Screening a Prognosis-Related Target Gene in Patients with HER-2-Positive Breast Cancer by Bioinformatics Analysis

Song Wang  Jian Deng  Xudong Gao  Hongying Lv  Yi Quan
Department of Breast Surgery, The Affiliated Hospital of Southwest Medical University, Luzhou, China

Highlights of the Study

• Based on bioinformatics principles, this study identified RRM2 as a new biomarker and a therapeutic target for HER-2-positive breast cancer.
• RRM2 is associated with cell cycle regulation in the occurrence and development of HER-2-positive breast cancer.

Keywords
HER-2-positive breast cancer · Differentially expressed genes · Bioinformatics analysis · Target genes · RRM2

Abstract

Objective: The objective of the present study was to determine a target gene and explore the molecular mechanisms involved in the pathogenesis of HER-2-positive breast cancer. Methods: Three RNA expression profiles obtained from the Gene Expression Omnibus (GEO) and the Cancer Genome Atlas (TCGA) were used to identify differentially expressed genes (DEGs) using the R software. A protein-protein interaction network was then constructed, and hub genes were determined. Subsequently, the relationship between clinical parameters and hub genes was examined to screen for target genes. Next, DNA methylation and genomic alterations of the target gene were evaluated. To further explore potential molecular mechanisms, a functional enrichment analysis of genes coexpressed with the target gene was performed. Results: The differential expression analysis revealed 217 DEGs in HER-2-positive breast cancer samples compared to normal breast tissues. RRM2 was the only hub gene closely associated with lymphatic metastasis and the patients’ prognosis. Additionally, RRM2 was found to be consistently amplified and negatively associated with the level of methylation. Functional enrichment analysis showed that the coexpressed genes were mainly involved in cell cycle regulation. Conclusions: RRM2 was identified as a target gene associated with the initiation, progression, and prognosis of HER-2-positive breast cancer, which may be considered as a new biomarker and therapeutic target.

Introduction

HER-2-positive breast cancer accounts for about 20% of all the breast cancer cases in China and is closely associated with a highly aggressive clinical course, due to amplification of the HER-2 or overexpression of the protein. Trastuzumab plus chemotherapy has become the first-
Predictive value with a p-value <0.05 indicated significant enrich-
ment of GO terms and KEGG pathways, and the GO annotation
analysis was carried out using the R software to display the results of differential expression analysis in 3 datasets. Finally, the common DEGs were identified by the Draw Venn Diagram tool.

**PPI Network Construction and Module Analysis**
Protein-protein interaction (PPI) analysis was carried out using the Search Tool for the Retrieval of Interacting Genes (STRING) database with a combined score of ≥0.7 as the threshold. Subsequently, the PPI network was visualized using the Cytoscape software. We used the Molecular Complex Detection (MCODE) and CytoHubba, two of Cytoscape plugin software, to perform module analysis of the PPI network and screen the hub genes, respectively. Moreover, degree cutoff = 2, node score cutoff = 0.2, k-core = 2, and max. depth = 100 were set as the cutoff criteria in module analysis. The highest scores of the degree algorithm were retained as the criteria to screen out hub genes with high connectivity in the most significant module.

**Survival and Clinicopathological Parameter Analysis of Hub Genes**
We used the GEPIA database to perform an OS analysis of the previously identified hub genes. Using the clinical data (age, pathologic T, pathologic N, and pathologic stage) of the 123 patients with HER-2-positive breast cancer, we further explored the correlation between hub genes and clinical parameters and identified a target gene most closely related to the development of HER-2-positive breast cancer.

**Target Gene Validation Based on Multiple Databases**
In the present study, using the Oncomine, TIMER, UALCAN, and Kaplan-Meier plotter database, we validated the overexpression and prognostic value of the target gene in HER-2-positive breast cancer.

**Target Gene Methylation Analysis**
We used UALCAN and cBioportal to assess the correlation between target gene mRNA expression and its methylation values. Moreover, MethSurv, a freely valuable platform, was used to assess the effects of target gene methylation on patient survival time.

**Genomic Alterations and Mutation Analysis of the Target Gene**
The cBioPortal was used to analyze the genomic alteration status of the target gene in breast cancer. Four breast cancer studies in the cBioPortal were utilized for this analysis. The 4 studies included Breast Invasive Carcinoma (TCGA, Cell 2015), Breast Invasive Carcinoma (TCGA, Firehose Legacy), Breast Invasive Carcinoma (TCGA, Nature 2012), and Breast Invasive Carcinoma (TCGA, Pan-Cancer Atlas). The COSMIC database was used to examine the different types of mutations affecting the target gene in breast cancer.

**Functional Enrichment Analysis of Coexpressed Genes Correlated with the Target Gene in Breast Cancer**
We used the cBioPortal database to screen genes coexpressed with the target gene in breast cancer, with Spearman scores >0.5 as unique criteria. We then conducted an analysis of the Gene Ontology (GO) annotation and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway using the clusterProfiler package [2]. A Benjamin-corrected p value <0.05 indicated significant enrichment of GO terms and KEGG pathways, and the GO annotation comprised the biological process, molecular function, and cellular component.
Results

Identification of DEGs

The microarray datasets (GSE29431 and GSE65194) were standardized, and the results were illustrated in a box plot (online suppl. Fig. 2). With a cutoff of adj \( p \) value <0.01 and fold change \( \geq 2 \), we eventually identified 503 DEGs (96 upregulated and 407 downregulated), 1,525 DEGs (809 upregulated and 716 downregulated), and 2,035 DEGs (1,000 upregulated and 1,035 downregulated) in GSE29431, GSE65194, and TCGA respectively, which were differentially expressed between HER-2-positive breast cancer samples and normal breast tissues. The volcano plot showed the DEGs (Fig. 1a–c). The VENN plot delineated the common 217 DEGs in the 3 datasets, including 50 upregulated genes and 167 downregulated genes (Fig. 1d, e).

Fig. 1. The volcano plot and Venn diagram of DEGs in the 3 datasets. a The volcano plot of DEGs for GSE29431. b The volcano plot of DEGs for GSE65194. c The volcano plot of DEGs for TCGA. d Venn diagram of common upregulated genes. e Venn diagram of common downregulated genes. In the volcano plot, each dot represents an individual gene (red dots denote upregulated genes with logFC > 2 and adj \( p \) value <0.01, while blue dots represent downregulated genes with logFC <2 and adj \( p \) value <0.01). FC, fold change; adj \( p \) value, adjusted \( p \) value; DEGs, differentially expressed genes.
Screening a Target Gene by Bioinformatics Analysis

PPI Network Construction and Module Analysis

The 217 DEGs were input into the STRING database for PPI network analysis, which was then visualized by the Cytoscape software, and achieved a PPI network of 111 nodes and 310 edges, with combined scores ≥0.7 (Fig. 2a). The most significant cluster 1 totally consisting of upregulated genes was identified using the MCODE plugin (Fig. 2b). Subsequently, with the highest connectivity degree = 16, 13 genes selected from cluster 1 were considered as hub genes, including UBE2C, KIF20A, PTTG1, CENPF, CCNB1, BIRC5, TOP2A, CEP55, AURKA, RRM2, BUB1B, NUSAP1, and MELK. These genes are presented in online suppl. Table 4.

Survival and Clinicopathological Parameter Analysis of Hub Genes

Using the GEPIA database, we assessed the prognostic value of hub genes in HER-2-positive breast cancer patients. The specific results of the survival analysis are shown in online suppl. Table 5. Among them, only the upregulated NUSAP1 and RRM2 genes were associated with a poor prognosis based on the OS (Fig. 2c, d). Clini-
copathological analysis of RRM2 and NUSAP1 using the χ² test demonstrated that only RRM2 was linked with pathologic N (p < 0.05; Table 1). As shown in online suppl. Table 6, there was no significant association of NUSAP1 expression with age (p = 1.000), pathologic T (p = 0.618), pathologic N (p = 0.054), and pathologic stage (p = 0.533). These results suggested that RRM2 may be involved in the occurrence, development, and prognosis of HER-2-positive breast cancer.

Validation of the Correlation between RRM2 mRNA Expression and Prognosis

Using the Oncomine and TIMER databases, we found that RRM2 was overexpressed in multiple cancer types (online suppl. Fig. 3a, b), suggesting that it might be relevant to the oncogenesis of these tumors. From the Oncomine and UALCAN databases, RRM2 overexpression was confirmed in HER-2-positive breast cancer (online suppl. Fig. 3c, d). By using the Kaplan-Meier plotter database, we further corroborated the prognostic significance of RRM2 (online suppl. Fig. 3e). Consequently, RRM2 was confirmed to be associated with HER-2-positive breast cancer.

RRM2 Methylation Analysis

From the UALCAN and cBioPortal databases, we observed that the RRM2 methylation levels in HER-2-positive breast cancer samples were apparently lower than those in normal breast tissues (p < 0.05) and showed a negative association with gene expression (p < 0.05, r = −0.30) (Fig. 3a, b). Using the MethSurv web tool, we found that hypermethylation was significantly correlated with a more favorable prognosis (Fig. 3c). We thus deduced that DNA hypomethylation was correlated with high gene expression levels, which may play an important role in the occurrence and development of HER-2-positive breast cancer.

Analysis of RRM2 Genomic Alterations and Mutations

Using 4 studies in the cBioPortal, we explored the genetic alterations of RRM2 in breast cancer patients. The results demonstrated that RRM2 was altered in 52 (1.4%) out of 3,798 patients (online suppl. Table 7). As depicted in online suppl. Figure 4a and b, these alterations including amplification, deep deletion, and mutation were associated with a poor prognosis and a lower OS in breast cancer patients. Furthermore, the pie chart showed the corresponding types of mutation of RRM2 in breast cancer patients from the COSMIC database, with missense and synonymous mutations being the most common types, and the mutations of the coding strand in RRM2 were as follows: G > A, A > T, G > T, A > G, and C > T (online suppl. Fig. 4c, d).

Functional Enrichment Analysis of RRM2 Coexpressed Genes

According to the gene coexpression modules obtained from the cBioPortal database, a total of 292 genes with Spearman’s correlation scores >0.5 were identified as be-

Table 1. The relationship between clinicopathological parameters and RRM2 gene expression

| Parameter       | Group   | Patients (n/%) | Low (n/%) | High (n/%) | p value |
|-----------------|---------|---------------|-----------|------------|---------|
| Age             | Age >60 | 63 (51.2)     | 33 (53.2) | 30 (49.2)  | 0.788   |
|                 | Age ≤60 | 60 (48.8)     | 29 (46.8) | 31 (50.8)  |         |
| Pathologic T    | T1      | 28 (22.8)     | 16 (25.8) | 12 (19.7)  | 0.472   |
|                 | T2      | 79 (64.2)     | 36 (58.1) | 43 (70.5)  |         |
|                 | T3      | 12 (9.8)      | 8 (12.9)  | 4 (6.6)    |         |
|                 | T4      | 4 (3.3)       | 2 (3.2)   | 2 (3.3)    |         |
| Pathologic N    | N0      | 45 (36.6)     | 25 (40.3) | 20 (32.8)  | 0.039   |
|                 | N1      | 52 (42.3)     | 24 (38.7) | 28 (45.9)  |         |
|                 | N2      | 13 (10.6)     | 3 (4.8)   | 10 (16.4)  |         |
|                 | N3      | 13 (10.6)     | 10 (16.1) | 3 (4.9)    |         |
| Pathologic Stage| I       | 13 (10.6)     | 9 (14.5)  | 4 (6.6)    | 0.236   |
|                 | II      | 76 (61.8)     | 36 (58.1) | 40 (65.6)  |         |
|                 | III     | 32 (26.0)     | 15 (24.2) | 17 (27.9)  |         |
|                 | IV      | 2 (1.6)       | 2 (3.2)   | 0 (0.0)    |         |
ing coexpressed with RRM2. The top 10 coexpressed genes are listed in online suppl. Table 8. For the GO analysis, the results pointed out that organelle fission was principally enriched in biological processes, and the chromosomal region was primarily enriched in the cellular component, whereas ATPase activity was chiefly enriched in molecular functions. Besides, KEGG pathway analysis indicated that these coexpressed genes were mainly enriched in the cell cycle. The results of the functional enrichment analysis are revealed in online suppl. Figure 5. As shown in Figure 4, the corresponding genes were colored red in the pathway mainly located in the S or G2/M phases of the cell cycle.

Discussion

HER-2 subtype breast cancer often presents with tumor recurrence and metastasis, and the patients’ OS rate is lower. In recent years, targeted drug therapies have improved the OS for patients with HER-2 subtype breast cancer [1]. Despite the advent of this standardized initial therapy, some patients rapidly develop disease progression, which is predominantly due to the development of primary and secondary resistance to targeted drugs. Thus, further studies are necessary to explore the pathogenesis of breast cancer and search for novel therapeutic targets.

In this study, we identified 217 DEGs in 3 datasets. With a highest connectivity degree of 16, 13 genes were considered as hub genes, including UBE2C, KIF20A, AURKA, BIRC5, PTTRG1, CENPF, CCNB1, TOP2A, CEP55, BUB1B, NUSAP1, MELK, and RRM2. Clinico-pathological analysis suggested that only RRM2 was positively correlated with a poor prognosis and associated with lymph node metastasis in HER-2-positive breast cancer. Additionally, using the Oncomine and TIMER databases, we found that RRM2 was overexpressed in multiple tumors (online suppl. Fig. 3). Thus, we deduced that RRM2 may play the role of oncogene in diverse carcinomas and is linked to the occurrence, development, diagnosis, treatment, and prognosis of HER-2-positive breast cancer.

High UBE2C expression contributes to estrogen-independent growth in HR+ breast cancer cells, which suggests that tamoxifen treatment combined with inhibitors targeting UBE2C could be a potential therapeutic strategy to overcome resistance in endocrine therapy [3]. KIF20A small interfering RNA has been reported to significantly inhibit the proliferation of pancreatic cancer cell lines [4]. Alisertib, a selective small-molecule AURKA inhibitor, was found to be able to promote preferential growth inhibition in AI-resistant cell lines [5]. Overexpression of BIRC5 has been shown to promote breast tumor proliferation by inducing genetic instability [6]. PTTRG1 is correlated with the regulation of cancer stem cells in breast cancer [7]. Sun et al. [8] demonstrated that CENPF promoted breast cancer bone metastasis by activating the PI3K-AKT-mTOR signaling pathway. CCNB1 silencing inhibits cell proliferation and promotes cell senescence.
via activation of the p53 signaling pathway in pancreatic cancer [9].

The PTEN/AKT signaling pathway affects the expression of TUBB3 and TOP2A, reducing cell growth and inducing apoptosis of human breast cancer MCF-7 cells through ATP and caspase-3 signaling pathways [10]. CEP55 has been shown to promote hepatocellular carcinoma cell migration and invasion by regulating the JAK2-STAT3-MMPs signaling pathway [11]. The oncogenic effect of BUB1B has been proven to be impaired by rapamycin via downregulation of the mTORC1 signaling pathway [12]. NUSAP1 is a potential prognostic marker for breast cancer patients that promotes the progression of breast cancer through the Wnt/β-catenin and epithelial-mesen-

chymal transition pathways [13]. Previous studies have demonstrated that MELK inhibition promotes cancer cell death through the p53 signaling pathway [14].

RRM2, a rate-limiting enzyme for DNA synthesis and repair, is involved in cell cycle regulation and overexpressed in various tumors, which is associated with cell proliferation, invasiveness, and migration. This is consistent with the findings observed in the present study. Previous studies have shown that small molecular antagonists of the RRM2 gene significantly reduce tamoxifen-resistant cell proliferation and decrease tumor growth [15]. Moreover, suppression of RRM2 synthesis could enhance the tumor’s chemosensitivity to Adriamycin, which has led to the recognition of RRM2 as a suitable biomarker.

Fig. 4. The position of RRM2-associated genes in the cell cycle. The corresponding genes in the pathway are marked with red circles.
er [16]. However, the studies focused only on estrogen-negative breast cancer, and the cancer-forming pathways underlying this process are yet to be discussed. In this study, we aimed to find an effective target for the treatment of HER-2-positive breast cancer, comprising estrogen-positive breast cancer (luminal B subtype [positive HER-2 expression]). In this study, we elucidated the molecular mechanisms and signal pathways and also performed further analysis of RRM2 genomic alterations and mutations. We inferred that methylation of the RRM2 gene may also serve as a potential drug target.

In recent years, some current therapeutic regimens or drugs use RRM2 as the therapeutic target. GTI2040, as an RRM2 expression silencer, is complementary to a region of the coding sequence in the RRM2 mRNA and shows strong antitumor activity. Additionally, GTI2040 decreases RRM2 mRNA expression in human tumor cell lines and arrests cells in the early G1/S phase. GTI2040 in combination with weekly doses of gemcitabine has shown good treatment outcomes for patients with advanced solid tumors in a phase I study [17]. Another RRM2 inhibitor, GW8510, suppresses the survival of lung squamous cell carcinoma cells and reverses their resistance to gemcitabine by inducing DNA damage and cell apoptosis [18].

Subsequently, we found a significantly low level of RRM2 methylation and identified trends for the negative correlation between its methylation and expression levels, thus confirming that DNA hypomethylation was primarily responsible for the increased RRM2 expression in HER-2-positive breast cancer. Survival analysis revealed that increasing levels of methylation were related to poor prognosis, which indicates that level of methylation of RRM2 can be used as a prognostic predictor with breast carcinoma.

Based on these observations, we speculated that most RRM2-related genes are linked to the cell cycle and promoted tumorigenesis and progression by regulating the G2/M phase transition in breast cancer. Previous studies have shown that multiple drugs can alter the proliferative activity of tumor cells by regulating the G2/M phase transition in the cell cycle. For example, CAPE could inhibit the cell cycle and eliminate breast cancer cells in the G2/M phase [19]. Moreover, cucurbitacin B has been shown to inhibit the proliferation of breast cancer cells by inducing cell cycle arrest at the G2/M phase, leading to apoptosis [20]. Therefore, RRM2 inhibition by regulating the G2/M phase in HER-2-positive breast cancer could possibly serve as a novel therapeutic strategy.

**Conclusion**

In summary, this study demonstrated that RRM2 was overexpressed in HER-2-positive breast cancer and was associated with lymph node metastasis and the prognosis of these patients, indicating that RRM2 expression is closely related to tumorigenesis and cancer progression. Based on functional enrichment analysis, we speculate that RRM2 promotes tumorigenesis principally via cell cycle regulation. These findings suggest that RRM2 could be used as a potential therapeutic target against HER-2-positive breast cancer in the future. However, animal experiments and population-based studies are necessary to confirm the results obtained in this study.

**Statement of Ethics**

TCGA and GEO belong to public databases. The patients involved in the database provided ethical approval. Users can download relevant data for free for research and publish relevant articles. Our study is based on open source data, so there are no ethical issues.

**Conflict of Interest Statement**

The authors declare that they have no competing interests regarding the publication of this paper.

**References**

1 Perez EA, Romond EH, Suman VJ, Jeong JH, Davidson NE, Geyer CE Jr, et al. Four-year follow-up of trastuzumab plus adjuvant chemotherapy for operable human epidermal growth factor receptor 2-positive breast cancer: joint analysis of data from NCCTG N9831 and NSABP B-31. J Clin Oncol. 2011 Sep 1; 29:3366–73.

2 Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012 May;16:284–7.

3 Kim YJ, Lee G, Han J, Song K, Choi JS, Choi YL, et al. UBE2C overexpression aggravates patient outcome by promoting estrogen-dependent/independent cell proliferation in early hormone receptor-positive and HER2-negative breast cancer. Front Oncol. 2019 Jan 23:9:1574.
4 Taniuchi K, Nakagawa H, Nakamura T, Eguchi H, Ohigashi H, Ishikawa O, et al. Down-regulation of RAB6KIFL/KIF20A, a kinesin involved with membrane trafficking of discs large homologue 5, can attenuate growth of pancreatic cancer cell. Cancer Res. 2005 Jan 1; 65:105–12.
5 Shang YY, Yao M, Zhou ZW, Jian-Cui, Li-Xia, Hu RY, et al. Alisertib promotes apoptosis and autophagy in melanoma through p38 MAPK-mediated aurora a signaling. Oncotarget. 2017 Nov 6;8:10706–88.
6 Boidot R, Veyran F, Jacob D, Chevrier S, Gangneux N, Taboureau J, et al. The expression of BIRC5 is correlated with loss of specific chromosomal regions in breast carcinomas. Genes Chromosomes Cancer. 2008 Apr;47:299–308.
7 Yoon CH, Kim MJ, Lee H, Kim RK, Lim EJ, Yoo KC, et al. PTTG1 oncogene promotes tumor malignancy via epithelial to mesenchymal transition and expansion of cancer stem cell population. J Biol Chem. 2012 Jun 1;287:19516–27.
8 Sun J, Huang J, Lan J, Zhou K, Gao Y, Song Z, et al. Overexpression of CENPF correlates with poor prognosis and tumor bone metastasis in breast cancer. Cancer Cell Int. 2019 Oct 11;19:264.
9 Zhang H, Zhang X, Li X, Meng WB, Bai ZT, Rui SZ, et al. Effect of CCNB1 silencing on cell cycle, senescence, and apoptosis through the p53 signaling pathway in pancreatic cancer. J Cell Physiol. 2018 Jan;234:619–31.
10 Yang Z, Liu Y, Shi C, Zhang Y, Lv R, Zhang R, et al. Suppression of FTEN/AKT signaling decreases the expression of TUBB3 and TOP2A with subsequent inhibition of cell growth and induction of apoptosis in human breast cancer MCF-7 cells via ATP and caspase-3 signaling pathways. Oncol Rep. 2017 Feb;37:1011–9.
11 Li M, Gao J, Li D, Yin Y. CEP55 promotes cell motility via JAK2-STAT3-MMPs cascade in hepatocellular carcinoma. Cells. 2018 Aug 8; 7:99.
12 Qiu J, Zhang S, Wang P, Wang H, Sha B, Peng H, et al. BUB1B promotes hepatocellular carcinoma progression via activation of the mTORC1 signaling pathway. Cancer Med. 2020 Nov;9:8159–72.
13 Sun L, Shi C, Liu S, Zhang E, Yan L, Ji C, et al. Overexpression of NuSAP1 is predictive of an unfavourable prognosis and promotes proliferation and invasion of triple-negative breast cancer cells via the Wnt/β-catenin/EMT signalling axis. Gene. 2020 Jul 15;747:144657.
14 Wang Q, Zhu Y, Li Z, Bu Q, Sun T, Wang H, et al. Up-regulation of SPC25 promotes breast cancer. Aging. 2019 Aug 10;11:5689–704.
15 Shah KN, Wilson EA, Malla R, Elford HL, Faridi JS. Targeting ribonucleotide reductase M2 and NF-κB activation with ddoc to circumvent tamoxifen resistance in breast cancer. Mol Cancer Ther. 2015 Nov;14:2411–21.
16 Zhang H, Liu X, Warden CD, Huang Y, Loera S, Xue L, et al. Prognostic and therapeutic significance of ribonucleotide reductase small subunit M2 in estrogen-negative breast cancers. BMC Cancer. 2014 Sep 11;14:664.
17 Malik L, Zwiebel A, Cooper J. A phase I pharmacokinetic and pharmacodynamic study of GTI-2040 in combination with gemcitabine in patients with solid tumors. Cancer Chemother Pharmacol. 2018 Sep;82:533–9.
18 Chen P, Wu JN, Shu Y, Jiang HG, Zhao XH, Qian H, et al. Gemcitabine resistance mediated by ribonucleotide reductase M2 in lung squamous cell carcinoma is reversed by GW8510 through autophagy induction. Clin Sci. 2018 Jul 9;132:1417–33.
19 Rzepecka-Stojko A, Kabala-Dzik A, Moździerz A, Kubina R, Wojtynca RD, Stojko R, et al. Caffeic acid phenethyl ester and ethanol extract of propolis induce the complementary cytotoxic effect on triple-negative breast cancer cell lines. Molecules. 2015 May 20;20:9242–62.
20 Duangmano S, Sae-Lim P, Suksamrarn A, Domann FE, Patmasiriwat P, Cucurbitacin B inhibits human breast cancer cell proliferation through disruption of microtubule polymerization and nucleophosmin/B23 translocation. BMC Complement Altern Med. 2012 Oct 12;12:185.