parental MDA-MB-231 cells (**Figure 8B**). These results suggested that *MRPL12*, *MRPL13* and *POP1* might acted as oncogenes in breast cancer.

Furthermore, we down-regulating the expression of endogenous *MRPL12*, *MRPL13* or *POP1* in 231-HM and 231-BO cells using the specific shRNAs (**Figures 8C, D**). Transwell assay indicated that down-regulating the expression of these three RBPs drastically decelerated the migration of breast cancer cells (**Figure 9**). Meanwhile, down-regulating the expression of these three RBPs also inhibited the cloning efficiency of breast

cancer cells, suggesting that these three RBPs played an important role in maintaining the vitality of breast cancer cells (**Figure 10**). All these results confirmed that *MRPL12*, *MRPL13* and *POP1* acted as oncogenes in breast cancer cells, therefore, they might be regarded as prognostic indicators in breast cancer, but also could be designed as therapeutic targets of breast cancer.

Target–Drug Interaction Network Analysis

To predict some available and potential compounds for targeting these three RBPs, a gene–compounds interaction network was constructed by using the Comparative Toxicogenomics Database

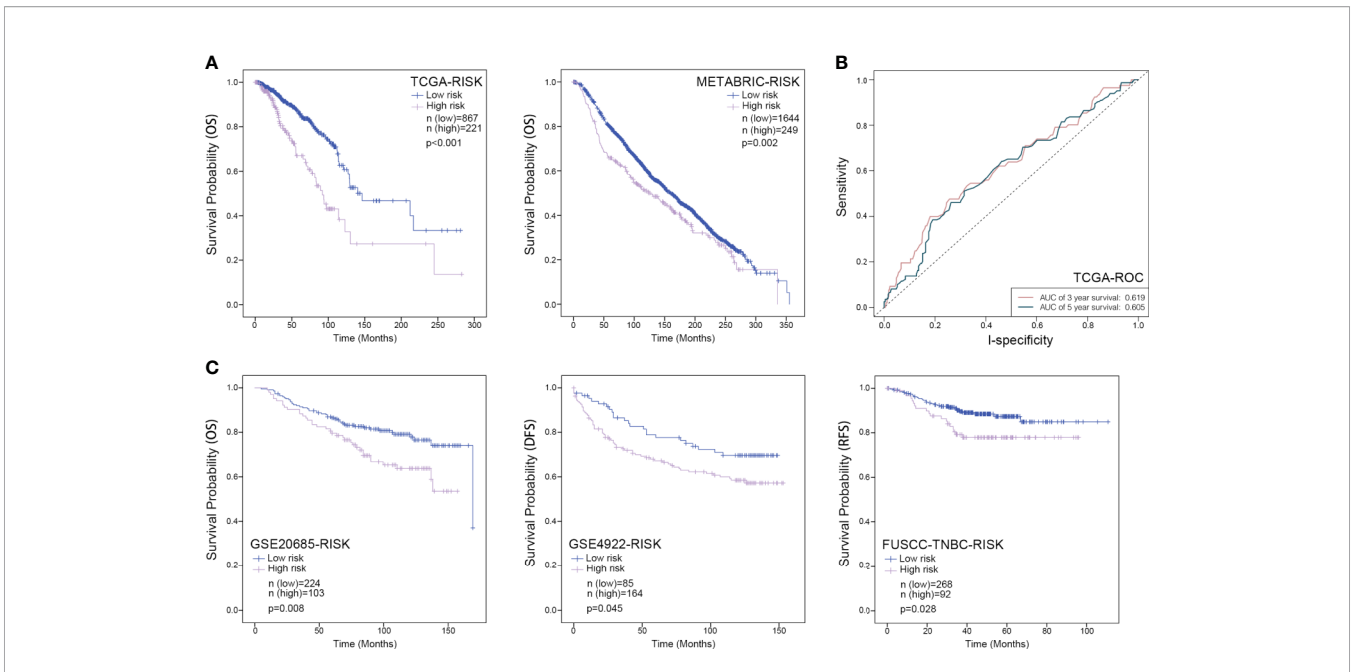


FIGURE 5 | Kaplan–Meier analysis for the three-RBPs prognostic signature in breast cancer. **(A)** Survival Kaplan–Meier estimate of the three-RBPs prognostic signature in the TCGA-BRCA database (n = 1,088) and METABRIC database (n = 1,893), in which clinical samples were grouped by Youden index of risk score; **(B)**, ROC curve for assessing the reliability of the model; **(C)** Survival Kaplan–Meier estimate of the three-RBPs prognostic signature in the GSE20685 database (n = 327), the GES4922 database (n = 249) as well as the FUSCC database (n = 360), in which samples were grouped by Youden index of risk score. OS, overall survival; DFS, disease-free survival; RFS, recurrence-free survival. *P* <0.05 was considered as statistically significant.

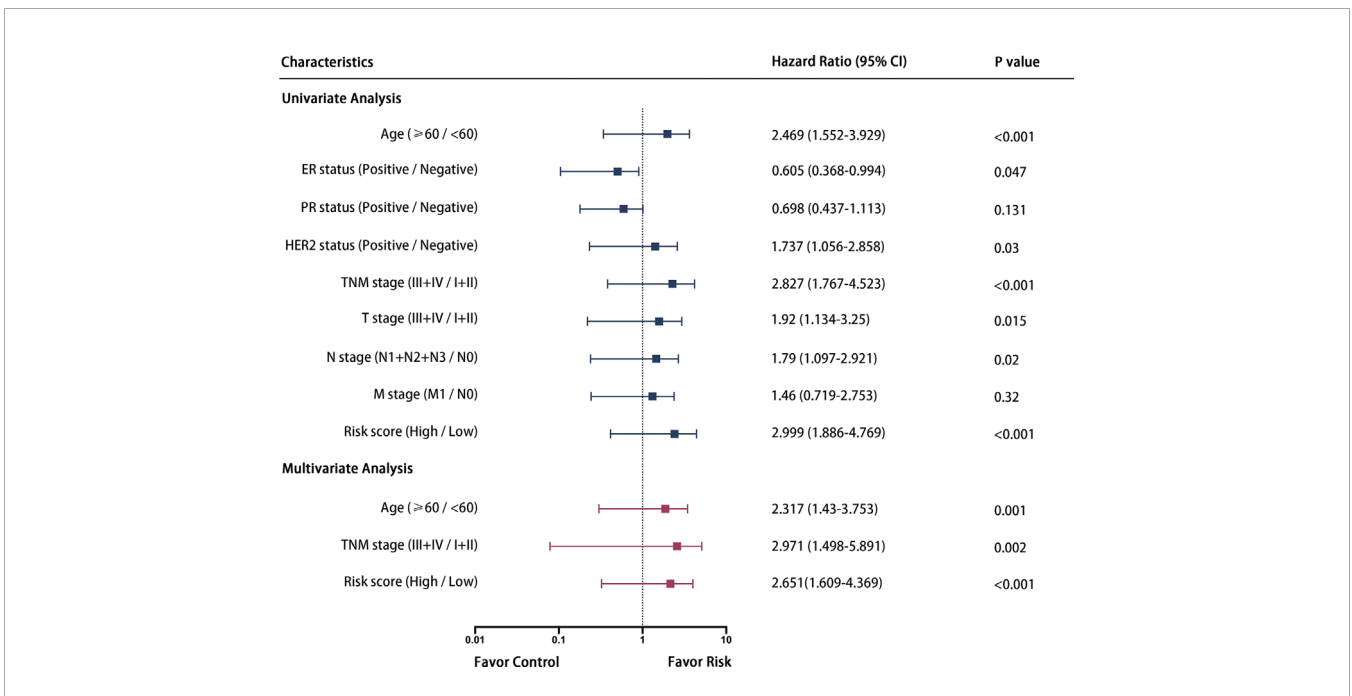


FIGURE 6 | Forrest plot of the univariate and multivariate Cox regression analysis in breast cancer.

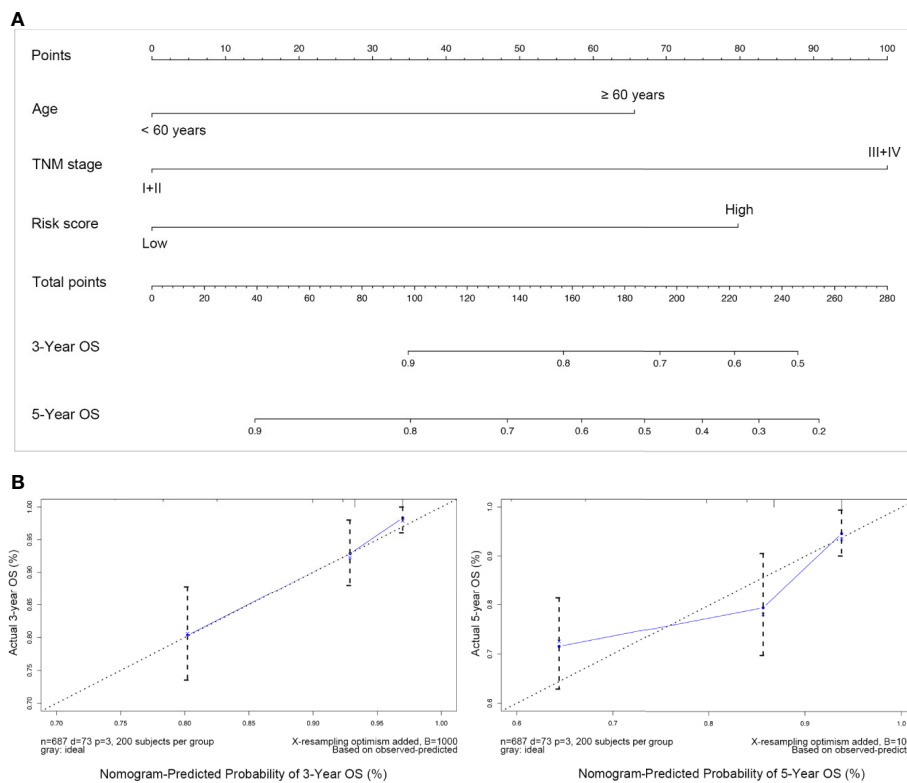


FIGURE 7 | Nomogram predicting overall survival for breast cancer. **(A)** For each patient suffering from breast cancer, three lines should be drawn upward to determine the points received from the three predictors in the nomogram, then the sum of these points should be located on the “Total points” axis, following a line be drawn downward to determine the possibility of 3-years and 5-years overall survival; **(B)** Calibration plot for validating this nomogram, in which the Y-axis represented actual survival while the X-axis represented predicted survival by the above nomogram.

(CTD) and *Cytoscape 3.8.0*. As shown in **Figure 11**, there were more than twenty compounds that could influence MRPL12, MRPL13 and/or POP1, wherein eight compounds could influence these three prognostic markers simultaneously. Thereinto, Ethinyl Estradiol could up-regulate the MRPL12, MRPL13 and POP1, while Acetaminophen, Bisphenol A, Urethane and Tunicamycin could negatively affect the physiological functions of these three RBPs. Flutamide and Valproic Acid might be used as agonist of MRPL12 and MRPL13, but might be used as antagonist against POP1. And Atrazine might be used as agonist of MRPL13 and POP1, but might be used as antagonist against MRPL12.

DISCUSSION

Currently, breast cancer remains the most prevalent malignancy in females, even though the advanced surgical and systemic treatments have successfully doubled the survival rate of patients (1). Breast cancer is characterized by multifactorial etiology and multigenetic alteration, which displays extraordinarily heterogeneity as well as capricious clinical phenotypes, confronts diverse clinical events and requires

different therapeutic strategies (20). Therefore, it is vital to consider the subtypes of breast cancer as a prerequisite to any basic research and clinical regimens.

In this study, we introduced the prevailing four molecular subtypes of breast cancer to define the universality and specificity of the hub genes which would be designed as biomarkers in clinical practice of breast cancer. Generally, breast cancer is mainly classified into the Luminal-A ($ER^+/PR^{high}/HER2^-$), Luminal-B ($ER^+/PR^{low}/HER2^-$), HER2-riched ($ER^-/PR^-/HER2^+$) and Basal-like ($ER^-/PR^-/HER2^-$, also known as TNBC) subtypes according to the expression pattern of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) (21, 22). As the retrieval of TCGA-BRCA database, a total of 323 differential expressed RBP-coding genes were identified by comparing the transcriptome between tumor and normal tissues (**Supplementary Table 3**). However, only a fraction of these differently expressed RBPs were communal in four subtypes of breast cancer, while others showed specificity in one or more subtypes (**Figure 2**). Here, to develop a universal predictive model that would be rudely applied in early diagnosis and prognosis of breast cancer, we merely focused on the differential expressed RBP-coding genes which were overlapped

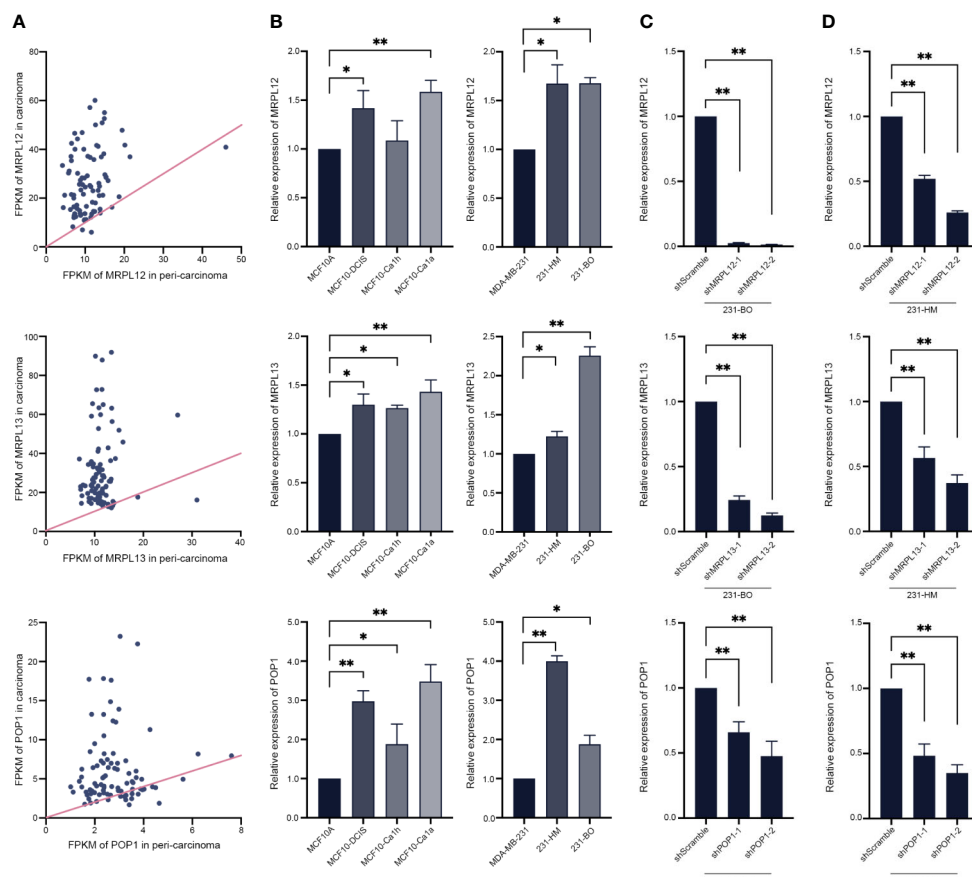


FIGURE 8 | Experimental validation and perturbation of the expression pattern of hub RBPs in breast cancer cells. **(A)** The total RNA was isolated from 88 tumor-tissues and its para-carcinoma tissues from TNBC patient, and realtime-qPCR was carried out to compare the expression pattern of MRPL12, MRPL13 and POP1 between tumor and normal tissues. The straight line was an isoline, which means the equal expression pattern of hub genes in carcinoma and para-carcinoma tissues; **(B)** Realtime-PCR assay for evaluating the expression pattern of these three RBPs in a series of immortalized breast cancer cell lines (such as MDA-MB-231, bone metastatic 231-BO, lung metastatic 231-HM), the mammary epithelial cell line MCF10A as well as its derived pre-malignant cell line MCF10^{DCIS}, malignant cell line MCF10^{Ca1h} and MCF10^{Ca1a}; **(C, D)** Expression of the hub RBPs was down-regulated by RNA interference in 231-BO and 231-HM cells. * $p < 0.05$; ** $p < 0.01$; Bars, \pm SD.

between the four molecular subtypes of breast cancer, wherein 62 RBP-coding genes were qualified for further study.

To increase feasibility and accuracy of the prognostic model, a succession of screening was performed as depicted in the flow chart (**Figure 1**), which identified three RBP-coding genes (*MRPL12*, *MRPL13* and *POP1*) that possessed statistical significance in prognosis assessment of patients with breast cancer. No matter in the TCGA-BRCA dataset or in the METABRIC dataset, a significant negative-relationship always existed between the survival probability and the expression level of *MRPL12*, *MRPL13* or *POP1*, meaning that patients should undergo high risk in recurrence and deterioration of breast cancer if they were abundant in the expression of *MRPL12*, *MRPL13* and/or *POP1*. Naturally, these three RBP-coding genes were constructed into the prognostic model for breast cancer, in addition, the risk score of this model was an independent prognostic factor. As shown in **Figure 5**, this prognostic model worked well not only in the TCGA-BRCA and METABRIC

datasets, but also be significant in the GSE20685, GSE4922 as well as the FUSCC-TNBC cohorts, in which patients in the high-risk group tended to undergo the more dismal survival probability than patients in the low-risk group. Moreover, a nomogram combining the 3-RBPs signature and conventional clinical parameters (such as age and TNM stage) performed the best in predicting survival probability for patients with breast cancer, indicating that this prognostic model developed from the three RBP-coding genes' signature could be designed as a useful indicator for survival in breast cancer.

Consistent with statistical analysis of the public databases, we found that *MRPL12*, *MRPL13* and *POP1* were over-expressed in most of the TNBC-tissues than their para-carcinoma specimens, and be more abundant in the differentiated and metastatic breast cancer cell lines (such as MCF10^{Ca1a}, 231-HM and 231-BO). Furthermore, down-regulating the expression of endogenous *MRPL12*, *MRPL13* or *POP1* in breast cancer cells, resulted in dramatically suppression of cellular viability and migration,

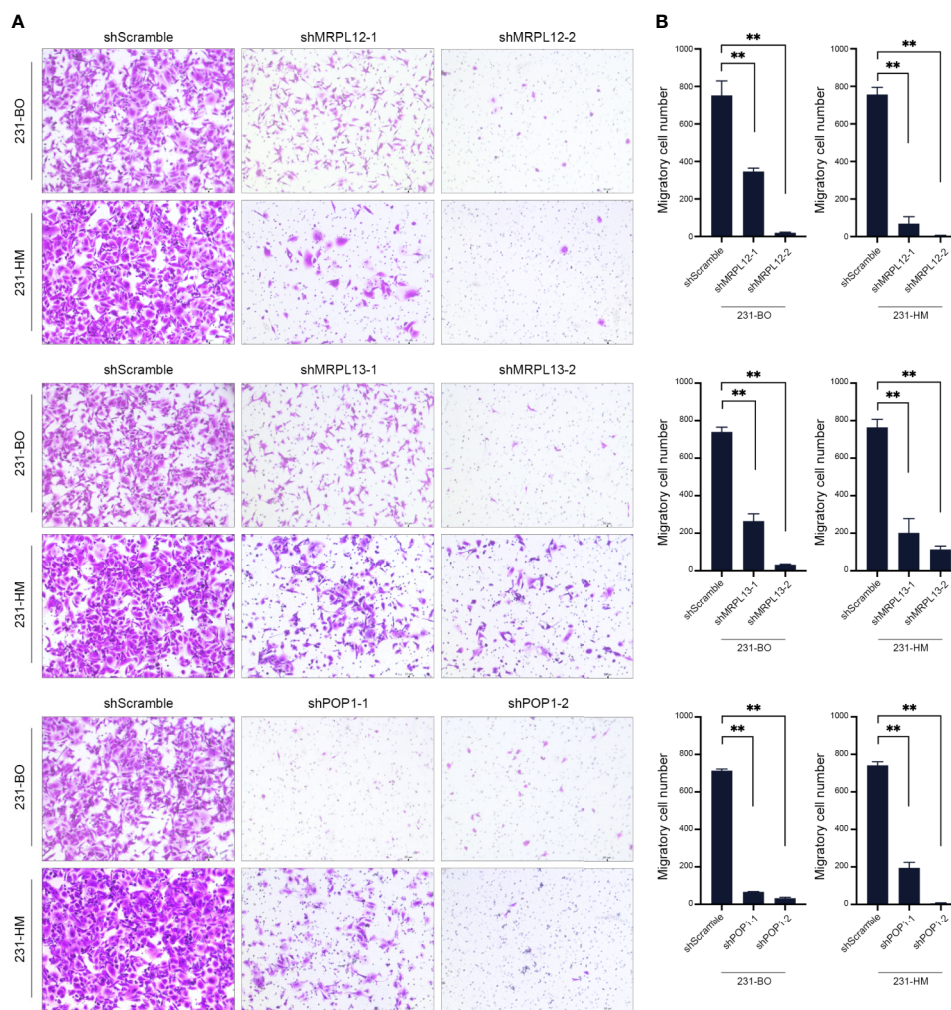


FIGURE 9 | Influence of the hub RBPs on migration of breast cancer cells. **(A)** Transwell assay for evaluating the influence of down-regulating expression of the hub RBPs on migration of 231-BO and 231-HM cells; **(B)** Counting and statistical analysis of transwell assay. ** $p < 0.01$; Bars, \pm SD.

suggesting that these three RBP-coding genes might act as oncogenes in accelerating the progress of breast cancer (**Figures 8, 9**). All these results indicated that *MRPL12*, *MRPL13* and *POP1* not only could be designed as prognostic factors for predicting survival probability of patients with breast cancer, but also could be used as potential targets for clinical intervention of breast cancer.

As far as we know, *MRPL12* and *MRPL13* are components of the mitochondrial large ribosomal subunit, while *POP1* has been identified as a communal component of the nuclear RNase P and the MRP ribonuclease complex (23–25). Mechanistically, *MRPL12* and *MRPL13* are major in maintaining the structural and functional integrity of mitoribosome, which is a dedicated apparatus for producing mtDNA-encoded multimeric oxidative phosphorylation enzymes and be essential in regulating the mitochondrial respiration and energy homeostasis (24, 26). Besides, *MRPL12* also interacts directly with the mitochondrial RNA polymerase (*POLRMT*) and facilitates the mitochondrial

transcription (27). Deficiency of either *MRPL12* or *MRPL13* can disrupt the mitoribosome and disturb the expression of mitochondrial genes, that eventually lead to loss of the functional mitochondrion (28). Accumulating evidence showed that the deformed and/or dysfunctional mitoribosome was responsible for some devastating malignancies, and a panel of mitoribosome components as well as its assembly factors could be regarded as hallmarks to plot the molecular portraits for specific tumors, such as *MRPS23* be significantly in the breast cancer, while *DDX28* acts as a risk factor for colorectal cancer (29, 30). Here, we had identified that *MRPL12* and *MRPL13* also could be regarded as potential risk factors for breast cancer. In this context, some strategies that targeting the *MRPL12*, *MRPL13* or even the mitoribosome might be practicable in curing breast cancer.

Moreover, analogous to *MRPL12* and *MRPL13*, our results also revealed that *POP1* functioned as an oncogene in breast cancer. In brief, majority of patients with abundant *POP1*

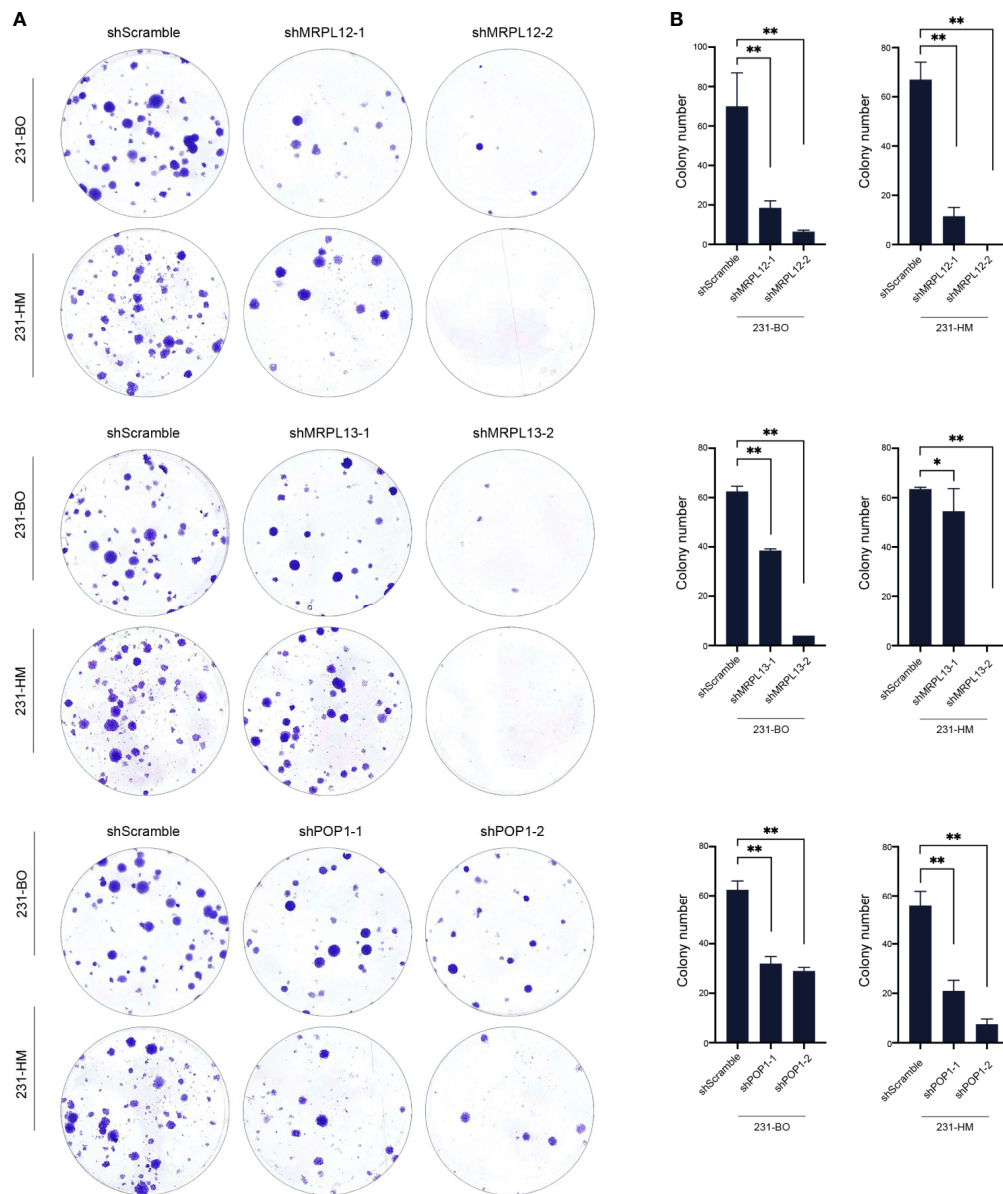


FIGURE 10 | Influence of the hub RBPs on cellular viability of breast cancer. **(A)** Colony efficiency assay for evaluating the influence of down-regulating expression of the hub RBPs on cellular viability of 231-BO and 231-HM cells; **(B)** Counting and statistical analysis of colony efficiency assay. * $p < 0.05$; ** $p < 0.01$; Bars, \pm SD.

expression showed poorer survival probability, and silencing the endogenous POP1 dramatically suppressed the viability and metastatic ability of breast cancer cells (**Figures 4, 9, 10**). Previous studies have reported that POP1 plays the role of a scaffold for stabilizing the global architecture of RNase P and MRP, which are primarily responsible for the maturation and metabolism of tRNA, rRNA as well as some mRNA (25, 31, 32). Biallelic mutation or haploinsufficiency of POP1 impairs the holoenzyme assembly and results in a severe skeletal dysplasia, but whether it is involved in tumorigenesis still is unsubstantiated (33).

Notably, by reviewing the Comparative Toxicogenomics Database (CTD), we excavated three available compounds (Acetaminophen, Urethane and Tunicamycin), which could inhibit the MRPL12, MRPL13 and POP1 simultaneously, implying their anti-neoplastic effects on breast cancer (**Figure 11**). Coinciding with our prediction, some studies have reported that Acetaminophen, one of the most widely used anti-inflammatory, antipyretic and analgesic drug, could inhibit the tumor growth by inducing the differentiation of breast cancer stem cells (34). Tunicamycin could induce apoptosis, decelerate growth and aggressiveness of breast cancer cells *via* the Akt/NF- κ B signaling

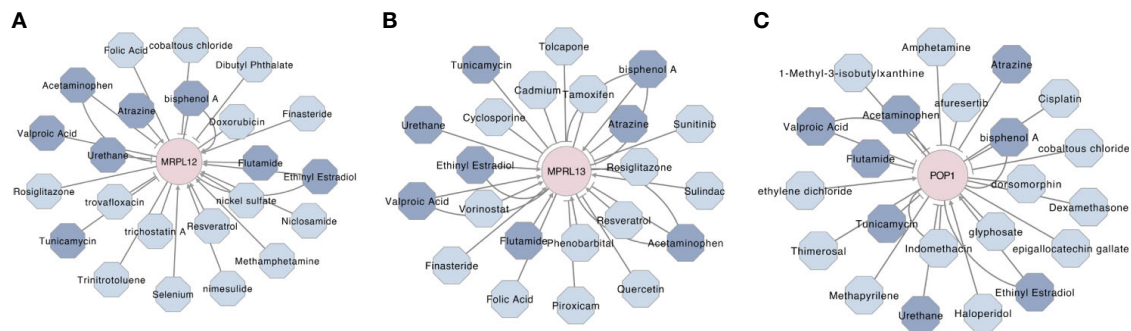


FIGURE 11 | Prognostic biomarker-drug interaction network of the hub RBPs in breast cancer. The network illustrated some available chemicals that would increase or decrease the expression levels and/or activity of MRPL12 (A), MRPL13 (B) and/or POP1 (C). The triangle arrow presented up-regulation, while the vertical arrow presented down-regulation. Moreover, the chemicals which be highlighted by navy blue were predicted as the communal chemical that might influence these three hub RBPs, simultaneously.

pathway (35). Meanwhile, Tunicamycin could enhance the anti-neoplastic activity of trastuzumab on breast cancer (36), suggesting a possibility that these compounds should be applied into the systemic therapeutic strategies for breast cancer.

CONCLUSION

Oval all, we identified three RBP-coding genes (*MRPL12*, *MRPL13* and *POP1*) that functioned as risk factors and independent prognostic factors for breast cancer. Naturally, a three-RBPs-signature based prognostic model was established, and be validated for predicting the survival probability of patients with breast cancer. Furthermore, we performed some preliminary experiments and provided concrete evidence that *MRPL12*, *MRPL13* and *POP1* might be oncogenes in maintaining cellular viability as well as accelerating metastasis of the breast cancer cells. All these results suggested that *MRPL12*, *MRPL13* and *POP1* should be designed as biomarkers and potential intervening targets for breast cancer. However, more in-depth studies are required to reveal the specific molecular mechanisms and to verify the related clinical practices.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Fudan University Shanghai

Cancer Center. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YL and HS conceived of and designed the study. QL, YZ, LL, and BX contributed to the literature search. YL and XL performed the data analysis. YL wrote the initial draft of this manuscript. YH and WJ reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the grants from National Natural Science Foundation of China (81972727 and 82002796).

ACKNOWLEDGMENTS

The authors appreciate the patients who have participated in FUSCC-TNBC and this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2021.663556/full#supplementary-material>

Supplementary Figure 1 | Lasso-penalized Cox analysis. (A) Lasso coefficient profiles; (B) Cross-validation for tuning parameter selection in the Lasso model.

Supplementary Figure 2 | Kaplan-Meier analysis for the prognostic significant of other RBPs in breast cancer. (A, B) Survival Kaplan-Meier estimate of RBPs identified by stepwise-multivariate Cox regression analysis in the TCGA-BRCA database (n=1088), clinical samples were grouped by Youden index of mRNA expression level; B, Kaplan-Meier plotter of *EF14E3* or *NOVA2* in the METABRIC database (n=1893), in which samples were grouped by Youden index of mRNA expression level. OS, overall survival and $p < 0.05$ was considered as statistically significant.

REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
- Gerstberger S, Hafner M, Tuschl T. A Census of Human RNA-Binding Proteins. *Nat Rev Genet* (2014) 15:829–45. doi: 10.1038/nrg3813
- Rissland OS. The Organization and Regulation of mRNA-Protein Complexes. *Wiley Interdiscip Rev RNA* (2017) 8:1369–86. doi: 10.1002/wrna.1369
- Muller-McNicoll M, Neugebauer KM. How Cells Get the Message: Dynamic Assembly and Function of mRNA-Protein Complexes. *Nat Rev Genet* (2013) 14:275–87. doi: 10.1038/nrg3434
- Gehring NH, Wahle E, Fischer U. Deciphering the Mrnp Code: RNA-Bound Determinants of Post-Transcriptional Gene Regulation. *Trends Biochem Sci* (2017) 42:369–82. doi: 10.1016/j.tibs.2017.02.004
- Hentze MW, Castello A, Schwarzl T, Preiss T. A Brave New World of RNA-Binding Proteins. *Nat Rev Mol Cell Biol* (2018) 19:327–41. doi: 10.1038/nrm.2017.130
- Beckmann BM, Horos R, Fischer B, Castello A, Eichelbaum K, Alleaume AM, et al. The RNA-Binding Proteomes From Yeast to Man Harbour Conserved Enigmrbps. *Nat Commun* (2015) 6:10127–36. doi: 10.1038/ncomms10127
- Yang SY, He XY, Isaacs C, Dobkin C, Miller D, Philipp M. Roles of 17Beta-Hydroxysteroid Dehydrogenase Type 10 in Neurodegenerative Disorders. *J Steroid Biochem Mol Biol* (2014) 143:460–72. doi: 10.1016/j.jsbmb.2014.07.001
- Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmann W. RNase P Without RNA: Identification and Functional Reconstitution of the Human Mitochondrial tRNA Processing Enzyme. *Cell* (2008) 135:462–74. doi: 10.1016/j.cell.2008.09.013
- Kang D, Lee Y, Lee JS. Rna-Binding Proteins in Cancer: Functional and Therapeutic Perspectives. *Cancers (Basel)* (2020) 12:2699–732. doi: 10.3390/cancers12092699
- Pereira B, Billaud M, Almeida R. Rna-Binding Proteins in Cancer: Old Players and New Actors. *Trends Cancer* (2017) 3:506–28. doi: 10.1016/j.trecan.2017.05.003
- Moore S, Jarvelin AI, Davis I, Bond GL, Castello A. Expanding Horizons: New Roles for non-Canonical RNA-binding Proteins in Cancer. *Curr Opin Genet Dev* (2018) 48:112–20. doi: 10.1016/j.gde.2017.11.006
- Gutschner T, Hammerle M, Pazaitis N, Bley N, Fiskin E, Uckelmann H, et al. Insulin-Like Growth Factor 2 mRNA-binding Protein 1 (IGF2BP1) is an Important Protumorigenic Factor in Hepatocellular Carcinoma. *Hepatology* (2014) 59:1900–11. doi: 10.1002/hep.26997
- Cheng F, Pan Y, Lu YM, Zhu L, Chen S. Rna-Binding Protein Dnd1 Promotes Breast Cancer Apoptosis by Stabilizing the Bim mRNA in a Mir-221 Binding Site. *BioMed Res Int* (2017) 2017:9596152–10. doi: 10.1155/2017/9596152
- Warzecha CC, Sato TK, Nabet B, Hogenesch JB, Carstens RP. ESRP1 and ESRP2 are Epithelial Cell-Type-Specific Regulators of FGFR2 Splicing. *Mol Cell* (2009) 33:591–601. doi: 10.1016/j.molcel.2009.01.025
- Yae T, Tsuchihashi K, Ishimoto T, Motohara T, Yoshikawa M, Yoshida GJ, et al. Alternative Splicing of CD44 mRNA by ESRP1 Enhances Lung Colonization of Metastatic Cancer Cell. *Nat Commun* (2012) 3:883–92. doi: 10.1038/ncomms1892
- Fan X, Liu L, Shi Y, Guo F, Wang H, Zhao X, et al. Integrated Analysis of RNA-binding Proteins in Human Colorectal Cancer. *World J Surg Oncol* (2020) 18:222–36. doi: 10.1186/s12957-020-01995-5
- Wang K, Li L, Fu L, Yuan Y, Dai H, Zhu T, et al. Integrated Bioinformatics Analysis the Function of RNA Binding Proteins (Rbps) and Their Prognostic Value in Breast Cancer. *Front Pharmacol* (2019) 10:140–51. doi: 10.3389/fphar.2019.00140
- Kentsis A, Topisirovic I, Culjkovic B, Shao L, Borden KL. Ribavirin Suppresses eIF4E-mediated Oncogenic Transformation by Physical Mimicry of the 7-Methyl Guanosine mRNA Cap. *Proc Natl Acad Sci U S A* (2004) 101:18105–10. doi: 10.1073/pnas.0406927102
- Harbeck N, Gnant M. Breast Cancer. *Lancet* (2017) 389:1134–50. doi: 10.1016/S0140-6736(16)31891-8
- Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene Expression Patterns of Breast Carcinomas Distinguish Tumor Subclasses With Clinical Implications. *Proc Natl Acad Sci USA* (2001) 98:10869–74. doi: 10.1073/pnas.191367098
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular Portraits of Human Breast Tumours. *Nature* (2000) 406:747–52. doi: 10.1038/35021093
- Surovtseva YV, Shutt TE, Cotney J, Cimen H, Chen SY, Koc EC, et al. Mitochondrial Ribosomal Protein L12 Selectively Associates With Human Mitochondrial RNA Polymerase to Activate Transcription. *Proc Natl Acad Sci USA* (2011) 108:17921–6. doi: 10.1073/pnas.1108852108
- Ke H, Dass S, Morrisey JM, Mather MW, Vaidya AB. The Mitochondrial Ribosomal Protein L13 is Critical for the Structural and Functional Integrity of the Mitochondrion in Plasmodium Falciparum. *J Biol Chem* (2018) 293:8128–37. doi: 10.1074/jbc.RA118.002552
- Fagerlund RD, Perederina A, Berezina I, Krasilnikov AS. Footprinting Analysis of Interactions Between the Largest Eukaryotic RNase P/MRP Protein Pop1 and RNase P/Mrp RNA Components. *RNA* (2015) 21:1591–605. doi: 10.1261/rna.049007.114
- Greber BJ, Ban N. Structure and Function of the Mitochondrial Ribosome. *Annu Rev Biochem* (2016) 85:103–32. doi: 10.1146/annurev-biochem-060815-014343
- Nouws J, Goswami AV, Bestwick M, McCann BJ, Surovtseva YV, Shadel GS. Mitochondrial Ribosomal Protein L12 is Required for POLRMT Stability and Exists as Two Forms Generated by Alternative Proteolysis During Import. *J Biol Chem* (2016) 291:989–97. doi: 10.1074/jbc.M115.689299
- Cheong A, Lingutla R, Mager J. Expression Analysis of Mammalian Mitochondrial Ribosomal Protein Genes. *Gene Expr Patterns* (2020) 38:119147. doi: 10.1016/j.gep.2020.119147
- Kim HJ, Maiti P, Barrientos A. Mitochondrial Ribosomes in Cancer. *Semin Cancer Biol* (2017) 47:67–81. doi: 10.1016/j.semcancer.2017.04.004
- Klaestad E, Opdahl S, Engstrom MJ, Ytterhus B, Wik E, Bofin AM, et al. MRPS23 Amplification and Gene Expression in Breast Cancer; Association With Proliferation and the Non-Basal Subtypes. *Breast Cancer Res Treat* (2020) 180:73–86. doi: 10.1007/s10549-020-05532-6
- Garcia PD, Leach RW, Wadsworth GM, Choudhary K, Li H, Aviran S, et al. Stability and Nuclear Localization of Yeast Telomerase Depend on Protein Components of RNase P/Mrp. *Nat Commun* (2020) 31:2173–292. doi: 10.1038/s41467-020-15875-9
- Garcia PD, Zakian VA. A New Role for Proteins Subunits of RNase P: Stabilization of the Telomerase Holoenzyme. *Microb Cell* (2020) 7:250–4. doi: 10.15698/mic2020.09.730
- Barraza-Garcia J, Rivera-Pedroza CI, Hisado-Oliva A, Belinchon-Martinez A, Senthordhi-Montane L, Duncan EL, et al. Broadening the Phenotypic Spectrum of POP1-Skeletal Dysplasias: Identification of POP1 Mutations in a Mild and Severe Skeletal Dysplasia. *Clin Genet* (2017) 92:91–8. doi: 10.1111/cge.12964
- Takehara M, Hoshino T, Namba T, Yamakawa N, Mizushima T. Acetaminophen-Induced Differentiation of Human Breast Cancer Stem Cells and Inhibition of Tumor Xenograft Growth in Mice. *Biochem Pharmacol* (2011) 81:1124–35. doi: 10.1016/j.bcp.2011.02.012
- Wang X, Xiong W, Tang Y. Tunicamycin Suppresses Breast Cancer Cell Growth and Metastasis Via Regulation of the Protein Kinase B/Nuclear Factor-Kappab Signaling Pathway. *Oncol Lett* (2018) 15:4137–42. doi: 10.3892/ol.2018.7874
- Han X, Zhang X, Li H, Huang S, Zhang S, Wang F, et al. Tunicamycin Enhances the Antitumor Activity of Trastuzumab on Breast Cancer In Vitro and In Vivo. *Oncotarget* (2015) 6:38912–25. doi: 10.18632/oncotarget.5334

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Liu, Sun, Li, Liu, Zhao, Li, Xu, Hou and Jin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.